Random Elements in a Banach Space

by

Édith Mourier.

Translated by Junhyung Park[1]

INTRODUCTION.

Until recently, the goal of Probability Calculus has been the study of random *numbers* or geometric probability, the study of a random point, defined by its coordinates, in a Euclidean space. The development of Probability Calculus and its applications calls for the study of more general *elements*: random series, vectors, functions, curves, transformations,... M. Fréchet [M. Fréchet, I, p.215-310][(1)] showed the necessity of a systematic study of abstract random elements; we will not go back to this point.

The set of realisations of an element X constitutes a space $\mathscr{X}$ of X and each realisation of X is a *point* of $\mathscr{X}$. The definition of abstract random elements implies the existence of a *measure* or *probability* on $\mathscr{X}$; we first need to introduce measurable sets forming a σ-algebra $\mathscr{F}$: the elements of $\mathscr{F}$ are subsets of $\mathscr{X}$ and $\mathscr{X}$ *itself is part of the σ-algebra $\mathscr{F}$*; then the measure μ is such that for all $A \in \mathscr{F}$, we have

$$\mu(A) \geqq 0, \qquad \mu(\mathscr{X}) = 1, \qquad \sum_i \mu(A_i) = \mu\left(\sum_i A_i\right)$$

if the countable collection A_i is pairwise disjoint. The generalisation of the notion of probability laws is then immediate: it is the collection of the probabilities of all the elements of $\mathscr{F}$. The notions of central values, dispersion, etc. imply that of "neighbourhood"; so there must be a topology on $\mathscr{X}$. Every metric defines a topology but there can be a topology in a space without a metric; however, we will restrict ourselves, like M. Fréchet, to metric spaces.

We can then, as M. Fréchet [M. Fréchet, I] and S. Doss [S. Doss, I] did, generalise the notion of mathematical expectation; however, if we want to generalise mathematical expectation as a *linear operator*, addition needs to be defined on $\mathscr{X}$, and so $\mathscr{X}$ needs to be vectorial; we can then directly generalise the classical

[1] This is not authoritative translation, and I do not claim any credit for the mathematical content of this document. Please send any corrections to `junhyung.park@tuebingen.mpg.de`.

(1) The square brackets [ ] can be found in the bibliography, page 65.

law of large numbers. Let us note that, considering the case where $\mathscr{X}$ is a metric space but not linear, S. Doss [S. Doss, I] defines the analogue of $\frac{1}{n}\sum_{i=1}^{n} X_i$ as any element Z_n, if it exists, such that

$$(Z_n, \lambda) \leqq \frac{1}{n}\sum_{i=1}^{n}(X_i, \lambda)$$

for any λ, where (a, b) denotes the distance between a and b.

He thus obtains a law of large numbers, but, except in a very particular case, he has neither an existence result nor a uniqueness result. So we see that we are led to consider metric vector spaces; more precisely, we will limit ourselves to *Banach* spaces, that is to say, to complete metric vector spaces in which the metric is defined through a norm. The linear functionals(2), real or complex, relative to $\mathscr{X}$ will be denoted x^*; they form the dual space $\mathscr{X}^*$ of $\mathscr{X}$ which is itself a Banach space.

We will assume that the measure defined on $\mathscr{X}$ is such that *all* linear functionals are measurable; we then say that the measure is an L-measure.

Under these conditions, we will study the definition and existence of a mathematical expectation (Chapter I), and we will establish laws of large numbers in expectation or almost surely (a.s.) with respect to the weak convergence and also with respect to the strong convergence (Chapter II). Chapter III is dedicated to the definition and study of the characteristic function of a random element, and finally in Chapter IV, we will define and study random Laplacian elements.

The main results shown in this Thesis are summarised in the following notes:

É. Mourier, *C. R. Acad. Sc.*, vol.229,1949, p.1300; vol.231, 1950, p.28; vol.232, 1951, p.923; vol.236, 1953, p.575.

I am happy to express here my respectful gratitude to Professor G. Darmois for the interest he has shown in this work, his observations and his very useful advice.

I equally address my strong gratitude to M. Fréchet whose recent works provided me with the subject of this study, who always takes an interest in my research, and from whom I learnt through numerous and instructive meetings.

Chapter I.

Definition and Study of a Mathematical Expectation in the case of an L-Measure.

In Probability Calculus, one defines the mathematical expectation of numerical random variables, but for a long time, in various applications, one has defined, and frequently uses, the "mean value" of elements which are not numbers. For example, nothing is more familiar to an artilleryman than the "average point" which he defines by the property that the coordinates of this point are the mathematical expectation of the corresponding coordinates; that is to say, if we

(2) In all that follows, "linear" implies *additive and continuous*.

denote by $X_i(M)$ the i^{th} coordinate of a point M, the *average point* E(M) is defined by the relation

$$X_i[E(M)] = E[X_i(M)] \qquad \text{for all } i.$$

Departing from the known notion of mathematical expectation for a numerical random variable, we generalise in an analogous way [É. Mourier, I] the definition of the mathematical expectation of a random element X whose values belong to some Banach space $\mathscr{X}$, which we will denote by $X \in \mathscr{X}$. Just as in the preceding example, the average point will only be defined when the mathematical expectation of each $X_i(M)$ exists, and likewise we will only define the mathematical expectation E(X) of $X \in \mathscr{X}$ if $E[x^*(X)]$ exists for each $x^* \in \mathscr{X}^*$.

Definition. – The mathematical expectation E(X) of $X \in \mathscr{X}$, if it exists, is the element of $\mathscr{X}$ such that

$$x^*[E(X)] = E[x^*(X)] \qquad \text{for every } x^* \in \mathscr{X}^*.$$

Let us note that $E[x^*(X)]$ is a mathematical expectation of an ordinary numerical random variable, so E(X), if it exists, is unique; indeed the knowledge of $x^*[E(X)]$ for every x^* determines E(X) [E. Hille, I, p.22].

This definition is equivalent ot that of the Pettis integral [Pettis I. p. 277-304].

Remark. – Given a set, we can sometimes choose different norms such that the space defined with either of these norms is a Banach space. Changing the norm can have the effect of modifying the class of (continuous) linear functionals which can be enlarged or shrunk (but there are always additive functions which are linear in all cases).

It is interesting to know whether E(X), which exists with one norm, also exists with another, and if these mathematical expectations are the same. In fact, in concrete applications, the choice of the norm is in general arbitrary.

Let $\|x\|_1$ and $\|x\|_2$ be two norms; if we assume that there exist two numbers a and b such that

$$(1) \qquad 0 < a < \frac{\|x\|_1}{\|x\|_2} < b,$$

the linear functionals are *the same* [E. Hille, I, Theorem 2.13.8] with $\|x\|_1$ and with $\|x\|_2$ so the mathematical expectation, if it exists in one case, exists in the other and with the same value.

The condition (1) can be interpreted qualitatively and is necessarily satisfied in the case of a finite-dimensional number (alongside the theorem E[U(X)]=U[E(X)], *cf.* page 5).

Property 1. – If E(X) exists, E(αX), where α is a fixed number, exists and equals αE(X)

$$\begin{aligned} x^*[E(\alpha X)] &= E[x^*(\alpha X)] = E[\alpha x^*(X)] = \alpha E[x^*(X)] \\ &= \alpha x^*[E(X)] = x^*[\alpha E(X)] \quad \text{for every } x^*, \end{aligned}$$

so E(αX) exists and E(αX)= αE(X).

Property 2. – If X is surely or almost surely equal to a fixed element x, E(X) exists and is equal to x.

$$x^*[\mathrm{E}(\mathrm{X})] = \mathrm{E}[x^*(\mathrm{X})] = x^*(x) \qquad \text{for every } x^*;$$

so E(X) exists and is equal to x.

Property 3. – If X is surely or almost surely equal to a fixed element x, if A is a numerical random variable and if E(A) exists, then E[AX] exists and is equal to xE[A].

$$\begin{aligned} x^*[\mathrm{E}(\mathrm{AX})] = \mathrm{E}[x^*(\mathrm{AX})] &= \mathrm{E}[\mathrm{A}x^*(\mathrm{X})] = x^*(x)[\mathrm{E}(\mathrm{A})] \\ &= x^*[x\mathrm{E}(\mathrm{A})] \quad \text{for every } x^*, \end{aligned}$$

so E(AX) exists and is equal to x[E(A)].

Property 4. – If X and Y are defined on the same $\mathscr{X}$ and if E(X) and E(Y) exist, E(X + Y) exists and we have:

$$\mathrm{E}(\mathrm{X}+\mathrm{Y}) = \mathrm{E}(\mathrm{X}) + \mathrm{E}(\mathrm{Y}).$$

LEMMA. – *Let us recall* [Banach, I, p.181] *that if $\mathscr{X}$ and $\mathscr{X}_1$ are Banach spaces and $x \in \mathscr{X}$ and $y \in \mathscr{X}_1$, if we denote by $\mathscr{X} \times \mathscr{X}_1$ the space of all ordered couples x, y where we define addition and multiplication by a scalar h by letting:*

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1),$$
$$h(x, y) = (hx, hy)$$

and the norm such that

(1)

$$\lim_{n\to\infty} x_n = x_0 \text{ and } \lim_{n\to\infty} y_n = y_0 \text{ is equivalent to } \lim_{n\to\infty} \|(x_n, y_n) - (x_0, y_0)\| = 0,$$

then $\mathscr{X} \times \mathscr{X}_1$ is a Banach space called the product of $\mathscr{X}$ and $\mathscr{X}_1$. (1) is fulfilled if, in particular, we take as the norm of $z = (x, y)$ one of the expressions

$$\|z\| = [\|x\|^p + \|y\|^p]^{\frac{1}{p}}$$

or

$$\|z\| = \max[\|x\|, \|y\|];$$

there are other possible norms, but by choosing any norm satisfying (1), we will always obtain isomorphic spaces. The product $\mathscr{X} \times \mathscr{X}$ is called the square of $\mathscr{X}$ and is denoted by $\mathscr{X}^2$.

We likewise define the product $\mathscr{X}_1 \times \mathscr{X}_2 \times \ldots \times \mathscr{X}_n$.

The product of a finite number of separable spaces(3) *is separable.*

Proof of property 4. – Let

$$\mathrm{X} \in \mathscr{X}, \qquad \mathrm{Y} \in \mathscr{Y} \equiv \mathscr{X}.$$

We assume that E(X) and E(Y) exist.

(3)We recall that a Banach space $\mathscr{X}$ is separable if there exists a countable sequence S of points of $\mathscr{X}$ such that every point of $\mathscr{X}$ is the limit of a partial sequence of S.

Let

$$\mathscr{Z} = \mathscr{X} \times \mathscr{Y} \qquad \text{and} \qquad (X, Y) = Z.$$

We have a measure on $\mathscr{Z}$, it induces a measure on $\mathscr{X}$ and another on $\mathscr{Y}$ that we assume are L-measures [since E(X) and E(Y) exist]. A linear functional applied to X + Y is of the form

$$x^*(\mathrm{X} + \mathrm{Y}) = x^*(\mathrm{X}) + x^*(\mathrm{Y}).$$

But $x^*(\mathrm{X})$ and $y^*(\mathrm{Y})$ [in other words, $x^*(\mathrm{Y})$] are measurable on $\mathscr{Z}$, so $x^*(\mathrm{X}+\mathrm{Y})$ is also measurable on $\mathscr{Z}$. But the measure on $\mathscr{Z}$ induces a measure on the space $\mathscr{X}'$ (identical to $\mathscr{X}$) of X + Y, and the latter is thus an L-measure. We can thus find out whether X + Y has a mathematical expectation E(X + Y), this mathematical expectation has to be such that

$$\begin{aligned} x^*[\mathrm{E}(\mathrm{X} + \mathrm{Y})] &= \mathrm{E}[x^*(\mathrm{X} + \mathrm{Y})] = \mathrm{E}[x^*(\mathrm{X}) + x^*(\mathrm{Y})], \\ &= \mathrm{E}[x^*(\mathrm{X})] + \mathrm{E}[x^*(\mathrm{Y})], \\ &= x^*[\mathrm{E}(\mathrm{X})] + x^*[\mathrm{E}(\mathrm{Y})], \\ &= x^*[\mathrm{E}(\mathrm{X}) + \mathrm{E}(\mathrm{Y})], \end{aligned}$$

and this holds for every $x^* \in \mathscr{X}^*$, so: E(X + Y) exists and E(X + Y) = E(X) + E(Y).

THEOREM. – *Let* U *be a linear operator defined in the Banach space* $\mathscr{X}$ *and whose codomain is a Banach space* $\mathscr{X}_1$*; letting* $\mathrm{X} \in \mathscr{X}$*, if* E(X) *exists,* E[U(X)] *exists and is equal to* U[E(X)].

LEMMA[4]. – *Let* $\mathscr{X}$ *and* $\mathscr{X}_1$ *be two Banach spaces,* $y = \mathrm{U}(x)$ *a linear operator defined in* $\mathscr{X}$ *whose codomain is contained in* $\mathscr{X}_1$ *and* x^* *and* y^* *linear functionals defined in* $\mathscr{X}$ *and* $\mathscr{X}_1$ *respectively. Let us consider the expression* $y^*[\mathrm{U}(x)]$*, where* y^* *is any linear functional defined on* $\mathscr{X}_1$*. This expression can be regarded as a functional defined on* $\mathscr{X}$*.*

Let us set

$$x^*(x) = y^*[\mathrm{U}(x)];$$

thus defined, the functional x^* *is additive and continuous, because we have*

$$|x^*(x)| = |y^*[\mathrm{U}(x)]| \leq \|y^*\| \cdot \|U\| \cdot \|x\|,$$

whence

$$\|x^*\| \leq \|y^*\| \cdot \|U\|.$$

Proof of theorem. – By definition,

$$x^*[\mathrm{E}(\mathrm{X})] = \mathrm{E}[x^*(\mathrm{X})];$$

E[X] exists, so $\mathrm{E}[x^*(\mathrm{X})]$ also exists for each x^*; by the lemma, for each $y^*[\mathrm{U}(\mathrm{X})]$ there is a $x^*(X)$ such that:

$$y^*[\mathrm{U}(\mathrm{X})] = x^*(\mathrm{X}),$$

(4) BANACH, *loc. cit.*, p.99.

so $E[y^*[U(X)]]$ exists and

$$E[y^*[U(X)] = E[x^*(X)],$$

$$\begin{aligned} y^*[E[U(X)]] &= E[y^*[U(X)]] = E[x^*(X)] = x^*[E(X)], \\ &= y^*[U[E(X)]] \end{aligned}$$

and this holds for every $y^* \in \mathscr{X}_1^*$ so:

$$E[U(X)] = U[E(X)].$$

Property 5. – If $\|X\|$ is measurable, if $E[\|X\|] < +\infty$ and if $E(X)$ exists, we have:

$$\|E(X)\| \leq E[\|X\|].$$

There exists [Banach, I, p.99], [E. Hille, I, Theorem 2.9.3] a linear functional $x_0^* \in \mathscr{X}^*$ such that:

$$(\alpha) \qquad \|x_0^*\| = 1,$$

$$(\beta) \qquad x_0^*[E(X)] = \|E(X)\|.$$

But

$$\begin{aligned} x_0^*[E(X)] &= E[x_0^*(X)], \\ |x_0^*[E(X)]| &= |E[x_0^*(X)]| \leq E|x_0^*(X)| \leq E[\|x_0^*\| \cdot \|X\|], \\ \|E(X)\| &\leq E[\|X\|]. \end{aligned}$$

The examination of this property poses the problem of studying the measurability of $\|X\|$.

Theorem. – *With an* L*-measure, if* $\mathscr{X}$ *is separable,* $\|x\|$ *is measurable.*

Indeed, an L-measure means that x is weakly Pettis-measurable, but $\mathscr{X}$ being separable means that x is Bochner-measurable [Pettis, I, p.279], that is to say, there exists a sequence of step functions[(5)] $\lambda_n(x)$ such that, for a fixed x, save for a set of zero measure,

$$\|x - \lambda_n(x)\| \to 0;$$

this implies

$$\|\lambda_n(x)\| \to \|x\|;$$

but $\|\lambda_n(x)\|$ is measurable, hence so is $\|x\|$. More generally, every continuous numerical function $f(x)$ of x is measurable; every function f with values in a Banach space and continuous will be Bochner-measurable.

Theorem. – *With an* L*-measure, if* $\mathscr{X}$ *is separable and reflexive and if* $E[\|X\|] = m < +\infty$*, then* $E(X)$ *exists.*

[(5)]"Step functions" are constant on each of a finite number of disjoint measurable sets whose sum is the entire space (*see* Pettis).

$\mathscr{X}$ is separable, so $\|X\|$ is measurable (preceding theorem). $\mathscr{X}$ being separable and reflexive implies that $\mathscr{X}^*$ is separable. Let $\{x_n^*\}$ be a countable dense sequence in $\mathscr{X}^*$; for $E(X) = y$ to exist, it is necessary and sufficient that

$$x_n^*(y) = E[x_n^*(X)] \qquad \text{for all } n. \tag{1}$$

Necessity is obvious. For sufficiency, if, for every n, $E[x_n^*(X)]$ exists, $E[x^*(X)]$ exists for every x^*; indeed:

$$E[x^*(X)] = E[x_n^*(X)] + E[(x^* - x_n^*)(X)],$$

and if $\|x^* - x_n^*\| \leq \varepsilon$:

$$|E[(x^* - x_n^*)(X)] \leq E[\|x^* - x_n^*\|\|X\|] \leq m\varepsilon.$$

So if y satisfies (1):

$$\begin{aligned} |E[x^*(X)] - x^*(y)| &\leq m\varepsilon + |x_n^*(y) - x^*(y)|, \\ &\leq \varepsilon(m + \|y\|). \end{aligned}$$

Here [E. Hille, I, p.21], a necessary and sufficient condition for (1) to have a solution y such that $\|y\| \leq M$ is that, for every y^* of the form

$$y^* = \sum_{n=1}^{k} \alpha_n x_n^*,$$

we have

$$\left| \sum_{n=1}^{k} \alpha_n E[x_n^*(X)] \right| \leq M \cdot \|y^*\|,$$

that is to say,

$$|E[y^*(X)]| \leq M \cdot \|y^*\|.$$

Now we have

$$|E[y^*(X)]| \leq E[\|y^*\| \cdot \|X\|] \leq \|y^*\| m.$$

In Chapter II, this theorem will be extended in the sense that the condition of $\mathscr{X}$ being *reflexive* will be dropped. We will then obtain this theorem as a consequence of the strong law of large numbers. It was, however, interesting to obtain a sufficient condition of the existence of E(X) at this stage.

Comparison with Fréchet's mathematical expectation

Fréchet's definition. – M. Fréchet [M. Fréchet, II] gave a constructive definition of the mathematical expectation of a random element. This mathematical expectation exists if two conditions are met. The first is about the measure. The measure has to satisfy a certain condition F, and if it does, we will say it is an F-measure; this condition is that, for any $\varepsilon > 0$, we can find a finite or countable number of sets $e_1, ..., e_k, ...$ such that:

α. $\sum_k e_k = \mathscr{X}$;

β. The e_k are pairwise disjoint and each e_k is measurable;

γ. For every e_k, the variation of x, when x varies in e_k, stays below ε in norm, that is to say, the diameter of each e_k is smaller than ε.

This condition implies that the space $\mathscr{X}$ is separable; and conversely, if $\mathscr{X}$ is separable, such a decomposition is possible. The condition γ was generalised by Fréchet himself and by Shafik Doss [S.Doss, I] and replaced by:

γ'. For every e_k of *finite measure* $m(e_k)$, the variation of x, as x varies in e_k, stays below ε in norm, that is to say, the diameter of each e_k such that $m(e_k) > 0$ is $< \varepsilon$, which is equivalent to saying that X is *almost surely* in a separable subset $\mathscr{X}_1$ of $\mathscr{X}$.

The second condition, condition F', for the existence of the mathematical expectation, is that there exists a value of ε, a choice of the e_k for this ε and a choice of ξ_k in e_k such that:

$$\sum_k \|\xi_k\| m(e_k) < +\infty.$$

This definition of the mathematical expectation is equivalent to the definition of integral of a function with values in a Banach space, it corresponds to the Bochner integral [Bochner, I], while the definition we gave corresponds to the Pettis integral [Pettis, I, p.277] which includes that of Bochner.

THEOREM 1. – *Every* F-*measure is an* L-*measure; more generally, if $f(x)$ is a* continuous and real numerical *function, the set $\mathscr{A}$ of the x for which $f(x) < a$ (for some real number a) is measurable (with the* F-*measure assumed to be given).*

1° $\mathscr{A}$ *is open.* – If $x_0 \in \mathscr{A}$, we can find a sphere with centre x_0 completely contained in $\mathscr{A}$; indeed, $x_0 \in \mathscr{A}$ implies that $a - f(x_0) > 0$.

Let $d = a - f(x_0)$. Since f is continuous, we can find η such that $\|x - x_0\| < \eta$ implies $|f(x) - f(x_0)| < \frac{d}{2}$, so the sphere with centre x_0 and radius η is included in $\mathscr{A}$.

2° $\mathscr{A}$ *is* F-*measurable.* – Let $\varepsilon_n = \frac{1}{n}$, and let e^n_k be the e_k for $\varepsilon = \varepsilon_n$; let us denote by e'^n_k the e^n_k which are contained in $\mathscr{A}$, and by e''^n_k the others.

Let $\mathscr{A}_n$ be the sum of the e'^n_k,

$$\mathscr{A}_n \subset \mathscr{A}.$$

Let $\mathscr{B}_n = \sum_{i=1}^n \mathscr{A}_i$ (the union of the $\mathscr{A}_i$),

$$\mathscr{B}_n \subset \mathscr{B}_{n+1}.$$

$\mathscr{A}_n$ *is* F-*measurable, hence so is* $\mathscr{B}_n$; denote by $(\overline{e''})$ the set of all the e^n_k, if they exist, which are of zero measure and have at least one point inside $\mathscr{A}$ and at least one point outside $\mathscr{A}$; there are at most countably many of them, and their union $\bar{e}$ is thus of zero measure.

Let us then set

$$\mathscr{A}' = \mathscr{A} - \mathscr{A}\bar{e} \qquad (\mathscr{A}' \subset \mathscr{A}).$$

If $x_0 \in \mathscr{A}'$, x_0 belongs to all the $\mathscr{B}_n$ as soon as n is larger than some number; indeed: for every n, there is a k, say k_n, such that $x_0 \in e^n_{k_n}$, then if n is large enough, every $e^n_{k_n}$ is an e'^n_k because if x varies in $e^n_{k_n}$,

$$\|x - x_0\| \leq \varepsilon_n = \frac{1}{n};$$

let η be the radius of a sphere with centre x_0 and completely contained in $\mathscr{A}$ (we saw that such an $\eta > 0$ exists); if $\frac{1}{n} < \eta$, $e^n_{k_n}$ is contained in the sphere, so contained in $\mathscr{A}$, and hence it is an e'^n_k. So

$$\mathscr{A}' \subset \mathscr{B} = \lim \mathscr{B}_n$$

(since every $x \in \mathscr{A}'$ is $\in \mathscr{B}_n$ for large enough n).

$$\mathscr{B}_n = \sum_{i=1}^{n} \mathscr{A}_i \subset \mathscr{A},$$

so

$$\mathscr{A}' \subset \mathscr{B} = \lim \mathscr{B}_n \subset \mathscr{A},$$

since $\mathscr{A} - \mathscr{A}'$ is of zero measure, $\mathscr{A} - \mathscr{B}$ is of zero measure. As $\mathscr{B}$ is measurable, $\mathscr{A}$ is also measurable, and $m(\mathscr{A}) = m(\mathscr{B})$.

This theorem can be immediately extended to complex functions $f(x)$ which are measurable if the real and imaginary parts are separately measurable.

The x^* are continuous numerical functions, so every F-measure is an L-measure.

Let us remark, moreover, that $\|x\|$ is also a continuous numerical function, so $\|x\|$ is F-measurable. In the case of an L-measure which is also an F-measure, $\|x\|$ is L-measurable; which we already knew, (c.f. page 6), because then $\mathscr{X}$ is separable. In the case of an F-measure, F' is a sufficient condition for $\mathrm{E}[\|x\|]$ to exist; it is also a necessary condition [M. Fréchet, II, p.494]. With an L-measure, we saw (page 6) that if $\mathscr{X}$ is not only separable but also reflexive, then $\mathscr{X}^*$ is separable, and the existence of $\mathrm{E}[\|\mathrm{X}\|] = m$ implies the existence of E(X).

Theorem 2. – *If we consider an* L*-measure which is an* F*-measure satisfying* F'*, the mathematical expectation of* X *exists and is equal to the mathematical expectation in the sense of Fréchet.*

Let $\mathrm{E_F(X)}$ be the mathematical expectation of X in the sense of Fréchet,

$$\mathrm{E_F(X)} = \lim_{\varepsilon \to 0} \sum_k m(e_k)\xi_k;$$

we have

$$\sum_k m(e_k)\|\xi_k\| < +\infty.$$

Let x^* be any linear functional (hence bounded):

$$\begin{aligned} x^*[\mathrm{E_F(X)}] &= x^* \left[\lim \sum_k m(e_k)\xi_k \right], \\ &= \lim \left[\sum_k m(e_k) x^*(\xi_k) \right], \\ &= \mathrm{E}[x^*(\mathrm{X})]. \end{aligned}$$

As a consequence, $\mathrm{E}[x^*(\mathrm{X})]$ exists and there exists an element $\mathrm{E_F(X)}$ such that

$$x^*[\mathrm{E_F(X)}] = \mathrm{E}[x^*(\mathrm{X})].$$

So E(X) exists and:

$$\boxed{\mathrm{E(X)} = \mathrm{E_F}(X).}$$

Chapter II.

Addition of Random Elements.

I. – Preliminaries.

We will always consider random elements X whose values belong to a Banach space $\mathscr{X}$ and such that for every linear functional – that is to say, additive and continuous, and hence bounded – x^*, $x^*(\mathrm{X})$ is measurable; we are now going to study the problem of the addition of "independent" random elements, or as we will say, "strictly stationary" random elements. We must thus define these terms; we will first recall the definition and some properties of the product space.

Product space. – If A and B are some sets (not necessarily subsets of the same space), the product $\mathrm{A} \times \mathrm{B}$ is the set of all the ordered couples x, y, where $x \in \mathrm{A}$ and $y \in \mathrm{B}$.

We defined (Chapter I, page 4) the product of two spaces $\mathscr{X}_1$ and $\mathscr{X}_2$, in particular the product of two Banach spaces.

