Universidad Nacional Autónoma de México

Facultad de Ciencias

Un acercamiento al Análisis Estocástico en espacios de Hilbert

TESIS

que para obtener el título de:

Actuario

presenta:

Hugo Guadalupe Reyna Castañeda

Directora del trabajo:

Dra. María de los Ángeles Sandoval Romero

Ciudad Universitaria, Cd. Mx.

2023

Dedicatoria

A mi papá, Hugo G. Reyna Zuñiga, que partió físicamente de este mundo mientras escribía este trabajo y quien fue pieza importante en mi vida para que me fuera posible terminarlo. Te dedico este escrito papá, en él está plasmado lo que me apasiona en esta vida y que estoy seguro te hubiera hecho muy feliz verlo terminado.

Agradecimientos

Agradezco a la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México por el apoyo brindado a través de la beca PAPIME, con clave de proyecto PE101521.

A mi asesora querida, la Dra. María de los Ángeles Sandoval Romero, por ser una inspiración y pieza clave para dedicarme a las matemáticas. Así mismo, le agradezco por toda la paciencia, consejos, observaciones y por haberme conseguido la ayuda necesaria para que me fuera posible terminar este trabajo.

A mi sinodales por sus observaciones y comentarios pues a través de ellos logré desarrollar un mejor escrito.

A mi mamá, Ma. Sofía Castañeda, por todos sus esfuerzos para que me fuera posible terminar la universidad. Gracias por estar a mi lado siempre, por cuidarme y por brindarme lo que necesitaba.

A mi tía y prima, Ma. Rocío Castañeda y Rocío Hernández, por apoyarme cada que lo necesitaba, por abrirme las puertas de su casa y recibirme con tanto cariño.

A mis hermanos, Andrea Reyna, Jacqueline Reyna y Miguel Reyna, por estar en mi vida y, de alguna forma, ser una inspiración.

A todas aquellas personas que con sus tantas preguntas e insistencias por estos temas me motivaban a hacerlo mejor.

Índice general

Introducción

La vida es una escuela de probabilidad.
Walter Bagehot

Muchos fenómenos de la física, la biología, la medicina, las finanzas, la ingeniería, etc. se modelan a través de un parámetro de tiempo pues interesa apreciar su evolución a través de este. Sin embargo, muchos de estos dan como resultado un conjunto de distintas alternativas cuando son realizados bajo las mismas condiciones iniciales y es por ello que reciben el nombre de fenómenos aleatorios. La rama de las matemáticas que se encarga del estudio detallado de los fenómenos aleatorios es la teoría de la probabilidad.

Uno de los primeros puntos de partida de la probabilidad fue el intentar resolver un problema particular concerniente a una apuesta de juegos de dados entre dos personas aproximadamente en 1654. Con el paso del tiempo y la formalidad en las matemáticas, se sentaron las bases y experiencias necesarias para la formulación de una teoría matemática que englobara los conceptos y metodología de solución de los problemas derivados de los juegos de azar resueltos a lo largo de varios años. En 1933, el matemático ruso A. N. Kolmogorov, propuso un sistema de axiomas para la teoría de la probabilidad basado en la teoría de conjuntos y en la teoría de la medida que había sido desarrollada años antes principalmente por los matemáticos franceses Henri Lebesgue y Émile Borel. Esta teoría es conocida como teoría de la probabilidad clásica y prevalece hasta hoy en día por las múltiples soluciones que brinda a problemas de distintas disciplinas técnicas y científicas. En este trabajo, supondremos que el lector tiene bien familiarizado los conceptos de un curso intermedio de probabilidad. Ver, por ejemplo, [26] y [66].

Uno de los conceptos que deriva en más aplicaciones en la teoría de la probabilidad es aquel en donde se considera a un sistema que puede caracterizarse por estar en cualquier valor de un conjunto dado. Es decir, dado un sistema, suponemos que éste evoluciona o cambia de un valor a otro a lo largo del tiempo de acuerdo a cierta ley de movimiento, por ejemplo, que un objeto esté en cierta posición en este momento y que en el siguiente instante de tiempo sabemos que puede retroceder o avanzar un valor de su posición. De manera general, no es posible conocer con absoluta certeza hacia donde evolucionará el sistema por lo que es preciso considerar aleatoriedad en él. De este modo, si suponemos que existe alguna indexación del tiempo entonces en cada valor del índice de tiempo podemos considerar que el estado del sistema ahí, queda representado por una variable aleatoria. A la familia de todas las variables aleatorias indexadas por el tiempo dado se le conoce como un proceso estocástico. Los procesos estocásticos, de acuerdo a [58], representan la mayoría de los ejemplos de evolución aleatoria dados por la naturaleza. Uno de los ejemplos más antiguos y trascendentales es el conocido movimiento Browniano registrado alrededor de 1827 por el botánico escocés Robert Brown. Este tipo de conceptos condujeron al desarrollo de una de las ramas de las matemáticas con mayor relevancia pues incluye la construcción del concepto de integral estocástica para el planteamiento de ecuaciones diferenciales estocásticas, el Cálculo Estocástico. Algunas referencias que resultan excelentes para aprender más sobre estos temas y algunas aplicaciones son, por ejemplo, [20] y [82].

El tema central de este trabajo está motivado por el siguiente problema:

Sea $(\Omega, \mathcal{F}, P)$ el espacio de probabilidad dado por $\Omega = (0,1)$, $\mathcal{F}$ la σ-álgebra de Borel del intervalo abierto $(0,1)$ y P la medida de Lebesgue en $\mathcal{F}$.

Para cada $\omega \in (0,1)$, consideremos la función característica:

$$1_{(0,\omega)}(x) = \begin{cases} 1 & \text{si } x \in (0,\omega), \\ 0 & \text{si } x \notin (0,\omega). \end{cases}$$

Problema I.1: *¿Para qué valores $\omega \in (0,1)$ es posible calcular la probabilidad de que el área bajo la curva descrita por la gráfica de $1_{(0,\omega)}$ sea menor que ε^2 para $\varepsilon \in (0,1)$?*

Precisemos esta pregunta. Denotemos por $\mathcal{I}((0,1),\mathcal{F})$ al conjunto de todas las funciones características de subconjuntos de $(0,1)$ que están en $\mathcal{F}$. Podemos entonces considerar a la relación descrita antes como una función definida en dicho conjunto, es decir,

$$\chi : (0,1) \to \mathcal{I}((0,1),\mathcal{F}), \qquad \chi(\omega) = 1_{(0,\omega)}$$

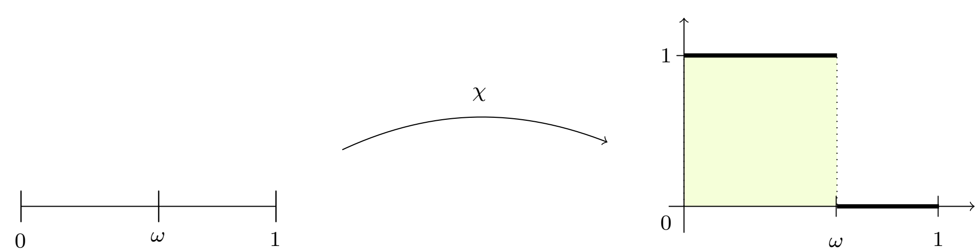


Figura 1: Representación gráfica

El concepto de área bajo la curva de una función sugiere que utilicemos, de alguna forma, el concepto de integral pues es inmediato que la función característica de cualquier intervalo abierto de $(0,1)$ es acotada e integrable. Teniendo esto en cuenta el Problema I.1 puede expresarse como sigue:

Problema I.2: *Sea $\varepsilon \in (0,1)$. ¿Cómo puede calcularse la probabilidad del conjunto $\{\omega \in (0,1) \,:\, \int_0^1 1_{(0,\omega)}(x)\,dx < \varepsilon^2\}$?*

En los cursos de Probabilidad de la Facultad de Ciencias de la Universidad Nacional Autónoma de México aparecen problemas similares que puedes resolverse a través de la construcción de variables aleatorias: Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Una función $X : \Omega \to \mathbb{R}^N$ es una variable aleatoria real si $X^{-1}(B) \in \mathcal{F}$ para todo subconjunto $B \in \mathcal{B}(\mathbb{R}^N)$ con $\mathcal{B}(\mathbb{R}^N)$ la σ-álgebra de Borel de $\mathbb{R}^N$. De este modo, a través de una variable aleatoria X se puede trasladar la medida de probabilidad P al espacio medible $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ considerando la fórmula $P_X(B) := P(X^{-1}(B))$.

Ahora bien, como $\mathcal{I}((0,1), \mathcal{F})$ es un conjunto de funciones y no de puntos en $\mathbb{R}^N$, no tiene sentido aplicar la solución anterior al problema que nos interesa. Sin embargo, es posible inspirarse en él. Esto es, ¿Es posible pensar a las funciones como si fueran puntos? ¿A qué espacios es posible definir el concepto de variable aleatoria? ¿Cuál es la pieza clave en $\mathbb{R}^N$ para definir a los subconjuntos Borel medibles? ¿En qué sentido la función $\chi : (0,1) \to \mathcal{I}((0,1), \mathcal{F})$ sería una variable aleatoria? ¿Cómo puede extenderse el concepto de esperanza para funciones que toman valores en $\mathcal{I}((0,1), \mathcal{F})$? ¿Es posible extender algunos conceptos fundamentales de la teoría clásica de la probabilidad a espacios de funciones?

Es bien sabido que el Análisis Matemático da respuesta a este tipo de preguntas pues en él se generalizan conceptos tales como el de continuidad y completitud los cuales se definen a través un concepto muy sencillo que es el de distancia y que generaliza el concepto de norma en un espacio vectorial. En Análisis Matemático se trabaja con espacios de funciones como lo es el conjunto de funciones características de subconjuntos de $(0,1)$. De acuerdo a Clapp [17] estos espacios aparecen de manera natural en muchos problemas de las matemáticas y de sus aplicaciones. Por otro lado, la Teoría de la Medida, se encarga de estudiar el concepto de integral (de Lebesgue) a la clase de funciones medibles mismas que resultan ser una generalización del concepto de variable aleatoria. De nuevo, en este trabajo supondremos que el lector está familiarizado con los conceptos de un curso básico de Análisis Matemático y de Teoría de la Medida, se recimenda consultar, por ejemplo, [2; 5; 7; 10; 13; 17; 18; 73].

De acuerdo a Mamporia [54], algunas de las preguntas derivadas del Problema I.2 empezaron a resolverse a mediados de 1960 con el estudio de las ecuaciones diferenciales estocásticas en espacios de dimensión infinita. En 1983, en la Conferencia Regional CBMSNSF sobre Ecuaciones Diferenciales Estocásticas en espacios de dimensión infinita y sus aplicaciones, celebrada en la Universidad Estatal de Luisiana, Itô [40] presentó sus investigaciones sobre estos temas en algunos espacios de Hilbert pues en ellos fue posible generalizar los métodos tradicionales dados en dimensión finita.

En este texto estudiaremos los problemas arriba mencionados y otros problemas interesantes para lo cual presentaremos una introducción a la teoría de la probabilidad y a los procesos estocásticos en espacios de Hilbert de dimensión infinita siguiendo la misma idea de Itô [40]. Es así que el objetivo general de este trabajo es desarrollar aplicaciones en Teoría de la Probabilidad a través de resultados clave del Análisis Funcional que permitan a los estudiantes de actuaría, matemáticas, matemáticas aplicadas y física de la Facultad de Ciencias de la Universidad Nacional Autónoma de México desarrollar conceptos más complejos así como resolver distintos problemas derivados de estos como lo son la integral estocástica en espacios de Hilbert de dimensión infinita y las ecuaciones de evolución estocástica [3; 21; 22; 62; 87].

