UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Series Aleatorias en Espacios de Funciones y Algunas Aplicaciones

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área de Ciencias Matemáticas.

Autor: Lic. Juan Miguel Medina

Director: Dr. Bruno Cernuschi Frías

Consejero de Estudios: Dr. Carlos Alberto Cabrelli

Lugar de Trabajo:
(I.A.M.) Instituto Argentino de Matemática
"Alberto P. Calderón"-CONICET

Marzo 2011

Series Aleatorias en Espacios de Funciones y Algunas Aplicaciones

Resumen. El objeto de este trabajo es el estudio de ciertas series aleatorias $\sum_i X_i$, con X_i variables aleatorias que toman valores en un espacio de funciones apropiado. Se le dará particular importancia al caso en que $X_i = a_i f_i$, donde $\{f_i\}_i$ es un conjunto de funciones fijas, por ejemplo una base de algún espacio apropiado, y los coeficientes a_i's son variables aleatorias. Este tipo de resultados está relacionado con la posible representación de procesos estocásticos mediante series. Por ejemplo, si los a_i's son ciertas variables aleatorias independientes y $\{f_i\}_i$ es un conjunto apropiado de funciones en $L^2[0,1]$, Itô de esta manera dió una construcción del proceso Browniano sobre el intervalo $[0,1]$ [39]. Se estudiarán los casos de series aleatorias con valores en espacios L^p separables y también se estudiará el caso de series convergentes en el espacio de distribuciones $\mathcal{D}'(\mathbb{R}^d)$. En el caso de los espacios L^p separable, se estudiarán algunas relaciones entre los distintos tipos de convergencia, casi segura con respecto a la norma del espacio subyacente que estamos considerando, convergencia en media y en casi todo punto respecto al espacio producto, que surge de considerar a la variable aleatoria que toma valores en L^p como una función de dos variables. La elección de estos espacios está motivada por algunas aplicaciones. Si lo deseado es utilizar este tipo de desarrollos para construir un proceso estocástico, puede ser que para algunos casos "patológicos", sea mas conveniente considerar por ejemplo series convergentes en $\mathcal{D}'(\mathbb{R}^d)$. Por ejemplo esto, finalmente, nos permitirá dar un desarrollo en serie para la familia de procesos $\frac{1}{f}$, que en los últimos años han recibido cierto interés en las aplicaciones. De alguna manera estas representaciones tienen una similitud con el clásico teorema de Karhunen-Loève [27]. Una propiedad del desarrollo de Karhunen-Loève es que se obtiene una base ortonormal del espacio lineal generado por el proceso. Esto permite escribir ciertas aproximaciones en forma de series incondicionalmente convergentes. Esta útil propiedad se puede obtener bajo otras condiciones. Para resolver éste problema, al final, estudiaremos condiciones para las cuales una sucesión estacionaria forma un frame o una base de Riesz.

Keywords: Series aleatorias de funciones, procesos estocásticos, convergencia.

CONICET

Random Series in Function Spaces and some Applications

Abstract. In this thesis we study certain random series of the form $\sum_i X_i$, where the X_i's are random variables taking values in an appropriate function space. We will give particular importance to the case when $X_i = a_i f_i$, where $\{f_i\}_i$ is an appropriate set of functions,for example a basis of some function space, and the coefficients a_i's are random variables. This type of result is related to the possible representation of random processes by series. For example, if the a_i's are suitable independent random variables and $\{f_i\}_i$ is an appropriate set of functions in $L^2[0,1]$, Itô in this way, gave a series representation of the Brownian process on the interval $[0,1]$ [39]. We will study the cases of random series taking values in separable L^p spaces, we will also study random series in $\mathcal{D}'(\mathbb{R}^d)$. In the case of the separable L^p spaces, we will study several relationships between different types of convergence: almost sure with respect to the norm of the underlying function space, convergence in the mean and convergence in the product space, as a consequence of considering L^p valued random variables as two variable functions. The election of these particular spaces was motivated by some applications. If we want to use these type of series expansion to construct random processes, for some "pathological" cases it could be more appropriate to consider convergent series in $\mathcal{D}'(\mathbb{R}^d)$. For example, this allows us to give a series representation of the $\frac{1}{f}$ family of stochastic processes, which in recent time has received special interest from the applications. In some way, this representations resemble the classic Karhunen-Loève theorem [27]. A property of the Karhunen-Loève expansion of a random process is that one obtains an orthonormal basis of the closed linear span of the whole process. This allows to write certain approximations as unconditional convergent series. This useful property could be obtained under other conditions. To solve this problem, finally, we study conditions under which a stationary sequence forms a frame or a Riesz basis of its closed linear span.

Keywords: Random series of functions, stochastic processes, convergence.

Agradecimientos y dedicatoria.

Este trabajo está dedicado a Juana Elsa Guerra Arce, Juan Carlos Medina y a mis abuelos que son los mas grandes contribuyentes en éste esfuerzo. A ellos no sólo agradezco su ayuda material, sinó su actitud y énfasis en la promoción de la idea de que el conocimiento es el mayor capital y fuente de progreso para el individuo y la sociedad. A la memoria de Eduardo y de Alfonso.

En primer lugar quiero agradecer a mi director Dr. Bruno Cernuschi Frías, por tener confianza en mí, su dedicación y paciencia, lo que no es poco. Bajo su dirección, no sólo fué posible realizar ésta tésis sinó que también me ví enriquecido en mi formación como persona. También agradezco al Dr. Carlos Cabrelli, mi consejero de estudios por ser un "buen consejero" y por hacerme participar de su "club" de analistas armónicos del DM de la FCEN-UBA, lo cual contribuyó a éste trabajo. Al Dr. Gustavo Corach, director del IAM-CONICET, agradezco que haya permitido que el instituto "A. Calderón" sea mi lugar de trabajo. Con éste trabajo espero saldar minimamente mi deuda por dejarme ser parte de una institución que lleva el nombre de alguien que dejó su huella en la matemática, y en particular, en la ciencia argentina.
Debo agradecer a Oscar H. Bustos, Enrique M. Cabaña, Liliana Forzani, José León y a Pablo Groisman haber aceptado ser jurados de mi tesis.

En los cimientos de éste trabajo hay muchas ideas relacionadas con, o respuestas parciales a algunos problemas que me llamaron la atención cuando pasé por la Facultad de Ingeniería de la Universidad de Buenos Aires, principalmente ligadas a problemas relacionados con el procesamiento de señales.
Bruno Cernuschi Frías, en la FI-UBA, de alguna manera supo como hacer que saque provecho de alguna de éstas inquietudes, favoreció el desarrollo de un espiritu crítico y la iniciativa a la hora de tratar diversos temas de investigación. Fué el primero me presentó algunos problemas relacionados con la representación de procesos $\frac{1}{f}$ y relacionados de memoria larga, y sus aplicaciones en procesamiento de señales; de alguna manera ésto sirvió como condición inicial a mis comienzos en investigación. Este tema, que en parte se toca en éste trabajo, también fué desarrollado en el contexto de un convenio entre UBATec y Repsol-YPF, por eso, debo agradecer a Mario Grinberg del área de exploración de YPF, que es susceptible de apoyar éste tipo de iniciativas.
Quiero agradecer también a: en la FCEN-UBA, nuevamente a Carlos Cabrelli, la naturaleza de éste trabajo hace necesaria la utilización de varias herramientas de Análisis Armónico y ésto se vió favorecido por mantener éste contacto. A todo el resto de los profesores como Mariela Sued y una lista interminable de personas que de alguna manera contribuyeron a mi formación. Mas personalmente, a amigos del "ambiente", alumnos, y compañeros de estudio y de trabajo, como Mariana Perez, Marina Fragalá y Alexia Yavícoli, que con sus, a veces, preguntas molestas, ideas, comentarios y sufridos oídos aportaron a mi formación. En ésta misma categoría de "cobayos matemáticos", que se prestaron a sufrirme con distintas dósis en su propio cuerpo, en la FI-UBA debo agradecer a Tomás Crivelli, Agustín Mailing, Trini Gonzales

Chaves y a Sebastián Facello entre otros. Entre los profesores, nuevamente, a Bruno Cernuschi por haber sabido dar vida a ésta 'criatura'. No me puedo olvidar en mis comienzos, de Pedro Vardanega y Rita Gamerman, si no hubiera aprendido a integrar bien (*in the Riemann sense*) no podría ni haber pensado en embarcarme en un proyecto como éste. Luego, el destino llevó a que compartieramos tareas docentes. Cuando volví a ésta pequeña familia de Análisis II, también se había agregado 'el cordobés' -al decir de los alumnos- Daniel Gonzalez. En el dpto. de matemática de la FI-UBA, también debo recordar al 'Dr. S' Sebastián Grynberg y al Dr. José Luis Mancilla Aguilar, con todos ellos, además tuve la oportunidad de compartir mas tarde tareas docentes que de alguna manera contribuyeron a mi formación. En el dpto. de electrónica, al Dr. José Luis Hamkalo, con quien compartí la oficina, le agradezco sus clases magistrales de "Bonvivantismo", las cuales indudablemente han contribuido a aumentar la producción científica . *Last but not least*, al exiliado en el Laboratorio de Física Teórica de la Univ. de París, Daniel 'Elvis' López; las discusiones con Daniel, que con el tiempo trascendieron la UBA y se volvieron intercontinentales, han tenido siempre una influencia que me es díficil de cuantificar, pues datan de una época "prehistórica" en el Colegio. Agradezco algunas ayudas con los detalles, no por eso menos importantes, de terminación: con el manejo del LaTex a Agustín y sugerencias varias sobre la presentación a Lucrecia Cúneo Libarona.

Para terminar, destaco la labor administrativa de la Dra. Alicia Dickenstein y de Mónica Lucas en la subcomisión de doctorado, que hicieron todo lo posible para agilizar el viaje de la tésis a través del oscuro valle de la burocracia.

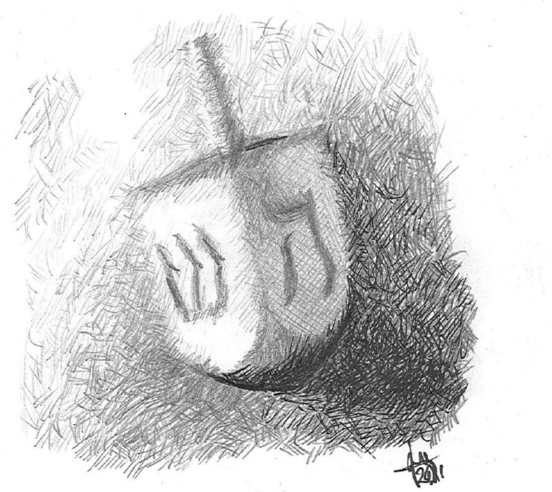

Figure 1: 'Dreidel'-pirinola.

" L'essentiel est invisible pour les yeux".

Antoine de Saint-Exupéry. *Le Petit Prince*, 1943.

Contents

Chapter 1

Prelude-Introduction

"Cuando se tiene algo que decir, se escribe en cualquier parte. Sobre una bobina de papel o en un cuarto infernal. Dios o el diablo están junto a uno dictándole inefables palabras (...)".

Roberto Arlt. *Los Lanzallamas*, 1932.

In order to give to the reader an impression of the problems which are going to be treated here, we are going to describe at a very informal level how these problems were chosen. Roughly speaking, it is possible to say that this work deals with the following problem: *Let $\{f_n(t)\}_{n\in\mathbb{N}}$ be a set of functions and let $\{Y_n\}_{n\in\mathbb{N}}$ be a sequence of random variables. We would like to find conditions under which*

$$X(t) = \sum_{n\in\mathbb{N}} Y_n f_n(t)$$

converges, in some sense to be specified.

To tackle this problem we can begin to consider the whole process, $\{X_t\}_t$ as a unique random variable taking values in an appropriate function space. The consideration of a random process as a random element (or a random variable taking values in some function space) by Doob, Phrohorov, Billingsley, Paley, Zygmund and Wiener and other has inspired the study of stochastic convergence properties for random elements. However, as we will see later a careful construction of the appropriate framework is needed in these considerations. One of the problems arising, is measurability. For example, as we will see, is easy to construct an example of a non measurable mapping from a probability space to $\mathbb{R}^T$, where T is a non countable parameter space. One way of solving this problem is by placing constrains on the parameter space. For example, a random process with a countable parameter space can be shown that it is a random element (i.e. a measurable mapping) in the space of sequences. Often the stochastic process will take values only in a "small" subspace of $\mathbb{R}^T$. Recall that a separable stochastic process may have sample paths which are Borel measurable functions from T into $\mathbb{R}$ (Loève 1963) and hence are restricted a.s. to a subspace of $\mathbb{R}^T$. Thus, the random processes may have properties that reduce the ranges of the mappings from Ω to interesting subspaces of $\mathbb{R}^T$ where different topological structures can be employed. In this thesis we will be concerned with the convergence problem of sums of certain classes of these random elements. The function spaces employed in these approaches are strongly influenced by the particular application which has inspired the problem. In the case of this work, it was influenced in some way by several results used in the applications, specially in engineering. The following are classic illustrative examples:

Theorem 1.0.1. *(Stochastic version of the Shannon- Kotelnikov sampling theorem.) Let $\{X_t\}_{t\in\mathbb{R}}$ be a wide sense stationary random process, with spectral measure supported on $[-B, B]$* [a] *Then $\{X_t\}_{t\in\mathbb{R}}$ admits the following series expansion*

$$X_t = \sum_{n\in\mathbb{Z}} X_{t_n} f_n(t)\,. \tag{1.0.1}$$

With $f_n(t) = \frac{sin(Bt-\pi n)}{Bt-\pi n}$ and $t_n = \frac{\pi n}{B}$, $n \in \mathbb{Z}$, and the convergence is in the mean square sense.

Theorem 1.0.2. *[27](Karhunen -Loève) Let $\{X_t\}_{t\in[a,b]}$ be a measurable process, of finite variance and mean square continuous, then $\{X_t\}_{t\in[a,b]}$ admits a series expansion*

$$X_t = \sum_{k\in\mathbb{N}} \chi_k f_k(t)$$

where the convergence is in the mean square sense. In this expansion the random variables χ_k are orthogonal and $\mathbb{E}|\chi_k|^2 = \lambda_k$, where λ_k and f_k are the eigenvalues and eigenfunctions, respectively, of the covariance operator.

The first one is a corner stone in communication theory. It allows the analogic-digital convertion of signals. The second one is also known in other contexts. These results admit several generalizations and variants, other type of convergence may be proven under other conditions. For example, an interesting generalization of theorem 1.0.1 is given in [45], for non stationary processes. There, some similar tools to those which we are going to use in chapter 5 are introduced. The idea in some way is to use the theory of generalized functions and certain Sobolev spaces as auxiliary tools to deal with some processes which have an spectral behaviour (i.e. in terms of the Fourier transform, in some sense, of a certain magnitude related to the problem which we are modelling) that falls out of the ordinary theory of stationary processes. In our case, in **Chapter 5** treating a different problem, we are going to give a construction which allows to give a series expansion representation of certain processes, such as the self-similar $\frac{1}{f}$ family of random processes. This type of process was first proposed by Kolmogorov in the context of turbulence [44]. In recent times, Karhunen-Loève like expansions for such processes have received special attention in many applications [82] [1]. There has been several attempts to represent $\frac{1}{f}$ and related processes in terms of a Karhunen- Loève like expansion, especially in the one dimensional case and using wavelet basis [82] [22] [36] [53] [60]. Here we shall prove that is possible to represent a d-dimensional $\frac{1}{f}$ processes by a series expansion using an arbitrary orthonormal basis. The proposed construction will converge with probability one to an element in $\mathcal{D}'(\mathbb{R}^d)$. Another, common fact between these representations is the use of basis of some functional space. In these examples, orthonormal basis. Orthonormal basis are particular cases of unconditional basis, this is related to stability and sometimes, in practice

[a] In some literature this hypothesis is called that the signal is band limited.

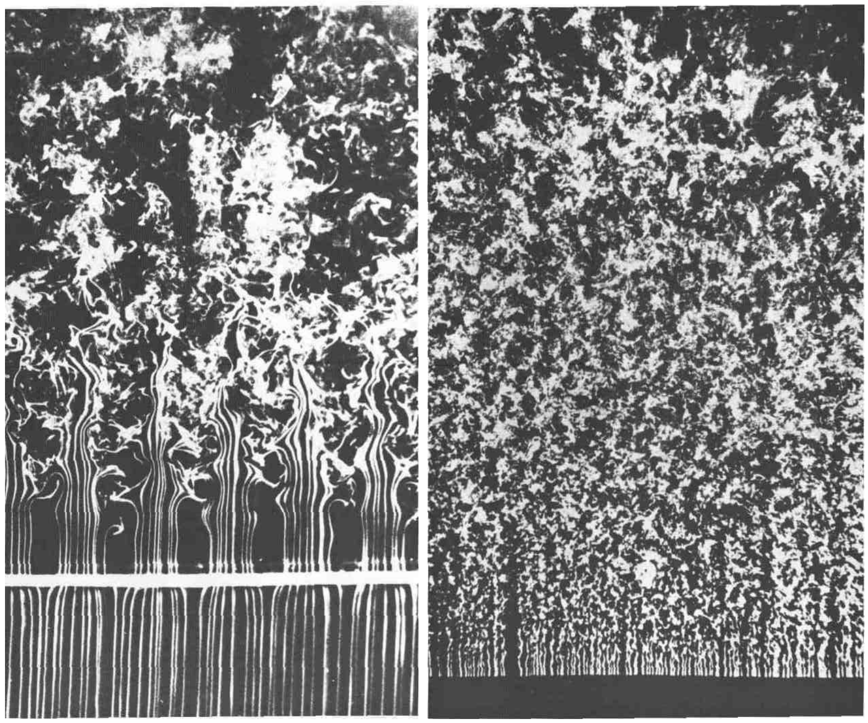

Figure 1.1: Two turbulent flows obtained from a uniform grid. Compare with figs. 5.1 and 5.2. Photographs by Thomas Corke and Hassan Nagib.

this could be the only condition required. Today, thanks to the advance of the theory and to the existence of more sophisticated devices, we have a wide range of tools to represent signals. Wavelets provide an example of a now widely used mathematical tool in this context. An important property of them, is that under mild conditions, they are unconditional basis of the $L^p(\mathbb{R}^d)$ spaces. Other systems also have this property. So, it would be interesting to study random series using unconditional basis. In **Chapter 4** we study random series in $L^p(X, \Sigma, \mu)$, with independent terms and/or using unconditional basis.

On the other hand, we will also study random "weighted" sums of vectors which have some structure, such as forming a basis of a subspace. This is also interesting, from the point of view of some applications. Since, it is a key problem in engineering to represent a signal (the random process) using the less information as possible. In mathematical language, this information is captured in the coordinate coefficients with respect to a fixed basis (or other set of vector with "good" properties) The prescribed basis is generally fixed. On the other hand, once we have recorded the information, captured in the coefficients, we would like to reconstruct the original signal. This operation, corresponds to writing the signal as a series using the coordinate coefficients. Naturally, at this point, is where some convergence problems could arise.

We will take special attention to sums of independent random elements. A typical example, is the Karhunen-Loève expansion of Gaussian processes. From an information theoretical point of view, to describe a signal with independent (or decorrelated, at least) random coefficients is very efficient. Some of the results resemble this original result. Additionally, Karhunen-Loève like expansions have, in general, proven to useful in development and interpretation of classical detection theory [77].

In **Chapter** 6 we will consider a related topic to theorems 1.0.1 and 1.0.2. One could note that this interpolation formula implies the weaker condition $X_t \in \overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}}$, $\forall\, t \in \mathbb{R}$ (with respect to the $L^2(\Omega, \mathcal{F}, \mathbf{P})$ norm) . In other circumstances, given a wide sense stationary process $\{X_t\}_t$ also appears the problem of approximating a random variable from $\overline{\text{span}}\{X_t\}_t$ by means of an element h of a closed subspace of $S \subseteq \overline{\text{span}}\{X_t\}_t$. It is interesting to find conditions under which there exists a basis or a "good" subset which permit us to write h as a convergent series. On the other hand, looking at theorem 1.0.2, a desirable property could be to have orthogonal random coefficients, i.e. they are an orthonormal basis of the closed linear span of the whole process. One would like to have at least another weaker condition which also assures unconditional convergence. This problem was first considered by Kolmogorov, Rozanov (e.g. [69]) among others. They studied conditions under which a wide sense stationary process $\{X_t\}_t$ is minimal, forms a basis, or even a Riesz basis. This conditions are generally given in terms of the spectral measure of the process. However, unconditional convergence could be also obtained if the stationary sequence is a frame of its closed linear span. In this chapter we will study this problem.

From, a practical point of view this could correspond to reconstructing in a stable form a continuous parameter process from discrete samples of a "filtered" or "measured" version of the original process.

Additional Comments.

This work in the beginning was inspired by this applications in mind. However, it is important to mention, that these topics on the convergence of L^p- valued, or sums of generalized random processes, are closely related to other problems, such as the existence of certain stochastic integrals [15],[42]. On the other hand, some of these tools have been used, with some success, in the study of the geometry of Banach spaces [73] [13].

Moreover, from other point of view, to study some of these problems may be interesting *per sé*. Some of the first problems on random series of functions were treated by Paley, Zygmund and Wiener. In particular, in [61] [62] [63], Paley and Zygmund posed a series of questions about the following random Fourier series:

$$\sum_{n=0}^{\infty} X_n \cos(nt + \Phi_n)\,,$$

where $\{X_n\}_n$ and $\{\Phi_n\}_n$ are sequences of real random variables such that the $X_n e^{i\Phi_n}$'s are independent random variables . They give conditions under one of these series, with probability one: i) is a Fourier-Stieltjes series ii) represent an $L^p(\mathbb{T})$ function. iii) Converges for almost all $t \in \mathbb{T}$ with respect to the Lebesgue measure. The idea behind this, in part, was the following [40]: Sometimes is difficult to exhibit a concrete function which full fills certain requirements, but it could become relatively easier, using a “randomization device” to prove the almost sure existence of such function. On the other hand, the study of these random series is also related

to the study of series representations of stationary and related processes [51], [41], [39]. Then it would be interesting to study random weighted sums using other systems different of the trigonometric.
The results exposed in this work, in my opinion, in some way lay somewhere between these motivations, as a consequence the series that we are going to study take values in $L^p(X, \Sigma, \mu)$ or the space of distributions $\mathcal{D}'(\mathbb{R}^d)$. However, is worth mention that other examples are the spaces $D[0,1]$ and $C[0,1]$ used in statistics (Billingsley 1968 [10]), which we are not going to treat here. From the point of view of these applications, other contrast with usual statistics is the following: one may ask the practical value of a result such as theorem 1.0.1, and the main doubt could come from the way on how such random coefficients are calculated or estimated. In general, the practical value of such a result relies on that, at some level, one assumes the existence of certain device (an A/D converter for example) which allows to capture the actual measurement of the random signal or a filtered version of it. So, this makes unnecessary to consider any estimation problems, at least up to certain level of the processing. So following this line of work, we will not make any direct mention of these sort of estimation problems. On the other hand to have a series representation of a process may provide a useful tool for modelling some problems.

Thesis organization.

In **Chapter 2** we review some basic characterizations of stochastic processes as random variables taking values in vector spaces. In particular we will be interested in normed and metric spaces. On the other hand we will make a brief review of some basic probability results and we will introduce most of the necessary analytical results which we are going to use. The great majority of these results are exposed without proof since most of them are more or less known or are mainly accessible when looking for a reference. We will give proofs for the few exceptions which are not in these cases.

In **Chapter 3** we review general results on the sum of independent random elements which are going to be used in the development of our results. In order to make this work as self-contained as possible in this case we prefer to give a complete proof or a sketch of it, at least. The majority of these results are spread in a variety of research articles and specialized literature of arguable accessibility and no unified appropriate source of reference was found. On the other hand, some of the results were adapted to the necessities of this work.

In **Chapter 4** we study the particular case of random variables with values in L^p spaces. We begin with some technical discussions which could arise when beginning to study this topic. For example, if $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $Y :\longrightarrow L^p(X, \Sigma, \mu)$ a random variable, then given $\omega \in \Omega$, $Y(\,.\,,\omega)$ represents an element of $L^p(X, \Sigma, \mu)$ thus a Σ-measurable function. So some technical questions about Y could arise when we want to treat Y as a two variables function defined over the product space $X \times \Omega$. This is related to the problem of treating Y as L^p valued random variable or as $\{Y_x\}_x$ scalar valued random variable indexed by $x \in X$. After this we introduce sums of independent random variables $\sum_i X_i$ where the X_i's take values in $L^p(X, \Sigma, \mu)$. Then we study the case when $X_i = a_i f_i$, with the a_i's being scalar valued random variables and the f_i's are fixed. We will study several conditions on the a_i's and on the f_i's, such as when the f_i's constitute an unconditional basis of $L^p(X, \Sigma, \mu)$. We study relationships between the a.s. convergence en $L^p(X, \Sigma, \mu)$ norm and the convergence in the p-th mean, i.e. respect to the norm $(\mathbb{E} \| \,.\, \|^p)^{1/p}$. The almost everywhere convergence with respect to the product space $(X \times \Omega, \Sigma \otimes \mathcal{F}, \mu \times \mathbf{P})$ is also studied. Finally, we discuss the application of some of these results to the construction of random process with a certain prescribed structure. An example of this is fractional Brownian Motion over a finite interval. The particular choice of this process, is related, to the spacial place which has taken in some applications in recent years. However, in other contexts this type of construction is no longer possible. For example, some problems may arise if we want to construct in this way some related processes, such as $\frac{1}{f}$ processes on the whole space $\mathbb{R}^d$.

In **Chapter 5**, we consider again the latter problems, but in this case we consider series taking values in the space of distributions. First we intruduce the class of generalized random processes, which play a similar role to that of the generalized functions. We discuss briefly some properties of the covariance functional of these processes. Then we introduced as an auxiliar tool the Sobolev spaces $H^s(\mathbb{R}^d)$. Then we study the construction of generalized random processes, with a prescribed covariance, by means of series. This approach seems to be more appropriate when dealing, for example, with some random process, such as "fractional" random fields, which exhibit long range dependence, and are defined over the whole space $\mathbb{R}^d$. As a final application we shall give a series expansion of these spatial processes or fields.

In **Chapter** 6 we study necessary and sufficient conditions for a stationary sequence to form a Riesz basis or a frame, then these results are related to the problem reconstructing a stationary random process by means of a convergent series using its samples.

1.1 Included Publications

Several results contained in this thesis have appeared as research articles in refereed journals and have been presented as individual contributions in conferences. Chapters 4, 5 and 6 include the following papers:

1) Medina J.M. Cernuschi-Frías B. "Random Series in $L^p(X, \Sigma, \mu)$ using unconditional basic sequences and l^p stable sequences: A result on almost sure almost everywhere convergence". *Proceedings of the American Mathematical Society*, 135(11), pp. 3561-3569. 2007.
Part of this work was also presented at the *2006 IEEE Information Theory Workshop, ITW 2006*, held at Punta del Este, Uruguay. Pages 342-344 of the conference proceedings.

2) Medina J.M. Cernuschi-Frías B. "On the a.s. convergence of certain random series to a fractional random field in $\mathcal{D}'(\mathbb{R}^d)$". *Statistics and Probability Letters*, 74(2005), pp. 39-49.

3) Medina J.M. Cernuschi-Frías B. "Wide Sense Stationary Processes forming Frames", To appear, accepted for publication in the *IEEE Transactions on Information Theory.* Part of this work was presented at the *International Symposium on Information Theory and its Applications 2010* as "Stationary Sequences and Stable Sampling" (pp.94-99 of the conference proceedings), and at the *2009 IEEE Statistical Signal Processing Workshop.*

Chapter 2

Preliminaries

"Prove all things; hold fast that which is good".

St. Paul in *The Bible, 1 Thess 5:21.*

In this chapter we review some mathematical tools which are going to be used throughout this thesis. The exposition given in this chapter is far from complete. Most of the results are presented without proof since many of them are more or less known results or are accessible in many textbooks. We will give proofs for the few exceptions which do not fall in this case. On the other hand, our intention is just to fix some notation and definitions. It aims to make the exposition of this work as self contained as possible for the reader. However, many important (and classic) results which are omitted in this chapter and are used in this work, will be introduced throughout the following chapters. Sometimes, we will just recall a result giving an appropriate reference.

2.1 Some Concepts of Probability Theory

The results, definitions and exposition of the theory in this section mainly follows [33], [16], [43], [9].

2.1.1 First definitions and basic properties

A probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is a measure space, with a measure $\mathbf{P}$ defined over a σ-algebra $\mathcal{F}$, of subsets of Ω. Such that $\mathbf{P}(\Omega) = 1$. This measure $\mathbf{P}$ is called a probability measure or just a *probability*. A random variable, is a measurable function $X : \Omega \longrightarrow \mathbb{R}$ (or $\mathbb{C}$). That is, for every Borelian subset A, $X^{-1}(A) \in \mathcal{F}$. Sometimes the subsets belonging to $\mathcal{F}$ are referred as "events". When a property holds a.e. $[\mathbf{P}]$ we say that this property holds almost surely (a.s.). Another important concept arising in probability is independence:

Definition 1. Let $\mathcal{A}$ be a collection of subsets in $\mathcal{F}$, we say the sets in the collection $\mathcal{A}$ are independent if

$$\mathbf{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbf{P}(A_i)\,,$$

for every finite sub collection $\{A_i\}_{i=1\ldots,n}$ of different sets in $\mathcal{A}$.

In a similar manner is possible to define:

Definition 2. Let $\mathcal{D}$ be a collection of random variables, we say that the random variables of $\mathcal{D}$ are independent if

$$\mathbf{P}\left(\bigcap_{i=1}^{n} X_i^{-1}(A_i)\right) = \prod_{i=1}^{n} \mathbf{P}(X_i^{-1}(A_i)),$$

for every finite subset of $\mathcal{D}$, $\{X_i\}_{i=1\ldots,n}$ of different random variables and every finite collection of Borel subsets $\{A_i\}_{i=1\ldots,n}$.

Given a random variable X we define its *Law* or distribution function by

$$F_X(x) = \mathbf{P}(X^{-1}(-\infty, x]).$$

In this way a probability measure is induced over $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, for every Borelian set A given by $\mu_X(A) = \mathbf{P}(X^{-1}(A))$. Sometimes, we will denote $\mathscr{L}(X)$ the law of X. In the same manner, considering several random variables it is possible to define a probability measure over $\mathbb{R}^n$. As usual, for every measurable function-random variable we introduce the Lebesgue integral respect to the probability $\mathbf{P}$, which in the context of probability is called *expected value of the random variable* X and is denoted as $\mathbb{E}(X)$ or $\mathbb{E}X$. Then:

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbf{P}.$$

Note, that the value of $\mathbb{E}(X)$ can be obtained as an integral over $\mathbb{R}$, that is

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbf{P} = \int_{\mathbb{R}} x d\mu_X,$$

moreover for every real Borel measurable function g, whenever this expressions exists, we have

$$\mathbb{E}(g(X)) = \int_{\Omega} g \circ X d\mathbf{P} = \int_{\mathbb{R}} g(x) d\mu_X.$$

Having defined the integral, as usual, for $0 < p < \infty$ we introduce the Lebesgue spaces of random variables $L^p(\Omega, \mathcal{F}, \mathbf{P}) = \{X : \mathbb{E}|X|^p < \infty\}$. Defining for $p = \infty$, $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ as the space of essentially bounded functions with respect to the measure $\mathbf{P}$. For $p \geq 1$ the L^p spaces are Banach spaces considering the norm $\|X\|_{L^p} = (\mathbb{E}|X|^p)^{1/p}$. For $p < 1$ this expression it is not a norm, however the L^p spaces are complete metric spaces with the distance $d(X, Y) = \mathbb{E}|X - Y|^p$.
A consequence of independence, is the following:

Theorem 2.1.1. *If X, Y are independent random functions, neither of which vanishes a.s. then a necessary and sufficient condition that both X and Y be integrable is that their product XY be integrable, if this condition is satisfied then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.*

2.1.2 Some results on convergence of sequences and series of independent random variables

First, let us give some definitions:

Definition 3. Let X_n, X be real random variables, then:
i) $\{X_n\}_n$ converges to X in law ($X_n \longrightarrow_{\mathscr{L}} X$) if $\mathscr{L}(X_n) \longrightarrow_w \mathscr{L}(X)$.
ii) $\{X_n\}_n$ converges to X in probability ($X_n \longrightarrow_{pr} X$) if for every $\epsilon > 0$, $\mathbf{P}(|X_n - X| > \epsilon) \longrightarrow 0$. i.e. $\{X_n\}_n$ converges to X in $\mathbf{P}$ measure.
iii) $\{X_n\}_n$ converges to X in L^p ($X_n \longrightarrow_{L^p} X$) or in the p-th mean if $\mathbb{E}|X_n - X|^p \longrightarrow 0$. i.e. $\{X_n\}_n$ converges to X in the L^p norm.
iv) $\{X_n\}_n$ converges to X almost surely ($X_n \longrightarrow X$ a.s.) if $X_n(\omega) \longrightarrow X(\omega)$ for almost all $\omega \in \Omega$ $[\mathbf{P}]$. i.e. the sequence converges a.e. with respect to the measure $\mathbf{P}$.

Remark 2.1.1. The notion of convergence in probability is compatible with the metric given by $d(X,Y) = \mathbb{E}\frac{|X-Y|}{1+|X-Y|}$ i.e. $X_n \longrightarrow_{pr} X$ if and only if $d(X_n, X) \longrightarrow 0$ whenever $n \longrightarrow \infty$. However a.s. convergence it is not compatible with any metric, moreover is not compatible with any topological notion. For further comments on this see section 2.5 at the end of this chapter.

Let us review the relationships between this types of convergence:
Convergence in $L^p \Longrightarrow$ convergence in probability (by Chevychev´s inequality) $\Longrightarrow$ Convergence in Law. But not conversely. Almost sure convergence $\Longrightarrow$ convergence in probability and hence in Law or in distribution. But not conversely.
In finite measure spaces there is a basic relationship between almost everywhere (almost sure) convergence and convergence in norm (mean convergence). For this purpose we need the following definition:

Definition 4. Let $\{X_n\}_n$ be a sequence of random variables, we say that $\{X_n\}_n$ is uniformly integrable if

$$\lim_{\alpha\to\infty} \sup_n \int_{\{|X_n|>\alpha\}} |X_n| \, d\mathbf{P} = 0 \,. \tag{2.1.1}$$

Then, it can be proved the following :

Theorem 2.1.2. *Let $p \geq 1$ and $\{X_n\}_n \subset L^p(\Omega, \mathcal{F}, \mathbf{P})$ be a sequence, such that $X_n \longrightarrow X$ a.s. as $n \longrightarrow \infty$ then: $\mathbb{E}|X_n - X|^p \longrightarrow 0$ when $n \longrightarrow \infty \iff \{|X_n|^p\}_n$ is uniformly integrable.*

It is easy to prove, that a sufficient condition for $\{X_n\}_n$ to be uniformly integrable is:

$$\exists\, \epsilon > 0\,, K > 0 \textit{ such that } \mathbb{E}|X_n|^{1+\epsilon} \leq K \,\forall\, n. \tag{2.1.2}$$

Now we are going to study the convergence of sums of independent random variables. Almost all of the results in this section are, more or less classics, so many of them are stated without proof. Nevertheless, in the following chapter we will give a proof for some of these results in a more abstract setting. Then, the reader might notice that many times, the proofs of the infinite dimensional versions of the results of this section, will not defer from the general idea of their real valued versions. Now, we give a result for sums of independent random variables, known as Kolmogorov inequality.

Theorem 2.1.3. *[43] (generalized Kolmogorov inequality) Let $X_1, X_2, \ldots$ be independent random variables with $\mathbb{E}X_i = 0\,\forall i$, and let $p \geq 1,\ \lambda > 0$ then*

$$\mathbf{P}\left(\max_{j=1,\ldots,n}\left|\sum_{i=1}^{j} X_i\right| > \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}\left|\sum_{i=1}^{n} X_i\right|^p .$$

With this inequality it is possible to prove the following:

Theorem 2.1.4. *[43] Let $X_1, X_2, \ldots$ be independent random variables with $\mathbb{E}X_i = 0\,\forall i$, and let $p \geq 1$. Suppose that $\sum_{i=1}^{\infty} X_i$ converges in $L^p(\Omega, \mathcal{F}, \mathbf{P})$. Then $\sum_{i=1}^{\infty} X_i$ converges a.s.*

In particular, we have seen that if a series of independent random variables converges en p-mean then it converges a.s.. In the L^2 case this can be written as:

Theorem 2.1.5. *[43], [16], [9] Let $X_1, X_2, \ldots$ be independent random variables with $\mathbb{E}X_i = 0\,\forall i$. Suppose that $\sum_{i=1}^{\infty} Var(X_i) < \infty$. Then $\sum_{i=1}^{\infty} X_i$ converges a.s.*

Now, let us state a partial converse of this result:

Theorem 2.1.6. *[41] If $\{X_n\}_n$ is a sequence of independent random variables and c is a positive constant such that $\mathbb{E}(X_n) = 0$ and $|X_n| \leq c$ a.s. , $n = 1, \ldots,$ and if $\sum_{i=1}^{\infty} X_i$ converges a.s. then $\sum_{i=1}^{\infty} Var(X_i) < \infty$.*

All the preceding results on series are included in the following very general assertion, known as Kolmogorov´s three series theorem:

Theorem 2.1.7. *[43], [16], [9] If $\{X_n\}_n$ is a sequence of independent random variables and c is a positive constant, and if $E_n = \{|X_n| \leq c\}$, $n = 1, \ldots$, then a necessary and sufficient condition for the a.s. convergence of the series $\sum_{i=1}^{\infty} X_i$ is the convergence of all the three series: i) $\sum_{n=1}^{\infty} \mathbf{P}(E_n^c)$. ii) $\sum_{n=1}^{\infty} \mathbb{E}X_n \mathbf{1}_{E_n}$. iii) $\sum_{n=1}^{\infty} Var(X_n \mathbf{1}_{E_n})$*

Theorem 2.1.3 is an example, of a general phenomenon: For sums of independent random variables, if $\max_{1\leq k\leq n} |S_k|$ is large, then $|S_n|$ is probably large as well. The previous theorem is an instance of this, and so is the following result :

Theorem 2.1.8. *[9], [16] Suppose that $X_1, \ldots, X_n$ are independent. For $\lambda > 0$:*
a)

$$\mathbf{P}\left(\max_{1\leq k\leq n} |\sum_{i=1}^{k} X_i| > 3\lambda\right) \leq 3 \max_{1\leq k\leq n} \mathbf{P}\left(|\sum_{i=1}^{k} X_i| > \lambda\right) .$$

b) Additionally if the X_i's are symmetric then,

$$\mathbf{P}\left(\max_{1\leq k\leq n} |\sum_{i=1}^{k} X_i| > \lambda\right) \leq 2\mathbf{P}\left(|\sum_{i=1}^{n} X_i| > \lambda\right) .$$

With this result, one can prove the following:

Theorem 2.1.9. *[9], [16] For an independent sequence $\{X_n\}_n$: $\sum_{i=1}^{\infty} X_i$ converges a.s. if and only if it converges in probability.*

Remark 2.1.2. Moreover, under this hypothesis, these modes of convergence are equivalent to convergence in law.

Let us prove another similar result to theorem 2.1.8:

Proposition 2.1.1. *Suppose that $X_1, \ldots, n$ are independent random variables. Then:*
a) For $\lambda_1, \lambda_2 \geq 0$,

$$\mathbf{P}\left(\max_{1\leq k\leq n} |X_k| > \lambda_2 + \lambda_1\right) \leq \frac{\mathbf{P}\left(\max_{1\leq k\leq n} |\sum_{i=1}^{k} X_i| > \lambda_1\right)}{\mathbf{P}\left(\max_{1\leq k\leq n} |\sum_{i=1}^{k} X_i| \leq \lambda_2\right)},$$

b) and if the X_i's are symmetric, then,

$$\mathbf{P}\left(\max_{1\leq k\leq n} |X_k| > \lambda\right) \leq 2\mathbf{P}(|S_n| > \lambda).$$

Proof. The proof is analogous to theorem 2.1.8. For fixed $\lambda_1, \lambda_2 \geq 0$, and $i = 1, \ldots, n$, let us denote $A_i = \{|X_i| > \lambda_1 + \lambda_2, |X_j| \leq \lambda_1 + \lambda_2, \forall\, i < j \leq n\}$. Then for $i \neq j$ the A_i's are disjoint, and $\bigcup_{i=1}^{n} A_i = \{\max_{1\leq k\leq n} |X_k| > \lambda_2 + \lambda_1\}$ and A_i are independent of $X_1, \ldots, X_{n-1}$. Now, if $i = 1, \ldots, n$,

$$A_i \bigcap \{|S_{i-1}| \leq \lambda_2\} \subset \{|S_i| > \lambda_1\} \subset \left\{\max_{1\leq k\leq n} |S_k| > \lambda_1\right\},$$

we have

$$\mathbf{P}\left(\max_{1\leq k\leq n} |X_k| > \lambda_2 + \lambda_1\right) \min_{1\leq i\leq n} \mathbf{P}(|S_i| \leq \lambda_2) \leq \mathbf{P}\left(\max_{1\leq k\leq n} |S_k| > \lambda_1\right)$$

this minimum is equal or greater than $\mathbf{P}(\max_{1\leq k\leq n} |S_k| \leq \lambda_2)$, then a) follows from this.
If the X_i's are symmetric, take the same A_i's as in part a), and write $\lambda_1 + \lambda_2 = \lambda$. Then

$$A_i \subset \left(A_i \bigcap \{|S_n| > \lambda\}\right) \bigcup \left(A_i \bigcap \{|S_n - 2X_i| > \lambda\}\right),$$

so that

$$\mathbf{P}(A_i) \leq \mathbf{P}\left(A_i \bigcap \{|S_n| > \lambda\}\right) + \mathbf{P}\left(A_i \bigcap \{|S_n - 2X_i| > \lambda\}\right).$$

Since the X_i's are independent and symmetric, the last two probabilities are equal. Then

$$\mathbf{P}(A_i) \leq 2\mathbf{P}\left(A_i \bigcap \{|S_n| > \lambda\}\right),$$

then summing over i, we get the desired result. □

Finally, let us discuss to results which are also a consequence of 2.1.8:

Proposition 2.1.2. *Let $X_1, \dots, X_n$ be independent and symmetric random variables, Then:*
a) For any $a_1, \dots, a_n \in \mathbb{R}$ and $\lambda > 0$:

$$\mathbf{P}\left(\left|\sum_{i=1}^{n} a_i X_i\right| > \lambda\right) \leq 2\mathbf{P}\left(\max_{1\leq i\leq n} |a_i||S_n| > \lambda\right) .$$

b) If in addition, we have a sequence $Y_1, \dots, Y_n$ of random variables such that $|Y_i| \leq 1$ a.s. and such that $X_1Y_1, \dots, X_nY_n$ is a sequence of independent and symmetric random variables. Then

$$\mathbf{P}\left(\left|\sum_{i=1}^{n} Y_i X_i\right| > \lambda\right) \leq 2\mathbf{P}(|S_n| > \lambda) .$$

Proof. a) Without loss of generality, suppose that $1 = a_1 \geq a_2 \dots, \geq a_n \geq 0$. And writing $a_{n+1} = 0$, we have:

$$\sum_{i=1}^{n} a_i X_i = \sum_{i=1}^{n} a_i (S_i - S_{i-1}) = \sum_{i=1}^{n} (a_i - a_{i+1}) S_i .$$

Since $\sum_{i=1}^{n} (a_i - a_{i+1}) = 1$,

$$\left\{ |\sum_{i=1}^{n} (a_i - a_{i+1}) S_i| > \lambda \right\} \subset \left\{ \max_{1\leq k\leq n} |S_k| > \lambda \right\} ,$$

now the result follows from proposition 2.1.8.
b) Let $(B_i)_i$ be a Bernoulli sequence, which is independent of the sequences $(Y_iX_i)_i$ and $(X_i)_i$. In view of the symmetry, the sequence $(X_iY_iB_i)_i$ has the same distribution as $(X_iY_i)_i$. The sequences $(X_i)_i$ and $(X_iB_i)_i$ also have identical distributions, then

$$\mathbf{P}\left(\left|\sum_{i=1}^{n} Y_i X_i\right| > \lambda\right) = \mathbf{P}\left(\left|\sum_{i=1}^{n} B_i Y_i X_i\right| > \lambda\right) .$$

Now, the result follows from applying part a) conditionally for $a_i = Y_i$ and $X_i' = X_i B_i$.

□

Khinchine´s inequalities and Rademacher functions

In the study of sums of independent random variables it is useful to introduce the Rademacher functions [41]. The investigation of the convergence problem of a series of Rademacher functions made by Khinchine and Kolmogorov motivated Kolmogorov to the further development of the theory of series of independent random variables. On the other hand, these functions provide a bridge between probability theory and some problems related to the study of Banach spaces [46].

Definition 5. Define the Rademacher functions $r_n(t)$, $n = 0, 1, \ldots$ over $[0,1]$ by

$$r_n(t) = \begin{cases} \text{sgn}(\sin(2^n \pi t)) \text{ if } t \neq \frac{k}{2^n} \\ 0 \text{ if } t = \frac{k}{2^n} \end{cases}$$

where $k = 0, \ldots, 2^n$.

The following properties are easy to show:
1) $\{r_n(t)\}_{n=0}^{\infty}$ is an orthonormal system, and,
2) $\int\limits_{[0,1]} r_n(t)dt = 0$.
3) The Rademacher system is not complete in $L^2[0,1]$.
4) Now if we take $\Omega = [0,1]$ for a probability space in which the σ-field of all measurable sets is considered and the Lebesgue measure is taken to be the probability, then each $r_n(t)$ is a sequence of independent random variables, with $\mathbb{E}r_n = 0$ and $Var(r_n) = 1$. A fundamental property of the Rademacher functions is the following:

Theorem 2.1.10. *(Khinchine´s inequality for Rademacher functions) For $p > 1$, there exists positive constants A_p, B_p, such that*

$$A_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2} \leq \left(\int\limits_{[0,1]} \left| \sum_{n=1}^{N} a_n r_n(t) \right|^p dt \right)^{1/p} \leq B_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2}.$$

Using a symmetrization argument and the the previous theorem, this can be generalized to the following result for general random variables:

Theorem 2.1.11. *Let $\{X_k\}_{k=1}^{n}$ be a subset of independent random variables. Let $p \geq 1$. Suppose that $X_k \in L^p(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathbb{E}(X_k) = 0$, for $k = 1, \ldots, n$. Then there exist positive constants A_p, B_p depending only on p, such that*

$$A_p \mathbb{E} \left(\sum_{k=1}^{n} |X_k|^2 \right)^{p/2} \leq \mathbb{E} \left| \sum_{k=1}^{n} X_k \right|^p \leq B_p \mathbb{E} \left(\sum_{k=1}^{n} |X_k|^2 \right)^{p/2}.$$

Corollary 2.1.1. *Let $\{X_k\}_{k \in \mathbb{N}}$ be a sequence of independent random variables. Let $p \geq 1$. Suppose that $X_k \in L^p(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathbb{E}(X_k) = 0$, for $k = 1, \ldots$. If $\sum\limits_{k \in \mathbb{N}} X_k$ converges a.s. , end if the sum is in $L^p(\Omega, \mathcal{F}, \mathbf{P})$, then*

$$A_p \mathbb{E} \left(\sum_{k \in \mathbb{N}} |X_k|^2 \right)^{p/2} \leq \mathbb{E} \left| \sum_{k \in \mathbb{N}} X_k \right|^p \leq B_p \mathbb{E} \left(\sum_{k \in \mathbb{N}} |X_k|^2 \right)^{p/2}.$$

2.1.3 Conditional Expectation and Discrete Martingales

Recall that if $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$, given $X \in L^1(\Omega, \mathcal{F}, \mathbf{P})$, we define the conditional expectation $\mathbb{E}[X|\mathcal{G}]$ of X relative $\mathcal{G}$ to the equivalence class of random variables satisfying:
1) $\int\limits_A \mathbb{E}[X|\mathcal{G}]d\mathbf{P} = \int\limits_A X d\mathbf{P}$, $\forall\, A \in \mathcal{G}$. In particular $\mathbb{E}[X|\mathcal{G}]$ is integrable.
2) $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable, i.e. $\mathbb{E}[X|\mathcal{G}]^{-1}(A) \in \mathcal{G}$, $\forall\, A$ in the Borel σ-algebra.
Using the Radon-Nykodym theorem it is possible to prove:

Theorem 2.1.12. *If X is integrable and $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$, then there exists a unique equivalence class of integrable, $\mathcal{G}$-measurable, random variables $\mathbb{E}[X|\mathcal{G}]$, such that $\int_A \mathbb{E}[X|\mathcal{G}]d\mathbf{P} = \int_A Xd\mathbf{P}$, $\forall\, A \in \mathcal{G}$.*

The conditional expectation defines a linear operator $\mathbb{E}[\,.\,|\mathcal{G}] : L^p \longrightarrow L^p$, and has the following properties:

Theorem 2.1.13. *Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$, and let $p \geq 1$. Then, given $X \in L^p(\Omega, \mathcal{F}, \mathbf{P})$:*
1) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$.
2) $\mathbb{E}|\mathbb{E}[X|\mathcal{G}]|^p \leq \mathbb{E}|X|^p$.
3) If $\mathcal{G} \subseteq \mathcal{H}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$.