We likewise define the product of a finite number of spaces $\mathscr{X}_1 \times \mathscr{X}_2 \times ... \times \mathscr{X}_n$, and, in particular, the product of a finite number of Banach spaces. The product of a finite number of separable spaces is separable.

Let $\mathscr{X}_1, \mathscr{X}_2, ..., \mathscr{X}_n$ be Banach spaces, and let $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2 \times ... \times \mathscr{X}_n$, and denote by $\mathscr{X}_1^*, ..., \mathscr{X}_n^*, \mathscr{X}^*$ the dual spaces of $\mathscr{X}_1, \mathscr{X}_2, ..., \mathscr{X}_n, \mathscr{X}$ respectively. The spaces $\mathscr{X}^*$ and $\mathscr{X}_1^* \times ... \times \mathscr{X}_n^*$ are isomorphic [Banach, I, p.192]. In particular, the dual of $\mathscr{X}^2$ is isomorphic to $(\mathscr{X}^*)^2$.

If, in addition to the spaces $\mathscr{X}_1$ and $\mathscr{X}_2$, we are given two σ-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$ of subsets of $\mathscr{X}_1$ and $\mathscr{X}_2$ respectively, we denote by $\mathscr{F}_1 \times \mathscr{F}_2$ the σ-algebra of subsets of $\mathscr{X}_1 \times \mathscr{X}_2$ generated by all the sets of the form $A_1 \times A_2$, where $A_1 \in \mathscr{F}_1$ and $A_2 \in \mathscr{F}_2$.

The measurable space $(\mathscr{X}_1 \times \mathscr{X}_2, \mathscr{F}_1 \times \mathscr{F}_2)$ is the product of two measurable spaces $(\mathscr{X}_1, \mathscr{F}_1)$ and $(\mathscr{X}_2, \mathscr{F}_2)$.

If μ_1 and μ_2 are two measures[(6)] defined on $\mathscr{F}_1$ and $\mathscr{F}_2$ respectively, there exists a unique measure λ such that, for every set $A_1 \times A_2 \in \mathscr{F}_1 \times \mathscr{F}_2$:

$$\lambda(A_1 \times A_2) = \mu_1(A_1) \times \mu(A_2).$$

λ is called the product of the measures μ_1 and μ_2, and we denote it by:

$$\lambda = \mu_1 \times \mu_2.$$

Independent random elements. – In the classical theory of Probability Calculus, we say that two numerical random variables X_1 and X_2 are independent if, for any x_1 and x_2:

$$\Pr[\mathrm{X}_1 < x_1; \mathrm{X}_2 < x_2] = \Pr[\mathrm{X}_1 < x_1] \times \Pr[\mathrm{X}_2 < x_2],$$

[(6)] We will consider measures which are probability measures, that is to say, such that $\mu(\mathscr{X}) = 1$. The theorem is true if μ_1 and μ_2 are any "σ-finite" measures [P.R.Halmos, I, p.144].

which is equivalent to saying that the events $X_1 < x_1$ and $X_2 < x_2$ are independent for any x_1 and x_2, that is to say, that the probability that $X_2 < x_2$ is not modified by the knowledge that the event $X_1 < x_1$ is realised, or put yet another way, that $\Pr[X_2 < x_2/X_1 < x_1] = \Pr[X_2 < x_2]$.

Let us now consider random elements X_1 and X_2 taking values in Banach spaces $\mathscr{X}_1$ and $\mathscr{X}_2$ respectively. We say again that these random elements are independent if the fact that one has information on the values taken by one of them does not modify the probability law of the second. Without studying when and how it will be possible to define conditional probabilities, the immediate generalisation of the definition of independence of two random variables gives a definition of random elements.

Let μ_1, μ_2, λ be measures defined on $\mathscr{F}_1$, $\mathscr{F}_2$ and $\mathscr{F}_1 \times \mathscr{F}_2$ respectively, and such that

$$\begin{aligned} \Pr[X_1 \in A_1] &= \mu_1(A_1) \qquad \text{for every } A_1 \in \mathscr{F}_1, \\ \Pr[X_2 \in A_2] &= \mu_2(A_2) \qquad \text{for every } A_2 \in \mathscr{F}_2, \\ \Pr[X_1 \in A_1; X_2 \in A_2] &= \lambda(A_1 \times A_2). \end{aligned}$$

Definition. – Two random elements X_1 and X_2 are independent if, for any $A_1 \in \mathscr{F}_1$ and $A_2 \in \mathscr{F}_2$:

$$\boxed{\Pr[X_1 \in A_1; X_2 \in A_2] = \Pr[X_1 \in A_1] \times \Pr[X_2 \in A_2]}$$

or

$$\boxed{\lambda(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2),}$$

that is to say, if, in the product space, the measure is the product of the measures.

The definition immediately extends to any number of random elements: $X_1, X_2, ..., X_n$ are mutually independent if

$$\lambda[A_1 \times A_2 \times ... \times A_n] = \mu_!(A_1) \times ... \times \mu_n(A_n), \quad \text{for } A_1 \times ... \times A_n \in \mathscr{F}_1 \times ... \times F_n.$$

Strictly stationary sequence. – By analogy with the case of random variables, we will say that random elements X_n form a strictly stationary sequence if, for any positive integer s, any integers $n_1, n_2, ..., n_s$ and any integer h, the probability law of the random element $X_{n_1+h}, X_{n_2+h}, ..., X_{n_s+h}$ does not depend on h.

II. – Strong law of large numbers with respect to weak convergence.

Theorem 1. – *If $\mathscr{X}^*$ is separable, if $X_1, X_2, ..., X_i, ...$ is an infinite sequence of mutually independent random elements in $\mathscr{X}$ with the same law, if $E[\|X_i\|] = M < +\infty$[7] and if $E[X_i]$ exists, – we can then, without loss of generality, assume $E[X_i] = 0$ –*

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

[7] $\mathscr{X}^*$ being separable means $\mathscr{X}$ is separable and, as a consequence, $\|X\|$ is measurable.

converges weakly almost surely to 0 when n converges to infinity.

As $\mathscr{X}^*$ is separable, let $\{x_k^*\}$ be a dense sequence in $\mathscr{X}^*$:

$$x_k^*(\mathrm{Y}_n) = \frac{1}{n}\sum_{i=1}^{n} x_k^*(\mathrm{X}_i);$$

$x_k^*(\mathrm{X}_i)$ is a numerical random variable with zero mathematical expectation; in fact, just from the definition:

$$\mathrm{E}[x_k^*(\mathrm{X}_i)] = x_k^*[\mathrm{E}(\mathrm{X}_i)].$$

But $\mathrm{E}(\mathrm{X}_i) = 0$, so

$$x_k^*[\mathrm{E}(\mathrm{X}_i)] = x_k^*(0) = 0,$$

so

$$\frac{1}{n}\sum_{i=1}^{n} x_k^*(\mathrm{X}_i) \to 0$$

almost surely for any given k (classical law of large numbers).

So it is almost sure that for every k, $x_k^*(\mathrm{Y}_n)$ converges to zero.

If x^* is any linear functional in $\mathscr{X}^*$, we have

$$x^* = x_k^* + y_k^*$$

and k can be chosen such that $\|y_k^*\|$ is as small as we want, since $\mathscr{X}^*$ is separable.

$$x^*(\mathrm{Y}_n) = x_k^*(\mathrm{Y}_n) + y_k^*(\mathrm{Y}_n),$$
$$|x^*(\mathrm{Y}_n)| \leq |x_k^*(\mathrm{Y}_n) + \|y_k^*\| \cdot \|\mathrm{Y}_n\|.$$

But here,

$$\|\mathrm{Y}_n\| \leq \frac{1}{n}\sum_{i=1}^{n} \|\mathrm{X}_i\|$$

and:

$$\frac{1}{n}\sum_{i=1}^{n} \|\mathrm{X}_i\| \quad \text{converges almost surely to} \quad \mathrm{E}[\|\mathrm{X}_i\|] = M$$

(the classical law of large numbers).

Hence, on an event of probability 1,

$$|x^*(\mathrm{Y}_n)| \leq |x_k^*(\mathrm{Y}_n) + \|x^* - x_k^*\| \cdot (\mathrm{M} + \varepsilon),$$

for any given $\varepsilon > 0$. It suffices to take k such that

$$\|x^* - x_k^*\| < \frac{\varepsilon}{2\mathrm{M}}$$

and then n large enough such that

$$|x_k^*(\mathrm{Y}_n)| < \frac{\varepsilon}{2}$$

to have

$$|x^*(\mathrm{Y}_n) \leq \varepsilon, \qquad \text{so } x^*(\mathrm{Y}_n) \to 0,$$

so Y_n converges weakly almost surely to 0.

THEOREM 2. – *If* $\mathscr{X}$ *is* separable *and* reflexive, *if* $..., \mathrm{X}_1, ..., \mathrm{X}_i, ...$ *form a strictly stationary infinite sequence in two directions – which contains the particular case where* X_i *are independent and identically distributed – if* $\mathrm{E}[\|\mathrm{X}_i\|] < +\infty$ *– and so* $\mathrm{E}(\mathrm{X}_i)$ *exists – :*

$$\mathrm{Y}_n = \frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i$$

converges weakly almost surely to a limit Y *as* n *converges to infinity.*

For every given x^*, the sequence of numerical random variables $x^*(\mathrm{X}_i)$ is stationary and $\mathrm{E}|x^*(\mathrm{X}_i)|$ exists, because

$$|x^*(\mathrm{X}_i)| \le \|x^*\| \cdot \|\mathrm{X}_i\|$$

and by the hypothesis that $\mathrm{E}[\|\mathrm{X}_i\|]$ exists; so, by Birkhoff's ergodicity theorem,

$$\frac{1}{n}\sum_{i=1}^{n} x^*(\mathrm{X}_i) = x^*\left[\frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i\right] = x^*(\mathrm{Y}_n)$$

converges almost surely to a limit $\mathscr{L}(x^*)$.

On the other hand,

$$\|\mathrm{Y}_n\| \le \frac{1}{n}\sum_{i=1}^{n}\|\mathrm{X}_i\|.$$

The sequence of random variables $\|\mathrm{X}_i\|$ is stationary, and $\mathrm{E}[\|\mathrm{X}_i\|]$ exists so $\frac{1}{n}\sum_{i=1}^{n}\|\mathrm{X}_i\|$ converges almost surely to a limit, so almost surely stays bounded; so on almost every outcome u, $\|\mathrm{Y}_n\|$ stays bounded.

As $\mathscr{X}$ is separable and reflexive, so is $\mathscr{X}^*$; let $\{x_j^*\}$ be a *dense* countable sequence in $\mathscr{X}^*$; according to above, it is almost sure that all the $x_j^*(\mathrm{Y}_n)$ converges to a limit $\mathscr{L}(x_j^*)$.

Let x^* be arbitrary in $\mathscr{X}^*$; there exists x_j^* such that $\|x^* - x_j^*\| < \varepsilon$; we have

$$x^*(\mathrm{Y}_n) = x_j^*(\mathrm{Y}_n) + [x^* - x_j^*](\mathrm{Y}_n). \tag{1}$$

Save for an event of probability zero, $x_j^*(\mathrm{Y}_n)$ has a limit, for any j, and $\|\mathrm{Y}_n\|$ is bounded, independently of j; as

$$|[x^* - x_j^*](\mathrm{Y}_n)| \le \varepsilon\|\mathrm{Y}_n\|$$

and as ε can be arbitrarily small, $x^*(\mathrm{Y}_n)$ converges.

So it is almost sure that all the $x^*(\mathrm{Y}_n)$ simultaneously has a limit $\mathscr{L}(x^*)$. To prove that Y_n converges weakly almost surely, we need to show moreover that there exists a $y \in \mathscr{X}$, which is random according to the outcome considered, such that

$$x^*(y) = \mathscr{L}(x^*) \qquad \text{for every } x^*.$$

But for the outcome in consideration [$\mathscr{L}(x^*)$ depends on this outcome], $\mathscr{L}(x^*)$ is an additive functional of x^* on $\mathscr{X}^*$, which is obvious. It is, moreover, a continuous functional; to show this, we need to prove that

$$|\mathscr{L}(x^*)| \to 0 \qquad \text{if } \|x^*\| \to 0.$$

This is true if x^* is an x_j^*, because:

$$\begin{aligned}|\mathscr{L}(x_j^*)| &= \lim_{n\to\infty}|x_j^*(\mathrm{Y}_n)| \le \limsup_{n\to\infty}|x_j^*(\mathrm{Y}_n)| \\ &\le \|x_j^*\| \limsup_{n\to\infty}\|\mathrm{Y}_n\|\end{aligned}$$

and we saw that $\limsup\|\mathrm{Y}_n\|$ is bounded, independently of j.

For an arbitrary x^*, we have

$$|\mathscr{L}(x^*)| \le \limsup_n|x_j^*(\mathrm{Y}_n)| + \varepsilon\|\mathrm{Y}_n\|,$$

by (1).

We just saw that $\limsup|x_j^*(y_n)|$ converges to zero if $\|x_j^*\| \to 0$, which is the case if $\|x^*\| \to 0$; as ε was arbitrarily small, $\mathscr{L}(x^*) \to 0$ as $\|x^*\| \to 0$.

So $\mathscr{L}(x^*)$ is a linear functional on $\mathscr{X}^*$ and then as $\mathscr{X}$ is reflexive, for every linear functional $\mathscr{L}(x^*)$ on $\mathscr{X}^*$ there exists a $y \in \mathscr{X}$ such that

$$\mathscr{L}(x^*) = x^*(y) \quad \text{for every } x^*,$$

so Y_n converges almost surely weakly to a limit Y.

Remark. – We saw that when the X_i are *independent*, $\mathscr{L}(x^*) = \mathrm{E}[x^*(\mathrm{X}_i)]$ for a given x^* is a fixed number; but when X_i form some stationary sequence for a given x^*, $\mathscr{L}(x^*)$ is a random variable.

III. – Study of an auxiliary space $\overset{\alpha}{\mathscr{X}}$.

Let $\mathscr{X}$ be any Banach space. Let us consider random elements X, taking their values in $\mathscr{X}$, such that $x^*(\mathrm{X})$ is measurable for any x^* – that is to say, the random elements are defined by an L-measure on $\mathscr{X}$ – additionally satisfying the condition C_α:

Condition C_α. – $\|\mathrm{X}\|$ is measurable, and for a given real number $\alpha \ge 1$, $\mathrm{E}(\|\mathrm{X}\|^\alpha) < +\infty$ (which implies that X is a random element in the proper sense).

We associate to $\mathscr{X}$ the space $\overset{\alpha}{\mathscr{X}}$ defined in the following way:

Every random element X whose values belong to a space $\mathscr{X}$ satisfying the above conditions will be considered as a point denoted by the same letter X with the letter a on top: $\overset{\alpha}{\mathrm{X}}$, of a normed space $\overset{\alpha}{\mathscr{X}}$, by letting:

$$\left\|\overset{\alpha}{\mathrm{X}}\right\| = [\mathrm{E}(\|\mathrm{X}\|^\alpha)]^{\frac{1}{\alpha}}. \tag{1}$$

The zero element of $\overset{\alpha}{\mathscr{X}}$ is $\overset{\alpha}{0}$ corresponding to X almost surely equal to 0.

Addition and multiplication by a scalar will be defined in $\overset{\alpha}{\mathscr{X}}$ in the following way:

1° If k is a number and if X defines $\overset{\alpha}{\mathrm{X}}$, $k\mathrm{X}$ defines $k\,\overset{\alpha}{\mathrm{X}}$.

2° If X_1 and X_2 define $\overset{\alpha}{\mathrm{X}}_1$ and $\overset{\alpha}{\mathrm{X}}_2$ respectively, $\mathrm{X}_1 + \mathrm{X}_2$ defines $\overset{\alpha}{\mathrm{X}}_1 + \overset{\alpha}{\mathrm{X}}_2$.

As $\mathscr{X}$ is a Banach space, $k\mathrm{X}$ and $\mathrm{X}_1 + \mathrm{X}_2$ are well-defined.

Finally, $\overset{\alpha}{\mathscr{X}}$ will be metrised by letting

$$d(\overset{\alpha}{\mathrm{X}}_1, \overset{\alpha}{\mathrm{X}}_2) = \left\| \overset{\alpha}{\mathrm{X}}_1 - \overset{\alpha}{\mathrm{X}}_2 \right\|.$$

With this definition, $\overset{\alpha}{\mathscr{X}}$ is hence vectorial and metric, and we will see that (1) does constitute a norm; indeed:

1° $\left\| \overset{\alpha}{\mathrm{X}} \right\|$ is a positive real number (because, of course, it is the positive value that will be taken in $[\mathrm{E}(\|\mathrm{X}\|^\alpha)]^{\frac{1}{\alpha}}$);

2° $\left\| \overset{\alpha}{\mathrm{X}} \right\| = 0$ if and only if $\|\mathrm{X}\| = 0$ almost surely, that is to say, if $\mathrm{X} = 0$ almost surely, so if $\overset{\alpha}{\mathrm{X}} = \overset{\alpha}{0}$;

3° $\left\| k \overset{\alpha}{\mathrm{X}} \right\| = [\mathrm{E}(\|k\mathrm{X}\|^\alpha)]^{\frac{1}{\alpha}} = |k|[\mathrm{E}(\|\mathrm{X}\|^\alpha)]^{\frac{1}{\alpha}} = |k| \cdot \left\| \overset{\alpha}{\mathrm{X}} \right\|$;

4° $\left\| \overset{\alpha}{\mathrm{X}}_1 + \overset{\alpha}{\mathrm{X}}_2 \right\| \leq \left\| \overset{\alpha}{\mathrm{X}}_1 \right\| + \left\| \overset{\alpha}{\mathrm{X}}_2 \right\|$, because

$$\left\| \overset{\alpha}{\mathrm{X}}_1 + \overset{\alpha}{\mathrm{X}}_2 \right\| = [\mathrm{E}(\|\mathrm{X}_1 + \mathrm{X}_2\|^\alpha)]^{\frac{1}{\alpha}};$$

here,

$$\|\mathrm{X}_1 + \mathrm{X}_2\| \leq \|\mathrm{X}_1\| + \|\mathrm{X}_2\|,$$

so

$$\mathrm{E}[\|\mathrm{X}_1 + \mathrm{X}_2\|] \leq \mathrm{E}[\|\mathrm{X}_1\|] + \mathrm{E}[\|\mathrm{X}_2\|]$$

and so

$$(\mathrm{E}[\|\mathrm{X}_1 + \mathrm{X}_2\|^\alpha])^{\frac{1}{\alpha}} \leq (\mathrm{E}[\|\mathrm{X}_1\|^\alpha])^{\frac{1}{\alpha}} + (\mathrm{E}[\|\mathrm{X}_2\|^\alpha])^{\frac{1}{\alpha}} \qquad \text{[M.Fréchet, III]}.$$

Finally, $\overset{\alpha}{\mathscr{X}}$ is *complete*, that is to say, that if we have sequence of $\overset{\alpha}{\mathrm{X}}_n$ satisfying the Cauchy condition:

$$(2) \qquad \left\| \overset{\alpha}{\mathrm{X}}_{n+p} - \overset{\alpha}{\mathrm{X}}_n \right\| = [\mathrm{E}(\|\mathrm{X}_{n+p} - \mathrm{X}_n\|^\alpha)]^{\frac{1}{\alpha}} \to 0, \qquad \text{with } \frac{1}{n}$$

uniformly in p, there exists an $\overset{\alpha}{\mathrm{X}}$ in $\overset{\alpha}{\mathscr{X}}$ such that

$$\left\| \overset{\alpha}{\mathrm{X}}_n - \overset{\alpha}{\mathrm{X}} \right\| \to 0.$$

In the course of the proof, we will use the following lemma:

FATOU'S LEMMA. – *If numerical random variables* X_n *are positive or zero, if* $\mathrm{Y} = \liminf_{n\to+\infty} \mathrm{X}_n$ *and if* $\liminf_{n\to+\infty} \mathrm{E}(\mathrm{X}_n) = \mathrm{M} < +\infty$, $\mathrm{E}(\mathrm{Y})$ *exists and is* $\leq \mathrm{M}$ [P.R.Halmos, I, p.113].

1° *Construction of* $\overset{\alpha}{\mathrm{X}}$. – Let s_k be numbers such that

$$s_k > 0, \qquad \sum_k s_k < +\infty$$

and ε_k positive numbers such that

$$\sum_k \frac{\varepsilon_k}{s_k^\alpha} < +\infty.$$

By virtue of (2), we can define a sequence of increasing integers n_k such that

$$\text{(3)} \qquad \mathrm{E}(\|\mathrm{X}_{n_k+p} - X_{n_k}\|^\alpha) < \varepsilon_k \quad \text{for all } p,$$

in particular,

$$\mathrm{E}(\|\mathrm{X}_{n_k+1} - \mathrm{X}_{n_k}\|^\alpha) < \varepsilon_k,$$

and by Bienaymé's inequality,

$$\Pr(\|\mathrm{X}_{n_k+1} - \mathrm{X}_{n_k}\| < s_k) > 1 - \frac{\varepsilon_k}{s_k^\alpha}.$$

By the Borel-Cantelli theorem, almost surely, the events

$$\|\mathrm{X}_{n_k+1} - \mathrm{X}_{n_k}\| < s_k$$

are realised for all sufficiently large values of k (because $\sum_k \frac{\varepsilon_k}{s_k^\alpha} < +\infty$).

As $\mathscr{X}$ is complete and $\sum_k s_k < +\infty$, X_{n_k} has a limit X as $k \to +\infty$; we have

$$\|\mathrm{X}_{n_k}\| \to \|\mathrm{X}\|.$$

The $\mathrm{E}(\|\mathrm{X}_{n_k}\|^\alpha)$ are bounded, because the $\overset{\alpha}{\mathrm{X}}_n$ are bounded, so, by Fatou's lemma, $\mathrm{E}(\|\mathrm{X}\|^\alpha) < +\infty$; moreover, $x^*(\mathrm{X})$ is the almost sure limit of $x^*(\mathrm{X}_{n_k})$, so $x^*(\mathrm{X})$ is measurable. So X defines an $\overset{\alpha}{\mathrm{X}}$ in $\overset{\alpha}{\mathscr{X}}$.

2° By passing to the limit in (3) ($p \to +\infty$), we have (by Fatou's lemma),

$$\text{(4)} \qquad \mathrm{E}(\|\mathrm{X} - \mathrm{X}_{n_k}\|^\alpha) < \varepsilon_k \quad \text{or} \quad \left\|\overset{\alpha}{\mathrm{X}} - \overset{\alpha}{\mathrm{X}}_{n_k}\right\| \leq (\varepsilon_k)^{\frac{1}{\alpha}}.$$

Studying $\left\|\overset{\alpha}{\mathrm{X}}_n - \overset{\alpha}{\mathrm{X}}\right\|$, we have

$$\left\|\overset{\alpha}{\mathrm{X}}_n - \overset{\alpha}{\mathrm{X}}\right\| \leq \left\|\overset{\alpha}{\mathrm{X}}_n - \overset{\alpha}{\mathrm{X}}_{n_k}\right\| + \left\|\overset{\alpha}{\mathrm{X}}_{n_k} - \overset{\alpha}{\mathrm{X}}\right\|.$$

Let us take $n_k > n$; if $n \to +\infty$, $\left\|\overset{\alpha}{\mathrm{X}}_{n_k} - \overset{\alpha}{\mathrm{X}}\right\| \to 0$ by (4) and also $\left\|\overset{\alpha}{\mathrm{X}}_{n_k} - \overset{\alpha}{\mathrm{X}}_n\right\| \to 0$ by hypothesis, so $\overset{\alpha}{\mathrm{X}}_n \to \overset{\alpha}{\mathrm{X}}$.

Q.E.D.

We can therefore conclude:

Theorem. – Under the conditions specified above, $\overset{\alpha}{\mathscr{X}}$ is a Banach space ($\alpha \geq 1$).

Definition. – An element $\overset{\alpha}{\mathrm{X}}$ of $\overset{\alpha}{\mathscr{X}}$ will be called countable if the corresponding X only takes countably many distinct values $x_1, x_2, ..., x_i, ...$ and if the event $\mathrm{X} = x_i$ can be given a probability.

The following terminology will be employed (measure theoretic instead of probability theoretic): u denotes an outcome; e a set of outcomes; $\mathscr{U}$ the set of all possible outcomes; $\mu(e)$ =probability that the realised outcome belongs to e; $\mu(e)$ is a measure on $\mathscr{U}$. If e_i is the set of u for which $X = x_i$, saying that the event $X = x_i$ can be given a probability is equivalent to saying that e_i is measurable (with respect to μ); a random element X in $\mathscr{X}$ can be defined as a function in u; if, at each u [except perhaps for u in a set e with $\mu(e) = 0$] we associate an x, say $x(u)$, of $\mathscr{X}$, this defines an X: the probability that X belongs to a set $h \subset \mathscr{X}$ is the measure μ of the set e of u for which $x(u) \in h$; for X to define an $\overset{\alpha}{X}$, we obviously need $x(u)$ to have suitable properties: that $x^*[x(u)]$ is measurable and

$$\int \|x(u)\|^\alpha d\mu < +\infty.$$

If the $x(u)$ only take countably many distinct values $x_1, ..., x_i, ...$ and if the set e_i of u such that $x(u) = x_i$ is measurable for any i, $x^*[x(u)]$ is automatically measurable; it then suffices to have

$$\sum_i \mu(e_i)\|x_i\|^\alpha = \int \|x(u)\|^\alpha d\mu < +\infty$$

for X to define a countable $\overset{\alpha}{X}$.

Let us suppose that $\mathscr{X}$ is separable; let $\{x_j\}$ be a countable dense sequence in $\mathscr{X}$; for an arbitrary $\varepsilon > 0$, let A_j be the set of x such that

$$\|x - x_j\| \leq \varepsilon$$

and let B_j be the set defined by

$$B_1 = A_1, \quad B_j = A_j - A_j(A_1 + A_2 + ... + A_{j-1}) \quad (j \geq 2).$$

The B_j are disjoint and naturally

$$\sum_j B_j = \sum_j A_j = \mathscr{X}.$$

For $\overset{\alpha}{X}$ (X), let $\overset{\alpha}{X'}$ (X′) be defined in the following way:

$$X' = x_j \qquad \text{when} \qquad X \in B_j.$$

X′ only takes countably many values; moreover, $\|X - x_j\|$ is measurable, that is to say that $\Pr[X \in A_j]$ exists, so $\Pr[X \in B_j]$ exists, so the set e_j of u for which $X' = x_j$ is measurable. Furthermore, we always have:

$$\|X' - X\| \leq \varepsilon,$$

so

$$E(\|X' - X\|^\alpha) \leq \varepsilon^\alpha,$$

which proves that $E(\|X'\|^\alpha) < +\infty$, that $\overset{\alpha}{X'}$ is countable and also that $\overset{\alpha}{X'}$ is *arbitrarily close to* $\overset{\alpha}{X}$ in $\overset{\alpha}{\mathscr{X}}$.