El desarrollo de este trabajo se divide en siete capítulos. En los primeros tres capítulos se detallarán algunos conceptos dados en un primer curso de Análisis Matemático como lo son los espacios vectoriales cuya norma está inducida por un producto escalar. Algunos ejemplos de estos espacios son los conocidos espacios de Lebesgue y de Lebesgue-Bochner que involucran el concepto de integral a cierta clase de funciones. Se construirá el concepto de mensurabilidad para funciones que toman valores en un espacio de Hilbert y se relacionará con las definiciones dadas en un curso básico de Teoría de la Medida. En lo último de esta primera parte se trabajará la teoría espectral de operadores compactos en espacios de Hilbert que permitirá extender algunos conceptos clave de la teoría clásica de probabilidad como el de distribución Gaussiana. A partir del cuarto capítulo, comienza el estudio de funciones que toman valores en espacios de Hilbert y de las cuales puede tratarse el concepto de variable aleatoria. El concepto central de esta segunda parte es el de medida Gaussiana pues permite introducir la noción de variable aleatoria Gaussiana o variable aleatoria normal como lo es en el caso de dimensión finita además estudiar el proceso de Wiener en espacios de Hilbert.

Considero pertinente mencionar que esta tesis es una recopilación bibliográfica de los conceptos y propiedades necesarios, presentados de manera ordenada, para introducir al lector al análisis estocástico en dimensión infinita. En cada capítulo cito la bibliografía considerada para su construcción. Si bien, algunas de las demostraciones aquí presentadas aparecen en los distintos libros de texto y artículos utilizados para la construcción de este trabajo, estas fueron replanteadas y desarrolladas, desde mi perspectiva, para una mayor claridad en los temas. Muchas otras son propias pues los resultados forman parte de las vastas listas de ejercicios de los textos consultados. Finalmente, es importante mencionar que la idea que me inspiró para desarrollar este trabajo surgió del artículo de integración estocástica en espacios de Hilbert escrito por Alvarado-Solano y Fonseca-Mora [3], las notas sobre análisis estocástico de Giussepe Da Prato [20] y las notas de ecuaciones de evolución estocástica de Jan van Neerven [87].

Capítulo 1

Espacios de Hilbert

Los *espacios de Hilbert* son la extensión más natural del espacio euclidiano a dimensión infinita pues en ellos se generalizan resultados bien conocidos de la geometría analítica tales como el teorema de pitágoras, ley del paralelogramo e incluso el concepto de proyección ortogonal. De manera formal, un espacio de Hilbert es un espacio vectorial, en este trabajo lo consideraremos sobre $\mathbb{R}$, con un producto escalar definido y tal que la norma inducida por él es completa. Por tanto, todo espacio de Hilbert es un espacio de Banach.

En este capítulo damos las propiedades que cumplen los espacios de Hilbert así como algunos ejemplos interesantes en la teoría. Mostraremos que el concepto de complemento ortogonal de un subespacio se extiende, en casos más generales, siempre que este sea cerrado. La existencia de la proyección ortogonal tiene una consecuencia muy importante, pues permite identificar al dual de un espacio de Hilbert H con él mismo, tal y como ocurre en $\mathbb{R}^N$. Esto es, cualquier función lineal y continua $\psi : H \to \mathbb{R}$ se puede expresar como el producto escalar por un único elemento de H. A este resultado se le conoce como el *teorema de representación de Fréchet-Riesz* y tiene aplicaciones muy importantes, por ejemplo, que todo espacio de Hilbert es reflexivo.

El concepto de ortogonalidad en espacios de Hilbert nos permite extender el concepto de base en cierto sentido al del álgebra lineal. En este caso, diremos que un subconjunto ortonormal de un espacio de Hilbert H es una *base de Hilbert* si el subespacio generado por él es denso H. Demostraremos que todo espacio de Hilbert separable siempre admite una base de Hilbert. La importancia de este resultado será fundamental para la teoría central de este trabajo.

Para el desarrollo de este capítulo nos basamos principalmente en [13], [17], [48] y [75].

1.1. Definiciones y propiedades básicas

Definición 1.1 *Sea V un espacio vectorial sobre $\mathbb{R}$. Un* ***producto escalar*** *en V es una función $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ con las siguientes propiedades:*

(PE1) $\langle \lambda\, v_1 + \mu\, v_2, w\rangle = \lambda\langle v_1, w\rangle + \mu\langle v_2, w\rangle$ *para cualesquiera $v_1, v_2, w \in V$, $\lambda, \mu \in \mathbb{R}$.*

(PE2) $\langle v, w\rangle = \langle w, v\rangle$ *para cualesquiera $v, w \in V$.*

(PE3) $\langle v, v\rangle \geq 0$ *para cualquier $v \in V$.*

(PE4) $\langle v, v\rangle = 0$ *si y sólo si $v = 0_V$.*

Veamos algunos ejemplos.

Ejemplo 1.2 *La función*

$$\langle v, w\rangle := v_1 w_1 + \ldots + v_N w_N$$

donde $v = (v_1, \ldots, v_N)$, $w = (w_1, \ldots, w_N) \in \mathbb{R}^N$ es un producto escalar en $\mathbb{R}^N$.

Demostración: De las propiedades de campo para $\mathbb{R}$ se tiene que, para cualesquiera $v_1, v_2, w \in V$, $\lambda, \mu \in \mathbb{R}$:

$$\begin{aligned} \langle \lambda\, v_1 + \mu\, v_2, w\rangle &:= (\lambda\, v_{11} + \mu\, v_{21})w_1 + \ldots + (\lambda\, v_{1N} + \mu\, v_{2N})w_N \\ &= (\lambda\, v_{11})w_1 + \ldots + (\lambda\, v_{1N})w_n + (\mu\, v_{21})w_1 + \ldots + (\mu\, v_{2N})w_N \\ &= \lambda(v_{11}w_1 + \ldots + v_{1N}w_N) + \mu(v_{21}w_1 + \ldots + v_{2N}w_N) \\ &= \lambda\langle v_1, w\rangle + \mu\langle v_2, w\rangle\,; \end{aligned}$$

$$\langle v_1, w\rangle := v_{11}w_1 + \ldots + v_{1N}w_N = w_1 v_{11} + \ldots + w_N v_{1N} := \langle w, v_1\rangle\,;$$

$$\langle v_1, v_1\rangle := v_{11}^2 + \ldots + v_{1N}^2 \geq 0, \,.$$

Por tanto, se cumplen (PE1), (PE2) y (PE3). La propiedad (PE4) es inmediata.

En consecuencia, $\langle\cdot,\cdot\rangle$ es un producto escalar en $\mathbb{R}^N$. $\blacksquare$

Ejemplo 1.3 *Sea $(\Omega, \mathcal{S}, \mu)$ un espacio de medida y el espacio $L^2(\Omega)$. Entonces la función:*

$$\langle f, g\rangle_{L^2(\Omega)} := \int_\Omega f g\, d\mu, \qquad f, g \in L^2(\Omega)$$

es un producto escalar.

Demostración: La desigualdad de Hölder-Riesz asegura que $fg \in L^1(\Omega)$ para cualesquiera $f, g \in L^2(\Omega)$ por lo que la función $\langle , \rangle_{L^2(\Omega)}$ está bien definida.

Dado que $|f|^2 \geq 0$ entonces $\int_\Omega |f|^2 \, d\mu = \langle f, f \rangle_{L^2(\Omega)} \geq 0$ y, se tiene que $\int_\Omega |f|^2 \, d\mu = 0$ si y sólo si $f = 0$ en $L^2(\Omega)$. Por tanto, se cumplen (PE3) y (PE4). Del mismo modo, como $fg = gf$ entonces $\langle f, g \rangle_{L^2(\Omega)} = \langle g, f \rangle_{L^2(\Omega)}$ para cualesquiera $f, g \in L^2(\Omega)$. Finalmente, La propiedad (PE1) es consecuencia inmediata de la linealidad de la integral de Lebesgue. ■

De los ejemplos anterior obtenemos que $\langle v, v \rangle = \|v\|_V^2$ para cuando el espacio vectorial es normado. Esto nos permite suponer que un producto escalar en algún espacio vectorial V define una norma en él a través de la fórmula:

$$\| v \| := \sqrt{\langle v, v \rangle}. \tag{1.1}$$

Para demostrar esto requerimos la siguiente proposición.

Proposición 1.4 *Si V es un espacio vectorial sobre $\mathbb{R}$ con un producto escalar definido, entonces se cumplen las siguientes relaciones:*

(a) Desigualdad de Cauchy-Schwarz:

$$|\langle v, w \rangle| \leq \| v \| \, \| w \| \qquad \forall v, w \in V. \tag{1.2}$$

(b) Desigualdad del triángulo:

$$\|v + w\| \leq \| v \| + \| w \| \quad \forall v, w \in V. \tag{1.3}$$

(c) Identidad del paralelogramo:

$$\|v + w\|^2 + \|v - w\|^2 = 2 \left(\| v \|^2 + \| w \|^2 \right) \quad \forall v, w \in V. \tag{1.4}$$

Demostración: (a) Para cualesquiera $v, w \in V$, $\lambda \in \mathbb{R}$, se cumple:

$$0 \leq \langle \lambda \, v + w, \lambda \, v + w \rangle = |\lambda|^2 \|v\|^2 + 2 \, \lambda \, \langle v, w \rangle + \|w\|^2.$$

Si $v = 0$ el resultado se sigue trivialmente, así es que supondremos que $v \neq 0$. Definiendo $\lambda := -\frac{\langle v, w \rangle}{\|v\|^2}$ obtenemos de lo anterior que:

$$0 \leq \frac{|\langle v, w \rangle|^2}{\|v\|^2} - 2 \, \frac{|\langle v, w \rangle^2|}{\|v\|^2} + \|w\|^2 = \|w\|^2 - \frac{|\langle v, w \rangle|^2}{\|v\|^2}.$$

Multiplicando la desigualdad anterior por $\|v\|^2$ obtenemos que:

$$0 \leq \|v\|^2 \, \|w\|^2 - |\langle v, w \rangle|^2,$$

y sacando raíz cuadrada se sigue la desigualdad deseada.

(b) De la desigualdad de Cauchy-Schwarz se sigue que:

$$\begin{aligned}\|v+w\|^2 &= \|v\|^2 + 2\langle v,w\rangle + \|w\|^2\\ &\leq \|v\|^2 + 2\,|\langle v,w\rangle| + \|w\|^2\\ &\leq \|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2 = (\|v\|^2 + \|w\|^2),\end{aligned}$$

sacando raíz cuadrada en la desigualdad anterior obtenemos el resultado.