Now, let us introduce a particular type of sequences. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of integrable random variables , and let $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subset \dots$ be an increasing sequence of sub σ-algebras of $\mathcal{F}$. Assuming that each X_n is $\mathcal{G}_n$-measurable, the sequence $\{X_n\}_n$ is said to be a *martingale* relative to the $\{\mathcal{G}_n\}_n$ if for all $n = 1, \dots$ $\mathbb{E}[X_{n+1}|\mathcal{G}_n] = X_n$. Martingales are the natural extension of sequences of partial sums of independent random variables.

Example. Let $\{X_n\}_n$ be a sequence of independent random variables, with $\mathbb{E}(X_n) = 0$. Then, $S_n = \sum_{k=1}^{n} X_k$, $\{S_n\}_n$ is a martingale relative to $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$, the σ-algebra generated by $X_1, \dots, X_n$.

Some inequalities and convergence

We will use some basic results on martingales. The following is a generalization of theorem 2.1.3

Proposition 2.1.3. *[32] Let $\{X_n\}_n$ be a martingale relative to $\{\mathcal{G}_n\}_n$, then for $p \geq 1$, and $\lambda > 0$,*

$$\mathbf{P}\left(\max_{1\leq k\leq n} |X_k|\right) \leq \frac{\mathbb{E}|X_n|^p}{\lambda^p}.$$

Another important result is the following:

Theorem 2.1.14. *(Doob´s inequality)[32] Let $\{X_n\}_n$ be a martingale relative to $\{\mathcal{G}_n\}_n$, then for $p > 1$,*

$$\mathbb{E}|X_n|^p \leq \mathbb{E}\left(\max_{1\leq k\leq n} |X_k|\right)^p \leq q^p\mathbb{E}|X_n|^p$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

The following result due to Lévy, was historically the first of the martingale convergence theorems.

Theorem 2.1.15. *[32] Let $\{\mathcal{G}_n\}_n$ be an increasing sequence of sub σ-algebras of $\mathcal{F}$, and let $\mathcal{G}_\infty$ be the σ-field generated by $\bigcup_n \mathcal{G}_n$. If Y is integrable, and $X_n = \mathbb{E}[Y|\mathcal{G}_n]$, then $X_n \longrightarrow \mathbb{E}[Y|\mathcal{G}_\infty]$ a.s. and in $L^1(\Omega, \mathcal{F}, \mathbf{P})$.*

For further references on this topic, read the end of this chapter.

2.2 Random Variables taking values on vector spaces

The main reference for this section is [76]. The idea of this section is to set a framework which among other things, will enable us to treat stochastic processes as random variables with values in an appropriate function space.

2.2.1 Basic results.

Here we will be concerned with the definition of random variables with values in linear spaces (or random elements in some literature). When possible the definitions and results will be given for topological spaces and linear spaces, which, of course, include linear metric spaces and Banach spaces. In this section, if a particular definition or result requires certain types of linear topological spaces such as separable Banach spaces, it will be stated. However, in the following chapters the results are mainly focused on separable Banach spaces. Throughout this chapter T will denote a topological space and d will denote a semimetric [a]. The class of Borel subsets of T will be denoted by $\mathscr{B}(T)$; that is, $\mathscr{B}(T)$ will be the smallest σ-algebra containing the open subsets of T. In the following, let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. In an analogous way for real valued random variables, we define:

Definition 6. A function $X : \Omega \longrightarrow T$ is said to be a random variable (or random element) in T if $X^{-1}(A) \in \mathcal{F}$ for every $A \in \mathscr{B}(T)$.

As in the real variable case a T valued random variable, induces a probability measure over T, given by $\mu_X(A) = \mathbf{P}(X \in A)$ for every $A \in \mathscr{B}(T)$. This measure will be called the Law of X, and sometimes is denoted by $\mathscr{L}(X)$.
Given an stochastic process say $\{X_t\}_t$ one could think it as a family of real random variables indexed by t, or as mapping $X : \Omega \longrightarrow \mathbb{R}^{\mathbb{R}}$. However, the following example illustrates that a careful construction is needed in this considerations.

2.2.2 A Counter Example.

Let us consider $\mathbb{R}^{\mathbb{R}}$, with the product topology, and consider the Borel σ-algebra $\mathscr{B}(\mathbb{R}^{\mathbb{R}})$ generated by its open sets. On the other hand let us consider an stochastic process $\{X_t\}_{t\in\mathbb{R}}$ with respect to a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Then for each $\omega \in \Omega$, the sample path $X_t(\omega)$ can be regarded as a real valued function of t. But, considering $\{X_t\}_t$ as a mapping from Ω into $\mathbb{R}^{\mathbb{R}}$ some measurability problems may occur. Define an identity function $X = \{X_t\}_{t\in\mathbb{R}}$ from $\Omega = \mathbb{R}^{\mathbb{R}}$ to $\mathbb{R}^{\mathbb{R}}$ by

$$X(\omega) = \omega\,; \ \ \forall\, \omega \in \Omega = \mathbb{R}^{\mathbb{R}}\,.$$

Let $\mathcal{A} = \bigotimes_{t\in\mathbb{R}} \mathscr{B}(\mathbb{R})$, the product space, and let $\mathbf{P}$ be the probability measure degenerate at the origin. Then $X_t(\omega) = \omega(t)$ for each $t \in \mathbb{R}$ and is a random variable since

$$\{\omega :\, X_t(\omega) \leq \alpha\} = \prod_{\mathbb{R}\setminus\{t\}} \mathbb{R} \times (-\infty, \alpha]$$

[a] A non empty set E is called semimetric space if there is a real-valued function d defined on $M \times M$ with the following properties: 1) $d(x, y) = d(y, x) \geq 0$ for all $(x, y) \in M \times M$. 2) $d(x, x) = 0$ for all $x \in M$. 3) $d(x, z) \leq d(x, y) + d(y, z)$.

for each $\alpha \in \mathbb{R}$. however,

$$\bigotimes_{t\in\mathbb{R}} \mathscr{B}(\mathbb{R}) \subsetneq \mathscr{B}(\mathbb{R}^{\mathbb{R}})$$

because $\mathbb{R}$ is uncountable [27]. Thus $X : \Omega \longrightarrow \mathbb{R}^{\mathbb{R}}$ may not be a measurable function.

Now, we list some basic properties and lemmas about T-valued random variables. Most of them, are generalizations of properties for real valued random variables.

Lemma 2.2.1. *If X is a T-valued random variable, and Y is a Borel Measurable function from T into a topological space T', then $Y \circ X$ is a T'-valued random variable.*

Lemma 2.2.2. *Let $T = M$ be a semimetric space with semimetric d. Let $\{X_n\}_n$ be a sequence of random variables in a semimetric space (M, d) such that $X_n(\omega) \longrightarrow X(\omega)$ when $n \to \infty$ for each $\omega \in \Omega$. Then X is a M-valued random variable.*

The previous lemmas can be used to prove that every T-valued random variable in a separable semimetric space $T = M$ is the uniform limit of a sequence of countably valued random variables.

Proposition 2.2.1. *Let M be a separable semimetric space. A mapping $X : \Omega \longrightarrow M$ is a random variable $\Longleftrightarrow \exists \{X_n\}_n$ a sequence of countably valued random variables which converge uniformly to X.*

Proof. (This proof follows closely [76].) $\Longrightarrow$) For each $\lambda > 0$ there exists a countably valued Borel measurable function $f_\lambda : M \longrightarrow M$ such that $\forall x \in M$, $d(f_\lambda(x), x) < \lambda$. Indeed, since M is separable, choose a countable dense subset $\{x_1, x_2, \dots\}$. For $\lambda > 0$ form a countable collection of λ-neighbourhoods $B_\lambda(x_i) = \{x : \ d(x, x_i) < \lambda\}$ that covers M. Define the countably valued Borel measurable function f_λ by

$$f_\lambda(x) = x_1 \text{ if } x \in \ B_\lambda(x_1)$$

and

$$f_\lambda(x) = x_n \text{ if } x \in \ B_\lambda(x_n) \setminus \left(\bigcup_{i=1}^{n} B_\lambda(x_i) \right)$$

for $n = 2, 3, \dots$.
Now, let $X_n = f_n \circ X$ where f_n is the the previous f_λ with $\lambda = \frac{1}{n}$. Then by lemma 6.3.4, X_n is a random variable in M. Moreover, by construction it is countably valued, and $d(X_n, X) < \frac{1}{n}$ uniformly. Thus X_n converges to X uniformly.
$\Longleftarrow$) Is immediate from lemma 2.2.2. □

Many authors define a random variable in a Banach space $T = E$ as a strongly measurable function from a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ to the Banach space. A function $X : \Omega \longrightarrow E$ is said to be strongly measurable if there exists a sequence of countably valued measurable functions X_n such that $X_n \longrightarrow X$ in the norm topology a.s. . For a separable Banach space, the previous lemma shows that the two definitions are (a.s.) the same. For nonseparable Banach spaces the range of a strongly measurable function must be (a.s.) a separable subset. Taking in account, the previous discussion, as in the following chapters we will restrain to separable Banach spaces, we can use one definition or the other.

Lemma 2.2.3. *If X is a random variable in a topological space T and A is a (scalar valued) random variable, then AX is a T valued random variable.*

Proof. See [76]. □

Let us discuss, some topological properties of random variables in topological spaces. Not all of the properties of scalar valued random variables can be extended to random variables taking values in topological spaces. For example sums of random variables are random variables, but sums of T-valued random variables may no be defined. Even when considering linear spaces , separability is often needed to extend the basic properties of random variables. Given a semimetric space (M, d), many of the results concerning random variables taking values in M depend on the fact, that given X, Y random variables, then $d(X, Y)$ is a real valued random variable.

Lemma 2.2.4. *For a separable semimetric space (M, d), if X and Y are M-valued random variables, then $d(X, Y)$ is a scalar valued random variable.*

Proof. See [76]. □

If M is not separable, then $d(X, Y)$ may not be a random variable [76].
Now, if M is a seminormed vector space with a seminorm $\| \, . \, \|$, then, by the previous argument $\|X\|$ is a random variable, if X is a random variable. Moreover, if X is a random variable in a linear topological space M, then $f(X)$ is a random variable for each $f \in M^*$. In a separable seminormed linear space, the converse is also true:

Lemma 2.2.5. *If $(M, \| \, . \, \|)$ is a separable seminormed linear space, then a function $X : \Omega \longrightarrow M$ is a random variable $\Longleftrightarrow$ $f(X)$ is a random variable for each $f \in M^*$.*

Proof. See [76]. □

Remark. Since $f(X+Y) = f(X) + f(Y)$ is a random variable, whenever X, Y are random variables in a seminormed linear space and $f \in M^*$, then the sum of two M-valued random variables is again a random variable. Now, let us define several modes of convergence:

Definition 7. Let (M, d) be a semimetric space, and let $\{X_n\}_n$ be a sequence of M-valued random variables. Then $\{X_n\}_n$ converges to a random variable X,
i) With probability one, or almost surely (a.s.) ($X_n \longrightarrow X$ a.s.) if

$$\mathbf{P}(\lim_{n\to\infty} d(X_n, X) = 0) = 1 \, .$$

ii) In probability ($X_n \longrightarrow_{Pr} X$), if for every $\epsilon > 0$,

$$\mathbf{P}(d(X_n, X) > \epsilon) \underset{n\to\infty}{\longrightarrow} 0 \, .$$

iii) In L^p or in the p-th mean, if for some $p > 0$,

$$\mathbb{E}(d(X_n, X)^p) \underset{n\to\infty}{\longrightarrow} 0 \, ,$$

where $\mathbb{E}(d(X_n, X)^p)$ is assumed to exist.

Other modes of convergence can be defined such as convergence in law, but we will be mostly concerned on results using these three types of convergence. Finally, for $d(X,Y)$, given $\epsilon > 0$, one obtains the following form of Chevychev´s inequality:

$$\mathbf{P}(d(X,Y) > \epsilon) \leq \frac{\mathbb{E}(d(X,Y)^p)}{\epsilon^p},$$

whenever $\mathbb{E}(d(x,y)^p)$ exists. Most of the relationships between the different modes of convergence of scalar valued random variables are also valid for random variables in semimetric spaces. Again we have: Convergence in $L^p \Longrightarrow$ convergence in probability (by Chevychev´s inequality) $\Longrightarrow$ Convergence in Law. But not conversely. Almost sure convergence $\Longrightarrow$ convergence in probability and hence in Law or in distribution. But not conversely.

Definition 8. Two T-valued random variables X, Y are said to be identically distributed if $\mathbf{P}(X \in B) = \mathbf{P}(Y \in B)$, for all $B \in \mathscr{B}(T)$.

Definition 9. A finite collection of random variables is $\{X_1, \ldots, X_n\}$ is said to be independent if

$$\mathbf{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbf{P}(X_1 \in B_1) \ldots \mathbf{P}(X_n \in B_n)$$

for every $B_1, \ldots, B_n \in \mathscr{B}(T)$. An arbitrary collection of random variables is said to be independent if every finite subcollection is independent.

Now we state the following results on the action of linear functionals over independent random variables.

Lemma 2.2.6. *Let $(M, \|\,.\,\|)$ be a separable seminormed linear space. The random variables X, Y are identically distributed $\Longleftrightarrow f(X), f(Y)$ are identically distributed random variables for each $f \in M^*$.*

Proof. See[76] □

Lemma 2.2.7. *Let $(M, \|\,.\,\|)$ be a separable seminormed linear space. The random variables X, Y are independent $\Longleftrightarrow f(X), f(Y)$ are independent random variables for each $f \in M^*$.*

Proof. See[76] □

Expected Value, The Bochner and Pettis Integrals

Let $(E, \|\,.\,\|)$ be a separable Banach space. Here we introduce two concepts of expected value for E-valued random variables.

Definition 10. [34], [4] Let $(E, \|\,.\,\|)$ be a separable Banach space. A random variable X, has an expected value in the sense of Pettis if there exists an element $\mathbb{E}X \in E$ such that $\mathbb{E}f(X) = f(\mathbb{E}X), \forall\, f \in E^*$.

It is simple to check that the Pettis expected value is well defined.

Definition 11. [34] [4] Let $(E, \|\,.\,\|)$ be a separable Banach space. And let X be a random variable with expected value $\mathbb{E}X$. Then we define the variance as $\mathbb{E}\,\|X - \mathbb{E}X\|^2$.

Definition 12. Let X be an E-valued random variable with Pettis expected value $\mathbb{E}X$. Suppose that for each $f \in E^*$, $\mathbb{E}(f(X-\mathbb{E}X))^2 < \infty$. Then the non negative bilinear form on E^* defined by $C(f,g) = (CovX)(f,g) = \mathbb{E}f(X-\mathbb{E}X)g(X-\mathbb{E}X)$ is called the covariance of X.

It can be proved that the covariance operator is a continuous bilinear form on E^*. Also, the covariance operator can be thought as a linear map from E^* to E^{**} the topological dual of E^*.: $(CovX)(f)(g) = (CovX)(f,g)$. It can be shown, that $Cov = A*A$ where A is a bounded operator from E^* into a dense subset of a Hilbert space H.

Definition 13. For countably valued random variables in a separable Banach space, the Bochner integral or expected value, is defined as follows: if $X = \sum_{i=1}^{\infty} f_i \mathbf{1}_{A_i}$, then $\mathbb{E}X = \int_{\Omega} X d\mathbf{P} = \sum_{i=1}^{\infty} f_i \mathbf{P}(A_i)$, when $\sum_{i=1}^{\infty} \|f_i\| \mathbf{P}(A_i) < \infty$. Let X be an arbitrary E-valued random variable. We say that X is Bochner integrable if there exists a sequence $\{X_n\}_n$ of countably valued random random variables, such that: 1) $\|X - X_n\| \longrightarrow 0$ a.s. whenever $n \to \infty$. 2) $\mathbb{E}\|X - X_n\| \longrightarrow 0$ whenever $n \to \infty$. Then we define the Bochner expected value or integral of X as:

$$\mathbb{E}X = \int_{\Omega} X d\mathbf{P} = \lim_{n\to\infty} \int_{\Omega} X_n d\mathbf{P}\,.$$

The following theorem gives a necessary and sufficient condition for the existence of the Bochner integral.

Theorem 2.2.1. *A E-valued random variable has a Bochner expected value* $\Longleftrightarrow$ $\mathbb{E}\|X\| < \infty$, *and in this case* $\|\mathbb{E}X\| \le \mathbb{E}\|X\|$.

Proof. [34] □

Remarks.[34] 1) An important fact is that the Bochner expected value is linear and commutes with linear continuous operators.
2) The existence of the Bochner integral implies the existence of the Pettis expected value, and both coincide. The converse is not true.
3) The Pettis expected value is also linear and commutes with linear continuous operators. Finally, let us prove the following simple result,

Lemma 2.2.8. *Let* X, Y *be independent E-valued random variables, such that* $\mathbb{E}\|X\|^p < \infty$, $\mathbb{E}\|Y\|^p < \infty$, $p \ge 1$ *and* $\mathbb{E}X = \mathbb{E}Y = 0$ *then* $\mathbb{E}\|Y\|^p \le \mathbb{E}\|Y+X\|^p$.

Proof. By the independence and as a consequence of theorem 2.2.1

$$\mathbb{E}\|Y+X\|^p = \mathbb{E}_X \int_{\mathbb{R}} \|y+X\|^p d\mathscr{L}(y) = \int_{\mathbb{R}} \mathbb{E}_X \|y+X\|^p d\mathscr{L}(y) \ge \int_{\mathbb{R}} \|y\|^p d\mathscr{L}(y)$$

□

2.3 Some concepts from the theory of Banach spaces

2.3.1 Bases and Unconditional Bases

The notion of a Schauder basis is due to S. Banach. Banach spaces with bases are presented in a natural way as sequences spaces and are the simplest among all Banach spaces. All results of this subsection are due to S. Banach. They can be found in the classic book of Lindenstrauss and Tzafriri [46].

Definition 14. Let $(E, \|\,.\,\|)$ be a Banach space.
1) A sequence $\{f_n\}_{n\in\mathbb{N}}$ is a basic sequence if for all $f \in \overline{\text{span}}\{f_n\}_{n\in\mathbb{N}}$ there exists a unique sequence $\{a_n\}_{n\in\mathbb{N}}$ of real numbers such that $f = \sum\limits_{n=1}^{\infty} a_n f_n$.
2) $\{f_n\}_{n\in\mathbb{N}}$ is a Schauder basis of E if it is a basic sequence and if $E = \overline{\text{span}}\{f_n\}_{n\in\mathbb{N}}$.

Here, we will not consider other type of bases for infinite dimensional spaces . So we we shall omit the word Schauder. Obviously, for the finite dimensional case the notions of algebraic bases and Schauder bases agree. A very useful class of bases with more precise properties is the class of unconditional bases.

Definition 15. A basic sequence $\{f_n\}_{n\in\mathbb{N}}$ is unconditional if any convergent series $\sum\limits_{i=1}^{\infty} a_i f_i$ converges unconditionally, that is, the series $\sum\limits_{i=1}^{\infty} a_{\pi(i)} f_{\pi(i)}$ converges to the same limit for all permutations π in $\mathbb{N}$.

The following theorem gives several characterizations of unconditional basis,

Theorem 2.3.1. *[25], [46] Let $\{f_n\}_{n\in\mathbb{N}}$ be a basic sequence in a Banach space E. Then the following are equivalent:*
1) $\{f_n\}_{n\in\mathbb{N}}$ is unconditional.
2) If $\sum\limits_{i=1}^{\infty} a_i f_i$ converges, then for all $\epsilon_n = \mp 1$, the series $\sum\limits_{i=1}^{\infty} \epsilon_i a_i f_i$ converges.
3) There exists $K_1 > 0$ such that, for all $\epsilon_n = \mp 1$ and $\{a_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$,

$$\left\| \sum_{i=1}^{\infty} \epsilon_i a_i f_i \right\| \leq K_1 \left\| \sum_{i=1}^{\infty} a_i f_i \right\| .$$

4) If a series $\sum\limits_{i=1}^{\infty} a_i f_i$ converges, then for all $A \subset \mathbb{N}$, the series, $\sum\limits_{i\in A} \epsilon_i a_i f_i$ converges.
5) There exists $K_2 > 0$ such that, for all $A \in \mathbb{N}$ and all $\{a_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$,

$$\left\| \sum_{i\in A} a_i f_i \right\| \leq K_2 \left\| \sum_{i=1}^{\infty} a_i f_i \right\| .$$

Unconditional basis in L^p spaces.

Since we shall deal with the Lebesgue L^p spaces, there is a simple and useful characterization of unconditional basis when we are dealing with σ-finite $L^p(X, \Sigma, \mu)$ spaces. In chapter 4 we will, sometimes, use this equivalence without referring to it, but it will become clear from the context.

Theorem 2.3.2. *$\{f_n\}_{n\in\mathbb{N}}$ is an unconditional basic sequence in $L^p(X,\Sigma,\mu)$ $\Longleftrightarrow$ there exists positive constants A, B, such that for every sequence $(a_n)_n \in \mathbb{R}^{\mathbb{N}}$, $\sum_n a_n f_n \in L^p(X)$:*

$$A\left\|\sum_{n\in\mathbb{N}} a_n f_n\right\| \le \left\|\left(\sum_{n\in\mathbb{N}} |a_n f_n|^2\right)^{1/2}\right\| \le B\left\|\sum_{n\in\mathbb{N}} a_n f_n\right\|.$$

Proof. $\Leftarrow$) Is immediate.
$\Rightarrow$) It suffices to prove the result for finite sequences. If $r_n(t)$ are the Rademacher functions over $[0,1]$, recall theorem 2.1.10, then there exists positive constants A', B' such that

$$A'\left(\sum_{n=1}^{N} |a_n f_n|^2\right)^{1/2} \le \left(\int_{[0,1]} \left|\sum_{n=1}^{N} a_n r_n(t) f_n\right|^p dt\right)^{1/p} \le B'\left(\sum_{n=1}^{N} |a_n f_n|^2\right)^{1/2}, \qquad (2.3.1)$$

now integrating on X and then applying Fubini's theorem we get,

$$\int_X \int_{[0,1]} \left|\sum_{n=1}^{N} a_n r_n(t) f_n\right|^p dt d\mu = \int_{[0,1]} \int_X \left|\sum_{n=1}^{N} a_n r_n(t) f_n\right|^p d\mu dt$$

$$= \int_{[0,1]} \left\|\sum_{n=1}^{N} a_n r_n(t) f_n\right\|_{L^p(X)}^p dt\,.$$

But since $\{f_n\}_n$ is an unconditional basic sequence, by theorem 2.3.1 there exists positive constants K_1, K_2 such that $K_1\left\|\sum_{n=1}^{N} a_n f_n\right\| \le \left\|\sum_{n=1}^{N} a_n r_n(t) f_n\right\| \le K_2\left\|\sum_{n=1}^{N} a_n f_n\right\|$ uniformly in t. Now, by equation 2.3.1, integrating in X we get:

$$\frac{K_1}{B'}\left\|\sum_{n=1}^{N} a_n f_n\right\| \le \left\|\left(\sum_{n=1}^{N} |a_n f_n|^2\right)^{1/2}\right\| \le \frac{K_2}{A'}\left\|\sum_{n=1}^{N} a_n f_n\right\|.$$

□

We shall not use any other results about unconditional bases in L^p spaces, but we refer the reader to [46] [83] for further results on this interesting topic.

2.3.2 Frames and Hilbert Spaces.

Let us review some of the basic results about frames and Hilbert spaces which will be used here. The main reference is [17]. Let $\mathcal{H}$ be a Hilbert space.

Definition 16. A Riesz basis for $\mathcal{H}$ is a family of the form $\{Ue_k\}_{k\in\mathbb{N}}$, where $\{e_k\}_{k\in\mathbb{N}}$ is an orthonormal basis for $\mathcal{H}$ and $U : \mathcal{H} \longrightarrow \mathcal{H}$ is a bounded bijective operator.

There is a very useful characterization of Riesz bases, we will use it in Chapter 6 several times without recalling it explicitly.

Theorem 2.3.3. *For a sequence $\{f_k\}_{k\in\mathbb{Z}}$ in $\mathcal{H}$, the following conditions are equivalent:*
i) $\{f_k\}_{k\in\mathbb{Z}}$ is a Riesz basis.
ii) $\{f_k\}_{k\in\mathbb{Z}}$ is complete in $\mathcal{H}$, and there exists constants, $A, B > 0$ such that for every finite scalar sequence $\{c_k\}_k$ one has:

$$A \sum_k |c_k|^2 \leq \left\| \sum_k c_k f_k \right\|^2 \leq B \sum_k |c_k|^2 .$$

The main feature of a basis is that any vector of the space admits a unique a unique representation as a linear combination of the elements of the basis. A frame, is also a sequence $\{f_k\}_k$ such that every $f \in \mathcal{H}$ admits a representation $f = \sum_k c_k(f) f_k$. However , the corresponding coefficients are not necessarily unique. Thus a frame might not be a basis.

Definition 17. A sequence $\{f_k\}_{k\in\mathbb{Z}}$ of elements in $\mathcal{H}$ is a frame for $\mathcal{H}$ if there exists constants $A, B > 0$ such that

$$A \|f\|^2 \leq \sum_k |\langle f, f_k\rangle|^2 \leq B \|f\|^2 .$$

Theorem 2.3.4. *([17], Chapter 5) If $\{g_k\}_k$ is a Riesz basis of its span then it is a frame.*

We recall the following definition:

Definition 18. Let $\{g_k\}_k$ be a sequence in a Hilbert space $\mathcal{H}$, we say that $\{g_k\}_k$ is minimal if for each j: $g_j \notin \overline{\text{span}}\{g_k\}_{k\neq j}$.

There is an interesting relationship between, minimal sequences and frames:

Theorem 2.3.5. *([17], Chapter 5) Let $\{g_k\}_k$ be a frame in a Hilbert space $\mathcal{H}$, then the following are equivalent:*
i) $\{g_k\}_k$ is a Riesz basis of $\mathcal{H}$.
ii) If $\sum_k c_k g_k = 0$ for $(c_k)_k \in l^2$ then $c_k = 0$ for all k.
iii) $\{g_k\}_k$ is minimal.

Given a frame $\{g_k\}_k$ in $\mathcal{H}$, we can define the associated *frame operator* S defined for every $f \in \mathcal{H}$ by: $Sf = \sum_k \langle f, g_k\rangle g_k$, which is a bounded invertible operator. Frames provide stable representations by means of series expansions. However to do this it is necessary to calculate the dual frame explicitly. Given a frame in a Hilbert space $\mathcal{H}$ with norm $\|.\|$ often it is more convenient and more efficient to employ an iterative reconstruction method:
Algorithm. ([31], Chapter 5) *Given a relaxation parameter $0 < \lambda < \frac{2}{B}$, set $\delta = max\{|1 - \lambda A|, |1 - \lambda B|\}$. Let $f_0 = 0$ and define recursively: $f_{n+1} = f_n + \lambda S(f - f_n)$. Then $\lim_{n\to\infty} f_n = f$, with a geometric rate of convergence, that is, $\|f - f_n\| \leq \delta^n \|f\|$.*

Some remarks on Basis and Banach space valued random variables

Let $(E, \|\,.\,\|)$ be a Banach space with a basis $\{f_n\}_{n\in\mathbb{N}}$ with biorthogonal - coordinate functionals [46] f_n^*. Random variables in E can be easily characterized in terms of the basis and the coordinate functionals. Let X be a E-valued random variable, then

$$X = \sum_{i=1}^{\infty} f_i^*(X) f_i \,.$$

Thus, for the sequences spaces c_0, and l^p $p \geq 1$, each random variable X is expressable as a sequence of scalar random variables $\{f_n^*(X)\}_n$. These are random variables, as a consequence of lemma 6.3.4 since the f_n^* are continuous and hence Borel measurable. A random variable may be constructed by the use of a Schauder basis. Let $\{Y_k\}_k$ be a sequence of scalar valued random variables such that,

$$\lim_{n\to\infty} \sum_{i=1}^{n} Y_i f_i$$

exists for each $\omega \in \Omega$ then by lemma 2.2.2, the limit is an E-valued random variable. We will use this fact in the forthcoming chapters 4 and 5. There is also a characterization of independence in terms of the coordinate functionals.

Lemma 2.3.1. *$(E, \|\,.\,\|)$ be a Banach space with a basis $\{f_n\}_{n\in\mathbb{N}}$ and coordinate functionals f_n^*. The random variables X, Y are independent $\iff$ the vectors $(f_1^*(X), \ldots, f_n^*(X))$ and $(f_1^*(Y), \ldots, f_n^*(Y))$ are independent, for each $n = 1, 2, \ldots$*

Proof. See [76] □

2.4 A review of Real and Harmonic Analysis

2.4.1 Some definitions

Remark: On the following, if $x \in \mathbb{C}^d$ $(d \geq 1)$ we will denote its usual norm by $|x|$ and $Supp(f) = \overline{\{x : f(x) = 0\}}$.

The Schwartz class of functions $\mathcal{S}(\mathbb{R}^d)$ is defined as the linear space of smooth functions rapidly decreasing at infinity, together with its derivatives, this means that $\phi \in \mathcal{S}(\mathbb{R}^d)$ whenever $\phi \in C^\infty\left(\mathbb{R}^d\right)$ and

$$\sup_{(x_1,\ldots x_d)\in\mathbb{R}^d} \prod_{i=1}^{d} |x_i|^{\alpha_i} \left| \frac{\partial}{\partial x_1^{\beta_1}} \cdots \frac{\partial}{\partial x_d^{\beta_d}} \phi(x_1, \ldots x_d) \right| < \infty \;\forall\; \alpha_j \, \beta_j \;\in\; \mathbb{N} \,,$$

endowed with its usual topology. We will denote $\mathcal{D}(\mathbb{R}^d)$ the space of functions which are in $C^\infty\left(\mathbb{R}^d\right)$ and have compact support. Both spaces are topological vector spaces (For more details see [75]), and their duals are denoted as: $\mathcal{S}'(\mathbb{R}^d)$ (*Tempered distributions*) and $\mathcal{D}'(\mathbb{R}^d)$ (*distributions*) respectively.
Clearly: $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$ and then $\mathcal{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d)$.

2.4.2 Fourier transforms.

The *Fourier Transform* $\widehat{f}$ of $f \in \mathcal{S}(\mathbb{R}^d)$ is defined as

$$\mathcal{F}(f)(\lambda) = \widehat{f}\,(\lambda) = \int_{\mathbb{R}^d} f\,(x) e^{-2\pi i \lambda . x} dx\,.$$

It is a known fact that $\widehat{f}$ also belongs to the space $\mathcal{S}(\mathbb{R}^d)$. $\mathcal{F}$ can be defined, as usual as a linear map $\mathcal{F} : L^1(\mathbb{R}^d) \longrightarrow L^\infty(\mathbb{R}^d)$ [b], or as an isometry on $L^2(\mathbb{R}^d)$ and by duality over the class of tempered distributions, that is $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \longrightarrow \mathcal{S}'(\mathbb{R}^d)$. For more references about Fourier transforms and series we refer the reader, for example, to [75] or [30].
Later, we will need a variant of the classic sampling theorem of Shannon, Nyquist and Kotelnikov for L^2 functions [c]:

Theorem 2.4.1. *(Variant of the Shannon-Kotelnikov theorem) If $f \in L^2(\mathbb{R}^d)$ is such that $Supp(f) \subset [-\lambda_o, \lambda_o]^d$ with $\lambda_o < \frac{1}{2}$. Then there exists $\theta \in \mathcal{S}(\mathbb{R}^d)$ such that*

$$\widehat{f}(\lambda) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) \theta(\lambda - k) \tag{2.4.1}$$

Proof. Let $\widetilde{f}(x) = \sum_{k \in \mathbb{Z}^d} f(x+k)$ be the periodization of f .The identification $\widetilde{f}$ with the torus verifies $\widetilde{f} \in L^2(\mathbb{T}^d) \subset L^1(\mathbb{T}^d)$, and, if $\widetilde{f} \sim \sum_{k \in \mathbb{Z}^d} a_k e^{-2\pi i x.k}$ then $\lim_{R\to\infty} \sum_{k \in D_R} a_k e^{-2\pi i x.k} = \widetilde{f}$ a.e. and in $L^1(\mathbb{T}^d)$ (and in L^2) norm for a suitable domain $D_R \in \mathbb{R}^d$. Now, we can take $\theta(x) \in \mathcal{S}(\mathbb{R}^d)$ such that [d]

$$\widehat{\theta}(\lambda) = \begin{cases} 1, |\lambda_i| < \lambda_0 \\ 0, |\lambda_i| \geq 1 - \lambda_0 \end{cases}$$

and define $S_R(x) = \widehat{\theta}(x) \sum_{k \in D_R} a_k e^{-2\pi i x.k}$ on the other hand $f = \widetilde{f}\,\widehat{\theta}$, then, is easy to show that $\lim_{R\to\infty} \|S_R - f\|_{L^1(\mathbb{R}^d)} = 0$. This implies $\lim_{R\to\infty} \sup_{\lambda \in \mathbb{R}^d} \left|\widehat{S_R}(\lambda) - \widehat{f}(\lambda)\right| = 0$, but (see [75]): $a_k = \widehat{f}(k)$, then

$$\widehat{S_R}(\lambda) = \sum_{k \in D_R} \widehat{f}(k) \theta(\lambda - k)\,.$$

Then (2.4.1) follows immediately from this. □

2.4.3 Some linear operators on L^p: Fractional integration.

In our examples we will sometimes use some fractional integration operators, let us review some of their properties.

[b] Indeed, the Fourier transform of an integrable function is continuous and tends to 0 as $|\lambda| \longrightarrow \infty$, by the Riemann-Lebesgue lemma.

[c] The original result can be found in some Harmonic Analysis books such as [30], or in the engineering literature related to signal analysis.

[d] In some literature more related to applications, such θ is called a low-pass or anti aliasing filter.

Let us consider the usual Laplacian of f [74]: $\Delta f = \sum_{j=1}^{d} \frac{\partial^2 f}{\partial x_j^2}$. Then, at least formally: $\widehat{\Delta f}(\lambda) = -(2\pi)^2 |\lambda|^2 \widehat{f}(\lambda)$. From this we could define the operators $(-\Delta)^{-\frac{\alpha}{2}}$ as:

$$(-\Delta)^{-\frac{\alpha}{2}} f = \mathcal{F}^{-1}(2\pi)^{-\alpha} |\,.\,|^{-\alpha} \mathcal{F} f. \tag{2.4.2}$$

The formal manipulations have a precise meaning [74]:

Definition 19. Let $0 < \alpha < d$. For $f \in \mathcal{S}(\mathbb{R}^d)$ we can define its *Riesz Potential*:

$$((-\Delta)^{-\frac{\alpha}{2}} f)(x) = \frac{1}{\gamma(\alpha)} \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}} dy \tag{2.4.3}$$

where $\gamma(\alpha) = \frac{\pi^{\frac{d}{2}} 2^\alpha \Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{d}{2} - \frac{\alpha}{2}\right)}$.

This linear operator has the following properties [74]:

Proposition 2.4.1. *Let $0 < \alpha < d$. Then: (a) The Fourier Transform of $|x|^{-d+\alpha}$ is $\gamma(\alpha)(2\pi)^{-\alpha} |\lambda|^{-\alpha}$ in the sense :*

$$\int_{\mathbb{R}^d} |x|^{-d+\alpha} \varphi(x) dx = \int_{\mathbb{R}^d} \gamma(\alpha)(2\pi)^{-\alpha} |\lambda|^{-\alpha} \widehat{\varphi}(\lambda) d\lambda$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$.
(b) The Fourier Transform of $((-\Delta)^{-\frac{\alpha}{2}} f)(x)$ is $(2\pi)^{-\alpha} |\lambda|^{-\alpha} \widehat{f}(\lambda)$ in the sense:

$$\int_{\mathbb{R}^d} ((-\Delta)^{-\frac{\alpha}{2}} f)(x) g(x) dx = \int_{\mathbb{R}^d} \widehat{f}(\lambda)(2\pi)^{-\alpha} |\lambda|^{-\alpha} \widehat{g}(\lambda) d\lambda$$

for all $f, g \in \mathcal{S}(\mathbb{R}^d)$.

It is easy to check that $\forall f \in \mathcal{S}(\mathbb{R}^d)$: If $\alpha + \beta < d$ then $(-\Delta)^{-\frac{\alpha}{2}}((-\Delta)^{-\frac{\beta}{2}} f) = (-\Delta)^{-\frac{(\alpha+\beta)}{2}}(f)$; and $\Delta((-\Delta)^{-\frac{\alpha}{2}} f) = (-\Delta)^{1-\frac{\alpha}{2}}(f)$.

We recall the following bound for these operators acting in $L^p(\mathbb{R}^d)$ [30], [74].

Theorem 2.4.2. *(Hardy,Littlewood and Sobolev) Let $0 < \alpha < d$, $1 \le p < q < \infty$ and $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$ then:*
(a) $\forall f \in L^p(\mathbb{R}^d)$, the integral that defines $(-\Delta)^{-\frac{\alpha}{2}} f$ converges a.e.
(b)If $p > 1$ then

$$\left\| (-\Delta)^{-\frac{\alpha}{2}} f \right\|_{L^q} \le C_{pq} \|f\|_{L^p} . \tag{2.4.4}$$

Remark. These operators are the inverses of the (positive) fractional powers of the Laplacian operator. On the class $\mathcal{S}(\mathbb{R}^d)$, $(-\Delta)^{\frac{\alpha}{2}}$ is given by

$$-(-\Delta)^{\frac{\alpha}{2}} f(x) = c \int_{\mathbb{R}^d} f(y) - f(x) - \frac{\nabla f(x).(y-x)}{1+|y-x|^2} \frac{dy}{|y-x|^{d+\alpha}} . \tag{2.4.5}$$

This expression follows from [74] section 6.10.
Now, introduce another fractional integration operator defined formally as:

$$(I-\Delta)^{s/2} f = \mathcal{F}^{-1}(1+|\,.|^2)^{s/2}\mathcal{F}f . \tag{2.4.6}$$

Theorem 2.4.3. *[30] If $s<0$ and $p\geq 1$, $(I-\Delta)^{s/2} : L^p(\mathbb{R}^d) \longrightarrow L^p(\mathbb{R}^d)$ defines a continuous linear operator i.e. there exists $C_p>0$ such that*

$$\left\|(I-\Delta)^{\frac{s}{2}} f\right\|_{L^p} \leq C_p \|f\|_{L^p} .$$

Fractional integral operators, as the Riesz integral operators can also be defined over a finite interval. In this case, they have similar properties to their $\mathbb{R}^d$ counterparts. This operators, for suitable parameters maps the L^p classes into the Lispschitz spaces Λ_β. For further references about this, see for example Zygmund's book ([85], chapter 12). We will use these properties briefly in chapter 4.

2.4.4 Fourier transform of measures. Applications to Probability Theory: Characteristic functions and stable distributions.

Let μ be a finite Borel measure on $\mathbb{R}^d$, We define its Fourier transform as:

$$\hat{\mu}(\lambda) = \int_{\mathbb{R}^d} e^{ix.\lambda} d\mu(x) .$$

This is the usual convention in Probability theory, this definition differs from the previous one by a conjugation operation and a multiplicative constant. This carries no problem. Let X be a random variable, then we define its characteristic function as $\phi_X(\lambda) = \mathbb{E}(e^{i\lambda X})$. Then, this can be written as

$$\phi_X(\lambda) = \int_{\mathbb{R}} e^{ix.\lambda} d\mu_X(x) .$$

Then the characteristic function is the Fourier transform of the measure μ_X. The definition can be extended for random vectors. One important fact about the characteristic function is that it defines uniquely a distribution function. For further references see for example [16].
Now, let us define an important class of distributions, we say that a random variable is symmetric α-*stable* [71], if for some $0<\alpha\leq 2$, and some $\sigma_X>0$:

$$\phi_X(\lambda) = e^{-\sigma_X^\alpha|\lambda|^\alpha} .$$

The parameters α and σ_X are unique. In this case we write $X \sim S_\alpha(\sigma_X, 0, 0)$. This definition can be extended to the non-symmetric case by introducing some additional parameters. Note that when $\alpha=2$ this corresponds to the gaussian case.

Some properties of stable distributions.

Let us list some properties of stable distributions, generally they can be obtained using the properties of characteristic functions. For more references see [71].

1) If $X_1, \ldots, X_n$ are independent and $X_i \sim S_\alpha(\sigma_{X_i}, 0, 0)$, then $\sum_{i=1}^{n} X_i \sim S_\alpha(\sigma', 0, 0)$, with $\sigma' = \|(\sigma_{X_i})_i\|_{l^\alpha}$.
2) Let $\alpha < 2$. If $X \sim S_\alpha(\sigma_X, 0, 0)$ then, if $0 < p < \alpha$: $(\mathbb{E}|X|^p)^{1/p} = C_p\, \sigma_X$. Where ${C_p}^p = \mathbb{E}|Y|^p$, with $Y \sim S_\alpha(1, 0, 0)$.
3) Let $\alpha < 2$, if $X \sim S_\alpha(\sigma_X, 0, 0)$, then for every $0 < p < \alpha$: $X \in L^p(\Omega, \mathcal{F}, \mathbf{P})$. But $X \notin L^\alpha(\Omega, \mathcal{F}, \mathbf{P})$.
4) Let $\alpha < 2$, if $X \sim S_\alpha(\sigma_X, 0, 0)$, then there exists positive constants x_0, A and B, such that:

$$A\, x^{-\alpha} \leq \mathbf{P}(|X| > x) \leq B\, x^{-\alpha}\,,$$

for all $x > x_0$.
5) Let $\alpha < 2$, if $X \sim S_\alpha(\sigma_X, 0, 0)$, then there exists positive constants x_0, A and B, such that:

$$A \log^+ \frac{x}{y} \leq \mathbb{E}|X|^\alpha \mathbf{1}_{\{y \leq |X| \leq x\}} \leq B \log^+ \frac{x}{y}\,,$$

for all $y, x > x_0$.
6) Let $\alpha < 2$ and $p > \alpha$, if $X \sim S_\alpha(\sigma_X, 0, 0)$, then there exists positive constants x_0, A and B, such that:

$$A\, x^{p-\alpha} \leq \mathbb{E}|X|^p \mathbf{1}_{\{|X| \leq x\}} \leq B\, x^{p-\alpha}\,,$$

for all $x > x_0$.
For further references about stable random variables and processes, see [71].

2.4.5 Wide sense stationary random processes

Now, let us review some basic facts about this important class of random processes. This brief introduction is mainly motivational, since we shall deal later, in chapter 5 with a more general, but closely related, class of random processes. Also, in chapter 6 we shall introduce a variant of this definition. This results are mainly from ([69], chapter I). Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, let $\mathscr{Y} = \{Y_x\}_{x \in \mathbb{A}}$ be a wide sense stationary random process, where $\mathbb{A} = \mathbb{R}^n$ or $\mathbb{Z}^n$. By this we mean a family of random variables in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ stationary correlated in the index x, i.e. $R(t, s) = \mathbb{E}(Y_t \overline{Y}_s) = R(t - s)$, for all t, s. We will also assume that $\mathbb{E}Y_x = 0$ for all x, and that the process is measurable in the wide sense, i.e. for any $Z \in \overline{\text{span}}\{Y_x\}_x$ the numerical valued function $\mathbb{E}(Y_x Z)$ is Lebesgue measurable. In the following, $\mathbb{G} = \mathbb{R}^n$ if $\mathbb{A} = \mathbb{R}^n$, or $\mathbb{G} = [0, 1)^n$ if $\mathbb{A} = \mathbb{Z}^n$.

Harmonic analysis of stationary processes:

Every stationary wide sense stationary random process $\mathscr{Y} = \{Y_x\}_{x \in \mathbb{A}}$ admits a spectral representation:

$$Y_x = \int_{\mathbb{G}} e^{i2\pi\lambda.x} d\Phi$$

in the form of an stochastic integral with respect to a random spectral measure Φ. Moreover for each $x \in \mathbb{R}^n$ or $\mathbb{Z}^n$, Y_x can be written as the result of the action of the (unitary) *shift operator* T on Y_0:

$$Y_x = T^x Y_0 \quad \text{where by Stone's spectral theorem: } T^x = \int_{\mathbb{G}} e^{i2\pi\lambda.x} dE(\lambda)\,, \tag{2.4.7}$$

where the $E(\lambda)$'s are orthogonal projection operators over $H(\mathscr{Y})$, the closed linear span of $\mathscr{Y} = \{Y_x\}_{x\in\mathbb{A}}$.
Set $\mu(A) = \mathbb{E}(|\Phi(A)|^2)$, then μ is a finite Borel measure called the *spectral measure* of the process $\mathscr{Y}$. On the other hand the spectral measures μ are also related by the following Fourier transform pairing

$$\mathbb{E}(Y_x\overline{Y}_0) = \int_{\mathbb{G}} e^{i2\pi\lambda.x} d\mu \tag{2.4.8}$$

In the case that μ is absolutely continuous with respect to the Lebesgue measure, the Radon-Nykodym derivative of μ, $\phi(\lambda)d\lambda = d\mu$, ϕ is called the *spectral density* of the process. In this case the correlation function is the ordinary Fourier transform of the spectral density.

2.5 Miscelanea: Additional comments, bibliographical and historical notes

2.5.1 On Remark 2.1.1. Almost sure convergence is not a topological notion.

Let us suppose that almost sure convergence is a topological notion and see what it would happen in this case. Let (X, Σ, μ) be a non trivial measure space, such that convergence in probability is not equivalent to a.e. convergence, for example $X = [0,1]$, $\Sigma = \mathscr{B}[0,1]$ the Borel sigma algebra and μ the Lebesgue measure. Then there exists a sequence of measurable functions $\{f_n\}_n$ and a measurable function f, such that $f_n \longrightarrow f$ in measure but not a.e. If a.e. convergence is a topological notion then there should exists a neighbourhood of f: $U(f)$ and infinitely many n's, say $\{n_k\}_k$, such that $\forall\, k \,:\, f_{n_k} \notin U(f)$. But the subsequence $f_{n_k} \longrightarrow f$ in measure, and by Riesz´s theorem there exists a subsubsequence $\{f_{n_{k_j}}\}_j$ such that $f_{n_{k_j}} \longrightarrow f$ a.e. and then there must exists j_0 such that $f_{n_{k_j}} \in U(f)$, for all $j \geq j_0$, which is a contradiction.

2.5.2 Martingales.

For more advance results about martingales we refer the reader to [32] and [23]. There is possible to find several classical results on convergence and their generalizations. Also, several generalizations of the Khinchine's inequalities are given there for martingale differences.

2.5.3 Construction of probability measures.

We will always assume that we have an underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Under the hypothesis that we are assuming in this work, the existence of a probability measure on the spaces where we are going to use is granted. However, the construction of a probability measure over a linear topological space is not a trivial topic. For further references see [27] and [26].

Chapter 3

Random Series in Banach Spaces

"Nanos gigantium humeris insidentes".