Study of linear functionals on $\overset{\alpha}{\mathscr{X}}$ when $\mathscr{X}$ is separable and reflexive, $\alpha > 1$. – Let an $\overset{\alpha}{\mathrm{X}}$ (X) be equivalent to a function $x(u)$ of u; to each u let us associate x_u^* as a function of u (which is equivalent to defining a random element X^* in $\mathscr{X}^*$); let us assume that x_u^* satisfies:

1° $x_u^*(x)$ for each fixed $x \in \mathscr{X}$ is a measurable function; this means, as we will see (page 21), that $\|x_u^*\|$ is measurable.
2° $\int \|x_u^*\|^{\frac{\alpha}{\alpha-1}} d\mu < +\infty$.

We then easily prove that $x_u^*[x(u)]$ is measurable, for every $\overset{\alpha}{\mathrm{X}} \in \overset{\alpha}{\mathscr{X}}$ and that

$$\int |x_u^*[x(u)]| d\mu < +\infty$$

by virtue of Hölder's inequality [F.Riesz, I, p.44].
So:

$$\int x_u^*[x(u)] d\mu = \mathrm{E}[\mathrm{X}^*(\mathrm{X})] \tag{5}$$

exists for every $\overset{\alpha}{\mathrm{X}} \in \overset{\alpha}{\mathscr{X}}$.

It is easy to see that it is a linear functional in $\overset{\alpha}{\mathrm{X}}$, the proof involves *only* the fact that $\mathscr{X}$ is *separable*.

Conversely, if $\mathscr{X}$ is *separable and reflexive*, every linear functional $\overset{\alpha}{\mathrm{X}}$ on $\overset{\alpha}{\mathscr{X}}$ is of the form (5). Indeed:

A. Let us take an $\overset{\alpha}{\mathrm{X}^*}$; let us consider an $\overset{\alpha}{\mathrm{X}}$ (X) such that $\mathrm{X} = x$ for $u \in e$, where e is any measurable set, and $\mathrm{X} = 0$ for $u \notin e$ (it is clear that such an X defines an $\overset{\alpha}{\mathrm{X}}$).

For a fixed e and varying x, $\overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}\right)$ is a numerical function in x; it is evidently an additive function, it is also a continuous function, because

$$\left|\overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}\right)\right| \le \left\|\overset{\alpha}{\mathrm{X}^*}\right\| \cdot \left\|\overset{\alpha}{\mathrm{X}}\right\| = \left\|\overset{\alpha}{\mathrm{X}^*}\right\| \cdot [\mathrm{E}(\|\mathrm{X}\|^\alpha)]^{\frac{1}{\alpha}} = \left\|\overset{\alpha}{\mathrm{X}^*}\right\| \cdot [\mu(e)]^{\frac{1}{\alpha}} \|x\|, \tag{6}$$

so if $\|x\| \to 0$, $\left|\overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}\right)\right| \to 0$, so $\overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}\right)$ is a continuous and additive functional, with $x^*(e)$ belonging to $\mathscr{X}^*$; naturally $x^*(e)$ depends on e; $x^*(e;x)$ will denote the number obtained by applying $x^*(e)$ to $x \in \mathscr{X}$. Thus, there exists a function of sets $x^*(e)$, with values in $\mathscr{X}^*$, defined for every measurable e: $x^*(e)$ *is an additive set function.*

Let e_1 and e_2 be disjoint; let x_1 and x_2 be two points in $\mathscr{X}$; let $\mathrm{X}_1 = x_1$ for $u \in e_1$, $\mathrm{X}_1 = 0$ for $u \notin e_1$; $\mathrm{X}_2 = x_2$ for $u \in e_2$ and $\mathrm{X}_2 = 0$ for $u \notin e_2$. $\mathrm{X} = \mathrm{X}_1 + \mathrm{X}_2$ defines $\overset{\alpha}{\mathrm{X}}$, X_1 and X_2 define $\overset{\alpha}{\mathrm{X}}_1$ and $\overset{\alpha}{\mathrm{X}}_2$, and

$$\overset{\alpha}{\mathrm{X}} = \overset{\alpha}{\mathrm{X}}_1 + \overset{\alpha}{\mathrm{X}}_2,$$

so

$$\overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}\right) = \overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}_1\right) + \overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{X}}_2\right)$$

and from above,

$$\overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{X}}\right) = x^*(e_1; x_1) + x^*(e_2; x_2). \tag{7}$$

But let us now take $x = x_1 = x_2$; then

$$\overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{X}}\right) = x^*(e_1 + e_2; x) \quad \text{[definition of } x^*(e)\text{]}$$

and by (7),

$$x^*(e_1 + e_2; x) = x^*(e_1; x) + x^*(e_2; x) \quad \text{for any } x,$$

which is to say

$$x^*(e_1 + e_2) = x^*(e_1) + x^*(e_2);$$

$x^*(e)$ is a completely additive set function.

Let $e_1, ..., e_j, ...$ be a countable family of pairwise disjoint sets and $x_1, ..., x_j, ...$ a sequence of points in $\mathscr{X}$; let X_j be the random element which takes the value x_j if $u \in e_j$ and 0 if $u \notin e_j$. Then

$$\overset{\alpha}{\mathrm{Y}}_n = \overset{\alpha}{\mathrm{X}}_1 + ... + \overset{\alpha}{\mathrm{X}}_n \quad \text{belongs to } \overset{\alpha}{\mathscr{X}}.$$

Let $\overset{\alpha}{\mathrm{X}}$ (X) be defined by $\mathrm{X} = \sum_j \mathrm{X}_j$, if we take the x_j such that

$$\sum_j \mu(e_j) \cdot \|x_j\|^{\alpha} = \mathrm{E}(\|\mathrm{X}\|^{\alpha}) < +\infty.$$

Then $\overset{\alpha}{\mathrm{X}}$ belongs to $\overset{\alpha}{\mathscr{X}}$, it is countable and

$$\overset{\alpha}{\mathrm{X}} = \lim_{n \to +\infty} \overset{\alpha}{\mathrm{Y}}_n,$$

$$\overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{X}}\right) = \lim_{n \to +\infty} \overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{Y}}_n\right) = \lim_{n \to +\infty} \sum_{j=1}^{n} x^*(e_j; x_j),$$

by above.

Choose x_j to be the summit of $x^*(e_j)$, i.e.

$$\|x_j\| = 1 \quad \text{and} \quad x^*(e_j; x_j) = \|x^*(e_j)\|.$$

such a summit exists if $\mathscr{X}$ is *reflexive*. We then have

$$\overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{Y}}_n\right) = \sum_{j=1}^{n} \|x^*(e_j)\|.$$

So the series with positive or zero terms $\sum_j \|x^*(e_j)\|$ converges. Let us now take the x_j to be all equal to some x. Then

$$\overset{\alpha}{\mathrm{X}}{}^{*}\left(\overset{\alpha}{\mathrm{X}}\right) = x^*\left(\sum_j e_j; x\right) \qquad \text{[definition of } x^*(e)\text{]}$$

and by above,

$$x^*\left(\sum_j e_j; x\right) = \sum_j x^*(e_j; x),$$

where the series on the right-hand side is absolutely convergent. Now this implies

$$x^*\left(\sum_j e_j\right) = \sum_j x^*(e_j), \quad \text{with } \sum_j \|x^*(e_j)\| < +\infty.$$

B. *Reminder of the Radon-Nikodym theorem* [P.R.Halmos, I, p.128]. – If $\mathrm{L}(e)$ is a bounded, completely additive and absolutely continuous numerical set function, there exists a numerical function $\lambda(u)$ such that we have, for every measurable e,

$$\mathrm{L}(e) = \int_e \lambda(u) d\mu.$$

$\lambda(u)$ is finite almost everywhere, and it is clear that we can change it on a set of measure 0 without any harm (which would prevent the proofs if $\mathscr{X}$ is not separable).

By A, $x^*(e)$ and, as a result, $x^*(e; x)$ for any fixed x, is a *completely additive* set function of e, and by (6), *absolutely continuous* [$\|x^*(e)\|$ and $|x^*(e; x)| \to 0$ if $\mu(e) \to 0$ for any fixed x].

By the Radon-Nikodym theorem, for any x, there exists a measurable numerical function in u, and obviously a function in x, $\mathrm{K}(u; x)$ such that for any measurable e, we have

$$x^*(e; x) = \int_e \mathrm{K}(u; x) d\mu;$$

As $\mathscr{X}$ is separable, let (x_j) be a dense countable sequence on a sphere of radius 1 in $\mathscr{X}$; assuming the numbers a_j appearing henceforth to be rational, every point of the form $\sum_{j=1}^k a_j x_j$ (with k finite) is said to be an x', and the set $\mathscr{X}'$ of the x' is *countable and dense* in $\mathscr{X}$; we can assume the x_j to be enumerated in a sequence $\{x'_j\}$.

Let $\mathrm{K}_j(u) = \mathrm{K}(u; x_j)$; for every $x \in \mathscr{X}'$, so for every x of the form $\sum_{j=1}^k a_j x_j$, let us set

$$h(u; x) = \sum_{j=1}^k a_j \mathrm{K}_j(u);$$

for every fixed u, $h(u; x)$ is obviously an additive functional in x, defined for $x \in \mathscr{X}$ and possibly infinite, but only for the u in a set of measure 0 and independent of x [it is the set of values of u for which some of the $\mathrm{K}_j(u)$ are infinite]; let us set

$$\lambda(u) = \sup_{x \in \mathscr{X}'} \frac{|h(u; x)|}{\|x\|} \qquad \text{(for any fixed } u).$$

Let us remark that for any fixed x in $\mathscr{X}'$, $\frac{|h(u;x)|}{\|x\|}$ is a measurable function in u since $\mathrm{K}_j(u)$ is; as $\mathscr{X}'$ is countable, $\lambda(u)$ is thus also measurable [P.R.Halmos, I, p.84].

Take any $\varepsilon > 0$. Then we can find an $x_1(u)$ such that (by definition of the supremum)

a. $x_1(u) \in \mathscr{X}'$;
b. $\|x_1(u)\| = 1$;
c. $|h[u; x_1(u)]| \geq \lambda(u) - \varepsilon$.

Now by setting

$$x(u) = x_1(u) \qquad \text{if} \qquad h[u; x_1(u)] > 0$$

and

$$x(u) = -x_1(u) \qquad \text{if} \qquad h[u; x_1(u)] \leq 0,$$

we obtain an $x(u)$ such that

a. $x(u) \in \mathscr{X}'$;
b. $\|x(u)\| = 1$;
c. $h(u; x(u)) \geq \lambda(u) - \varepsilon$.

Let $\overset{\alpha}{\mathrm{X}}$ (X) be defined by $x(u)$; I say that this is countable. Firstly, $\mathrm{E}(\|\mathrm{X}\|^{\alpha}) = 1 < +\infty$ ($\|\mathrm{X}\|$ being equal to 1 is measurable); then X only takes countably many values, namely the x'_j; we now need to check that the set e_j of the u for which $x(u) = x'_j$ is measurable; let e'_j be the set of the u for which

$$\frac{h(u; x'_j)}{\|x'_j\|} \geq \lambda(u) - \varepsilon.$$

As the two sides of this inequality are measurable functions of u, e'_j is measurable; let e''_j be the sets defined by

$$e''_1 = e'_1, \qquad e''_j = e'_j - e'_j(e'_1 + ... + e'_{j-1}).$$

The e''_j are measurable. Choose $x(u) = x'_j$ if $u \in e''_j$; by above, the e''_j are pairwise disjoint and $\sum_j e''_j = \mathscr{U}$, which justifies the operation; then e''_j is measurable.

We saw previously that when X is countable, we have

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}\right) = \sum_j x^*(e_j; x'_j).$$

But by the definition of h,

$$x^*(e_j; x'_j) = \int_{e_j} h(u; x'_j) d\mu$$

and since, on e_j, $x'_j = x(u)$, we have

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}\right) = \sum_j \int_{e_j} h[u; x(u)] d\mu = \int_{\mathscr{U}} h[u; x(u)] d\mu \geq \int_{\mathscr{U}} [\lambda(u) - \varepsilon] d\mu;$$

as ε is arbitrarily small and $\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}\right)$ is finite, we have

$$\int_{\mathscr{U}} \lambda(u) d\mu < +\infty$$

[$\lambda(u)$ is obviously positive or zero], so $\lambda(u)$ is finite almost everywhere. So, apart from some exceptional values of u, $h(u;x)$ for fixed u is a continuous and additive functional on $\mathscr{X}'$; if $x \notin \mathscr{X}'$ and if $x_j \in \mathscr{X}'$ with $x_j \to x$ since $\mathscr{X}'$ is dense in $\mathscr{X}$, $h(u,x_j)$ converges according to Cauchy; let $h(u;x)$ be its limit. Then it is clear that $h(u;x)$ is a continuous and additive functional defined on $\mathscr{X}$ with norm $\lambda(u)$.

C. We are now going to show that

$$\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}} d\mu < +\infty.$$

The ideas behind the proof are borrowed from Landau [F.Riesz, I, p.44] and from above.

Suppose that $\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}} d\mu = +\infty$; we can then find a sequence of increasing positive numbers b_k such that

$$\text{(8)} \qquad \mathrm{B}_r = \int_{b_r}^{b_{r+1}} \lambda(u)^{\frac{\alpha}{\alpha-1}} d\mu \geq 2^{\frac{2}{\alpha-1}} \qquad \text{for every } r.$$

Let $x_1(u)$ be the summit of $h(u;x)$ considered as a linear functional in x; and $x_2(u)$ the point, if u is such that $b_r \leq \lambda(u) < b_{r+1}$, equal to

$$\frac{\|h(u)\|^{\frac{1}{\alpha-1}}}{\mathrm{B}_r} x_1(u).$$

We remark that $\|x_2(u)$ is measurable and that, by (8),

$$\int_{\mathscr{U}} \|x_2(u)\|^{\alpha} d\mu < +\infty,$$

and moreover,

$$h[u;x_2(u)] = \frac{\|h(u)\|^{\frac{\alpha}{\alpha-1}}}{\mathrm{B}_r}.$$

Let e'_j be the set of u for which we have *both*

$$\text{(9)} \qquad h(u;x'_j) \geq \frac{1}{2} \frac{\|h(u)\|^{\frac{\alpha}{\alpha-1}}}{\mathrm{B}_r}$$

$$\text{(10)} \qquad \|x_2(u)\|^{\alpha} \geq \|x'_j\|^{\alpha} - \varepsilon \quad (\text{any } \varepsilon > 0)$$

The e'_j are measurable since $h[u;x'_j]$, $\|h(u)\|$ and $\|x_2(u)\|$ are measurable; every u belongs to at least one of the e'_j since, for any given u (save for exceptional u), there is an x'_j arbitrarily close to $x_2(u)$; let e''_j be defined by

$$e''_1 = e'_1, \quad e''_j = e'_j - e'_j(e'_1 + ... + e'_{j-1}).$$

The e''_j are measurable, pairwise disjoint and

$$\sum_j e''_j = \mathscr{U}.$$

If $u \in e''_j$, let us set $x(u) = x'_j$ and let $\overset{\alpha}{\mathrm{X}}$ (X) be the random element defined by $x(u)$; by above, it is countable since, thanks to (10),

$$\int_{\mathscr{U}} \|x(u)\|^{\alpha} d\mu \leq \varepsilon + \int_{\mathscr{U}} \|x_2(u)\|^{\alpha} d\mu < +\infty.$$

Since $x(u) = x'_j$ on e''_j, we have

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}} \right) = \sum x^*[e''_j; x'_j] = \sum_j \int_{e''_j} h[u; x'_j] d\mu$$

$$\sum_j \int_{e''_j} h[u; x(u)] d\mu = \int_{\mathscr{U}} h[u; x(u)] d\mu = \frac{1}{2} \int_{\mathscr{U}} \frac{\|h(u)\|^{\frac{\alpha}{\alpha-1}}}{\mathrm{B}_r},$$

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}} \right) \geq \frac{1}{2} \sum_r \frac{1}{\mathrm{B}_r} \int_{b_r}^{b_{r+1}-0} \|h(u)\|^{\frac{\alpha}{\alpha-1}} d\mu = \frac{1}{2} \sum_r 1 = +\infty,$$

which is impossible, so

$$\int_{\mathscr{U}} \|h(u)\|^{\frac{\alpha}{\alpha-1}} d\mu < +\infty \qquad [\|h(u)\| = \lambda(u)].$$

D. The above work shows that if $\overset{\alpha}{\mathrm{X}}$ $-\mathrm{X} - x(u)$ is countable, $h[u; x(u)]$ is a measurable function in u; take any $\overset{\alpha}{\mathrm{X}}$ $-\mathrm{X} - x(u)$; let $\overset{\alpha}{\mathrm{X}}_n$ $-\mathrm{X}_n - x_n(u)$ be countable and converge to $\overset{\alpha}{\mathrm{X}}$ as it was indicated before, that is to say, in such a way that $\|x_n(u) - x(u)\|$ converges to zero uniformly in u; then $h[u; x_n(u)]$ is measurable and converges to $h[u; x(u)]$ which is thus measurable (as a limit of measurable functions).

We conclude from this that

$$\int_{\mathscr{U}} h[u; x(u)] d\mu$$

makes sense for any $\overset{\alpha}{\mathrm{X}}$ $-\mathrm{X} - x(u)$, because

$$\int_{\mathscr{U}} |h[u; x(u)]| d\mu \leq \int_{\mathscr{U}} \lambda(u) \|x(u)\| d\mu.$$

It then suffices to apply Hölder's inequality to see that

$$\int_{\mathscr{U}} |h[u; x(u)]| d\mu < +\infty$$

knowing that

$$\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}} d\mu < +\infty \quad \text{and} \quad \int_{\mathscr{U}} \|x(u)\|^{\alpha} d\mu < +\infty.$$

We have

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}} \right) = \int_{\mathscr{U}} h[u; x(u)] d\mu;$$

this is true if $\overset{\alpha}{\mathrm{X}}$ is countable, as we have seen before. If $\overset{\alpha}{\mathrm{X}}$ is arbitrary, let $\overset{\alpha}{\mathrm{X}}_n$ be countable and converge to $\overset{\alpha}{\mathrm{X}}$.

$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}_n\right)$ is equal to $\int_{\mathscr{U}} h[u; x_n(u)]d\mu$ and converges to $\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}\right)$. But if $x_n(u) \to x(u)$ uniformly in u, which is possible since

$$\int_{\mathscr{U}} \lambda(u)d\mu < +\infty,$$

we then have that

$$\int_{\mathscr{U}} h[u; x_n(u)]d\mu \to \int_{\mathscr{U}} h[u; x(u)]d\mu,$$

so

$$\overset{\alpha}{\mathrm{X}}{}^* \left(\overset{\alpha}{\mathrm{X}}\right) = \int_{\mathscr{U}} h[u; x(u)]d\mu. \tag{11}$$

Q.E.D.

The case $\alpha = 1$ is only different from paragraph C onwards, but the proof is simpler; in the case $\alpha = 1$ and under the hypothesis that $\mathscr{X}^*$ is separable, the result was obtained by Dieudonné [J.Dieudonné, I, p.38]. In collaboration with M. Fortet [R.Fortet and E.Mourier, I], we could extend the above results for $\alpha \geq 1$ under the hypothesis that just $\mathscr{X}$ is separable.

IV. – Strong law of large numbers with respect to strong convergence

Reminder of Birkhoff's ergodicity theorem[(8)]. – If ordinary numerical random variables X_s form a strictly stationary discontinuous chain and if $\mathrm{E}(\mathrm{X}_s)$ exists, the arithmetic mean $\mathrm{Y}_n = \frac{\mathrm{X}_1 + \ldots + \mathrm{X}_n}{n}$ converges, as n increases indefinitely, almost surely to a limiting random variable Y.

Now let X be a random element taking its values in a separable Banach space $\mathscr{X}$, such that $x^*(\mathrm{X})$ is measurable for any fixed x^* and such that

$$\mathrm{E}(\|\mathrm{X}\|) < +\infty.$$

Let us consider a strictly stationary and discontinuous sequence X_s.

1st *Case.* – The X_s only take a *finite* number of distinct values $x_1, ..., x_k$ (the same values whatever s is, because of stationarity).

Let A_s^j the (ordinary) random variable, which equals 1 if $\mathrm{X}_s = x_j$ and 0 otherwise. The set e_j^s of events for which $\mathrm{X}_s = x_j$ is assumed to be measurable;

$$\mu(e_j^s) = \Pr(\mathrm{X}_s = x_j) = \mathrm{P}_j$$

is independent of s because of stationarity. For a given j, the A_s^j form a strictly stationary sequence of random variables, so

$$\text{almost surely, } \frac{1}{n}\sum_{s=1}^{n} \mathrm{A}_s^j \underset{n\to\infty}{\longrightarrow} \text{a limit } \mathrm{L}^j.$$

(8) In [Kolmogorov, II, p.367], one can find a proof of Birkhoff's theorem that simplifies Birkhoff's original proof [Khintchine, I, p.485], as well as a simple probabilistic proof.

It immediately follows that almost surely, $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s$ converges strongly to $\sum_{j=1}^{k} \mathrm{L}^j x_j$ when $n \to \infty$.

More precisely, we see that

$$\mathrm{E}\left[\left\|\frac{1}{n}\sum_{1}^{n} \mathrm{X}_s - \sum_{1}^{k} \mathrm{L}^j x_j\right\|\right] \to 0.$$

2nd *Case.* – The X_s only take a *countable* number of distinct values $x_1, ..., x_j, ...$ (independently of s) and the event $\mathrm{X}_s = x_j$ can be given a probability.

Let X_s^t be the random element defined by $\mathrm{X}_s^t = \mathrm{X}_s$ if $\mathrm{X}_s = x_1$ or x_2 or ... or x_t; $\mathrm{X}_s^t = 0$ if $\mathrm{X}_s = x_{t+1}$ or x_{t+2} or ... The X_s^t are of type studied in the first case; so, except for the s in a set e_t such that $\mu(e_t) = 0$, we have

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s^t = \mathrm{A}_t$$

(some limit dependent on the t under consideration).

Let us set

$$\mathrm{X}_s = \mathrm{X}_s^t + \mathrm{R}_s^t;$$

R_s^t is equal to 0 or to $x_{t+1}, x_{t+2}, ...$

If $\mathrm{P}_j = \Pr[\mathrm{X}_s = x_j]$,

$$\mathrm{E}(\|\mathrm{R}_s^t\|) = \sum_{j>t} \mathrm{P}_j \|x_j\|.$$

The hypothesis is that

$$\sum_{j} \mathrm{P}_j \|x_j\| < +\infty,$$

so

$$\mathrm{E}(\|\mathrm{R}_s^t\|) \to 0 \qquad \text{if} \qquad t \to +\infty,$$

$$\text{(12)} \qquad \frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s = \frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s^t + \frac{1}{n}\sum_{s=1}^{n} \mathrm{R}_s^t,$$

$$\left\|\frac{1}{n}\sum_{s=1}^{n} \mathrm{R}_s^t\right\| \leq \frac{1}{n}\sum_{s=1}^{n} \|\mathrm{R}_s^t\|.$$

The $\|\mathrm{R}_s^t\|$ for any given t form a stationary sequence of ordinary random variables; so, unless $u \in e'_t$,

$$\mu(e'_t) = 0, \qquad \frac{1}{n}\sum_{s=1}^{n} \|\mathrm{R}_s^t\| \underset{n\to+\infty}{\longrightarrow} M_t(\geq 0)$$

and, by Fatou's lemma, since

$$\mathrm{E}\left(\frac{1}{n}\sum_{s=1}^{n} \|\mathrm{R}_s^t\|\right) = \mathrm{E}(\|\mathrm{R}_s^t\|) = \sum_{j>t} \mathrm{P}_j \|x_j\| \quad \text{independently of } n,$$

we have

$$\mathrm{E}(\mathrm{M}_t) \leq \sum_{j>t} \mathrm{P}_j \|x_j\| = K_t \geq 0;$$

we can thus find an increasing sequence of t, $t_1, t_2, ..., t_\nu, ...$, such that

$$\sum_\nu \mathrm{K}_{t_\nu} < +\infty,$$

so, by the Borel-Cantelli theorem, the sequence of random variables M_{t_ν} ($\nu = 1, 2, ...$) converge almost surely to 0.

Let

$$e = \sum_t e_t, \qquad e' = \sum_t e'_t;$$

let e''_t be the set of u for which M_{t_ν} does not tend to 0. Then

$$\mu(e) = \mu(e') = \mu(e'') = 0.$$

Finally, let $e''' = e + e' + e''$ and let u be any outcome with $u \notin e'''$. Then we can take ν sufficiently large so that, for the u considered, $\mathrm{M}_{t_\nu} \leq \varepsilon$, and with this fixed ν, we have

$$\frac{1}{n} \sum_{s=1}^{n} \|\mathrm{R}_s^{t_\nu}\| \leq 2\varepsilon$$

and, as a result,

$$\left\| \frac{1}{n} \sum_{s=1}^{n} \mathrm{R}_s^{t_\nu} \right\| \leq 2\varepsilon$$

for all n sufficiently large.

On the other hand,

$$\frac{1}{n} \sum_{s=1}^{n} \mathrm{X}_s^{t_\nu} \to \mathrm{A}_{t_\nu} \qquad \text{when } n \to +\infty;$$

by (12) with $t = t_\nu$ we have, as ε is arbitrary, that $\frac{1}{n} \sum_{s=1}^{n} \mathrm{X}_s$ has a limit $y(u)$ (strong convergence) for this u. Hence, this is true for all the $u \notin e'''$ where $\mu(e''') = 0$, so $\frac{1}{n} \sum_{s=1}^{n} \mathrm{X}_s$ converges strongly almost surely. If Y is the random element defined by $y(u)$, we have, for each fixed x^*, a measurable function $x^*[y(u)]$ of u, as an almost everywhere limit of $x^* \left[\frac{1}{n} \sum_{s=1}^{n} \mathrm{X}_s \right]$, which is measurable. Then as $\mathscr{X}$ is separable, $\|y(u)\|$ is measurable, and by Fatou's lemma,

$$E(\|\mathrm{Y}\|) = \int_{\mathscr{U}} \|y(u)\| d\mu \leq \liminf_n \int_{\mathscr{U}} \left\| \frac{1}{n} \sum_{s=1}^{n} \mathrm{X}_s \right\| d\mu \leq \mathrm{E}(\|\mathrm{X}_s\|) < +\infty,$$

and we have the following theorem.