(c) Por cálculos directos obtenemos:

$$\begin{aligned}\|v+w\|^2 &= \|v\|^2 + 2\langle v,w\rangle + \|w\|^2,\\ \|v-w\|^2 &= \|v\|^2 - 2\langle v,w\rangle + \|w\|^2,\end{aligned}$$

sumando amabas expresiones concluimos que:

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2 + 2\langle v,w\rangle - 2\langle v,w\rangle = 2\left(\| v\|^2 + \| w\|^2\right),$$

como afirma el enunciado. ■

Proposición 1.5 *Sea V un espacio vectorial con un producto escalar $\langle\cdot,\cdot\rangle$. Entonces $\|v\| := \sqrt{\langle v,v\rangle}$ define norma en V.*

Demostración: La desigualdad del triángulo se probó en la Proposición 1.4. Luego, por (PE4) se tiene que $\|v\| := \sqrt{\langle v,v\rangle} = 0$ si y sólo si $\langle v,v\rangle = 0$ si y sólo si $v = 0$. Finalmente, para cualesquiera $v \in V$ y $\lambda \in \mathbb{R}$, de las propiedades (PE1) y (PE2) se sigue que:

$$\|\lambda v\| := \sqrt{\langle \lambda v,\lambda v\rangle} = \sqrt{\lambda^2\langle v,v\rangle} = |\lambda|\sqrt{\langle v,v\rangle} := |\lambda|\|v\|.$$

En consecuencia, $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$ es una norma. ■

El siguiente resultado establece que un producto escalar es una aplicación continua.

Proposición 1.6 *Si (w_k) es una sucesión en V tal que $w_k \to w$ en V, entonces*

$$\lim_{k\to\infty}\langle v,w_k\rangle = \langle v,w\rangle \qquad \forall v \in V.$$

Demostración: De la desigualdad de Cauchy-Schwarz (ver Proposición 1.4) se sigue que:

$$|\langle v,w_k\rangle - \langle v,w\rangle| \leq |\langle v,w_k - w\rangle| \leq \|v\|\|w_k - w\|$$

para cualquier $v \in V$. Dado que $w_k \to w$ entonces $\langle v,w_k\rangle \to \langle v,w\rangle$ como afirma el enunciado. ■

Nos preguntamos ahora si un espacio normado puede inducir un producto escalar a través de su norma. En realidad esto sólo es posible cuando la norma cumple con la identidad del paralelogramo (ver Proposición 1.4) como lo muestra el siguiente resultado debido a P. Jordan y John von Neumann (ver [42]).

Teorema 1.7 (Jordan-Von Neumann) *Si* $(V, \|\cdot\|_V)$ *es un espacio normado tal que:*

$$\|v+w\|_V^2 + \|v-w\|_V^2 = 2\left(\|\, v\,\|_V^2 + \|\, w\,\|_V^2\right) \quad \forall v, w \in V$$

entonces la función dada por:

$$\langle v, w\rangle := \frac{1}{4}\left(\|v+w\|_V^2 - \|v-w\|_V^2\right) \tag{1.5}$$

es un producto escalar en V *tal que* $\|v\|_V^2 = \langle v, v\rangle$ *para todo* $v \in V$.

Demostración: Sean $v, w \in V$. Dado que $\|v+w\|_V = \|w+v\|_V$ y $\|v-w\|_V = \|-(w-v)\|_V = \|w-v\|_V$ entonces:

$$\langle v, w\rangle = \frac{1}{4}\left(\|v+w\|_V^2 - \|v-w\|_V^2\right) = \frac{1}{4}\left(\|w+v\|_V^2 - \|w-v\|_V^2\right) = \langle w, v\rangle.$$

Notemos que $\langle v, v\rangle = \frac{1}{4}\left(\|2v\|_V^2\right) = \|v\|_V^2 \geq 0$ para todo $v \in V$. Luego, $\langle v, v\rangle = 0$ si y sólo si $\|v\|_V^2 = 0$ si y sólo si $v = 0_V$. Esto prueba (PE2), (PE3) y (PE4).

Sean $u, v, w \in V$. Usando la identidad del paralelogramo obtenemos que:

$$\|(u+w)+v\|_V^2 + \|(u+w)-v\|_V^2 = 2\left(\|u+w\|_V^2 + \|v\|_V^2\right) \tag{1.6}$$

$$\|(u-w)+v\|_V^2 + \|(u-w)-v\|_V^2 = 2\left(\|u-w\|_V^2 + \|v\|_V^2\right) \tag{1.7}$$

Restando la expresión (1.7) a (1.6) obtenemos:

$$\|(u+w)+v\|_V^2 - \|(u-w)+v\|_V^2 + \|(u+w)-v\|_V^2 - \|(u-w)-v\|_V^2 = 2\left(\|u+w\|_V^2 - \|u-w\|_V^2\right)$$

y, en consecuencia:

$$\|(u+v)+w\|_V^2 - \|(u+v)-w\|_V^2 + \|(u-v)+w\|_V^2 - \|(u-v)-w\|_V^2 = 2\left(\|u+w\|_V^2 - \|u-w\|_V^2\right).$$

Por tanto,

$$4\langle u+v, w\rangle + 4\langle u-v, w\rangle = 8\langle u, w\rangle$$

de donde se sigue que:

$$\langle u+v, w\rangle + \langle u-v, w\rangle = 2\langle u, w\rangle \qquad \forall\, u, v, w \in V. \tag{1.8}$$

Observemos que $\langle 0_V, w\rangle = \frac{1}{4}\left(\|w\|_V^2 - \|-w\|_V^2\right) = 0$ pues $\|w\|_V = \|-w\|_V$ para todo $w \in V$. De este modo, en (1.8) se sigue que:

$$2\langle u, w\rangle = \langle 2u, w\rangle + \langle 0_V, w\rangle = \langle 2u, w\rangle \qquad \forall\, u, w \in V. \tag{1.9}$$

De (1.8) y (1.9) concluimos que:

$$\langle u+v,w\rangle + \langle u-v,w\rangle = \langle 2u,w\rangle \qquad \forall\, u,v,w \in V. \tag{1.10}$$

Sean $u,v,w \in V$. Definiendo $v_1 := \frac{1}{2}(u+v)$ y $v_2 := \frac{1}{2}(u-v)$ se sigue de (1.10) que:

$$\langle u+v,w\rangle = \langle 2v_1,w\rangle = \langle v_1+v_2,w\rangle + \langle v_1-v_2,w\rangle = \langle u,w\rangle + \langle v,w\rangle. \tag{1.11}$$

Sean $v,w \in V$ y $\lambda \in \mathbb{R}$. Es inmediato que si $\lambda = 0$, $\lambda = 1$ ó $\lambda = -1$ entonces $\lambda\langle v,w\rangle = \langle \lambda v,w\rangle$. Así pues, consideremos los siguientes casos:

Caso 1. $\lambda \in \mathbb{N}$.

Argumentando por inducción, supongamos que para $j-1 \in \mathbb{N}$ se tiene que:

$$(j-1)\langle v,w\rangle = \langle (j-1)v,w\rangle.$$

En consecuencia, para $j \in \mathbb{N}$, de (1.11) e hipótesis de inducción concluimos que:

$$\langle jv,w\rangle = \langle jv-v+v,w\rangle = \langle (j-1)v,w\rangle + \langle v,w\rangle = (j-1)\langle v,w\rangle + \langle v,w\rangle = j\langle v,w\rangle.$$

Caso 2. $\lambda \in \mathbb{Z}$.

Del caso anterior basta suponer que $\lambda = -j$ con $j \in \mathbb{N}$. Entonces,

$$-j\langle v,w\rangle = \langle 0_V,w\rangle - j\langle v,w\rangle = \langle 0_V,w\rangle - \langle jv,w\rangle = \langle 0_V - jv,w\rangle = \langle -jv,w\rangle.$$

Caso 3. $\lambda \in \mathbb{Q}$ con $\lambda = \frac{p}{q}$ con $p,q \in \mathbb{Z}$ y $q \neq 0$.

Del caso anterior se sigue que:

$$p\langle v,w\rangle = \langle p\,v,w\rangle = \left\langle \frac{pq}{q}\,v,w\right\rangle = q\left\langle \frac{p}{q}\,v,w\right\rangle$$

y, en consecuencia:

$$\frac{p}{q}\langle v,w\rangle = \left\langle \frac{p}{q}\,v,w\right\rangle.$$

Caso 4. $\lambda \in \mathbb{R}$.

Definamos $\xi,\vartheta : \mathbb{R} \to \mathbb{R}$ como $\xi(\lambda) := \lambda\langle v,w\rangle$ y $\vartheta(\lambda) := \langle \lambda\, v,w\rangle$. Es inmediato que ξ y ϑ son funciones continuas para $v,w \in V$ fijos.

Sea $\lambda \in \mathbb{R}$ y elijamos una sucesión (p_k) de elementos de $\mathbb{Q}$ tal que $p_k \to \lambda$ en $\mathbb{R}$. Del caso anterior concluimos que $\xi(p_k) = \vartheta(p_k)$ para todo $k \in \mathbb{N}$ y, usando la continuidad de las funciones obtenemos que:

$$\xi(\lambda) = \lim_{k\to\infty} \xi(p_k) = \lim_{k\to\infty} \vartheta(p_k) = \vartheta(\lambda).$$

Es decir, $\lambda\langle v, w\rangle = \langle \lambda v, w\rangle$.

Esta última afirmación junto con la expresión (1.11) prueban que $\langle\cdot,\cdot\rangle$ satisface la propiedad (PE1) y, por tanto, que es un producto escalar en V tal que $\|v\|_V^2 = \langle v, v\rangle$ para todo $v \in V$. ■

Definición 1.8 *Un* ***espacio de Hilbert*** *es un espacio vectorial H sobre $\mathbb{R}$ con un producto escalar $\langle\cdot,\cdot\rangle$ que es completo con la norma inducida* (1.1)*, es decir, tal que toda sucesión de Cauchy converge en H con la norma* (1.1)*. Lo denotaremos por $(H, \langle\cdot,\cdot\rangle, \|\cdot\|)$ o simplemente por H cuando no haga falta especificar su producto escalar.*

Veamos algunos ejemplos.

Ejemplo 1.9 $\mathbb{R}^N$ *con el producto escalar*

$$\langle v, w\rangle := v_1w_1 + \ldots + v_Nw_N, \quad v = (v_1, \ldots, v_N),\, w = (w_1, \ldots, w_N)$$

es un espacio de Hilbert.

Demostración: Se sigue del hecho que, si $v = (v_1, \ldots, v_N) \in \mathbb{R}^N$ entonces:

$$\|v\| := \sqrt{\langle v, v\rangle} = \sqrt{(v_1)^2 + \ldots + (v_N)^2}.$$

Es decir, la norma inducida por el producto escalar es la norma Euclidiana en $\mathbb{R}^N$ y, sabemos que este espacio es completo con dicho norma. ■

Ejemplo 1.10 *Sea $(\Omega, \mathcal{S}, \mu)$ un espacio de medida. El espacio $L^2(\Omega)$ con el producto escalar:*

$$\langle f, g\rangle_{L^2(\Omega)} := \int_\Omega fg\, d\mu, \quad f, g \in L^2(\Omega),$$

es un espacio de Hilbert.

Demostración: La norma asociada a este producto escalar es:

$$\|f\| = \sqrt{\langle f, f\rangle_{L^2(\Omega)}} = \left(\int_\Omega |f|^2\, d\mu\right)^{1/2} = \|f\|_{L^2(\Omega)}.$$

Es decir, la norma inducida por $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ es la norma en $L^2(\Omega)$. Del teorema de Riesz-Fischer sabemos que $(L^2(\Omega), \|\cdot\|_{L^2(\Omega)})$ es completo. ■

Un caso particular del ejemplo anterior lo podemos apreciar en el siguiente resultado.