"[Dwarfs] Standing on the shoulders of giants".

Metaphor attributed to Bernard of Chartres, also used by Isaac Newton. Inscription bear by the British two pound coin on its edge.

3.1 Introduction.

In this chapter we study random series with values in Banach spaces. We study the a.s. convergence and convergence in the p-mean of these series. We begin with a review of some existing inequalities which will be useful during the development of this work. This inequalities are useful on its own. Then we will use this results to state general conditions for the convergence of random series in Banach spaces.
This chapter mainly contains some existing results on the convergence of Banach space valued random series which will be used in the forthcoming chapters. So it may be regarded also as a summary of them. The proofs have been included in order to make the exposition as self contained as possible, since no unified source of reference was found. Also some auxiliary results will be presented. Many of these results, in a form or another, are spread in different textbooks and research papers. The proofs, of existing results, have been chosen or adapted to be as close as possible to our necessities. The exposition of the proofs, also have three purposes: Some of them may serve as a review of the proof of the original scalar valued case. Also, at one point one may get the intuition of how some of these results have been developed from their scalar valued cousins. Finally, directly or indirectly, many of the ideas behind the results of this chapter will appear further. And sometimes a complete proof of a previous lemma will later "be worth more than a thousand words", since the idea of the forthcoming results will become clearer looking at the previous ones. The main references are [35], [39], [28], [42], [4], [40] and [51].

3.1.1 Basic Inequalities.

This section contains general basic inequalities for sums of independent random variables with values in a separable Banach space $(E, \|.\|)$. These inequalities will be used through out the reminder of this work. In the forthcoming we will denote $L^p(E) = L^p(\Omega, \mathcal{F}, \mathbf{P}, E)$ to the space of all E valued random variables X, such that $\mathbb{E}\,\|X\|^p < \infty$.

Proposition 3.1.1. *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of non negative random variables and let c_i, $i = 0, 1, \dots,$ be a sequence of non negative numbers. If for each $t \in \mathbb{R}^+$,*

$$\sum_{i=1}^{\infty} c_i \mathbf{P}(X_i > t) \le c_0 \mathbf{P}(X_0 > t)\,,$$

then, for any non decreasing function $\phi : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ with $\phi(0) = 0$,

$$\sum_{i=1}^{\infty} c_i \mathbb{E}\phi(X_i) \le c_0 \mathbb{E}\phi(X_0)\,,$$

moreover, if $\sum_{i=1}^{\infty} c_i \le c_0$ then the above is true for all non decreasing functions $\phi : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$.

The following simple result will be very useful:

Lemma 3.1.1. *Let $0 < \lambda < 1$, $p > 1$, $X \in L^p(\Omega, \mathcal{F}, \mathbf{P})$ and $X \ge 0$ a.s. then*

$$\mathbf{P}\left(X \ge \lambda \mathbb{E}(X)\right) \ge (1-\lambda)^q \frac{(\mathbb{E}(X))^q}{(\mathbb{E}(X^p))^{\frac{q}{p}}}\,,$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Define $Y = X\mathbf{1}_{\{X > \lambda\mathbb{E}(X)\}}$. By Hölder's inequality:

$$(\mathbb{E}(Y))^p \le \mathbb{E}(Y^p)(\mathbf{P}(Y \ne 0))^{p-1} \le \mathbb{E}(Y^p)\left(\mathbf{P}(X > \lambda\mathbb{E}(X))\right)^{p-1} .$$

Moreover $\mathbb{E}(X) \le \mathbb{E}(Y) + \lambda\mathbb{E}(X)$, then as $\mathbb{E}(X^p) \ge \mathbb{E}(Y^p)$,

$$(1-\lambda)^p(\mathbb{E}(X))^p \le \mathbb{E}(X^p)\left(\mathbf{P}(X > \lambda\mathbb{E}(X))\right)^{p-1} .$$

The result follows noting that $\frac{p}{p-1} = q$. □

This type of condition appear also in [40] [41] dealing with random series. With this, if we suppose now that $E = H$ is a separable Hilbert space, we can prove the following inequality which plays a fundamental role in the work of Paley and Zygmund on random Fourier series:

Proposition 3.1.2. *Let $X_1, \dots, X_n \in L^4(\Omega, \mathcal{F}, \mathbf{P}, E)$, $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(\|X_k\|^4) \le CVar^2(X_k)$, for $k = 1, \dots, n$. Then $\exists\, C_0(\lambda) > 0$ independent of n, such that:*

$$\mathbf{P}\left(\left\|\sum_{i=1}^{n} X_i\right\|^2 > \lambda \sum_{i=1}^{n} Var(X_i)\right) > C_0$$

Proof. Let $U = \left\| \sum_{i=1}^{n} X_i \right\|^2$, then $\mathbb{E}(U) = \sum_{i=1}^{n} Var(X_i)$, and

$$\mathbb{E}(U^2) = \sum_{i\,j\,k\,l} \mathbb{E}(\langle X_i, X_j \rangle \langle X_k X_l \rangle)$$

If one of the indexes is different from the other: $\mathbb{E}(\langle X_i, X_j \rangle \langle X_k X_l \rangle) = 0$. Thus

$$\mathbb{E}(U^2) \leq \sum_{k=1}^{n} \mathbb{E}(\|X_n\|^4) + 6 \sum_{1 \leq k < l \leq n} Var(X_n) Var(X_m)$$

$$C \sum_{i=1}^{n} Var^2(X_i) + 6 \sum_{1 \leq k < l \leq n} Var(X_n) Var(X_m) \leq \sup(3, C) \left(\sum_{i=1}^{n} Var(X_i) \right)^2 .$$

And then apply the previous lemma 3.1.1 with $p = 2$. □

Very similarly to the real valued case, one can proves:

Proposition 3.1.3. *If $\{X_i\}_{i=1}^{n}$ are independent random variables in E, then for each $\lambda \leq 0$,*

$$\mathbf{P}\left(\max_{1 \leq k \leq n} \|S_k\| > \lambda \right) \leq 3 \max_{1 \leq k \leq n} \mathbf{P}\left(\|S_k\| > \frac{\lambda}{3} \right)$$

and if the X_i´s are symmetric then

$$\mathbf{P}\left(\max_{1 \leq k \leq n} \|S_k\| > \lambda \right) \leq 2\mathbf{P}\left(\|S_n\| > \lambda \right)$$

Proof. for fixed $t, s \geq 0$ and $k = 1, 2, \ldots, n$, define:

$$A_k = \{\|S_1\| \leq s + t; \ \ldots, \|S_{k-1}\| \leq s + t, \ \|S_k\| > s + t\} \ ,$$

then the A_k's are disjoint and $A = \left\{ \max_{1 \leq k \leq n} \|S_k\| > t + s \right\} = \bigcup_{k=1}^{n} A_k$. And for each $i = 1, \ldots, n$, the A_i's are independent of $X_{i+1}, \ldots, X_n$, and

$$A_i \bigcap \{\|S_n - S_i\| \leq s\} \subseteq A_i \bigcap \{\|S_n\| > t\} \ ;$$

so that

$$\mathbf{P}\left(\|S_n\| > t \right) \geq \sum_{i=1}^{n} \mathbf{P}(A_i) \mathbf{P}\left(\|S_n - S_i\| \leq s \right) \geq \min_{1 \leq i \leq n} \mathbf{P}\left(\|S_n - S_i\| \leq s \right) \mathbf{P}\left(\max_{1 \leq k \leq n} \|S_k\| > t + s \right) ,$$

then

$$\mathbf{P}\left(\max_{1 \leq k \leq n} \|S_k\| > t + s \right) \leq \frac{\mathbf{P}\left(\|S_n\| > t \right)}{1 - \max_{1 \leq i \leq n} \mathbf{P}\left(\|S_n - S_i\| > s \right)} ,$$

in particular, if $\max_{1\le i\le n} \mathbf{P}\left(\|S_i\| > \frac{t}{3}\right) < \frac{1}{3}$ then,

$$\mathbf{P}\left(\max_{1\le k\le n}\|S_k\| > t\right) \le \frac{\mathbf{P}\left(\|S_n\| > \frac{t}{3}\right)}{1-\max_{1\le i\le n}\mathbf{P}\left(\|S_n - S_i\| > \frac{2t}{3}\right)}$$

$$\le \frac{\max_{1\le i\le n}\mathbf{P}\left(\|S_i\| > \frac{t}{3}\right)}{1-2\max_{1\le i\le n}\mathbf{P}\left(\|S_i\| > \frac{t}{3}\right)} \le 3\max_{1\le i\le n}\mathbf{P}\left(\|S_i\| > \frac{t}{3}\right) ,$$

which proves the first part of the theorem since, in the case when

$$\max_{1\le i\le n}\mathbf{P}\left(\|S_i\| > \frac{t}{3}\right) \ge \frac{1}{3} ,$$

is automatically satisfied. To prove the second statement of the theorem, denoting $A_i = \{\|S_j\| \le t \,\forall\, j < i\,, \|S_i\| > t\}$, for $i = 1, \ldots, n$, and observe that

$$A_i \subseteq (A_i \cap \{\|S_n\| > t\}) \bigcup (A_i \cap \{\|2S_i - S_n\| > t\}) ,$$

so that

$$\mathbf{P}(A_i) \le \mathbf{P}\left(A_i \cap \{\|S_n\| > t\}\right) + \mathbf{P}\left(A_i \cap \{\|2S_i - S_n\| > t\}\right) .$$

Since $X_1, \ldots, X_n$ are assumed to be symmetric and independent, the last two probabilities are equal. Then, $\mathbf{P}(A_i) \le 2\mathbf{P}(A_i \cap \{\|S_n\| > t\})$, and a summation over $i = 1, \ldots, n$ gives

$$\mathbf{P}(\max_{1\le k\le n}\|S_k\| > t) \le 2\mathbf{P}(\|S_n\| > t) .$$

□

Proposition 3.1.4. *Let $\{X_i\}_{i=1}^n$ be independent random variables with values in E, then for any $s, t, u \ge 0$ we have*

$$\mathbf{P}\left(\max_{1\le k\le n}\|S_k\| > s+t+u\right)$$

$$\le \mathbf{P}\left(\max_{1\le k\le n}\|X_k\| > u\right) + \mathbf{P}\left(\max_{1\le k\le n}\|S_k\| > t\right)\mathbf{P}\left(\max_{k,l\le n}\left\|\sum_{i=k}^{l} X_i\right\| > s\right) ,$$

and if the X_i's are symmetric then

$$\mathbf{P}\left(\max_{1\le k\le n}\|S_k\| > s+t+u\right)$$

$$\le \mathbf{P}\left(\max_{1\le k\le n}\|X_k\| > u\right) + 2\mathbf{P}\left(\max_{1\le k\le n}\|S_k\| > t\right)\mathbf{P}\left(\|S_n\| > s\right) .$$

Proof. Again, as in the previous, define

$$A_k = \{\|S_1\| \le t; \ldots, \|S_{k-1}\| \le t, \|S_k\| > t\} ,$$

then

$$\left\{\max_{1\le k\le n} \|S_k\| > s+t+u\right\} \bigcap \left\{\max_{1\le k\le n} \|X_k\| \le u\right\} \subseteq \bigcup_{i=1}^{n} A_i \cap \left\{\max_{i\le k\le n} \|S_k - S_i\| > s\right\}$$

The A_i's and $\left\{\max_{i\le k\le n} \|S_k - S_i\| > s\right\}$ are independent, then $\forall\, i = 1, \ldots, n$:

$$\mathbf{P}\left(A_i \cap \left\{\max_{i\le k\le n} \|S_k - S_i\| > s\right\}\right) \le \mathbf{P}(A_i) \max_{1\le j\le n} \mathbf{P}\left(\max_{j\le k\le n} \|S_k - S_j\| > s\right) ,$$

then summing over i:

$$\mathbf{P}\left(\max_{1\le k\le n} \|S_k\| > s+t+u\, , \max_{1\le k\le n} \|X_k\| \le u\right)$$

$$\le \mathbf{P}\left(\max_{1\le k\le n} \|S_k\| > t\right) \max_{1\le j\le n} \mathbf{P}\left(\max_{j\le k\le n} \|S_k - S_j\| > s\right) .$$

Now the proposition follows by estimating from above with

$$\mathbf{P}\left(\max_{k,l\le n} \left\|\sum_{i=k}^{l} X_i\right\| > s\right) .$$

And for the symmetric case, recalling proposition 3.1.3, it is estimated by $2\mathbf{P}(\|S_n\| > s)$. □

Some moment inequalities.

The following two results, first appeared in [35]. We will follow its main idea.

Lemma 3.1.2. *Let $\{X_j\}_{j=1}^n$ be independent E-valued random variables with $\mathbb{E}X_j = 0$. Then $\forall\, (a_1, \ldots, a_n) \in \mathbb{R}^n$:*

$$\left(\mathbb{E}\left\|\sum_{j=1}^{n} a_j X_j\right\|^p\right)^{\frac{1}{p}} \le 2 \max_{1\le j\le n} |a_j| \left(\mathbb{E}\left\|\sum_{j=1}^{n} X_j\right\|^p\right)^{\frac{1}{p}}$$

Proof. Let $K = \{x \in \mathbb{R}^n : \|x\|_\infty \le 1\}$ and let $f : K \longrightarrow \mathbb{R}_{\ge 0}$ be defined by $f(x) = \mathbb{E}\left\|\sum_{j=1}^{n} x_j X_j\right\|^p$. Then f is convex and continuous so by Caratheodory's theorem it attains its maximum for points of the form $x = \mp 1$, for such a point let $\Sigma_+ = \{j : x_j = 1\}$ and let $\Sigma_- = \{j : x_j = -1\}$. Then

$$f(x) = \left(\mathbb{E}\left\|\sum_{j\in\Sigma_+} X_j + \sum_{j\in\Sigma_-} X_j\right\|^p\right)^{\frac{1}{p}} \le \left(\mathbb{E}\left(\left\|\sum_{j\in\Sigma_+} X_j\right\| + \left\|\sum_{j\in\Sigma_-} X_j\right\|\right)^p\right)^{\frac{1}{p}}$$

$$\leq \left(\mathbb{E}\left\|\sum_{j\in\Sigma_+} X_j\right\|^p\right)^{\frac{1}{p}} + \left(\mathbb{E}\left\|\sum_{j\in\Sigma_-} X_j\right\|^p\right)^{\frac{1}{p}} \leq 2\left(\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p\right)^{\frac{1}{p}} \quad \text{(By lemma 2.2.8)}$$

Now, given $a = (a_1, \ldots, a_n) \in \mathbb{R}^n \setminus \{0\}$, take $x = \frac{a}{\|a\|_\infty}$, then,

$$f(x) = f\left(\frac{a}{\|a\|_\infty}\right) = \frac{1}{\|a\|_\infty}\left(\mathbb{E}\left\|\sum_{j=1}^n a_j X_j\right\|^p\right)^{\frac{1}{p}} \leq 2\left(\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p\right)^{\frac{1}{p}}$$

□

Lemma 3.1.3. *Let $\{X_j\}_{j=1}^n$ be independent E-valued random variables with $\mathbb{E}X_j = 0$ and $\mathbb{E}\|X_j\|^p < \infty$ and let $\{a_j\}_{j=1}^n$ be independent of $\{X_j\}_j$, real valued random variables in $L^p(\Omega, \mathcal{F}, \mathbf{P})$, then:*

$$2^{-p}\mathbb{E}\left(\min_{1\leq j\leq n}|a_j|^p\right)\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^n a_j X_j\right\|^p \leq 2^p\mathbb{E}\left(\max_{1\leq j\leq n}|a_j|^p\right)\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p$$

Proof. (Sketch.) Denote $\mathbb{E}_X$ the expected value with respect to $\{X_j\}_j$, and $\mathbb{E}_a$ to the expected value with respect to $\{a_j\}_j$. We can prove the right hand side of the inequality, using independence and the previous lemma:

$$\mathbb{E}\left\|\sum_{j=1}^n a_j X_j\right\|^p = \mathbb{E}_a\mathbb{E}_X\left\|\sum_{j=1}^n a_j X_j\right\|^p$$

$$\leq \mathbb{E}_a\left(2^p\max_{1\leq j\leq n}|a_j|\mathbb{E}_X\left\|\sum_{j=1}^n X_j\right\|^p\right) = 2^p\mathbb{E}\left(\max_{1\leq j\leq n}|a_j|^p\right)\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p.$$

The other inequality is obtained writing, $Y_j = a_j X_j$ and $b_j = \frac{\min|a_j|}{a_j}\mathbf{1}_{\mathbb{R}\setminus\{0\}}(a_j)$. By the independence,

$$\mathbb{E}\left\|\sum_{j=1}^n b_j Y_j\right\|^p = \mathbb{E}\left(\min_{1\leq j\leq n}|a_j|^p\right)\mathbb{E}\left\|\sum_{j=1}^n X_j\right\|^p.$$

The result follows from this.

□

Proposition 3.1.5. *Let $p > 0$, let $\{X_i\}_{i=1}^n$ be independent random variables with values in E then, for any $t \geq 0$:*

$$\mathbb{E}\left(\max_{1\leq k\leq n}\|S_k\|\right)^p \leq \frac{\mathbb{E}\left(\max_{1\leq k\leq n}\|X_k\|\right)^p + t^p}{3^{-p} - \mathbf{P}\left(\max_{k,l\leq n}\left\|\sum_{i=k}^l X_i\right\| > t\right)}.$$

And if the X_i's are symmetric then and independent then,

$$\mathbb{E}\left(\max_{1\le k\le n}\|S_k\|\right)^p \le \frac{\mathbb{E}\left(\max_{1\le k\le n}\|X_k\|\right)^p + t^p}{3^{-p} - 2\mathbf{P}\left(\|S_n\| > t\right)}.$$

Proof. write $s = t = u$ in prop. 3.1.4. And let us call $X = \max_{1\le k\le n}\|S_k\|$, $Y = \max_{1\le k\le n}\|X_k\|$ and $Z = \max_{k,l\le n}\left\|\sum_{i=k}^{l} X_i\right\|$. Then

$$\mathbf{P}(X > 3s) \le \mathbf{P}(Y > s) + \mathbf{P}(X > s)\mathbf{P}(Z > s).$$

Then for all $t \ge 0$:

$$3^{-p}\,\mathbb{E}(X^p) = \int_0^\infty p\,s^{p-1}\mathbf{P}(X > 3s)ds$$

$$\le \int_0^\infty p\,s^{p-1}(\mathbf{P}(Y > s) + \mathbf{P}(X > s)\mathbf{P}(Z > s))ds$$

$$\le \mathbb{E}(Y^p) + \mathbb{E}(X^p)\mathbf{P}(Z > t) + \int_0^t p\,s^{p-1}ds = \mathbb{E}(Y^p) + \mathbb{E}(X^p)\mathbf{P}(Z > t) + t^p.$$

If the X_i's are symmmetric then, we can proceed as above and using the last part of proposition 3.1.4 we get the last inequality. □

For a variant on the proof of this result see [4].

Proposition 3.1.6. *Let $\{X_i\}_{i=1}^n$ be independent random variables with values in E, then, for any $t \ge 0$:*

$$\sum_{i=1}^n \mathbf{P}(\|X_i\| > t) \le \frac{\mathbf{P}\left(\max_{1\le k\le n}\|X_k\| > t\right)}{\mathbf{P}\left(\max_{1\le k\le n}\|X_k\| \le t\right)}$$

Proof.

$$\mathbf{P}\left(\max_{1\le k\le n}\|X_k\| > t\right) = 1 - \prod_{k=1}^n(1 - \mathbf{P}(\|X_k\| > t)) \ge 1 - e^{-\sum_{k=1}^n \mathbf{P}(\|X_k\|>t)}, \tag{3.1.1}$$

and since $1 - e^{-x} \ge xe^{-x}$, then

$$\mathbf{P}\left(\max_{1\le k\le n}\|X_k\| > t\right) \ge \sum_{k=1}^n \mathbf{P}(\|X_k\| > t)e^{-\sum_{k=1}^n \mathbf{P}(\|X_k\|>t)},$$

on the other hand, taking in account eq. 3.1.1, the last term is estimated from below by $1 - \mathbf{P}(\max_{1\le k\le n}\|X_k\| > t)$.

□

Finally, we have Rosenthal´s inequality:

Theorem 3.1.1. *Let $2 \leq p < \infty$. Then there exists constants $K_p, K'_p > 0$ so that for any sequence of independent real valued random variables $\{X_k\}_{k=1\ldots n} \subset L^p(\Omega, \mathcal{F}, \mathbf{P})$, and $\mathbb{E}X_j = 0$ for all $j = 1 \ldots n$. Then for all $n \in \mathbb{N}$:*

$$K'_p \max\left\{\left(\sum_{j=1}^n \mathbb{E}|X_j|^p\right)^{\frac{1}{p}}, \left(\sum_{j=1}^n \mathbb{E}|X_j|^2\right)^{\frac{1}{2}}\right\} \leq \left(\mathbb{E}\left|\sum_{j=1}^n X_j\right|^p\right)^{\frac{1}{p}} \tag{3.1.2}$$

$$\leq K_p \max\left\{\left(\sum_{j=1}^n \mathbb{E}|X_j|^p\right)^{\frac{1}{p}}, \left(\sum_{j=1}^n \mathbb{E}|X_j|^2\right)^{\frac{1}{2}}\right\}$$

Proof. First, let us prove the right hand side inequality. Take $q = \frac{p}{2}$, and write $Y_i = |X_i|^2$, then for $i = 1, \ldots, n$:

$$\left(\sum_{j=1}^n Y_j\right)^{q-1} = \left(\sum_{j\neq i} Y_j + Y_i\right)^{q-1} \leq 2^{q-1}\left(\left(\sum_{j\neq i} Y_j\right)^{q-1} + (Y_i)^{q-1}\right),$$

now, by the independence:

$$\mathbb{E}\left(\left(\sum_{j=1}^n Y_j\right)^{q-1} Y_i\right) \leq 2^{q-1}\left(\mathbb{E}\left(\sum_{j\neq i} Y_j\right)^{q-1} \mathbb{E}Y_i + \mathbb{E}Y_i{}^q\right).$$

$$\leq 2^{q-1}\left(\mathbb{E}\left(\sum_{j=1}^n Y_j\right)^{q-1} \mathbb{E}Y_i + \mathbb{E}Y_i{}^q\right).$$

Summing over i:

$$\sum_{i=1}^n \mathbb{E}\left(\left(\sum_{j=1}^n Y_j\right)^{q-1} Y_i\right) \leq 2^{q-1}\left(\mathbb{E}\left(\sum_{j=1}^n Y_j\right)^{q-1} \sum_{i=1}^n \mathbb{E}Y_i + \sum_{i=1}^n \mathbb{E}Y_i{}^q\right).$$

By Hölder´s inequality $\mathbb{E}\left(\sum_{j=1}^n Y_j\right)^{q-1} \leq \left(\mathbb{E}\left(\sum_{j=1}^n Y_j\right)^q\right)^{\frac{q-1}{q}}$, then,

$$\mathbb{E}\left(\sum_{i=1}^n Y_i\right)^q \leq 2^{q-1}\left(\left(\mathbb{E}\left(\sum_{j=1}^n Y_j\right)^q\right)^{\frac{q-1}{q}} \sum_{i=1}^n \mathbb{E}Y_i + \sum_{i=1}^n \mathbb{E}Y_i{}^q\right).$$

Write $\sum_{j=1}^n \mathbb{E}Y_j^q = \left(\sum_{j=1}^n \mathbb{E}Y_j^q\right)^{1-\frac{1}{q}+\frac{1}{q}} \leq \left(\sum_{j=1}^n \mathbb{E}Y_j^q\right)^{\frac{1}{q}}\left(\mathbb{E}\left(\sum_{i=1}^n Y_i\right)^q\right)^{1-\frac{1}{q}}$, since $\sum_i Y_i^q \leq \left(\sum_i Y_i\right)^q$.
Then,

$$\mathbb{E}\left(\sum_{i=1}^{n} Y_i\right)^q \leq 2^{q-1}\left(\left(\mathbb{E}\left(\sum_{j=1}^{n} Y_j\right)^q\right)^{\frac{q-1}{q}} \sum_{i=1}^{n} \mathbb{E}Y_i\right)$$
$$+2^{q-1}\left(\left(\sum_{j=1}^{n} \mathbb{E}Y_j^q\right)^{\frac{1}{q}}\left(\mathbb{E}\left(\sum_{i=1}^{n} Y_i\right)^q\right)^{1-\frac{1}{q}}\right).$$

Thus,

$$\left(\mathbb{E}\left(\sum_{i=1}^{n} Y_i\right)^q\right)^{\frac{1}{q}} \leq 2^{q-1}\left(\left(\sum_{j=1}^{n} \mathbb{E}Y_j^q\right)^{\frac{1}{q}} + \sum_{i=1}^{n} \mathbb{E}Y_i\right)$$

Recalling that $q = \frac{p}{2}$ and $|X_j|^2 = Y_j$. since if $a, b \geq 0$, $(a+b)^{\frac{1}{2}} \leq a^{\frac{1}{2}} + b^{\frac{1}{2}}$:

$$\left(\mathbb{E}\left(\sum_{i=1}^{n} |X_i|^2\right)^{\frac{p}{2}}\right)^{\frac{1}{p}} \leq 2^{\frac{p}{4}-\frac{1}{2}}\left(\left(\sum_{i=1}^{n} \mathbb{E}|X_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} \mathbb{E}|X_i|^2\right)^{\frac{1}{2}}\right),$$

but by Khinchine's theorem 2.1.11 we have that,

$$\mathbb{E}\left|\sum_{j=1}^{n} X_j\right|^p \leq B_p \mathbb{E}\left(\sum_{j=1}^{n} |X_j|^2\right)^{\frac{p}{2}},$$

combining this, we get:

$$\left(\mathbb{E}\left|\sum_{j=1}^{n} X_j\right|^p\right)^{\frac{1}{p}} \leq K_p\left(\left(\sum_{i=1}^{n} \mathbb{E}|X_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} \mathbb{E}|X_i|^2\right)^{\frac{1}{2}}\right),$$

and the inequality follows immediately from this.
For the left hand side inequality, first note that by Hölders's inequality

$$\mathbb{E}\left(\sum_{j=1}^{n} |X_j|^2\right)^{\frac{p}{2}} \geq \left(\sum_{i=1}^{n} \mathbb{E}|X_i|^2\right)^{\frac{p}{2}},$$

on the other hand, since $\left(\sum_i |X_i|^2\right)^{\frac{p}{2}} \geq \sum_i |X_i|^p$. Then,

$$2\left(\mathbb{E}\left(\sum_{i=1}^{n} |X_i|^2\right)^{\frac{p}{2}}\right)^{\frac{1}{p}} \geq \left(\sum_{i=1}^{n} \mathbb{E}|X_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} \mathbb{E}|X_i|^2\right)^{\frac{1}{2}}.$$

Now, the desired inequality follows from applying again theorem 2.1.11. □

Here we gave a proof using Khinchine's inequalities. The result can also be derived using lemma 3.1.3. The constants, however, may differ.

3.1.2 Convergence of series.

We begin this section with an auxiliary result on the convergence of series of real random variables. Let $c = \frac{1}{3^p}$. And for $0 < \delta < \frac{c}{8}$ define: given X a $\mathbb{R}$-valued random variable ,

$$\Psi(X) = X\mathbf{1}_{\{|X|\leq 1\}} + \text{sign}(X)\mathbf{1}_{\{|X|>1\}}$$

and for every subset of indexes J :

$$\triangle = \triangle((X_i)_{i\in J}) = \inf\left\{ s : \sum_{i\in J} \Psi^2\left(\frac{X_i}{s}\right) \leq \delta \right\} .$$

Lemma 3.1.4. *Let $X = \sum_{i=1}^{\infty} X_i$ be an a.s. convergent series of independent, symmetric, real valued random variables then:*
i)$\mathbf{P}(|X| > \triangle) \leq 2\delta$.
ii) $\mathbf{P}\left(|X| > \frac{\delta^{\frac{1}{2}}\triangle}{c}\right)$.
iii) $\exists\, C_1, C_2 > 0$ *depending only on c and δ, such that*

$$C_1\mathbb{E}|X|^p \leq \triangle^p + \sum_{i=1}^{\infty} \mathbb{E}|X_i|^p \mathbf{1}_{\{|X|>\triangle\}} \leq C_2\mathbb{E}|X|^p .$$

Proof. (i) Since $\triangle = \lim_{n\to\infty} \triangle((X_i)_{i=1}^n)$, it is enough to prove the result for finite sums. Denoting $\Phi((X_i)_{i=1}^n) = \mathbb{E}\left(\sum_{i=1}^n \Psi\left(X_i\right)\right)^2$ we have that, if $S_n = \sum_{i=1}^n X_i$ then

$$\mathbf{P}(|S_n| > 1) \leq \mathbf{P}\left(\left|\sum_{i=1}^n \Psi\left(X_i\right)\right| > 1\right) + \mathbf{P}\left(\max_{1\leq k\leq n} |X_k| > 1\right)$$

$$\leq \mathbb{E}\left(\sum_{i=1}^n \Psi\left(X_i\right)\right)^2 + \sum_{i=1}^n \mathbf{P}(|X_i| > 1) \leq 2\Phi((X_i)_{i=1}^n) ,$$

by Chevychev's inequality.
As $\triangle$ is an homogeneous function , then $\mathbf{P}(|S_n| > \triangle) < 2\delta$.
(ii) By prop. 3.1.5 with $p = 2$, and the symmetry of the X_i's we can write,

$$\Phi((X_i)_{i=1}^n) \leq \frac{\mathbb{E}\left(\max_{1\leq k\leq n} |\Psi(X_k)|\right)^2 + t^2}{3^{-2} - 2\mathbf{P}\left(\left|\sum_{i=1}^n \Psi(X_i)\right| > t\right)} \leq \frac{\mathbf{P}\left(\max_{1\leq k\leq n} |X_k| > t\right) + 2t^2}{3^{-2} - 2\mathbf{P}\left(\left|\sum_{i=1}^n \Psi(X_i)\right| > t\right)}$$

$$\leq \frac{2\mathbf{P}(|S_n| > t) + 2t^2}{3^{-2} - 4\mathbf{P}(|S_n| > t)}$$

then by proposition 2.1.2

$$\mathbf{P}\left(\left|\sum_{i=1}^{n}\Psi\left(X_i\right)\right|>t\right)\leq \mathbf{P}(|S_n|>t)$$

and from prop. 2.1.1: $\mathbf{P}\left(\max_{1\leq k\leq n}|X_k|>t\right)\leq \mathbf{P}(|S_n|>t)$. So if $\Phi((X_i)_{i=1}^n)=\delta \Longrightarrow \mathbf{P}(|S_n|> t)\geq (\frac{\delta}{18}-t^2)/(2\delta+1)$, putting $t=\frac{\delta^{1/2}}{6}$, we get that $\mathbf{P}(|S_n|>\frac{\delta^{1/2}}{6})\geq \frac{1}{36(2\delta+1)}>\frac{\delta}{45}$, since $\delta>\frac{1}{8}$. Then by the homogenity of δ (ii) holds.
(iii) By prop. 3.1.5 putting $t=\triangle$ and by (i) of this lemma:

$$\mathbb{E}|S_n|^p\leq \frac{2\triangle^p+\mathbb{E}\left(\max_{1\leq k\leq n}|X_k|\,\mathbf{1}_{\{\max_{1\leq k\leq n}|X_k|>\triangle\}}\right)^p}{c-8\delta}\leq \frac{1}{c-8\delta}2\left(\triangle^p+\sum_{i=1}^{n}\mathbb{E}|X_i|^p\mathbf{1}_{\{X_i>\triangle\}}\right)$$

which proves the first part of the inequality in (iii). By part (ii)

$$\mathbb{E}|S_n|^p\geq\left|\frac{\delta^{1/2}\triangle}{6}\right|^p\mathbf{P}\left(|S_n|>\frac{\delta^{1/2}\triangle}{6}\right)\geq c'\frac{\delta\triangle^p}{45}$$

Now, recalling prop. 3.1.5 we have by prop. 2.1.1 and prop. 3.1.1 that

$$\sum_{i=1}^{n}\mathbb{E}|X_i|^p\mathbf{1}_{\{|X_i|>\triangle\}}\leq \frac{\mathbb{E}\left(\max_{1\leq k\leq n}|X_k|\right)^p\mathbf{1}_{\{\max_{1\leq k\leq n}|X_k|>\triangle\}}}{1-\mathbf{P}\left(\max_{1\leq k\leq n}|X_k|>\triangle\right)}$$

$$\leq \frac{2}{1-4\delta}\mathbb{E}|S_n|^p\,.$$

□

For further references on similar results see, for example, Giné and Zinn's work [28].
The following Itô Nisio theorem is the infinite dimensional analogue of the important result of Lèvy on the convergence of series of independent real random variables:

Theorem 3.1.2. *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent E-valued random variables, and $S_n=\sum_{i=1}^{n}X_i$, then the following are equivalent:*
i) $\{S_n\}_n$ converges a.s.
ii) $\{S_n\}_n$ converges in probability.
iii) The distributions $\mathscr{L}(S_n)$ converge weakly.

For the original proof, see [39].

Proof. ii)$\Longrightarrow$i) If $S_n\underset{n\to\infty}{\longrightarrow}S$ in probability then from the Cauchy condition, for every $\epsilon>0$: $\sup_{j>1}\mathbf{P}(\|S_{n+j}-S_n\|\geq\epsilon)\underset{n\to\infty}{\longrightarrow}0$ but by prop. 3.1.3 this implies:

$$\mathbf{P}\left(\sup_{j>1}\|S_{n+j}-S_n\|\geq\epsilon\right)\leq 3\sup_{j>1}\mathbf{P}\left(\|S_{n+j}-S_n\|\geq\frac{\epsilon}{3}\right)\underset{n\to\infty}{\longrightarrow}0\,.$$

Now, write $\bigcup_{\epsilon\in\mathbf{Q}_{>0}} \bigcap_{n=1}^{\infty} \left\{ \sup_{j,k\geq n} \|S_j - S_k\| > 2\epsilon \right\}$ Taking in account the equation above, then

$$\mathbf{P}\left(\bigcap_{n=1}^{\infty} \left\{ \sup_{j,k\geq n} \|S_j - S_k\| > 2\epsilon \right\} \right) = 0 ,$$

and the a.s. convergence follows from this.
iii)$\Longrightarrow$ ii) First, observe that $\{S_n\}_n$ converges in probability if and only if

$$\mathbf{P}(\|S_n - S_m\| > \epsilon) \underset{n>m\to\infty}{\longrightarrow} 0$$

or equivalently, if $\mu_{m\,n} = \mathscr{L}(S_n - S_m) \underset{n>m\to\infty}{\longrightarrow} \delta_0$ weakly. Let $K \subset E$ be a compact set, then $\mathbf{P}(S_n - S_m \notin K - K) \leq \mathbf{P}(S_n \notin K) + \mathbf{P}(S_m \notin -K)$, and since $\mu_n = \mathscr{L}(S_n)$ converges weakly to a measure μ, then $\{\mu_{m\,n}\}_{n>m}$ is relatively compact, so there exists a measure ν and a subsequence of measures $\{\mu_{m_k\,n_k}\}_{m_k\,n_k}$, such that $\mu_{m_k\,n_k} \underset{k\to\infty}{\longrightarrow} \nu$, with $n_k > m_k$. Then the convolution of the measures verifies:

$$\mu_{m_k} * \mu_{m_k\,n_k} \underset{k\to\infty}{\longrightarrow} \nu * \mu ,$$

but on the other hand, the X_i's are independent so that the measures associated to the distributions of $S_{n_k} = (S_{n_k} - S_{m_k}) + S_{m_k}$ are the convolutions $\mu_{m_k} * \mu_{m_k\,n_k} = \mu_{n_k}$, thus $\mu * \nu = \mu$ which implies $\nu = \delta_0$.
i)$\Longrightarrow$ii)$\Longrightarrow$iii) is immediate. □

3.1.3 Convergence in the p-th mean

Often is easier to investigate the convergence in the mean, or L^p metric than the almost sure convergence. Also in many cases, it is important to have information about moments of series of independent random variables in E, a separable Banach space. We have seen in the previous theorem that for the case of sums of independent random variables, as in the real variable case a.s convergence, convergence in probability and convergence in distribution are equivalent. However, these type of convergences are not equivalent to convergence in mean (or L^p norm). Some additional conditions have to be considered. Let us first study some relationships between the a.s. convergence and moment inequalities for maximal functions.

Theorem 3.1.3. *Let $p > 0$, and let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent E-valued random variables. If the series $\sum_{i=1}^{\infty} X_i$ converges almost surely to Z then the following are equivalent:*
i) $\mathbb{E}(\sup_{1\leq k\leq\infty} \|S_k\|)^p < \infty$.
ii) $\mathbb{E}(\max_{1\leq k\leq\infty} \|X_k\|)^p < \infty$.
iii) $\forall t > 0$ *(equivalently, for some $t > 0$):* $\sum_{i=1}^{\infty} \mathbb{E}(\|X_i\|^p\, \mathbf{1}_{\{\|X_i\|>t\}}) < \infty$.
iv) $\mathbb{E}\,\|Z\|^p < \infty$.
And, if one of these conditions is fulfilled then $\lim_{n\to\infty} \mathbb{E}\,\|S_n - Z\|^p = 0$.

For an alternative proof, see [4].

Proof. ii)$\Longrightarrow$ i) This implication holds because $\sum\limits_{k=1}^{\infty} X_k$ is a.s. convergent and, therefore, given $a < 1$ and sufficient large t, $\mathbf{P}(\sup\limits_{1\leq k\leq\infty} \|S_k\| > t) < 3^{-p}a$, so that by prop 3.1.5, taking limit as $n \longrightarrow \infty$, we have

$$\mathbb{E}\left(\sup_{1\leq k\leq\infty} \|S_k\|\right)^p \leq \frac{3^p\,\mathbb{E}\left(\max\limits_{1\leq k\leq\infty} \|X_k\|\right)^p + t^p}{1-a}.$$

iii)$\Longrightarrow$ ii) follows from the following inequality:

$$\forall\, t \geq 0 \ \ (\max_{1\leq k\leq\infty} \|X_k\|)^p \leq t^p + \sum_{i=1}^{\infty} \|X_i\|^p \,\mathbf{1}_{\{\|X_i\|>t\}}.$$

ii)$\Longrightarrow$iii) Let $a > 0$, and let t be such that $\mathbf{P}(\max\limits_{1\leq k\leq\infty} \|X_k\| > t) \leq a < 1$ then, by prop. 3.1.6, for $s > t$: $\mathbf{P}(\max\limits_{1\leq k\leq\infty} \|X_k\| > s) \geq (1-a)\sum\limits_{i=1}^{\infty} \mathbf{P}(\|X_i\| > s)$, and by prop. 3.1.1,

$$\sum_{i=1}^{\infty} \mathbb{E}(\|X_i\|^p \,\mathbf{1}_{\{\|X_i\|>t\}}) \leq \frac{1}{1-a}\mathbb{E}\left(\max_{1\leq k\leq\infty} \|X_k\|\,\mathbf{1}_{\{\max\limits_{1\leq k\leq\infty}\|X_k\|>t\}}\right)^p.$$

From this, we have that if iii) is fulfilled for some $t_0 > 0$ then it also holds for $t > t_0$, and finally, for $t < t_0$, we have,

$$\sum_{i=1}^{\infty} \mathbb{E}(\|X_i\|^p \,\mathbf{1}_{\{\|X_i\|>t\}}) \leq \sum_{i=1}^{\infty} \mathbb{E}(\|X_i\|^p \,\mathbf{1}_{\{\|X_i\|>t_0\}}) + {t_0}^p \sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > t)$$

but $\sum\limits_{i=1}^{\infty} \mathbf{P}(\|X_i\| > t) < \infty$ by the Borel-Cantelli lemma, thus iii) holds.
i)$\Longrightarrow$iv) is immediate.
iv)$\Longrightarrow$ i) it follows from inequality 3.1.3 because, for a given $a > 0$ and t large enough to satisfy $\mathbf{P}\left(\sup\limits_{1\leq k\leq\infty} \|S_k\| > \frac{t}{2}\right) < a < 1$, then

$$\mathbf{P}\left(\sup_{1\leq k\leq\infty} \|S_k\| > s+t\right) \leq \frac{1}{1-a}\mathbf{P}(\|Z\| > s).$$

Recalling prop. 3.1.1 this yields

$$\mathbb{E}\left(\left(\sup_{1\leq k\leq\infty} \|S_k\| - t\right)^p \mathbf{1}_{\{\sup\limits_{1\leq k\leq\infty}\|S_k\|>t\}}\right) \leq \frac{\mathbb{E}\,\|Z\|^p}{1-a}.$$

Finally, by the dominated convergence theorem, $\mathbb{E}\left(\sup\limits_{1\leq k\leq\infty} \|S_k\|\right)^p < \infty$, and from $\|S_n - Z\| \leq 2\sup\limits_{1\leq k\leq\infty} \|S_k\|$, then $\lim\limits_{n\to\infty} \mathbb{E}\,\|S_n - Z\|^p = 0$. □

In order to get convergence in the p-mean from the a.s. convergence, an additional condition could be the following: suppose, there exists positive constants C, λ, such that, $\mathbb{E}\,\|X_n\|^p\,\mathbf{1}_{\{\|X_n\|>\lambda\}} \leq C\mathbf{P}(\|X_n\| > \lambda)$, then from theorem 3.1.3 and the Borel-Cantelli lemma, we have that the a.s. convergence of $\sum_i X_i$ implies convergence in the p-mean. Later, we shall discuss some other conditions, in particular when E is the Lebesgue space $L^p(X, \Sigma, \mu)$.

Corollary 3.1.1. *let $p > 0$, and let $\{X_i\}_{i=1}^{\infty}$ be a sequence of E-valued independent random variables, then the series $\sum_{i=1}^{\infty} X_i$ converges a.s. $\iff \exists\,\lambda > 0$ (or equivalently, $\forall\,\lambda > 0$) for which the following conditions are satisfied:*
i) $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > \lambda) < \infty$.
ii) $\sum_{i=1}^{\infty} X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$ converges in $L^p(E)$.

Proof. ($\Longrightarrow$)If $\sum_{i=1}^{\infty} X_i$ converges a.s. then by the Borel-Cantelli lemma $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > \lambda) < \infty$. Hence $\sum_{i=1}^{\infty} X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$ converges a.s., and by theorem 3.1.3 it converges in $L^p(E)$ (since $\left\|X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}\right\| \leq \lambda$).
($\Longleftarrow$) by the Itô-Nisio theorem 3.1.2 the series $\sum_{i=1}^{\infty} X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$ converges a.s. and by i) and the Borel-Cantelli lemma , $\sum_{i=1}^{\infty} X_i\mathbf{1}_{\{\|X_i\|>\lambda\}}$ converges a.s. Thus $\sum_{i=1}^{\infty} X_i$ converges a.s.. □

Now, let $E = H$ be a separable Hilbert space, from the previous we can get the following analogue of Kolmogorov's three series theorem.

Corollary 3.1.2. *Let H be a separable Hilbert space, and let $\{X_n\}_n$ be a sequence of independent H valued random variables. Then $\sum_{i=1}^{\infty} X_i$ converges a.s. $\iff \exists\,\lambda > 0$ (or equivalently $\forall\,\lambda > 0$) the following series are convergent:*
i) $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > \lambda)$, ii) $\sum_{i=1}^{\infty} \mathbb{E}X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$, iii) $\sum_{i=1}^{\infty} Var\left(X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}\right)$.

Using a symmetrization argument it is possible to prove the following useful result:

Corollary 3.1.3. *Let $\{X_n\}_n$ be a sequence of independent E-valued random variables. Then the series $\sum_{i=1}^{\infty} X_i$ converges a.s. $\iff$ for some (equivalently for all) $\lambda > 0$, the following three conditions hold:*
i) $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > \lambda) < \infty$ ii) $\sum_{i=1}^{\infty} \mathbb{E}X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$ converges in E. iii) $\sum_{i=1}^{\infty} Y_i$ converges a.s. where Y_i is a symmetrization of $X_i\mathbf{1}_{\{\|X_i\|\leq\lambda\}}$.

Proof. By corollary 3.1.1, it suffices to prove that the series $\sum_{i=1}^{\infty} Y_i$ converges in the 1-mean if and only if conditions $ii)$ and $iii)$ hold. Convergence in the 1-mean obviously implies $ii)$ and

it also implies *iii*) in view of theorem 3.1.2.
On the other hand by corollary 3.1.2, the series $\sum_{i=1}^{\infty} Y_i$ converges in the 1-mean . Therefore,

$$\sum_{i=1}^{\infty} \int_{\Omega} Y_i(.,\omega')\mathbf{P}(\omega') = \sum_{i=1}^{\infty}(X_i \mathbf{1}_{\{\|X_i\| \leq \lambda\}} - \mathbb{E} X_i \mathbf{1}_{\{\|X_i\| \leq \lambda\}})$$

converges in the 1-mean. Since, by *ii*), the series $\sum_{i=1}^{\infty} \mathbb{E} X_i \mathbf{1}_{\{\|X_i\| \leq \lambda\}}$ converges in E, the series $\sum_{i=1}^{\infty} X_i$ converges in the 1-mean □

From the previous corollary 3.1.2, as in the real valued case, one can derive the following very useful result:

Corollary 3.1.4. *Let H be a separable Hilbert space, and let $\{X_n\}_n$ be a sequence of independent H valued random variables. Such that $X_n \in L^2(\Omega, \mathcal{F}, \mathbf{P}, H)$ and moreover $\sum_{i=1}^{\infty} Var(X_i) < \infty$. Then $\sum_{i=1}^{\infty} X_i$ converges a.s..*

Alternatively, this result as in the real valued case, can be obtained from a maximal inequality, which is useful on its own:

Theorem 3.1.4. *(Hilbert space version of Kolmogorov's inequality) Let H be a separable Hilbert space and let $\{X_k\}_{k=1}^n$ be a sequence of independent H valued random variables. Such that $X_n \in L^2(\Omega, \mathcal{F}, \mathbf{P})$, then given $\lambda > 0$:*

$$\mathbf{P}\left(\max_{1 \leq k \leq n} \left\| \sum_{i=1}^{k} X_i \right\| > \lambda \right) \leq \frac{\sum_{i=1}^{n} Var(X_i)}{\lambda^2} . \tag{3.1.3}$$

Proof. (sketch) The proof of this result is very similar to the original version for real valued random variables. In a similar manner to props. 3.1.3 and 3.1.4 it starts defining events

$$A_k = \{\|S_1\| \leq \lambda; \ldots, \|S_{k-1}\| \leq \lambda, \|S_k\| > \lambda\} ,$$

then the $A_k' s$ are disjoint and $A = \left\{ \max_{1 \leq k \leq n} \|S_k\| > \lambda \right\} = \bigcup_{k=1}^{n} A_k$. Now estimate $\mathbf{P}(A)$ as before. □

Note that, for example, corollary 3.1.4 gives a sufficient condition for a mean convergent series to converge a.s. However to obtain a kind of reciprocal, we have to add some conditions on the moments of the $\|X_n\|' s$.

Theorem 3.1.5. *Let $\{X_n\}_n$ be a sequence of independent H-valued random variables such that $X_n \in L^4(\Omega, \mathcal{F}, \mathbf{P}, H)$, for $n = 1, \ldots$ and suppose that there exists a positive constant C, such that $\mathbb{E} \|X_n\|^4 \leq C Var^2(X_n)$, for all n, if $\sum_{n=1}^{\infty} X_n$ converges a.s. then $\sum_{n=1}^{\infty} Var(X_n) < \infty$, in particular $\sum_{n=1}^{\infty} X_n$ converges in $L^2(H)$.*

Proof. (sketch.) Take $\lambda \in (0,1)$. And define:

$$A_n = \left\{ \left\| \sum_{k=1}^{n} X_k \right\|^2 > \lambda^2 \sum_{k=1}^{n} Var(X_k) \right\} , \quad A = \overline{\lim_{n\to\infty}} A_n \,.$$

On the other hand, apply prop. 3.1.2, to get: $\exists\, k(\lambda) > 0$ independent of n, such that,

$$\mathbf{P}\left(\left\| \sum_{i=1}^{n} X_i \right\|^2 > \lambda \sum_{i=1}^{n} Var(X_i) \right) > k \,.$$

From this estimate $\mathbf{P}(\overline{\lim_{n\to\infty}} A_n)$ □

In the previous result, we have imposed a "reverse Hölder inequality" condition on the random variables. In the following, alternatively, the additional condition to ensure convergence in the p-th mean, is similar to a "reverse Tchevicheff inequality".