Theorem. – *If the X_s in the stationary sequence only take a countable number of distinct values $x_1, ..., x_j, ...$ and if the event $X_s = x_j$ can be given a probability, almost surely $\frac{1}{n} \sum_{s=1}^{n} X_s$ converges strongly to a limit Y which is a random element of type X_s, that is to say, that $x^*(Y)$ is measurable and*

$$\mathrm{E}(\|\mathrm{Y}) < +\infty.$$

3^{rd} *Case. – The* X_s *take any values.*
Since $\mathscr{X}$ is separable, let $[x_j]$ be a dense countable sequence in $\mathscr{X}$.
Let ν be any integer and A_j^ν the set of the $x \in \mathscr{X}$ for which $\|x - x_j\| \leq \frac{1}{\nu}$.
Let B_j^ν be the sets defined in the following way:

$$\mathrm{B}_1^\nu = \mathrm{A}_1^\nu, \qquad \mathrm{B}_1^\nu = \mathrm{A}_j^\nu - \mathrm{A}_j^\nu(\mathrm{A}_1^\nu + ... + \mathrm{A}_{j-1}^\nu) \quad (j \geq 2).$$

The B_j^ν are disjoint and $\sum_j \mathrm{B}_j^\nu = \sum_j \mathrm{A}_j^\nu = \mathscr{X}$.

Let $\mathscr{Z}_\nu$ be the transformation which associates to every $x \in \mathscr{X}$ some $y = \mathscr{Z}_\nu(x)$ by the rule

$$y = x_j \qquad \text{if } x \in \mathrm{B}_j^\nu.$$

We always have

$$\|\mathscr{Z}_\nu(x) - x\| \leq \frac{1}{\nu}.$$

Let X be a random element, of the type considered, on $\mathscr{X}$. We associate a random element X^ν with it, by

$$\mathrm{X}^\nu = \mathscr{Z}_\nu(\mathrm{X}).$$

It is clear that X^ν only takes a countable number of distinct values, that $\mathrm{X}^\nu = x_j$ can be given a probability, and that

$$\mathrm{E}(\|\mathrm{X}^\nu - \mathrm{X}\|) \leq \frac{1}{\nu},$$

more precisely, we *surely* have

$$\|\mathrm{X}^\nu - \mathrm{X}\| \leq \frac{1}{\nu}.$$

As $\mathrm{X}_1, \mathrm{X}_2, ..., \mathrm{X}_s, ...$ is a strictly stationary sequence, the sequence of the $\mathrm{X}_s^\nu = \mathscr{Z}_\nu(\mathrm{X}_s)$ is strictly stationary, and we have

$$\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s = \frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s^\nu + \frac{1}{n}[\mathrm{X}_s^\nu - \mathrm{X}_s].$$

We have *surely*

$$\left\| \frac{1}{n}\sum_{s=1}^{n}[\mathrm{X}_s^\nu - \mathrm{X}_s] \right\| \leq \frac{1}{n}\sum_{s=1}^{n}\|\mathrm{X}_s^\nu - \mathrm{X}_s\| \leq \frac{1}{\nu} \qquad \text{for any } n.$$

For any, but fixed, given ν, we saw that $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s^\nu$ converges strongly almost surely (2^{nd} case above), that is to say, with the exception of

$$u \in e_\nu, \qquad \text{with} \qquad \mu(e_\nu)0.$$

If we consider

$$u \notin e = \sum_\nu e_\nu, \qquad \mu(e) = 0,$$

$\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s^\nu$ converges for any ν, and as a result, we have that $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s$ converges strongly almost surely to a limit Y [it is clear that $x^*(\mathrm{Y})$ is measurable and $\mathrm{E}(\|\mathrm{Y}\|) < +\infty$ as in the above case]. Hence:

Strong law of large numbers. – If $\mathscr{X}$ is separable, if every $x^*(\mathrm{X}_s)$ is measurable and $\mathrm{E}(\|\mathrm{X}_s\|) < +\infty$, and if $\mathrm{X}_1, \mathrm{X}_2, ..., \mathrm{X}_s, ...$ form a strictly stationary discontinuous sequence, then almost surely the mean $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s$ converges *strongly* to a limit Y which is a random element of the same type as the X_s.

Remark I. – We did not assume that $\mathscr{X}$ is reflexive, and we did not assume the existence of $\mathrm{E}(\mathrm{X}_s)$. Remark that if $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s$ converges *strongly*, *a fortiori* it converges *weakly*. Moreover, if the X_s are *independent* and have the same law (a particular case of a stationary sequence), the $x^*(\mathrm{X}_s)$ will also be independent and of the same law; also,

$$|x^*(\mathrm{X}_s)| \leq \|x^*\| \cdot \|\mathrm{X}_s\|,$$

so $\mathrm{E}\|\mathrm{X}_s\| < +\infty$ implies the existence of $\mathrm{E}[x^*(\mathrm{X}_s)]$. But, by the preceding theorem,

$$x^*\left[\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s\right] = \frac{1}{n}\sum_{s=1}^{n} x^*(\mathrm{X}_s) \to x^*(\mathrm{Y}) \qquad \text{almost surely,}$$

but, by Kolmogorov's theorem,

$$\frac{1}{n}\sum_{s=1}^{n} x^*(\mathrm{X}_s) \to \mathrm{E}[x^*(\mathrm{X}_s)] \qquad \text{almost surely,}$$

so

$$\mathrm{E}[x^*(\mathrm{X}_s)] = x^*(\mathrm{Y}).$$

Thus, just by the definition of $\mathrm{E}(\mathrm{X}_s)$, X_s has a mathematical expectation Y, so:

Theorem. – *When* $\mathscr{X}$ *is* separable, $\mathrm{E}(\|\mathrm{X}\|) < +\infty$ *implies the existence of* E(X).

This theorem is an interesting existence theorem of Pettis integrals; it generalises the theorem of Chapter I (page 6), where $\mathscr{X}$ was assumed to be not only separable, but also reflexive.

Remark II. – When the X_s are independent and of the same law, the limit Y of $\frac{1}{n}\sum_{s=1}^{n} \mathrm{X}_s$ is the mathematical expectation $\mathrm{E}(\mathrm{X}_s)$.

V. – Strong law of large numbers in mean of order α.

Ergodic theorem of Yosida and Kakutani [Yosida and Kakutani, I]. – Let T be a bounded linear operator which maps from a Banach space $\mathscr{X}$ into itself and such that $\|\mathrm{T}^n\| \leq C$ for $n = 1, 2, ...$ (where C is a fixed number independent of n).

If, for $x \in \mathscr{X}$, the sequence $\{x_n\}$, where

$$x_n = \frac{1}{n}(\mathrm{T} + \mathrm{T}^2 + ... + \mathrm{T}^n)x,$$

$n = 1, 2, ...$, contains a subsequence which converges weakly to a point $\bar{x} \in \mathscr{X}$, the sequence $\{x_n\}$ converges *strongly* to this point $\bar{x}$.

If we denote by T_1 the operation $x \mapsto \bar{x}$, T_1 is a bounded linear operator which maps from $\mathscr{X}$ into itself and

$$TT_1 = T_1T = T_1^2 = T_1.$$

We will apply the following theorem:

Let $\{X_i\}_{i=(0\pm1\pm2\pm\ldots)}$ be a strictly stationary sequence of random elements with values in some Banach space $\mathscr{X}$ such that

$$E(\|X_i\|^\alpha) < +\infty;$$

let $\overset{\alpha}{X}_i$ be the corresponding elements in $\overset{\alpha}{\mathscr{X}}$. Then these $\overset{\alpha}{X}_i$ define in $\overset{\alpha}{\mathscr{X}}$ a closed linear manifold $\overset{\alpha}{\mathscr{X}}{}'$ which is itself a Banach space (it is a subset of $\overset{\alpha}{\mathscr{X}}$, it is even a separable subset).

We will need to know the form of the linear functionals on this linear manifold; now, in Paragraph III, we found the general form of linear functionals on $\overset{\alpha}{\mathscr{X}}$; but, taking into account the extension theorem of linear functionals on a linear subspace to the whole space with conserved norm [E. Hille, I], we will be able to use this result.

The $\overset{\alpha}{X}$ of the form $\sum_{i=1}^k a_i \overset{\alpha}{X}_i$, where k is finite and arbitrary and the a_i are some numbers, form a non-closed subset $\overset{\alpha}{\mathscr{X}}{}''$ of $\overset{\alpha}{\mathscr{X}}{}'$; $\overset{\alpha}{\mathscr{X}}{}'$ is the closure of $\overset{\alpha}{\mathscr{X}}{}''$. If $\overset{\alpha}{X}$ is in $\overset{\alpha}{\mathscr{X}}{}''$, so of the form

$$\sum_{i=1}^k a_i \overset{\alpha}{X}_i,$$

let us set

$$\overset{\alpha}{Z} = T\left(\overset{\alpha}{X}\right) = \sum_{i=1}^k a_i \overset{\alpha}{X}_{i+1}; \tag{13}$$

$T\left(\overset{\alpha}{X}\right)$ is an operation which is obviously additive, and by virtue of stationarity, we have

$$E(\|Z\|^\alpha) = E\left(\left\|\sum_{i=1}^k a_i X_{i+1}\right\|^\alpha\right) = E\left(\left\|\sum_{i=1}^k a_k X_i\right\|^\alpha\right) = E(\|X\|^\alpha),$$

so

$$\left\|\overset{\alpha}{Z}\right\| = \left\|\overset{\alpha}{X}\right\|. \tag{14}$$

We need to see whether $\overset{\alpha}{Z}$ is defined uniquely. If $\overset{\alpha}{X}$ only admits one representation of the form $\sum_{i=1}^k a_i \overset{\alpha}{X}_i$, uniqueness is evident; but suppose that $\overset{\alpha}{X}$ admits two distinct representations of this form, for example,

$$\sum_{i=1}^k a_i \overset{\alpha}{X}_i \qquad \text{and} \qquad \sum_{i=1}^k a_i' \overset{\alpha}{X}_i$$

(we can always assume that k is the same for both forms); with (13), we will obtain two transformations of $\overset{\alpha}{\mathrm{X}}$, $\overset{\alpha}{\mathrm{Z}}$ and $\overset{\alpha}{\mathrm{Z}}'$ given by

$$\overset{\alpha}{\mathrm{Z}}=\sum_{i=1}^{k} a_i \overset{\alpha}{\mathrm{X}}_{i+1}, \qquad \overset{\alpha}{\mathrm{Z}}'=\sum_{i=1}^{k} a_i' \overset{\alpha}{\mathrm{X}}_{i+1} .$$

But $\overset{\alpha}{\mathrm{Z}}$ and $\overset{\alpha}{\mathrm{Z}}'$ are not distinct; indeed,

$$\overset{\alpha}{\mathrm{X}}=\sum_{i=1}^{k} a_i \overset{\alpha}{\mathrm{X}}_i=\sum_{i=1}^{k} a_i' \overset{\alpha}{\mathrm{X}}_i$$

implies that, by setting $b_i = a_i - a_i'$,

$$\mathrm{E}\left(\left\|\sum_{i=1}^{k} b_i \mathrm{X}_i\right\|^{\alpha}\right) = 0.$$

So, by virtue of stationarity,

$$\mathrm{E}\left(\left\|\sum_{i=1}^{k} b_i \mathrm{X}_{i+1}\right\|^{\alpha}\right) = 0,$$

which is to say that $\overset{\alpha}{\mathrm{Z}}=\overset{\alpha}{\mathrm{Z}}'$.

Now take any $\overset{\alpha}{\mathrm{X}}$ in $\overset{\alpha}{\mathscr{X}}$; there is a sequence $\overset{\alpha}{\mathrm{X}}_i \in \overset{\alpha}{\mathscr{X}}''$ converging to $\overset{\alpha}{\mathrm{X}}$, which, by Cauchy, implies that

$$\text{when } n \to \infty, \qquad \left\|\overset{\alpha}{\mathrm{X}}_{n+p} - \overset{\alpha}{\mathrm{X}}_n\right\| \to 0,$$

with $\frac{1}{n}$, uniformly in p. We then have

$$\left\|\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}_{n+p}\right) - \mathrm{T}\left(\overset{\alpha}{\mathrm{X}}_n\right)\right\| = \mathrm{T}\left(\overset{\alpha}{\mathrm{X}}_{n+p} - \overset{\alpha}{\mathrm{X}}_n\right) = \left\|\overset{\alpha}{\mathrm{X}}_{n+p} - \overset{\alpha}{\mathrm{X}}_n\right\| \qquad \text{by (14).}$$

So $\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}_n\right)$ converges to a limit which, by definition, we will call $\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)$.

$\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)$ depends on $\overset{\alpha}{\mathrm{X}}$, but obviously not on the sequence $\overset{\alpha}{\mathrm{X}}_n$; $\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)$ is then a transformation defined on $\overset{\alpha}{\mathscr{X}}$; it is immediate that it is linear, continuous and that

$$\|\mathrm{T}\| = 1 \qquad \left(\text{more precisely, } \left\|\mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)\right\| = \left\|\overset{\alpha}{\mathrm{X}}\right\|\right)$$

and, as a result, for all n,

$$\|\mathrm{T}^n\| = 1.$$

The theorem of Yosida and Kakutani then gives us the following statement: let $\overset{\alpha}{\mathrm{X}}\in\overset{\alpha}{\mathscr{X}}'$; if $\frac{1}{n}\sum_{i=1}^{n}\mathrm{T}^i\left(\overset{\alpha}{\mathrm{X}}\right)$ converges weakly, it converges *strongly*.

In particular, let us take

$$\overset{\alpha}{\mathrm{X}}=\overset{\alpha}{\mathrm{X}}_1,\quad \mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)=\overset{\alpha}{\mathrm{X}}_2,\quad ,...,\quad \mathrm{T}\left(\overset{\alpha}{\mathrm{X}}\right)=\overset{\alpha}{\mathrm{X}}_{n+1}\ .$$

Then we have that, if

$$\frac{1}{n}\sum_{i=1}^{n}\mathrm{T}^i\left(\overset{\alpha}{\mathrm{X}}\right)=\frac{1}{n}\left(\overset{\alpha}{\mathrm{X}}_2+\overset{\alpha}{\mathrm{X}}_3+...+\overset{\alpha}{\mathrm{X}}_{n+1}\right)$$

converges *weakly*, it converges *strongly*, which is to say that there exists a random element L such that

$$\mathrm{E}\left(\left\|\frac{\mathrm{X}_1+...+\mathrm{X}_n}{n}-\mathrm{L}\right\|^{\alpha}\right)\to 0. \tag{15}$$

It suffices, moreover, that there exists a subsequence of $\frac{1}{n}\sum_{i=2}^{n+1}\overset{\alpha}{\mathrm{X}}_i$ which converges weakly; this will be the case, for example, if $\overset{\alpha}{\mathscr{X}}'$ is weakly compact, so in particular if $\overset{\alpha}{\mathscr{X}}'$ is uniformly convex.

Note that the existence of $\mathrm{E}(\mathrm{X}_i)$ is not assumed, likewise E(L) is not assumed to exist, but if $\mathscr{X}$ is separable and reflexive, $\mathrm{E}(\mathrm{X}_i)$ and E(L) exist (cf. Chapter I, page 6); in any case, when E(L) exists, L does not reduce to the almost sure element E(L). In all cases, we have the following properties.

Property 1. – If $\mathrm{E}(\mathrm{X}_i)$ and E(L) exist – so in particular if $\mathscr{X}$ is separable and reflexive – we have $\mathrm{E(L)}=\mathrm{E}(\mathrm{X}_i)$. In fact, we do not lose generality by assuming $\mathrm{E}(\mathrm{X}_i)$; then

$$\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{X}_i-\mathrm{L}\right)=-\mathrm{E(L)}$$

exists, but

$$\left\|\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{X}_i-\mathrm{L}\right)\right\|\leq\mathrm{E}\left\|\frac{1}{n}\sum_{i=1}^{n}\mathrm{X}_i-\mathrm{L}\right\|$$

which tends to zero by (15); so

$$\mathrm{E(L)}=0=\mathrm{E}(\mathrm{X}_i).$$

Property 2. – When (15) holds, for any $x^*\in\mathscr{X}^*$,

$$x^*\left[\frac{\mathrm{X}_1+...+\mathrm{X}_n}{n}\right]=\frac{1}{n}\sum_{i=1}^{n}x^*(\mathrm{X}_i)$$

converges in mean of order α, hence also in probability, to $x^*(\mathrm{L})$; $\mathrm{E}[x^*(\mathrm{X}_i)]$ exists, because $\mathrm{E}\|\mathrm{X}_i\|^{\alpha}<+\infty$; if the X_i are *independent* the $x^*(\mathrm{X}_i)$ are also independent. Then, by Kolmogorov's theorem,

$$\frac{1}{n}\sum_{i=1}^{n} x^*(\mathrm{X}_i) \to \mathrm{E}[x^*(\mathrm{X}_i)] \qquad \text{almost surely}$$

so

$$x^*(\mathrm{L}) = \mathrm{E}[x^*(\mathrm{X}_i)] \qquad \text{almost surely,}$$

in other words,

$$x^*(\mathrm{L}) = \mathrm{E}[x^*(\mathrm{X}_i)],$$

except for the u belonging to an event e of zero μ measure, but e can depend on x^*.

But let us assume that $\mathscr{X}^*$ is separable and let $\{x_j^*\}$ be a dense countable sequence in $\mathscr{X}^*$. There exists e independent of j with $\mu(e) = 0$ such that for all $u \notin e$, we have

$$x_j^*(\mathrm{L}) = \mathrm{E}[x_j^*(\mathrm{X}_i)].$$

Let x^* be arbitrary and x_j^* such that $\|x^* - x_j^*\| \leq \varepsilon$. There exists a set e' such that $\mu(e') = 0$ and such that if $u \notin e'$, $\|\mathrm{L}\|$ is finite [since $\mathrm{E}(\|\mathrm{L}\|) < +\infty$]. Then we have

$$|\mathrm{E}[x^*(\mathrm{X}_i)] - \mathrm{E}[x_j^*(\mathrm{X}_i)]| < \varepsilon \mathrm{E}(\|\mathrm{X}_i\|),$$

$$\begin{aligned} |x^*(\mathrm{L}) - \mathrm{E}[x^*(\mathrm{X}_i)]| &\leq |(x^* - x_j^*)(\mathrm{L}) + x_j^*(\mathrm{L}) - \mathrm{E}[x_j^*(\mathrm{X}_i)] \\ &\quad + \mathrm{E}[x_j^*(\mathrm{X}_i)] - \mathrm{E}[x^*(\mathrm{X}_i)]| \\ &\leq \varepsilon \mathrm{E}\|\mathrm{X}_i\| + \varepsilon \mathrm{E}\|\mathrm{L}\| + |x_j^*(\mathrm{L}) - \mathrm{E}[x_j^*(\mathrm{X}_i)]|; \end{aligned}$$

if $u \notin e' + e$,

$$x_j^*(\mathrm{L}) - \mathrm{E}[x_j^*(\mathrm{X}_i)] = 0,$$

which means

$$|x^*(\mathrm{L}) - \mathrm{E}[x^*(\mathrm{X}_i)]| \leq \varepsilon \|\mathrm{L}\| + \varepsilon \mathrm{E}(\|\mathrm{X}_i\|)$$

and as ε is arbitrarily small,

$$x^*(\mathrm{L}) = \mathrm{E}[x^*(\mathrm{X}_i)].$$

As a point L in $\mathscr{X}$ is entirely determined by the set of the values of $x^*(\mathrm{L})$, L is an almost sure random element l satisfying

$$x^*(l) = \mathrm{E}[x^*(\mathrm{X}_i)]$$

for every x^*, which proves that the X_i have mathematical expectation l and that $\frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i$ tends "in mean of order α", hence in probability, to l.

Application 1. – $\mathscr{X}$ is the real line, $\overset{\alpha}{\mathscr{X}}$ and $\overset{\alpha}{\mathscr{X}}{}'$ are then the space L^α which is uniformly convex if $\alpha > 1$, because a random variable is equivalent to a numerical function $f(u)$, then (15) can be applied (Yosida's theorem reduces in this case to Birkhoff's theorem); hence, we have the following theorem:

Let $\{\mathrm{X}_i\}$ *be a strictly stationary sequence of random variables, with*

$$\mathrm{E}(\|\mathrm{X}_i\|^\alpha) < +\infty \qquad (\alpha > 1).$$

Then its temporal mean $\frac{1}{n}\sum_{i=1}^{n} X_i$ *converges in mean of order* α.

The result is known for $\alpha = 2$ under the less strict condition of only stationarity of order 2, but here we can also avoid strict stationarity for $\alpha = 2$, or even α, and maybe for any α. For an arbitrary α, the theorem is perhaps novel.

It can immediately be extended to the case where the X_i are random variables in several dimensions.

Application 2. – Assume that $\mathscr{X}$ is the space L^α; an X_i is a *numerical* function $f(t)$ of a variable t such that

$$\int |f(t)|^\alpha dt < +\infty,$$

but this function f varies with u; we can thus consider a function $f(u;t)$ in two variables u and t; we have

$$\int |f(u;t)|^\alpha dt < +\infty,$$

but the hypothesis $E(\|X\|\alpha) < +\infty$, given

$$\|X\|^\alpha = \int |f(u;t)|^\alpha dt,$$

is equivalent to

$$\int\int |f(u;t)|^\alpha dt d\mu < +\infty$$

which shows that $\overset{\alpha}{\mathscr{X}}$ is the space L^α of numerical functions in two variables u and t, so it is a uniformly convex space and (15) can be applied.

This case is an extremely important case of random functions, and moreover is a particular case of the following, since $\mathscr{X} = L^\alpha$ is reflexive, for $\alpha > 1$.

Application 3. – Assume that $\mathscr{X}$ is *separable* and *reflexive*, which means that (11) holds, let

$$Y_n = \frac{X_2 + ... + X_{n+1}}{n}$$

and let $y_n(u)$ be the value of Y_n for the outcome u. We are going to show that $\overset{\alpha}{Y}_n$ converges weakly. Save for some exceptional u, by a preceding theorem, $y_n(u)$ converges strongly, so weakly to a limit $y(u)$, which implies that

$$h[u; y_n(u)] \to h[u; y(u)];$$

$h[u; y(u)]$ is a measurable function in u, as the limit of $h[u; y_n(u)]$ which is measurable (*see* above). We will show that

$$\overset{\alpha}{X}{}^* (\overset{\alpha}{Y}_n) = \int_{\mathscr{U}} h[u; y_n(u)] d\mu \to \int_{\mathscr{U}} h[u; y(u)] d\mu. \tag{16}$$

Let $\mathscr{A}$ be the set of u for which we do not have an upper bound $\|y_n(u)\| \leq A$; as $y_n(u)$ converges, save for exceptional u, towards a limit $y(u)$, the $y_n(u)$ are uniformly bounded in n; so if we take A large enough, $\mu(\mathscr{A}) \leq \varepsilon$; we have

$$\left| \int_{\mathscr{A}} h[u; y_n(u)] d\mu \right| \leq \int_{\mathscr{A}} \lambda(u) \left\| y_n(u) \right\| d\mu$$

and by Hölder's inequality,

$$(17)\qquad \int_{\mathscr{A}} \lambda(u)\,\|y_n(u)\|\,d\mu \le \left[\int_{\mathscr{A}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu\right]^{\frac{\alpha-1}{\alpha}} \left[\int_{\mathscr{A}} \|y_n(u)\|^{\alpha}\,d\mu\right]^{\frac{1}{\alpha}}$$
$$\le \left[\int_{\mathscr{A}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu\right]^{\frac{\alpha-1}{\alpha}} \left[\int_{\mathscr{U}} \|y_n(u)\|^{\alpha}\,d\mu\right]^{\frac{1}{\alpha}}.$$

But

$$\left[\int_{\mathscr{U}} \|y_n(u)\|^{\alpha}\,d\mu\right]^{\frac{1}{\alpha}} = \|\overset{\alpha}{\mathrm{Y}}_n\| = [\mathrm{E}(\|\mathrm{X}_i\|^{\alpha})]^{\frac{1}{\alpha}} < +\infty$$

independently of n; moreover,

$$\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu < +\infty;$$

if we replace A by $\mathrm{A}' > \mathrm{A}$, $\mathscr{A}$ is replaced by $\mathscr{A}' \subset \mathscr{A}$ and $\mu(\mathscr{A}) \to 0$ if $\mathrm{A} \to +\infty$; this means that by taking A large enough (ε small enough), $\int_{\mathscr{A}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu$ is arbitrarily small, hence so is $\int_{\mathscr{A}} h[u, y_n(u)]d\mu$ and this holds uniformly *in n*. Simultaneously and for the same reasons, $\int_{\mathscr{A}} h[u; y(u)]d\mu$ is small, because for any n:

$$\int_{\mathscr{A}} \big|h[u; y_n(u)]\big|\,d\mu \le \left[\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu\right]\frac{\alpha-1}{\alpha}[\mathrm{E}(\|\mathrm{X}_i\|^{\alpha})]^{\frac{1}{\alpha}} = \mathrm{C} \qquad \text{(Hölder)},$$
$$\big|h[u; y_n(u)]\big| \to \big|h[u; y(u)]\big| \qquad \text{almost everywhere;}$$

so by Fatou's lemma,

$$(18)\qquad \int_{\mathscr{U}} \big|h[u; y(u)]\big|\,d\mu < \mathrm{C},$$

so as $\mu(\mathscr{A}) \to 0$,

$$\int_{\mathscr{A}} h[u; y(u)]d\mu \to 0.$$

Likewise, let $\mathscr{B}$ be the set of the u for which we have $\lambda(u) > \mathrm{B}$, where B is a positive number; if we replace B by $\mathrm{B}' > \mathrm{B}$, $\mathscr{B}$ is replaced by $\mathscr{B}' \subset \mathscr{B}$ and $\mu(\mathscr{B}) \to 0$ if $\mathrm{B} \to +\infty$ because

$$\int_{\mathscr{U}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu < +\infty$$

then

$$\int_{\mathscr{B}} \lambda(u)^{\frac{\alpha}{\alpha-1}}\,d\mu \to 0.$$

It then follows from (17) that

$$\int_{\mathscr{B}} h[u; y_n(u)]d\mu$$

is small, uniformly in n when B is large, and likewise by (18),

$$\int_{\mathscr{B}} h[u; y(u)]d\mu \qquad \text{is small.}$$

Let $\mathscr{D} = \mathscr{A} + \mathscr{B}$. On $\mathscr{U} - \mathscr{B}$, we have $\lambda(u) \leq \text{B}$,

$$\text{the upper bound } \left\|y_n(u)\right\| \leq \text{A}, \qquad \text{so } |h[u; y_n(u)]| \leq \text{AB}.$$

We can then pass to the limit under the integral sign $\int$ and write

$$\int_{\mathscr{U}-\mathscr{D}} h[u; y(u)]d\mu = \lim_{n\to+\infty} \int_{\mathscr{U}-\mathscr{D}} h[u; y_n(u)]d\mu.$$

But our work above then shows that, since A and B can be arbitrarily large,

$$\int_{\mathscr{U}} h[u; y(u)]d\mu = \lim_{n\to+\infty} \int_{\mathscr{U}} h[u; y_n(u)]d\mu = \lim_{n\to+\infty} \overset{\alpha}{\text{X}}{}^* \, (\overset{\alpha}{\text{Y}}_n).$$

Now, let Y be the random element defined by $\text{Y} = y(u)$ for the outcome u. Let us show that Y defines an $\overset{\alpha}{\text{Y}} \in \overset{\alpha}{\mathscr{X}}$, in other words,, that

$a.$
$$x^*(\text{Y}) = x^*[y(u)]$$

is a measurable function in u for every fixed x^* and that $\|\text{Y}\|$ is measurable. The first point results from the fact that almost everywhere, $x^*[y(u)] = \lim x^*[y_n(u)]$ and that $x^*[y_n(u)]$ is measurable. As $\mathscr{X}$ is separable, it follows that $\|\text{Y}\|$ is measurable (*cf.* Chapter I).