Ejemplo 1.11 *El espacio $\ell_2(\mathbb{R})$ de todas las sucesiones $\overline{x} = (x_k)$ de números reales tales que la serie $\sum_{k=1}^{\infty} |x_k|^2$ converge, con el producto escalar:*

$$\langle \overline{x}, \overline{y} \rangle := \sum_{k=1}^{\infty} x_k\, y_k, \qquad \overline{x} = (x_k),\, \overline{y} = (y_k) \in \ell_2,$$

es un espacio de Hilbert.

Demostración: En efecto, $\ell_2(\mathbb{R})$ es el espacio $L^2(\Omega)$ con $\Omega = \mathbb{N}$, $\mathcal{S} = \mathcal{P}(\mathbb{N})$ y $\mu = \mu^\sharp$ medida de conteo. ■

Ejemplo 1.12 (Suma directa de espacios de Hilbert) *Sean $(H_1, \langle \cdot, \cdot \rangle_1)$, $(H_2, \langle \cdot, \cdot \rangle_2)$ dos espacios de Hilbert. Entonces, la función:*

$$\langle v_1 + v_2, w_1 + w_2 \rangle := \langle v_1, w_1 \rangle_1 + \langle v_2, w_2 \rangle_2, \;\; v_1, w_1 \in H_1,\, v_2, w_2 \in H_2$$

es un producto escalar en la suma directa $H_1 \oplus H_2$ y, $H_1 \oplus H_2$ es un espacio de Hilbert.

Demostración: Sean $u, v, w \in H_1 \oplus H_2$, $\lambda, \mu \in \mathbb{R}$. Existen únicos $u_1, v_1, w_1 \in H_1$, $u_2, v_2, w_2 \in H_2$ tales que:

$$u = u_1 + u_2,\; v = v_1 + v_2 \text{ y } w = w_1 + w_2.$$

Dado que $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$ son productos escalares, obtenemos:

$$\begin{aligned}
\langle \lambda\, u + \mu\, v, w \rangle &:= \langle \lambda\, u_1 + \mu\, v_1, w_1 \rangle_1 + \langle \lambda\, u_2 + \mu\, v_2, w_2 \rangle_2 \\
&= (\lambda \langle u_1, w_1 \rangle_1 + \mu \langle v_1, w_1 \rangle_1) + (\lambda \langle u_2, w_2 \rangle_2 + \mu \langle v_2, w_2 \rangle_2) \\
&:= \lambda \langle u, w \rangle + \mu \langle v, w \rangle.
\end{aligned}$$

$$\begin{aligned}
\langle v, w \rangle &:= \langle v_1, w_1 \rangle_1 + \langle v_2, w_2 \rangle_2 \\
&= \langle w_1, v_1 \rangle_1 + \langle w_2, v_2 \rangle_2 \\
&= \langle w, v \rangle
\end{aligned}$$

$$\langle v, v \rangle := \langle v_1, v_1 \rangle_1 + \langle v_2, v_2 \rangle_2 \geq 0.$$

Ahora bien, $\langle v, v \rangle := \langle v_1, v_1 \rangle_1 + \langle v_2, v_2 \rangle_2 = 0$ si y sólo si $v_1 = 0_1$ y $v_2 = 0_2$, así pues, $v = 0$.

En consecuencia, $\langle \cdot, \cdot \rangle$ es un producto escalar en $H_1 \oplus H_2$. La norma inducida está dada por:

$$\|v\| := \sqrt{\langle v, v \rangle} = \sqrt{\|v_1\|_1^2 + \|v_2\|_2^2}.$$

Sea (v_k) una sucesión de Cauchy en $H_1 \oplus H_2$ en donde $v_k = v_{1k} + v_{2k}$ con $v_{1k} \in H_1$, $v_{2k} \in H_2$ para todo $k \in \mathbb{N}$. Dada $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que

$$\|v_j - v_k\| = \sqrt{\|v_{1j} - v_{1k}\|_1^2 + \|v_{2j} - v_{2k}\|_2^2} < \varepsilon \qquad \forall j, k \geq k_0.$$

De lo anterior obtenemos:

$$\|v_{ij} - v_{ik}\|_i < \varepsilon \qquad \forall j, k \geq k_0, \ \ i = 1, 2,$$

lo que establece que (v_{ik}) es una sucesión de Cauchy en H_i y, por tanto, converge a v_i con $i = 1, 2$. Existen $k_1, k_2 \in \mathbb{N}$ tales que:

$$\|v_{ik} - v_i\|_i < \frac{\varepsilon}{\sqrt{2}} \qquad \forall k \geq k_i, \ \ i = 1, 2.$$

Definiendo $k_* := \max\{k_1, k_2\}$ concluimos que:

$$\|v_k - v\| = \sqrt{\|v_{1k} - v_1\|_1^2 + \|v_{2k} - v_2\|_2^2} < \varepsilon \qquad \forall k \geq k_*.$$

Esto prueba el resultado. ■

Ejemplo 1.13 *El espacio $\mathcal{C}^0([-1,1], \mathbb{R})$ de las funciones continuas de $[-1,1]$ a $\mathbb{R}$ con el producto escalar:*

$$\langle f, g \rangle := \int_{-1}^{1} f(x) g(x) \, dx$$

no es un espacio de Hilbert.

Demostración: Es inmediato comprobar que $\langle \cdot, \cdot \rangle$ es un producto escalar. Consideremos la sucesión de funciones $f_k : [-1, 1] \to \mathbb{R}$

$$f_k(x) := \begin{cases} 0 & \text{si } \ -1 \leq x \leq 0, \\ kt & \text{si } \ 0 \leq x \leq \frac{1}{k}, \\ 1 & \text{si } \ \frac{1}{k} \leq x \leq 1. \end{cases}$$

Veamos que (f_k) es de Cauchy en $\mathcal{C}^0([-1,1], \mathbb{R})$ con la norma inducida por el producto escalar pero que no converge en $\mathcal{C}^0([-1,1], \mathbb{R})$ con la misma.

Para $j \geq k$ se tiene que:

$$\begin{aligned} \int_{-1}^{1} |f_j(x) - f_k(x)|^2 \, dt = \int_{0}^{1} |f_j(x) - f_k(x)|^2 \, dx &= \int_{0}^{1/j} |(j-k)x|^2 \, dx + \int_{1/j}^{1/k} |1 - kx|^2 \, dx \\ &= \frac{k}{3j^2} - \frac{2}{3j} + \frac{1}{3k} \\ &\leq \frac{1}{k} - \frac{1}{j} \, . \end{aligned}$$

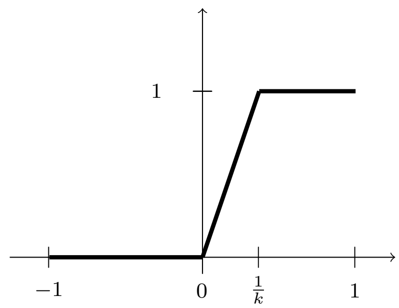


Figura 1.1: Sucesión de funciones (f_k)

En consecuencia, dada $\varepsilon > 0$ se cumple que:

$$\|f_j - f_k\| := \sqrt{\langle f_j - f_k, f_j - f_k\rangle} < \varepsilon \qquad \forall j,k > \frac{2}{\varepsilon^2}.$$

Por tanto, (f_k) es de Cauchy.

Supongamos que $f_k \to f$ en $\mathcal{C}^0([-1,1],\mathbb{R})$. Sea $\delta \in (0,1)$. Si $k \geq \frac{1}{\delta}$ entonces $f_k(t) = 1$ para todo $t \in [\delta, 1]$. Entonces:

$$0 \leq \int_\delta^1 |1 - f(x)|^2\, dx \leq \int_{-1}^1 |f_k(x) - f(x)|^2\, dx := \|f_k - f\|^2 \to 0$$

lo que establece que:

$$\int_\delta^1 |1 - f(x)|^2\, dx = 0,$$

y por lo tanto $f(t) = 1$ para todo $t \in [\delta, 1]$. De manera similar obtenemos:

$$0 \leq \int_{-1}^{-\delta} |f(x)|^2\, dx \leq \int_{-1}^1 |f_k(x) - f(x)|^2\, dx := \|f_k - f\|^2 \to 0$$

por lo que necesariamente $f(x) = 0$ para todo $[-1, -\delta]$. Dado que $\delta \in (0,1)$ es arbitraria, concluimos que:

$$f(x) = \begin{cases} 0 & \text{si } x \in [-1, 0), \\ 1 & \text{si } x \in (0, 1]. \end{cases}$$

Por tanto, f no es continua en $[-1,1]$ y esto contradice nuestra suposición. En consecuencia, $\mathcal{C}^0([-1,1],\mathbb{R})$ con el producto escalar definido, no es un espacio de Hilbert. ∎

El Teorema de Henri Lebesgue asegura que $\mathcal{C}^0([-1,1],\mathbb{R}) \subset L^2([-1,1])$ por lo que $\mathcal{C}^0([-1,1],\mathbb{R})$ con el producto escalar definido en el ejemplo anterior es un subespacio vectorial del espacio de Hilbert $L^2([-1,1])$. Así pues, del Ejemplo 1.13 concluimos que no todo subespacio vectorial de un espacio de Hilbert resulta ser un espacio de Hilbert. Sin embargo, dado que todo subespacio cerrado de un espacio normado completo es también completo podemos establecer el siguiente resultado.

Proposición 1.14 *Sea H un espacio de Hilbert. Un subespacio vectorial V de H es un espacio de Hilbert si y sólo si V es cerrado.*

Demostración: $\Rightarrow)$: Sea $v \in \overline{V}$. Existe una sucesión (v_k) de elementos de V tal que $v_k \to v$ en H. Es claro que (v_k) es una sucesión de Cauchy en V de modo que existe $\widetilde{v} \in V$ tal que $v_k \to \widetilde{v}$ en V. De la unicidad del límite concluimos que $v = \widetilde{v}$ y, por tanto, que V es cerrado.

$\Leftarrow)$: Sea (v_k) una sucesión de Cauchy en V. Luego, (v_k) es de Cauchy en H de modo que, existe $v \in H$ tal que $v_k \to v$. Dado que V es cerrado, entonces $v \in V$. En consecuencia V es de Banach y, por tanto, un espacio de Hilbert. ■

1.2. Ortogonalidad

Sea $H = (H, \langle \cdot, \cdot \rangle, \| \cdot \|)$ un espacio de Hilbert sobre $\mathbb{R}$.

Definición 1.15 *Sean v, w dos elementos de H. Diremos que v y w son* ***ortogonales****, que denotamos por $v \perp w$, si y sólo si $\langle v, w \rangle = 0$.*

El siguiente resultado afirma que todo elemento del espacio H es ortogonal al vector cero.

Proposición 1.16 *En un espacio de Hilbert H se cumple que, $v = 0_H$ si y sólo si $\langle v, w \rangle = 0$ para todo $w \in H$.*

Demostración: Supongamos que $v = 0_H$, entonces $\langle v, w \rangle = \langle w - w, w \rangle$ y, de la propiedad (PE1), concluimos que $\langle v, w \rangle = \langle w, w \rangle - \langle w, w \rangle = 0$. Inversamente, si $\langle v, w \rangle = 0$ para todo $w \in H$, en particular para $w = v$, y de la propiedad (PE4) se sigue que $\langle v, v \rangle = 0$ si y sólo si $v = 0_H$. Esto concluye la demostración. ■

Uno de los resultados más trascendentales que asociamos a la ortogonalidad en espacios Euclidianos es el teorema de Pitágoras, sin embargo, podemos extenderlo a espacios de Hilbert más generales.