Lemma 3.1.5. *Let $p > 0$. If $\{X_n\}_{n=1}^{\infty}$ are real, independent, random variables such that for some $\alpha,\ \beta, \lambda > 0$, $\mathbf{P}(|X_i| > \alpha) \geq \beta$, and*

$$\mathbb{E}|X_i|^p \mathbf{1}_{\{|X_i|>t\}} \leq \lambda t^p \mathbf{P}(|X_i| > t)\,,$$

for all $t > \alpha$ and $i \in \mathbb{N}$, then for each sequence $f_1, f_2, \cdots \in E$, the series $\sum_{i=1}^{\infty} X_i f_i$ converges a.s. $\iff$ it converges in the p-th mean. Moreover $\forall\, q < p$, there exists a constant $C > 0$ such that:

$$\left(\mathbb{E} \left\| \sum_{i=1}^{\infty} X_i f_i \right\|^p \right)^{\frac{1}{p}} \leq C \left(\mathbb{E} \left\| \sum_{i=1}^{\infty} X_i f_i \right\|^q \right)^{\frac{1}{q}} \,.$$

Proof. If the series converges a.s. then $\sum_i \mathbf{P}(\|X_i f_i\| > 1) < \infty$. Then, by the hypothesis $1/\|f_i\| > \alpha$ for i large enough. For these $i's$ we have:

$$\mathbb{E}\, \|X_i f_i\|^p \, \mathbf{1}_{\{\|X_i f_i\|>1\}} = \|f_i\|^p \, \mathbb{E}|X_i|^p \mathbf{1}_{\{|X_i|>1/\|f_i\|\}} \leq \lambda \mathbf{P}\left(|X_i| > \frac{1}{\|f_i\|} \right)$$

which, as we have seen before, implies the convergence of the series. The proof of the inequality is immediate. □

Example For example, if the X_i's are i.i.d's symmetric α-stable random variables, $\sum_{i=1}^{\infty} X_i f_i$ converges a.s. if and only if it converges in the p-th mean.

Chapter 4

L^p Valued Random Series

4.1 Introduction

Here we will concentrate in convergence problems when the random variables take values in the Lebesgue spaces $L^p(X)$.
First, we will study sums of independent random variables, and we will give necessary and sufficient conditions for the a.s. convergence. Afterwords, we study the particular case of stable series, since for this case some more tractable conditions can be given. Finally, we will consider another particular case, when the series comes from an unconditional basis. For this case, we will study some relationships between different types of convergence. In this case, sometimes, the somewhat strong condition of a basis being unconditional allows us to drop the independence hypothesis on the terms of the sums. This may be regarded as the case when we have the representation of a random element with respect to a given basis (see Chapter 2). Recalling that given a function in $L^p(X)$ (or equivalence class) its series expansion, with respect to a basis, always converges in norm, but it may not converge a.e.. So for this case we will also consider some pointwise convergence problems, namely, we will study if given a random element (or series), when this representation, with respect to an unconditional basis, converges a.e. $[\mu]$ respect to the underlying measure space (X, Σ, μ), with probability one .

4.2 Auxiliary Results

Here we will be considering two measure spaces: a probability space, say $(\Omega, \mathcal{F}, \mathbf{P})$ and a σ-finite measure space (X, Σ, μ). As usual, we define the Lebesgue spaces $L^p(X, \Sigma, \mu)$. We talk about properties that hold *almost everywhere* $[\mu]$ *almost surely.* This must be understood without ambiguity meaning that such a property holds for almost all pairs (x, ω) in a measurable subset of $X \times \Omega$ with respect to the complete measure $\mu \times \mathbf{P}$ [40]. Some results of this section remain true in a general separable Banach space with arbitrary norm $\|.\|$. In this case we will denote it just $(E, \|.\|)$. Recall that by convergence in the p- mean, we mean that $\mathbb{E}\,\|X_n - X\|^p \longrightarrow 0$ whenever $n \to \infty$. In the particular case that $E = L^p(X, \Sigma, \mu)$ is σ-finite, one would expect this type of convergence to be equivalent to convergence in the norm of the space $L^p(X \times \Omega, \Sigma \otimes \mathcal{F}, \mu \times \mathbf{P})$. Moreover, if $Y : \Omega \longrightarrow L^p(X, \Sigma, \mu)$ is a random variable, then for each $\omega \in \Omega$, $Y(\omega)$ represents a (an equivalence class) function of $L^p(X, \Sigma, \mu)$. Then if for each ω we select a particular function $f(\omega,.)$ of this equivalence class, we obtain a function $f(\omega,x) : X\times\Omega \longrightarrow \mathbb{R}$. This resulting function is called a *representation* of the random variable Y. However, it is not immediate that this representation is a $\Sigma\otimes\mathcal{F}$-measurable function. So, we first give some conditions that identify $Y(\omega)$ with a $\Sigma\otimes\mathcal{F}$-measurable function $f(\omega,.)$.

Theorem 4.2.1. *Let (X,Σ,μ) be a σ-finite measure space, $(\Omega,\mathcal{F},\mathbf{P})$ a probability space and let $Y:\Omega\longrightarrow L^p(X,\Sigma,\mu)$ be a $\mathbf{P}$-integrable random variable. Let $1\le p\le\infty$, and let $(X\times\Omega,\Sigma\otimes\mathcal{F},\mu\times\mathbf{P})$ be the product measure space. Then there exists a $\Sigma\otimes\mathcal{F}$-measurable function, $f:X\times\Omega\longrightarrow\mathbb{R}$, which is uniquely determined except for a set of $\mu\times\mathbf{P}$-measure zero, and such that $f(.,\omega)=Y(\omega)$ for almost all $\omega\in\Omega$ $[\mathbf{P}]$. Moreover, $f(x,.)$ is $\mathbf{P}$-integrable on Ω for almost all $x\in X$ $[\mu]$ and the integral $\int\limits_\Omega f(x,\omega)d\mathbf{P}$ as a function of x, is equal to the element of $L^p(X,\Sigma,\mu)$, $\int\limits_\Omega Y(\omega)d\mathbf{P}$.*

Proof. See [20] □

Remark 4.2.1. In chapter 2 we introduced the definition of the Bochner integral or expected value. It would be desirable that this notion coincides with an ordinary Lebesgue integral for the case of L^p valued random variables. In ([20], chapter III) is proved that if $f:X\times\Omega\longrightarrow\mathbb{R}$ is a measurable function, such that it belongs to $L^1(\Omega,\mathcal{F},\mathbf{P},L^p(X))$, then the the Bochner integral or expected value $\mathbb{E}(f)\in L^p(X,\Sigma,\mu)$ equals a.e. the Lebesgue integral $\int_\Omega f(x,\omega)d\mathbf{P}(\omega)$.

Taking in account these results, on the following, we will sometimes make the following abuse of notation: given a vector valued random variable $Y:\Omega\longrightarrow L^p(X,\Sigma,\mu)$, we will denote also as Y its $\Sigma\otimes\mathcal{F}$-measurable representation. So for a fixed $x\in X$ it is possible, and useful, to think this measurable representation $Y(x,.)$ as an ordinary real valued random variable. Then, we would like to know if these real valued random variables inherit from the original, some characteristic properties, such as independence and symmetry. Before answering these questions, we will prove an auxiliary result on real valued random variables:

Lemma 4.2.1. *Let $(\Omega,\mathcal{F},\mathbf{P})$ be a probability space and let $\{X_1,\dots,X_m\}\subset L^p(\Omega,\mathcal{F},\mathbf{P})$, with $1\le p<\infty$. Then, there exists a family of finite sub σ-algebras $\{\mathcal{G}_n\}_{n\in\mathbb{N}}$, such that $\mathcal{G}_{n-1}\subseteq\mathcal{G}_n$, and $\forall\, j=1,\dots,m$*

$$(\mathbb{E}|X_j-\mathbb{E}[X_j|\mathcal{G}_n]|^p)^{\frac{1}{p}}<\frac{1}{n}\,.$$

Moreover, $\mathbb{E}[X_j|\mathcal{G}_n]\longrightarrow X_j$ a.s, as $n\longrightarrow\infty$.

Proof. If $\{X_1,\dots,X_m\}\subset L^p(\Omega,\mathcal{F},\mathbf{P})$, given $n\in\mathbb{N}$, for each X_j there exists a simple random variable Y_j^n, such that $(\mathbb{E}|X_j-Y_j^n|^p)^{\frac{1}{p}}<\frac{1}{2n}$. The sigma algebra generated by the set of random variables $A_n=\{Y_j^k\}_{j=1,\dots,m}^{k=1,\dots,n}$, $\sigma(A_n)$ is finite and verifies $\sigma(A_{n-1})\subseteq\sigma(A_n)$, so we take $\mathcal{G}_n=\sigma(A_n)$. It remains to prove that the $\mathcal{G}_n$'s verify the last property. First note that $\mathbb{E}[Y_j^n|\mathcal{G}_n]=Y_j^n$ a.s. then,

$$(\mathbb{E}|X_j-\mathbb{E}[X_j|\mathcal{G}_n]|^p)^{\frac{1}{p}}\le\left(\mathbb{E}|X_j-Y_j^n|^p\right)^{\frac{1}{p}}+\left(\mathbb{E}|\mathbb{E}[X_j|\mathcal{G}_n]-Y_j^n|^p\right)^{\frac{1}{p}}$$

$$=\left(\mathbb{E}|X_j-Y_j^n|^p\right)^{\frac{1}{p}}+\left(\mathbb{E}|\mathbb{E}[X_j-Y_j^n|\mathcal{G}_n]|^p\right)^{\frac{1}{p}}\le 2(\mathbb{E}|X_j-Y_j^n|^p)^{\frac{1}{p}}<\frac{1}{n}\ \text{(by theorem 2.1.13)}$$

By Lévy´s martingale theorem 2.1.15 $\mathbb{E}[X_j|\mathcal{G}_n] \longrightarrow \mathbb{E}[X_j|\mathcal{G}_\infty]$ a.s. But on the other hand, by the previous arguments $\mathbb{E}[X_j|\mathcal{G}_n] \longrightarrow X_j$ in L^p and then in probability, thus by Riesz´s theorem there exists a subsequence $\mathbb{E}[X_j|\mathcal{G}_{n_k}] \longrightarrow X_j$ a.s. From this fact, $\mathbb{E}[X_j|\mathcal{G}_n] \longrightarrow X_j$ a.s. □

Theorem 4.2.2. *Let (X, Σ, μ) be a σ-finite measure space, $(\Omega, \mathcal{F}, \mathbf{P})$ a probability space, and $1 \leq p < \infty$, and let $Y : \Omega \longrightarrow L^p(X, \Sigma, \mu)$ be an $\mathbf{P}$-integrable random variable. Then:*
a) If Y is symmetric, i.e. $\mathbf{P}(Y \in B) = \mathbf{P}(-Y \in B)$ for every B in the Borel σ-algebra of $L^p(X, \Sigma, \mu)$, then for almost all $x \in X$ $[\mu]$, $Y(x, .)$ is a symmetric real valued random variable.
b) If Z is another $L^p(X, \Sigma, \mu)$-valued integrable random variable, independent of Y, then $Y(x, .)$ and $Z(x, .)$ are independent for almost all $x \in X$ $[\mu]$.

Proof. a) Without loss of generality we may assume that (X, Σ, μ) is another probability space. If $Y : \Omega \longrightarrow L^p(X, \Sigma, \mu)$ verifies $\mathbf{P}(Y \in B) = \mathbf{P}(-Y \in B)$ for every B in the Borel σ-algebra of $L^p(X, \Sigma, \mu)$. Then for every $g \in L^q(X, \Sigma, \mu)$ and $t \in \mathbb{R}$, with $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\mathbf{P}(\langle g, Y \rangle \leq t) = \mathbf{P}(-\langle g, Y \rangle \leq t)\,, \tag{4.2.1}$$

where $\langle ., . \rangle$ denotes the usual duality pairing. On the other hand, if $L^p(X, \Sigma, \mu)$ is separable there exists a dense subset $\{f_n\}_{n \in \mathbb{N}}$. By theorem 4.2.1, for each $k = 1, \ldots$ there exist finite sub σ-algebras of Σ, $\{\mathcal{G}_m^k\}_m$, such that $\mathcal{G}_m^k \subseteq \mathcal{G}_{m+1}^k$ and $\mathbb{E}_\mu[f_k|\mathcal{G}_m^k] \underset{m \to \infty}{\longrightarrow} f_k = \mathbb{E}_\mu[f_k|\mathcal{G}_\infty^k]$ a.e. $[\mu]$ and in $L^p(X, \Sigma, \mu)$. Here $\mathbb{E}_\mu[\,.\,|\,.\,]$ denotes the conditional expectation with respect to (X, Σ, μ). Take $\mathcal{G}_m = \sigma\left(\bigcup_{k=1}^{m} \mathcal{G}_m^k\right)$, then $\mathcal{G}_m$ is finite and $\mathcal{G}_m \subseteq \mathcal{G}_{m+1}$. Note that since each $\mathcal{G}_m$ is finite, then $\mathcal{G}_m = \sigma(\{E_1, \ldots, E_l\})$, for some measurable partition $\{E_1, \ldots, E_l\}$ of X, and for every $f \in L^p(X, \Sigma, \mu)$:

$$\mathbb{E}_\mu[f|\mathcal{G}_m](x) = \sum_{i=1}^{l} \frac{1}{\mu(E_i)} \int_{E_i} f d\mu \mathbf{1}_{E_i}(x) \;\text{ for almost all } x \in X\ [\mu]\,.$$

But this equation can be rewritten as

$$\mathbb{E}_\mu[f|\mathcal{G}_m](x) = \langle f, K_m(\,.\,, x) \rangle\,, \tag{4.2.2}$$

where $K_m(y, x) = \sum_{i=1}^{l} \frac{1}{\mu(E_i)} \mathbf{1}_{E_i}(y) \mathbf{1}_{E_i}(x)$. Fix $x \in X$, $K_m(\,.\,, x)$ represents an element of $L^q(X, \Sigma, \mu)$ and is independent of the particular choice of f.
On the other hand, given $\epsilon > 0$, there exists f_{k_0} such that $\|Y - f_{k_0}\|_{L^p(X,\Sigma,\mu)} < \frac{\epsilon}{2}$, and by theorem 2.1.15,

$$\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m] \longrightarrow \mathbb{E}_\mu[f_{k_0}|\mathcal{G}_\infty] \quad \text{a.e. } [\mu] \text{ and in } L^p(X, \Sigma, \mu)\,.$$

But $f_{k_0} = \mathbb{E}_\mu[f_{k_0}|\mathcal{G}_\infty^{k_0}]$ a.e. and $\mathcal{G}_\infty^{k_0} \subseteq \mathcal{G}_\infty$ then $\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_\infty] = f_{k_0}$ a.e. Thus, $\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m] \longrightarrow f_{k_0}$ a.e. and in $L^p(X, \Sigma, \mu)$. Now,

$$\begin{aligned} &\|\mathbb{E}_\mu[Y|\mathcal{G}_m] - Y\|_{L^p(X)} \\ &\leq \|Y - f_{k_0}\|_{L^p(X)} + \|\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m] - f_{k_0}\|_{L^p(X)} + \|\mathbb{E}_\mu[Y|\mathcal{G}_m] - \mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m]\|_{L^p(X)} \end{aligned}$$

$$\leq 2\,\|Y - f_{k_0}\|_{L^p(X,\Sigma,\mu)} + \|\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m] - f_{k_0}\|_{L^p(X)} \quad \text{(by theorem 2.1.13)}$$
$$< \epsilon + \|\mathbb{E}_\mu[f_{k_0}|\mathcal{G}_m] - f_{k_0}\|_{L^p(X)} \underset{m\to\infty}{\longrightarrow} \epsilon\,.$$

Again, theorem 2.1.15 asserts that the sequence $\{\mathbb{E}_\mu[Y|\mathcal{G}_m]\}_m$ converges a.e. $[\mu]$. Then for almost all x,

$$\mathbb{E}_\mu[Y|\mathcal{G}_m](x) \longrightarrow Y(x)\,. \tag{4.2.3}$$

Note that since Y takes values in $L^p(X,\Sigma,\mu)$ and $K_m(.,x) \in L^q(X,\Sigma,\mu)$, then, for every $x \in X$, $\langle Y, K_m(\,.\,,x)\rangle$ represents, a well defined, real valued random variable. And then all the previous arguments hold for almost all $\omega \in \Omega$. Thus recalling equation 4.2.2, we have that, for almost all $x \in X$ $[\mu]$,

$$\langle Y, K_m(\,.\,,x)\rangle = \mathbb{E}_\mu[Y|\mathcal{G}_m](x) \longrightarrow Y(x) \text{ a.s. } [\mathbf{P}]\,.$$

Recalling eq. 4.2.1, we have that for almost all $x \in X$ $[\mu]$, and every $t \in \mathbb{R}$:

$$\mathbf{P}(\langle K_m(.,x), Y\rangle \leq t) = \mathbf{P}(-\langle K_m(.,x), Y\rangle \leq t)\,,$$

and since a.s. convergence implies convergence in distribution the result follows taking $m \longrightarrow \infty$.

b) The idea of the proof is very similar to a). If X, Y are independent, then for every g, h in $L^q(X,\Sigma,\mu)$, with $\frac{1}{p}+\frac{1}{q}=1$ and $t_1, t_2 \in \mathbb{R}$,

$$\mathbf{P}(\{\langle g, Z\rangle \leq t_1\} \cap \{\langle h, Y\rangle \leq t_2\}) = \mathbf{P}(\langle g, Z\rangle \leq t_1)\mathbf{P}(\langle h, Y\rangle \leq t_2)\,,$$

then taking again $g = h = K_m(\,.\,,x)$ as in a) and taking $m \longrightarrow \infty$, the result follows. □

Let J be any finite subset of $\mathbb{N}$ then, the same argument leads to the following, if $\{Y_j\}_{j\in J}$ are independent as vector valued random variables then there exists (P) A_J, such that $\mu(A_J^c) = 0$ and for all $x \in A_J$: given $t_j \in \mathbb{R}$, $\mathbf{P}\left(\bigcap_{j\in J}\{Y_j(x) \leq t_j\}\right) = \prod_{j\in J}\mathbf{P}(Y_j(x) \leq t_j)$. If $\{Y_j\}_{j\in\mathbb{N}}$ are independent as vector valued random variables, then for every finite subcollection of $\{Y_j\}_{j\in J} \subset \{Y_j\}_{j\in\mathbb{N}}$ property (P) holds, so taking $A^c = \bigcup_J A_j^c$, then the scalar valued random variables $\{Y_j(x)\}_{j\in\mathbb{N}}$ are independent for all $x \in A$.

4.3 Convergence of series of L^p valued random variables

Now, let us prove a result which is a consequence of theorem 2.1.3:

Theorem 4.3.1. *Let $\{X_i\}_i$ be a sequence on independent random elements in $L^p(X,\Sigma,\mu)$, $p \in [1,\infty)$ such that $\mathbb{E}X_i = 0$, then: if $\sum_i X_i$ converges in the norm topology of $L^p(X \times \Omega)$ then it converges $[\mu]$-a.e. a.s..*

Remark. The main idea of this result is to transfer a maximal inequality to the product space. We could use a weaker condition on the X_i's. For example if the sequence of partial sums forms a martingale, one could use proposition 2.1.3 or theorem 2.1.14. For further references about this type of argument see section 4.5 at the end of this chapter.

Proof. First note that, for almost all $x \in X$ $[\mu]$, by theorem 4.2.2, the $X_i(x)$'s can be seen as independent real valued random variables. Now let us transfer theorem 2.1.3 for random variables to this context:

$$\int_X \mathbf{P}\left((x,\omega) \in X \times \Omega : \max_{j=1,\dots,n} \left|\sum_{i=m+1}^{m+j} X_i(x,\omega)\right| > \delta\right) d\mu$$

$$\leq \frac{1}{\delta^p} \int_X \mathbb{E}\left|\sum_{i=m+1}^{m+n} X_i\right|^p d\mu \qquad \text{(By theorem 2.1.3)}$$

$$= \frac{1}{\delta^p}\mathbb{E}\left\|\sum_{i=m+1}^{m+n} X_i\right\|^p_{L^p(X,\Sigma,\mu)} \qquad \text{(By Fubini's theorem)}\,. \tag{4.3.1}$$

Now, with this maximal inequality we have: in $X \times \Omega$ write $\nu = \mu \times \mathbf{P}$; take $\delta > 0$ and $m \in \mathbb{N}$ then:

$$\left\{(x,\omega) : \sup_{j\in\mathbb{N}} \left|\sum_{i=m+1}^{m+j} X_i(x,\omega)\right| > \delta\right\} \subset \bigcup_{n\in\mathbb{N}} D_n\,,$$

where $D_n = \left\{(x,\omega) : \max_{j=1,\dots,n} \left|\sum_{i=m+1}^{m+j} X_i(x,\omega)\right| > \delta\right\}$. Clearly $D_n \subset D_{n+1}$ then:

$$\nu\left\{(x,\omega) : \sup_{j\in\mathbb{N}} \left|\sum_{i=m+1}^{m+j} X_i(x,\omega)\right| > \delta\right\} \leq \nu\left(\bigcup_{n\in\mathbb{N}} D_n\right) = \lim_{n\to\infty} \nu(D_n)$$

$$\leq \frac{K_p^p}{\delta^p} \lim_{n\to\infty} \mathbb{E}\left\|\sum_{i=m+1}^{n} X_i\right\|^p = C(m,\delta) < \infty \qquad \text{(By equation 4.3.1)}\,.$$

Since $\sum_{i=1}^{n} X_i$ is Cauchy in $L^p(X \times \Omega)$, this implies:

$$\lim_{m\to\infty} \nu\left\{(x,\omega) : \sup_{j\in\mathbb{N}} \left|\sum_{i=m+1}^{m+j} X_i(x,\omega)\right| > \delta\right\} = 0\,. \tag{4.3.2}$$

Define $E_{n\,\delta} = \left\{(x,\omega) : \sup_{j,k>n} \left|\sum_{i=k+1}^{j} X_i(x,\omega)\right| > 2\delta\right\}$, then:

$$E_{n\,\delta} \subset \left\{\sup_{j\in\mathbb{N}} \left|\sum_{i=n+1}^{n+j} X_i(x,\omega)\right| > \delta\right\}\,,$$

so that $E_{n+1\,\delta} \subset E_{n\,\delta}$. From this and equation 4.3.2 we have:

$$\nu\left(\bigcap_{n\in\mathbb{N}} E_{n\,\delta}\right) = \lim_{n\to\infty} \nu(E_{n\,\delta}) = 0 \Longrightarrow \nu\left(\bigcup_{\delta\in\mathbb{Q}_{>0}} \bigcap_{n\in\mathbb{N}} E_{n\,\delta}\right) = 0\,.$$

□

4.3.1 General conditions for sums of independent random variables

Theorem 4.3.2. *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent and symmetric $L^p(X,\Sigma,\mu)$-valued random variables, and let $c>0$. Then, the series $\sum_{i=1}^{\infty} X_i$ converges a.s. in $L^p(X,\Sigma,\mu) \iff$ the following conditions are satisfied:*
i) $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\|_{L^p} > c) < \infty$.
ii) $\triangle((X_i)_i) \in L^p(X,\Sigma,\mu)$.
iii)

$$\int_X \sum_{i=1}^{\infty} \mathbb{E}|X_i(x)|^p \mathbf{1}_{\{|X_i(x)|>\triangle((X_i)_i)(x), \|X_i\|_{L^p}\le c\}} d\mu(x) < \infty .$$

Proof. By the Borel-Cantelli lemma and by corollary 3.1.3 we have that $\sum_{i=1}^{\infty} X_i$ converges a.s. in $L^p(X,\Sigma,\mu) \iff \sum_{i=1}^{\infty} \mathbf{P}(\|X_i\|_{L^p} > c) < \infty$ and $\sum_{i=1}^{\infty} X_i \mathbf{1}_{\{\|X_i\|_{L^p}\le c\}}$ converges a.s. So it will be sufficient to show that these conditions are equivalent.
$\Longrightarrow$) Given $c>0$, if $Y_i = X_i \mathbf{1}_{\{\|X_i\|_{L^p}\le c\}}$, then $\sum_{i=1}^{\infty} \mathbb{E}\|Y_i\|^p \mathbf{1}_{\{\|Y_i\|>c\}} = 0$. Thus by theorem 3.1.3, calling S' the a.s. limit of $\sum_{i=1}^{\infty} Y_i$, we have that $\mathbb{E}\|S'\|^p < \infty$. Since the Y_n are symmetric, then by theorem 4.2.2 we have that for almost all $x \in X$ the $Y_n(x)$ are symmetric real valued random variables. If we denote, for $k = 1\ldots$, $S'_k(x) = \sum_{i=1}^{k} Y_i(x)$, then also by theorem 3.1.3[a], S'_n converges in $L^p(\Omega,\mathcal{F},\mathbf{P},L^p(X))$. Now, define $(S'_k)^*(x) = \max_{1\le j\le k} |S'_j(x)|$, then for almost all x $[\mu]$, by the symmetry of the Y_j's recalling theorem 2.1.8, we have:

$$\mathbf{P}(|(S'_n)^*(x)| > \lambda) \le 2\mathbf{P}(|S'_n(x)| > \lambda) .$$

Now, let us consider the product measure $\nu = \mu \times \mathbf{P}$, so by the previous inequality, Chevychev´s inequality over (X,Σ,μ) and Fubini´s theorem, we have:

$$\nu(|(S'_n)^*| > \lambda) \le 2\nu(|S'_n| > \lambda) = \mathbb{E}\left(\mu(|S'_n| > \lambda)\right) \le \frac{\mathbb{E}\|S'_n\|^p}{\lambda^p} ,$$

thus $S'_n(x)$ converges a.s. for almost all $x \in X$ $[\mu]$ by a similar argument to the proof of theorem 4.3.1[b]. Then for almost all $x \in X$ we can apply lemma 3.1.4:

$$C_1 \mathbb{E}|S'(x)|^p \le \left(\triangle((Y_i)_i)\right)^p(x) + \sum_{i=1}^{\infty} \mathbb{E}|Y_i(x)|^p \mathbf{1}_{\{|Y_i|>\triangle(Y_i)_i\}} \le C_2 \mathbb{E}|S'(x)|^p \ ,$$

then integrating, and by Fubini´s theorem:

$$C_1 \mathbb{E}\left\|S'\right\|^p \le \|\triangle((Y_i)_i)\|^p + \sum_{i=1}^{\infty} \int_X \mathbb{E}|Y_i(x)|^p \mathbf{1}_{\{|Y_i|>\triangle(Y_i)_i\}} d\mu \le C_2 \mathbb{E}\left\|S'\right\|^p \ ,$$

[a]Although this is a result for normed spaces this theorem can be easily adapted for the L^p spaces, with $p<1$.
[b]If we restrain to the case $p \ge 1$ one could appeal to theorem 4.3.1 directly.

from this

$$\int_X \sum_{i=1}^{\infty} \mathbb{E}|Y_i(x)|^p \mathbf{1}_{\{|Y_i(x)|>\triangle((Y_i)_i)(x)\}} d\mu(x) < \infty\,, \tag{4.3.3}$$

and $\triangle(Y_i)_i \in L^p(X)$. but by the definition of $\triangle$, we always have that

$$\triangle((Y_i(x))_i) \leq \triangle((X_i(x))_i)\,,$$

and, from the definition of Y_i, we have that over $\{\|X_i\| \leq c\}$: $|Y_i| > \triangle((Y_i)_i) \iff |X_i| > \triangle((Y_i)_i)$. Then

$$\{|X_i| > \triangle((X_i)_i), \|X_i\|_{L^p} \leq c\} \subseteq \{|Y_i| > \triangle((Y_i)_i)\}\,,$$

and by eq. 4.3.3 this readily implies that

$$\int_X \sum_{i=1}^{\infty} \mathbb{E}|X_i(x)|^p \mathbf{1}_{\{|X_i(x)|>\triangle(x), \|X_i\|_{L^p} \leq c\}} d\mu(x) < \infty\,.$$

For $\delta' < \delta$, we have that $\triangle_\delta((X_i(x))) \leq \triangle_{\delta'}((Y_i(x)))$,[c] whenever $\sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > c) < \delta - \delta'$. This fact follows from the inequality:

$$\sum_{i=1}^{\infty} \Psi^2(X_i(t)/s) \leq \sum_{i=1}^{\infty} \Psi^2(Y_i(t)/s) + \sum_{i=1}^{\infty} \mathbf{P}(\|X_i\| > c)\,.$$

And then, $\triangle(Y_i)_i \in L^p(X) \Longrightarrow \triangle(X_i)_i \in L^p(X)$.

$\Longleftarrow$) By the definition of $\triangle_\delta((X_i)_i)$, given $x \in X$ and $\epsilon > 0$, there exists $s = s(x)$ such that $\sum_i \left(\mathbb{E}\left(\frac{X_i(x)}{s}\right)^2 \mathbf{1}_{\{|X_i| \leq s\}} + \mathbf{P}(|X_i(x)| > s) \right) \leq \delta$, and $\triangle \leq s < \triangle + \epsilon$. Thus for every $\epsilon > 0$, $\left(\sum_i \mathbb{E}\left(X_i(x)\right)^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{1}{2}} < \delta^{1/2}(\triangle((X_i)_i) + \epsilon)$. So if $\triangle((X_i)_i) \in L^p(X)$, then

$$\left(\sum_i \mathbb{E}\left(X_i(x)\right)^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{1}{2}} \in L^p(X)\,,$$

and then

$$\left(\sum_i \mathbb{E}\left(Y_i(x)\right)^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{1}{2}} \in L^p(X) \text{ with } Y_i = X_i \mathbf{1}_{\{\|X_i\| \leq c\}}\,. \tag{4.3.4}$$

Now, using eq. 4.3.4 and condition (iii) let us prove that $\sum_i Y_i \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}}$ and $\sum_i Y_i \mathbf{1}_{\{|X_i| > \triangle((X_i)_i)\}}$ converge a.s.

Case $1 \leq p < 2$: By Hölder´s inequality, we have

$$\left(\mathbb{E} \sum_{i=n}^{m} \left(Y_i(x)\right)^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{p}{2}} \geq \mathbb{E} \left(\sum_{i=n}^{m} \left(Y_i(x)\right)^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{p}{2}}$$

[c] Remember that $\triangle$ depends also of a positive constant δ.

$$\geq \frac{1}{B_p}\mathbb{E}\left|\sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}}\right|^p \quad \text{(By theorem 2.1.11)}$$

Then, by eq. 4.3.4 and by Fubini´s theorem:

$$\mathbb{E}\int_X \left|\sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}}\right|^p d\mu \underset{n,m\to\infty}{\longrightarrow} 0\,,$$

so $\sum_i Y_i \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}}$ converges in probability and then by theorem 3.1.2 it converges a.s. On the other hand, using that $a^r + b^r \geq (a+b)^r$, if $a, b \geq 0$ and $0 < r < 1$, with $r = \frac{p}{2}$:

$$\mathbb{E}\left(\sum_{i=n}^{m} |Y_i|^p \mathbf{1}_{\{|X_i|>\triangle((X_i)_i)\}}\right) \geq \mathbb{E}\left(\sum_{i=n}^{m} |Y_i|^2 \mathbf{1}_{\{|X_i|>\triangle((X_i)_i)\}}\right)^{\frac{p}{2}}$$

$$\geq \frac{1}{B_p}\mathbb{E}\left|\sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i|>\triangle((X_i)_i)\}}\right|^p \quad \text{(By theorem 2.1.11)}$$

Then, by condition (iii) and by Fubini´s theorem:

$$\mathbb{E}\int_X \left|\sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i|>\triangle((X_i)_i)\}}\right|^p d\mu \underset{n,m\to\infty}{\longrightarrow} 0\,,$$

so $\sum_i Y_i \mathbf{1}_{\{|X_i|>\triangle((X_i)_i)\}}$ converges in probability and then by theorem 3.1.2 it converges a.s. Thus, $\sum_i Y_i$ converges a.s.

Case $2 \leq p < \infty$: By Rosenthal´s theorem 3.1.1 and Fubini´s theorem,

$$K_p\left(\int_X \sum_{i=n}^{m} \mathbb{E}|Y_i|^p \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} d\mu + \int_X \left(\mathbb{E}\sum_{i=n}^{m} (Y_i)^2\, \mathbf{1}_{\{|X_i|\leq\, \triangle((X_i)_i)\}}\right)^{\frac{p}{2}} d\mu\right) \tag{4.3.5}$$

$$\geq \mathbb{E}\int_X \left|\sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}}\right|^p d\mu\,,$$

Now, let us bound $\int_X \sum_{i=n}^{m} \mathbb{E}|Y_i|^p \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} d\mu$. First, note that

$$\mathbb{E}|Y_i|^p \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} \leq \mathbb{E}|Y_i|^2 \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} \left(\triangle((X_i)_i)\right)^{p-2}.$$

Then by Hölder´s inequality with exponent $\frac{p}{2}$ and conjugate exponent $\frac{p}{p-2}$:

$$\int_X \sum_{i=n}^{m} \mathbb{E}|Y_i|^p \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} d\mu \leq \int_X \sum_{i=n}^{m} \mathbb{E}|Y_i|^2 \mathbf{1}_{\{|X_i|\leq \triangle((X_i)_i)\}} (\triangle((X_i)_i))^{p-2} d\mu$$

$$\leq \left\| \left(\sum_{i=n}^{m} \mathbb{E}|Y_i|^2 \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right)^{\frac{1}{2}} \right\|_{L^p(X)}^2 \|\triangle((X_i)_i)\|_{L^p(X)}^{p-2} .$$

Then taking in account equation 4.3.5, again as in the previous case, we have that,

$$\mathbb{E} \int_X \left| \sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}} \right|^p d\mu \underset{n,m\to\infty}{\longrightarrow} 0 ,$$

so $\sum_i Y_i \mathbf{1}_{\{|X_i| \leq \triangle((X_i)_i)\}}$ converges in probability and then by theorem 3.1.2 it converges a.s. Now, we prove that $\sum_i Y_i \mathbf{1}_{\{|X_i| > \triangle((X_i)_i)\}}$ converges a.s. Again we use theorem 3.1.1:

$$K_p \left(\int_X \sum_{i=n}^{m} \mathbb{E}|Y_i|^p \mathbf{1}_{\{|X_i| > \triangle((X_i)_i)\}} d\mu + \int_X \left(\mathbb{E} \sum_{i=n}^{m} (Y_i)^2 \mathbf{1}_{\{|X_i| >, \triangle((X_i)_i)\}} \right)^{\frac{p}{2}} d\mu \right) \tag{4.3.6}$$

$$\geq \mathbb{E} \int_X \left| \sum_{i=n}^{m} Y_i \mathbf{1}_{\{|X_i| > \triangle((X_i)_i)\}} \right|^p d\mu ,$$

But in this case we will bound $\int_X \left(\mathbb{E} \sum_{i=n}^{m} (Y_i)^2 \mathbf{1}_{\{|X_i| >, \triangle((X_i)_i)\}} \right)^{\frac{p}{2}} d\mu$. Indeed, write $A_i = \{(x,\omega) : |X_i| >, \triangle((X_i)_i)\}$, so by Minkowski´s inequality

$$\int_X \left(\mathbb{E} \sum_{i=n}^{m} (Y_i)^2 \mathbf{1}_{\{|X_i| >, \triangle((X_i)_i)\}} \right)^{\frac{p}{2}} d\mu \, \|\triangle((X_i)_i)\|_{L^p(X)}^{\frac{(p-2)p}{2}}$$

$$\leq \left(\sum_{i=n}^{m} \mathbb{E} \left(\int_X |Y_i|^p \mathbf{1}_{A_i} d\mu \right)^{\frac{2}{p}} \right)^{\frac{p}{2}} \|\triangle((X_i)_i)\|_{L^p(X)}^{\frac{(p-2)p}{2}} = \left(\sum_{i=n}^{m} \mathbb{E} \left\| Y_i \mathbf{1}_{A_i} \right\|_{L^p(X)}^2 \right)^{\frac{p}{2}} \|\triangle((X_i)_i)\|_{L^p(X)}^{\frac{(p-2)p}{2}}$$

Since $|X_i| \geq \triangle((X_i)_i)$ over A_i, we have that

$$\|Y_i \mathbf{1}_{A_i}\|_{L^p(X)}^p \geq \|Y_i \mathbf{1}_{A_i}\|_{L^p(X)}^2 \|\triangle((X_i)_i)\|^{(p-2)} .$$

Then the last inequality can be bounded by: $\left(\sum_{i=n}^{m} \mathbb{E} \left\| Y_i \mathbf{1}_{A_i} \right\|_{L^p(X)}^p \right)^{\frac{p}{2}}$. But from condition (iii) this term tends to 0 whenever $n, m \longrightarrow \infty$. Then,

$$\sum_i Y_i \mathbf{1}_{\{|X_i| > \triangle((X_i)_i)\}} ,$$

converges in probability and then by theorem 3.1.2 it converges a.s. □

Remark. Case $0 < p < 1$: Note, that if $0 < p < 1$ the L^p spaces are metric spaces but not normed spaces. The proofs of the previous cases rely on corollary 3.1.1, which is a result for normed spaces. However, these general results for normed spaces can be modified for the case of the L^p metric. For further references about theorem 4.3.2 and its consequence, theorem 4.3.3, see section 4.5 at the end of this chapter.

4.3.2 a.s. convergence of stable sequences in L^p

In the following theorem, we consider series of the form $\sum_{i=1}^{\infty} a_i f_i$, where a_i are symmetric α-stable ($\alpha \in (0,2)$) independent random variables, and $f_i \in L^p(X, \Sigma, \mu)$.

Theorem 4.3.3. *Let $0 < p < \infty$, $\{f_j\}_{j\in\mathbb{N}} \subset L^p(X, \Sigma, \mu)$, and let $\{a_j\}_{j\in\mathbb{N}}$ be a sequence of real, independent, identically distributed, α-stable symmetric random variables. Then the series $\sum_{i=1}^{\infty} a_i f_i$ converges in $L^p(X, \Sigma, \mu)$ a.s. $\iff$*
i) In the case $\alpha = 2$ or $p < \alpha < 2$,

$$\left\| \left(\sum_{i=1}^{\infty} |f_i|^{\alpha} \right)^{1/\alpha} \right\|_{L^p(X,\Sigma,\mu)} < \infty \,. \tag{4.3.7}$$

ii) In the case $p > \alpha$ and $\alpha < 2$,

$$\sum_{i=1}^{\infty} \|f_i\|^{\alpha}_{L^p(X,\Sigma,\mu)} < \infty \,;$$

iii) In the case $p = \alpha$ and $\alpha < 2$,

$$\sum_{i=1}^{\infty} \int_X |f_i|^{\alpha} \left(1 + log^{+} \frac{|f_i|^{\alpha}}{\sum_j |f_j|^{\alpha} \, \|f_i\|^{\alpha}_{L^{\alpha}}} \right) d\mu < \infty$$

Moreover, the a.s. convergence of the series $\sum_{i=1}^{\infty} a_i f_i$ implies its convergence en the r-mean, $r < \alpha$.

Proof. The equivalence of the a.s. convergence and convergence in the r mean for $r < \alpha$ was established in corollary 3.1.5. We will use theorem 4.3.2, and we begin estimating $\triangle((X_i)_i)$ by $\left(\sum_{i=1}^{\infty} |f_i|^{\alpha} \right)^{1/\alpha}$. Indeed, first note, that if $\Psi(x) = x\mathbf{1}_{\{|x|\le 1\}} + \operatorname{sgn}(x)\mathbf{1}_{\{|x|>1\}}$, then $(\Psi(x))^2 = x^2\mathbf{1}_{\{|x|\le 1\}} + \mathbf{1}_{\{|x|>1\}}$. Thus,

$$\mathbb{E}\left(\Psi\left(\frac{a_i f_i}{s} \right) \right)^2 = f_i^2 \mathbb{E}\left(\frac{a_i}{s} \right)^2 \mathbf{1}_{\{|a_i f_i| \le s\}} + \mathbf{P}(|a_i f_i| > s) \,.$$

Recalling the asymptotic properties of stable distributions from chapter 2, section 2.4.4, we have, that for some positive constants $A, B > 0$:

$$A \frac{|f_i|^{\alpha-2}}{s^{\alpha}} \le \mathbb{E}\left(\frac{a_i}{s} \right)^2 \mathbf{1}_{\{|a_i f_i| \le s\}} \le B \frac{|f_i|^{\alpha-2}}{s^{\alpha}} \,,$$

and, for some $A', B' > 0$,

$$A' \frac{|f_i|^{\alpha}}{s^{\alpha}} \le \mathbf{P}(|a_i f_i| > s) \le B' \frac{|f_i|^{\alpha}}{s^{\alpha}} \,,$$

if $\frac{s}{|f_i|} \geq s_0$, for some s_0 sufficiently large. So,

$$\sum_{i=1}^{\infty} f_i^2 A \frac{|f_i|^{\alpha-2}}{s^\alpha} + A' \frac{|f_i|^\alpha}{s^\alpha} \leq \sum_{i=1}^{\infty} \mathbb{E}\left(\Psi\left(\frac{a_i f_i}{s}\right)\right)^2 \leq \sum_{i=1}^{\infty} f_i^2 B \frac{|f_i|^{\alpha-2}}{s^\alpha} + B' \frac{|f_i|^\alpha}{s^\alpha},$$

then, from the definition of $\triangle((a_i f_i)_i)$, we have that for an appropriate small δ, given $\epsilon > 0$, there exists $s = s(x)$ $(x \in X)$, such that $\triangle((a_i f_i)_i)(x) \leq s(x) < \triangle((a_i f_i)_i)(x) + \epsilon$ and for some constant $C > 0$: $C \sum_{i=1}^{\infty} |f_i(x)|^\alpha \leq \delta s^\alpha(x)$. But if $\triangle((a_i f_i))(x) > s(x) - \epsilon$, $C^{\frac{1}{\alpha}} \|(f_i)_i\|_{l^\alpha} > \delta^{\frac{1}{\alpha}}(s(x) - \epsilon)$. Thus

$$\|(f_i)_i\|_{l^\alpha} \in L^p(X, \Sigma, \mu) \iff \triangle((a_i f_i)_i) \in L^p(X, \Sigma, \mu)\,, \tag{4.3.8}$$

with $\alpha \in (0,2)$ and $0 < p < \infty$. In particular we have from theorem 4.3.2 that,

$$\sum_{i=1}^{\infty} a_i f_i \text{ converges a.s.} \Longrightarrow \|(f_i)_i\|_{l^\alpha} \in L^p(X, \Sigma, \mu)\,. \tag{4.3.9}$$

Now, let us consider different separate cases:

Case $p < \alpha < 2$: Let $Y_n = \sum_{i=1}^{n} a_i f_i$, then since the $a_i's$ are independent symmetric α-stable random variables: there exists a positive constant K, such that for every x (again recall chapter 2, section 2.4.4), $n > m$:

$$\mathbb{E}|Y_n(x) - Y_m(x)|^p = K \left(\sum_{i=m}^{n} |f_i(x)|^\alpha\right)^{\frac{p}{\alpha}}.$$

So, integrating in the variable x:

$$\mathbb{E}\,\|Y_n - Y_m\|^p = K \int_X \left(\sum_{i=m}^{n} |f_i(x)|^\alpha\right)^{\frac{p}{\alpha}} d\mu(x)\,.$$

But if $\|(f_i)_i\|_{l^\alpha} \in L^p(X, \Sigma, \mu)$ then $\int_X \left(\sum_{i=m}^{n} |f_i(x)|^\alpha\right)^{\frac{p}{\alpha}} d\mu(x) \underset{n,m\to\infty}{\longrightarrow} 0$. Thus Y_n converges in the p-mean and then in probability but by theorem 3.1.2 this is equivalent to a.s. convergence. This, together with the implication 4.3.9 proved before concludes the proof of the theorem for this case.

Case $\alpha < 2$, $\alpha < p$: Condition i) of theorem 4.3.2 reads as follows, for $\alpha \neq 2$:

$$\sum_{i=1}^{\infty} \|f_i\|^\alpha_{L^p(X,\Sigma,\mu)} < \infty\,, \tag{4.3.10}$$

Indeed, this follows from the estimate for the i.i.d. stable a_i's :

$$A \frac{\|f_i\|^\alpha_{L^p(X)}}{c^\alpha} \leq \mathbf{P}(|a_i| > c\,\|f_i\|^{-1}_{L^p(X)}) \leq B \frac{\|f_i\|^\alpha_{L^p(X)}}{c^\alpha}$$

for large enough c.
In particular, in view of theorem 4.3.2, we have that eq. 4.3.10 is a necessary condition for the a.s. convergence of $\sum_{i=1}^{\infty} a_i f_i$. On the other hand, by Minkowski's inequality, since $\frac{p}{\alpha} > 1$:

$$\sum_{i=1}^{\infty} \|f_i\|^{\alpha}_{L^p(X,\Sigma,\mu)} = \sum_{i=1}^{\infty} \left(\int_X |f_i|^p d\mu \right)^{\frac{\alpha}{p}} \geq \left(\int_X \left(\sum_{i=1}^{\infty} |f_i|^{\alpha} \right)^{\frac{p}{\alpha}} d\mu \right)^{\frac{\alpha}{p}} . \tag{4.3.11}$$

Then condition 4.3.10 together with eq. 4.3.8 implies condition ii) of theorem 4.3.2. Now, let us prove that condition iii) of theorem 4.3.2 is verified. In this case condition iii) reads

$$\sum_{i=1}^{\infty} \int_X |f_i|^p \mathbb{E}|a_i|^p \mathbf{1}_{A_i} d\mu \,, \tag{4.3.12}$$

where

$$A_i(x) = \left\{ \omega \in \Omega : \frac{\|(f_i(x))\|_{l^\alpha}}{|f_i(x)|} < |a_i| \leq \frac{c}{\|f_i\|_{L^p}} \right\} .$$

If $\alpha \neq 2$ and $\alpha \neq p$, then we majorize the sum in 4.3.12 by

$$\sum_{i=1}^{\infty} \int_X |f_i|^p \mathbb{E}|a_i|^p \mathbf{1}_{\{|a_i| \leq \frac{c}{\|f_i\|_{L^p}}\}} d\mu \,,$$

and since the $a_i' s$ are α-stable and $p > \alpha$, we have the following estimate on the distributions,

$$A \left(\frac{c}{\|f_i\|_{L^p}} \right)^{p-\alpha} \leq \mathbb{E}|a_i|^p \mathbf{1}_{\{|a_i| \leq \frac{c}{\|f_i\|_{L^p}}\}} \leq B \left(\frac{c}{\|f_i\|_{L^p}} \right)^{p-\alpha} .$$

Then the sum 4.3.12 is bounded by:

$$\sum_{i=1}^{\infty} \int_X |f_i|^p \mathbb{E}|a_i|^p d\mu \; \|f_i\|^{\alpha-p}_{L^p} = C \sum_{i=1}^{\infty} \|f_i\|^{\alpha}_{L^p} \;\; ,$$

thus this concludes the proof the theorem for this case.