$b.$
$$\text{E}[\|\text{Y}\|^\alpha] = \int_{\mathscr{U}} \|y(u)\|^\alpha d\mu < +\infty.$$

We saw that if $h(u) \in \mathscr{X}^*$ is such that

1° $h[u; x]$ is measurable in u for every fixed x, which means that $\|h(u)\|$ is measurable, since $\mathscr{X}^*$ is separable (*cf.* Chapter I);

2°
$$\int_{\mathscr{U}} \|h(u)\|^{\frac{\alpha}{\alpha-1}} d\mu < +\infty,$$

$h(u)$ defines a continuous linear functional $\overset{\alpha}{\text{X}}{}^*$ on $\overset{\alpha}{\mathscr{X}}$ by the formula

$$\overset{\alpha}{\text{X}}{}^* \left(\overset{\alpha}{\text{X}}\right) = \int_{\mathscr{U}} h[u; x(u)]d\mu,$$

where $x(u)$ is the value of the X corresponding to $\overset{\alpha}{\text{X}}$ for the outcome u. As a consequence of the above, we have

$$\int_{\mathscr{U}} |h[u; y(u)]|d\mu < +\infty$$

for every $h(u)$ satisfying the conditions 1° and 2° above. It suffices to use the proof c on page 19 again to deduce that

$$\int_{\mathscr{U}} \|y(u)\|^\alpha d\mu < +\infty,$$

the role of α being played by $\frac{\alpha}{\alpha-1}$ and *vice versa*, $h(u)$ will play the role of $x(u)$ on page 19, $y(u)$ that of $h(u)$ on page 19; this "exchange" is normal, as $\mathscr{X}$ is reflexive.

Conclusion. – Since Y defines an $\overset{\alpha}{\mathrm{Y}} \in \overset{\alpha}{\mathscr{X}}$,

$$\int_{\mathscr{U}} h[u; y(u)] d\mu = \overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{Y}}\right),$$

and we proved that, for every $\overset{\alpha}{\mathrm{X}^*} \in \overset{\alpha}{\mathscr{X}}{}^*$,

$$\lim_{n \to +\infty} \overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{Y}}_n\right) = \overset{\alpha}{\mathrm{X}^*}\left(\overset{\alpha}{\mathrm{Y}}\right),$$

which shows that $\overset{\alpha}{\mathrm{Y}}_n$ converges weakly towards $\overset{\alpha}{\mathrm{Y}}$; then as a result, $\overset{\alpha}{\mathrm{Y}}$, which we know belongs to $\overset{\alpha}{\mathscr{X}}$, actually belongs to $\overset{\alpha}{\mathscr{X}}{}'$ more precisely, by virtue of a theorem that says that if a point belonging to a fixed closed bounded manifold $\mathscr{V}$ converges weakly to a limit, this limit belongs to $\mathscr{V}$.

Yosida's theorem can then be applied and tells us that $\overset{\alpha}{\mathrm{Y}}_n$ converges *strongly* to $\overset{\alpha}{\mathrm{Y}}$, i.e.

$$\lim_{n \to +\infty} \left\| \overset{\alpha}{\mathrm{Y}}_n - \overset{\alpha}{\mathrm{Y}} \right\| = \lim_{n \to +\infty} \left[\mathrm{E}\left(\left\| \frac{1}{n} \sum_{i=1}^{n} \mathrm{X}_i - \mathrm{Y} \right\| \right)^{\alpha} \right]^{\frac{1}{\alpha}} = 0.$$

Theorem. – *Law of large numbers in mean of order α. – If $\mathscr{X}$ is separable and reflexive, and if $\alpha > 1$, $\{X_i\}$ ($\alpha = 0 \pm 1 ...$) denote a strictly stationary sequence of elements X_i of $\overset{\alpha}{\mathscr{X}}$, there exists an element Y in $\mathscr{X}$ such that*

$$\lim_{n \to +\infty} \mathrm{E}\left(\left\| \frac{1}{n} \sum_{i=1}^{n} \mathrm{X}_i - \mathrm{Y} \right\|^{\alpha} \right) = 0.$$

Remark. – Let us return to theorem 2 (Section II); there, we proved that if Y is the random element defined by $y(u)$, save exceptional outcomes, $y_n(u)$ tends weakly to $y(u)$, which allows us to say that $\mathrm{Y}_n \to \mathrm{Y}$ almost surely. But we left the following point in the shade: whether Y is a random element of the type considered, that is to say, such that

a. $x^*[y(u)]$ is measurable, for any fixed x^*; $\|y(u)\|$ is measurable;
b. $\int_{\mathscr{U}} \|y(u)\|^{\alpha} d\mu < +\infty$ (in theorem 2, Section II, α could equal 1).

For *a*, $x^*[y(u)]$ is necessarily measurable as a limit of $x^*[y_n(u)]$ and then $\|y(u)\|$ is measurable because $\mathscr{X}$ is separable.

For *b*, if $\alpha > 1$, this was just shown above; for $\alpha = 1$, we can prove it directly by using the procedure of Landau. $\overset{\alpha}{\mathscr{X}}$ is always defined, we cannot find all the functionals of it[(9)], but some obvious ones suffice to prove *b*.

[(9)] This question is now resolved [R. Fortet and E. Mourier, I].

Chapter III.

Characteristic Function of a Random Element in a Banach Space.

Definition. – Let x^* be any real linear functional; $x^*(\mathrm{X})$ is a numerical random variable X^* equipped with a distribution function; the random variable $e^{i\mathrm{X}^*}$ is equally measurable if, as it is bounded in modulus, it has a mathematical expectation – the usual characteristic function of X^* – $\varphi(x^*)$:

$$\varphi(x^*) = \mathrm{E}[e^{i\mathrm{X}^*}] = \mathrm{E}[e^{ix^*(\mathrm{X})}].$$

However, this assumes that X^* *is a random variable in the proper sense.* By definition, X will be a random variable in the proper sense if there exists a sequence of bounded measurable sets $e_k \in \mathscr{X}$ tending to $\mathscr{X}$ and such that

$$\lim_{k\to+\infty} \mathrm{mes}(e_k) = 1.$$

If X is a random element in the proper sense, X^* is a random variable in the proper sense. In the case of the measure introduced by M. Fréchet (F-measure, with condition F'), X is always a random element in the proper sense. In all that follows we will assume that X is a random element in the proper sense.

By definition, $\varphi(x^*)$*, considered as a function of* x^* *in* $\mathscr{X}^*$*, is the characteristic function of* X. [E.Mourier, I and III] [L. Le Cam, I].

Remark. – In the case of a Euclidean space R_n with n dimensions, with X having coordinates $\mathrm{X}_1, ..., \mathrm{X}_n$, every linear functional is of the form

$$x^*(\mathrm{X}) = \nu_1\mathrm{X}_1 + ... + \nu_n\mathrm{X}_n,$$

where $\nu_1, ..., \nu_n$ are the constants defining x^* and vice versa.

In this case, the characteristic function of X has long been

$$\begin{aligned}\varphi(\nu_1, ..., \nu_n) &= \mathrm{E}[e^{i(\nu_1\mathrm{X}_1+...+\nu_n\mathrm{X}_n)}]\\ &= \mathrm{E}[e^{ix^*(X)}].\end{aligned}$$

For an arbitrary Banach space, the definition $\varphi(x^*) = \mathrm{E}[e^{ix^*(\mathrm{X})}]$ is thus the immediate generalisation of the classical characteristic function.

Theorem 1. – *If* X *and* Y *are independent random elements defined on the same* $\mathscr{X}$*, the characteristic function of* X+Y *is the product of the characteristic functions of* X *by that of* Y.

Let φ_{X}, φ_{Y} and $\varphi_{\mathrm{X+Y}}$ be the characteristic functions of X, Y and X + Y respectively. Then

$$\begin{aligned}\varphi_{\mathrm{X+Y}} &= \mathrm{E}[e^{ix^*(\mathrm{X+Y})}]\\ &= \mathrm{E}[e^{ix^*(\mathrm{X})}e^{ix^*(\mathrm{Y})}]\\ &= \varphi_X(x^*)\varphi_Y(x^*).\end{aligned}$$

Indeed, $e^{ix^*(\mathrm{X})}$ and $e^{ix^*(\mathrm{Y})}$ are two independent numerical random variables and we can apply the classical theorem of the mathematical expectation of the product of two independent random variables.

Theorem 2. – *$\varphi(x^*)$ is a uniformly continuous function in x^*, that is to say, that if a positive ε is given, we can find positive η depending only on ε such that*

$$|\varphi(x_1^*) - \varphi(x_2^*)| \leq \varepsilon,$$

provided that

$$\|x_1^* - x_2^*\| \leq \eta.$$

Indeed:

$$\begin{aligned} |\varphi(x_1^*) - \varphi(x_2^*)| &= |\mathrm{E}[e^{ix_1^*(\mathrm{X})}] - \mathrm{E}[e^{ix_2^*(\mathrm{X})}]| \\ &\leq \mathrm{E}[|e^{ix_1^*(\mathrm{X})} - e^{ix_2^*(\mathrm{X})}|] \\ &\leq \mathrm{E}[|e^{ix_1^*(\mathrm{X})}(1 - e^{iy^*(\mathrm{X})})|], \end{aligned}$$

where

$$x_2^* = x_1^* + y^*.$$

We thus have $\|y^*\| \leq \eta$; let e_k be such that

$$\mathrm{mes}(e_k) > 1 - \frac{\varepsilon}{3}.$$

We have

$$\mathrm{E}[|e^{ix_1^*(\mathrm{X})}(1 - e^{iy^*(\mathrm{X})})|] = \int_{x \in e_k} |1 - e^{iy^*(\mathrm{X})}| dm + \int_{x \in \mathscr{X} - e_k} |1 - e^{iy^*(\mathrm{X})}| dm.$$

The second term is smaller than $2\frac{\varepsilon}{3}$, because

$$|1 - e^{iy^*(x)}| \leq 2$$

and $m(\mathscr{X} - e_k)$ is smaller than $\frac{\varepsilon}{3}$.

As for

$$\int_{x \in e_k} |1 - e^{iy^*(x)}| dm,$$

if $\|y^*\| \leq \eta$,

$$|1 - e^{iy^*(x)}| \leq \frac{\varepsilon}{3},$$

because e_k is bounded, so there exists M such that on e_k, $\|x\| \leq \mathrm{M}$; so

$$|\varphi(x_1^*) - \varphi(x_2^*)| \leq \varepsilon.$$

Theorem 3. – *$\varphi(x^*)$ is continuous in the sense of weak convergence in $\mathscr{X}^*$.*

As in the previous theorem, we have

$$|\varphi(x_1^*) - \varphi(x_2^*)| \leq \mathrm{E}|1 - e^{i(x_2^* - x_1^*)(\mathrm{X})}| = \int_{e_k} + \int_{\mathscr{X} - e_k},$$

where we again have e_k bounded and such that $m(e_k) > 1 - \frac{\varepsilon}{3}$, so the second term is $\leq 2\frac{\varepsilon}{3}$ as in Theorem 2.

Saying that x_2^* converges weakly to x_1^* is to say that for every x^{**} we have

$$|x^{**}(x_2^* - x_1^*)| \to 0.$$

So for every fixed x, the weak convergence of x_2^* to x_1^* implies that

$$(x_2^* - x_1^*)(x) \to 0.$$

The measurable and bounded function $e^{i(x_2^*-x_1^*)(x)}$ thus tends to 1 *everywhere* and

$$\int_{e_k} |1 - e^{i(x_2^*-x_1^*)(x)}| dm \to 0.$$

THEOREM 4. – *If* E(X) *and* $\mathrm{E}[\|\mathrm{X}\|^2]$ *exist*[10] *the characteristic function* $\varphi(x^*)$ *of* X *takes the form*

$$\varphi(x^*) = 1 + ix^*[\mathrm{E(X)}] - \frac{1}{2}\mathrm{E}[|x^*(\mathrm{X})|^2] + \|x^*\|^2\omega(x^*),$$

where

$$\omega(x^*) \to 0 \qquad \text{if} \qquad \|x^*\| \to 0.$$

Proof. – Let $z^* \in \mathscr{X}^*$ be such that $\|z^*\| = 1$ and $x^* = \lambda z^*$, which implies that $\lambda = \pm\|x^*\|$. Let us set $\mathrm{U} = z^*(\mathrm{X})$. Then

$$\mu_1 = \mathrm{E(U)} = \mathrm{E}[z^*(\mathrm{X})] = z^*[\mathrm{E(X)}]$$

exists since E(X) exists. Likewise, $\mathrm{E(U^2)}$ exists because $|\mathrm{U}| \leq \|\mathrm{X}\|$ and $\mathrm{E}[\|\mathrm{X}\|^2]$ exists.

$$\varphi(x^*) = \mathrm{E}[e^{ix^*}(\mathrm{X})] = \mathrm{E}[e^{i\lambda z^*}(\mathrm{X})] = \mathrm{E}[e^{i\lambda \mathrm{U}}].$$

As a function of λ, $\varphi(x^*)$ is thus the characteristic function of the random variable U, and as a consequence,

$$\text{(1)} \qquad \varphi(x^*) = 1 + i\mu_1 - \frac{1}{2}\mathrm{E(U^2)}\lambda^2 + \lambda^2\omega_1(z^*, \lambda),$$

where, for every fixed z^*, $\omega(z^*;\lambda) \to 0$ if $\lambda \to 0$. But (1) can be written as

$$\varphi(x^*) = 1 + ix^*[\mathrm{E(X)}] - \frac{1}{2}\mathrm{E}[x^*(\mathrm{X})^2] + \|x^*\|^2\omega_1(z^*;\lambda).$$

The convergence of $\omega_1(z^*;\lambda)$ to 0 when $\lambda \to 0$ is uniform with respect to z^*. Indeed, with $\mathrm{F}_{z^*}(\alpha)$ denoting the distribution function of U, the mean value theorem gives

$$\omega_1(z^*;\lambda) = \frac{1}{2}\int_{-\infty}^{+\infty} \alpha^2[1 - e^{i\lambda_0\alpha}]d\mathrm{F}_{z^*}(\alpha), \qquad \text{where} \qquad 0 \leq \lambda_0 \leq \lambda,$$

and if M is any positive number,

$$\omega_1(z^*;\lambda) = \frac{1}{2}\int_{|\alpha|>\mathrm{M}} \alpha^2[1 - e^{i\lambda_0\alpha}]d\mathrm{F}_{z^*}(\alpha) + \frac{1}{2}\int_{-M}^{+M}.$$

Denoting by m the l-measure on $\mathscr{X}$, the first term of the right-hand side is bounded by

$$\int_{|\mathrm{U}|>\mathrm{M}} \mathrm{U}^2 dm \leq \int_{|\mathrm{U}|>\mathrm{M}} \|\mathrm{X}\|^2 dm \leq \int_{\|\mathrm{X}\|>\mathrm{M}} \|\mathrm{X}\|^2 dm,$$

[10] We know that if $\mathscr{X}$ is separable and reflexive, $\mathrm{E}[\|\mathrm{X}\|] < +\infty$, which is implied by $\mathrm{E}[\|\mathrm{X}\|^2] < +\infty$, implies the existence of E(X).

because $|\mathrm{U}| \leq \|\mathrm{X}\|$ and the domain $(|\mathrm{U}| > \mathrm{M})$ is contained in the domain $(\|\mathrm{X}\| > \mathrm{M})$.

By hypothesis, we can find M independent of z^* such that

$$\int_{\|\mathrm{X}\|>\mathrm{M}} \|\mathrm{X}\|^2 dm < \frac{\varepsilon}{2}.$$

Having thus chosen M, it suffices that λ_0, and so λ, is small enough to have $|e^{i\lambda_0\alpha} - 1| < \frac{\varepsilon}{\mathrm{M}^2}$ for any α in $(-\mathrm{M}, +\mathrm{M})$ and then

$$\frac{1}{2}\int_{-\mathrm{M}}^{+\mathrm{M}} \alpha^2[1 - e^{i\lambda_0\alpha}]d\mathrm{F}_{z^*}(\alpha) < \frac{\varepsilon}{2} \qquad \text{for any } \mathrm{F}_{z^*}(\alpha).$$

So the convergence of $\omega_1(z^*;\lambda) \to 0$ as $\lambda \to 0$ is uniform in z^*, which establishes the theorem.

Definition. – Given a real or complex numerical function $\varepsilon(x^*)$, we will say that it is *positive definite* if:

1° it is continuous in the sense of strong convergence in $\mathscr{X}^*$;
2° for any n, $x_1^*, ..., x_n^*$ and complex numbers $\alpha_1, ..., \alpha_n$, we have that

$$\sum_{j,k}^{n} \varphi(x_j^* - x_k^*)\alpha_j\bar{\alpha}_k \qquad \text{is real and non-negative.}$$

THEOREM 5. – *Every characteristic function is positive definite.*

We saw that it is continuous (Theorem 2). The second condition is also satisfied; indeed,

$$\begin{aligned} 0 \leq |\alpha_1 u_1 + ... + \alpha_n u_n|^2 &= (\alpha_1 u_1 + ... + \alpha_n u_n)(\overline{\alpha_1 u_1 + ... + \alpha_n u_n}) \\ &= (\alpha_1 u_1 + ... + \alpha_n u_n)(\bar{\alpha}_1 \bar{u}_1 + ... + \bar{\alpha}_n \bar{u}_n) \\ &= \sum_{j,k}^{n} \alpha_j \bar{\alpha}_k u_j \bar{u}_k; \end{aligned}$$

it then suffices to take $u_j = e^{ix_j^*(\mathrm{X})}$:

$$u_j\bar{u}_k = e^{ix_j^*(\mathrm{X})}e^{-ix_k^*(\mathrm{X})} = e^{i[x_j^*(\mathrm{X}) - x_k^*(\mathrm{X})]}$$

and then

$$\sum_{j,k} \alpha_j\bar{\alpha}_k e^{i[x_j^* - x_k^*](\mathrm{X})} \qquad \text{being real and non-negative}$$

means that

$$\mathrm{E}\left[\sum_{j,k} \cdots\right] = \sum_{j,k} \alpha_j\bar{\alpha}_k\varphi(x_j^* - x_k^*) \qquad \text{is real and non-negative.}$$

Remark. – Straight from the definition of the characteristic function we have that $\varphi(0^*) = 1$ and that

$$\varphi(-x^*) = \overline{\varphi(x^*)}.$$

We are going to see that if a function $\varphi(x^*)$ satisfies the condition

$$\sum_{j,k} \alpha_j \bar{\alpha}_k \varphi(x_j^* - x_k^*) \qquad \text{is real and non-negative}$$

for any n, $x_1^*, ..., x_n^*$ and $\alpha_1, ..., \alpha_n$, then $\varphi(0^*)$ is real and non-negative, and is zero only if $\varphi \equiv 0$:

$$\varphi(-x^*) = \overline{\varphi(x^*)};$$

if φ is continuous at the origin, it is continuous everywhere and also if $\varphi(x^*) \to \varphi(0^*)$ when x^* tends weakly to 0^*, then $\varphi(y^* - x^*)$ tends to $\varphi(y^*)$ for any y^* when $x^* \to 0^*$.

Let us take $n = 2$:

$$\alpha_1 \bar{\alpha}_2 \varphi(x_1^* - x_2^*) + \bar{\alpha}_1 \alpha_2 \varphi(x_2^* - x_1^*) + [|\alpha_1|^2 + |\alpha_2|^2] \varphi(0^*) \text{ is real and non-negative;}$$

if $x_2^* = 0^*$:

$$\alpha_1 \bar{\alpha}_2 \varphi(x_1^*) + \alpha_2 \bar{\alpha}_1 \varphi(-x_1^*) + [|\alpha_1|^2 + |\alpha_2|^2] \varphi(0^*) \text{ is real and non-negative;}$$

$\alpha_2 = 0$ shows that $\varphi(0^*)$ *is real and non-negative.*

$\alpha_1 = \alpha_2 = 1$:

$$\varphi(x_1^*) + \varphi(-x_1^*) \qquad \text{is real,}$$

so

$$\mathcal{I}\varphi(x_1^*) = -\mathcal{I}\varphi(-x_1^*).$$

$\alpha_1 = i, \alpha_2 = 1$:

$$i\varphi(x_1^*) - i\varphi(-x_1^*) \qquad \text{is real,}$$

so

$$\mathcal{R}\varphi(x_1^*) = \mathcal{R}\varphi(-x_1^*),$$

so

$$\varphi(-x^*) = \overline{\varphi(x^*)}.$$

Let us take $n = 3$:

$$\alpha_1 = 1, \quad \alpha_2 = -1, \quad \alpha_3 = \lambda, \quad x_1^* = 0^*, \quad x_2^* = x^*, \quad x_3^* = y^*.$$

We have

$$[2 + |\lambda|^2]\varphi(0^*) - \varphi(x^*) - \varphi(-x^*) \\ + \lambda[\varphi(y^*) - \varphi(y^* - x^*)] + \bar{\lambda}[\varphi(-y^*) - \varphi(x^* - y^*)] \geq 0.$$

By letting $\mathrm{A} = \varphi(y^*) - \varphi(y^* - x^*)$, we have

$$[2 + |\lambda|^2]\varphi(0^*) - \varphi(x^*) - \varphi(-x^*) + \lambda\mathrm{A} + \overline{\lambda\mathrm{A}} \geq 0$$

and by taking $\lambda = -\frac{\bar{\mathrm{A}}}{\varphi(0^*)}$,

$$\frac{|\mathrm{A}|^2}{\varphi(0^*)} - \frac{|\mathrm{A}|^2}{\varphi(0^*)} - \frac{|\mathrm{A}|^2}{\varphi(0^*)} + 2\varphi(0^*) - \varphi(x^*) - \varphi(-x^*) \geq 0.$$

Then

$$|\varphi(y^*) - \varphi(y^* - x^*)|^2 \leq \varphi(0^*)[2\varphi(0^*) - \varphi(x^*) - \varphi(-x^*)],$$

which shows that $\varphi(0^*)$ cannot be zero unless φ is identically zero, and that if φ is continuous at the origin, it is continuous everywhere and even uniformly continuous everywhere, and also that if $\varphi(x^*) \to \varphi(0^*)$ when x^* converges weakly to 0^*, we have $\varphi(y^* - x^*) \to \varphi(y^*)$ for any y^* as $x^* \to 0^*$ weakly.

Definition of a cylindrical set. – A cylindrical set $\mathscr{E}_n$ of $\mathscr{X}$ is defined as the set of all the $x \in \mathscr{X}$ such that

$$[x_1^*(x), ..., x_n^*(x)] \in \mathrm{E}_n$$

where E_n is a measurable set in the Euclidean space R_n with n dimensions.

Definition of the field $\mathscr{B}$. – All the cylindrical sets are L-measurable; they form a field $\mathscr{C}$; if $\mathscr{F}$ is the σ-algebra in the definition of the L-measure m, we have

$$\mathscr{C} \subset \mathscr{F}.$$

Let $\mathscr{B}$ be the smallest σ-algebra containing $\mathscr{C}$, we necessarily have $\mathscr{B} \subset \mathscr{F}$. Now, a completely additive function on a field can be uniquely extended to the smallest σ-algebra which contains it [Kolmogoroff]; so: the knowledge of m on $\mathscr{C}$ determines m on $\mathscr{B}$. Let us note that $\mathscr{C}$ and $\mathscr{B}$ have a geometric meaning independent of m and that they do not depend on the norm in the sense of page 3 (Chapter I).

THEOREM. – *If $\mathscr{X}$ is separable, $\mathscr{B}$ contains the spheres and all the sets defined by $f(x) < a$, where a is a real number and f a continuous real function defined for $x \in \mathscr{X}$. More generally, $\mathscr{B}$ contains all the open sets $\subset \mathscr{X}$.*

1° *Case of* $f(x) = \|x\|$. – The proof consists of proving that every sphere S with centre 0 belongs to $\mathscr{B}$. The proof is analogous to that of Pettis [Pettis, I].

Let us denote by $\bar{x^*}$ the x^* such that $\|x^*\| = 1$ and let Γ be the set of $\bar{x^*}$. Let $\mathrm{A}_{\bar{x^*}}$ be the set defined by $\left|\bar{x^*}(x)\right| \leq a$; the sphere S defined by $\|x\| \leq a$ is the product $\prod \mathrm{A}_{\bar{x^*}}$, where the product is extended to every $\bar{x^*} \in \Gamma$ (there are uncountably many $\mathrm{A}_{\bar{x^*}}$).