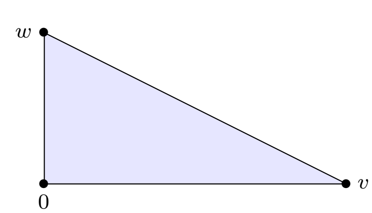


Figura 1.2: Teorema de Pitágoras

Teorema 1.17 (Pitágoras) *Sean $v, w \in H$. Si $v \perp w$, entonces:*

$$\|v\|^2 + \|w\|^2 = \|w - v\|^2. \tag{1.12}$$

Demostración: Aplicando directamente las propiedades del producto escalar obtenemos que:

$$\|w - v\|^2 := \langle w - v, w - v \rangle = \langle v, v \rangle + \langle w, w \rangle - \langle w, v \rangle - \langle v, w \rangle = \|v\|^2 + \|w\|^2,$$

como afirma el enunciado. ■

Definición 1.18 *Sea V un subconjunto no vacío de H. Definimos el* ***espacio ortogonal a V en H****, que denotaremos por $V^\perp$, como el subconjunto de H dado por:*

$$V^\perp := \{w \in H \,:\, \langle v, w \rangle = 0 \quad \forall v \in V\}.$$

Por ejemplo, la Proposición 1.16 afirma que $H^\perp = \{0_H\}$. Veamos que el conjunto ortogonal resulta ser un subespacio vectorial cerrado de H.

Proposición 1.19 *Sea V un subespacio vectorial de H. Entonces $V^\perp$ es un subespacio vectorial cerrado de H.*

Demostración: Sean $v, w \in V^\perp$ y $\lambda, \mu \in \mathbb{R}$, entonces $\langle u, v \rangle = 0 = \langle u, w \rangle$ para todo $u \in V$. De las propiedades (PE1) y (PE2) se sigue que:

$$\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle = 0 \quad \forall u \in V.$$

En consecuencia, $\lambda v + \mu w \in V^\perp$. Ahora bien, como $0_H \in V$ por la Proposición 1.16 concluimos que $0_H \in V^\perp$ y, por tanto, $V^\perp$ es un subespacio vectorial de H.

Sea (w_k) una sucesión de elementos de $V^\perp$ tal que $w_k \to w$ en H. Aplicando la continuidad del producto escalar (ver Proposición 1.6) obtenemos que:

$$\langle w, v \rangle = \lim_{k \to \infty} \langle w_k, v \rangle = 0 \quad \forall v \in V,$$

por tanto, $w \in V^\perp$.

Así pues, $V^\perp$ es cerrado en H. ■

Veamos algunas propiedades.

Proposición 1.20 *Sean V y W espacios vectoriales de H tales que $V \subset W$, entonces:*

(a) $W^\perp \subset V^\perp$.

(b) $V \cap V^{\perp} = \{0_H\}$.

(c) $(\overline{V})^{\perp} = V^{\perp}$.

Demostración: (a) Si $u \in W^{\perp}$ entonces $\langle u, w\rangle = 0$ para todo $w \in W$. En particular, $\langle u, v\rangle = 0$ para todo $v \in V$. Por tanto, $u \in V^{\perp}$.

(b) Dado que V y $V^{\perp}$ son subespacios vectoriales de H entonces $\{0_H\} \subset V \cap V^{\perp}$. Ahora, si $v \in V \cap V^{\perp}$ entonces $\langle v, u\rangle = 0$ para todo $u \in V$. En particular, $\langle v, v\rangle = 0$ lo cual ocurre si y sólo si $v = 0_H$.

(c) Como $V \subset \overline{V}$ entonces por el inciso *(a)* tenemos que $(\overline{V})^{\perp} \subset V^{\perp}$. Ahora bien, si $u \in V^{\perp}$ entonces $\langle u, v\rangle = 0$ para todo $v \in V$. Sea $w \in \overline{V}$. Existe una sucesión (v_k) de elementos de V tal que $v_k \to w$ en H y, por la continuidad del producto escalar (ver Proposición 1.6) obtenemos:

$$\lim_{k\to\infty} \langle v_k, u\rangle = \langle w, u\rangle.$$

Dado que $\langle v_k, u\rangle = 0$ para todo $k \in \mathbb{N}$ entonces $\langle w, u\rangle = 0$. En consecuencia, $u \in (\overline{V})^{\perp}$. ■

Notemos que si V es un subespacio vectorial de H, entonces para $v \in V$, se tiene que $\langle v, w\rangle = 0$ para todo $w \in V^{\perp}$. En consecuencia $v \in (V^{\perp})^{\perp}$ y, por tanto, $V \subset (V^{\perp})^{\perp}$. Resulta natural preguntarse si $V = (V^{\perp})^{\perp}$ y, la respuesta no es cierta en general como lo muestra el siguiente ejemplo.

Ejemplo 1.21 *Sea $(\Omega, \mathcal{S}, \mu)$ espacio de medida y $\mathcal{S}_I(\Omega, \mathcal{S})$ el subespacio de todas las funciones simples e integrables. Entonces, el espacio ortogonal a $\mathcal{S}_I(\Omega, \mathcal{S})$ en $L^2(\Omega)$ es $\{0\}$.*

Demostración: Dado que $\mathcal{S}_I(\Omega, \mathcal{S})$ es denso en $L^2(\Omega)$ entonces $\overline{\mathcal{S}_I(\Omega, \mathcal{S})} = L^2(\Omega)$ y, por tanto, $(\overline{\mathcal{S}_I(\Omega, \mathcal{S})})^{\perp} = (L^2(\Omega))^{\perp} = \{0\}$.

Por otro lado, la Proposición 1.20 asegura que $(\mathcal{S}_I(\Omega, \mathcal{S}))^{\perp} = (\overline{\mathcal{S}_I(\Omega, \mathcal{S})})^{\perp}$. Esto concluye la demostración. ■

Los siguientes resultados establecen una condición necesaria y suficiente para que un subespacio vectorial V de H cumpla que $V = (V^{\perp})^{\perp}$.

Una forma de describir al espacio ortogonal es la siguiente.

Proposición 1.22 *Sea V un subespacio vectorial de H y $w \in H$. Entonces, $w \in V^{\perp}$ si y sólo si $\|w\| = \inf_{v\in V} \|w - v\|$.*

Demostración: $\Rightarrow)$: Sea $w \in V^{\perp}$. Aplicando el Teorema de Pitágoras (ver Teorema 1.17) obtenemos que, $\|w - v\|^2 = \|w\|^2 + \|v\|^2$ para cada $v \in V$ y, por tanto $\|w\| \leq \|w - v\|$

para cada $v \in V$. Luego, como $0 \in V$, entonces $\|w\| = \|w - 0\| \geq \inf_{v \in V} \|w - v\|$ por lo que necesariamente:

$$\|w\| = \inf_{v \in V} \|w - v\|.$$

$\Leftarrow$) : Supongamos que $\|w\| = \inf_{v \in V} \|w - v\|$. Sea $v \in V$ arbitrario. Para cada $0 < t < 1$ se tiene que:

$$\|w - tv\|^2 = \|w\|^2 - 2t\langle w, v\rangle + t^2\|v\|^2.$$

Dado que $tv \in V$ para cada $0 < t < 1$, entonces $\|w - tv\| \geq \inf_{v \in V} \|w - v\| = \|w\|$ y, de la identidad anterior obtenemos:

$$\|w\|^2 \leq \|w - tv\|^2 = \|w\|^2 - 2t\langle w, v\rangle + t^2\|v\|^2.$$

Por tanto,

$$2t\langle w, v\rangle \leq t^2\|v\|^2, \tag{1.13}$$

y tomando el límite cuando $t \to 0$ concluimos que $\langle w, v\rangle \leq 0$. Si ahora consideramos a $-v \in V$ en la desigualdad (1.13) obtenemos que $0 \leq \langle w, v\rangle$. Así pues, $\langle w, v\rangle = 0$ para todo $v \in V$ y, en consecuencia $w \in V^{\perp}$. ■

El siguiente resultado establece ahora que, si V es un subconjunto no vacío cerrado y convexo entonces, para cada elemento u de H existe un único elemento v de V cuya distancia de u es la menor de entre todos los elementos de V. Recordemos que $V \subset H$ es convexo si para cada $v, w \in V$, el segmento:

$$[v, w] := \{t\, v + (1 - t)\, w \,:\, t \in [0, 1]\},$$

está contenido en V.

Teorema 1.23 *Sea V un subconjunto no vacío cerrado y convexo de H. Entonces, para cada $u \in H$, existe un único $v \in V$ tal que:*

$$\|u - v\| = \inf_{w \in V} \|u - w\|.$$

Demostración: Sea $u \in H$. Denotemos por

$$\rho := \operatorname{dist}(u, V) = \inf_{w \in V} \|u - w\|.$$

Aplicando la identidad del paralelogramo (ver Proposición 1.4) obtenemos que:

$$\|w_1 - w_2\|^2 = \|(w_1 - u) - (w_2 - u)\|^2 = 2\|w_1 - u\|^2 + 2\|w_2 - u\|^2 - \|(w_1 - u) + (w_2 - u)\|^2,$$

para cualesquiera $w_1, w_2 \in V$. Y, además se cumple que:

$$\|(w_1 - u) + (w_2 - u)\| = \|w_1 + w_2 - 2u\| = 2 \left\| \frac{w_1 + w_2}{2} - u \right\| \geq 2\rho,$$

pues $\frac{1}{2}(w_1 + w_2) \in V$ por ser convexo. Así pues, de las dos identidades anteriores concluimos que:

$$\|w_1 - w_2\|^2 \leq 2\|w_1 - u\|^2 + 2\|w_2 - u\|^2 - 4\rho^2, \tag{1.14}$$

para cualesquiera $w_1, w_2 \in V$. Por definición de ínfimo, consideremos una sucesión (v_k) en V tal que $\lim_{k\to\infty} \|u - v_k\| = \rho$. Entonces, dada $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que:

$$\|u - v_k\|^2 < \rho^2 + \frac{\varepsilon^2}{4} \qquad \forall k \geq k_0.$$

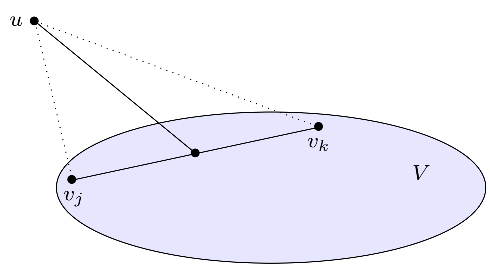


Figura 1.3: Conjunto convexo y cerrado

De la desigualdad (1.14) se sigue que:

$$\begin{aligned} \|v_j - v_k\|^2 &\leq 2\|v_j - u\|^2 + 2\|u - v_k\| - 4\rho^2 \\ &< 4\rho^2 + \varepsilon^2 - 4\rho^2 = \varepsilon^2 \qquad \forall j, k \geq k_0. \end{aligned}$$

En consecuencia, (v_k) es de Cauchy en V. Como V es un espacio de Hilbert (ver Proposición 1.14), existe $v \in V$ tal que $v_k \to v$ en H y, por tanto, concluimos que:

$$\|u - v\| = \lim_{k\to\infty} \|u - v_k\| = \rho.$$

Veamos ahora la unicidad. Supongamos que existe $\tilde{v} \in V$ tal que $\|u - \tilde{v}\| = \rho$. Aplicando la desigualdad (1.14) se sigue que:

$$\|v - \tilde{v}\|^2 \leq 2\|v - u\|^2 + 2\|\tilde{v} - u\|^2 - 4\rho^2 = 4\rho^2 - 4\rho^2 = 0,$$

por lo que necesariamente $v = \tilde{v}$. Esto concluye la demostración. ∎

Sea $u \in H$. Observamos lo siguiente: Supongamos que $v \in V$ para V cerrado y convexo no vacío en H satisface la identidad del teorema anterior, es decir, $\|u - v\| = \inf_{w\in V} \|u - w\|$ y sea $w \in V$ arbitrario. Entonces,

$$\|u - v\| \leq \|u - [(1 - t)v + t\, w]\| = \|(u - v) - t(w - v)\|$$

y, por tanto:

$$\|u-v\|^2 \leq \|u-v\|^2 - 2t\langle u-v, w-v\rangle + t^2\|w-v\|^2,$$

lo que implica que $2t\langle u-v, w-v\rangle \leq t^2\|w-v\|^2$ para todo $t \in (0,1]$. Tomando el límite cuando $t \to 0$ obtenemos que:

$$\langle u-v, w-v\rangle \leq 0 \qquad \forall w \in V.$$

Inversamente, supongamos que v satisface la identidad anterior. Así pues,

$$\|v-u\|^2 - \|w-u\|^2 = 2\langle u-v, w-v\rangle - \|v-w\|^2 \leq 0 \qquad \forall w \in V.$$

lo que implica que $\|u-v\| = \operatorname{ínf}_{w\in V} \|u-w\|$.