Case $p = \alpha < 2$: If $\alpha < 2$ and $p = \alpha$ then, in a similar way to the previous cases, but using the property 5 of chapter 2, section 2.4.4, since the a_i´s are α-stable, the sum 4.3.12 can be estimated by

$$C \sum_{i=1}^{\infty} |f_i(x)|^p log^+ \frac{|f_i(x)|}{\|f_i\|_{L^p} \, \|(f_i(x))_i\|_{l^\alpha}} \,,$$

since $\|(f_i(x))_i\|_{l^\alpha} \leq |f_i(x)|$ and conditions 4.3.10 and 4.3.7 coincide with the condition

$$\sum_{i=1}^{\infty} \int_X |f_i(x)|^p d\mu < \infty \,,$$

which proves the corollary in this case. Finally, if $\alpha = 2$ then condition 4.3.7 implies 4.3.10 and 4.3.12. Indeed, since for $r \geq \max\{2, \text{p}\}$,

$$\left(\sum_{i=1}^{\infty} \|f_i\|_{L^p}^r\right)^{\frac{1}{r}} \leq \left\|\left(\sum_i |f_i|^2\right)^{\frac{1}{2}}\right\|_{L^p}$$

since the a_i's are stable, condition 4.3.7 implies 4.3.10. If $p < 2$, we estimate the sum in the condition 4.3.12, by

$$\sum_{i=1}^{\infty} \int_X |f_i(x)|^p \mathbb{E}|a_i|^p \mathbf{1}_{\{\|(f_i(x))_i\|_{l^2} < |a_i f_i(x)|\}} d\mu$$

$$\leq C \sum_{i=1}^{\infty} \int_X \frac{|f_i(x)|^p \, \|(f_i(x))_i\|_{l^2}^{2-p}}{|f_i(x)|^{2-p}} d\mu = C \left\|\left(\sum_i |f_i|^2\right)^{\frac{1}{2}}\right\|_{L^p}^p$$

the follows from $\mathbb{E}|a_i|^p \mathbf{1}_{\{|a_i| \leq t\}} \leq \mathbb{E}|a_i|^p$ and since $\|(f_i(x))_i\|_{l^2} \geq |f_i(x)|$. As a result, we obtain that 4.3.7 implies 4.3.10. If $p \geq 2$, then the sum in 4.3.12 is estimated by

$$\mathbb{E}|a_1|^p \sum_{i=1}^{\infty} |f_i|^p d\mu = \mathbb{E}|a_1|^p \sum_{i=1}^{\infty} \|f_i\|_{L^p}^p \leq \mathbb{E}|a_1|^p \left\|\left(\sum_i |f_i|^2\right)^{\frac{1}{2}}\right\|_{L^p}^p .$$

This concludes the proof of the corollary in the case $\alpha = 2$. □

4.3.3 Series expansions with respect to an Unconditional Basis. Convergence in p-th mean and a.s. almost everywhere convergence.

Unconditional basic sequences are very important in the theory of Banach spaces. Another concept which is related with them is l^p stability and both are important topics in wavelet analysis, shift invariant subspaces and their applications such as sampling of signals. It is also interesting to study the pointwise, almost everywhwere convergence properties of these expansions. Here, we will consider random series of the form

$$\sum_{i=1}^{\infty} a_i f_i \,, \tag{4.3.13}$$

where the a_i's are independent (or not) zero mean random variables and the f_i's are an unconditional basic sequence or an l^p stable sequence in a Lebesgue space $L^p(X, \Sigma, \mu)$, where $p \in [1, \infty)$. In this section we give conditions under which, if one of these series 4.3.13 converges in the norm topology almost surely then converges almost everywhere almost surely and converges in the p mean.

We begin with a concept from the theory of Banach spaces. l^p stability is defined as an equivalence of norms :

Definition 20. [25] Let $(E, \|.\|)$ be a Banach space, then $\{f_j\}_{j \in \mathbb{N}} \subset E$ is a l^p-stable sequence ($p \in [1, \infty)$), if there exist positive constants c_p and K_p, such that:

$$c_p \|a\|_{l^p} \leq \left\|\sum_i a_i f_i\right\| \leq K_p \|a\|_{l^p} \,, \quad \forall \, a \in l_0 \,.$$

Where l_0 is the linear space of all real (or complex) valued sequences which have all its coordinates equal to 0 but for a finite set of indexes.

l^p- stable sequences provide examples of bounded unconditional basic sequences. That is, if the $\{f_j\}_j$ is a l^p-stable sequence then $0 < \inf_j \|f_j\| \leq \sup_j \|f_j\| < \infty$ and it is a basis. Indeed, the first assertion is obvious and for the second take $f \in \overline{\text{span}}\{f_j\}_j$, then for each $n \in \mathbb{N}$ there exists a sequence $a_n = (a_{j\,n})_{j\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, with zero coordinates except for a finite set of indexes j, such that $\left\| f - \sum_{j\in\mathbb{N}} a_{j\,n} f_j \right\| < \frac{1}{n}$. We shall see, that the sequence $(a_n)_n$ is a Cauchy sequence in l^p. Note, that from the definition, and the triangle inequality:

$$c_p \sum_{j\in\mathbb{N}} |a_{j\,n} - a_{j\,m}|^p \leq \left\| \sum_{j\in\mathbb{N}} a_{j\,m} f_j - \sum_{j\in\mathbb{N}} a_{j\,n} f_j \right\|^p < \frac{1}{n} + \frac{1}{m}\,.$$

Then there exists $b = (b_j)_j \in l^p$ such that $a_n \longrightarrow b$ in l^p. Let us verify that $\sum_j b_j f_j$ converges in $(E, \|\,.\,\|)$. Given $n \in \mathbb{N}$: $\exists\, j_0$, $a_{j\,n} = 0$ for all $j \geq j_0$. Taking $m \geq k \geq j_0$:

$$\left\| \sum_{j=k}^{m} b_j f_j \right\| = \left\| \sum_{j=k}^{m} (b_j - a_{j\,n}) f_j \right\| \leq K_p \|b - a_n\|_{l^p}\ .$$

Then $\sum_j b_j f_j$ converges in $(E, \|\,.\,\|)$. On the other hand,

$$\left\| \sum_{j\in\mathbb{N}} (b_j - a_{j\,n}) f_j \right\| \leq K_p \|b - a_n\|_{l^p} \longrightarrow 0\,.$$

Hence $f = \sum_j b_j f_j$. The uniqueness of this representation follows again from the definition of the l_p stability. Then the $\{f_j\}_j$ is a basis of $\overline{\text{span}}\{f_j\}_j$.

Recalling the characterization of unconditional bases in $L^p(X)$ spaces of theorem 2.3.2, we can prove the first result of this section: a kind of analogue of a result of Paley and Zygmund.

Proposition 4.3.1. *a) Let $\{f_j\}_{j\in\mathbb{N}} \subset E$ be an l^p-stable sequence, $0 < \lambda < 1$; and $\{a_j\}_{j\in\mathbb{N}}$ a sequence of random variables such that there exist a constant $C > 0$ and $r \in (1, \infty)$ such that $\mathbb{E}|a_j|^{rp} \leq C(\mathbb{E}|a_j|^p)^r$, $\forall\, j$, then equation 4.3.14 holds.*
b) Let $\{f_j\}_{j\in\mathbb{N}} \subset L^p(X, \Sigma, \mu)$, $(\infty > p \geq 2)$ be basic unconditional sequence, $0 < \lambda < 1$; and $\{a_j\}_{j\in\mathbb{N}}$ a sequence of random variables such that there exists a constant $C > 0$ such that $\mathbb{E}|a_j|^{2p} \leq C(\mathbb{E}|a_j|^2)^p$, $\forall\, j$, then equation 4.3.14 holds with $r' = 2$. If $1 \leq p < 2$, the last assertion remains true with the additional condition: $(\mathbb{E}|a_i|^p)^{\frac{1}{p}} \geq c(\mathbb{E}|a_i|^2)^{\frac{1}{2}}$.

$$\mathbf{P}\left(\left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p > \lambda \mathbb{E} \left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p \right) \geq (1-\lambda)^{r'} k \qquad (4.3.14)$$

where k is a positive constant independent of n and $\frac{1}{r} + \frac{1}{r'} = 1$.

Remark. The hypothesis $\mathbb{E}|a_j|^{rp} \leq C(\mathbb{E}|a_j|^p)^r \ \forall j$ may look artificial, but this regularity condition is necessary in order to control the values of the a_i's. Similar conditions can be found in [40], [18], [62] and [63] dealing, for example, with random Fourier series.

Proof. Part a) First, by lemma 3.1.1 we have:

$$\mathbf{P}\left(\left\|\sum_{j=1}^n a_j\, f_j\right\|^p > \lambda \mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^p\right) \geq (1-\lambda)^{r'} \frac{\left(\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^p\right)^{r'}}{\left(\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^{pr}\right)^{\frac{r'}{r}}}, \tag{4.3.15}$$

on the other hand, by the definition of l^p-stability and Minkowski's inequality:

$$\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^{rp} \leq K_p^{rp}\, \mathbb{E}\left(\sum_{j=1}^n |a_j|^p\right)^r \leq K_p^{rp}\left(\sum_{j=1}^n (\mathbb{E}|a_j|^{pr})^{\frac{1}{r}}\right)^r,$$

using the condition $\mathbb{E}|a_j|^{rp} \leq C(\mathbb{E}|a_j|^p)^r \ \forall j$, one gets,

$$\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^{rp} \leq K_p^{rp} C \left(\sum_{j=1}^n \mathbb{E}|a_j|^p\right)^r \tag{4.3.16}$$

Clearly, from eq. 4.3.16 :

$$\frac{\left(\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^p\right)^{r'}}{\left(\mathbb{E}\left\|\sum_{j=1}^n a_j\, f_j\right\|^{pr}\right)^{\frac{r'}{r}}} \geq \frac{c_p^{pr'}\left(\sum_{j=1}^n \mathbb{E}|a_j|^p\right)^{r'}}{K_p^{r'p} C^{\frac{r'}{r}}\left(\sum_{j=1}^n \mathbb{E}|a_j|^p\right)^{r'}}.$$

This, together with equation 4.3.15 implies the desired result.

Part b) To bound $\mathbb{E}\left\|\left(\sum_{i=1}^n |a_i f_i|^2\right)^{\frac{1}{2}}\right\|^p$ we must consider two separate cases: first $\infty > p \geq 2$ and then $1 \leq p \leq 2$. The rest of the proof is valid for all $\infty > p \geq 1$.

<u>If $p \geq 2$</u> then,

$$\mathbb{E}\left\|\left(\sum_{i=1}^n |a_i f_i|^2\right)^{\frac{1}{2}}\right\|^p = \int_\Omega \int_X \left(\sum_{i=1}^n |a_i f_i|^2\right)^{\frac{p}{2}} d\mu\, d\mathbf{P}$$

$$= \int_X \int_\Omega \left(\sum_{i=1}^n |a_i f_i|^2\right)^{\frac{p}{2}} d\mathbf{P} d\mu = \int_X \mathbb{E}\left(\sum_{i=1}^n |a_i f_i|^2\right)^{\frac{p}{2}} d\mu\,. \tag{4.3.17}$$

But by Hölder's inequality $\mathbb{E}\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{p}{2}} \geq \left(\mathbb{E}\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{p}{2}}$ and clearly, from this and 4.3.17 we have:

$$\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{1}{2}}\right\|^p \geq \left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^p . \tag{4.3.18}$$

Now, <u>if $1 \leq p < 2$</u> as a direct consequence of Minkowski's integral inequality and Fubini's theorem:

$$\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{1}{2}}\right\|^p = \int_X \mathbb{E}\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{p}{2}} d\mu \geq \int_X \left(\sum_{i=1}^{n}(\mathbb{E}|a_i|^p)^{\frac{2}{p}}|f_i|^2\right)^{\frac{p}{2}} d\mu$$

$$\geq c\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^p . \tag{4.3.19}$$

On the other hand,

$$\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{1}{2}}\right\|^{2p} = \int_\Omega\left(\int_X\left(\sum_{i=1}^{n}|a_if_i|^2\right)^{\frac{p}{2}} d\mu\right)^2 d\mathbf{P} . \tag{4.3.20}$$

If we define $g(x,\omega) = \left(\sum_{i=1}^{n}|a_i(\omega)f_i(x)|^2\right)^{\frac{p}{2}}$, then by, Minkowski's inequality we have the following bound on 4.3.20:

$$\int_\Omega\left(\int_X g(x,\omega)d\mu\right)^2 d\mathbf{P} \leq \left(\int_X\left(\int_\Omega |g(x,\omega)|^2 d\mathbf{P}\right)^{\frac{1}{2}} d\mu\right)^2 .$$

Now, $\int_\Omega |g(x,\omega)|^2 d\mathbf{P} = \mathbb{E}\left(\sum_{i=1}^{n}|a_if_i(x)|^2\right)^p$ then by the triangle inequality:

$$\mathbb{E}\left(\sum_{i=1}^{n}|a_if_i(x)|^2\right)^p \leq \left(\sum_{i=1}^{n}\left(\mathbb{E}|a_if_i(x)|^{2p}\right)^{\frac{1}{p}}\right)^p \tag{4.3.21}$$

$$= \left(\sum_{i=1}^{n}(\mathbb{E}|a_i|^{2p})^{\frac{1}{p}}|f_i(x)|^2\right)^p \leq C\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i(x)|^2\right)^p , \tag{4.3.22}$$

where the last inequality follows from $\mathbb{E}|a_j|^{2p} \leq C(\mathbb{E}|a_j|^2)^p,\ \forall\, j$.
Hence:

$$\int_X\left(\int_\Omega |g(x,\omega)|^2 d\mathbf{P}\right)^{\frac{1}{2}} d\mu \leq C^{\frac{1}{2}}\int_X\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{p}{2}} d\mu = C^{\frac{1}{2}}\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|_{L^p(X)}^p ,$$

and from this it is immediate that:

$$\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_i f_i|^2\right)^{\frac{1}{2}}\right\|^{2p} \leq C\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{2p}. \tag{4.3.23}$$

By equations 4.3.23 and 4.3.18 or 4.3.19 we have the following bounds:

$$\mathbb{E}\left\|\sum_{j=1}^{n}a_j\, f_j\right\|^{p} \geq A_p^p\,\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_i f_i|^2\right)^{\frac{1}{2}}\right\|^{p} \geq k_p A_p^p\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{p}, \tag{4.3.24}$$

and:

$$\mathbb{E}\left\|\sum_{j=1}^{n}a_j\, f_j\right\|^{2p} \leq B_p^{2p}\,\mathbb{E}\left\|\left(\sum_{i=1}^{n}|a_i f_i|^2\right)^{\frac{1}{2}}\right\|^{2p} \leq CB_p^{2p}\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{2p}. \tag{4.3.25}$$

Recall 4.3.15 with $r = 2$, so as from 4.3.24 and 4.3.25, we have:

$$\frac{\left(\mathbb{E}\left\|\sum_{j=1}^{n}a_j\, f_j\right\|^{p}\right)^2}{\mathbb{E}\left\|\sum_{j=1}^{n}a_j\, f_j\right\|^{2p}} \geq \frac{A_p^{2p}\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{2p}}{B_p^{2p}C\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{2p}},$$

and then we get the desired result. □

4.3.4 Convergence in the p mean and almost sure $[\mu]$-a.e. convergence.

First, let us note that if

$$\sup_{n\in\mathbb{N}}\left\|\left(\sum_{i=1}^{n}\mathbb{E}|a_i|^2|f_i|^2\right)^{\frac{1}{2}}\right\|^{p} < \infty,$$

under the conditions of proposition 4.3.1 part b), then it is rather straightforward to see that $S_n = \sum_{i=1}^{n} a_i f_i$ is a Cauchy sequence in $L^p(X\times\Omega)$ (equations 4.3.29 and 4.3.30) and since for $\lambda > 0$: $\mathbf{P}(\|S_n - S_m\| > \lambda) \leq \frac{\mathbb{E}\|S_n-S_m\|^p}{\lambda^p}$, then it is a Cauchy sequence in probability but if the a_i's are independent then convergence in probability of sums of independent Banach space valued random implies a.s. convergence (By theorem 3.1.2). A similar argument holds for the case of l^p-stable sequences. Next we prove a result which is a partial converse of this fact.

Proposition 4.3.2. *a) Let $\{f_j\}_{j\in\mathbb{N}}$ be a l^p-stable sequence and $\{a_j\}_{j\in\mathbb{N}}$ be a sequence of random variables such that there exist a constant $C > 0$ and $r \in (1,\infty)$ such that $\mathbb{E}|a_j|^{rp} \leq C(\mathbb{E}|a_j|^p)^r\ \forall, j$; then if $\sum_{i=1}^{\infty} a_i f_i$ converges in the norm topology of E a.s. then*

$$\sum_{i=1}^{\infty}\mathbb{E}|a_i|^p < \infty.$$

b) Let $\{f_j\}_{j\in\mathbb{N}} \subset L^p(X,\Sigma,\mu)$, $(p \geq 2)$ be a basic unconditional sequence and $\{a_j\}_{j\in\mathbb{N}}$ a sequence of independent random variables such that there exists a constant $C > 0$ $\mathbb{E}|a_j|^{2p} \leq C(\mathbb{E}|a_j|^2)^p \;\forall\, j$; if $\sum\limits_{i=1}^{\infty} a_i f_i$ converges in the norm topology of $L^p(X,\Sigma,\mu)$ a.s. then

$$\left\| \left(\sum_{i=1}^{\infty} \mathbb{E}|a_i|^2 |f_i|^2 \right)^{\frac{1}{2}} \right\|^p < \infty \, .$$

If $1 \leq p < 2$, the last assertion remains true with the additional condition: $(\mathbb{E}|a_i|^p)^{\frac{1}{p}} \geq c(\mathbb{E}|a_i|^2)^{\frac{1}{2}}$.
In particular, a) or b) imply that $\sum\limits_{i=1}^{\infty} a_i f_i$ converges in the p-mean.

Proof. part a) Take $\lambda \in (0,1)$, define:

$$D_n = \left\{ \omega \in \Omega : \left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p > \lambda \mathbb{E} \left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p \right\}$$

and define:

$$D = \overline{\lim_{n\to\infty}} D_n = \bigcap_{p=1}^{\infty} \bigcup_{n=p}^{\infty} D_n \, , \tag{4.3.26}$$

By proposition 4.3.1 $\exists\, k > 0$ such that $\mathbf{P}(D_n) \geq k(1-\lambda)^{r'}$ for all n, where $\frac{1}{r} + \frac{1}{r'} = 1$, but $\mathbf{P}(\overline{\lim\limits_{n\to\infty}} D_n) \geq \overline{\lim\limits_{n\to\infty}} \mathbf{P}(D_n) \geq k(1-\lambda)^{r'} > 0$ then, $\mathbf{P}(D) > 0$.
From this last fact: $D \bigcap \{\omega \in \Omega : \sum\limits_i a_i f_i$ *converges in* $E\} \neq \oslash$, equivalently $\exists\, \omega \in D$ such that $\sum\limits_i a_i(\omega) f_i$ converges in $(E, \|.\|)$ and this implies: $\exists\, M > 0$ such that

$$\sup_{n\in\mathbb{N}} \left\| \sum_{j=1}^{n} a_j(\omega)\, f_j \right\|^p \leq M \, .$$

By equation 4.3.26 there exist infinetly many $n's$, such that for this $\omega \in D$:

$$\infty > M \geq \left\| \sum_{j=1}^{n} a_j(\omega)\, f_j \right\|^p > \lambda \mathbb{E} \left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p \geq \lambda c_p^p \sum_{i=1}^{n} \mathbb{E}|a_i|^p \tag{4.3.27}$$

$\Longrightarrow \sum\limits_{i=1}^{\infty} \mathbb{E}|a_i|^p < \infty$, and the proof of a) is complete.
part b) The proof is almost the same as for part a) but with $r = r' = 2$. Instead of the bound 4.3.27, recalling the bound of equation 4.3.24 we have:

$$\infty > M \geq \left\| \sum_{j=1}^{n} a_j(\omega)\, f_j \right\|^p > \lambda \mathbb{E} \left\| \sum_{j=1}^{n} a_j\, f_j \right\|^p \geq \lambda k_p A_p^p \left\| \left(\sum_{i=1}^{n} \mathbb{E}|a_i|^2 |f_i|^2 \right)^{\frac{1}{2}} \right\|^p \tag{4.3.28}$$

for infinetly many $n's$ and, then by Beppo Levi's theorem:

$$\infty > \lim_{n\to\infty} \left\| \left(\sum_{i=1}^{n} \mathbb{E}|a_i|^2|f_i|^2 \right)^{\frac{1}{2}} \right\|^p = \left\| \left(\sum_{i=1}^{\infty} \mathbb{E}|a_i|^2|f_i|^2 \right)^{\frac{1}{2}} \right\|^p .$$

The convergence in the mean follows from eq. 4.3.23 of the previous result. □

Almost Sure a.e. $[\mu]$ convergence

Now, we relate the previous result with a.e. convergence: let us give a result on convergence over the product space. However, some additional conditions on the random coefficients a_i's, such as independence, should be added.

Theorem 4.3.4. *a) Let $\{a_j\}_{j\in\mathbb{N}}$ be a sequence of independent random variables such that there exist a constant $C > 0$ and $r \in (1,\infty)$ such that $\mathbb{E}|a_j|^{rp} \leq C(\mathbb{E}|a_j|^p)^r$, $\forall j$, and $\mathbb{E}(a_j) = 0$; and let $\{f_j\}_{j\in\mathbb{N}} \subset L^p(X,\Sigma,\mu)$, $1 \leq p < \infty$, be a l^p-stable sequence. If $\sum_{i=1}^{\infty} a_i f_i$ converges in $L^p(X,\Sigma,\mu)$ a.s. $\Longrightarrow \sum_{i=1}^{\infty} a_i f_i$ converges $[\mu]$ almost everywhere a.s.*
b) Let $\{a_j\}_{j\in\mathbb{N}}$ be a sequence of independent random variables such that there exists a constant $C > 0$ such that $\mathbb{E}|a_j|^{2p} \leq C(\mathbb{E}|a_j|^2)^p$, $\forall j$, and $\mathbb{E}(a_j) = 0$; and let $\{f_j\}_{j\in\mathbb{N}} \subset L^p(X,\Sigma,\mu)$ be an unconditional basic sequence. If $\sum_{i=1}^{\infty} a_i f_i$ converges in $L^p(X,\Sigma,\mu)$ a.s. $\Longrightarrow \sum_{i=1}^{\infty} a_i f_i$ converges $[\mu]$-almost everywhere a.s. If $1 \leq p < 2$, the last assertion remains true with the additional condition: $(\mathbb{E}|a_i|^p)^{\frac{1}{p}} \geq c(\mathbb{E}|a_i|^2)^{\frac{1}{2}}$.

Proof. Under the hypothesis of a) or b) , we can see that $\sum_{i=1}^{n} a_i f_i$ is a Cauchy sequence in $L^p(X \times \Omega)$. Then both assertions will follow as a consequence of theorem 4.3.1:
Part a) We have:

$$\mathbb{E} \left\| \sum_{i=m}^{n} a_i f_i \right\|^p \leq K_p^p \, \mathbb{E} \sum_{i=m}^{n} |a_i|^p = K_p^p \sum_{i=m}^{n} \mathbb{E}|a_i|^p \longrightarrow 0 \,, \tag{4.3.29}$$

when $n, m \longrightarrow \infty$, as a consequence of proposition 4.3.2 part a).
Part b) Again, from Hölder's inequality and equation 4.3.25:

$$\mathbb{E} \left\| \sum_{i=m}^{n} a_i f_i \right\|^p \leq \left(\mathbb{E} \left\| \sum_{i=m}^{n} a_i f_i \right\|^{2p} \right)^{\frac{1}{2p}} \leq C^{\frac{1}{2p}} B_p \left\| \left(\sum_{i=m}^{n} \mathbb{E}|a_i|^2|f_i|^2 \right)^{\frac{1}{2}} \right\| \longrightarrow 0 \,, \tag{4.3.30}$$

when $n, m \longrightarrow \infty$, since $\sum_{i=1}^{n} \mathbb{E}|a_i|^2|f_i|^2$ is a Cauchy sequence in $L^{\frac{p}{2}}(X,\Sigma,\mu)$ as a consequence of proposition 4.3.2 part b).

□

As a consequence we obtain that almost all permutations of sign of the coefficients of expansions using unconditional basis converges a.e. $[\mu]$:

Theorem 4.3.5. *Let $f \in Span\{f_j\}_{j\in\mathbb{N}} \subset L^p(X,\Sigma,\mu)$ with (X,Σ,μ) σ-finite, $\{\theta_j\}_j$ a sequence of i.i.d. r.v's taking values in $\{+1,-1\}$ with equal probability, and $\{f_j\}_{j\in\mathbb{N}}$ an unconditional basic (l_p stable) sequence, if $f = \sum_i a_i f_i$ is the expansion of f in this basis then the random series $\sum_i \theta_i a_i f_i$ converges a.e. $[\mu]$ a.s.*

Proof. This result a direct application of theorem 4.3.4 above and the definition of unconditional basic (l^p stable) sequence. □

It is remarkable that the case $p = 2$ for theorem 4.3.4 can be easily derived from the results in chapter 3 using theorems 3.1.4 and 3.1.5, with no more assumptions on the sequence $\{f_j\}_{j\in\mathbb{N}}$ than the convergence in norm of the series 4.3.13. This is a consequence of the fact that $L^2(X,\Sigma,\mu)$ is a Hilbert space, and that the independence of the a_i's makes the $a_i f_i$ behave in some sense as orthogonal elements, moreover they are independent and have O mean (in $L^2(X\times\Omega) = L^2(\Omega,\mathcal{F},\mathbf{P},L^2(X))$). Since unconditional basis are good basis and keep some of the properties of orthogonal basis, it is reasonable that this result can be extended to the case $p \neq 2$ when $\{f_j\}_{j\in\mathbb{N}}$ is an unconditional basis. Finally, as an example of application: consider the case of a wavelet basis $\{\psi_{j\,k}\}_{j\,k\in\mathbb{Z}\times\mathbb{Z}}$ of $L^p(\mathbb{R})$, $\infty > p > 1$ with the usual Lebesgue measure. It is known that under some conditions on the mother wavelet ψ then $\{\psi_{j\,k}\}_{j\,k}$ is an unconditional basis of $L^p(\mathbb{R})$ (for example see [80] or similar references). Take an arbitrary $f \in L^p(\mathbb{R})$ and $\theta_{j\,k}$ i.i.d. r.v.'s taking values in $\{1,-1\}$ with equal probability, then if we consider the expansion of f in this basis, $f \sim S(f) = \sum_{j\,k\in\mathbb{Z}\times\mathbb{Z}} < f,\psi_{j\,k} > \psi_{j\,k}$, we can construct the following random series in $L^p(\mathbb{R})$: $X(\omega) = \sum_{j\,k\in\mathbb{Z}\times\mathbb{Z}} \theta_{j\,k}(\omega) < f,\psi_{j\,k} > \psi_{j\,k}$ as $\{\psi_{j\,k}\}_{j\,k}$ is an unconditional basis. This series converges in $L^p(\mathbb{R})$ for every $\omega\in\Omega$. Then, we have convergence in norm almost surely, but we also have a.e. convergence a.s. as a consequence of theorem 4.3.5. Hence, $S(f)$ is just *one* realization of this random element defined by this series, but one may intuitively expect this series to converge not only in the norm of $L^p(\mathbb{R})$ but also almost everywhere. Despite these facts it can be shown that the exceptional set of zero probability is not necessarily void for an arbitrary unconditional basic sequence.Moreover consider the following deep result: orthonormal basis in a Hilbert space are unconditional basis, but Menchoff [59] showed that if (X,Σ,μ) is $[0,1]$ with Lebesgue Measure then there exists an orthonormal basis $\{f_j\}_{j\in\mathbb{N}}$ of $L^2[0,1]$ and an $f_0 \in L^2[0,1]$ such that the sequence $P_k f_0$ of projections of f_0 on the subspaces spanned by $\{f_1,...,f_k\}$ diverges a.e..

4.4 Some Applications and considerations.

4.4.1 About the representation of continuous parameter processes without loss of information.

In the previous section 4.3.3 me have studied several relationships and conditions between different types of convergences for series of the form 4.3.13. However, when studying a series representation of a process by a series relative to a basis, one may be concerned if the coefficients a_i's carry all the information about the whole process. This situation may be the following, given a process $\{Y_x\}_{x\in\mathbb{R}^d}$ with sample paths in $L^p(\mathbb{R}^d,d\mu)$ with an unconditional basis $\{f_i\}_i$, where μ is some equivalent measure to Lebesgue measure, in some applications such as detection theory, one wishes to express a likelihood ratio or a posteriori probability , given the sample path $\{Y_x\}_{x\in\mathbb{R}^d}$ as the limit of the corresponding n-dimensional quantity, given n functionals of Y. This is granted if the σ-algebras $\sigma((Y_x)_x)$ and $\sigma((a_i)_i)$ coincides. Let us discuss briefly this problem under the assumptions of proposition 4.3.2. In many applications the process is assumed to be continuous in probability. Let us see that under this condition $\sigma((Y_x)_x) = \sigma((a_i)_i)$. Indeed, if $D = \{x_i\}_i$ is any countable dense subset of $\mathbb{R}^d$ is easy to verify that if Y_x is continuous in probability, then $\sigma((Y_{x_i})_i) = \sigma((Y_x)_x)$. Under the assumptions of proposition 4.3.2, we have that $\mathbb{E}\left\|Y - \sum_{i=1}^{n} a_i f_i\right\|^p \underset{n\to\infty}{\longrightarrow} 0$ and then this implies convergence in measure, with respect to the product, i.e. for every $\epsilon > 0$: $\lim_{n\to\infty} \mu\times\mathbf{P}\left(|Y - \sum_{i=1}^{n} a_i f_i| > \epsilon\right) = 0$. So there exists a measurable set $A \subseteq \mathbb{R}^d \times \Omega$ and a subsequence such that $\lim_{n\to\infty} \sum_{i=1}^{m_n} a_i(\omega) f_i(x) = Y(x,\omega)$ for every $(x,\omega) \in A$ and $\mu\times\mathbf{P}(A^c) = 0$. Hence almost every x section of A has $\mathbf{P}$ measure 0. Namely, there exists a Lebesgue measurable set C, such that $\mathcal{L}(C^c) = 0$, and $\mathbf{P}(A_x) = 0$ for all $x \in C$. Then $Y(x) = \lim_{n\to\infty} \sum_{i=1}^{m_n} a_i f_i(x)$ a.s. for every $x \in C$. Since C is dense, also there exists a countable subset $D \subset C$ which is also dense, such that $Y(x_k) = \lim_{n\to\infty} \sum_{i=1}^{m_n} a_i f_i(x_k)$ a.s. for every $x_k \in D$. This implies that each $Y(x_k)$ is $\sigma((a_i)_i)$-measurable, and then $\sigma((Y_{x_k})_k) \subset \sigma((a_i)_i)$. But as Y is continuous in measure then $\sigma((Y_x)_{x\in\mathbb{R}^d}) \subseteq \sigma((a_i)_i)$. The reverse inclusion follows from this: Let $\{f_i^*\}_i$ be the dual basis then $a_i = \int_{\mathbb{R}^d} Y f_i^* d\mu$, and since the integrand is $\sigma((Y_x)_{x\in\mathbb{R}^d}) \otimes \Sigma$- measurable, by Fubini's theorem the a_i's are $\sigma((Y_x)_{x\in\mathbb{R}^d})$-measurable.

4.4.2 Construction of Random Processes. Fractional Brownian Motion.

Let us discuss some constructions of random process by means of series. This type of constructions where first given for Brownian Motion by N. Wiener using Fourier expansions, and later in a more general form by Laurent Schwartz, Itô, Nisio and others.

Theorem 4.4.1. *Let (X,Σ,μ) be a finite measure space and $T : L^2 \longrightarrow L^2$ a bounded linear operator defined by $T(f)(x) = \int_X k(x,y)f(y)d\mu(y)$, where $k : X\times X :\longrightarrow \mathbb{R}$ is a measurable function such that k is symmetric and $\sup_{x\in X}\int_X |k(x,y)|^2 d\mu(y) < \infty$. If $\{a_i\}_{i\in\mathbb{N}}$ is a sequence of independent identically distributed, zero mean $\mathbb{R}$-valued random variables, with $a_i \in L^p(\Omega,\mathcal{F},\mathbf{P})$, $p\in[1,\infty)$ and $\{f_i\}_{i\in\mathbb{N}}$ is an orthonormal basis. Then, $X = \sum_{i\in\mathbb{N}} a_i T f_i$ converges in p- mean and a.s. in $L^p(X,\Sigma,\mu)$. Moreover this series converges a.e. μ a.s.*
And if $p \geq 2$ the covariance functional for this process is given by $\Gamma(\phi,\psi) = \langle\phi, T\circ T^*\psi\rangle$.

Proof. By Fubini's theorem

$$\mathbb{E}\left\|\sum_{i=n}^{m} a_i T f_i\right\|_{L^p(X)}^p = \int_X \mathbb{E}\left|\sum_{i=n}^{m} a_i T f_i\right|^p d\mu\,,$$

but since the a_i's are independent, by theorem 2.1.11 for each $x \in X$:

$$\mathbb{E}\left|\sum_{i=n}^{m} a_i T f_i(x)\right|^p \leq C_p \left(\sum_{i=n}^{m} |T f_i(x)|^2\right)^{\frac{p}{2}}$$

Since $Tf_i(x) = \int\limits_X k(x,y) f_i(y) d\mu(y) = \langle k(x,.), f_i\rangle$, by Parseval's identity,

$$\sum_{i=n}^{m} |T f_i(x)|^2 \leq \sum_{i=1}^{\infty} |T f_i(x)|^2 = \int\limits_X |k(x,.)|^2 d\mu \leq \sup_{x\in X} \int\limits_X |k(x,.)|^2 d\mu\,,$$

and since $\mu(X) < \infty$, by Lebegue's theorem, then

$$\mathbb{E}\left\|\sum_{i=n}^{m} a_i T f_i\right\|_{L^p(X)}^p \leq C_p \int\limits_X \left(\sum_{i=n}^{m} |T f_i|^2\right)^{\frac{p}{2}} d\mu \underset{n,m\to\infty}{\longrightarrow} 0\,.$$

Then, we have convergence in the p-mean, and then in probability, but from theorem 3.1.2, this implies a.s. convergence. Recalling theorem 4.3.1 we get that a.s. this series converges a.e. $[\mu]$. Now, note that since the a_i's are uncorrelated $\mathbb{E}(X(x)X(x')) = \sum\limits_{i\,j=1}^{\infty} Tf_i(x)Tf_j(x')\mathbb{E}(a_i a_j) = \sigma^2 \sum\limits_{i=1}^{\infty} Tf_i(x)Tf_i(x')$. Then for fixed $\phi, \psi \in L^2(X, \Sigma, \mu)$:

$$\int\limits_X \int\limits_X \phi(x)\mathbb{E}(X(x)X(x'))\psi(x') d\mu(x) d\mu(x')$$

$$= \sum_{i=1}^{\infty} \langle T^*\phi, f_i\rangle\langle T^*\psi, f_i\rangle = \langle T^*\psi, T^*\phi\rangle\ .$$

But this last term equals $\langle \psi, T \circ T^*\phi\rangle$. □

In some sense, this gives a method for constructing processes with a prescribed correlation functional. However, there are some limitations, such as the the finite measure hypothesis. In the case when $X = \mathbb{R}^d$, in the following chapter we shall give a similar construction but in the space of Schwartz's distributions, although this convergence will be in a weaker sense, it gives some flexibility. Finally, let us give an example, the Brownian Motion.

Brownian motion (Bm) and fractional Brownian motion of type II(fBm II)

In this case set $X = [0,1]$ and μ the ordinary Lebesgue measure. Let $\{f_i\}_i$ be any orthonormal basis of $L^2[0,1]$, and let $\{a_i\}_i$ be a sequence of $N(0,1)$ i.i.d's random variables, and consider the fractional integration operators over $[0,1]$:

$$I_\alpha(f)(x) = \Gamma^{-1}(\alpha) \int\limits_0^x (x-y)^{\alpha-1} f(y) dy\,.$$

This integral operators have similar properties to their $\mathbb{R}^d$ versions, defined in chapter 2. It is easy to verify that I_α defines a continuous linear operator from $L^2[0,1]$ to $L^p[0,1]$, for any $p < \infty$, provided that $\alpha > \frac{1}{2}$. Indeed, by the Cauchy-Schwartz inequality,

$$\|I_\alpha f\|_{L^p}^p = \int\limits_{[0,1]} |I_\alpha f(x)|^p dx$$

$$\leq \int\limits_{[0,1]} \left(\int\limits_0^x |f(y)|^2 dy \right)^{\frac{p}{2}} \left(\int\limits_0^x \left|\Gamma^{-1}(\alpha)(x-y)^{\alpha-1}\right|^2 dy \right)^{\frac{p}{2}} dx \leq C \, \|f\|_{L^2}^p \ ,$$

where C is a constant depending only on α and p. Now if we denote $k_\alpha(x,y) = \Gamma^{-1}(\alpha)(x-y)^{\alpha-1}\mathbf{1}_{\{y\leq x\}}$. We have by the previous results that $X(x) = \sum\limits_{i=1}^{\infty} a_i I_\alpha(f_i)(x)$ converges in p-mean and a.s. to a process with covariance given by:

$$\mathbb{E}(X(x)X(x')) = \int\limits_{[0,1]} k_\alpha(x,y)k_\alpha(x',y)dy = K(\alpha) \int\limits_0^{x\wedge x'} (x-y)^{\alpha-1}(x'-y)^{\alpha-1}dy \, .$$

Note that, we have not used that the a_i´s are Gaussian r.v.'s. So we have constructed a process with the same second order properties of *fBm* II. If the the a_i's are Gaussian then this is a *fBm*. This approach is also useful to study the regularity of the trajectories. Indeed, [85], as I_β also maps continuously L^r into the Hölder class of continuous functions $\Lambda_{\alpha-1/r}$, if $1/r < \alpha < 1 + 1/r$, then, taking $\frac{1}{2} < \gamma < \alpha$ write $\alpha = (\alpha - \gamma) + \gamma$, and since ,[85], $I_\alpha = I_{\alpha-\gamma} \circ I_\gamma$. Then, by the previous arguments, we have that $\sum\limits_{i=1}^{\infty} a_i I_\gamma f_i \in L^p[0,1]$ a.s., and then $X \in \Lambda_{\alpha-\gamma-\frac{1}{p}}$, $\forall\, 1 \leq p < \infty$. Thus $X \in \Lambda_\beta$ a.s. for all $\beta < \alpha - \frac{1}{2}$.

4.5 Bibliographical and Historical Notes

4.5.1 The product space, Fubini´s theorem and almost everywhere convergence. Theorem 4.3.1 its consequences and related results.

An early, and brief, result in the direction of theorems 4.3.4 and 4.3.1 is given in Alexits' book [2]:

Theorem 4.5.1. *Let $\{f_k\}_k$ be an orthonormal system in $L^2((a,b),dx)$. If $\sum_k c_k^2 < \infty$, and $\{a_k\}_k$ is a sequence of independent $Ber(p=0.5)$ random variables, taking values in $\{-1,1\}$, then with probability one: $\sum_k a_k c_k f_k(x)$ converges for almost all $x \in (a,b)$.*

This result is a consequence of the a.e. convergence of series of Rademacher functions, Parseval´s theorem and, finally, Fubini´s theorem. In theorem 4.3.1 we proved an almost everywhere convergence result by a device of using an already known result over a fixed measure space, in this case Kolmogorov's inequality 2.1.3 (or similar results could have been used), and by means of Fubini´s theorem, by transferring this known result to the product measure space. This device can be found in several results of ergodic theory, for example, in Wiener´s local ergodic theorem or Calderón´s proof of the a.e. convergence of the Hilbert ergodic transform, namely:

Theorem 4.5.2. *(N. Wiener)[65] Let $\{T_t\}_t$ be a measure preserving flow on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and suppose that $X \in L^1(\Omega, \mathcal{F}, \mathbf{P})$. Then,*

$$\lim_{\epsilon \to 0^+} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} X \circ T_t dt = X \ a.s.$$

This rather short result is a consequence of Lebesgue´s differentiation theorem and Fubini´s theorem. On the other hand, a similar, but more complicated idea is behind Calderón´s proof of the convergence of the ergodic Hilbert transform. There, using Fubini´s theorem he transfers a known result for real variable functions, the weak -type $(1,1)$ inequality for the ordinary Hilbert transform, to the ergodic context. Using this device he proves:

Theorem 4.5.3. *(A.P. Calderón)[65] Let $\{T_t\}_t$ be a measure preserving flow on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and suppose that $X \in L^1(\Omega, \mathcal{F}, \mathbf{P})$. Then,*

$$\lim_{\epsilon \to 0^+} \int_{\epsilon \leq |t| \leq \frac{1}{\epsilon}} \frac{X \circ T_t}{t} dt \quad exists \ \ a.s.$$

4.5.2 On theorems 4.3.2 and 4.3.3

An alternative proof of points $i)$ and $ii)$ of theorem 4.3.3 can be found, in the book of Samorodnitsky and Taqqu [71]. This completely different proof, in the context of stochastic integrals, relies in some specific representations of stable processes by means of Poisson processes. Another different proof, of the complete result, can be found in [15]. Here, theorem 4.3.3 was presented as a consequence of a more general result on the convergence of L^p valued random

series, theorem 4.3.2. This is the way in which this result is presented, in for example, [42]. There, a similar result to theorem 4.3.2, for Orlicz spaces, is presented previously. However, it is worth mention, that the proof presented there is not complete. The "if-$\Leftarrow$" part of the analogue of theorem 4.3.2 presented there is almost missing.
Other rather similar conditions to theorem 4.3.2, for series of $L^p(X)$-valued independent random variables, with $p \geq 2$, are studied in [29]. These results are based on Rosenthal's inequality 3.1.1.

Chapter 5

Some Random Series in $\mathcal{D}'(\mathbb{R}^d)$

5.1 Introduction

In this chapter we prove the convergence in the sense of distributions of certain random series, this result is useful to construct some type of random processes which cant be defined from their pointwise values. Here we construct series such that given $\{\xi_n\}_n$ a set of independent, identically distributed random variables, and if $\{g_k\}_k$ is a set of appropriate functions, then

$$\sum_{n=0}^{\infty} \xi_n g_n = X\,, \tag{5.1.1}$$

converges a.s. in the sense of Schwartz distributions. First, we shall give some definitions and general results related to the theory of generalized random processes. Then, we will discuss a method for the construction by series of generalized random fields with prescribed correlation functional. This construction resembles the classic Karhunen-Loéve theorem. In particular, this method is useful for the construction of stationary generalized random processes, with a spectral behaviour which fall out of the ordinary theory of stationary processes, such as the $\frac{1}{f}$ (where f is "frequency") family of stochastic processes . For more references and physical motivations on constructing such a process, and other possible applications read sec. 5.5 at the end of this chapter.

5.2 Generalized random processes

5.2.1 Some generalities.

In this chapter we will consider the class of random variables taking values in the space $\mathcal{D}'(\mathbb{R}^d)$ (or in $\mathcal{S}'(\mathbb{R}^d)$).
Our results, are aimed at the construction of certain random variables taking values in $\mathcal{D}'(\mathbb{R}^d)$. In this case, every $\mathcal{D}'(\mathbb{R}^d)$- valued random variable, say X, takes the form of a random linear functional defined on $\mathcal{D}(\mathbb{R}^d)$. Previously, we will also need to define the class of *generalized random processes* [a], of which these $\mathcal{D}'(\mathbb{R}^d)$- valued random variables are particular cases.

[a] In some literature, the term "process" is reserved for the case $d = 1$, and "field" for the other cases.

Following [25], we will say that a generalized random functional is defined on $\mathcal{D}(\mathbb{R}^d)$ if for every $\phi \in \mathcal{D}(\mathbb{R}^d)$ there is associated a real valued random variable $X(\phi)$. In accordance with the way that one usually specifies the probability distributions of a countable set of real random variables, given $n \in \mathbb{N}$, $\phi_1, \ldots, \phi_n \in \mathcal{D}(\mathbb{R}^d)$ one gives the probability of the events,

$$\{a_k \leq X(\phi_k) < b_k\}, \quad k = 1, \ldots, n\ ,$$

and these probability distributions are compatible in the usual sense.
On the other hand, the linearity means that for any $a, b \in \mathbb{R}$, ϕ, $\psi \in \mathcal{D}(\mathbb{R}^d)$:

$$X(a\phi + b\psi) = aX(\phi) + bX(\psi) \quad \text{a.s.}$$

Finally, the random functional X is called continuous if given $n \in \mathbb{N}$, and if for $j = 1, \ldots, n$ the functions $\phi_{k\,j}$ converge to ϕ_j in $\mathcal{D}(\mathbb{R}^d)$, when $k \longrightarrow \infty$ then for any continuous bounded function $f \in C(\mathbb{R}^n)$:

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) d\mathscr{L}_k = \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) d\mathscr{L}\ ,$$

where $\mathscr{L}$ is the law of corresponding to $(X(\phi_1), \ldots, X(\phi_n))$ and $\mathscr{L}_k$ the law corresponding to $(X(\phi_{k\,1}), \ldots, X(\phi_{k\,n}))$. Note, that in particular, a $\mathcal{D}'(\mathbb{R}^d)$-valued random variable is a generalized random process. Following the usual convention, we will just refer to them as generalized random processes.

5.2.2 Two examples

"Continuous time" white noise.
We associate to the linearly independent functions $f_1, \ldots, f_n \in \mathcal{D}(\mathbb{R})$, the random variable having the probability distribution

$$\mathbf{P}\left((X(f_1), \ldots, X(f_n)) \in A\right) = \frac{(\det(\mathrm{R}))^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \int_A e^{-\frac{1}{2}Rx.x} dx\,, \tag{5.2.1}$$

where $R^{-1}_{i\,j} = \int\limits_{\mathbb{R}} f_i(x) f_j(x) dx$. It can be shown that these random variables are compatible, and are continuous and linear in f. The generalized random process defined by eq. 5.2.1, is called the unit process or white noise.
Brownian motion.
In this instance, one may specify in the usual way the probabilities of

$$(W(t_1), \ldots, W(t_n))\,,$$

where $W(t)$ is the trajectory of a Brownian process at an instant t. Namely if $W(0) = 0$, then the probability that $(W(t_1), \ldots, W(t_n)) \in A$ is expressed by

$$\mathbf{P}((W(t_1), \ldots, W(t_n)) \in A)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \prod\limits_{j=1}^{n} (t_j - t_{j-1})^{1/2}} \int\limits_A e^{-\frac{1}{2}\sum\limits_{j=1}^{n} \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}}} dx_1 \ldots dx_n\,, \tag{5.2.2}$$

with $t_0 = x_0 = 0$. It can be shown that the generalized random process corresponding to the random process $\{W(t)\}_{t\in\mathbb{R}}$, with probability distribution as eq. 5.2.2, associates with the linearly independent functions $f_1, \ldots, f_n$ the random variable $(W(f_1), \ldots, W(f_n))$ with probability distribution

$$\mathbf{P}((W(f_1), \ldots, W(f_n)) \in A) = \frac{(\det(\mathrm{R}))^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \int_A e^{-\frac{1}{2}Rx.x} dx \,,$$

where, $R^{-1}_{i\,j} = \int\limits_{\mathbb{R}} (g_i(x) - g_i(\infty))(g_j(x) - g_j(\infty))dx$, $g_i(x) = \int\limits_0^x f_i(t)dt$, this random process is called the Wiener process.

For more information about these facts and how measures can be constructed over a space such as $\mathcal{D}'(\mathbb{R}^d)$ we refer the reader to Gel'fand and Vilenkin's book [26], chapters III and IV.

5.2.3 Basic operations on Generalized Random Processes

The operations which can be performed on generalized random processes are defined in a manner analogous to that by which they are defined for generalized functions. For example, by a linear combination $aX_1 + bX_2$ is understood the generalized random process which associates with every test function $f \in \mathcal{D}(\mathbb{R}^d)$ the random variable $aX_1(f) + bX_2(f)$. Thus, the set of all generalized random processes forms a linear space.
The ordinary operations on generalized random processes are defined by means of the corresponding operations on the test functions f. Thus the product fX of an infinitely differentiable function f with a generalized process X is defined as the process for which there corresponds to the function $g \in \mathcal{D}(\mathbb{R}^d)$ the random variable $X(fg)$. In the same way it is possible to define the derivative X' of X, as the process for which there corresponds to the test function f the random variable $-X(f')$. We note that the derivative of a generalized random process always exists and is also a generalized random process, however the (generalized) derivative of an ordinary process may no longer be an ordinary random process. As an example it is easy to verify that the unit process or "white noise" process is the (generalized) derivative of the the Wiener process, although it is well known that this process has no ordinary derivative.

5.2.4 Expected value of Generalized Random Processes and correlation functional

In chapter 2 we introduced several notions of expected value for vector valued random variables. Let us discuss this notion in the context of generalized random processes.