Indeed, $\prod \mathrm{A}_{\bar{x^*}} \subset \mathrm{S}$: let $x \in \prod \mathrm{A}_{\bar{x^*}}$, and let $\bar{x_0^*}$ be a functional $\bar{x^*}$ such that $\bar{x_0^*}(x) = \|x\|$. Then by hypothesis, we have: $\bar{x_0^*}(x) \leq a$, so $\|x\| \leq a$ and $\mathrm{S} \subset \prod \mathrm{A}_{\bar{x^*}}$, because if $\|x\| \leq a$,

$$\left|\bar{x^*}(x)\right| \leq \|\bar{x^*}\| \cdot \|x\| = 1 \cdot a = a.$$

As $\mathscr{X}$ is separable, there exists a countable sequence $\left\{\bar{x_i^*}\right\}$ of $\bar{x^*}$ weakly dense in Γ [Banach, I, p.124]. Let A_i be the set defined by

$$\left|\bar{x_i^*}(x)\right| \leq a$$

and $\mathrm{A} = \prod_i \mathrm{A}_i$. A is a product of countably many cylindrical sets A_i, so it is in $\mathscr{B}$.

We have

$$\mathrm{A} = \prod \mathrm{A}_{\bar{x^*}} = \mathrm{S}.$$

Firstly,

$$\prod \mathrm{A}_{\bar{x^*}} \subset \mathrm{A} = \prod_i \mathrm{A}_i,$$

which is evident since $\prod A_{\bar{x^*}}$ is extended to all $\bar{x^*} \in \Gamma$ and since $\prod_i A_i$ is extended to $\bar{x^*} \in \{\bar{x_i^*}\}$, a countable subset of Γ.

On the other hand, if $x_1 \in A$ does not belong to $\prod A_{\bar{x^*}}$, we would have:

$$\left|\bar{x_i^*}(x_1)\right| \leq a \qquad \text{for all } i$$

and

$$\left|\bar{x^*}(x_1)\right| > a \quad \text{for at least one } \bar{x^*};$$

let $\bar{x_0^*}$ be such a $\bar{x^*}$, and $\{\bar{x_{i'}^*}\}$ a subsequence of $\{\bar{x_i^*}\}$ converging weakly to $\bar{x_0^*}$:

$$\left|\bar{x_0^*}(x)\right| = \lim \left|\bar{x_{i'}^*}(x)\right| \quad \text{for every } x \in \mathscr{X},$$

so

$$\left|\bar{x_0^*}(x_1)\right| = \lim \left|\bar{x_{i'}^*}(x_1)\right| \leq a, \qquad \text{contradiction.}$$

So

$$A \subset \prod A_{\bar{x^*}}.$$

As a consequence,

$$A = \prod A_{\bar{x^*}}.$$

We saw that

$$\prod A_{\bar{x^*}} = S,$$

so

$$A = \prod A_{\bar{x^*}} = S,$$

so S is in $\mathscr{B}$.

2° *Case of* $f(x) = \|x - x_0\|$. – Spheres of centre x_0, same proof.

3° *Case of arbitrary* $f(x)$. – The set defined by $f(x) < a$ is open. Now, every open subset Ω of a separable space $\mathscr{X}$ is the union of countably many spheres.

Indeed, as $\mathscr{X}$ is separable, let $\{x_n\}$ be a countable dense sequence in $\mathscr{X}$ and let x'_n those of x_n that are in the interior of Ω. Let $S\left(x'_n; \frac{1}{k}\right)$ be the sphere with centre x'_n and of radius $\frac{1}{k}$ (with k an integer). Among these spheres, some are completely contained in Ω: these are the S_1; the others have some points outside Ω. Let A be the union of the S_1. There are a countably infinite number of S_1 and every sphere is in $\mathscr{B}$, so A is in $\mathscr{B}$.

We have $A = \Omega$; to prove this, it suffices to show that every $x \in \Omega$ belongs to an S_1. Now, x belongs to $S\left(x'_n; \frac{1}{k}\right)$ if $\|x'_n - x\| \leq \frac{1}{k}$ and this sphere is an S_1 if $\frac{2}{k} < \delta$, where we denoted by δ the lower bound of $\|x - y\|$ as y runs through the boundary of Ω. So it suffices to take $\frac{1}{k} < \frac{\delta}{2}$, then x'_n close enough to x so that $\|x'_n - x\| \leq \frac{1}{k}$, which is possible since we have a dense sequence.

So every set defined by $f(x) < a$ and more generally every open set is in $\mathscr{B}$.

Remark. – From the very definition of the characteristic function,

$$\varphi(x^*) = E[e^{ix^*(X)}]$$

we have that every L-measure on $\mathscr{X}$, strictly speaking, defines a unique characteristic function; conversely, we are going to see that

THEOREM 6. – *The characteristic function determines the* L*-measure on every cylindrical set, and so on* $\mathscr{B}$*.*

Let $X^* = x^*(X)$ if we know $\varphi(x^*) = E[e^{ix^*(X)}]$, we know the characteristic function φ_1 of X^*:

$$\varphi_1(\nu) = E[e^{i\nu X^*}] = E[e^{i\nu x^*(X)}] = \varphi(\nu x^*),$$

so we know $\Pr[x^*(X) \in E_1]$, where E_1 is a measurable set on the real line.

Let $x_1^*, ..., x_n^*$ and $X_1^* = x_1^*(X), ..., X_n^* = x_n^*(X)$, we know $\varphi(x^*)$, so we know the characteristic function $\varphi_1(\nu_1, ..., \nu_n)$ of the random variable with n dimensions $\{X_1^*, ..., X_n^*\}$:

$$\varphi_1(\nu_1, ..., \nu_n) = E[e^{i(\nu_1 X_1^* + ... + \nu_n X_n^*)}] = \varphi(\nu_1 x_1^* + ... + \nu_n x_n^*).$$

So we know the measure on every cylindrical set, and so on $\mathscr{B}$.

Remark. – The proof of the preceding theorem relies on the well-known property that in a Euclidean space with a finite number of dimensions, the characteristic function completely defines the probability law of the random point. It is likewise true in the case of a Banach space $\mathscr{X}$ possessing a base, that is to say, such that there exists a countable sequence of distinct elements $e_1, ..., e_n, ...$ of $\mathscr{X}$ such that, for any $x \in \mathscr{X}$, there is a unique sequence of numbers $x_1, ..., x_n, ...$ such that

$$x = x_1 e_1 + ... + x_n e_n + ...,$$

with $\lim_{n\to\infty} \|x - x(n)\| = 0$ by letting

$$x(n) = x_1 e_1 + ... + x_n e_n.$$

Knowing $\varphi(x^*)$ for every x^*, we know in particular:

$$\varphi(x_n^*) = E[e^{i[t_1 X_1 + ... + t_n X_n]}]$$

which is the characteristic function of the point $X(n)$ in R_n and determines the distribution function of $X(n)$. Let

$$F_n(x_1, ..., x_n) = \Pr[X_1 < x_1, ..., X_n < x_n];$$

if we set

$$F(x_1, ..., x_n, ...) = \Pr[X_1 < x_1, ..., X_n < x_n, ...],$$

we have

$$F(x_1, ..., x_n, ...) = \lim_{n\to\infty} F_n(x_1, ..., x_n).$$

So the characteristic function $\varphi(x^*)$ then completely determines the law of X.

Generalisation of Bochner's theorem [Bochner, II, p.239]. – We know that the characteristic function of an ordinary random variable is a positive definite function which equals 1 at the origin, and we also know that, conversely, every positive definite function φ such that $\varphi(0) = 1$ is the characteristic function of a probability law. In the case of a random element X with values in a Banach space $\mathscr{X}$, we just saw that the characteristic function of X, $\varphi(x^*) = E[e^{ix^*(X)}]$, is again a positive definite function, and $\varphi(0^*) = 1$. We are thus led to ask ourselves whether every positive definite function $\varphi(x^*)$ such that $\varphi(0^*) = 1$ is the characteristic function of a random element (in the proper sense) on $\mathscr{B}$.

So let $\varphi(x^*)$ be a positive definite function on $\mathscr{X}^*$ and such that $\varphi(0^*) = 1$. Let $x_1^*, ..., x_n^*$ be n deterministic linear functionals that are linearly independent, and let

$$\bar{x^*} = \nu_1 x_1^* + ... + \nu_n x_n^*,$$

where $\nu_1, ..., \nu_n$ are any n real numerical variables; $\varphi(\bar{x^*})$ considered as a function of $\nu_1, ..., \nu_n$ is a positive definite function with n dimensions, in a Euclidean space with n dimensions, and it takes value 1 if $\nu_1 = \nu_2 = ... = \nu_n = 0$; so it is a characteristic function in n dimensions (Bochern-Weil) and it defines, as a consequence, a distribution function in n dimensions: $\mathrm{F}_n(x_1^*, x_1; x_2^*, x_2; ...; x_n^*, x_n)$.

F_n has the following properties:

1° As a function of the pairs (x_j^*, x_j), it is a symmetric function;
2° As a function of $x_1, ..., x_n$, it is a distribution function;
3° Moreover, we have

$$\mathrm{F}_{n+1}(x_1^*, x_1; ...; x_n^*, x_n; x_{n+1}^*, +\infty) = \mathrm{F}_n(x_1^*, x_1; ...; x_n^*, x_n).$$

Indeed,

$$\varphi(\bar{x^*} + \nu_{n+1} x_{n+1}^*) = \int e^{i(\nu_1 x_1 + ... + \nu_{n+1} x_{n+1})} d\mathrm{F}_{n+1}(x_1^*, x_1; ...; x_n^*, x_n; x_{n+1}^*, x_{n+1});$$

if $\nu_{n+1} = 0$,

$$\varphi(\bar{x^*} + 0) = \varphi(\bar{x^*}) = \int e^{i(\nu_1 x_1 + ... + \nu_n x_n)} d\mathrm{F}_{n+1}(x_1^*, x_1; ...; x_{n+1}^*, x_{n+1}).$$

By integrating with respect to x_{n+1},

$$\varphi(\bar{x^*}) = \int e^{i(\nu_1 x_1 + ... + \nu_n x_n)} d\Phi(x_1^*, x_1, ..., x_n^*, x_n)$$

by letting

$$\Phi(x_1^*, x_1; ...; x_n^*, x_n) = \mathrm{F}_{n+1}(...x_n^*, x_n; x_{n+1}^*, +\infty),$$

but by the definition of F_n:

$$\varphi(\bar{x^*}) = \int e^{i(\nu_1 x_1 + ... + \nu_n x_n)} d\mathrm{F}_n(x_1^*, x_1; ...; x_n^*, x_n),$$

so

$$\begin{aligned} \mathrm{F}_n(x_1^*, x_1; ...; x_n^*, x_n) &= \Phi(x_1^*, x_1; ...; x_n^*, x_n) \\ &= \mathrm{F}_{n+1}(x_1^*, x_1; ...; x_n^*, x_n; x_{n+1}^*, +\infty). \end{aligned}$$

Let $\mathscr{E}_n$ be a cylindrical set in $\mathscr{X}$, defined as the set of all the $x \in \mathscr{X}$ such that

$$\{x_1^*(x), ..., x_n^*(x)\} \in \mathrm{E}_n,$$

where E_n is a Borel measurable set in the Euclidean space R_n with n dimensions. F_n defines a measure in R_n for which E_n is measurable; let $\mu(\mathrm{E}_n)$ be its measure and let

$$m(\mathscr{E}_n) = \mu(\mathrm{E}_n),$$

we thus define a set function $m(\epsilon)$ on the field of cylindrical sets in $\mathscr{X}$; we see without difficulty that

1° $m(\mathscr{X}) = 1$;
2° $m(\mathscr{E}_n) \geq 0$;
3° $m(\mathscr{E}) + m(\mathscr{E}') = m(\mathscr{E} + \mathscr{E}')$ if $\mathscr{E}$ and $\mathscr{E}'$ are two disjoint cylindrical sets [Kolmogoroff, I, p.27].

It remains to see if $m(\mathscr{E})$ is completely additive (on the field of cylindrical sets) or, equivalently, to try to prove that if $\mathscr{E}^1, \mathscr{E}^2, ..., \mathscr{E}^k, ...$ are cylindrical sets such that

$$\mathscr{E}^1 \supset \mathscr{E}^2 \supset ... \supset \mathscr{E}^k \supset ...$$

and if $\lim_{k\to\infty} m(\mathscr{E}^1) = \mathrm{L} > 0$, $\mathscr{E}^2 = \lim_{k\to\infty} \mathscr{E}^k$ is not empty.

Let $\{x_j^*\}(j = 1, 2, ...)$ be any sequence of linear functionals of norm 1 and such that any finite number of these functionals are linearly independent. Let $\mathrm{S}(a)$ be the sphere of $\mathscr{X}$ of centre 0 and of radius a. Let A_n be the set of points of R_n of coordinates $\alpha_1, ..., \alpha_n$ such that the system

$$x_1^*(x) = \alpha_1, \qquad ..., \qquad x_n^*(x) = \alpha_n$$

has at least one solution x such that $\|x\| \leq \alpha$.

If $\mathscr{X}$ is reflexive, A_n *is closed.* – First, A_n is bounded, which is obvious since

$$\alpha_j = x_j^*(x) \leq \|x_j^*\| \cdot \|x\| \leq \alpha;$$

then, let $\mathrm{P}^k(\alpha_j^k)$ be a sequence of points in A_n converging to a point P in R_n with coordinates (α_j), so $\alpha_j^k \to \alpha_j$. We have $\mathrm{P} \in \mathrm{A}_n$, that is to say, there exists an $x \in \mathrm{S}(a)$ such that

$$x_j^*(x) = \alpha_j, \qquad (j = 1, 2, ..., n).$$

A necessary and sufficient condition for that (E.Hille, I, p.21) is that, for any numbers a_j $(j = 1, 2, ..., n)$, we have

$$\left|\sum_{j=1}^{n} a_j x_j\right| \leq a \left\|\sum_{j=1}^{n} a_j x_j^*\right\|, \tag{1}$$

but the same theorem implies that, since $\mathrm{P}^k \in \mathrm{A}_n$,

$$\left|\sum_{j=1}^{n} a_j \alpha_j^k\right| \leq a \left\|\sum_{j=1}^{n} a_j x_j^*\right\|. \tag{2}$$

To obtain (1), it suffices to let k tend to $+\infty$ in (2).

Let $\mathscr{A}_n$ be the set of the $x \in \mathscr{X}$ such that the point P in $\mathscr{R}_n$ with coordinates $[x_1^*(x), ..., x_n^*(x)]$ belong to A_n. We obviously have

$$\mathscr{A}_n \supset \mathrm{S}(a).$$

$\mathscr{A}_n$ is a cylindrical set, closed in $\mathscr{X}$; it is not bounded in general; since it is closed, if $x_1, ..., x_k, ... \in \mathscr{A}_n$ and if $x_k \to x$, x belongs to $\mathscr{A}_n$.

Every cylindrical set $\mathscr{E}_n$ defined by $x_1^*, ..., x_n^*$ which contains $\mathrm{S}(a)$ contains $\mathscr{A}_n$.

On the other hand, $\mathscr{A}_{n+1} \subset \mathscr{A}_n$; indeed, if $x_0 \in \mathscr{A}_{n+1}$, the numbers α_j $(j = 1, 2, ..., n+1)$ defined by

$$x_j^*(x_0) = \alpha_j \qquad (j = 1, 2, ..., n+1)$$

are such that the equations

$$x_j^*(x) = \alpha_j \qquad (j = 1, 2, ..., n+1)$$

admit at least one solution x' such that $\|x'\| \leq a$, so the equations

$$x_j^*(x) = \alpha_j \qquad (j = 1, 2, ..., n)$$

admit a solution $x' \in \mathrm{S}(a)$, so $x \in \mathscr{A}_n$.

Remark. – If R_{n+1} corresponding to $x_1^*, ..., x_n^*, x_{n+1}^*$ is constructed as a product of R_n corresponding to $x_1^*, ..., x_n^*$ and a straight line, A_n is the projection parallel to this straight line, from R_{n+1} onto R_n.

Let

$$\mathscr{L} = \lim_{n \to +\infty} \mathscr{A}_n, \qquad \mathscr{L} = \prod_i \mathscr{A}_i \qquad \text{and} \qquad \mathscr{L} \supset \mathrm{S}(a).$$

If the sequence $\{x_i^*\}$ is dense on the sphere of radius 1 on $\mathscr{X}^*$ (which implies that $\mathscr{X}^*$, and so $\mathscr{X}$ is not only reflexive, but also separable), $\mathscr{L} = \mathrm{S}(a)$.

Indeed, if $x_0 \in \mathscr{L}$ is not in $\mathrm{S}(a)$ we would have $\|x_0\| > a$ and

$$|x_i^*(x_0)| = |x_i| \leq a \qquad \text{for all } i;$$

but, on the other hand, there exists [E. Hille, I, theorem 2.9.3] $x^* \in \mathscr{X}^*$ such that

$$x^*(x_0) = \|x_0\| \qquad \text{and} \qquad \|x^*\| = 1.$$

As the sequence $\{x_i^*\}$ is dense on the unit sphere, let $\{x_{i'}^*\}$ be a sequence tending to x^*

$$|x^*(x_0)| = \lim |x_{i'}^*(x_0)| \leq a,$$

which contradicts

$$|x^*(x_0)| = \|x_0\| > a.$$

So $\mathscr{L} \subset \mathrm{S}(a)$, and since we always have $\mathscr{L} \supset \mathrm{S}(a)$:

$$\mathscr{L} = \mathrm{S}(a).$$

Let us now take these conditions, and let $\rho(a)$ be the limit as $n \to +\infty$ of $m(\mathscr{A}_n)$; then $m(\mathscr{A}_{n+1}) \leq m(\mathscr{A}_n)$. We will say that $\varphi(x^*)$ satisfies the condition C if, for any sequence $\{x_i^*\}$ dense on the unit sphere of $\mathscr{X}^*$, $\lim_{a \to +\infty} \rho(a) = 1$. It is clear that

$$\rho(a) \leq \rho(a') \qquad \text{if } a < a'.$$

Let us consider the cylindrical sets $\mathscr{E}^k$ such that

$$\mathscr{E}^1 \supset \mathscr{E}^2 \supset ...\mathscr{E}^k \supset \mathscr{E}^{k+1} \supset ... \qquad \text{and} \qquad \lim m(\mathscr{E}^k) = l.$$

We are going to show that

$$\lim_k \mathscr{E}^k = \lim \left[\mathscr{B}_k = \prod_{j=1}^{k} \mathscr{E}^j \right]$$

is not empty.

We do not lose generality by considering a sequence of linear functionals $\{x_k^*\}$ of norm 1, such that any finite number of them are linearly independent and such that the functionals defining ε^k are $x_1^*, ..., x_k^*$. Let E_n be the corresponding set in R_n of $\mathscr{E}_n$; we set

$$m(\mathscr{E}_n) = \mu(\mathrm{E}_n).$$

We can find in R_n a bounded closed set D_n contained in E_n such that

$$\mu(\mathrm{D}_n) \geq \mu(\mathrm{E}_n) - \frac{\varepsilon}{2^n};$$

D_n defines in $\mathscr{X}$ a cylindrical set $\mathscr{D}_n \subset \mathscr{E}_n$, closed but not necessarily bounded and

$$m(\mathscr{D}_n) \geq m(\mathscr{E}_n) - \frac{\varepsilon}{2^n}.$$

Let $\mathscr{C}_n = \prod_{k=1}^n \mathscr{D}_k$ and let w_n be in R_n defining $\mathscr{C}_n$; w_n is closed, bounded, contained in D_n:

$$m(\mathscr{C}_n) \geq m(\mathscr{B}_n) - \varepsilon \geq l - \varepsilon,$$

so

$$\mu(w_n) \geq l - \varepsilon.$$

Let $\mathrm{S}(a)$ be the sphere in $\mathscr{X}$ of radius a and let us take a sequence of linear functionals of norm 1, dense on the sphere of radius 1 in $\mathscr{X}^*$, such that any finite number of these functionals are linearly independent, and such that it contains $\{x_n^*\}$ as a partial sequence; we do not lose generality by assuming that these are the x_n^* themselves. Let us consider the sets $\mathscr{C}_n \cdot \mathrm{S}(a)$; if none of them is empty, they have a non-empty limit. Indeed, if $\mathscr{C} \cdot \mathrm{S}(a)$ is assumed to be non-empty for any n, let $x_n \in \mathscr{C}_n \cdot \mathrm{S}(a)$, then the sequence x_n is bounded since $x_n \in \mathrm{S}(a)$; as $\mathscr{X}$ is reflexive, a subsequence, which we will assume to be the sequence itself, has a weak limit x; let $\alpha_j^n = x_j^*(x_n)$ and let P_k^n be the point in R_k with coordinates $(\alpha_1, ..., \alpha_k^n)$, then $\mathrm{P}_k^n \in w_k$ because $x_n \in \mathscr{C}_k \cdot \mathrm{S}(a)$ for $k \leq n$ at least, so $x_n \in \mathscr{C}_k$. If P_k is the point in R_k with coordinates $\alpha_j = x_j^*(x)$ $(j = 1, 2, ..., k)$, $\mathrm{P}_k = \lim_{n\to\infty} \mathrm{P}_k^n$, so $\mathrm{P}_k \in w_k$, so $x \in \mathrm{C}_k$, and this holds for all k, so x belongs to $\lim_{n\to\infty} \mathrm{C}_n$, so *a fortiori* x belongs to $\lim_{n\to\infty} \mathscr{E}_n$ which is thus not empty.

Let us assume that $\mathrm{C}_N \cdot \mathrm{S}(a)$ is empty, so $\mathrm{C}_n \cdot \mathrm{S}(a)$ is empty for any $n \geq N$. I say that if $\mathrm{C}_n \cdot \mathrm{S}(a)$ is empty, $\mathrm{C}_n \mathscr{A}_l$ is empty, at least for $l \geq n$. C_n can be represented in R by an "extension" w_n' of w_n, and if w_n' and A_l have a common point P there is a $x \in \mathrm{S}(a)$ which is in C_n, which contradicts the hypothesis, so w_n' and A_l do not have a common point, so $\mathrm{C}_n \mathscr{A}_l = 0$ and then as C_n is outside $\mathscr{A}_l$ (for any $l \geq n$)

$$m(\mathscr{C}_n) = m[\mathscr{X} - \mathscr{A}_l] = 1 - m(\mathscr{A}_l),$$

so

$$m(\mathscr{A}_l) \leq 1 - m(\mathrm{C}_n) \leq 1 - l + \varepsilon,$$

so

$$\rho(a) \leq 1 - l + \varepsilon.$$

If we took a sufficiently large so that $\rho(a) > 1 - \varepsilon$ – which is possible by virtue of condition C – and $\varepsilon < \frac{l}{2}$, there is a contradiction, so $C_n \cdot S(a)$ is not empty and we can state the following theorem:

THEOREM 7. – *If $\mathscr{X}$ is separable and reflexive, a necessary and sufficient condition for a positive definite function $\varphi(x^*)$ with $\varphi(0^*) = 1$ is a characteristic function of a random element in the strict sense, is that $\varphi(x^*)$ satisfies condition C.*

That condition C is sufficient results from the preceding proof, and the necessity is obvious since:

a. If $\mathscr{X}$ is separable and reflexive, $\|x\|$ is measurable, that is to say, that $m[S(a)]$ exists for any a;
b. and that $m[S(a)] \to 1$ if $a \to \infty$ if X is a random variable in the strict sense. Condition C poses two problems.

Problem 1. – We can ask ourselves if every positive definite function $\varphi(x^*)$ with $\varphi(0^*) = 1$ satisfies condition C. We know that it is so if $\mathscr{X}$ is a Euclidean space with any finite dimension, say n. An example will show that this does not remain true in the general case.

Let $\mathscr{X}$ be the Hilbert space, reflexive and separable, of the sequence of real numbers $x = (x_1, ..., x_k, ...)$ such that $\sum_k |x_k|^2 < +\infty$; let x_k^* be the linear functional defined by $x_k^*(x) = x_k$; every linear functional is of the form $x^* = \sum_k a_k x_k^*$.

Set

$$\varphi(x^*) = e^{-\frac{\|x^*\|^2}{2}},$$

we see immediately that

a. $\varphi(0^*) = 1$;
b. $\varphi(x^*) \to 1$ if $x^* \to 0^*$.

Moreover, for any $x^*_{(1)}, ..., x^*_{(h)}$ and any real or complex numbers $\alpha_1, ..., \alpha_h$, we have that

$$\sum_{gj} \varphi[x^*_{(g)} - x^*_{(j)}]\alpha_g \bar{\alpha}_j \qquad \text{is real and non-negative.}$$

Let $X_1, ..., X_k, ...$ be mutually independent Laplacian ordinary random variables with zero mathematical expectation and variance 1; to every

$$x^* = \sum_k a_k x_k^* \qquad \text{with } \sum_k |a_k|^2 < +\infty,$$

we associate the random variable $Y = \sum_k a_k X_k$. This series is convergent almost surely, and also in square-mean since $\sum_k |a_k|^2 < +\infty$. Y is Laplacian, and

$$\|x^*\|^2 = \sum |a_k|^2,$$

so

$$E(e^{iY}) = e^{-\frac{\sum |a_k|^2}{2}} = \varphi(x^*).$$

If $\mathrm{Y}_{(j)}$ is the random variable corresponding to $x_{(j)}$, we see that

$$\varphi[x^*_{(g)} - x^*_{(j)}] = \mathrm{E}[e^{\mathrm{Y}_{(g)} - \mathrm{Y}_{(j)}}],$$

$$\sum_{gj} \varphi[x^*_{(g)} - x^*_{(j)}]\alpha_g \bar{\alpha}_j = \sum_{gj} \mathrm{E}[\alpha_g \bar{\alpha}_j e^{j(\mathrm{Y}_{(g)} - \mathrm{Y}_{(j)})}] = \sum_{gj} \mathrm{E}[\alpha_g e^{i\mathrm{Y}_{(g)}} \overline{\alpha_j e^{i\mathrm{Y}_{(j)}}}]$$

$$= \mathrm{E}\left[\sum_{gj} \alpha_g e^{i\mathrm{Y}_{(g)}} \overline{\alpha_j e^{i\mathrm{Y}_{(j)}}}\right] = \mathrm{E}\left[\left|\sum_{gj} \alpha_j e^{i\mathrm{Y}_{(j)}}\right|^2\right] \geq 0.$$

As a consequence, $\varphi(x^*)$ is positive definite. But $\varphi(x^*)$ does not satisfy condition C. Indeed, with the sequence $[x^*_k]$, let us construct the $\mathscr{A}_k$ of the proof of the preceding theorem; we remark that A_n is defined by

$$\sum_{k=1}^{n} |x_k|^2 \leq a^2 \qquad \text{or} \qquad \sum_{k=1}^{n} |x^*_k(x)|^2 \leq a^2.$$

The above interpretation of $\varphi(x^*)$ with the random variables Y shows that

$$m(\mathscr{A}_n) = \mathrm{Pr}\left[\sum_{k=1}^{n} |\mathrm{X}_k|^2 \leq a^2\right] = \mathrm{P}(a, n).$$

The calculation of $\mathrm{P}(a, n)$ is elementary; it shows that

$$\lim_{n\to\infty} \mathrm{P}(a, n) = 0 \qquad \text{for any } a.$$

Moreover, this is obvious because the series $\sum_k |\mathrm{X}_k|^2$ with positive terms is almost surely divergent; for every definition of $\rho(a)$ that affects all the x^*_k, we have $\rho(a) \leq \lim_{n\to\infty} \mathrm{P}(a, n)$, so $\rho(a) = 0$ for any a.