En el resultado anterior, a v se le conoce como la *proyección* de u sobre V. Escribimos la definición formal a continuación para cualquier subespacio vectorial V cerrado de H pues éste resulta ser siempre convexo y no vacío. En efecto, sean $v, w \in V$ arbitrarios. Entonces, $tv, (1-t)w \in V$ para todo $t \in [0,1]$ y, por tanto, $tv + (1-t)w \in V$ para todo $t \in [0,1]$ por ser subespacio vectorial de H. En consecuencia, $[u,v] \subset V$.

Definición 1.24 *Sea V un subespacio vectorial cerrado de H.* ***La proyección ortogonal de H sobre V*** *es la función $p_V : H \to V$ que a cada elemento $u \in H$ le asocia el único elemento $p_V(u) \in V$ tal que:*

$$\|u - p_V(u)\| = \operatorname*{ínf}_{w\in V} \|u-w\|.$$

Teorema 1.25 *Si V es un subespacio vectorial cerrado de H y $u \in H$, entonces la proyección ortogonal de u sobre V es el único $v \in V$ tal que $u - v \in V^\perp$.*

Demostración: Sea $p_V(u)$ la proyección ortogonal u sobre V. Si $v \in V$, entonces $u-v \in V^\perp$ si y sólo si $\|u-v\| = \operatorname*{ínf}_{w\in V} \|u-v-w\| = \operatorname*{ínf}_{z\in V} \|u-z\|$ (ver Proposición 1.22). El Teorema 1.23 segura que $p_V(u) \in V$ y, por definición $u - p_V(u) \in V^\perp$.

Sea $v \in V$ tal que $u - v \in V^\perp$. Probaremos que $\|u-v\| = \operatorname*{ínf}_{w\in V} \|u-w\|$ lo que implicará que $v = p_V(u)$. Es claro que $\operatorname*{ínf}_{w\in V} \|u-w\| \leq \|u-v\|$ pues $v \in V$.

Sea $w \in V$. Dado que $(u-v) \in V^\perp$ y $(v-w) \in V$, del teorema de Pitágoras obtenemos que:

$$\|u-w\|^2 = \|(u-v)+(v-w)\|^2 = \|u-v\|^2 + \|v-w\|^2 \geq \|u-v\|^2$$

y, en consecuencia:

$$\|u-w\| \geq \|u-v\|.$$

Por tanto, $\|u-v\| \leq \operatorname*{ínf}_{w\in V} \|u-w\|$ y esto concluye la demostración. ■

Veamos algunos ejemplos.

Ejemplo 1.26 *Sea* $(\Omega, \mathcal{S}, \mu)$ *un espacio de medida finita y* $h : \Omega \to [0, \infty)$ *una función medible tal que* $h \in L^\infty(\Omega)$. *Definamos el conjunto:*

$$\mathcal{K} := \{u \in L^2(\Omega) \, : \, |u(x)| \leq h(x) \;\; p.c.t. \; x \in \Omega\}.$$

Entonces $\mathcal{K}$ *es un conjunto cerrado no vacío y convexo de* $L^2(\Omega)$.

Demostración: Dado que $0(x) \leq h(x)$ para todo $x \in \Omega$ y claramente $0 \in L^2(\Omega)$ entonces $0 \in \mathcal{K}$ y así probamos que es no vacío. Sean $u, v \in \mathcal{K}$ y denotemos por $[u, v]$ al segmento:

$$[u, v] = \{\alpha u + (1 - \alpha)v \, : \, \alpha \in [0, 1]\}.$$

Sea $w \in [u, v]$ entonces $w = \alpha_0 u + (1 - \alpha_0)v$ para algún $\alpha_0 \in [0, 1]$. Por tanto,

$$\begin{aligned} |w(x)| &= |\alpha_0 u(x) + (1 - \alpha_0)v(x)| \leq \alpha_0 |u(x)| + (1 - \alpha_0)|v(x)| \\ &\leq \alpha_0 h(x) + (1 - \alpha_0)h(x) = h(x) \; p.c.t. \; x \in \Omega \end{aligned}$$

y, en consecuencia $\mathcal{K} \subset L^2(\Omega)$ es convexo. Sea $u \in \overline{\mathcal{K}}$. Existe una sucesión (u_k) en $\mathcal{K}$ tal que $u_k \to u$ en $L^2(\Omega)$. Dado que $|u_k(x)| \leq h(x)$ p.c.t. $x \in \Omega$ y para todo $k \in \mathbb{N}$ entonces $|u(x)| \leq h(x)$ p.c.t. $x \in \Omega$. Por tanto, $u \in \mathcal{K}$ y esto prueba que $\mathcal{K}$ es cerrado en $L^2(\Omega)$.

Definamos $p_{\mathcal{K}} : L_2(\Omega) \to \mathcal{K}$ como sigue:

$$p_{\mathcal{K}}(u) := \begin{cases} u & \text{si } |u| \leq h, \\ \frac{uh}{|u|} & \text{si } |u| > h. \end{cases}$$

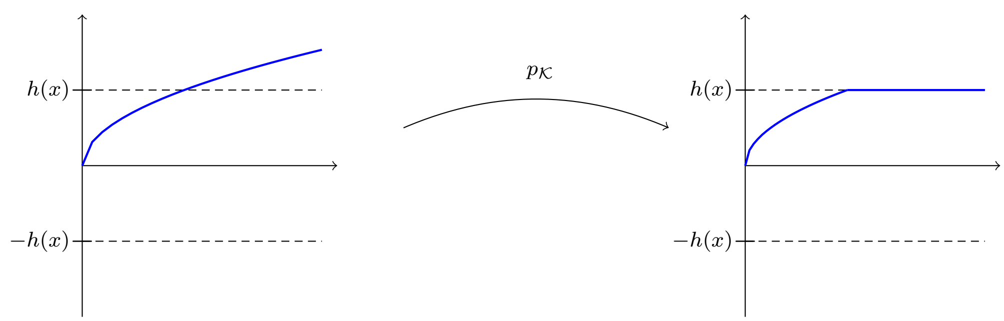


Figura 1.4: Proyección de $L^2(\Omega)$ en $\mathcal{K}$

Caso 1. Si $u \in L^2(\Omega)$ es tal que $|u(x)| \leq h(x)$ p.c.t $x \in \Omega$ es claro que $p_{\mathcal{K}}(u) \in \mathcal{K}$ y que:

$$\|u - p_{\mathcal{K}}(u)\|_2 \leq \|u - v\|_2 \qquad \forall v \in \mathcal{K}.$$

Caso 2. Si $u \in L_2(\Omega)$ es tal que $|u(x)| > h(x)$ p.c.t $x \in \Omega$ entonces:

$$|p_{\mathcal{K}}(u)| = \frac{|u(x)||h(x)|}{|u(x)|} = h(x) \leq h(x)$$

lo que establece que $p_{\mathcal{K}}(u) \in \mathcal{K}$. Sea $v \in \mathcal{K}$ entonces $|v(x)| \leq h(x)$ p.c.t. $x \in \Omega$. Es suficiente demostrar que:

$$|u(x) - p_{\mathcal{K}}(u(x))| \leq |u(x) - v(x)| \quad p.c.t.\ x \in \Omega.$$

Entonces:

$$\begin{aligned} |u(x) - p_{\mathcal{K}}(u(x))| &= \left| u(x) - \frac{u(x)h(x)}{|u(x)|} \right| = \left| \frac{u(x)|u(x)| - u(x)h(x)}{|u(x)|} \right| \\ &= |u(x)| - h(x) \leq |u(x)| - |v(x)| \leq |u(x) - v(x)| \quad p.c.t.\ x \in \Omega. \end{aligned}$$

Así pues, $p_{\mathcal{K}}$ cumple con ser la proyección. ■

Ejemplo 1.27 *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{G} \subset \mathcal{F}$ una sub σ-álgebra. Sea $X : \Omega \to \mathbb{R}$ una variable aleatoria tal que $\|X\|_{L^2(\Omega)} < \infty$. La variable aleatoria $\xi : \Omega \to \mathbb{R}$ es llamada **esperanza condicional de X con respecto de la σ-álgebra** $\mathcal{G}$ si y sólo si:*

(i) ξ es $\mathcal{G}$-medible.

(ii) Para cada $A \in \mathcal{G}$,

$$\int_A \xi \, dP = \int_A X \, dP.$$

Demostraremos que la esperanza condicional siempre existe y es única en el espacio de Hilbert $L^2(\Omega, \mathcal{F}, P)$ (ver Ejemplo 1.10*).*

Demostración: El producto escalar definido en el Ejemplo 1.10, para cuando tenemos un espacio de probabilidad, resulta ser la esperanza del producto de dos variables aleatorias. Es decir:

$$\langle X, Y \rangle_{L^2(\Omega)} = \int_\Omega XY \, dP = E(XY) \qquad X, Y \in L^2(\Omega, \mathscr{F}, P).$$

El conjunto de las variables aleatorias $\mathcal{G}$-medibles con valores reales tales que $E(X^2) < \infty$ es el subespacio vectorial $L^2(\Omega, \mathcal{G}, P)$ de $L^2(\Omega, \mathcal{F}, P)$.

Como la demostración es larga, la subdividimos en dos pasos

Paso 1: Sea (X_k) una sucesión en $L^2(\Omega, \mathcal{G}, P)$ tal que $X_k \to X$ en $L^2(\Omega, \mathcal{F}, P)$. Existe (X_{k_j}) una subsucesión de (X_k) y $N \in \mathcal{G}$ con $P(N) = 0$ tal que $X_{k_j}(\omega) \to X(\omega)$ para todo $\omega \in \Omega \setminus N$.

Definiendo $\widetilde{X}_j := X_{k_j} \cdot 1_{\Omega \setminus N}$ y $\widetilde{X}(\omega) = \lim_{k \to \infty} \widetilde{X}_j(\omega)$ para todo $\omega \in \Omega$, es claro que $\widetilde{X}$ es una función $\mathcal{G}$-medible y $X = \widetilde{X}$ c.d. En consecuencia, $X \in L^2(\Omega, \mathcal{G}, P)$ lo que prueba que $L^2(\Omega, \mathcal{G}, P)$ es un subespacio vectorial cerrado de $L^2(\Omega, \mathcal{F}, P)$.