Definition 21. *1. Expected value.* Let X be a generalized random process, then for each $f \in \mathcal{D}(\mathbb{R}^d)$, $X(f)$ is a random variable, if for each f, $X(f)$ has an expected value $m(f) = \mathbb{E}X(f)$, which is continuous in f. Then is easy to see that m is a continuous linear functional on $\mathcal{D}(\mathbb{R}^d)$ i.e. a generalized function.
2. Correlation functional. If the mean of the random variable $X(f)X(g)$ exists for all $f, g \in \mathcal{D}(\mathbb{R}^d)$ and is continuous in each of the arguments f, g, we call it the *correlation functional*

of X [b], Thus the correlation functional of the generalized random processes X is given by $\Gamma_X(f,g) = \mathbb{E}(X(f)X(g))$.

From the linearity of the random functional X , it follows that Γ_X is a bilinear functional. Moreover since $X(f)^2$ is positive, its mean $\Gamma_X(f,f)$ is also positive. Therefore, the correlation functional is positive-definite. Also it is possible to define the covariation functional $C(f,g) = \mathbb{E}(X(f)X(g)) - \mathbb{E}(X(f))\mathbb{E}(X(g))$. Which is also positive-definite.

Recall that in the finite dimensional case, given the correlation matrix we can associate to it a positive-definite bilinear form. We shall see, that in this case we can find a distribution which plays a similar role of this matrix.

Theorem 5.2.1. *(The Kernel Theorem) i) Every bilinear functional $B(\,,\,)$, defined over $\mathcal{D}(\mathbb{R}^d) \times \mathcal{D}(\mathbb{R}^d)$, which is continuous in each of the arguments, has the form:*

$$B(f,g) = (F, f(x)g(x'))\,,$$

where $F \in \mathcal{D}'(\mathbb{R}^{2d})$.
ii)Every bilinear functional $B(\,,\,)$, defined over $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d)$, which is continuous in each of the arguments, has the form:

$$B(f,g) = (F, f(x)g(x'))\,,$$

where $F \in \mathcal{S}'(\mathbb{R}^{2d})$.

Proof. [26] □

In both cases, when $B = \Gamma_X$ is the correlation functional of a processes X, we will sometimes call to F the correlation functional. From the theorem above, it is clear that there is no ambiguity in using this term both for F or B.
Example. Recall the example of white noise, defined by eq. 5.2.1, in this case it is clear that the correlation functional is given by $F(x,x') = \delta(x-x')$. This means, that

$$\Gamma_X(f,g) = (\delta(x-x'), f(x)g(x')) = \int_{\mathbb{R}}\int_{\mathbb{R}} \delta(x-x')f(x)g(x')dxdx'$$

$$= \int_{\mathbb{R}} f(x)g(x)dx\,.$$

Where the identity with the double integral is only formal, since this integrals are not defined. Finally, as in the finite dimensional case, we have:

Theorem 5.2.2. *In order that a linear continuous functional m on $\mathcal{D}(\mathbb{R}^d)$ and a bilinear continuous functional Γ on $\mathcal{D}(\mathbb{R}^d) \times \mathcal{D}(\mathbb{R}^d)$ be respectively the mean and the correlation functional of a generalized random process X, it is necessary and sufficient that the bilinear functional,*

$$C(f,g) = \Gamma(f,g) - m(f)m(g)$$

be positive-definite, in which case the process can be chosen to be gaussian.

Proof. [26] □

[b] For simplicity, here we are considering real random processes, for complex random processes the correlation is defined by $\mathbb{E}(X(f)\overline{X}(g))$.

5.2.5 Stationary Generalized Random Process.

In chapter 2, section 2.4.5, we introduced the class of wide sense stationary random process. Here, we define the corresponding notion for generalized random processes. First, let us discuss stationarity in the strict sense. A generalized random process is called *stationary* if for any functions $f_1, \ldots, f_n \in \mathcal{D}(\mathbb{R}^d)$ and any $h \in \mathbb{R}^d$: $(X(f_1(\,.+h)), \ldots, X(f_n(\,.+h)))$ and $(X(f_1), \ldots, X(f_n))$ are identically distributed. If X is stationary, then its expected value is invariant under translation, thus

$$\mathbb{E}(X(f)) = \mathbb{E}(X(f(.+h)))\,. \tag{5.2.3}$$

But the only linear functionals on $\mathcal{D}'(\mathbb{R}^d)$ which are invariant under translation are those of the form:

$$\mathbb{E}(X(f)) = a \int_{\mathbb{R}^d} f(x)dx$$

for some constant a. Also from the stationarity of the process it follows that

$$\Gamma_X(f,g) = \Gamma_X(f(\,.+h), g(\,.+h))\,, \tag{5.2.4}$$

for any two functions $f, g \in \mathcal{D}(\mathbb{R}^d)$. Thus, the correlation functional of a stationary generalized random process is a positive- definite bilinear Hermitian functional which is translation invariant. However, in the following we shall deal, with generalized processes which only verify the weaker conditions of eqs. 5.2.3 and 5.2.4. Following the definition given in 2, section 2.4.5, if a generalized process verifies these conditions it will be called wide sense stationary. And to simplify, without loss of generality we will assume that $a = 0$ and from here when we talk about stationarity we will be referring to wide sense stationarity. Let us see the form that the correlation functional takes under this assumptions. We need two results,

Theorem 5.2.3. *Every translation invariant bilinear functional B on $\mathcal{D}'(\mathbb{R}^d)$ has the form:*

$$B(f,g) = (F, f * g)$$

where $F \in \mathcal{D}'(\mathbb{R}^d)$. If B is positive definite then F is positive definite, i.e. $(F, f * f^*) \geq 0$ for every $f \in \mathcal{D}(\mathbb{R}^d)$, where $f^*(x) = f(-x)$.

Proof. See [26]. □

Using a generalization of Bochner's theorem: the Bochner-Schwartz theorem is possible to give an expression in terms of Fourier transforms.

Theorem 5.2.4. *(Bochner-Schwartz) The class of positive definite generalized functions on $\mathcal{D}(\mathbb{R}^d)$ coincides with the class of positive tempered measures. (A measure μ is called tempered if for some $m \geq 0$, $\int_{\mathbb{R}^d} \frac{d\mu}{(1+|\lambda|^2)^m} < \infty$)*

Proof. See [26]. □

From this, one gets,

Theorem 5.2.5. *Every translation invariant positive definite bilinear functional B on $\mathcal{D}(\mathbb{R}^d)$ has the form*

$$B(f,g) = \int_{\mathbb{R}^d} \hat{f}(\lambda)\overline{\hat{g}}(\lambda)d\mu(\lambda)$$

where μ is a positive tempered measure.

Proof. See [26]. □

These results have the following immediate consequence:

Corollary 5.2.1. *The correlation functional of a stationary generalized random process X, has the form*

$$\Gamma_X(f,g) = (R, f * g^*)\,,$$

where R is the Fourier transform of some positive tempered measure μ_X. Moreover,

$$\Gamma_X(f,g) = \int_{\mathbb{R}^d} \hat{f}(\lambda)\overline{\hat{g}}(\lambda)d\mu_X(\lambda)\,.$$

Proof. [26] Is immediate consequence of the previous theorems 5.2.3 and 5.2.5. □

So in view of this, the measure μ_X of corollary 5.2.1 will be called *spectral measure* of the process. This measure is uniquely defined by the process.

Some examples. Let X be a white noise or unit process, then as we have seen, the correlation functional can be written as

$$\Gamma_X(f,g) = (\delta(x-x'), f(x)g(x')) = (\delta, f * g^*) = \int_{\mathbb{R}^d} \hat{f}(\lambda)\overline{\hat{g}}(\lambda)d\lambda\,,$$

so, the spectral measure of this process is the Lebesgue measure , i.e. $d\mu_X = d\lambda$. Intuitively, this says that in the process X, all the frequencies are present and are "equidistributed".
Let $d > s > 0$, then as we have seen in chapter 2, the Fourier Transform of $|x|^{-d+s}$ is $\gamma(s)(2\pi)^{-s}\,|\lambda|^{-s}$ in the distributional sense. So in view of the previous results and theorem 6.3.2 there exists a generalized random process with spectral measure defined by the generalized function,

$$\gamma(s)(2\pi)^{-s}\,|\lambda|^{-s}\;.$$

This corresponds to the case of the so called $\frac{1}{f}$-processes. The correlation functional is given by $|x|^{-d+s}$.
If we can write $d\mu_X = \phi_X d\lambda$ for some $\phi_X \in L^1_{loc}(\mathbb{R}^d)$, this ϕ_X will be called the spectral density of the process X.
Note that in chapter 2, section 2.4.5, we noted that every (ordinary) w.s.s. process has a *finite* spectral measure. So spectral measures such as those above, are not valid within the theory of ordinary stationary random process.
So, if we would like to construct a process with a spectral behaviour such as those described before, in principle one could specify or construct a probability measure over the space $\mathcal{D}'(\mathbb{R}^d)$. A detailed discussion about the construction of measures in linear topological spaces can be

found in [26]. However, from the point of view of some applications, to describe the desired process this way could become less valuable. From a practical view, it could be difficult to handle such measures. Moreover in many times only is required a model which fullfills certain requisites related with its spectral behaviour or second order properties. So, if we find an expansion like 5.1.1, which converges to a process with the desired properties we would be done. Note that, in contrast, if we adopt this approach to construct the process, and if the f_n's are fixed deterministic functions we would only need to construct the real valued random variables ξ_n.

5.3 Representation of some generalized random processes by random series.

5.3.1 Some auxiliary results and definitions.

As in chapter 3 we gave several results on random series in Banach spaces, but $\mathcal{D}'(\mathbb{R}^d)$ is not a normed space, so it will be necessary to introduce as an auxiliary tool the Sobolev H^s spaces. At an intermediate step they will allow us to introduce the results for Banach spaces into this context.

Definition 22. The Sobolev spaces H^s [30] are defined as:

$$H^s(\mathbb{R}^d) = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \int\limits_{\mathbb{R}^d} \left|\widehat{f}(\lambda)\right|^2 \left(1 + |\lambda|^2\right)^s d\lambda < \infty \right\} \tag{5.3.1}$$

Remark 5.3.1. Let $s \in \mathbb{R}$ then $H^s(\mathbb{R}^d)$ is a Hilbert space with the product $(\ ,\ )_{H^S} : H^s(\mathbb{R}^d) \times H^s(\mathbb{R}^d) \longmapsto \mathbb{C}$

$$(h, g)_{H^s} = \int\limits_{\mathbb{R}^d} \widehat{h}(\lambda)\overline{\widehat{g}}(\lambda) \left(1 + |\lambda|^2\right)^s d\lambda\,. \tag{5.3.2}$$

For $f,\, g \in \mathcal{D}(\mathbb{R}^d)$ we define the pairing $\langle\, .,\, .\rangle : \mathcal{D}(\mathbb{R}^d) \times \mathcal{D}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ as

$$\langle f, g\rangle = \int\limits_{\mathbb{R}^d} f(x)g(x)dx\,.$$

This can be extended by a density argument over $L^p \times L^q$, $\frac{1}{p} + \frac{1}{q} = 1$ (when $p = 2$ this is the usual inner product) or $H^s \times H^{-s}$.

Now recalling the variant of the Shannon Nyquist Kotelnikov theorem of chapter 2, we can prove:

Proposition 5.3.1. *Let $f \in L^2(\mathbb{R}^d)$ be with the same hypotheses of theorem 2.4.1, then*

$$\|f\|_{H^s} \le K(s) \left(\sum_{k\in\mathbb{Z}^d} \left|\widehat{f}(k)\right|^2 \left(1 + |k|^2\right)^s \right)^{1/2}\,.$$

Proof. Recall Peetre's inequality: $\left(1+(a+b)^2\right)^s \leq 2^{|s|}\left(1+a^2\right)^{|s|}\left(1+b^2\right)^s$ and by theorem 2.4.1 we have:

$$\int\limits_{\mathbb{R}^d}\left|\widehat{f}(\lambda)\right|^2\left(1+|\lambda|^2\right)^s d\lambda \leq \int\limits_{\mathbb{R}^d}\left(\sum_{k\in\mathbb{Z}^d}\left|\widehat{f}(k)\theta(\lambda-k)\right|\left(1+|\lambda|^2\right)^{s/2}\right)^2 d\lambda$$

$$\leq \int\limits_{\mathbb{R}^d}\sum_{k\in\mathbb{Z}^d}u_k^2\left(\lambda\right)\sum_{k\in\mathbb{Z}^d}v_k^2\left(\lambda\right)d\lambda\ ,$$

where $v_k\left(\lambda\right)=\left|\theta(\lambda-k)\right|^{1/2}$ and

$$u_k\left(\lambda\right)=\left|\widehat{f}(k)\right|\left(1+|k|^2\right)^{s/2}2^{|s|/2}\left(1+||\lambda|-|k||^2\right)^{|s|/2}\left|\theta(\lambda-k)\right|^{1/2}\ .$$

Since $\theta(x)\in\mathcal{S}(\mathbb{R}^d)$ we have $\sum\limits_{k\in\mathbb{Z}^d}v_k^2\left(\lambda\right)=\sum\limits_{k\in\mathbb{Z}^d}\left|\theta(\lambda-k)\right|\leq C$, for some $0<C<\infty$, and:

$$K\left(s\right)2^{-|s|}=\int\limits_{\mathbb{R}^d}\left(1+||\lambda|-|k||^2\right)^{|s|}\left|\theta(\lambda-k)\right|d\lambda$$

$$\leq\int\limits_{\mathbb{R}^d}\left(1+|\lambda-k|^2\right)^{|s|}\left|\theta(\lambda-k)\right|d\lambda<\infty\ ,$$

then:

$$\int\limits_{\mathbb{R}^d}\left|\widehat{f}(\lambda)\right|^2\left(1+|\lambda|^2\right)^s d\lambda\leq CK\left(s\right)\sum_{k\in\mathbb{Z}^d}\left|\widehat{f}(k)\right|^2\left(1+|k|^2\right)^s$$

□

5.3.2 Main Results.

Now, we can prove the desired result,

Theorem 5.3.1. *Let $\{\xi_n\}_{n\in\mathbb{N}}\subset L^4(\Omega,\mathcal{F},\mathbf{P})$ be a sequence of independent identically distributed random variables such that $\mathbb{E}\xi_n=0$. If $\{f_n\}_{n\in\mathbb{N}}$ is an orthonormal basis of $L^2\left(\mathbb{R}^d\right)$ and let $T:L^2(\mathbb{R}^d)\longrightarrow L^p(\mathbb{R}^d)$ be a bounded linear operator, with $p\geq 1$; then:*
I)

$$X=\sum_{n=0}^{\infty}\xi_n Tf_n \tag{5.3.3}$$

converges to a generalized process a.s.

II) The covariance functional of X, $\Gamma_X:\mathcal{D}(\mathbb{R}^d)\times\mathcal{D}(\mathbb{R}^d)\longrightarrow\mathbb{R}$ is $\Gamma_X(\phi,\psi)=\langle\phi,T\circ T^*\psi\rangle$
III) Given $\varphi\in\mathcal{D}(\mathbb{R}^d)$ then

$$X(\varphi)=\sum_{n=0}^{\infty}\xi_n\left\langle Tf_n,\varphi\right\rangle \text{ in the } L^2 \text{ sense.} \tag{5.3.4}$$

Proof. (Part I) Let $\{Q_p\}_p$ be a denumerable family of disjoint cubes such that by some translation τ_p equals $\left(-\frac{1}{4},\frac{1}{4}\right]^d$ and $\mathbb{R}^d=\bigsqcup_p Q_p$. then

$$\left\|(Tf_n)\,\mathbf{1}_{Q_p}\right\|_{H^s}\le K(s)\left(\sum_{k\in\mathbb{Z}^d}\left|\widehat{(Tf_n)\,\mathbf{1}_{Q_p}}(k)\right|^2\left(1+|k|^2\right)^s\right)^{1/2}$$

with $\left|\widehat{(Tf_n)\,\mathbf{1}_{Q_p}}(k)\right|=\left|\left\langle (Tf_n)\,\mathbf{1}_{Q_p},e_k\right\rangle\right|$ and $e_k=e^{i2\pi k.x}\mathbf{1}_{\tau_p^{-1}[-\frac{1}{4},\frac{1}{4}]^d}$ then by proposition 5.3.1

$$\sum_n\left\|(Tf_n)\,\mathbf{1}_{Q_p}\right\|^2_{H^s}\le\sum_n K(s)\sum_{k\in\mathbb{Z}^d}\left|\widehat{(Tf_n)\,\mathbf{1}_{Q_p}}(k)\right|^2\left(1+|k|^2\right)^s$$

Taking $s=-d$, and as $e_k\in L^{p'}(\mathbb{R}^d)$, $\frac{1}{p}+\frac{1}{p'}=1$ and $Supp((Tf_n)\,\mathbf{1}_{Q_p})=Supp(e_k)$ then the last term equals:

$$\sum_{k\in\mathbb{Z}^d}K(s)\left(1+|k|^2\right)^{-d}\sum_n\left|\left\langle (Tf_n)\,\mathbf{1}_{Q_p},e_k\right\rangle\right|^2$$

$$=\sum_{k\in\mathbb{Z}^d}K(s)\left(1+|k|^2\right)^{-d}\sum_n\left|\langle f_n,T^*e_k\rangle\right|^2\le\sum_{k\in\mathbb{Z}^d}K(s)\left(1+|k|^2\right)^{-d}\left\|T^*e_k\right\|^2_{L^2}\qquad(5.3.5)$$

$$\le\sum_{k\in\mathbb{Z}^d}K(s)\left(1+|k|^2\right)^{-d}K''\left\|e_k\right\|^2_{L^{p'}}\le K'''\int_{\mathbb{R}^d}(1+|x|^2)^{-d}dx\,|Q_p|^{2/p'}<\infty\,,$$

Independently of p. Since $\{\xi_n\}_{n\in\mathbb{N}}$ are independent random variables we can assume, without loss of generality, $\mathbb{E}\,|\xi_n|^2=1$ then:

$$\sum_n\mathbb{E}\,|\xi_n|^2\left\|(Tf_n)\,\mathbf{1}_{Q_p}\right\|^2_{H^{-d}}=\sum_n\left\|(Tf_n)\,\mathbf{1}_{Q_p}\right\|^2_{H^{-d}}<\infty$$

By corollary 3.1.4 and remark 5.3.1 we have $\left\|\sum_n\xi_n\,(Tf_n)\,\mathbf{1}_{Q_p}\right\|_{H^{-d}}<\infty$ a.s.

Take $\Omega'=\bigcap_{p=1}^{\infty}\left\{\omega\in\Omega:\left\|\sum_n\xi_n(\omega)\,(Tf_n)\,\mathbf{1}_{Q_p}\right\|_{H^{-d}}<\infty\right\}$. Now, for a fixed $\omega\in\Omega'$, let us see that 5.3.3 converges in $\mathcal{D}'(\mathbb{R}^d)$. For this it suffices to show that given $\varphi\in\mathcal{D}'(\mathbb{R}^d)$ then $\lim_{n\to\infty}\sum_{i=1}^n\langle Tf_i,\varphi\rangle\xi_i(\omega)$ exists [72]. Indeed, if $\varphi\in\mathcal{D}'(\mathbb{R}^d)$, then it has compact support, so for some fixed q, $Supp(\varphi)\subset\bigcup_{p=1}^q Q_p$. So,

$$\left|\sum_{i=m}^n\langle Tf_i,\varphi\rangle\xi_i(\omega)\right|\le\sum_{p=1}^q\left|\sum_{i=m}^n\langle (Tf_i)\mathbf{1}_{Q_p},\varphi\rangle\xi_i(\omega)\right|$$

$$=\sum_{p=1}^q\left|\langle\sum_{i=m}^n(Tf_i)\mathbf{1}_{Q_p},\varphi\rangle\xi_i(\omega)\right|$$

$$= \sum_{p=1}^{q} \left| \int_{\mathbb{R}^d} \left(\sum_{i=m}^{n} \xi_n(\omega) \widehat{(Tf_i)\mathbf{1}_{Q_p}}(\lambda) \right) (1+|\lambda|^2)^{\frac{d}{2}} \overline{\hat{\varphi}}(\lambda)(1+|\lambda|^2)^{\frac{-d}{2}} d\lambda \right|$$

Where the last equation follows from Parseval's theorem. Now, by the Cauchy-Schwartz inequality, this expression is less or equal than

$$\sum_{p=1}^{q} \left\| \sum_{i=m}^{n} \xi_i(\omega) \, (Tf_i) \, \mathbf{1}_{Q_p} \right\|_{H^{-d}} \|\varphi\|_{H^d} \underset{n,m\to\infty}{\longrightarrow} 0 \,.$$

(Part II) If $X : \Omega \longrightarrow \mathcal{D}'(\mathbb{R}^d)$ is the limit field then its covariance is $\Gamma_X(\phi,\psi) := \mathbb{E}X(\phi)X(\psi)$. (note that $X(\phi)$ is $\mathcal{F}-$ measurable since $X_m(\phi) = \sum_{n=0}^{m} \xi_n \langle Tf_n, \phi\rangle$ is $\mathcal{F}$-measurable)
In order to prove that $\mathbb{E}X_m(\phi)X_m(\psi) \longrightarrow \mathbb{E}X(\phi)X(\psi) = \langle \phi, T\circ T^*\psi\rangle$ when $m \longrightarrow \infty$, with $\phi\,,\psi \in \mathcal{D}(\mathbb{R}^d)$; first we prove the uniform integrability of the sequence $\{X_m(\phi)X_m(\psi)\}_m$, which will follow if we find $\epsilon > 0,\ K > 0$ such that

$$\mathbb{E}\,|X_m(\phi)X_m(\psi)|^{1+\epsilon} \le K \ \ \forall\, m\,. \tag{5.3.6}$$

Given $\phi \in \mathcal{D}(\mathbb{R}^d)$ let us call $c_m := \langle Tf_m, \phi\rangle = \langle f_m, T^*\phi\rangle$, $c := (c_m)_m \in \mathbb{R}^{\mathbb{N}}$. Then

$$\mathbb{E}\,|X_m(\phi)|^4 = \mathbb{E}\left(\sum_{i\,j\,k\,l=0}^{m} c_i c_j c_k c_l \xi_i \xi_j \xi_k \xi_l \right)$$

but, since the ξ_m's are independent then we have the following factorization: $d_{i\,j\,k\,l} := \mathbb{E}(\xi_i\xi_j\xi_k\xi_l) = \mathbb{E}(\xi_i)\mathbb{E}(\xi_j\xi_k\xi_l) = 0$ whenever $i \ne j\,,k\,,l$. From this fact and since the ξ_m's are identically distributed, we get

$$d_{i\,j\,k\,l} = \begin{cases} \left(\mathbb{E}|\xi_1|^2\right)^2 & \text{whenever two pairs of indexes are equal.} \\ \mathbb{E}|\xi_1|^4 & \text{if} \qquad i=j=k=l \\ 0 & \text{whenever only one index differs from the others.} \end{cases}$$

From this,

$$\begin{aligned} \mathbb{E}\,|X_m(\phi)|^4 &= \sum_{i=0}^{m} c_i^4 \mathbb{E}|\xi_1|^4 + 3 \sum_{i,j=0\ i\ne j}^{m} c_i^2 c_j^2 \left(\mathbb{E}|\xi_1|^2\right)^2 \\ &\le \mathbb{E}|\xi_1|^4 \left\|c^2\right\|_{l^\infty} \|T^*\phi\|^2{}_{L^2(\mathbb{R}^d)} + 3\|T^*\phi\|^2{}_{L^2(\mathbb{R}^d)} \qquad (5.3.7) \\ &\le \left(\mathbb{E}|\xi_1|^4 \left\|c^2\right\|_{l^\infty} + 3\right) \|T^*\| \, \|\phi\|^2{}_{L^{p'}(\mathbb{R}^d)} < \infty \qquad (5.3.8) \end{aligned}$$

Now, since $\mathbb{E}\,|X_m(\phi)X_m(\psi)|^2 \le \left(\mathbb{E}\,|X_m(\phi)|^4\right)^{\frac{1}{2}} \left(\mathbb{E}\,|X_m(\psi)|^4\right)^{\frac{1}{2}}$ and from (5.3.7) condition (5.3.6) is verified for $\epsilon = 1$, then recalling theorem 2.1.2, $\mathbb{E}X_m(\phi)X_m(\psi) \longrightarrow \mathbb{E}X(\phi)X(\psi) = \Gamma_X(\phi,\psi)$ when $m \longrightarrow \infty$.
Let us prove that $\Gamma_X(\phi,\psi) = \langle \phi, T\circ T^*\psi\rangle$. Given m let us define the bilinear form Γ_m :

$\mathcal{D}(\mathbb{R}^d) \times \mathcal{D}(\mathbb{R}^d) \longmapsto \mathbb{R}$ as follows: Let $k_m(x,y) := \sum_{j\,k=0}^{m} \mathbb{E}\xi_j\xi_k Tf_k(x)Tf_j(y)$. Then given $\phi,\ \psi \in \mathcal{D}(\mathbb{R}^d)$,

$$\Gamma_m(\phi,\psi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k_m(x,y)\phi(y)\psi(x)dxdy$$

Since $\{\xi_n\}_{n\in\mathbb{N}}$ is a sequence of independent random variables with $Var(\xi_n) = 1$ and $\mathbb{E}[\xi_n] = 0$ then $\mathbb{E}\xi_n\xi_m = \delta_{nm}$ from this it follows that $k_m(x,y) = \sum_{k=0}^{m} Tf_k(x)Tf_k(y)$ then,

$$\Gamma_m(\phi,\psi) = \int_{\mathbb{R}^d}\left(\sum_{k=0}^{m}\int_{\mathbb{R}^d} Tf_k(x)\psi(x)dxTf_k(y)\right)\phi(y)dy$$

$$= \int_{\mathbb{R}^d} T\left(\sum_{k=0}^{m}\int_{\mathbb{R}^d} f_k(x)T^*\psi(x)dxf_k(y)\right)\phi(y)dy$$

$$= \int_{\mathbb{R}^d}\left(\sum_{k=0}^{m}\int_{\mathbb{R}^d} f_k(x)T^*\psi(x)dx\, f_k(y)\right)T^*\phi(y)dy,. \tag{5.3.9}$$

Then, if P_m is the orthogonal projection over span $\{f_0, ... f_m\}$ (5.3.9) equals $\langle P_m \circ T^*\psi, T^*\phi\rangle_{L^2(\mathbb{R}^d)}$ and since $\{f_n\}_n$ is complete, given $\epsilon > 0$ there exists $M(\epsilon) \in \mathbb{N}$ such that $\|P_m \circ T^*\phi - T^*\phi\|_{L^2} < \frac{\epsilon}{\|T^*\psi\|_{L^2}}$ if $m \geq M$. On the other hand, $\langle\phi, T\circ T^*\psi\rangle = \langle T^*\phi, T^*\psi\rangle$ and from these facts, taking for example $m \geq M(\epsilon)$ it follows that

$$|\langle\phi, T\circ T^*\psi\rangle - \Gamma_m(\phi,\psi)| = |\langle T^*\phi, T^*\psi\rangle - \langle P_m\circ T^*\psi, T^*\phi\rangle| \tag{5.3.10}$$

$$= |\langle P_m\circ T^*\phi - T^*\phi, T^*\psi\rangle| \leq \|T^*\psi\|_{L^2}\,\|P_m\circ T^*\phi - T^*\phi\|_{L^2} < \epsilon\,.$$

(Part III) From equations (5.3.7), (5.3.8), given $\varphi \in \mathcal{D}(\mathbb{R}^d)$ we have that $\left\{|X_n(\varphi)|^2\right\}_n$ is uniformly integrable, since condition (2.1.2) is verified for $\epsilon = 2$; and since $X_n(\varphi) \longrightarrow X(\varphi)$ a.s. from Part I, then from theorem 2.1.2 we have

$$\lim_{n\to\infty} \mathbb{E}|X_n(\varphi) - X(\varphi)|^2 = 0\,.$$

□

5.4 Some consequences and applications. Construction of a $\frac{1}{f}$ process.

We will need the following well known result [75]:

Theorem 5.4.1. *Let $T : L^p(\mathbb{R}^d) \longrightarrow L^q(\mathbb{R}^d)$ be a bounded linear operator, if T is translation invariant then there exists a unique tempered distribution h such that for every $f \in \mathcal{S}(\mathbb{R}^d)$: $Tf = h * f$.*

Then, from the above theorem and the definition of Fourier transform of a distribution, we have the following immediate and intuitive result on the covariance functional of de limit process X of proposition 5.3.1: If T is translation invariant and $h \in \mathcal{S}'(\mathbb{R}^d)$ is the distribution of theorem 5.4.1 then,

$$\Gamma_X(\varphi,\psi) = \int\limits_{\mathbb{R}^d} \hat{\varphi}(\lambda)\,|\hat{h}(\lambda)|^2\overline{\hat{\psi}}(\lambda)d\lambda\,. \tag{5.4.1}$$

In this way, taking T as a translation invariant operator we obtain a series which converges to a generalized (wide sense) stationary process. Moreover, its spectral density is given by $\phi_X(\lambda) = |\hat{h}(\lambda)|^2$.

Construction of a fractional random field

Random processes with $1/f$ spectral behaviour, first introduced by Kolmogorov in the context of turbulent flows (For a brief reference about this read the end of this chapter, sec. 5.5.1), have numerous applications in engineering , general science and wherever strong long-range (Long Memory) dependence (LRD) phenomena appear.
A long memory process or field X with spectral density $\phi_{\mathrm{X}}(\lambda)$ verifies the condition (see [7] and [66]): there exists $\beta > 0$ and $c_f > 0$ such that

$$\lim_{\lambda\to 0}\frac{\phi_{\mathrm{X}}(\lambda)}{c_f\,|\lambda|^{-\beta}} = 1\ , \tag{5.4.2}$$

this suggests (for example, [3], [11]) to look for a relation between these processes and certain fractional integration differerencing operators (see equations 2.4.2, 2.4.3). Considering, these processes not yet as point processes, but as random elements in a space of distributions using the previous results, we can construct series which converges a.s. to a generalized fractional random field which show LRD or more generally with spectral density of the form:

$$\phi_X(\lambda) = (1+|\lambda|^2)^{-\gamma}|\lambda|^{-2\alpha} \quad \gamma, \in \mathbb{R}_{>0}\,, 0<\ \alpha\ < d/2. \tag{5.4.3}$$

The term $(1+|\lambda|^2)$, suggested in [3], is intended to give intermittency to the model. This characteristic is sometimes observed in turbulence. For this purpose we need to recall the results of section 2.4 on fractional integration.

From theorems 2.4.2 and 2.4.3, we can claim that if we define $T = (-\Delta)^{-\alpha/2}(I-\Delta)^{-\gamma/2}$ with $0<\alpha<d/2$, $\gamma>0$, then T defines a bounded linear operator from $L^p(\mathbb{R}^d)$ into $L^q(\mathbb{R}^d)$, with $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$. From this, 5.3.3 converges to a generalized random process with spectral density as 5.4.3. Note that the case $\alpha=\gamma=0$ corresponds to the case of white noise.

Remark 5.4.1. Is straightforward to see from the proof of proposition 5.3.1 that this assertion may fail if $\alpha \geq d/2$. In [3] by means of the operators 2.4.6, 2.4.3 and Fourier transform methods a nice proof of the existence over a bounded domain $U \subset \mathbb{R}^d$ of a process with spectral density as 5.4.3 with $\alpha \in (0,d)$ is given. However, this result is based on the following assertion: If $D \subset \mathbb{R}^d$ is a measurable bounded domain there exists $C>0$ such that for every $f \in L^2(\mathbb{R}^d)$ such that $Supp(f) \subset D$

$$\int\limits_{\mathbb{R}^d} |\widehat{(-\Delta)^{-\alpha/2}f}(\lambda)|^2 d\lambda \leq C\int\limits_{\mathbb{R}^d}|\hat{f}(\lambda)|^2 d\lambda\,. \tag{5.4.4}$$

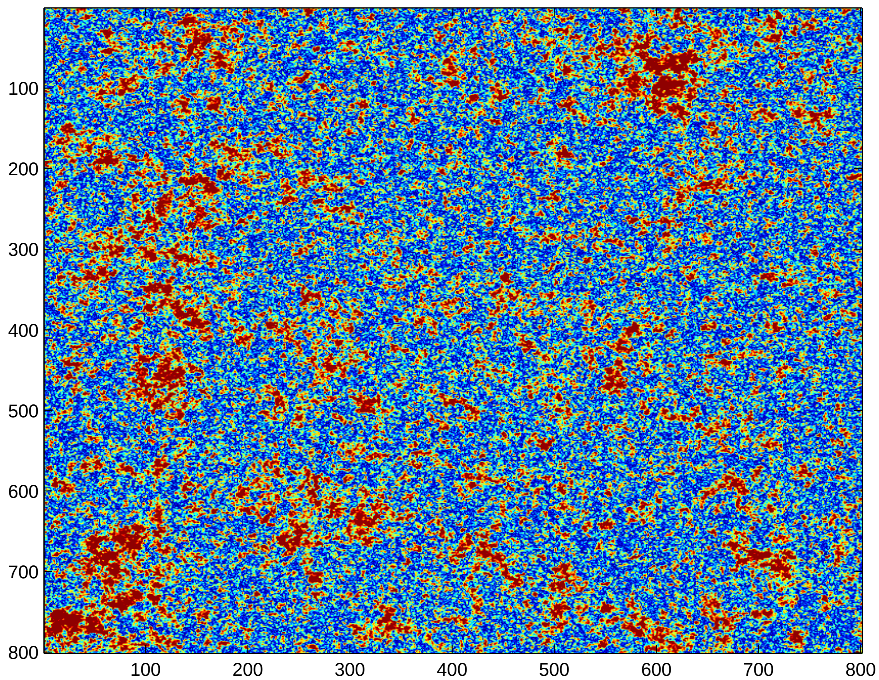

Figure 5.1: A sample of a 2-dimensional process, with spectral density given by eq. 5.4.3 ($\alpha = 0.5$).

But this is false for $\alpha \geq \frac{d}{2}$: Take $D = B(0,1)$ the ball of radius 1 and $f = \mathbf{1_D}$, we prove that for this f, $(-\Delta)^{-\alpha/2} f$ does not belong to $L^2(\mathbb{R}^d)$. From eq. 2.4.3 we have

$$(-\Delta)^{-\alpha/2} f(x) = \frac{1}{\gamma(\alpha)} \int\limits_{B(0,1)} \frac{dy}{|x-y|^{d-\alpha}} \,,$$

but $|x-y| \leq |x| + |y| \leq |x| + 1$ then $(|x|+1)^{-d+\alpha} \leq |x-y|^{-d+\alpha}$ if $|y| \leq 1$, so

$$|(-\Delta)^{-\alpha/2} f(x)| \geq K \frac{\mathcal{L}(B(0,1))}{(|x|+1)^{-d+\alpha}}$$

for all $x \in \mathbb{R}^d$. Then we have the following bound:

$$\left\| (-\Delta)^{-\alpha/2} f \right\|^2 \geq \int\limits_{\mathbb{R}^d} (|x|+1)^{-2d+2\alpha} dx \, \mathcal{L}(B(0,1))^2$$

$$= K \int\limits_0^\infty (r+1)^{-2d+d\alpha} r^{d-1} dr = k\beta(d, d-2\alpha) \,,$$

but this expression for Euler's Beta function converges if and only if $d > 0$ and $d - 2\alpha > 0$.

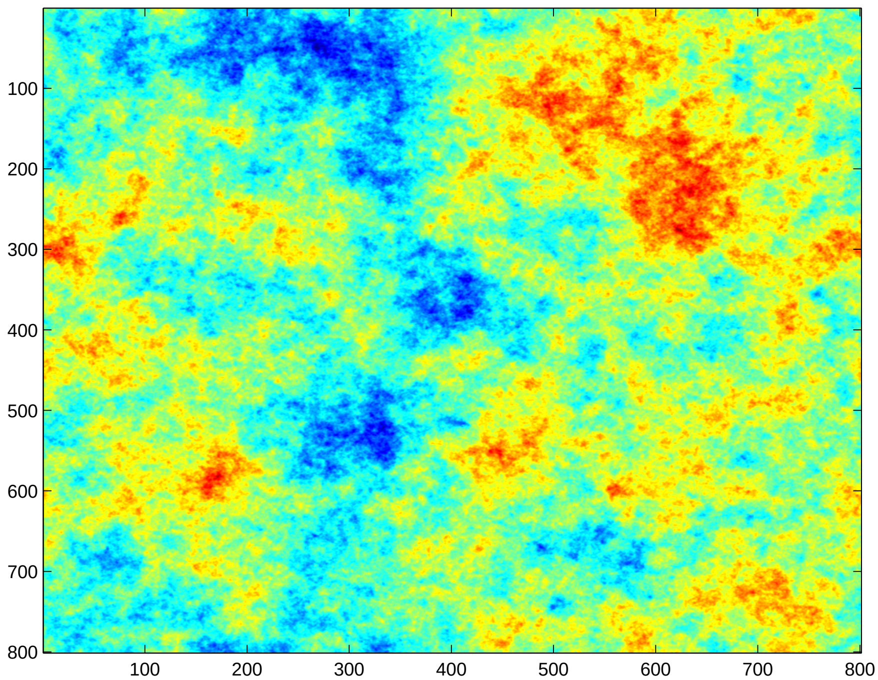

Figure 5.2: A sample of a 2-dimensional process, with spectral density given by eq. 5.4.3 ($\alpha = 0.99$).

Remark 5.4.2. *About Self-similarity.* The self-similarity property for this process is understood in the following manner: for $0 < s < d$, given $f, g \in \mathcal{D}(\mathbb{R}^d)$, $a > 0$, then by a change of variables:

$$\Gamma_X(f(a\,.), g(a\,.)) = \int_{\mathbb{R}^d} \frac{1}{a^{2d}} \hat{f}\left(\frac{\lambda}{a}\right) \overline{\hat{g}}\left(\frac{\lambda}{a}\right) \gamma(s)(2\pi)^{-s} \frac{d\lambda}{|\lambda|^s}$$

$$= \frac{1}{a^{d-s}} \int_{\mathbb{R}^d} \hat{f}\,(\lambda)\, \overline{\hat{g}}\,(\lambda)\, \gamma(s)(2\pi)^{-s} \frac{d\lambda}{|\lambda|^s} = \frac{1}{a^{d-s}} \Gamma_X(f, g)\,.$$

If X is gaussian this relation on the covariance functional is equivalent to the corresponding notion for the finite dimensional distributions of the process.

5.5 Bibliographical and Historical Notes

5.5.1 More on $\frac{1}{f}$ random fields.

A. N. Kolmogorov's $\frac{1}{f}$ model for turbulence. The term turbulence as used in fluid mechanics refers to a flow whose characteristics are so chaotic that statistical and probabilistic methods are used to describe many of its aspects. In many studies the velocity field $v(x,t)$ of the flow is considered function of the position $x = (x_1, x_2, x_3)$ and the time t that is a random solution, in some sense, of the Navier-Stokes equation

$$v_t(x,t) + (v.\bigtriangledown)v = -\frac{1}{\rho} \bigtriangledown p + \nu \bigtriangleup v$$

and the continuity equation $\bigtriangledown.v = 0$. Where ρ is the density, p the thermodynamic pressure, ν the kinematic viscosity. The dimensionless quantity $R = ul/\nu$, with u a characteristic velocity and l a characteristic length, represents an important aspect of the flow and its for large Reynold's number R that the flow takes on the erratic and unstable character of turbulence. he Reynold's number R represents the ratio of inertial to viscous forces. An idealized model is that of turbulence in which the process $v(x,t)$ is considered a stationary random process in the spatial coordinates x with finite second order moments. It is thought that the turbulence generated downstream in a wind tunnel by passing a uniform fluid flow through a regular grid of bars held at right angles to the flow is approximated reasonably by such a model locally. Another idea, to model turbulence, is that the motion can be regarded qualitatively as a summation of turbulent eddies (corresponding to harmonics). Idealized models like that of homogeneous turbulence have been proposed where the random velocity field is assumed to be spatially stationary. This leads to a Fourier representation of the covariance function and the random field in homogeneous turbulence. Before continuing with this informal discussion, let us consider another heuristics with which turbulence is often analyzed. One introduces the concept of Reynold's number numbers for turbulent eddies of different sizes. Let h be the magnitude of a given eddy and v_h the corresponding velocity. The corresponding Reynold's number is taken to be $R_h = \frac{v_h h}{\nu}$. In the case of large eddies, R_h is large and the viscosity is thought to be unimportant. The viscosity is important for small eddies of magnitude h_0. Energy transfer and dissipation is thought to have the following character in turbulent flow. The energy passes from large eddies to the smaller ones with essentially no dissipation. Energy is dissipated in the smallest eddies and there the kinetic energy is transformed to heat. Many arguments in fluid mechanics use dimensional analysis. Let us discuss briefly Kolmogorov's argument [44]. The object is in part to specify the parameters relevant in turbulent flow over regions small relative to the overall scale l of the turbulence, but large relative to the distance h_0 at which viscosity is important. In the model of homogeneous turbulence with the implied spatial harmonic analysis, this corresponds to wave numbers λ in the range $l^{-1} << \lambda << h_0^{-1}$, the inertial subrange (Here, only for this section, $<<$ is "very much bigger than..."). One argues that the relevant parameters are wave-number and the mean dissipation of energy per unit time per unit mass of fluid ϵ. Consider estimating the order of magnitude v_h of turbulent velocity variation over distances of the order of h in the range $l > h > h_0$. The only obvious quantity in terms of ρ, ϵ and h having the dimensions of velocity is $(\epsilon h)^{1/3}$. In the model of homogeneous turbulence this suggests an approximate form of the spectrum in the inertial range. That is $\mathbb{E}(v(x)v(x+h)) \sim |\epsilon h|^{-1/3}$. This implies that the spectral density should have the form of the so-called Kolmogorov spectrum: $c|\lambda|^{-5/3}$ (averaged about a sphere of radius $|\lambda|$ in wave number space, $\lambda = (\lambda_1, \lambda_2, \lambda_3)$).

Mandelbrot's interpretation of $1/f$ *processes.* Motivated by some problems related to the design of electronic devices B. Mandelbrot studied in [50] the self-similar processes with spectrum $\sim \frac{1}{|\lambda|^\alpha}$. This type of noise appears in some measurements of current in thin metallic films and semiconductors. He observed that the ordinary theory for stationary processes of Wiener and Khinchine is not applicable to such a process. To reinterpret these spectral measurements without paradox, he choose an alternative approach and introduced the concept of "conditional spectrum".

Many elegant results on fractional processes, can be obtained using properties of fractional powers of the Laplacian operator or fractional integrals. Just one example, from formula 2.4.5 and the results of section 5.2.1 it can be given a short proof of the existence of the fractional Brownian field with exponent $\alpha/2$ [11].

Chapter 6

Stationary sequences and stable sampling

"Après moi, le déluge".

Louis XV.

6.1 Introduction

Let $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{A}}^{r=1\dots n}$ be a wide sense stationary (w.s.s.) n-dimensional random process, indexed in a "time" set $\mathbb{A}$. Many statistical problems such as linear prediction, interpolation or extrapolation, or computing conditional expectations for gaussian processes [69]-[21]-[38]-[68] are reduced to the problem of obtaining the best approximation of a random Variable $Y \in \overline{\mathrm{span}}\mathscr{X} = H(\mathscr{X})$ in terms of an element of a closed subspace of $H(\mathscr{X})$. Some authors e.g. [52] made clear the relationship between some of these subjects and the theory of shift-invariant subspaces. A related problem is that of reconstructing a continuous time random process/signal from discrete time samples. Analogous results to the classic Shannon sampling theorem, or its generalizations, can be given for stationary or related processes [47]-[45]-[69]. For example, the problem 1.0.1, stated in the introduction admits the following weaker reformulation: to find conditions under which $\overline{\mathrm{span}}\{X_k\}_{k\in\mathbb{Z}} = \overline{\mathrm{span}}\{X_t\}_{t\in\mathbb{R}}$. However, that a process is linearly determined by its samples is a necessary condition for a process to have a series representation like eq. 1.0.1. So it is natural to ask under what conditions Y can, in some reasonable sense, be represented in the form of a series

$$Y \sim \sum_{r=1}^{n}\sum_{t} a_r(t) X_t^r \ .$$

In the context of stationary processes, it is natural to formulate conditions on the spectral density, or spectral measure of the process. To treat some of these approximation problems Rozanov [69] introduced the concept of *Conditional basis*. Giving sufficient conditions [69]-[70] on the spectral density, this can be strengthened to an *unconditional basis* or Riesz Basis, that makes the series unconditionally convergent. In these derivations much care is given to

minimality, i.e. no element X_t^r of this system belongs to the closed linear span of the remaining elements. Minimal 1-dimensional processes were first introduced by Kolmogorov, and their structure can be characterized in terms of the spectral measure of the process [69]-[78]-[49]. Many interpolation or extrapolation problems are easier to handle when the processes involved are minimal. In the context of sampling, sufficient conditions on the spectral density are also used in e.g. [84] to construct a Wavelet Karhunen-Loéve like expansion for a w.s.s. process. The random variables obtained in this case are uncorrelated, i.e. they are orthogonal. Wavelet type expansions of random signals, and some of their properties, are also studied in [14] and [37]. However, in some applications mainly related to signal analysis, redundancy could be useful. As for example, a natural way to achieve this is by means of frames. The main results are given in section 3. First, in theorem 6.3.1 we give necessary and sufficient conditions in terms of the spectral measure for a stationary sequence to form a frame. This result on stability generalizes the results of Rozanov ([69], Chapter II sec. 7 and sec. 11). Second, the problem of reconstructing a random signal from its samples can be viewed as a problem of completeness of the sequence of samples in the closed linear span of the whole random process [47]-[45]. So, we study conditions for the stability of these sequences of samples. We relate the previous result on frames to the case of a sequence obtained from sampling a continuous parameter process. The conditions are obtained in terms of the periodized spectral density of the process (e.g. theorem 6.3.2). Finally, in theorem 6.3.4 we give conditions for a sequence of samples to be a frame of the closed linear span of the whole process.
As we will see the study of conditions for stationary sequences to form frames is similar, in some way, to the problem of the characterization of frames for shift invariant subspaces (SIS) of $L^2(\mathbb{R}^d)$ ([19]-[12]-[6]-[17]). Theorem 6.3.1 is an analogue to the results of sec. 3 of [6] or sec. 2 of [12]. In [12] the SIS are generated by more than one generator, then the conditions are given in terms of the Gramian matrix or the dual Gramian matrix. In our context the spectral density matrix of a stationary process will play a similar role. Finally, we note that the study conditions of frames and Riesz basis is very useful for the theory of sampling of signals, since these conditions are related to stability. For example in [48], several conditions for stable sampling in an union of shift invariant subspaces, are studied using Riesz basis. The theory of sampling in an union of subspaces gives an appropriate framework for problems related to sparse representations [48] and spectrum-blind sampling of multiband signals among other applications.

6.2 Auxiliary Results

6.2.1 Stationary processes revisited.

Recalling the definitions of section 2.4.5, let us see the form that these definitions and results take for the case of vector valued processes. This brief review is mainly borrowed from ([69], Chapter I). Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, let $\mathscr{X} = \{X_t^j\}_{t\in\mathbb{A}}^{j=1\ldots n}$ be a $n-dimensional$ wide sense stationary random process, where $\mathbb{A} = \mathbb{R}$ or $\mathbb{Z}$. By this we mean a family of random variables in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ (i.e. with finite variance) stationary correlated in the index t, i.e. : $R_{ij}(t,s) = \mathbb{E}(X_t^i X_s^j) = R_{ij}(t-s)$, for all t, s. We will also assume that $\mathbb{E}X_t^k = 0$ for all t, k. In the following, $\mathbb{G} = \mathbb{R}$ if $\mathbb{A} = \mathbb{R}$ or $\mathbb{G} = [0,1)$ if $\mathbb{A} = \mathbb{Z}$.