Remark. – In the above example, let us make the following modification: let us assume that the variance of X_k is $\frac{1}{\sqrt{k}}$ instead of 1 and set

$$\varphi(x^*) = \mathrm{E}[e^{i\mathrm{Y}}], \qquad \text{where } \mathrm{Y} = \sum a_k \mathrm{X}_k.$$

Then Y is Laplacian and

$$\mathrm{E}(\mathrm{Y}) = 0, \qquad \mathrm{E}(\mathrm{Y}^2) = \left[\sum_k \frac{a_k^2}{k}\right],$$

so

$$\varphi(x^*) = e^{-\sum \frac{a_k^2}{2k}}.$$

$\varphi(x^*)$ is continuous with respect to weak convergence in $\mathscr{X}^*$; to show this, it suffices to show that $\rho(x^*) \to 1$ if $x^* \to 0^*$ weakly; if x^* converges weakly to 0^*, $\sum_k a_k^2$ is bounded, say by M, and a_k for a fixed k converges to zero:

$$\sum_k \frac{a_k^2}{k} = \underbrace{\sum_{k=1}^{\mathrm{P}} \frac{a_k^2}{k}}_{\mathrm{A}} + \underbrace{\sum_{k>\mathrm{P}} \frac{a_k^2}{k}}_{\mathrm{B}}.$$

We have $\mathrm{B} \leq \frac{1}{\mathrm{P}}\mathrm{M}$, $\mathrm{A} \to 0$ for any fixed P.

If we took P to be large enough, the exponent of $e^{-\sum \frac{a_k^2}{k}}$ is arbitrarily small.

C.Q.F.D.

But, on the other hand, as the series $\sum_k \frac{1}{k}$ is divergent, $\sum_k |X_k|^2$ is almost surely divergent, so $P(a, n) \to 0$ for any a.

So the fact that positive definite $\varphi(x^*)$ is continuous with respect to weak convergence for x^* does not suffice to imply that $\varphi(x^*)$ satisfies condition C.

Problem 2. – Given these examples, we should investigate simple criteria allowing us to recognise if a positive definite function $\varphi(x^*)$ satisfies condition C or not.

Let $\mathscr{X}$ be a reflexive and separable Banach space, and $\varphi_1(x^*), \varphi_2(x^*), ..., \varphi_n(x^*), ...$ a sequence of characteristic functions such that:

a. There exist positive α and s such that, for any n,

$$E_n[\|X\|^\alpha] = s_n^\alpha < s^\alpha;$$

b. $\lim_{n\to+\infty} \varphi_n(x^*) = \varphi(x^*)$;

c. There exists A such that, for any x^* satisfying $\|x^*\| \leq A$:

$$|\varphi_n(x^*) - \varphi(x^*)| \to 0 \qquad \text{uniformly in } x^*.$$

$\varphi(x^*)$ is obviously positive definite.

Let us consider the law $\mathcal{L}_n$ corresponding to the characteristic function φ_n, then Bienaymé's inequality gives

$$\Pr^{(n)}[\|X\| < a] > 1 - \frac{s_n^\alpha}{a^\alpha} \qquad (a \text{positive}).$$

Using the notations of generalised Bochner's theorem, we have, for any k,

$$\Pr^{(n)}[X \in \mathscr{A}_k] = m_n(\mathscr{A}_k) = \mu_n(A_k) > 1 - \frac{s_n^\alpha}{a^\alpha},$$

$$\lim_{n\to+\infty} m_n(\mathscr{A}_k) = m(\mathscr{A}_k) > 1 - \frac{s^\alpha}{a^\alpha}$$

and, as a consequence,

$$\rho(a) = \lim_{k\to+\infty} m(\mathscr{A}_k) > 1 - \frac{s^\alpha}{a^\alpha},$$

$$\lim_{a\to+\infty} \rho(a) = 1.$$

Condition C is satisfied, so $\varphi(x^*)$ is a characteristic function and defines a law $\mathcal{L}$; from which we have:

THEOREM 8. – *If a sequence of characteristic functions $\varphi_n(x^*)$ converges uniformly in x^*, for every x^* such that $\|x^*\| \leq A$, to a function $\varphi(x^*)$, if, moreover, there exists $\alpha > 0$ such that $E_n[\|X\|^\alpha]$ is uniformly bounded, then $\varphi(x^*)$ is a characteristic function.*

The question now is to know whether $\mathcal{L}$ is the limit of the laws $\mathcal{L}_n$ and first the problem of defining the convergence of a sequence of laws to a limit law.

Example. – Let be a separable Hilbert space and let $x_1, ..., x_n, ...$ be orthonormal vectors, let m_n be the l-measure formed by placing mass 1 on x_n and nowhere else; its characteristic function $\varphi_n(x^*)$ is

$$\varphi_n(x^*) = e^{ix^*(x_n)}.$$

If x^* is fixed and arbitrary and n tends to $+\infty$, $x^*(x_n) \to 0$, so $\varphi_n(x^*) \to 1$ which is the characteristic function of a mass 1 concentrated at 0: we see that m_n does not tend to this distribution although there is the convergence of the characteristic functions.

Chapter IV.

Laplacian Random Elements.

I. – Definition of Laplacian Random Elements.

Definition. – A random element X in the strict sense, with values in a Banach space $\mathscr{X}$ is a Laplacian random element if $x^*(\mathrm{X})$ is a Laplacian random number for any functional $x^* \in \mathscr{X}^*$ [E. Mourier, IV].

This definition of a Laplacian random element has already been proposed by M. Fréchet [M. Fréchet, IV] and compared by this author to a second definition, deduced from the generalisation of a theorem of S. Bernstein, which assumes the existence of $\mathrm{E}[\|\mathrm{X}\|^2]$. A recent result of M.G.Darmois [G.Darmois, I] allows the removal of this hypothesis. It is then possible to show the equivalence of the two definitions for any Banach space $\mathscr{X}$. It is this that we propose to do in what follows. For this, we will need to establish some characteristic properties of the independence of two random elements whose definition was given in Chapter II (page 10).

Theorem 1. – *For two random elements* X_1 *and* X_2 *to be independent, it is necessary and sufficient that for every* $x_1^*, x_2^* \in \mathscr{X}_1^* \times \mathscr{X}_2^*$ *the characteristic function of the pair* $\mathrm{X}_1, \mathrm{X}_2$ *to be the product of the characteristic functions of* X_1 *and* X_2*, that is to say, that*

$$\Phi(x_1^*, x_2^*) = \varphi_1(x_1^*)\varphi_2(x_2^*),$$

where $\varphi_1(x_1^*), \varphi_2(x_2^*)$ *and* $\Phi(x_1^*, x_2^*)$ *are the characteristic functions of* X_1, X_2 *and of the pair* $\mathrm{X}_1, \mathrm{X}_2$ *respectively.*

From the definition of the characteristic function of a random element[(11)],

$$\varphi_1(x_1^*) = \mathrm{E}[e^{ix_1^*(\mathrm{X}_1)}] = \int_{\mathscr{X}_1} e^{ix_1^*(\mathrm{X}_1)} d\mu_1(\epsilon_1),$$

a function of x_1^* defined for every $x_1^* \in \mathscr{X}_1^*$.

Likewise, the characteristic function of X_2 is

$$\varphi_2(x_2^*) = \mathrm{E}[e^{ix_2^*(\mathrm{X}_2)}] = \int_{\mathscr{X}_2} e^{ix_2^*(\mathrm{X}_2)} d\mu_2(\varepsilon_2),$$

a function of x_2^* defined for every $x_2^* \in \mathscr{X}_2^*$.

And the characteristic function of $\mathrm{X}_1, \mathrm{X}_2 \in \mathscr{X}_1 \times \mathscr{X}_2$ is

$$\begin{aligned}\Phi(x_1^*, x_2^*) &= \mathrm{E}e^{i[x_1^*(\mathrm{X}_1)+x_2^*(\mathrm{X}_2)]}\\ &= \int_{\mathscr{X}_1 \times \mathscr{X}_2} e^{i[x_1^*(x_1)+x_2^*(x_2)]} d\lambda(\varepsilon_1 \times \varepsilon_2)\end{aligned}$$

(11) The notations follow those of Chapter II.

defined for all $x_1^*, x_2^* \in (\mathscr{X}_1 \times \mathscr{X}_2)^* \simeq \mathscr{X}_1^* \times \mathscr{X}_2^*$ (Chapter II, page 10). Indeed, every linear functional z^* on $\mathscr{X}_1 \times \mathscr{X}_2$ is of the form

$$z^*(x_1, x_2) = x_1^*(x_1) + x_2^*(x_2); \tag{1}$$

first, (1) is indeed a linear functional on $\mathscr{X}_1 \times \mathscr{X}_2$. It is additive since $x_1^*(x_1)$ and $x_2^*(x_2)$ are additive, and

$$|z^*(x_1, x_2)| \le \|x_1^*\| \cdot \|x_1\| + \|x_2^*\| \cdot \|x_2\|.$$

If $\|x_1, x_2\|$ stays finite, then $\|x_1\|$ and $\|x_2\|$ stay finite, so $|z^*(x_1, x_2)|$ stays finite, so the functional is bounded.

Then, every linear functional on $\mathscr{X}_1 \times \mathscr{X}_2$ is of type (1). Indeed,

$$z^*(x_1, x_2) = z^*(x_1 + 0, 0 + x_2) = z^*(x_1, 0) + z^*(0, x_2).$$

If X_1 and X_2 are independent,

$$\lambda(\varepsilon_1 \times \varepsilon_2) = \mu_1(\varepsilon_1) \times \mu_2(\varepsilon_2)$$

for any $\varepsilon_1 \in \mathscr{F}_1$ and $\varepsilon_2 \in \mathscr{F}_2$, so [P.R.Halmos, I, p.146]

(2)

$$\begin{aligned}\Phi(x_1^*, x_2^*) &= \int\int_{\mathscr{X}_1 \times \mathscr{X}_2} e^{ix_1^*(x_1)} e^{ix_2^*(x_2)} d\mu_1(\varepsilon_1) d\mu_2(\varepsilon_2)\\ &= \int_{\mathscr{X}_1} e^{ix_1^*(x_1)} d\mu_1(\varepsilon_1) \int_{\mathscr{X}_2} e^{ix_2^*(x_2)} d\mu_2(\varepsilon_2),\ \ \Phi(x_1^*, x_2^*) = \varphi_1(x_1^*)\varphi_2(x_2^*).\end{aligned}$$

Conversely, if the characteristic function of X_1, X_2 is of the form (2) for any $x_1^*, x_2^* \in \mathscr{X}_1^* \times \mathscr{X}_2^*$, X_1 and X_2 are independent random elements, indeed the characteristic function determines the measure on the smallest σ-algebra which contains the collection of cylindrical sets.

Let us remark that if a characteristic function $\Phi(x_1^*, x_2^*)$ is the product of a function in x_1^* by a function in x_2^*,

$$\Phi(x_1^*, x_2^*) = f(x_1^*) g(x_2^*),$$

$f(x_1^*)$ and $g(x_2^*)$ are necessarily characteristic functions of X_1 and X_2 respectively; to see this, it suffices to let $x_1^* = 0^*$ or $x_2^* = 0^*$.

Theorem 2. – *For two random elements X_1 and X_2 to be independent, it is necessary and sufficient that, for any $x_1^* \in \mathscr{X}_1^*$ and $x_2^* \in \mathscr{X}_2^*$, $x_1^*(X_1)$ and $x_2^*(X_2)$ to be independent random variables.*

That the condition is necessary is obvious; that it is sufficient follows from the preceding theorem and the theorem on the mathematical expectation of the product of independent random variables.

Indeed,

$$\begin{aligned}\Phi(x_1^*, x_2^*) &= \mathrm{E}e^{i[x_1^*(X_1) + x_2^*(X_2)]}\\ &= \mathrm{E}[e^{ix_1^*(X_1)} e^{ix_2^*(X_2)}]\\ &= \varphi_1(x_1^*)\varphi_2(x_2^*).\end{aligned}$$

THEOREM 3. – *If* X *is a random element in* $\mathscr{X}$ *such that* $E(X) = 0$, *if* X^* *is a random element in* $\mathscr{X}^*$ *and if* X *and* X^* *are independent,*

$$E[X^*(X)] = 0.$$

Indeed,

$$E[X^*(X)] = E[E[X^*(X)/X^*]];$$

and for any fixed x^*,

$$E[X^*(X)/X^* = x^*] = x^*[E(X)] = 0,$$

since, by hypothesis, $E(X) = 0$.

THEOREM 4. – *If* X^* *is a random element in* $\mathscr{X}^*$ *such that* $E(X^*) = 0^*$, *if* X *a random element in* $\mathscr{X}$ *and if* X *and* X^* *are independent,*

$$E[X^*(X)] = 0.$$

Indeed, we know that [E. Hille, I, p.22] that to all $x \in \mathscr{X}$ we can associate $x^{**} \in \mathscr{X}^{**}$ such that for all $x^* \in \mathscr{X}^*$ we have

$$x^*(x) = x^{**}(x^*).$$

We thus have

$$E[X^*(X)] = E[X^{**}(X^*)];$$

but, according to the preceding theorem, if $E(X^*) = 0^*$, we have

$$E[X^{**}(x)] = 0$$

and, as a consequence,

$$E[X^*(X)] = 0.$$

Properties of Laplacian elements. – If $X_1, ..., X_n$ are independent Laplacian random elements taking values in $\mathscr{X}$, x_0 an element in $\mathscr{X}$ and $a_0, a_1, ..., a_n$ some numbers, $Z = a_0x_0 + a_1X_1 + ... + a_nX_n$ is Laplacian.

Indeed,

$$x^*(Z) = a_0x^*(x_0) + a_1x^*(X_1) + ... + a_nx^*(X_n),$$

and according to a classical theorem of calculus of probabilities, $x^*(Z)$ is Laplacian and this holds for all x^*, so Z is a Laplacian element.

Conversely, if $Z = aX + bY$ is Laplacian and if ab is not zero, then if X and Y are independent they are also Laplacian; indeed,

$$x^*(Z) = ax^*(X) + bx^*(Y);$$

$x^*(Z)$ is a Laplacian variable, $x^*(X)$ and $x^*(Y)$ are two independent variables since X and Y are independent, they are thus Laplacian (Lévy-Cramer theorem), so X and Y are Laplacian elements.

Let X be a Laplacian element and x^* an arbitrary element in $\mathscr{X}^*$. Denote by $\sigma^2_{x^*}$ the fluctuation of $x^*(X)$,

$$\sigma^2_{x^*} = E\{x^*(X) - E[x^*(X)]\}^2;$$

$x^*(\mathrm{X})$ is a Laplacian random variable so $\mathrm{E}[x^*(\mathrm{X})]$ and $\sigma^2_{x^*}$ exist and its characteristic function $\varphi_{x^*}(v)$ is

$$\varphi_{x^*}(v) = \mathrm{E}[e^{ivx^*(\mathrm{X})}] = e^{iv\mathrm{E}[x^*(\mathrm{X})]-\frac{1}{2}v^2\sigma^2_{x^*}}, \tag{1}$$

but, denoting by $\varphi(x^*)$ the characteristic function of X,

$$\varphi_{x^*}(v) = \mathrm{E}[e^{ivx^*(\mathrm{X})}] = \phi(vx^*),$$

(1) can be written as

$$\varphi_{x^*}(v) = e^{i\mathrm{E}[vx^*(\mathrm{X})]-\frac{1}{2}\sigma^2_{vx^*}} = \varphi(vx^*)$$

and so for every $x^* \in \mathscr{X}^*$,

$$\boxed{\varphi_{(x^*)} = e^{i\mathrm{E}[x^*(\mathrm{X})]-\frac{1}{2}\sigma^2_{x^*}}.} \tag{2}$$

Conversely, assume that X is a random element whose characteristic function is of the form (2). For any $x^* \in \mathscr{X}^*$, the characteristic function of $x^*(\mathrm{X})$ is

$$\varphi_{x^*}(v) = \varphi(vx^*) = e^{i\mathrm{E}[vx^*(\mathrm{X})]-\frac{1}{2}\sigma^2_{vx^*}} = e^{iv\mathrm{E}[x^*(\mathrm{X})]-\frac{1}{2}v^2\sigma_{x^*}},$$

so $x^*(\mathrm{X})$ is a Laplacian random variable and X is a Laplacian random element.

The theorem of S. Bernstein on which M. Fréchet is based is the following:

For two independent random numbers X *and* Y*, each with a probability density defined everywhere and having the same non-zero and finite fluctuation* σ^2*, to be two Laplacian variables, it is necessary and sufficient that the two variables* $\mathrm{X}+\mathrm{Y}$ *and* $\mathrm{X}-\mathrm{Y}$ *are independent.*

Using a proof that is different to that of S. Bernstein, M. Fréchet extends this proposition to the case where we do not assume the existence of a density, he defines the fluctuation of a random element with values in a complete metric space as being the lower bound of the square of the mathematical expectation of the distance between X and a as a ranges through the space. He shows that if X, taking values in a Banach space $\mathscr{X}$, is such that

1° its fluctuation is finite;

2° there exists in the same space $\mathscr{X}$ another element Y independent of X and such that $\mathrm{X}+\mathrm{Y}$ and $\mathrm{X}-\mathrm{Y}$ are independent, X is Laplacian.

Conversely, M. Fréchet shows that if $\mathscr{X}$ possesses a base and if Y is Laplacian there exists a random element Y independent of X such that $\mathrm{X}+\mathrm{Y}$ and $\mathrm{X}-\mathrm{Y}$ are independent.

Theorem of G. Darmois. – *Two independent random variables* X *and* Y *are necessarily Laplacian if* $\mathrm{X}+\mathrm{Y}$ *and* $\mathrm{X}-\mathrm{Y}$ *are independent.*

Theorem. – *If* $\mathscr{X}$ *is any Banach space, the necessary and sufficient condition for a random element* X *with values in* $\mathscr{X}$ *to be Laplacian is that there exists a random element* Y *with values in* $\mathscr{X}$*, independent of* X *and such that* $\mathrm{X}+\mathrm{Y}$ *and* $\mathrm{X}-\mathrm{Y}$ *are independent.*

The condition is sufficient. Indeed, if X and Y are independent, as well as $\mathrm{X}+\mathrm{Y}$ and $\mathrm{X}-\mathrm{Y}$, $x^*(\mathrm{X})$ and $x^*(\mathrm{Y})$ are likewise independent as well as $x^*(\mathrm{X}+\mathrm{Y})$ and $x^*(\mathrm{X}-\mathrm{Y})$ (c.f. Theorem 2, page 53).

Now,

$$x^*(\mathrm{X}+\mathrm{Y}) = x^*(\mathrm{X}) + x^*(\mathrm{Y}),$$

for any $x^* \in \mathscr{X}^*$ and

$$x^*(\mathrm{X}-\mathrm{Y}) = x^*(\mathrm{X}) - x^*(\mathrm{Y}).$$

So, by G. Darmois's theorem, the random number $x^*(\mathrm{X})$ is Laplacian for any $x^* \in \mathscr{X}^*$, and as a consequence, X is Laplacian.

The condition is necessary. Indeed, let X be a Laplacian random element and Y a random element independent of X with the same law. Let $\varphi(x^*)$ be the characteristic function of X,

$$\varphi(x^*) = e^{i\mathrm{E}[x^*(\mathrm{X})] - \frac{1}{2}\sigma^2_{x^*}} \qquad (x^* \in \mathscr{X}^*).$$

Let $\mathrm{U} = \mathrm{X}+\mathrm{Y}$ and $\mathrm{V} = \mathrm{X}-\mathrm{Y}$. Then U and V are two Laplacian random elements with values in the same space $\mathscr{X}$. Denote their characteristic functions by φ_U and φ_V and by $\varphi_{\mathrm{U},\mathrm{V}}$ that of the pair U,V. As X and Y are independent and of the same law,

$$\begin{aligned}
\varphi_\mathrm{U}(x^*) &= [\varphi(x^*)]^2 = e^{i2\mathrm{E}[x^*(\mathrm{X})] - \sigma^2_{x^*}},\\
\varphi_\mathrm{V}(y^*) &= e^{-\sigma^2_{y^*}},
\end{aligned}$$

so

$$\varphi_\mathrm{U}(X^*) = \varphi_\mathrm{V}(y^*) = e^{i2\mathrm{E}[x^*(\mathrm{X})] - \sigma^2_{x^*} - \sigma^2_{y^*}}.$$

On the other hand,

$$\begin{aligned}
\varphi_{\mathrm{U},\mathrm{V}}(x^*, y^*) &= \mathrm{E}\{e^{i[x^*(\mathrm{U}) + y^*(\mathrm{V})]}\}\\
&= \mathrm{E}\{e^{i[x^*(\mathrm{X}) + x^*(\mathrm{Y}) + y^*(\mathrm{X}) - y^*(\mathrm{Y})]}\}\\
&= \mathrm{E}\{e^{i[(x^*+y^*)(\mathrm{X}) + (x^*-y^*)(\mathrm{Y})]}\}\\
&= \mathrm{E}[e^{i(x^*+y^*)(\mathrm{X})}]\mathrm{E}[e^{i(x^*-y^*)(\mathrm{Y})}]\\
&= \varphi(x^*+y^*)\varphi(x^*-y^*),
\end{aligned}$$

because, as X and Y are independent, $(x^*+y^*)(\mathrm{X})$ and $(x^*-y^*)(\mathrm{Y})$ are also independent for any $x^*, y^* \in \mathscr{X}^*$, so

$$\varphi_{\mathrm{U},\mathrm{V}}(x^*, y^*) = e^{i\mathrm{E}[(x^*+y^*)(\mathrm{X})] - \frac{1}{2}\sigma^2_{x^*+y^*} + i\mathrm{E}[(x^*-y^*)(\mathrm{Y})] - \frac{1}{2}\sigma^2_{x^*+y^*}}.$$

But

$$\begin{aligned}
&\mathrm{E}[(x^*+y^*)(\mathrm{X})] = \mathrm{E}[(x^*(\mathrm{X}) + y^*(\mathrm{X})] = \mathrm{E}[x^*(\mathrm{X})] + \mathrm{E}[y^*(\mathrm{X})],\\
&\mathrm{E}[(x^*-y^*)(\mathrm{Y})] = \mathrm{E}[(x^*-y^*)(\mathrm{X})] = \mathrm{E}[x^*(\mathrm{X})] - \mathrm{E}[y^*(\mathrm{X})],\\
&\sigma^2_{x^*+y^*} = \sigma^2_{x^*} + \sigma^2_{y^*},\\
&\sigma^2_{x^*-y^*} = \sigma^2_{x^*} + \sigma^2_{y^*},
\end{aligned}$$

so

$$\varphi_{\mathrm{U},\mathrm{V}}(x^*, y^*) = e^{i2\mathrm{E}[x^*(\mathrm{X})] - \sigma^2_{x^*} - \sigma^2_{y^*}} = \varphi_\mathrm{U}(x^*)\varphi_\mathrm{V}(y^*).$$

So U and V are independent.

If X has zero fluctuation, X is an almost certain element of $\mathscr{X}$, and conversely, in this case for every $x^* \in \mathscr{X}^*$, $x^*(\mathrm{X})$ is an almost certain number. Conversely, M. Fréchet showed [M. Fréchet, IV] that, if $\mathscr{X}$ possesses a base, $x^*(\mathrm{X})$ can only be almost certainly constant if X is an almost certain element. In the preceding theorem, if X has zero fluctuation, it suffices to take for Y any certain element of $\mathscr{X}$ for $\mathrm{X}+\mathrm{Y}$ and $\mathrm{X}-\mathrm{Y}$ to be independent. Just as we do in the case of random variables, we will think of an almost certain element satisfying a singular Laplacian law.

The preceding study poses a certain number of problems: we saw that if X is a Laplacian random element, $\mathrm{E}[x^*(\mathrm{X})]$ exists for any $x^* \in \mathscr{X}^*$. It is a necessary but not sufficient condition for the existence of E(X), so the first problem is: what can we say about E(X)?

Then, what can we say about $\|\mathrm{X}\|$? And about $\mathrm{E}[\|\mathrm{X}\|^2]$?

We saw that the characteristic function of a Laplacian element X is

$$\varphi(x^*) = e^{i\mathrm{E}[x^*(\mathrm{X})] - \frac{1}{2}\mathrm{E}[x^*(\mathrm{X}) - \mathrm{E}(x^*(\mathrm{X}))]^2};$$

conversely, given a function of the form

$$f(x^*) = e^{i\mathrm{E}[x^*(\mathrm{X})] - \frac{1}{2}\mathrm{E}[x^*(\mathrm{Y}) - \mathrm{E}(x^*(\mathrm{Y}))]^2},$$

where Y is a random element with values in $\mathscr{X}$, does there exist a random element X with values in $\mathscr{X}$ whose characteristic function is $f(x^*)$?