Paso 2: Como $L^2(\Omega, \mathcal{G}, P)$ es un subespacio vectorial cerrado de $L^2(\Omega, \mathcal{F}, P)$, el Teorema 1.23 asegura que existe una única $\xi \in L^2(\Omega, \mathcal{G}, P)$ tal que:

$$\|X - \xi\|_{L^2(\Omega)} = \inf \left\{ \|X - Y\|_{L^2(\Omega)} \ : \ Y \in L^2(\Omega, \mathcal{F}, P) \right\}.$$

Por el Teorema 1.25, ξ es la proyección ortogonal de $L^2(\Omega, \mathcal{F}, P)$ sobre $L^2(\Omega, \mathcal{G}, P)$ y, es la única variable aleatoria tal que $E(X - \xi, Y) = 0$ para toda $Y \in L^2(\Omega, \mathcal{G}, P)$. Además, ξ claramente es $\mathcal{G}$-medible.

Sea $A \in \mathcal{G}$, entonces

$$\int_A X\, dP - \int_A \xi\, dP = \int_\Omega (X - \xi) \cdot 1_A\, dP = E(X - \xi, 1_A) = 0,$$

pues $\int |1_A|^2\, dP = P(A) \leq 1$. En consecuencia, ξ es la esperanza condicional de X con respecto de la σ-álgebra $\mathcal{G}$. ■

El Ejemplo 1.21 muestra que la suma directa de V y $V^\perp$ en general no es todo H. Lo anterior ocurre cuando el subespacio V es cerrado. Es decir, si V es cerrado entonces $V \oplus V^\perp = H$ y en este caso al espacio $V^\perp$ lo llamamos el **complemento ortogonal de** V **en** H**.** Probamos a continuación tales afirmaciones.

Antes damos las siguientes definiciones: Si V y W son espacios de Banach, con normas $\|\cdot\|_V$ y $\|\cdot\|_W$ respectivamente, el espacio

$$\mathcal{L}_c(V, W) := \{T : V \to W \ :\ T \text{ es lineal y continua}\}$$

con la norma:

$$\|T\|_{\mathcal{L}_c(V,W)} := \sup_{v \in V \smallsetminus \{0_V\}} \frac{\|T(v)\|_W}{\|v\|_V} := \sup_{\|v\|_V \leq 1} \|T(v)\|_W := \sup_{\|v\|_V = 1} \|T(v)\|_W$$

resulta ser un espacio de Banach. De hecho, $\|T\|_{\mathcal{L}_c(V,W)}$ es el mínimo valor $c \in \mathbb{R}$ que satisface que $\|T(v)\|_W \leq c\, \|v\|_V$ para todo $v \in V$. Si $W = \mathbb{R}$ escribimos

$$V^* := \mathcal{L}_c(V, \mathbb{R})$$

y lo llamamos el **dual (topológico) de** V.

Además, una función $\iota : V \to W$ es una **isometría** si:

$$\|\iota(u) - \iota(v)\|_W = \|u - v\|_V \qquad \forall\, u, v \in V.$$

En el caso en que ι sea lineal y biyectiva diremos que es un **isomorfismo isométrico**.

Teorema 1.28 (Complemento ortogonal) *Sea V un subespacio vectorial cerrado de H. Entonces:*

(a) La proyección ortogonal $p_V : H \to V$ es la única función lineal de H en V que cumple:

(a.1) $p_V \circ p_V = p_V$.
(a.2) $\ker p_V = V^\perp$.

(b) p_V *es continua y* $\|p_V\|_{\mathcal{L}_c(H,V)} \leq 1$.

(c) La función $\iota : V \oplus V^{\perp} \to H$ *dada por* $\iota(w) := u + v$ *es un isomorfismo lineal y una isometría.*

Demostración: (a) Sean $u, \tilde{u} \in H$ y $\lambda, \mu \in \mathbb{R}$. Entonces $p_V(u), p_V(\tilde{u}) \in V$ y se sigue entonces que $\lambda p_V(u) + \mu p_V(\tilde{u}) \in V$. Notemos que:

$$(\lambda u + \mu \tilde{u}) - (\lambda p_V(u) + \mu p_V(\tilde{u})) = \lambda(u - p_V(u)) + \mu(\tilde{u} - p_V(\tilde{u})) \in V^{\perp}.$$

Por tanto, de la unicidad de la proyección ortogonal (ver Teorema 1.25) se sigue que $p_V(\lambda u + \mu \tilde{u}) = \lambda p_V(u) + \mu p_V(\tilde{u})$.

(a.1) Dado que para todo $v \in V$, $v - v = 0_H \in V^{\perp}$ entonces $v = p_V(v)$ por el Teorema 1.25. En consecuencia, $p_V(p_V(u)) = p_V(u)$ para todo $u \in H$.

(a.2) Si $v \in \ker p_V$ entonces $p_V(v) = 0_H$. El Teorema 1.25 asegura entonces que $v - p_V(v) = v \in V^{\perp}$. Inversamente, si $v \in V^{\perp}$ entonces $\langle v, u\rangle = 0$ para todo $u \in V$. Nuevamente, por el Teorema 1.25 se sigue que $v - p_V(v) \in V^{\perp}$ por lo que $\langle v - p_V(v), u\rangle = 0$ para todo $u \in V$. Por tanto,

$$\langle p_V(v), u\rangle = \langle p_V(v) - v, u\rangle + \langle v, u\rangle = 0 \qquad \forall\, u \in V$$

lo que establece que $p_V(v) \in V \cap V^{\perp}$. La Proposición 1.20 asegura que $p_V(v) = 0_H$.

Si $T : H \to V$ es una función lineal que satisface *(a.1)* entonces, para todo $u \in H$, se cumple que $T(u - T(u)) = 0_H$. Si además T satisface *(a.2)* entonces $u - T(u) \in V^{\perp}$ y el Teorema 1.25 asegura que $T(u) = p_V(u)$ lo que demuestra la unicidad.

(b) Como $u = p_V(u) + (u - p_V(u))$ y $\langle p_V(u), u - p_V(u)\rangle = 0$ para todo $u \in H$. Aplicando el Teorema 1.17 obtenemos:

$$\|u\|^2 = \|p_V(u)\|^2 + \|u - p_V(u)\|^2 \qquad \forall u \in H. \tag{1.15}$$

Por tanto, $\|p_V(u)\| \leq \|u\|$ para todo $u \in H$ y esto implica que p_V es continua. Además, obtenemos que:

$$\|p_V\|_{\mathcal{L}_c(H,V)} := \sup_{u \in H \smallsetminus \{0_H\}} \frac{\|p_V(u)\|}{\|u\|} \leq 1.$$

(c) La función $\iota : V \oplus V^{\perp} \to H$ dada por $\iota(w) := u + v$ es claramente lineal y la función $j : H \to V \oplus V^{\perp}$ dada por $j(u) := (p_V(u)) + (u - p_V(u))$ es su inverso. En efecto, el Teorema 1.25 asegura que j está bien definida. Ahora, si $w \in V \oplus V^{\perp}$ entonces existe únicos $u \in V$ y $v \in V^{\perp}$ tales que $w = u + v$ y, por tanto,

$$(j \circ \iota)(w) = j(\iota(w)) = j(u + v) = (p_V(u + v)) + ((u + v) - p_V(u + v)) = u + v = w.$$

Por otro lado, si $u \in H$ el Teorema 1.25 asegura que $p_V(u) \in V$ y $u - p_V(u) \in V^{\perp}$. Dado que $u = p_V(u) + (u - p_V(u))$ y $p_V(u)$ es único, entonces $u \in V \oplus V^{\perp}$. En consecuencia,

$$(\iota \circ j)(u) = \iota(j(u)) = p_V(u) + (u - p_V(u)) = u.$$

Por otro lado, sea $w := u + v \in V \oplus V^{\perp}$ entonces por el Teorema 1.17 y el Ejemplo 1.12 obtenemos:

$$\|\iota(w)\|^2 = \|u+v\|^2 = \|u\|^2 + \|v\|^2 = \|w\|_{V\oplus V^{\perp}}.$$

En consecuencia, $\iota : V \oplus V^{\perp} \to H$ es un isomorfismo isométrico. ■

Corolario 1.29 *Si V un subespacio vectorial cerrado de H, entonces:*

(a) $p_V(u) + p_{V^{\perp}}(u) = u$ *para todo* $u \in H$.

(b) $V = (V^{\perp})^{\perp}$.

Demostración: (a) Sea $u \in H$. El Teorema 1.28 asegura que $u = p_V(u) + (u - p_V(u))$. Luego, como $p_V(u) = u-(u-p_V(u))$ y $V \subset (V^{\perp})^{\perp}$ entonces $p_V(u) = u-(u-p_V(u)) \in (V^{\perp})^{\perp}$. Por el Teorema 1.25 concluimos que $u - p_V(u) = p_{V^{\perp}}(u)$. En consecuencia, $u = p_V(u) + (u - p_V(u)) = p_V(u) + p_{V^{\perp}}(u)$ como afirma el enunciado.

(b) Sea $v \in (V^{\perp})^{\perp}$. El Teorema 1.28 implica que $v = p_{V^{\perp}}(v) + (v - p_{V^{\perp}}(v))$. Del inciso *(a.2)* del Teorema 1.28 se tiene que $p_{V^{\perp}}(v) = 0_H$. Del inciso anterior concluimos que $v = v - p_{V^{\perp}}(v) = p_V(v)$ y esto prueba que $v \in V$. ■

Notemos que, de manera general, para cualquier subespacio vectorial V de H se tiene que $\overline{V} = (V^{\perp})^{\perp}$ y si V es denso en H entonces $V^{\perp} = \{0_H\}$.

Recordemos que para U subconjunto de V con V espacio vectorial sobre $\mathbb{R}$, denotamos por $\mathrm{lin}(U)$ al subespacio vectorial de V dado por:

$$\mathrm{lin}(U) := \left\{ \sum_{j=1}^{m} \alpha_j \, v_j \; : \; \alpha_j \in \mathbb{R}, \, v_j \in U, \, m \in \mathbb{N} \right\}.$$

Es decir, el conjunto de todas las combinaciones lineales finitas de elementos de U. Notemos que $\mathrm{lin}(U)$ es el subespacio vectorial más pequeño que contiene a U. En este trabajo convenimos que $\mathrm{lin}(\varnothing) := \{0_V\}$.

Corolario 1.30 *Para X cualquier subconjunto no vacío de H se tiene que:*

$$(X^{\perp})^{\perp} = \overline{\mathrm{lin}(X)}.$$

Demostración: Dado que $X \subset \mathrm{lin}(X)$ entonces $(\mathrm{lin}(X))^{\perp} \subset X^{\perp}$. Si $v \in X^{\perp}$ entonces $\langle v, x\rangle = 0$ para todo $x \in X$. Sea $u := \sum_{i=1}^{m} \alpha_i x_i$ con $\alpha_i \in \mathbb{R}$ y $x_i \in X$ para todo $i = 1, \ldots, m$ y $m \in \mathbb{N}$. En consecuencia,

$$\langle x, u\rangle = \left\langle x, \sum_{i=1}^{m} \alpha_i x_i \right\rangle = \sum_{i=1}^{m} \alpha_i \langle v, x_i\rangle = 0,$$

es decir, $v \in (\mathrm{lin}(X))^{\perp}$. Por tanto, $X^{\perp} = (\mathrm{lin}(X))^{\perp} = (\overline{\mathrm{lin}(X)})^{\perp}$.