Harmonic Analysis of Stationary n-dimensional Random Processes:

Every stationary multidimensional process $\mathscr{X} = \{X_t^j\}_{t\in\mathbb{A}}^{j=1\ldots n}$ admits a spectral representation:

$$X_t^j = \int_{\mathbb{G}} e^{i2\pi\lambda t} d\Phi_j$$

in the form of an stochastic integral [42] with respect to a random spectral measure $\Phi = (\Phi_1, \ldots, \Phi_n)$. Moreover for each $t \in \mathbb{R}$ or $\mathbb{Z}$, X_t^r can be written as the result of the action of the (unitary) *time shift operator* T on X_0^r:

$$X_t^r = T^t X_0^r \quad \text{where by Stone's spectral theorem: } T^t = \int_{\mathbb{G}} e^{i\lambda 2\pi t} dE(\lambda)\,, \tag{6.2.1}$$

where the $E(\lambda)$'s are orthogonal projection operators over $H(\mathscr{X})$.
Set $\mu_{i\,j}(A) = \mathbb{E}(\Phi_i(A)\overline{\Phi_j(A)})$, $i, j = 1\ldots, n$. This matrix of complex measures is positive definite and we call it the *spectral measure* of the process $\mathscr{X}$. On the other hand the (cross) spectral measures $\mu_{r\,j}$ are also related by the following Fourier transform pairing

$$\mathbb{E}(X_t^r \overline{X_0^j}) = \int_{\mathbb{G}} e^{i\lambda 2\pi t} d\mu_{r\,j} \tag{6.2.2}$$

We study some properties of the Hilbert space $H(\mathscr{X})$ which is the closed linear span of $\mathscr{X}$ in $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Some properties are more easily characterized over an isometrically isomorphic space, defined as:

Definition 23.

$$\mathbf{L}^2(\mathbb{G}) = \left\{ f : \mathbb{G} \to \mathbb{C}^n, f \text{ is measurable and } \int_{\mathbb{G}} f \frac{d\mathbf{M}}{d\nu} f^* d\nu < \infty \right\} ,$$

where $\mathbf{M}$ is a $n \times n$ positive semi definite matrix of complex measures defined by $(\mathbf{M})_{i\,j}(A) = \mu_{i\,j}(A)$ for each $A \in \mathscr{B}(\mathbb{G})$, $\nu(A) = tr(\mathbf{M})(A)$ and $\frac{d\mathbf{M}}{d\nu}$ is the matrix of Radon Nykodym derivatives $\frac{d\mu_{i\,j}}{d\nu}$.

First note that always $\mu_{i\,j} << \nu$, so this assures the existence of the $\frac{d\mu_{i\,j}}{d\nu}$. In our case $\mathbf{M}$ will be the spectral measure of the process. Unless $\frac{d\mathbf{M}}{d\nu}(\lambda)$ is of full rank for almost all λ $[\nu]$, there may exist measurable f such that $f \neq 0$ over a set of positive ν measure with $\int_{\mathbb{G}} f \frac{d\mathbf{M}}{d\nu} f^* d\nu = 0$. Moreover if we do not distinguish between two vector functions f, g such that $\int_{\mathbb{G}} (f - g) \frac{d\mathbf{M}}{d\nu} (f - g)^* d\nu = 0$, then we can treat $\mathbf{L}^2(\mathbb{G})$ as a Hilbert space. More precisely $\mathbf{L}^2(\mathbb{G})/\{f \in \mathbf{L}^2 : f^* \in Nul(\frac{d\mathbf{M}}{d\nu})$ *a.e.* $[\nu]\}$ is a Hilbert space with the norm $\|f\|^2_{\mathbf{L}^2} = \int_{\mathbb{G}} f \frac{d\mathbf{M}}{d\nu} f^* d\nu$. The isomorphism between $\mathbf{L}^2(\mathbb{G})$ and $H(\mathscr{X})$ is given by an integral respect to the random measure Φ. That is, for every $Y \in H(\mathscr{X})$ there exists $f \in \mathbf{L}^2(\mathbb{G})$ such that

$Y = \sum_{j=1}^{n} \int_{\mathbb{G}} f_j d\Phi_j$. In the case that all the elements $\mu_{ij} << \mathcal{L}$, where $\mathcal{L}$ is Lebesgue measure, then we call the Radon-Nykodym derivatives ϕ_{ij} with respect to $\mathcal{L}$ *spectral densities*, and we say that $\mathscr{X}$ has a spectral density (matrix) D of elements ϕ_{ij}. Then, the integrals involving $\mathbf{M}$ introduced before can be written as $\int_{\mathbb{G}} fDf^* d\lambda$ and so on. The same discussion made for ν and $\mathbf{M}$ also holds for this case.

Finally, it is clear that a corresponding concept of stationarity can be considered for any process $\{X_g\}_{g\in G}$, with $\mathbb{E}X_g = 0$ and $\mathbb{E}|X_g|^2 < \infty$ and the index set G a group. The results, for w.s.s. processes, involving Fourier transforms are also extended to the case when G is a LCA group [68], Chapter I.

There exists a useful spectral characterization of minimal sequences for stationary sequences . This result also admits an extension to processes indexed over LCA groups [49]:

Theorem 6.2.1. *([69], Chapter II, sec. 11, and [49]) Let* $\mathscr{X} = \{X_k^j\}_{k\in\mathbb{Z}}^{j=1\ldots n}$ *be a w.s.s. process with spectral density* D*, if*

$$\exists\, D^{-1} \quad a.e.\ [\mathcal{L}] \quad and \int_{[0,1)} tr(D^{-1}) d\lambda < \infty \Longrightarrow \mathscr{X} \text{ is minimal}. \tag{6.2.3}$$

Remark. In [69] it is claimed, that condition 6.2.3 is necessary and sufficient, however the proof of the necessity part contains an error, see [49]. Theorem 6.2.1 is an immediate consequence of Corollary 4.9 of [49].

Some facts about periodic functions and Sampling

We will see that some properties of uniform sampling can be derived from the properties of periodic functions and measures, For this purpose it is useful to consider the quotient space $\mathbb{R}/\mathbb{Z}$. We will denote the *canonical* projection $\Pi : \mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Z}$, the map which assigns to every $x \in \mathbb{R}$ its equivalence class $\Pi(x)$. In our derivations it is useful to make the following convention: to identify $\Pi(x)$ with its *unique* representative in the interval $[0, 1)$. That is to consider Π as the following map: $\Pi : \mathbb{R} \longrightarrow [0, 1)$, $\Pi(x) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[0,1)+k}(x)(x - k)$.

Let $\mathbb{U} \subset \mathbb{R}$ be a Borel measurable set, then, the class of Borel subsets of $\mathbb{U}$ will be denoted by $\mathscr{B}(\mathbb{U})$. If μ is a complex measure, we recall that the induced measure Π is the measure defined for every Borel set $U \subset [0, 1)$ by the formula $\mu_\Pi(U) = \mu(\Pi^{-1}(U))$. Let f be a Borel measurable 1-periodic function, i.e. $f(x) = f(x + 1)$ for every $x \in \mathbb{R}$ (we will not distinguish between two functions which are equal at almost every x μ-a.e.). Then, if we denote $f|_{[0,1)}$ the restriction of f to the interval $[0, 1)$, i.e. $(f|_{[0,1)} \circ \Pi)(x) = (f \circ \Pi)(x) = f(\Pi(x)) = f(x)$ for every $x \in \mathbb{R}$.

Given a continuous time process $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}$, with random measure Φ over $\mathbb{R}$, if we consider the sequence of samples $\{X_k^r\}_{k\in\mathbb{Z}}$ as a discrete stationary sequence, then we have two possible spectral representations for every $k \in \mathbb{Z}$:

$$X_k^j = \int_{\mathbb{R}} e^{i2\pi\lambda k} d\Phi_j = \int_{[0,1)} e^{i2\pi\lambda k} d\Phi'_j \text{ a.s.}$$

where Φ'_j is a random measure over $[0,1)$. From this, by a density argument this is also true for 1-periodic functions, in particular if $A \in \mathscr{B}[0,1)$:

$$\int_{\mathbb{R}} \mathbf{1}_{\Pi^{-1}(A)} d\Phi_r \int_{\mathbb{R}} \mathbf{1}_{\Pi^{-1}(A)} d\overline{\Phi_j} = \int_{[0,1)} \mathbf{1}_A d\Phi'_r \int_{[0,1)} \mathbf{1}_A d\overline{\Phi'_j} \text{ a.s.}$$

Then, taking expected values in both sides of the equality, if we denote $\mu'_{r\,j}$ the cross spectral measure of the discrete sequence of samples, we have: $\mu'_{r\,j}(A) = \mu_{r\,j}(\Pi^{-1}(A))$, where $\mu_{r\,j}$ is the cross spectral measure of the original process. On the other hand, as we have seen before, given a continuous time process $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}$, $H(\mathscr{X})$ is isomorphic to $\mathbf{L}^2(\mathbb{R})$, and this induces an isomorphism between the closed subspace $\overline{\text{span}}\{X_k^r\}_{k\in\mathbb{Z}}$ and the closed subspace of $\mathbf{L}^2(\mathbb{R})$: $\{f \in \mathbf{L}^2(\mathbb{R}) : f \textit{ is } 1-\textit{periodic}\}$. Additionally, taking in account the previous discussion about periodic functions, both subspaces are isometrically isomorphic to the Hilbert space $\mathbf{L}^2[0,1) = \left\{ f : [0,1) \to \mathbb{C}^n, f \text{ is measurable and } \int_{[0,1)} f \frac{d\mathbf{M_\Pi}}{d\nu_\Pi} f^* d\nu_\Pi < \infty \right\}$. This suggests that some properties may be characterized in terms of induced measures. We begin with the following result:

Proposition 6.2.1. *Let μ be a complex measure (Borel) over $\mathbb{R}$ If $(\mu_\Pi)_s$ and $(\mu_\Pi)_{ac}$ denote the singular and absolutely continuous parts of μ_Π respectively (with respect to Lebesgue measure). Then:*
a) For every Borel set U in $[0,1)$: $(\mu_\Pi)_s(U) = \mu_s(\Pi^{-1}(U))$ and $(\mu_\Pi)_{ac}(U) = \mu_{ac}(\Pi^{-1}(U))$. The measures μ_s and μ_{ac} denote the singular and absolutely continuous parts of μ respectively. That is, the singular part of the induced measure by Π is the induced measure by Π through the singular part of μ, the same with the absolutely continuous part.
b) If $w \in L^1(\mathbb{R})$ is the Radon-Nykodym (R-N) derivative of μ_{ac} respect to Lebesgue measure, then $\sum_{k\in\mathbb{Z}} w(.+k)$ is the R-N derivative of $(\mu_\Pi)_{ac}$.

Notation. Given a complex measure μ, the total variation of μ is $|\mu|$.

Proof. a) It is sufficient to find $Z, Z' \in \mathscr{B}([0,1))$ such that $Z \cup Z' = [0,1)$, $Z \cap Z' = \varnothing$ and $|(\mu_s)_\Pi|(Z) = \mathcal{L}(Z') = 0$. We have, that there exists $U, U' \in \mathscr{B}(\mathbb{R})$ such that $U \cup U' = \mathbb{R}$, $U \cap U' = \varnothing$ and $|(\mu_s)|(U) = \mathcal{L}(U') = 0$. We claim that $Z' = \Pi(U')$ and $Z = [0,1) \setminus Z'$.
To prove this fact, first note that: if μ is a complex measure, $\{A_j\}_{j=1\ldots n}$ is any (measurable) partition of $[0,1)$ and E any measurable subset of $[0,1)$, then $\{\Pi^{-1}(A_j)\}_{j=1\ldots n}$ is a partition of $\mathbb{R}$. So we have

$$\sum_{j=1}^n |\mu_\Pi(A_j \cap E)| = \sum_{j=1}^n |\mu(\Pi^{-1}(A_j) \cap \Pi^{-1}(E)| \le |\mu|(\Pi^{-1}(E)) = |\mu|_\Pi(E) .$$

Then if we take the supremum over all possible partitions of $[0,1)$, in the left-hand side of the chain of inequalities we get

$$|\mu_\Pi|(E) \le |\mu|_\Pi(E) .$$

In particular, $|(\mu_s)_\Pi|(Z) \le |\mu_s|_\Pi(Z)$. On the other hand, $\Pi^{-1}(Z) = \mathbb{R}\setminus\Pi^{-1}(Z') = (\Pi^{-1}(Z'))^c$, but $U' \subseteq \Pi^{-1}(Z')$ then $(\Pi^{-1}(Z'))^c \subseteq (U')^c = U$. From this we have,

$$|(\mu_s)_\Pi|(Z) \le |\mu_s|_\Pi(Z) = |\mu_s|(\Pi^{-1}(Z)) \le |\mu_s|(U) = 0 .$$

On the other hand, let $I_k = [k, k+1)$, $k \in \mathbb{Z}$, then $U' = \bigcup_{k\in\mathbb{Z}} I_k \cap U'$, hence:

$$Z' = \Pi(U') = \bigcup_{k\in\mathbb{Z}} [(I_k \cap U') - k] \,.$$

Since the Lebesgue measure is invariant under translations, then $\mathcal{L}[(I_k \cap U') - k] = \mathcal{L}(I_k \cap U') \leq \mathcal{L}(U') = 0$, and:

$$\mathcal{L}(Z') = \mathcal{L}\left(\bigcup_{k\in\mathbb{Z}} [(I_k \cap U') - k]\right) \leq \sum_{k\in\mathbb{Z}} \mathcal{L}[(I_k \cap U') - k] = 0 \,,$$

so that $(\mu_s)_\Pi \perp \mathcal{L}$.
Now, we prove that $(\mu_{ac})_\Pi << \mathcal{L}$. Take $W \in \mathscr{B}[0,1)$ such that $\mathcal{L}(W) = 0$. Again, by the translation invariance property of the Lebesgue measure: $\mathcal{L}(W+k) = \mathcal{L}(W) = 0$, so $\mu_{ac}(W+k) = 0$ for every $k \in \mathbb{Z}$, since $\mu_{ac} << \mathcal{L}$ over $\mathbb{R}$. Then:

$$(\mu_{ac})_\Pi(W) = \mu_{ac}(\Pi^{-1}(W)) = \mu_{ac}\left(\bigcup_{k\in\mathbb{Z}} W + k\right) \leq \sum_{k\in\mathbb{Z}} \mu_{ac}(W+k) = 0 \,. \tag{6.2.4}$$

Finally, we have:

$$\mu_\Pi = (\mu_{ac})_\Pi + (\mu_s)_\Pi \,. \tag{6.2.5}$$

The equations given above, together with the uniqueness of the Lebesgue decomposition of a measure, show that 6.2.5 must be the Lebesgue decomposition of μ_Π. The result of part a) follows from this.
Part b)
Is immediate. □

The following is proved in the appendix:

Lemma 6.2.1. *Let $\mathscr{X} = \{X_k^j\}_{k\in\mathbb{Z}}^{j=1\ldots n}$ be a w.s.s. stationary in the variable k, if D is the matrix of spectral densities and if $P_{Col(D)}(\lambda)$ is the orthogonal projection matrix over $Col(D)$ for each λ, then: a) The Moore Penrose pseudo-inverse $D^\sharp$ and $P_{Col(D)}$ are measurable. b) If $0 \in \sigma(D)(\lambda)$ for all λ in a set of positive Lebesgue measure, then there exists a column of $P_{Nul(D)}$ and a measurable set A, such that: $\mathcal{L}(A) > 0$ and $(P_{Nul(D)})_j(\lambda) \neq 0$ for all $\lambda \in A$. c) If $Y = H(\mathscr{X})$ admits the following representation for some $f \in \mathbf{L}^2$: $Y = \sum_{j=1}^{n} \int_{[0,1)} f_j d\Phi_j$ then for all $g \in \mathbf{L}^2$ such that $g \in Nul(D)$ a.e. : $Y = \sum_{j=1}^{n} \int_{[0,1)} (f_j + g_j) d\Phi_j$, in particular Y can be written as $Y = \sum_{j=1}^{n} \int_{[0,1)} (P_{Col(D)} f^*)_j d\Phi_j$.*

Remark. The same holds in $\mathbb{R}$ (continuous parameter case) or for the derivative of the matrix measure $\mathbf{M}$ with respect to ν.

6.3 Main Results.

Now, we can give a necessary and sufficient condition for a stationary sequence to form a frame in terms of its spectral measure. In e.g. ([69], Chapter II, sec. 7, [70]), it is proved that if a stationary sequence has an spectral density (matrix) which has all its eigenvalues inside an interval $[A, B]$ a.e. then it is a Riesz basis of its span. This stability condition entails many consequences [79]. For example: the linear predictor, for any time lag, and the innovation process are expressible as the sum of mean-convergent infinite series.
The following theorem generalizes the result of [69] in two ways. First, we obtain a similar condition, on the spectral density, for frames. And second, it is proved that Rozanov's sufficient condition of the spectral measure being absolutely continuous is also necessary. The original result of [69]-[70] is contained in one of the implications of part b) of theorem 6.3.1, however in our case, the result will be derived from the first result for frames. On the other hand, this theorem resembles theorem 3.4 and prop. 3.2 of [6], or theorem 2.5. of [12] for shift invariant subspaces of $L^2(\mathbb{R}^d)$.

Theorem 6.3.1. *Let $\mathscr{X} = \{X_k^j\}_{k\in\mathbb{Z}}^{j=1\ldots n}$ be stationary in the variable k. Then: a) $\mathscr{X}$ is a frame of its span $H(\mathscr{X})$ (in $L^2(\Omega, \mathcal{F}, \mathbf{P})$) with constants $A, B \iff$ the (cross) spectral measures $\mu_{i\,j}$ verify the following conditions:(i) $\mu_{i\,j} << \mathcal{L}$, (ii) the spectral densities matrix $D_{i\,j} = \frac{d\mu_{i\,j}}{d\mathcal{L}}$ verifies $\sigma(D)(\lambda) \subseteq \{0\} \cup [A, B]$ for almost all $\lambda \in [0, 1)$ $[\mathcal{L}]$.*
b) $\mathscr{X}$ is a Riesz basis with constants $A, B \iff$ (i) $\mu_{i\,j} << \mathcal{L}$, (ii) the spectral densities matrix $D_{i\,j} = \frac{d\mu_{i\,j}}{d\mathcal{L}}$ verifies $\sigma(D)(\lambda) \subseteq [A, B]$ for almost all $\lambda \in [0, 1)$ $[\mathcal{L}]$

Proof. *Part a)*($\Rightarrow$) If we suppose that $\mathscr{X}$ is a frame, then, given $Y \in H(\mathscr{X})$,

$$A\mathbb{E}|Y|^2 \leq \sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} |\mathbb{E}(X_k^j \overline{Y})|^2 \leq B\mathbb{E}|Y|^2 \ , \tag{6.3.1}$$

On the other hand we know that Y admits the following representation: $Y = \sum_{j=1}^{n} \int_{[0,1)} f_j d\Phi_j$, where $f = (f_1, \ldots, f_n) \in \mathbf{L}^2[0, 1)$ and the spectral random measures verify $X_k^j = \int_{[0,1)} e^{i2\pi\lambda k} d\Phi_j$, and $\mathbb{E}(\overline{Y} X_k^j) = \sum_{j=1}^{n} \int_{[0,1)} \overline{f}_j e^{i2\pi\lambda k} d\mu_{r\,j}$. Recall that the latter can be written as,

$$\int_{[0,1)} e^{i2\pi\lambda k} \sum_{j=1}^{n} \overline{f}_j \frac{d\mu_{r\,j}}{d\nu} d\nu \ , \tag{6.3.2}$$

then

$$\sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} |\mathbb{E}(X_k^j \overline{Y})|^2 = \sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} \left| \int_{[0,1)} e^{i2\pi\lambda k} \sum_{j=1}^{n} \overline{f}_j \frac{d\mu_{r\,j}}{d\nu} d\nu \right|^2 \leq B\mathbb{E}|Y|^2 < \infty \ ,$$

Hence 6.3.2 gives the Fourier coefficients of a function belonging to $L^2[0, 1)$ (thus in $L^1[0, 1)$). In particular, if we take $f = e_r = (0, \ldots, 0\,, 1\,, 0 \ldots)$, the vector which is zero in all its coordinates but for the r-th. coordinate, we get that $(\hat{\mu}_{r\,r}(k))_k \in l^2(\mathbb{Z})$ are the Fourier coefficients of an integrable function. Then by the uniqueness theorem of the Fourier transform of measures, we have that $\mu_{r\,r} << \mathcal{L}$ and then $\nu = \sum_{k=1}^{n} \mu_{k\,k} << \mathcal{L}$. From this: $\mu_{r\,j} << \mathcal{L}$, so by the Radon-Nykodym theorem there exist derivatives (spectral densities) $\frac{d\mu_{r\,j}}{d\mathcal{L}} \in L^1[0,1)$. So if $D_{i\,j}(\lambda) = \frac{d\mu_{i\,j}}{d\mathcal{L}}(\lambda)$ is the matrix of spectral densities, we have that:

$$\mathbb{E}(YX_k^j) = \int_{[0,1)} e^{i2\pi\lambda k} e_j D f^* d\lambda \quad \text{and} \quad \mathbb{E}|Y|^2 = \int_{[0,1)} f D f^* d\lambda \,.$$

Then,

$$A \int_{[0,1)} f D f^* d\lambda \leq \sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} \left| \int_{[0,1)} e^{i2\pi\lambda k} e_j D f^* d\lambda \right|^2 \leq B \int_{[0,1)} f D f^* d\lambda \,,$$

and from Parseval's identity we have,
$\sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} \left| \int_{[0,1)} e^{i2\pi\lambda k} e_j D f^* d\lambda \right|^2 = \int_{[0,1)} \|Df^*\|^2 \, d\lambda$. Hence we can write,

$$A \int_{[0,1)} f D f^* d\lambda \leq \int_{[0,1)} \|Df^*\|^2 \, d\lambda \leq B \int_{[0,1)} f D f^* d\lambda \tag{6.3.3}$$

Now, given $x \in \mathbb{C}^n$ we can take $\epsilon > 0$, $\lambda_0 \in [0,1]$ and then if we replace in eq. 6.3.3 $f = x\mathbf{1}_{B(\lambda_0,\epsilon)}$, by Lebesgue's differentiation theorem, we have that $\forall x \in \mathbb{C}^n$ there exists a measurable set F_x such that $\mathcal{L}(F_x^c) = 0$ and

$$AxD(\lambda)x^* \leq \|D(\lambda)x^*\|^2 \leq BxD(\lambda)x^* \;\; \forall\, \lambda \in \; F_x. \tag{6.3.4}$$

Take $\mathcal{D}$ a countable dense subset of $\mathbb{C}^n$, and define $F = \bigcap_{x\in\mathcal{D}} F_x$, then clearly $\mathcal{L}(F^c) = 0$ and given any $\lambda \in F$ inequality 6.3.4 holds for every $x \in \mathcal{D}$. On the other hand, for each $\lambda \in F$, $h(x) = CxDx^* - \|Dx^*\|^2$ is continuous, so given $\lambda \in F$ the inequality must hold for every $x \in \mathbb{C}^n$. But since D is self adjoint, then $Nul(D) = Col(D)^{\perp}$, so $Col(D) \oplus Nul(D) = \mathbb{C}^n$, a.e. Now, if $v^* \notin Nul(D)$ is an eigenvector associated to $z \in \sigma(D) \setminus \{0\}$, $\|Dv^*\|^2 = z^2 \|v^*\|^2$ and $vDv^* = z\|v\|^2$, then as we have seen, eq. 6.3.4 holds for every $\lambda \in F$ and $x \in \mathbb{C}^n$, thus, $A \leq z \leq B$.
($\Leftarrow$) It is easy to see that assuming eq. 6.3.4, from eqs. 6.3.4 and 6.3.1 one can reverse the argument .
Part b) ($\Rightarrow$) If $\mathscr{X}$ is a Riesz basis then it is a frame (Theorem 2.3.4). From part (a) we have that $\mu_{i\,j} << \mathcal{L}$ and $\sigma(D) \subseteq \{0\} \cup [A,B]$ a.e.. Let us check that $0 \notin \sigma(D)$ a.e., if this is not the case, by lemma 6.2.1 we can take a column $g^* = (P_{Nul(D)})_j \neq 0$ over a set of positive $\mathcal{L}$ measure, moreover we can suppose $\|g\| = 1$ over some A of positive measure. If $c_k^r(g) = \int_{[0,1)} e^{-i\lambda 2\pi k} g_r(\lambda) d\lambda$ then there exists k, r such that $c_k^r(g) \neq 0$, and on the other hand $(c_k^r(g))_k \in l^2(\mathbb{Z})$. Now we can define $Y \in H(\mathscr{X})$ as $Y = \sum_{j=1}^{n} \int_{[0,1)} g_j d\Phi_j$, clearly $\mathbb{E}|Y|^2 = \int_{[0,1)} gDg^* d\lambda = 0$. If we show that $Y = \sum_{k\in\mathbb{Z}} \sum_{r=1}^{n} c_k^r(g) X_k^r = 0$ we are done, since from theorem

2.3.5 then $\{X^r_k\}^r_{k\in\mathbb{Z}}$ can't be a frame. Let us define for $r = 1,\dots,n$, $g_{N\,r} = \sum_{|k|\le N} c^r_k(g)e^{i2\pi\lambda k}$, $g_N = (g_{N\,1},\dots,g_{N\,n})$, and finally $Y_N = \sum_{j=1}^n \int_{[0,1)} g_{N\,j}d\Phi_j = \sum_{|k|\le N}\sum_{r=1}^n c^r_k(g)X^r_k$. The result will follow if we show that $\mathbb{E}|Y_N - Y|^2 \underset{N\to\infty}{\to} 0$. This is true since $M(\lambda) = \sup_{\|x\|=1} xD(\lambda)x^*$ is measurable and $M \le B$ a.e. and on the other hand $g_{N,r} \to g_r \in L^2[0,1)$ converges in $L^2[0,1)$ and a.e. from which we have, $\mathbb{E}|Y_N - Y|^2$

$$= \int_{[0,1)} (g-g_N)D(g-g_N)^*d\lambda \le \int_{[0,1)} \|g_N - g\|^2 M d\lambda \le B\int_{[0,1)} \|g_N - g\|^2 d\lambda \underset{N\to\infty}{\to} 0\,.$$

($\Leftarrow$) Under these conditions, again from part (a), we know that $\mathscr{X}$ is a frame. On the other hand $\sigma(D) \subseteq [A,B]$ a.e. implies that $\exists\, D^{-1}$ a.e.. Moreover $\sigma(D^{-1}) \subseteq [B^{-1}, A^{-1}]$. Then if we call Λ_i $i = 1\dots n$ to the eigenvalues of D taking in account their multiplicity, we have that

$$\frac{1}{B} \le tr(D^{-1}) = \sum_{i=1}^n \Lambda_i^{-1} \le \frac{n}{A} \ \text{ a.e. in } [0,1)\ [\mathcal{L}] \implies \int_{[0,1)} tr(D^{-1})d\lambda < \infty\,.$$

This proves the result, since by theorem 6.2.1 this implies that $\mathscr{X}$ is minimal, and then a Riesz basis, by theorem 2.3.5.

□

Sampling.

Given a continuous time stationary process we can give conditions for samples taken at an uniform rate to form a frame. This is related to reconstructing a continuous parameter process from its samples. Recall in section 6.2 we noted that the spectral measure of a sampled process is $(\mu_{i\,j})_\Pi$, where $\mu_{i\,j}$ is the spectral measure of the original continuous time process. Let us prove the following result:

Lemma 6.3.1. *Given the spectral measures $\mu_{i\,j}$ over $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$: $\mu_{i\,j} << \mathcal{L} \iff (\mu_{i\,j})_\Pi << \mathcal{L}$.*

Proof. In general, if $f : (X,\Sigma) \longrightarrow (Y,\Sigma')$ is a map between measurable spaces we have that given a (positive) measure ν defined over Σ then $\nu \equiv 0 \iff \nu_f \equiv 0$, where ν_f is the induced measure by f. Since given $A' \in \Sigma'$ then $\nu_f(A') = \nu(f^{-1}(A'))$ and on the other hand given $A \in \Sigma$ then $\nu(A) \le \nu(f^{-1}(f(A))) = \nu_f(f(A))$. In our case, taking $f = \Pi$, and $\nu = \sum_j \mu_{j\,j}$ as in def. 23, we have that $(\nu_s)_\Pi \equiv 0 \iff \nu_s \equiv 0$, but by prop. 6.2.1 $(\nu_s)_\Pi = (\nu_\Pi)_s$ then $(\nu_\Pi)_s \equiv 0 \iff \nu_s \equiv 0$ which is equivalent to the result, since $\mu_{i\,j} << \nu$ and $(\mu_{i\,j})_\Pi << \nu_\Pi$. □

Theorem 6.3.2. *Let $\mathscr{X} = \{X^r_t\}^{r=1,\dots,n}_{t\in\mathbb{R}}$ be a continuous time stationary process, and let $\mathscr{Y} = \{X^r_k\}^{r=1,\dots,n}_{k\in\mathbb{Z}}$ then the following are equivalent:*
a) $\mathscr{Y}$ is a frame of its span $H(\mathscr{Y})$.
b) $\mu_{i\,j} << \mathcal{L}$, where $\mu_{i\,j}$ are the spectral measures of $\mathscr{X}$, and there exists $A,B > 0$ such that D_Π the matrix of periodized spectral densities verifies $\sigma(D_\Pi) \subseteq \{0\} \cup [A,B]$ a.e $[\mathcal{L}]$.

Proof. This is immediate from the previous lemma 6.3.1 and theorem 6.3.1. □

Let us study under which conditions, given a w.s.s. process $\mathscr{X} = \{X^r_t\}^{r=1\dots n}_{t\in\mathbb{R}}$, there exists another stationary process (indexed by $k \in \mathbb{Z}$) $\mathscr{W} = \{W^r_k\}^{r=1\dots m}_{k\in\mathbb{Z}}$ such that $\overline{\text{span}}\{W^r_k\}^{r=1\dots m}_{k\in\mathbb{Z}} = \overline{\text{span}}\{X^r_k\}^{r=1\dots n}_{k\in\mathbb{Z}}$. For this purpose, let us review some facts about linear transformations over the space $H(\mathscr{X})$ spanned by a process $\mathscr{X}$. For more details see for example, [69], Chapter I, sec. 8.
Let $\mathscr{X} = \{X^k_t\}^{k=1\dots n}_{t\in\mathbb{A}}$ be a stationary random process with an spectral representation

$$X^k_t = \int_{\mathbb{G}} e^{i2\pi\lambda t} d\Phi_{\mathscr{X}\,k}\,,$$

we say that a process $\mathscr{Y} = \{Y^k_t\}^{k=1\dots m}_{t\in\mathbb{A}}$ is obtained by a *linear transformation*, if each of its components admits the following representation:

$$Y^k_t = \int_{\mathbb{G}} e^{i2\pi\lambda t} \phi^{\mathscr{Y}\,\mathscr{X}}_k\, d\Phi_{\mathscr{X}}\,, \tag{6.3.5}$$

for some vector functions $\phi^{\mathscr{Y}\,\mathscr{X}}_k = (\phi^{\mathscr{Y}\,\mathscr{X}}{}_{k\,j})_{j=1\dots n} \in \mathbf{L}^2(\mathbb{G})$. The equation 6.3.5 can be written in the matrix form:

$$Y_t = \int_{\mathbb{G}} e^{i2\pi\lambda t} \phi^{\mathscr{Y}\,\mathscr{X}}\, d\Phi_{\mathscr{X}}\,, \tag{6.3.6}$$

where the matrix $\phi^{\mathscr{Y}\,\mathscr{X}}$ acts symbolically over the vector random measure $\Phi_{\mathscr{X}}$. Being a stationary process $\mathscr{Y}$ has its own spectral representation, $Y^k_t = \int_{\mathbb{G}} e^{i2\pi\lambda t} d\Phi_{\mathscr{Y}}$, but it follows from eq. 6.3.5 that the random spectral measure of $\mathscr{Y}$ verifies

$$\Phi_{\mathscr{Y}\,k}(A) = \int_A \phi^{\mathscr{Y}\,\mathscr{X}}_k\, d\Phi_{\mathscr{X}}\,.$$

From this we have that the spectral measures of $\mathscr{Y}$ are related by

$$\mu^{\mathscr{Y}}{}_{ij}(A) = \int\limits_A \phi^{\mathscr{Y}\,\mathscr{X}}_i \frac{d\mathbf{M}_{\mathscr{X}}}{d\nu} (\phi^{\mathscr{Y}\,\mathscr{X}}_j)^* d\nu\,.$$

The measurable matrix $\phi^{\mathscr{Y}\,\mathscr{X}}$ is called the spectral characteristic of the linear transformation being considered. The processes $\mathscr{X}$ and $\mathscr{Y}$ are obviously stationarily correlated. If we write in a matrix form $\mathbf{M}_{\mathscr{X}\mathscr{Y}}$ their joint spectral (matrix) measure and the spectral (matrix) measure of $\mathscr{Y}$ then we have the following relations:

$$\mathbf{M}_{\mathscr{Y}\,\mathscr{X}}(A) = \int_A \phi^{\mathscr{Y}\,\mathscr{X}} \frac{d\mathbf{M}_{\mathscr{X}}}{d\nu} d\nu, \quad \mathbf{M}_{\mathscr{X}}(A) = \int_A \phi^{\mathscr{Y}\,\mathscr{X}} \frac{d\mathbf{M}_{\mathscr{X}}}{d\nu} (\phi^{\mathscr{Y}\,\mathscr{X}})^* d\nu\,.$$

It can be proved that this conditions are also sufficient in order that $\mathscr{Y}$ be obtainable from $\mathscr{X}$ from a linear transformation [69]. Let us introduce some notation. If $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal non negative matrix, for $\alpha \in \mathbb{R}$ we define $\Lambda^{(\alpha)}$ as $\Lambda^{(\alpha)}_{t\,l} = \Lambda^{\alpha}_l$ if $t = l$ and $\Lambda_l \neq 0$, or $\Lambda^{(\alpha)}_{t\,l} = 0$ otherwise. If $A \in \mathbb{R}^{n\times n}$ is a (symmetric) non negative definite matrix, A admits a diagonal decomposition $A = P\Lambda P^*$, from this, we define $A^{(\alpha)} = P\Lambda^{(\alpha)}P^*$.

Lemma 6.3.2. *Let $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n}$ be a w.s.s. process. $\mathscr{X}$ contains a stationary sequence which is a frame of the closed linear span of $\mathscr{Y} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$, $H(\mathscr{Y}) \Longleftrightarrow$ the spectral measures verify, for all $i\,j$: $\mu_{i\,j} << \mathcal{L}$.*

Proof. ($\Rightarrow$) We denote $\mathscr{Y} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$. If $\mathscr{Y}$ contains a stationary sequence, say $\mathscr{W}$, which is a frame of its span, then by the previous results, theorem 6.3.1, the (matrix) spectral measure associated to $\mathscr{W}$ must be absolutely continuous, so there exists a spectral density matrix $D^{\mathscr{W}}$, defined over $[0,1)$. On the other hand if $\mathscr{Y}$ is obtainable by a linear transformation from $\mathscr{W}$, say a $n \times m$ measurable matrix function $\varphi^{\mathscr{Y}\,\mathscr{W}}$, then the spectral measures of $\mathscr{Y}$ are given by [69]:

$$\forall\, A \in \mathscr{B}[0,1)\ :\ \mu'_{i\,j}(A) = \int_A \varphi_i^{\mathscr{Y}\,\mathscr{W}} D^{\mathscr{W}} (\varphi_j^{\mathscr{Y}\,\mathscr{W}})^* d\lambda\ ,$$

so $\mu'_{i\,j} << \mathcal{L}$, but since $\mu'_{i\,j} = (\mu_{i\,j})_\Pi$ then by the claim 6.3.1: $\mu_{i\,j} << \mathcal{L}$.
($\Leftarrow$) The spectral density matrix of $\mathscr{Y}$, $D_\Pi^{\mathscr{X}}$ exists since $\mu_{i\,j} << \mathcal{L}$ is diagonalizable in a measurable form, i.e. there exist a measurable [67] P orthogonal matrix and a diagonal matrix of eigenvalues Λ such that $\Lambda = P^* D_\Pi^{\mathscr{X}} P$. Take $\mathscr{W}$ the process obtained from $\mathscr{Y}$ by the linear invertible transformation[a] [69] induced by $G^{(-\frac{1}{2})} = P\Lambda^{(-\frac{1}{2})}P^*$. Then $G^{(-\frac{1}{2})} D_\Pi^{\mathscr{X}} (G^{(-\frac{1}{2})})^* = P\Lambda^{(-\frac{1}{2})}\Lambda\Lambda^{(-\frac{1}{2})}P^* = B$, where B is a diagonal matrix which takes the values 0 or 1 in the diagonal. From this we have that this linear operation is well defined since each column of $G^{(-\frac{1}{2})}$ is in $\mathbf{L}^2[0,1)$. On the other hand by the result [69] which relates a linear transformation and the spectral measure: $D^{\mathscr{W}} = B \Rightarrow \sigma(D^{\mathscr{W}}) \subseteq \{0,1\}$ a.e., then, by theorem 6.3.1, $\mathscr{W}$ is a frame of $H(\mathscr{Y})$. □

In previous works e.g.[69]-[84] a common condition for the existence of orthogonal or minimal stationary sequences, is that the eigenvalues of the spectral density be not null a.e. In theorem 6.3.1 we proved that the spectral density matrix of a stationary sequence which is a Riesz basis must have all its eigenvalues inside a positive bounded interval $[A,B]$. However there are some limitations on the supports of the spectral densities:

Lemma 6.3.3. *Let $\mathscr{X} = \{X_k\}_{k\in\mathbb{Z}}$ be a w.s.s. sequence, with absolutely continuous spectral measure, and let $\mathscr{Z} = \{Z_k\}_{k\in\mathbb{Z}}$ be another w.s.s. sequence such that $H(\mathscr{X}) = H(\mathscr{Z})$, then $\mathscr{Z}$ has an absolutely continuous spectral measure, and if $\phi^{\mathscr{X}}$ and $\phi^{\mathscr{Z}}$ are the spectral densities of $\mathscr{X}$ and $\mathscr{Z}$ respectively, then*

$$\mathcal{L}(\mathrm{supp}(\phi^{\mathscr{X}}) \triangle \mathrm{supp}(\phi^{\mathscr{Z}})) = 0.$$

Proof. From 6.3.2 there exists $\mathscr{Y} = \{Y_k\}_{k\in\mathbb{Z}}$ a stationary sequence which is a frame of $H(\mathscr{Z})$. Let us call $\Phi_{\mathscr{Z}}$ and $\Phi_{\mathscr{X}}$ the random spectral measures of $\mathscr{Z}$ and $\mathscr{X}$ respectively. Then as we have seen in lemma 6.3.2,

$$Y_k = \int_{[0,1)} (\phi^{\mathscr{Z}})^{-\frac{1}{2}} \mathbf{1}_{\mathrm{supp}(\phi^{\mathscr{Z}})} e^{i2\pi\lambda k} d\Phi_{\mathscr{Z}}\ .$$

On the other hand, as $Z_k \in H(\mathscr{X})$, Z_k admits the following representation:

$$Y_k = \int_{[0,1)} f e^{i2\pi\lambda k} d\Phi_{\mathscr{X}}\ ,$$

[a]Note that the operation is always defined over the closed linear span of the process

for some $f \in L^2(\mathbb{R}, \phi^{\mathscr{X}} d\lambda)$. Then $\phi^{\mathscr{Y}} = \mathbf{1}_{\text{supp}(\phi^{\mathscr{Z}})}$ a.e. $[\mathcal{L}]$. But [69] $\phi^{\mathscr{Z}} = |f|^2\phi^{\mathscr{X}}$ a.e. (in particular $\mathscr{Z}$ has an absolutely continuous spectral measure), then $\phi^{\mathscr{Y}} = \frac{|f|^2}{\phi^{\mathscr{Z}}}\mathbf{1}_{\text{supp}(\phi^{\mathscr{Z}})}\phi^{\mathscr{X}} = \mathbf{1}_{\text{supp}(\phi^{\mathscr{Z}})}$ a.e. From this $\mathcal{L}(\text{supp}(\phi^{\mathscr{Z}}) \setminus \text{supp}(\phi^{\mathscr{X}})) = 0$. Interchanging the roles of $\mathscr{Z}$ and $\mathscr{X}$ we get $\mathcal{L}(\text{supp}(\phi^{\mathscr{X}})\triangle\text{supp}(\phi^{\mathscr{Z}})) = 0$. □

We have seen that the closed linear span of a stationary sequence always contains a stationary sequence which is a frame. An arbitrary stationary sequence may not contain any *stationary* Riesz basis. Moreover, from lemmas 6.3.3, 6.3.2 and theorem 6.3.1, it is immediate that a necessary and sufficient condition for a 1-dimensional stationary sequence to contain a (stationary) Riesz basis of its span, is to have a spectral measure equivalent to $\mathcal{L}$.

Fundamental Frame

We give conditions for a sequence of samples to be a fundamental frame, i.e. to be a frame and to be complete with respect to the closed linear span of the continuous time process, i.e.

$$\overline{\text{span}}\{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n} = \overline{\text{span}}\{X_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n} = H(\mathscr{X}). \tag{6.3.7}$$

This is related to reconstructing a signal from its samples, since the sampling problem can be formulated in terms of finding conditions for which eq. 6.3.7 holds [47]-[45]. The following theorem 6.3.3 was proved by Lloyd [47], and it is an analogue result of the result of [5] for $L^2(\mathbb{R})$ functions:

Theorem 6.3.3. *[47] Let $\mathscr{X} = \{X_t\}_{t\in\mathbb{R}}$ be a w.s.s. process with spectral measure μ. The following are equivalent:*
i) $\overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$, (in $L^2(\Omega, \mathcal{F}, \mathbf{P})$).
ii) $\overline{\text{span}}\{e^{i2\pi kx}\}_{k\in\mathbb{Z}} = \{f \in L^2(\mu) : f \text{ is } 1-periodic\}$ [b] $= L^2(\mu)$.
iii) There exists $A \in \mathscr{B}(\mathbb{R})$ such that $\mu(A^c) = 0$ and $A \cap A + k = \varnothing$, $k \in \mathbb{Z}\setminus\{0\}$.

Proof. The equivalence $i) \iff ii)$ is immediate, so let us prove $i)$ (or $ii)$)$\Longrightarrow iii)$. Let $\mathbb{V} = \overline{\text{span}}\{e^{i2\pi kx}\}_{k\in\mathbb{Z}}$, and let $P_{\mathbb{V}} : L^2(\mu) \longrightarrow \mathbb{V}$ be the orthogonal projection on $\mathbb{V}$. We will first derive some useful expressions for $P_{\mathbb{V}}$. Given $f \in L^2(\mu)$, define for every Borel subset $A \subseteq \mathbb{R}$ the auxiliary measures: $\nu_0(A) = \sum\limits_{n\in\mathbb{Z}} \mu(A+n)$,

$$\nu_1(A) = \sum_{n\in\mathbb{Z}} \int\limits_{A+n} f d\mu\,, \qquad \nu_2(A) = \sum_{n\in\mathbb{Z}} \int\limits_{A+n} |f|^2 d\mu\,.$$

Is easy to verify that $\nu_1 \ll \nu_0$ and $\nu_2 \ll \nu_0$. We claim that, $P_{\mathbb{V}}(f) = \frac{d\nu_1}{d\nu_0}$ a.e. $[\mu]$, where $\frac{d\nu_1}{d\nu_0}$ is the Radon-Nykodym derivative of ν_1 with respect to ν_0. Indeed if $\text{diam}(A) \leq 1$, by the definition of ν_1 and the Cauchy-Schwartz inequality:

$$|\nu_1(A)|^2 \leq \int\limits_{\bigcup\limits_{n\in\mathbb{Z}} A+n} |f|^2 d\mu\, \mu\left(\bigcup_{n\in\mathbb{Z}} A+n\right) = \nu_2(A)\nu_0(A)\,.$$

[b] Indeed, every f coincides a.e. $[\mu]$ with a periodic function.
Obviously, here $L^2(\mu)$ coincides with the previously defined space $\mathbf{L}^2(\mathbb{R})$.

Recalling the Lebesgue differentation theorem for Radon measures, this implies that $\left|\frac{d\nu_1}{d\nu_0}\right|^2 \leq \frac{d\nu_2}{d\nu_0}$ a.e. $[\mu]$. Then as the $\frac{d\nu_i}{d\nu_0}$ are taken as periodic functions

$$\|P_{\mathbb{V}}(f)\|^2 = \int_{\mathbb{R}} \left|\frac{d\nu_1}{d\nu_0}\right|^2 d\mu \leq \int_{\mathbb{R}} \frac{d\nu_2}{d\nu_0} d\mu = \sum_{k\in\mathbb{Z}} \int_{[0,1)+k} \frac{d\nu_2}{d\nu_0} d\mu = \int_{[0,1)} \frac{d\nu_2}{d\nu_0} d\nu_0 = \nu_2([0,1)) = \|f\|^2 .$$

Then $P_{\mathbb{V}}$ is bounded. Let $f \in \mathbb{V}$, thus $f = g$ a.e. $[\mu]$ for some periodic g. Then

$$\nu_1(A) = \sum_{n\in\mathbb{Z}} \int_{A+n} g d\mu = \int_A g d\nu_0 \,,$$

and from this $\frac{d\nu_1}{d\nu_0} = g$ a.e. $[\nu_0]$ and since $\mathrm{Ran}(P_{\mathbb{V}}) \subseteq \mathbb{V}$, then this is an expression for the projection over $\mathbb{V}$. Now, let us write $P_{\mathbb{V}}(f)$ in another way. Define, for every $n \in \mathbb{Z}$ and $A \in \mathscr{B}(\mathbb{R})$, $\mu_n(A) = \mu(A+n)$, then μ_n admits the following Lebesgue decomposition, $n \neq 0$:

$$\mu_n(A) = \int_A \varphi_n d\mu + \mu_n(A \cap \Gamma_n), \quad \mu(\Gamma_n) = 0, \quad \varphi_n \in L^1(\mu) \,, \tag{6.3.8}$$

and $\varphi_0 = 1$ a.e. $[\mu]$, $\Gamma_0 = \varnothing$. Define $\Gamma = \bigcup_{n\in\mathbb{Z}} \Gamma_n$, then $\mu(\Gamma) = 0$ and for every $n \in \mathbb{Z}$ we can write:

$$\mu_n(A) = \int_A \varphi_n d\mu + \mu_n(A \cap \Gamma) \,,$$

thus ν_0 and ν_1 can be rewritten as

$$\nu_0(A) = \int_A \sum_{n\in\mathbb{Z}} \varphi_n d\mu + \nu_0(A \cap \Gamma) \,, \qquad \nu_1(A) = \int_A \sum_{n\in\mathbb{Z}} f(\,.\,+n)\varphi_n d\mu + \nu_1(A \cap \Gamma) \,,$$

then, on Γ^c (see [33] sec. 32),

$$P_{\mathbb{V}}(f) = \frac{d\nu_1}{d\nu_0} = \frac{f + \sum\limits_{n\neq 0} f(\,.\,+n)\varphi_n}{1 + \sum\limits_{n\neq 0} \varphi_n} \quad a.e.\ [\mu] \,. \tag{6.3.9}$$

If $X_t \in \overline{\mathrm{span}}\{X_k\}_{k\in\mathbb{Z}}$ for every $t \in \mathbb{R}$, in particular if $t \in \mathbb{R}\setminus\mathbb{Q}$ this implies that $P_{\mathbb{V}}(e^{i2\pi\lambda t}) = e^{i2\pi\lambda t}$ a.e. $[\mu]$. Then

$$e^{i2\pi\lambda t} = \frac{e^{i2\pi\lambda t} + \sum\limits_{n\neq 0} e^{i2\pi(\lambda+n)t}\varphi_n(\lambda)}{1 + \sum\limits_{n\neq 0} \varphi_n(\lambda)} \quad a.e.\ [\mu] \,,$$

equivalently, $\sum\limits_{n\neq 0} (1 - e^{i2\pi nt})\varphi_n(\lambda) = 0$ a.e. $[\mu]$. Thus $\mathbb{R}e((1 - e^{i2\pi nt})) > 0$, since $t \in \mathbb{R}\setminus\mathbb{Q}$. And since the φ_n are non negative then $\mu_n \perp \mu$, $n \neq 0$. Thus there exists complementary supports for μ and μ_n, $n \neq= 0$; let A_n be a support of μ such that $\mu_n(A_n) = 0$, $n \neq 0$. The intersection

$N = \bigcap_{n\neq 0} A_n$ is a support of μ such that $\mu_n(N) = 0$, $n \neq 0$. From the definition of the μ_n, we have that the translates $N_n = N - n$ is a support of μ_n which has the property $\mu_m(N_n) = 0$, $n \neq m$, $n, m \in \mathbb{Z}$; In particular, $\mu(N_n)$, $n \neq 0$. Finally, the set $A = N \bigcap \left(\bigcup_{n\neq 0} N_n \right)^c$ is a support of μ which is disjoint of each of its translates $A - n$, $n \neq 0$.

iii) $\Longrightarrow$ *i)*, Suppose A is a support of μ which is disjoint from each of its translates. $A - n$ is a support of the previously defined μ_n, so that μ_n and μ have disjoint supports, $n \neq 0$, thus μ and μ_n are mutually singular. From eq. 6.3.8 we have that $\varphi_n = 0$ a.e. μ, $n \neq 0$. This implies that each $f \in L^2(\mu)$ is equivalent to its projection $P_{\mathbb{V}}(f)$, using eq. 6.3.9. Hence $\mathbb{V} = L^2(\mu)$ or equivalently $\overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$. □

Remark 6.3.1. Condition i) of the theorem above, does not necessarily imply that every X_t can be written as an (infinite, mean square convergent) linear combination like the eq. 1.0.1 of theorem 1.0.1.