II. – Laplacian Random Elements in a Hilbert Space.

We are going to study these problems in the particular case where $\mathscr{X}$ is a separable Hilbert space $\mathscr{H}$. We know that, then, $\|\mathrm{X}\|$ is measurable (cf. Chapter I, page 6). There exists an orthogonal system $\{x_i\}$ such that every $x \in \mathscr{H}$ is of the form

$$x = \sum_i a_i x_i$$

with the a_i being numbers such that

$$\sum_i |a_i|^2 < +\infty$$

and

$$\|x\|^2 = \sum_i |a_i|^2.$$

Every linear functional $x^*(x)$ is of the form

$$x^*(x) = \sum_i \alpha_i a_i,$$

with the α_i being numbers such that

$$\sum_i |\alpha_i|^2 < +\infty$$

and

$$\|x^*\| = \sum_i |\alpha_i|^2.$$

If $X = \sum_i A_i x_i$ is a Laplacian random element, then $x^*(X)$ is, for every $x^* \in \mathscr{H}^*$, a Laplacian random variable. Consider a functional x_i^* defined by $x_i^*(x) = a_i$.

$x_i^*(X) = A_i$ is a Laplacian random variable, so for any i, $E(A_i)$ and $E[|A_i|^2]$ exist and are finite.

Set

$$m_i = E(A_i), \qquad \sigma_i^2 = \sigma^2(A_i).$$

On the other hand,

$$\sum_i |A_i|^2 = \|X\|^2 = \rho$$

converges almost surely to a random variable in the strict sense.

Let us consider the point P_n with coordinates $(A_i, A_2, ..., A_n)$ (*cf.* figure, page 59), it is a Laplacian random point in a Euclidean space E_n with n dimensions, with orthogonal axes; let M_n be the point in E_n with coordinates $(m_1, m_2, ..., m_n)$. Then M_n is the central point of the distribution in E_n. We have

$$\rho_n^n = \overline{OP}_n^2 = \sum_{i=1}^n |A_i|^2 \leq \rho^2.$$

Let Π be the plane (in E_n) passing through M_n and perpendicular to OM_n, it is the diametrical plane for the ellipsoid of equidensity, so there is a probability $\frac{1}{2}$ for P_n to be beyond Π, so a probability greater than or equal to $\frac{1}{2}$ that $\rho^2 \geq \overline{OM}_n^2$.

If $\overline{OM}_n^2 = \sum_{i=1}^n |m_i|^2$ was not bounded, there would be a probability greater than or equal to $\frac{1}{2}$ that ρ^2 is not bounded, which is impossible. So

$$\sum_i |m_i|^2 < +\infty.$$

Let m be the point in $\mathscr{H}$ with coordinates m_i. Then X and $X - m$ are Laplacian at the same time, and if one has a mathematical expectation, the other also has it (Chapter I, page 4); so it suffices for us to study $X - m$, or else to assume that all the m_i are zero, which we will do in what follows.

Let $x^*(\alpha_1, \alpha_2, ..., \alpha_i, ...)$, $\sum_i |\alpha_i|^2 < +\infty$, be an arbitrary functional in $\mathscr{H}^*$, and let $x_n^*(\alpha_1, ..., \alpha_n, 0, 0, 0, ...)$. Then x_n^* tends strongly to x^* when n tends to $+\infty$, so denoting the characteristic function of X by $\varphi(x^*)$, $\varphi(x_n^*)$ tends, uniformly in x^*, to $\varphi(x^*)$ (*cf.* Chapter III, page 37).

Now, $x_n^*(X)$ is a Laplacian random variable with zero mathematical expectation and $x^*(X)$ is a Laplacian random variable, so

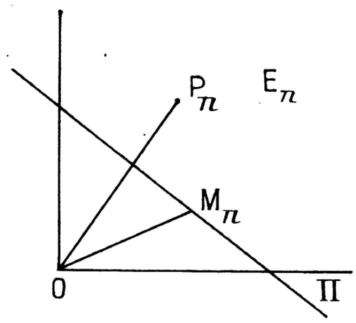
$$E[x^*(X)] = \lim_{n \to +\infty} E[x_n^*(X)] = 0,$$

from which it results that 0 is the mathematical expectation of X, so

Theorem 1. – *If* X, *taking values in a separable Hilbert space, is Laplacian,* E(X) *exists.*

If $X = \sum_i A_i x_i$,

$$E(X) = \sum_i E(A_i) x_i.$$



Preliminary calculations. – Let Z be a Laplacian random variable with zero mathematical expectation and variance σ. Let us calculate the characteristic function φ_{Z^2} of Z^2.

$$\varphi_{Z^2} = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{ivz^2} e^{-\frac{z^2}{2\sigma^2}} dz = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2\sigma^2}(1-2iv\sigma^2)} dz.$$

Set

$$z = \frac{\sigma}{\sqrt{1-2iv\sigma^2}} u,$$

$$\varphi_{Z^2} = \frac{1}{\sqrt{2\pi}\sqrt{1-2iv\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du$$

$$\varphi_{Z^2} = \frac{1}{\sqrt{1-2iv\sigma^2}}.$$

Let $X = \sum_i A_i x_i$ be a Laplacian random element in $\mathscr{H}$ such that $E(A_i) = 0$ for any i, and let P_n be the point in E_n with coordinates $(A_1, ..., A_n)$; the A_i are Laplacian random variables, but not necessarily independent. In E_n we can take new orthogonal axes $(x'_{n,1}, ..., x'_{n,n})$ such that the $A'_{n,1}, ..., A'_{n,n}$ are independent Laplacian random elements with $E(A'_{n,j}) = 0$ for any j.

We have

$$\rho_n^2 = \sum_{j=1}^{n} |A'_{n,j}|^2.$$

Set

$$\lambda_{n,j}^2 = E|A'_{n,j}|^2, \qquad \text{thence} \qquad E(\rho_n^2) = \sum_{j=1}^{n} \lambda_{n,j}^2;$$

we are going to show that the $\lambda_{n,j}$ are bounded.

Let $A'_{n,k}$ be the $A'_{n,j}$ which, for fixed n, has the largest $\lambda_{n,j}$. If $|A'_{n,k}|^2 > C^2$, *a fortiori* we have $\rho_n^2 > C^2$. Assuming that $\lambda_{n,k}$ is not bounded when n tends to $+\infty$,

$$\Pr[A'^2_{n,k} > C^2] = \frac{2}{\sqrt{2\pi}\lambda_{n,k}} \int_{C}^{+\infty} e^{-\frac{u^2}{2\lambda_{n,k}^2}} du \qquad \text{(C positive)}$$

$$= \frac{2}{\sqrt{2\pi}} \int_{\frac{C}{\lambda_{n,k}}}^{+\infty} e^{-\frac{t^2}{2}} dt.$$

Taking $C = \lambda_{n,k}$,

$$\Pr[A'^2_{n,k} > \lambda^2_{n,k}] = \frac{2}{\sqrt{2\pi}} \int_1^{+\infty} e^{-\frac{t^2}{2}} dt,$$

a quantity which is not small, so if $\lambda_{n,k}$ is not bounded, there is not a small probability that $\rho_n^2 \geq C^2$ (with C^2 very large), which is impossible since

$$\rho^2 = \sum_i |A_i|^2 \geq \rho_n^2$$

is almost surely convergent.

So the $\lambda_{n,j}$ are bounded, so the characteristic function of ρ_n^2

$$\varphi_n(\nu) = \prod_{j=1}^{n} (1 - 2i\lambda^2_{n,j}\nu)^{-\frac{1}{2}}$$

is regular in a circle (γ) with centre 0 and a fixed non-zero radius, and as a consequence, it is legitimate to use a limited growth of $\varphi_n(\nu)$ in the neighbourhood of the origin.

Set

$$\begin{aligned}
u_j &= 2i\lambda^2_{n,j} v, \\
\psi_n(v) = \log \varphi_n(v) &= -\frac{1}{2} \sum \log(1 - u_j) \\
&= \frac{1}{2} \left[\sum_1^n u_j + \frac{1}{2} \sum_1^n u_j^2 + \frac{1}{3} \sum_1^n u_j^3 + ... \right], \\
\varphi_n(v) &= 1 + \frac{1}{2} \sum_{j=1}^n u_j + \frac{1}{4} \sum_1^n u_j^2 + \frac{1}{8} \left(\sum_{j=1}^n u_j \right)^2 + ..., \\
\varphi_n(v) &= 1 + i \sum_{j=1}^n \lambda^2_{n,j} v - \left[\sum_{j=1}^n \lambda^4_{n,j} + \frac{1}{2} \left(\sum_{j=1}^n \lambda^2_{n,j} \right)^2 \right] v^2 + (...)v^3.
\end{aligned}$$

But, on the other hand, denoting by M_1 and M_2 the moments of order 1 and 2 of ρ_n^2, we have

$$\varphi_n(v) = 1 + iM_1 v - \frac{1}{2} M_2 v^2 + (...)v^3$$

and as a consequence,

$$\begin{aligned}
M_1 &= \sum_{j=1}^n \lambda^2_{n,j}, \\
M_2 &= 2 \sum_{j=1}^n \lambda^4_{n,j} + \left(\sum_{j=1}^n \lambda^2_{n,j} \right)^2,
\end{aligned}$$

from which we get

$$\sigma^2(\rho_n^2) = \mathrm{M}_2 - \mathrm{M}_1^2 = 2\sum_{j=1}^{n} \lambda_{n,j}^4.$$

Imagine that the $\lambda_{n,j}$ are all equal to 1, we then have

$$\mathrm{E}[\rho_n^2] = n, \qquad \sigma^2(\rho_n^2) = 2n.$$

Bienaymé's inequality gives, for any β,

$$\Pr[|\rho_n^2 - \mathrm{E}(\rho_n^2)| < \beta_n] \geq 1 - \frac{2n}{\beta^2 n^2}$$

and *a fortiori*,

$$\Pr[\rho_n^2 > n - \beta n] > 1 - \frac{2n}{\beta^2 n^2};$$

$1 - \frac{2n}{\beta^2 n^2}$ tends to 1 when n tends to $+\infty$ so, taking β to be smaller than 1, the probability that ρ_n^2 is large is not small; this result is *a fortiori* exact if the $\lambda_{n,j}$ are all at least equal to 1, without being equal between them, and this proves that the $\lambda_{n,j}$ cannot all be larger than or equal to 1. Let q_n be the number of the $\lambda_{n,j}$ $(j = 1, 2, ..., n)$ which are larger than or equal to 1, and assume that q_n is not bounded when n tends to $+\infty$. In

$$\rho_n^2 = \sum_{i=1}^{n} |\mathrm{A}'_{n,j}|^2$$

let us neglect all the terms for which the corresponding $\lambda_{n,j}$ is smaller than 1, which is to say that we consider a $\rho_n'^2$ for which all the $\lambda_{n,j}$ are larger than or equal to 1; according to above, for any C^2, as large as we want it to be, it suffices that n is large enough for the probability of ρ_n^2 being greater than C^2 not to be small, which is impossible since ρ_n^2 is almost surely convergent. So q_n is *bounded*.

Let us assume that $\mathrm{E}(\rho_n^2) = \sum \lambda_{n,j}^2$ tends to $+\infty$ when n tends to $+\infty$.

Let β be smaller than 1. Then Bienaymé's inequality gives

$$\Pr[\rho_n^2 > \mathrm{E}(\rho_n^2)(1-\beta)] = \Pr[|\rho_n^2 - \mathrm{E}(\rho_n^2)| < \beta \mathrm{E}(\rho_n^2)] \geq 1 - \frac{2\sum_{j=1}^{n} \lambda_{n,j}^4}{\beta^2 \left(\sum_{j=1}^{n} \lambda_{n,j}^2\right)^2}.$$

Let us study the quantity

$$\mathrm{Q}_n = \frac{2\sum_{j=1}^{n} \lambda_{n,j}^4}{\beta^2 \left(\sum_{j=1}^{n} \lambda_{n,j}^2\right)^2} = \frac{2}{\beta^2} \frac{\sum_{j=1}^{n} \lambda_{n,j}^4}{\sum_{j=1}^{n} \lambda_{n,j}^2} \frac{1}{\sum_{j=1}^{n} \lambda_{n,j}^2}.$$

We showed that when n tends to $+\infty$, the $\lambda_{n,j}$ are bounded and that the number of $\lambda_{n,j}$ larger than or equal to 1 is finite, so

$$\lim_{n\to+\infty} \frac{\sum_{k_i=k_1}^{k} \lambda_{n,k_i}^4}{\sum_{j=1}^{n} \lambda_{n,j}^2} \to 0.$$

For $j \neq k_i$,

$$\lambda_{n,j} < 1, \qquad \text{so } \lambda_{n,j}^4 < \lambda_{n,j}^2,$$

so

$$\frac{\sum_{j\neq k_i} \lambda_{n,j}^4}{\sum_{j=1}^n \lambda_{n,j}^2} < 1.$$

So

$$\lim_{n\to+\infty} \mathrm{Q}_n = 0,$$

which proves that the probability that $\rho_n^2 > \mathrm{E}(\rho_n^2)(1-\beta)$ is not small; now, by hypothesis,, $\mathrm{E}(\rho_n^2)$ tends to $+\infty$, so the probability that ρ_n^2 is larger than C^2 for any C^2 is not small, which is impossible, so $\mathrm{E}(\rho_n^2)$ is *bounded*.

$\rho^2 = \lim_{n\to+\infty} \rho_n^2$, so, by Fatou's Lemma, $\mathrm{E}(\rho^2)$ exists and

$$\mathrm{E}(\rho^2) \leq \lim_{n\to+\infty} \mathrm{E}(\rho_n^2).$$

THEOREM 2. – *If* X *with values in a separable Hilbert space is Laplacian,*

$$\mathrm{E}(\|\mathrm{X}\|^2) < +\infty.$$

We used the well-known property that if P_n is a Laplacian random point in a Euclidean space E_n with n dimensions, there exists an orthogonal system of axes in E_n such that the coordinates of P_n are independent Laplacian variables. Does this property remain true in the case of a Laplacian random point P in a separable Hilbert space $\mathscr{H}$? Put otherwise, does there exist an orthogonal system $\{x_j'\}$ in $\mathscr{H}$ such that if X is a Laplacian random element with values in $\mathscr{H}$, $\mathrm{X} = \sum_j \mathrm{A}_j' x_j'$, with A_j' being independent Laplacian random variables?

With the orthogonal system $\{x_j\}$, let $\mathrm{X} = \sum_j \mathrm{A}_j x_j$. Then every linear functional is of the form

$$x^*(\mathrm{X}) = \sum_j \alpha_j \mathrm{A}_j,$$

with the α_j being numbers such that $\sum|\alpha_j|^2 < +\infty$,

$$\mathrm{E}\{[x^*(\mathrm{X})]^2\} = \mathrm{E}\left[\sum_{jk} \alpha_j\alpha_k \mathrm{A}_j\mathrm{A}_k\right] = \sum_{jk} r_{jk}\alpha_j\alpha_k = \Phi$$

by setting

$$\mathrm{E}(\mathrm{A}_j\mathrm{A}_k) = r_{jk};$$

Φ is a quadratic Hermitian form.

Let us consider the reduction

$$\Phi_n = \sum_{j,k\leq n} r_{jk}\alpha_j\alpha_k;$$

$$\Phi - \Phi_n = \sum_{j,k,\text{not both } \leq n} r_{jk}\alpha_j\alpha_k = \sum_{j=1}^{n}\left[\sum_{k>n} r_{jk}\alpha_j\alpha_k\right] + \sum_{i>n}\sum_k r_{jk}\alpha_j\alpha_k.$$

Let us assume that $\sum|\alpha_j|^2 \leq 1$,

$$\left|\sum_j \alpha_j \mathrm{A}_j\right| = |x^*(\mathrm{X})| \leq \|x^*\|\cdot\|\mathrm{X}\| \leq \|\mathrm{X}\|,$$

$$\sum_{j=1}^{n} \sum_{k>n} r_{jk} \alpha_j \alpha_k = \mathrm{E}\left[\sum_{j=1}^{n} \sum_{k>n} \alpha_j \alpha_k \mathrm{A}_j \mathrm{A}_k\right].$$

Now,

$$\sum_{j=1}^{n} \sum_{k>n} \alpha_j \alpha_k \mathrm{A}_j \mathrm{A}_k = \sum_{j=1}^{n} \alpha_j \mathrm{A}_j \sum_{k>n} \alpha_k \mathrm{A}_k,$$

$$\left|\sum_{j=1}^{n} \alpha_j \mathrm{A}_j\right| \leq 1\|\mathrm{X}\|,$$

$$\left|\sum_{k>n} \alpha_k \mathrm{A}_k\right| \leq \varepsilon_n \left[\sum_{k>n} |\mathrm{A}_k|^2\right]^{\frac{1}{2}},$$

where ε_n is a number bounded by 1.

We likewise have

$$\left|\sum_{j>n} \sum_{k} \alpha_j \alpha_k \mathrm{A}_j \mathrm{A}_k\right| \leq \|\mathrm{X}\| \varepsilon_n \left[\sum_{j>n} |\mathrm{A}_j|^2\right]^{\frac{1}{2}}$$

and since $\sum_k |\mathrm{A}_k|^2 = \|\mathrm{X}\|^2$ is almost surely convergent, $\left[\sum_{k>n} |\mathrm{A}_k|^2\right]^{\frac{1}{2}}$ tends to zero almost surely when n tends to $+\infty$.

As a consequence, Φ_n tends to Φ uniformly in α_j, provided that $\sum |\alpha_j|^2 \leq 1$, and it results that [F. Riesz, I, p.113] that Φ is completely continuous and so [F. Riesz, I, p.146] admits a decomposition of the form

$$\Phi = \sum_{j=1}^{\infty} s_j \left(\sum_{k=1}^{\infty} l_{jk} \alpha_k\right)^2,$$

the linear forms being normalised and pairwise orthogonal, that is to say, there exist $s_1, s_2, ..., s_j, ...$ and points $x'_1, ..., x'_j, ...$, pairwise orthogonal with $\|x'_j\| = 1$, such that

$$\Phi = \sum_j s_j [x^*(x'_j)]^2;$$

if we take the x'_j as the new axes, we have

$$\Phi = \sum_j s_j \alpha_j^2.$$

Now,

$$\Phi = \mathrm{E}\{[x^*(\mathrm{X})]^2\} = \sum_{jk} [\alpha_j \alpha_k \mathrm{E}(\mathrm{A}'_j \mathrm{A}'_k)],$$

so

$$\mathrm{E}(\mathrm{A}'_j \mathrm{A}'_k) = 0, \qquad \text{if } j \neq k$$
$$\mathrm{E}(\mathrm{A}'^2_j) = s_j,$$

that is to say, that A'_j and A'_k $(j \neq k)$ are non-correlated random variables, which means, as A'_j are Laplacian random variables, that they are independent, and that one of them A'_k is independent of $\mathrm{A}'_{j_1}, ..., \mathrm{A}'_{j_n}$ for any n, $j_1, ..., j_n$ and $k \neq j_i$ $(i = 1, ..., n)$, so

THEOREM 3. – *If* X *is a random element with values in a separable Hilbert space* $\mathscr{H}$*, there exists in* $\mathscr{H}$ *an orthogonal system* $\{x'_j\}$ *such that*

$$X = \sum_j x'_j A'_j,$$

with the A'_j *being independent Laplacian random variables.*

Finally, let us see whether every function of the form

$$f(x^*) = e^{i\mathrm{E}[x^*(\mathrm{Y})] - \frac{1}{2}\mathrm{E}\{[x^*(\mathrm{Y}) - \mathrm{E}(x^*(\mathrm{Y}))]^2\}},$$

where Y is a random element in $\mathscr{H}$ is the characteristic function of a random element with values in $\mathscr{H}$.

We will assume that Y is such that $\mathrm{E}(\|\mathrm{Y}^2) = s^2$, which implies that $\mathrm{E}(\|\mathrm{Y}\|)$ and hence, since $\mathscr{H}$ is separable, that E(Y) exists. We can then, without loss of generality, assume that $\mathrm{E}(\mathrm{Y}) = 0$; this then reduces to

$$f(x^*) = e^{-\frac{1}{2}\mathrm{E}([x^*(\mathrm{Y})]^2)}.$$

Let us consider n independent random elements $\mathrm{Y}_1, \mathrm{Y}_2, ..., \mathrm{Y}_n$ with the same law as Y and let

$$\mathrm{Z}_n = \frac{1}{\sqrt{n}}(\mathrm{Y}_1 + \mathrm{Y}_2 + ... + \mathrm{Y}_n).$$

In $\mathscr{H}$, the square of the norm, $\|\mathrm{Z}_n\|^2$, is equal to the scalar product $(\mathrm{Z}_n, \bar{\mathrm{Z}}_n)$, so

$$\mathrm{E}[\|\mathrm{Z}_n\|^2] = \frac{1}{n}\sum_{ij} \mathrm{E}[(\mathrm{Y}_i, \bar{\mathrm{Y}}_j)];$$

if $i = j$,

$$\mathrm{E}[(\mathrm{Y}_i, \bar{\mathrm{Y}}_j)] = \mathrm{E}[\|\mathrm{Y}_i\|^2] = s^2;$$

if $i \neq j$, $(\mathrm{Y}_i, \bar{\mathrm{Y}}_j)$ is a linear functional of Y_i and $\mathrm{E}(\mathrm{Y}_i) = 0$, so by Theorem 3 of Paragraph 1 of this Chapter,

$$\mathrm{E}[(\mathrm{Y}_i, \bar{\mathrm{Y}}_j)] = 0$$

and, as a consequence,

$$\mathrm{E}[\|\mathrm{Z}_n\|^2] = \frac{ns^2}{n} = s^2.$$

Let $\varphi(x^*)$ be the characteristic function of Y. Then that of Z_n is

$$\Phi_n(x^*) = \left[\varphi\left(\frac{x^*}{\sqrt{n}}\right)\right]^n = \left\{1 - \frac{1}{2n}\mathrm{E}[|x^*(\mathrm{Y})|^2] + \frac{1}{n}\|x^*\|^2\omega\left(\frac{x^*}{\sqrt{n}}\right)\right\}^n,$$

where $\omega\left(\frac{x^*}{\sqrt{n}}\right) \to 0$ when $\left\|\frac{x^*}{\sqrt{n}}\right\| \to 0$ (Theorem 4, Chapter III). Then

$$\log \Phi_n(x^*) = -\frac{1}{2}\mathrm{E}[|x^*(\mathrm{Y})|^2] + \|x^*\|^2\omega\left(\frac{x^*}{\sqrt{n}}\right).$$

So if $\|x^*\| < \mathrm{A}$, for any $\mathrm{A} > 0$, $\Phi_n(x^*)$ converges uniformly to $f(x^*)$ and, as a consequence, (Theorem 8, Chapter III), $f(x^*)$ is a characteristic function; it is thus (Section 1, Chapter IV) the characteristic function of a Laplacian element X, so

Theorem 4. – *Every function*

$$f(x^*) = e^{i\mathrm{E}[x^*(\mathrm{Y})] - \frac{1}{2}\mathrm{E}\{[x^*(\mathrm{Y}) - \mathrm{E}(x^*(\mathrm{Y}))]^2\}},$$

where Y *is a random element with values in a separable Hilbert space, such that* $\mathrm{E}(\|\mathrm{Y}\|^2) = s^2$, *is the characteristic function of a Laplacian random element.*

The proof of the above theorem equally shows us that:

Theorem 5. – *If* $\mathrm{Y}_1, \mathrm{Y}_2, ..., \mathrm{Y}_n$ *are independent random elements with the same law, with values in a separable Hilbert space, and if*

$$\mathrm{E}(\|\mathrm{Y}_i\|^2) = s^2, \qquad \frac{1}{\sqrt{n}}(\mathrm{Y}_1 + ... + \mathrm{Y}_n)$$

converges in the sense of Bernoulli, when $n \to +\infty$, *to a Laplacian random element.*

Bibliography

Banach:

I. *Théorie des opération linéaires.*

Bochner:

I. *Integration von Funktionen deren Werte die Elemente eines Vektorraumes sind* (Fund. Math., t.20, 1933)

II. *Vorlesung über Fouriersche Integrale (Akad. Verlagsges.,* Leipzig, 1939)

Darmois (G.):

I. *C.R.Acad. Sc.,* t. 232, 1951, p.1999.

Dieudonné (J.):

I. *Annales Univ. Grenoble, (2),* t.23, 1947-1948.

Doss (S.):

I. *Sur la moyenne d'un élément aléatoire dans un espace distancié. (Bull. Sc. Math.,* 2e série, t. 73, mars-avril 1949).

Fortet (R.) and Mourier, (E.):

I. *Loi des grands nombres et théorie ergodique (C.R.Acad.Sc.,* t. 232, 1951, p.923).

Fréchet (M.):

I. *Les éléments aléatoires de nature quelconque dans un espace distancié (Annales de l'Institut H. Poincaré,* X, fasc. IV).

II. *Intégrale abstraite d'une fonction abstraite d'une variable abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque. (Revue Scientifique,* fasc. 8)

III. *Les éléments aléatoires de nature quelconque (Annales de l'Institut H. Poincaré,* t. XIV, 1948).

IV. *Généralisation de la loi de probabilité de Laplace (Annales de l'Institut H. Poincaré,* t. XII, fasc. I, 1951).

Halmos (P.-R.):

I. *Measure Theory (D. van Nostrand Company, Inc.,* New-York).

Hille (E.):

I. *Functional Analysis and semi-groups (Amer. Math. Soc.,* édit.).

Khintchine:

I. *Zur Birkhofflosung der Ergodein-problemes (Math. Ann.,* t. 107, 1932).

Kolmogoroff:

I. *Grundbegriffe der Wahrscheinlichkeitsrechnung.*

II. *Ein vereinfachter Beweis des Birkhoffschen Ergodensatzes (Rec. Math. de Moscou*, t. 44, 1937).

Le Cam (L.):

I. *Un instrument d'étude des fonctions aléatoires: la fonctionnelle caractèristique (C.R. Acad. Sc.,* t.224, 1947, p.710).

Mourier (E.):

I. *Sur l'espérance mathématique d'un élément aléatoire dans un espace de Banach (C.R. Acad. Sc.,* t. 229, 1949, p.1300).

II. *Lois des grands nombres et théorie ergodique (C.R. Acad. sc.,* t. 232, 1951, p.923).

III. *Propriétés des caractèristiques d'un élément aléatoire dans un espace de Banach (C.R. Acad. Sc.,* t.231, 1950, p.28).

IV. *Éléments aléatoires laplaciens dans un espace de Banach (C.R. Acad. Sc.,* t.236, 1953, p.575)

Pettis:

I. *On integration in vector spaces (Trans. Amer. Math. Soc.,* vol.44, 1938).

Riesz (F.):

I. *Systèmes d'équations linéaires à une infinité d'inconnues,* Gauthier-Villars, Paris.

Yosida and Kakutani:

I. *Operator theoretical treatment of Markoff process and mean ergodic theorem (Annals of Math.,* vol. 42, no. 1, 1941).