For further comments on this result read at the end of this chapter. The following is an extension of theorem 6.3.3 :

Lemma 6.3.4. *Let $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n}$ be w.s.s. process and let $\mathscr{Y} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$ be the discrete time 'sampled' process. Then:*
a) $H(\mathscr{X}) = H(\mathscr{Y}) \iff$ for each eigenvalue Λ_j of $\frac{d\mathbf{M}}{d\nu}$ there exists $A_j \in \mathscr{B}(\mathbb{R})$, such that $\Lambda_j = 0$ a.e. $[\nu]$ on A_j^c and $A_j \cap A_j + k = \varnothing$ for every $k \in \mathbb{Z} \setminus \{0\}$.
b) Let $\mathscr{X}$ be such that the spectral measures $\mu_{ij} << \mathcal{L}$, then $H(\mathscr{X}) = H(\mathscr{Y}) \iff$ for each eigenvalue Λ_j of D there exists $A_j \in \mathscr{B}(\mathbb{R})$, such that $\Lambda_j = 0$ a.e. $[\mathcal{L}]$ on A_j^c and $A_j \cap A_j + k = \varnothing$ for every $k \in \mathbb{Z} \setminus \{0\}$.

Remark. Recall that $\nu = tr(\mathbf{M})$

Proof. a) Since $\frac{d\mathbf{M}}{d\nu}$ is non negative definite and self adjoint, there exists a measurable [67] P orthogonal matrix and a diagonal matrix of eigenvalues Λ such that $\Lambda = P^* \frac{d\mathbf{M}}{d\nu} P$. Let us introduce the process $\mathscr{Z} = \{Z_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n}$ defined by the linear operation on $\mathscr{X}$ ([69], Chapter I, sec. 8):

$$Z_t^r = \sum_{j=1}^{n} \int_{\mathbb{R}} e^{i\lambda 2\pi t} P_{rj}^* d\Phi_j. \tag{6.3.10}$$

This operation is well defined since P_i, the i-th column of P, is in $\mathbf{L}^2(\mathbb{R})$:

$$\int_{\mathbb{R}} P_i^* \frac{d\mathbf{M}}{d\nu} P_j d\nu \leq \int_{\mathbb{R}} \|P_i\|^2 \left\| \left(\frac{d\mathbf{M}}{d\nu} \right)^{\frac{1}{2}} \right\|_{op}^2 d\nu \leq \int_{\mathbb{R}} tr \left(\frac{d\mathbf{M}}{d\nu} \right) d\nu = \nu(\mathbb{R}) < \infty \ ,$$

and if we denote μ'_{ij} the spectral measures of the process $\mathscr{Z}$, then by the result [69] which relates a linear transformation and the spectral measure:

$$\forall\ A \in \mathscr{B}(\mathbb{R}) :\ \mu'_{ij}(A) = \int_A P_i^* \frac{d\mathbf{M}}{d\nu} P_j d\nu = \int_A \Lambda_{ij} d\nu. \tag{6.3.11}$$

The linear operation induced by P over $H(\mathscr{X})$ in eq. 6.3.10 is invertible, since P^{-1} exists for almost all λ, then $H(\mathscr{X}) = H(\mathscr{Z})$. Equation 6.3.10 also implies that for each $k \in \mathbb{Z}$, the X_k^r are obtainable from the Z_k^r, for every $r = 1 \ldots n$ and reciprocally, since P^{-1} exists. If we introduce another process of samples from $\mathscr{Z}$: $\mathscr{S} = \{Z_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$, from the latter discussion we have that $H(\mathscr{X}) = H(\mathscr{Y}) \iff H(\mathscr{Z}) = H(\mathscr{S})$. But eq. 6.3.11 means that $\mu_{i\,j} \equiv 0$ for $i \neq j$, or equivalently $\mathbb{E}(Z_{t'}^i Z_t^j) = 0$ for $i \neq j$, so that we have the orthogonal sum $H(\mathscr{Z}) = \bigoplus_{j=1}^{n} \overline{\operatorname{span}}\{Z_t^j\}_{t\in\mathbb{R}}$. Hence it will suffice to study when $\overline{\operatorname{span}}\{Z_t^j\}_{t\in\mathbb{R}} = \overline{\operatorname{span}}\{Z_k^j\}_{k\in\mathbb{Z}}$, for each j. This holds if and only if (by theorem 6.3.3) there exists $A_j \in \mathscr{B}(\mathbb{R})$ such that $\mu'_{j\,j}(A^c) = 0$ and $A_j \cap A_j + k = \varnothing$ for integer $k \neq 0$. But from eq. 6.3.11 $\mu'_{j\,j}(A^c) = 0$ is equivalent to the condition on the eigenvalue $\Lambda_j = \Lambda_{j\,j} = 0$ a.e. $[\nu]$ on A_j^c.
b) Is an immediate consequence of a) □

Now, we can characterize fundamental frames of uniform samples, in other words, given $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n}$ a w.s.s. process and $\mathscr{Y} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$ the discrete time 'sampled' process, we want to give conditions when $\mathscr{Y}$ is a fundamental frame for $H(\mathscr{X})$ in terms of the spectral measure of the process. For this purpose, in an analogue way to definition 23 for each $\lambda \in [0,1)$ we introduce the following sequence space:

$$l^2(D) = \left\{ x : \mathbb{Z} \to \mathbb{C}^n \;,\; \sum_{k\in\mathbb{Z}} x_k D(\lambda+k) x_k^* < \infty \right\} .$$

In a similar manner to that of definition 23, $l^2(D)$ can be identified with a Hilbert space with norm $\|x\|_{l^2(D)}^2 = \sum_{k\in\mathbb{Z}} x_k D(\lambda+k) x_k^*$. We give a condition in terms of $l^2(D)$ and an alternative condition combining the preceding results.

Theorem 6.3.4. *Let $\mathscr{X} = \{X_t^r\}_{t\in\mathbb{R}}^{r=1\ldots n}$ be a w.s.s. process, with spectral measure $\mu_{i\,j}$, and let $\mathscr{Y} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\ldots n}$ be the discrete time 'sampled' process. Then:*
a) $\mathscr{Y}$ is a frame with constants $A, B > 0$ for $H(\mathscr{X}) \iff \mu_{i\,j} << \mathcal{L}$ and for almost all $\lambda \in [0,1)$:

$$A\,\|x\|_{l^2(D)}^2 \leq \left\| \sum_{k\in\mathbb{Z}} D(\lambda+k) x_k^* \right\|_{\mathbb{C}^n}^2 \leq B\,\|x\|_{l^2(D)}^2 \;\; \forall x \in\, l^2(D) \;.$$

b) $\mathscr{Y}$ is a frame with constants $A, B > 0$ for $H(\mathscr{X}) \iff$ the following conditions hold simultaneously: i) $\mu_{i\,j} << \mathcal{L}$ and there exists $A, B > 0$ such that D_Π the matrix of periodized spectral densities verifies $\sigma(D_\Pi) \subseteq \{0\} \cup [A,B]$ a.e $[\mathcal{L}]$. ii) For each eigenvalue Λ_j of D there exists $A_j \in \mathscr{B}(\mathbb{R})$, such that $\Lambda_j = 0$ a.e. $[\mathcal{L}]$ on A_j^c and $A_j \cap A_j + k = \varnothing$ for every $k \in \mathbb{Z} \setminus \{0\}$.

Proof. (*Part a.*) ($\Rightarrow$) First note that by theorem 6.3.2, $\mu_{i\,j} << \mathcal{L}$ and then there exists D the spectral density matrix. Recall eq. 6.3.1, if $Y \in H(\mathscr{X})$ then there exists $f \in \mathbf{L}^2(\mathbb{R})$ such that $Y = \sum_{j=1}^{n} \int_{\mathbb{R}} f_j d\Phi_j$. Then for such Y and f we have:

$$\sum_{k\in\mathbb{Z}} \sum_{j=1}^{n} |\mathbb{E}(X_k^j \overline{Y})|^2 = \int_{[0,1)} \left\| \sum_{k\in\mathbb{Z}} D(\lambda+k) f^*(\lambda+k) \right\|_{\mathbb{C}^n}^2 d\lambda \;,$$

on the other hand,

$$\mathbb{E}|Y|^2 = \int_{[0,1)} \sum_{k\in\mathbb{Z}} f(\lambda+k)D(\lambda+k)f^*(\lambda+k)d\lambda\ .$$

Calling the sequence $\mathbf{f}_k = f(.+k)$, then by a similar argument to that of the proof of theorem 6.3.1 we can rewrite condition eq. 6.3.1 as

$$A\int_{[0,1)} \|\mathbf{f}\|^2_{l^2(D)}\, d\lambda \le \int_{[0,1)} \left\|\sum_{k\in\mathbb{Z}} D(\lambda+k)f^*(\lambda+k)\right\|^2_{\mathbb{C}^n} d\lambda \le B\int_{[0,1)} \|\mathbf{f}\|^2_{l^2(D)}\, d\lambda\ . \tag{6.3.12}$$

Let us call $\mathcal{D}$ the set of all sequences in $(\mathbb{C}^n)^{\mathbb{Z}}$ such that x^j_k belongs to $\mathcal{D}$ if $x^j_k = 0$ but for finitely many $j\,k$'s. $\mathcal{D}$ is a countable set which is dense in $l^2(D)$ for every λ. Given $x \in \mathcal{D}$, $\epsilon > 0$ and $\lambda_0 \in [0,1)$ set $\mathbf{f}(\lambda) = \sum_{k\in\mathbb{Z}} x_k \mathbf{1}_{B(\lambda_0,\epsilon)}(\lambda-k)\mathbf{1}_{[k,k+1)}$. In this case eq. 6.3.12 becomes:

$$A\int_{B(\lambda_0,\epsilon)\cap[0,1)} \|x\|^2_{l^2(D)}\, d\lambda \le \int_{B(\lambda_0,\epsilon)\cap[0,1)} \left\|\sum_{k\in\mathbb{Z}} D(.+k)x^*_k\right\|^2_{\mathbb{C}^n} d\lambda \le B\int_{B(\lambda_0,\epsilon)\cap[0,1)} \|x\|^2_{l^2(D)}\, d\lambda\ .$$

Then for each $x \in \mathcal{D}$, there exists F_x, such that $\mathcal{L}(F^c_x) = 0$ and,
$A\,\|x\|^2_{l^2(D)} \le \left\|\sum_{k\in\mathbb{Z}} D(\lambda+k)x^*_k\right\|^2_{\mathbb{C}^n} \le B\,\|x\|^2_{l^2(D)}$, for all $\lambda \in F_x$. Then taking $F = \bigcap_{x\in\mathcal{D}} F_x$, we have $\mathcal{L}(F^c) = 0$ and that for all $\lambda \in F$ it holds

$$A\,\|x\|^2_{l^2(D)} \le \left\|\sum_{k\in\mathbb{Z}} D(\lambda+k)x^*_k\right\|^2_{\mathbb{C}^n} \le B\,\|x\|^2_{l^2(D)} \quad \forall\, x \in \mathcal{D}\,. \tag{6.3.13}$$

We want to prove that eq. 6.3.13 holds for every $x \in l^2(D)$. For such x let us define $y_k = D^{\frac{1}{2}}(\lambda+k)x^*_k$, then $\sum_{k=N+1}^{M} D(\lambda+k)x^*_k = \sum_{k=N+1}^{M} D^{\frac{1}{2}}(\lambda+k)y_k$. First we shall see that given $x \in l^2(D)$: $\lim_{N,M\to\infty} \left\|\sum_{k=N+1}^{M} D(\lambda+k)x^*_k\right\| = 0$, for $j = 1\ldots n$. Denoting $D^{\frac{1}{2}}_j(\lambda+k)$ the j th= row of $D^{\frac{1}{2}}(\lambda+k)$, we have: $\left\|\sum_{k=N+1}^{M} D^{\frac{1}{2}}(\lambda+k)y_k\right\|^2 = \sum_{j=1}^{n}\left|\sum_{k=N+1}^{M} D^{\frac{1}{2}}_j(\lambda+k)y_k\right|^2 \le$ $\sum_{j=1}^{n}\left(\sum_{k=N+1}^{M}\left\|D^{\frac{1}{2}}_j(\lambda+k)\right\|_{\mathbb{C}^n}\|y_k\|_{\mathbb{C}^n}\right)^2$. Applying again the Cauchy-Schwartz inequality, we get:

$$\left\|\sum_{k=N}^{M} D^{\frac{1}{2}}(\lambda+k)y_k\right\|^2 \le \sum_{j=1}^{n}\left(\sum_{k=N+1}^{M}\left\|D^{\frac{1}{2}}_j(\lambda+k)\right\|^2_{\mathbb{C}^n}\right)\left(\sum_{k=N+1}^{M}\|y_k\|^2_{\mathbb{C}^n}\right)\ .$$

If we denote $\|.\|_{\mathbb{C}^{n\times n}}$ the Froebenius/euclidean norm , denoting $x_N(k) = x_k\mathbf{1}_{(-\infty,N]}(k)$ and recalling the definition of y_k we can rewrite the last equation as:

$$\sum_{k=N+1}^{M}\left\|D^{\frac{1}{2}}(\lambda+k)\right\|^2_{\mathbb{C}^{n\times n}}\|x_N-x_M\|^2_{l^2(D)} = \sum_{k=N+1}^{M} tr(D^{\frac{1}{2}}(\lambda+k)D^{\frac{1}{2}}(\lambda+k))\,\|x_N-x_M\|^2_{l^2(D)}$$

$$= tr\left(\sum_{k=N+1}^{M} D(\lambda+k))\right)\|x_N - x_M\|^2_{l^2(D)} \ .$$

But by theorem 6.3.1: $tr\left(\sum_{k=N+1}^{M} D(\lambda+k))\right) \leq tr(D_\Pi(\lambda)) \leq nB$ a.e. in $[0,1)$. Equivalently this holds for all $\lambda \in F'$ with $\mathcal{L}((F')^c) = 0$, so we can take $G = F \cap F'$. Then, the result follows, since for $\lambda \in G$ and $x \in l^2(D)$: $\lim_{M\to\infty} \left\|\sum_{k\in\mathbb{Z}} D(\lambda+k)x^*_M(k)\right\| = \left\|\sum_{k\in\mathbb{Z}} D(\lambda+k)x^*_k\right\|$ and $\lim_{M\to\infty} \|x_M\|_{l^2(D)} = \|x\|_{l^2(D)}$.
($\Leftarrow$) It is easy to reverse the proof using the condition of the hypothesis and equation 6.3.12.
(*Part b.*) It follows from combining theorem 6.3.2 and lemma 6.3.4. □

In particular we have that if $\mathscr{Y}$ is a fundamental frame, then for almost all λ $l^2(D)$ must be isomorphic to a finite dimensional space.

6.4 Canonical Dual Frame and a.s. Convergence

6.4.1 Canonical Dual Frame.

In the case of shift invariant subspaces of $L^2(\mathbb{R})$ some useful formulations for the dual frame are obtained in terms of the Fourier transforms of the generators. In the following we consider a similar problem for the frame formed by the stationary sequence: $\{X^r_k\}^{r=1\ldots n}_{k\in\mathbb{Z}}$. In this case is possible to give conditions in terms of the spectral density. Recall that for each $t \in \mathbb{R}$ or $\mathbb{Z}$, X^r_t can be written as the result of the action of the (unitary) *time shift operator* T on X^r_0 [69]. In our case, the *frame operator* $S : H(\mathscr{X}) \longrightarrow H(\mathscr{X})$ is given by:

$$Y \longmapsto S(Y) = \sum_{r=1}^{n}\sum_{k\in\mathbb{Z}} \mathbb{E}(\overline{Y}X^r_k)X^r_k \ .$$

Recall that S has a bounded inverse and on the other hand each $Y \in H(\mathscr{X})$ admits the following representation:

$$Y = \sum_{r=1}^{n}\sum_{k\in\mathbb{Z}} \mathbb{E}(\overline{Y}S^{-1}X^r_k)X^r_k \ .$$

Taking into account eq. 6.2.1 and if we suppose that T commutes with S^{-1} we have:

$$Y = \sum_{r=1}^{n}\sum_{k\in\mathbb{Z}} \mathbb{E}(\overline{Y}S^{-1}T^kX^r_0)X^r_k = \sum_{r=1}^{n}\sum_{k\in\mathbb{Z}} \mathbb{E}(\overline{Y}T^kS^{-1}X^r_0)X^r_k \ ,$$

so, in this case the canonical dual frame is a new stationary sequence given by $W^r_k = T^k(S^{-1}X^r_0)$, $k \in \mathbb{Z}$, and it would suffice to show that $W^r_0 = S^{-1}X^r_0$. We need the following lemma to solve this problem:

Lemma 6.4.1. *Let $\mathscr{X} = \{X^r_k\}^{r=1\ldots n}_{k\in\mathbb{Z}}$ be a stationary sequence which is a frame of its span $H(\mathscr{X})$. Then, for all $Y \in H(\mathscr{X})$ the following holds: $ST^kY = T^kSY$ and $S^{-1}T^kY = T^kS^{-1}Y$.*

Proposition 6.4.1. *Let $\mathscr{X} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\dots n}$ be a stationary sequence which is a frame of its span $H(\mathscr{X})$. Then the canonical dual frame $\{W_k^r\}_{k\in\mathbb{Z}}^{r=1\dots n}$ is given by*

$$W_k^r = \sum_{j=1}^{n} \int_{[0,1)} e^{i\lambda 2\pi k}(D^\sharp e_r^*)_j d\Phi_j \ . \tag{6.4.1}$$

Proof. Let us define $Z_0^m = \sum_{j=1}^{n} \int_{[0,1)} (D^\sharp e_m^*)_j d\Phi_j$. Where $D^\sharp$ is measurable by lemma 6.2.1. If we show that $SZ_0^m = X_0^m$ we are done, since S is invertible. Define $M_N = \sum_{r=1}^{n} \sum_{|k|\le N} \mathbb{E}(\overline{Z_0^m} X_k^r) X_k^r$ and $h_{N,r}(\lambda) = \sum_{|k|\le N} \mathbb{E}(\overline{Z_0^m} X_k^r) e^{i\lambda 2\pi k}$, and recall that $SZ_0^m = \sum_{r=1}^{n} \sum_{k\in\mathbb{Z}} \mathbb{E}(\overline{Z_0^m} X_k^r) X_k^r$. By lemma 6.2.1 we have that $X_0^m = \sum_{j=1}^{n} \int_{[0,1)} (e_m)_j d\Phi_j = \sum_{j=1}^{n} \int_{[0,1)} (P_{Col(D)} e_m^*)_j d\Phi_j$ almost surely. On the other hand:

$$\mathbb{E}(\overline{Z_0^m} X_k^r) = \int_{[0,1)} e^{i\lambda 2\pi k} e_r D D^\sharp e_m^* d\lambda = \int_{[0,1)} e^{i\lambda 2\pi k} e_r P_{Col(D)} e_m^* d\lambda \ .$$

So $h_{N,r} \longrightarrow e_r P_{Col(D)} e_m^*$ in $L^2[0,1)$ and a.e. But: $\mathbb{E}|X_0^m - M_N|^2 =$

$$\int_{[0,1)} (h_N - P_{Col(D)} e_m^*)^* D (h_N - P_{Col(D)} e_m^*) d\lambda \le B \int_{[0,1)} \left\| h_N - P_{Col(D)} e_m^* \right\|^2 d\lambda$$

where again, $M(\lambda) = \sup_{\|x\|=1} x D(\lambda) x^*$ is measurable and $M \le B$ a.e. which proves the result. □

On the expansion coefficients

Observing equation 6.4.1 we expect that the coefficients of the expansion may be rewritten in terms of ordinary Fourier transforms. Let us discuss the case for fundamental frames and when $n = 1$. In this case, given a stationary process $Y_t \in H(\mathscr{X})$, we know that $Y_t = \int_{\mathbb{R}} f\, e^{i\lambda 2\pi t} d\Phi$ for some $f \in L^2(\mathbb{R}, \phi d\lambda)$, where ϕ is the spectral density. On the other hand, eq. 6.4.1 becomes $W_k = \int_{\mathbb{R}} (\phi)^{-1} \mathbf{1}_S e^{i\lambda 2\pi k} d\Phi$, where $S = \{\lambda : \phi(\lambda) \neq 0\}$, then $\mathbb{E}(Y_t \overline{W_k}) = \int_S f(\lambda) e^{-i2\pi\lambda(t-k)} d\lambda = \widehat{f}(t-k)$. Note that we can take f such that $f = 0$ outside S, and moreover $f \in L^2(\mathbb{R})$ since $B \int_{\mathbb{R}} |f|^2 d\lambda \le \int_{\mathbb{R}} |f|^2 \phi d\lambda < \infty$. Then $Y_t \in H(\mathscr{X})$, provided that $\mathscr{Y} = \{X_k\}_{k\in\mathbb{Z}}$ is a fundamental frame, admits the following representation:

$$Y_t = \sum_{k\in\mathbb{Z}} \widehat{f}(t-k) X_k \,.$$

Let us discuss the case when another dual frame is used. First we need an auxiliary lemma:

Lemma 6.4.2. *Let $f, g \in L^2(\mathbb{R})$ and let $\widehat{f}, \widehat{g}$ be their Fourier transforms, then*

$$\int_{\mathbb{R}} |f * g|^2 dt = \int_{\mathbb{R}} |\widehat{f}\,\widehat{g}|^2 d\lambda \ .$$

When one side of the above equation is finite, then $\widehat{f * g} = \widehat{f}\,\widehat{g}$ a.e.

Proof. In the appendix.

In the general case, let $\{W_k\}_{k\in\mathbb{Z}}$ be the dual frame. If $g \in L^2(\mathbb{R}, \phi d\lambda)$ is such that $W_k = \int_{\mathbb{R}} g e^{i2\pi\lambda k} d\Phi$, supposing that f and Y_t are as in the previous discussion, then $\mathbb{E}(\overline{Y_t} W_k) = \int_{\mathbb{R}} \overline{f(\lambda)} g(\lambda)\phi(\lambda) e^{-i2\pi\lambda(t-k)} d\lambda$. First, we have that $fg\phi \in L^1(\mathbb{R})$. And, again as in the previous case $f,\ g \in L^2(\mathbb{R})$. We also have that $\|g\phi\|^2_{L^2(\mathbb{R})} \leq A \int_{\mathbb{R}} |g|^2 \phi d\lambda < \infty$, since $g \in L^2(\mathbb{R}, \phi d\lambda)$. Now by lemma 6.4.2 $\widehat{h} := \widehat{g\phi} = \widehat{g} * \widehat{\phi}$ and $\widehat{h} = \widehat{g} * \widehat{\phi} \in L^2(\mathbb{R})$. Then, $\widehat{h} * \widehat{f}$ is well defined and then $\widehat{hf} = \widehat{h} * \widehat{f}$ (in $\mathscr{S}'(\mathbb{R})$) but $hf \in L^1(\mathbb{R})$ so $\widehat{h} * \widehat{f}$ is well defined a.e. and then $\mathbb{E}(Y_t W_k) = \widehat{g} * \widehat{\phi} * \widehat{f}$. Finally, from this:

$$Y_t = \sum_{k\in\mathbb{Z}} (\widehat{g} * \widehat{\phi} * \widehat{f})(t-k) X_k \,.$$

6.4.2 Almost Sure Convergence

The representations given above converge in norm. Let us discuss briefly the problem of almost sure convergence for these representations. Note that point-wise convergence strongly depends on the summation method. First let us examine what happens in this context of the frame algorithm described before in section 2. Here, given $Y \in H(\mathscr{X})$ we can write: $Y_0 = 0$ and define $Y_{n+1} = Y_n + \lambda S(Y - Y_n)$, with λ and δ defined as before, then $\mathbb{E}|Y - Y_n|^2 \leq \delta^{2n} \mathbb{E}|Y|^2$. But given $\epsilon > 0$ by Chevyshev's inequality,

$$\mathbf{P}(|Y - Y_n| > \epsilon) \leq \frac{\mathbb{E}|Y - Y_n|^2}{\epsilon^2} \leq \frac{\delta^{2n}\mathbb{E}|Y|^2}{\epsilon^2} \,,$$

then $\sum_n \mathbf{P}(|Y - Y_n| > \epsilon) < \infty$, and then by the first Borel-Cantelli Lemma $Y_n \underset{n\to\infty}{\longrightarrow} Y$ a.s. The Menchoff-Rademacher theorem [43] gives a sufficient condition for the a.s. convergence of orthogonal expansions. A similar result holds for sequences which form a frame. This result could be obtained in a similar manner to [36] adapting the classical proof [43] of the original theorem. We include in the appendix a sketch of a shorter argument, similar to that of [81] involving absolutely summing operators. To prove such a result one would like to bound the maximal operator $\sup\limits_{N\leq s} \left| \sum\limits_{r=1}^{n} \sum\limits_{|k|\leq N} (\log(k+1))^{-1} c_k^r X_k^r \right|$. Note that this also can be written in terms of norms of sequences, and combining props. 33 and 21. From [81] it is possible to summarize these results in the following:

Lemma 6.4.3. *Let $(c_k)_k \in l^2(\mathbb{Z})$ be such that $\sum_k c_k^2 \log^2(k+1) < \infty$, if we define $T : l^2 \longrightarrow l^\infty$ as $(T\xi)_j = \sum_{|k|\leq j} c_k \xi_k$ and given a random vector $\Theta = (X_{-s}, \dots, X_0, \dots, X_s)$ we have $\mathbb{E}\,\|T\Theta\|^2_{l^\infty} \leq C \sum_k c_k^2 \log^2(k+1) \sup\limits_{x\in l^2\ \|x\|=1} \mathbb{E}|\Theta x^*|^2$.*

This lemma contains all we need to prove:

Proposition 6.4.2. *$\mathscr{X} = \{X_k^r\}_{k\in\mathbb{Z}}^{r=1\dots n}$ be a stationary sequence which is a frame of its span $H(\mathscr{X})$. If $(c_k^r)_{k\,r} \in l^2$ then $\sum\limits_{r=1}^{n} \sum\limits_{|k|\leq N} (\log(k+1))^{-1} c_k^r X_k^r$ converges almost surely as $N \longrightarrow \infty$.*

Proof. See the appendix

6.5 Appendix-Proofs of Some Auxiliary Results

Here are the proofs of some short and auxiliary results.

Proof of lemma 6.2.1 (a) In [67] it is proved that every non negative self adjoint matrix is diagonalizable in a measurable form $D = P\Lambda P^*$, then $D^\sharp = D^{(-1)}$ and $DD^\sharp = P_{Col(D)}$ are measurable.

(b) From the previous $I - P_{Col(D)} = P_{Nul(D)}$ is measurable. Now if $\mathcal{L}(\{\lambda : \mathrm{rg}(D)(\lambda) < n\}) > 0$ and if $A_j = \{\lambda : (P_{Nul(D)})_j(\lambda) \neq 0\}$, then $\{\lambda : \mathrm{rg}(D)(\lambda) < n\} = \bigcup_{j=1}^{n} A_j$,. From this, there exists a column $(P_{Nul(D)})_j(\lambda) \neq 0$ for every λ in some measurable A_j, with $\mathcal{L}(A_j) > 0$.

Part c) For such g, put $Z = \sum_{j=1}^{n} \int_{[0,1)} (f_j + g_j) d\Phi_j$, then $\mathbb{E}|Y - Z|^2 = \int_{[0,1)} gDg^* d\lambda = 0$ since $g \in Nul(D)$ a.e. so $Y = Z$ a.s. □

Proof of lemma 6.4.1 Given $Y \in H(\mathscr{X})$, $k \in \mathbb{Z}$: $S(T^jY) = \sum_{r=1}^{n} \sum_{k\in\mathbb{Z}} \mathbb{E}(T^jY\overline{T^kX_0^r})T^kX_0^r$. Recall that T is unitary so $T^* = T^{-1}$ and then $S(T^jY) = \sum_{r=1}^{n} \sum_{k\in\mathbb{Z}} \mathbb{E}(Y\overline{T^{k-j}X_0^r})T^kX_0^r$. Making a change of variables $k - j =: m$ we have:

$$S(T^jY) = \sum_{r=1}^{n} \sum_{m\in\mathbb{Z}} \mathbb{E}(Y\overline{T^mX_0^r})T^{m+j}X_0^r = T^j(SY) \ .$$

Finally, $S^{-1}T^kSS^{-1}Y = (S^{-1}S)T^kSY = T^kS^{-1}Y$. □

Proof of lemma 6.4.2. Note that $(f * g)(t)$ is well defined for all t and is a bounded function as a consequence of the Cauchy Schwartz inequality . Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a sequence in the Schwartz space $\mathcal{S}(\mathbb{R})$ such that $\varphi_n \longrightarrow g$ in $L^2(\mathbb{R})$ as $n \longrightarrow \infty$, then by Cauchy-Schwartz: $|f * g(t) - f * \varphi_n(t)| \leq \|f\|_{L^2(\mathbb{R})} \|g - \varphi_n\|_{L^2(\mathbb{R})}$ so that $f * \varphi_n \underset{n\to\infty}{\longrightarrow} f * g$ uniformly. Taking $\psi \in \mathscr{S}(\mathbb{R})$ we have that $|f * (g - \varphi_n)(t)||\widehat{\psi}(t)| \leq \|f\|_{L^2(\mathbb{R})} M|\widehat{\psi}(t)| \in L^1(\mathbb{R})$ then by the dominated convergence theorem:

$$\lim_{n\to\infty} \int_{\mathbb{R}} (f * \varphi_n)(t)\widehat{\psi}(t)dt = \int_{\mathbb{R}} (f * g)(t)\widehat{\psi}(t)dt \ . \tag{6.5.1}$$

But we also have that:

$$\forall\, n: \quad \int_{\mathbb{R}} (f * \varphi_n)(t)\widehat{\psi}(t)dt = \int_{\mathbb{R}} (\widehat{f}\widehat{\varphi}_n)(t)\psi(t)dt \ , \tag{6.5.2}$$

and again by the Cauchy-Schwartz inequality:

$$\left| \int_{\mathbb{R}} \widehat{f}\psi(\widehat{\varphi}_n - \widehat{g})dt \right| \leq \left\| \widehat{f}\psi \right\|_{L^2(\mathbb{R})} \|\widehat{g} - \widehat{\varphi}_n\|_{L^2(\mathbb{R})} \underset{n\to\infty}{\longrightarrow} 0 \ . \tag{6.5.3}$$

Combining eqs. 6.5.1, 6.5.2 and 6.5.3 we have that:

$$\langle f * g, \widehat{\psi} \rangle = \lim_{n\to\infty} \int_{\mathbb{R}} (f * \varphi_n)(t)\widehat{\psi}(t)dt = \lim_{n\to\infty} \int_{\mathbb{R}} (\widehat{f}\widehat{\varphi}_n)(t)\psi(t)dt = \langle \widehat{f}\widehat{g}, \psi \rangle \ .$$

On the other hand,

$$\left\| \widehat{f}\widehat{g} \right\| = \sup_{\|\psi\|=1\ \psi\in\mathscr{S}(\mathbb{R})} \int_{\mathbb{R}} \widehat{f}\widehat{g}\psi dt = \sup_{\|\psi\|=1\ \psi\in\mathscr{S}(\mathbb{R})} \int_{\mathbb{R}} (f * g)\widehat{\psi} dt = \|f * g\| \ .$$

□

Proof of proposition 6.4.2 The result follows if we bound the expected value of the square of the maximal function
$\sup_{N\le s}\left|\sum_{r=1}^{n}\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right| \le \sum_{r=1}^{n}\sup_{N\le s}\left|\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|$ then

$$\left(\mathbb{E}\sup_{N\le s}\left|\sum_{r=1}^{n}\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|^2\right)^{\frac{1}{2}} \le \sum_{r=1}^{n}\left(\mathbb{E}\sup_{N\le s}\left|\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|^2\right)^{\frac{1}{2}} .$$

Now, from lemma 6.4.3, for each r:
$\mathbb{E}\sup_{N\le s}\left|\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|^2 \le C_r \|c^r\|_{l^2} \sup_{a\in l^2(\mathbb{Z})\ \|a\|=1} \mathbb{E}\left|\sum_{|k|\le s} a_k X_k^r\right|^2$. Since the sequence $\mathscr{X}$ is Besselian $\mathbb{E}\left|\sum_{|k|\le s} a_k X_k^r\right|^2 \le B\|a\|^2_{l^2(\mathbb{Z})} \le B$. From this:
$\mathbb{E}\sup_{N\le s}\left|\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|^2 \le C_r B$. Then using the Cauchy-Schwartz inequality:

$$\mathbb{E}\sup_{N\le s}\left|\sum_{r=1}^{n}\sum_{|k|\le N}(\log(k+1))^{-1}c_k^r X_k^r\right|^2 \le \left(\sum_{r=1}^{n} C_r B \|c^r\|_{l^2(\mathbb{Z})}\right)^2 \le C' \sum_{r=1}^{n}\sum_{k\in\mathbb{Z}} |c_k^r|^2$$

□

6.6 Some additional comments

6.6.1 About theorem 6.3.3 and ergodic theory.

[47] Note that the result contained in this theorem about $P_{\mathbb{V}}$, i.e. eq. 6.3.9 and related is an ergodic theorem. Recall that a subset W of some measure space is wandering with respect to a $1-1$ point transformation T if its transforms $\{T^n(W)\}_n$ are mutually disjoint, and a set Y is dissipative with respect to T if Y can be written as $Y = \bigcup_n T^n(W)$ for some wandering set W. In the present case, $[0,1)$, for instance, is wandering with respect to $T(\lambda) = \lambda - 1$ and its translates $\{[0,1) - n\}_n$ cover $\mathbb{R}$. In the usual case of an ergodic theorem an assumption is made to the effect that all (measurable) dissipative sets have measure zero. Finally, such μ of condition iii) theorem 6.3.3, in the language of ergodic theory, has a wandering support.

Bibliography

[1] Abry P., Flandrin P., Taqqu M., Veitch D.,"Wavelets for the anlysis, estimation and synthesis of scaling data". In : Park, Willinger (eds.) *Self-similar network traffic and performance evaluation.* Wiley, New York, 2000, pp. 39-88.

[2] Alexits G., *Convergence Problems of Orthogonal Series*, Pergamon Press, 1961.

[3] Anh V.V., Angulo J.M., Ruiz-Medina M.D. , "Possible Long Range Dependance in fractional Random Fields", *Journal of Statistical Planning and Inf.* 80(1999) p. 95-110.

[4] Araujo A. Giné, E., *The Central Limit Theorem for real and Banach valued random variables.* J. Wiley, 1980.

[5] Beaty M.G. and Dodson. M.M. " The WKS Theorem, Spectral Translates and Plancherel's Formula", *J. Fourier Anal. Appl., 10(2) pp. 179-199*, 2004.

[6] Benedetto J.J., Shidong Li." The theory of multiresolution analysis frames and applications to filter banks" *Appl. Comput. Harmon. Anal.*, 5, pp. 389-427, 1998.

[7] Beran J., *Statistics for Long-Memory Processes*, Chapman and Hall/CRC (1994)

[8] Bharucha B.H. Kadota, T.T. "On the representation of continuous parameter processes by a sequence of random variables", *IEEE Trans. Inf. Theory*, IT-16 (2), 1970, pp.139-141.

[9] Billingsley P., *Probability and Measure* 3rd. edition, John Wiley and sons (Wiley series in probability and Mathematical Statics) (1994).

[10] Billingsley P., *Convergence of Probability Measures* 1st. edition, John Wiley and sons (Wiley series in probability and Mathematical Statics) (1968).

[11] Bojdecki T.,Gorostiza L., "Fractional Brownian motion via fractional Laplacian", *Statistics and Probability Letters*, 44(1), August 1999, pp. 107-108.

[12] Bownik M.: The structure of shift invariant subspaces of $L^2(\mathbb{R})$. J. Funct. Anal., 177(2), pp. 282-309, 2000.

[13] Braverman M. *Independent Random Variables and Rearrangement Invariant Spaces.* Cambridge Univ. Press. 1994.

[14] Cambanis S. Masry E." Wavelet Approximation of determinstic and random signals: Convergence properties and rates". *IEEE Trans. Inf. Theory*, Vol. 40, pp.1013-1029, 1994.

[15] Cambanis S., Rosinski J. Woyczynski W. "Convergence of quadratic forms in p-stable random variables and θ_p-Radonifying operators". *Ann. Prob.* Vol. 13, 3, pp.885-897.

[16] Chung, K. L. *A Course in Probability Theory*, Academic Press, 2001.

[17] Christensen O.*Frames and Bases.* Applied and Numerical Harmonic Analysis Series, Birkhäuser, 2008.

[18] Cuszic J. Tze Leung Lai. "On Random Fourier Series", *Trans. of the A.M.S.* 261(1),pp. 53-80 1980.

[19] De Boor C., DeVore R., Ron A." The structure of finitely generated shift-invariant spaces in $L^2(\mathbb{R}^d)$", *J. Funct. Anal.*, 119(1), pp. 37-78, 1994.

[20] Dunford N., Schwartz J.T.,*Linear Operators. Part I* , Interscience, 1958.

[21] Dym H., McKean H.P. *Gaussian Processes, Function Theory and The Inverse Spectral Problem.* Academic Press, 1976.

[22] Flandrin P., "Wavelet analysis and synthesis of fractional Brownian motion", *IEEE Trans. Inf. Theory.* IT 38(2), pp. 910-917, 1992.

[23] Garsia A. M., *Martingale Inequalities*, W.A. Benjamin, Reading Mass., 1973.

[24] Garsia A. M., *Topics in almost everywhere convergence*, Markham, 1970.

[25] Guerré, S., *Classical Sequences in Banach Spaces*, Marcel Dekker, 1992.

[26] Gel'fand I.M. Vilenkin N. Ya. *Generalized Functions* Vol. IV. Fizmatgiz, Moscow, 1961.(Russian). English trnsl. Academic Press, New York, 1964.

[27] Gikhman I.I. Skorokhod A.V., *The Theory of Stochastic Processes.* Vol. I. Springer, 1980.

[28] Giné E. Zinn J., "Central limit theorems and weak laws of large numbers in certain Banach spaces", *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 62, pp. 323-354.

[29] Giné E. Mandrekar V. Zinn J., "On sums of independent random variables with values in L_p $(2 \leq p < \infty)$", In *Probability in Banach Spaces II*, Ed. A. Beck, LNM 709, Springer, pp. 111-124.

[30] Grafakos L. *Classical and Modern Fourier Analysis.* Pearson/Prentice-Hall, 2004.

[31] Gröchenig K. *Time Frequency Analysis.* Birkhäuser, 2001.

[32] Hall P., Heyde C.C., *Martingale limit theory and its application.* Academic Press, 1980.

[33] Halmos P., *Measure Theory*, Van Nostrand, 1954.

[34] Hille E. Phillips R.S., *Functional Analysis and Semi-groups.* Amer.Math. Soc. publ. 1996.

[35] Hoffmann-Jorgensen J., "Sums of independent Banach space valued random variables", *Studia Math.* T. LII, pp. 159-186, 1974.

[36] Houdré C., “ Wavelets, probability and statistics: some bridges”, In: Benedetto J. and Frazier M. (eds.) *Wavelets: Mathematics and applications*, CRC press, Boca ratón, FL., pp. 361-399, 1993.

[37] Houdré C. “Wavelets, probability and statistics, some bridges”. Chapter 9 in Wavelets: Mathematics and Applications, Vol. 52. EDs. J. Benedetto and M. Frazier, CRC press, 1994.

[38] Ibragimov I.A., Rozanov Y.A. *Gaussian Random Processes.* Springer Verlag, 1978.

[39] Itô K. Nisio M., “On the convergence of sums of independent Banach space valued random variables”, *Osaka J. Math.* 5, 1968, pp. 35-48.

[40] Kahane J.P., *Some Random Series of Functions*, Cambridge, 1993.

[41] Kawata T., *Fourier Analysis in Probability Theory*, Academic Press, 1972.

[42] Kwapień S. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple.* Birkhäuser, 1992.

[43] Loéve M., *Probability Theory* Vol. I, Springer Verlag, 1977.

[44] Kolmogorov A.N., “The local structure of turbulence in incompressible viscous flow for very large Reynolds numbers´´, C.R. Acad. Sci. U.S.S.R. 3, 301, 1941.

[45] Lee A. J., “Sampling theorems for nonstationary processes”, *Trans. of the A.M.S.* V. 242, 1978, pp.225-241.

[46] Lindenstrauss J., Tzafriri L., *Classical Banach Spaces* Vol.II. Springer, 1979,

[47] Lloyd S.P.“ A sampling theorem for stationary (wide sense) stochastic processes”. *Trans. Amer. Math. Soc.* 92(1) pp.1-12, 1959.

[48] Lu Y.M., Do M.N. “ A Theory for sampling signals from a union of subspaces”. *IEEE Trans. Signal Proc.* 56(6), pp. 2334-2345, 2008.

[49] Makagon A. Weron A.“ q-variate minimal stationary processes”, *Studia Math.* 59, pp.41-52, 1976.

[50] Mandelbrot B.B., “Some noise with $1/f$ spectrum, a bridge between direct current and white noise”, *IEEE Transactions on Inf. Theory*, Vol. IT13, NO.2, Apr.1967, pp.289-298.

[51] Marcus, M. B. Pisier G. *Random Fourier Series with applications to Harmonic Analysis.* Princeton Univ. Press, 1981.

[52] Masani P.“ Shift invariant spaces and prediction theory”. *Acta Math.* 107, pp.275-290, 1962.

[53] Masry E. “The wavelet transform of stochastic processes with stationary increments and its applications to fractional Brownian motion”. *IEEE Trans. on Inf. Theory.* IT 34(1), pp. 260-264, 1993.

[54] Masry E. Cambanis S. "The representation of stochastic processes without loss of information", *SIAM J. Appl. Math.* Vol. 25(4), 1973, pp. 628-633.

[55] Medina J.M. Cernuschi-Frías, B. "Random Series in $L^p(X, \Sigma, \mu)$ using unconditional basic sequences and l^p stable sequences: A result on almost sure almost everywhere convergence". *Proc. of the Amer. Math. Soc.*, 135(11), pp. 3561-3569. 2007.

[56] Medina J.M. Cernuschi-Frías B. "On the a.s. convergence of certain random series to a fractional random field in $\mathcal{D}'(\mathbb{R}^d)$". *Statistics and Probability Letters*, 74(2005), pp. 39-49.

[57] Medina J.M Cernuschi-Frías B. "Wide Sense Stationary Processes forming Frames´´, accepted for publication. To appear in the *IEEE Trans. Inf. Theory.*

[58] Medina J.M Cernuschi-Frías B. "On the prediction of a class of wide sense stationary Random Process´´, *IEEE Trans. Signal Proc.* TSP 59(1) , pp. 70-77. 2011.

[59] Menchoff D. *Sur les séries de fonctions orthogonales I.* Fund. Math. 4, 1923, pages 82-105.

[60] Meyer Y., Sellan F., Taqqu M.S., "Wavelets, genaralized white noise and fractional integration: The synthesis of Fractional Brownian Motion" *The Journal Of Fourier Analysis and Applications*, Vol 5, Issue 5,1999.

[61] Paley R.E.A.C., Zygmund A., "A note on analytic functions on the circle", *Proc. Camb. Phil. Soc.* 28, 1932, pp. 266-272.

[62] Paley R.E.A.C., Zygmund A., "On some series of functions (1)(2)(3)", *Proc. Camb. Phil. Soc.* 26, 1930, pp. 337-357. 26, 1930, pp. 458-470. 28, 1932, 190-205.

[63] Paley R.E.A.C., Wiener N., Zygmund A., "Notes on Random Functions", *Math. Z.* 37, 1932, pp. 647-668.

[64] Parzen E. "Statistical Inference on Time Series By Hilbert Space Methods" (I and II), in *Time Series Analysis Papers*, ed. Parzen E., Holden-Day, 1968.

[65] Petersen K., *Ergodic Theory*, Cambridge Univ. Press, 1989.

[66] Reed I.S., Lee P.C., and Truong T.K., "Spectral Representation of Fractional Brownian Motion in n Dimensions and its Properties", *IEEE Transactions on Inf. Theory*, Vol. 41, NO.5, Sept 1995.

[67] Ron A., Shen Z.“ Frames and stable bases for shift -invariant subspaces of $L^2(\mathbb{R}^d)$". *Canad. J. Math.* 47, pp. 1051-1094, 1995.

[68] Rosenblatt M. *Stationary Sequences and Random Fields.* Birkhäuser, 1985.

[69] Rozanov Y., *Stationary random processes.* Holden-Day, 1967.

[70] Rozanov Y.A.“ On stationary sequences forming a basis", *Soviet Math.-Doklady* 1, 91-93 (1960).

[71] Samorodnitsky G. and Taqqu M.S., *Stable Non-Gaussian Random Processes* Chapman and Hall/CRC, 1994.

[72] Rudin W., *Análisis Funcional*, Revertè, Barcelona, 1979.

[73] Schwartz L., *Geometry and Probability in Banach Spaces.* LNM 852, Springer, 1980.

[74] Stein E.M., *Singular Integrals and Differentiability Properties of functions*, Princeton Univ. Press (1970)

[75] Stein E.M. and Weiss G., *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton Univ. Press (1970)

[76] Taylor R.L. *Stochastic Convergence of Weighted Sums of Random Elements in Linear Spaces.* Lecture Notes in Mathematics No.672, Springer-Verlag, 1978.

[77] Van Trees H.L. *Detection, estimation and modulation theory.* Part I. Wiley, New York, 1968.

[78] Weron A.“ On characterizations of interpolable and minimal stationary processes.” *Studia Math.* 49, pp. 165-183, 1974.

[79] Wiener N, Masani P.“ The prediction theory of multivariate stochastic processes,II. The linear predictor”. *Acta Math.* 99, pp.94-137, 1959.

[80] Wojtaszczyk P., “Wavelets as Unconditional Basis in $L^p(\mathbb{R})$”, *J. Fourier Anal. Appl.* 5(1) 1999, pages 73-85.

[81] Wojtaszczyc P. *Banach Spaces for Analysts.* Cambridge, 1991.

[82] Wornell, G. *Signal Processing with Fractals- A Wavelet based approach* Prentice Hall, 1996.

[83] Yang L., “Unconditional Basic Sequence in $L^p(\mu)$ and its l^p stability”, *Proc. A.M.S.* Vol. 127(2), 1999, pages. 455-464.

[84] Zhao P., Lin G., Zhao C.“ A matrix-valued wavelet KL-like expansion for wide sense stationary processes”. *IEEE Trans. Signal Proc.* Vol. 52(4), pp. 914-920, 2004.

[85] Zygmund, A. *Trigonometric Series*, vol. 2, Cambridge Univ. Press,1968.