Así pues, como $\overline{\mathrm{lin}(X)}$ es un subespacio vectorial cerrado de H entonces (ver Corolario 1.29):

$$(X^{\perp})^{\perp} = ((\overline{\mathrm{lin}(X)})^{\perp})^{\perp} = \overline{\mathrm{lin}(X)}$$

como afirma el enunciado. ■

1.3. El teorema de representación de Fréchet-Riesz

Sea $H = (H, \langle \cdot, \cdot \rangle, \| \cdot \|)$ un espacio de Hilbert sobre $\mathbb{R}$. En esta sección describiremos al espacio dual del espacio de Hilbert H.

Proposición 1.31 *Para cada $w \in H$ la función $T_w : H \to \mathbb{R}$ dada por:*

$$T_w(u) := \langle w, u \rangle, \tag{1.16}$$

es lineal, continua y cumple que

$$\|T_w\|_{H^*} = \|w\|. \tag{1.17}$$

Demostración: Sea $w \in H$. Es inmediato de las propiedades (PE1) y (PE2) de la Definición 1.1 que T_w es lineal. Por otro lado, la Proposición 1.6 asegura que T_w es lineal. Si $w = 0_H$ entonces la igualdad (1.17) se satisface trivialmente. Supongamos que $w \neq 0_H$. Usando la desigualdad de Cauchy-Schwarz obtenemos que:

$$\|w\| = \frac{\|w\|^2}{\|w\|} = \frac{|T_w(w)|}{\|w\|} \leq \sup_{u \in H \setminus \{0_H\}} \frac{|T_w(u)|}{\|u\|} \leq \sup_{u \in H \setminus \{0_H\}} \frac{\|w\|\|u\|}{\|u\|} = \|w\|$$

por lo que $\|w\| = \|T_w\|_{H^*}$. Esto concluye la prueba. ■

Antes de probar el teorema de representación de Fréchet-Riesz que nos brinda una descripción precisa del espacio dual de un espacio de Hilbert H requerimos el siguiente lema.

Lema 1.32 *Sea H un espacio de Hilbert y $T : H \to \mathbb{R}$ lineal no constante igual a cero. Entonces son equivalentes:*

(i) *T es continua.*

(ii) $\ker T := \{u \in H \,:\, T(u) = 0\}$ *es un subespacio vectorial cerrado en H.*

Demostración: *(i)* $\Rightarrow$ *(ii)* : Sea $v \in \ker T$. Existe una sucesión (v_k) tal que $v_k \in \ker T$ para todo $k \in \mathbb{N}$ y $v_k \to v$ en H. Usando la continuidad de T $T(v_k) \to T(v)$ en $\mathbb{R}$ donde $T(v_k) = 0$ para todo $k \in \mathbb{N}$. Por tanto, $v \in \ker T$ y esto prueba que $\ker T$ es cerrado.

$(ii) \Rightarrow (i)$: Supongamos que T no es continua. Existe $v_k \in \mathbb{S}_H(0_H, 1) := \{u \in H : \|u\| = 1\}$ tal que $|T(v_k)| > k$ para cada $k \in \mathbb{N}$. Sea $u \in H$ tal que $T(u) \neq 0$ y definamos $u_k := u - \frac{v_k}{T(v_k)} T(u) \in H$ para cada $k \in \mathbb{N}$. Es claro que $T(u_k) = 0$ para todo $k \in \mathbb{N}$ por lo que (u_k) es una sucesión de elementos de $\ker T$ y $u_k \to u$ pues:

$$\lim_{k\to\infty} \|u_k - u\| = \lim_{k\to\infty} \left\| -\frac{v_k}{T(v_k)} T(u) \right\| = \lim_{k\to\infty} |T(u)| \frac{\|v_k\|}{|T(v_k)|} \leq \lim_{k\to\infty} \frac{|T(u)|}{k} = 0.$$

Por tanto, $u \in \overline{\ker T}$ pero $T(u) \neq 0$ lo que demuestra que $\ker T$ no es cerrado. ■

Notemos que en el lema anterior no usamos alguna característica propia del espacio de Hilbert mas que las propiedades de la norma. Por tanto, si en la hipótesis sustituimos al espacio de Hilbert H por cualquier espacio vectorial normado, la demostración es la misma. Además, si $T : H \to \mathbb{R}$ es la función constante con valor 0 para toda $u \in H$ se sigue trivialmente que es lineal, continua y $\ker T = H$.

Teorema 1.33 (de representación de Fréchet-Riesz) *Sea H un espacio de Hilbert y $T : H \to \mathbb{R}$ una función lineal continua. Entonces existe un único $w \in H$ tal que:*

$$T(u) = \langle w, u \rangle := T_w(u) \qquad \forall\, u \in H. \tag{1.18}$$

Más aún, la función $\iota : H \to H^*$ dada por $\iota(w) := T_w$ es un isomorfismo lineal y una isometría.

Demostración: Sea $T \in H^*$ y denotemos por $V := \ker T$. Dado que T es lineal y continua, el Lema 1.32 asegura que V es un subespacio vectorial cerrado de H. Notemos que si $V = H$ entonces $T = 0$ y $w = 0_H$ por lo que (1.18) se satisface trivialmente. Así pues, supongamos que $V \neq H$. El Teorema 1.28 asegura que $V^\perp \neq \{0\}$. Escojamos $w_0 \in V^\perp$ tal que $\|w_0\| = 1$. Entonces $T(w_0) \neq 0$ y definamos $w := (T(w_0))w_0 \in H$. En consecuencia,

$$T\left(u - \frac{T(u)}{T(w_0)} w_0\right) = T(u) - \frac{T(u)}{T(w_0)} T(w_0) = 0 \qquad \forall u \in H.$$

Por tanto, $u - \frac{T(u)}{T(w_0)} w_0 \in V$ y se sigue que:

$$\begin{aligned} \langle w, u \rangle &= \left\langle w, u - \frac{T(u)}{T(w_0)} w_0 + \frac{T(u)}{T(w_0)} w_0 \right\rangle = \left\langle w, u - \frac{T(u)}{T(w_0)} w_0 \right\rangle + \frac{T(u)}{T(w_0)} \langle w, w_0 \rangle \\ &= \frac{T(u)T(w_0)}{T(w_0)} \langle w_0, w_0 \rangle = T(u) \quad \text{para todo } \ u \in H. \end{aligned}$$

Supongamos que existe $\widetilde{w} \in H$ tal que cumple (1.18). Entonces $\langle w - \widetilde{w}, u \rangle = 0$ para todo $u \in H$ y, en particular, para $u = w - \widetilde{w}$. Por tanto, $\|w - \widetilde{w}\|^2 = 0$ lo que implica que $w = \widetilde{w}$.

Por otro lado,

$$\iota(\lambda w_1 + \mu w_2)(u) = \langle \lambda w_1 + \mu w_2, u\rangle = \lambda\langle w_1, u\rangle + \mu\langle w_2, u\rangle = (\lambda\, \iota(w_1) + \mu\, \iota(w_2))(u)$$

para todo $u \in H$. Por tanto, ι es lineal y es biyectiva por la discusión de la primera parte de este teorema. La Proposición 1.31 asegura que ι es una isometría. ■

El teorema de representación de Fréchet-Riesz afirma que el espacio dual de un espacio de Hilbert se identifica con él mismo. Notemos que, dado un espacio de medida $(\Omega, \mathcal{S}, \mu)$ y el espacio de Hilbert $H = L^2(\Omega)$, el teorema anterior resulta ser el teorema de representación de Riesz (ver [13]).

Sea V un espacio de Banach con norma $\|\cdot\|_V$. Al espacio $\mathcal{L}_c(V^*, \mathbb{R})$ lo llamamos el espacio **bidual (topológico) de** V y resulta ser un espacio de Banach. Decimos que V es **reflexivo** si y sólo si para toda $f \in V^{**}$ existe $v \in V$ tal que $f(T) = T(v)$ para toda $T \in V^*$. De manera intuitiva, V es reflexivo si se puede identificar bajo un isomorfismo isométrico con su espacio bidual.

Una consecuencia intuitiva del teorema de Fréchet-Riesz es que todo espacio de Hilbert H es reflexivo. En efecto, sea $f \in H^{**}$. El Teorema 1.33 asegura que la función $\iota : H \to H^*$ es un isomorfismo lineal isométrico. Así, la función $f \circ \iota : H \to \mathbb{R}$ es lineal y continua pues $|f(\iota(w))| \leq \|f\|_{H^{**}}\|w\|$ para todo $w \in H$. El teorema de representación de Fréchet-Riesz asegura que existe $v \in H$ tal que $f(\iota(u)) = \langle v, u\rangle$ para todo $u \in H$. Sea $T \in H^*$, nuevamente el teorema de representación de Fréchet-Riesz implica que existe un único $w \in H$ tal que $T(u) = \langle w, u\rangle = T_w(u) = \iota(w)$. Por tanto, $T(v) = \langle v, w\rangle = T_w(v) = f(T)$.

A continuación damos una demostración alterna al hecho de que todo espacio de Hilbert es reflexivo.

Teorema 1.34 *Sea H un espacio de Hilbert. Entonces H es uniformemente convexo.*

Demostración: Sean $\varepsilon > 0$ y $u, v \in H$ tales que $\|u\|, \|v\| \leq 1$ y $\|u - v\| > \varepsilon$. Aplicando la identidad del paralelogramo (ver Teorema 1.4) obtenemos que:

$$\left\|\frac{u+v}{2}\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right) \leq 1.$$

Por tanto,

$$\left\|\frac{u+v}{2}\right\|^2 \leq 1 - \left\|\frac{u-v}{2}\right\|^2 < 1 - \frac{\varepsilon^2}{4}$$

y definiendo $\delta := 1 - (1 - \frac{\varepsilon^2}{4})^{1/2} > 0$ se sigue el resultado. ■

Corolario 1.35 *Sea H un espacio de Hilbert. Entonces H es reflexivo.*

Demostración: Es consecuencia del Teorema de Milman-Pettis (ver [13]) y del Teorema 1.34. ■

1.4. Bases de Hilbert

Sea H un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle\cdot,\cdot\rangle$.

Definición 1.36 *Un subconjunto $\mathcal{O}$ de un espacio de Hilbert H se llama un **conjunto ortogonal** si*

$$\langle u,v\rangle = 0 \qquad \forall\, u,v\in\mathcal{O},\; u\neq v.$$

*Si además $\|u\|=1$ para todo $u\in\mathcal{O}$, se dice que $\mathcal{O}$ es un **conjunto ortonormal**.*

Lema 1.37 *Si $\mathcal{O}$ es un conjunto ortonormal de H entonces $\mathcal{O}$ es linealmente independiente.*

Demostración: Sean $k\in\mathbb{N}$, $e_1,\ldots,e_k\in\mathcal{O}$ con $e_i\neq e_j$ para toda $i\neq j$ y $\alpha_1,\ldots\alpha_k\in\mathbb{R}$ tales que:

$$0_H = \sum_{j=1}^{k}\alpha_j\, e_j.$$

La Proposición 1.16 asegura que $\langle 0_H, e_i\rangle = 0$ para toda $i=1,\ldots,k$. Por tanto,

$$0 = \langle 0_H, e_i\rangle = \left\langle \sum_{j=1}^{k}\alpha_j e_j\, e_i, \right\rangle = \sum_{j=1}^{k}\alpha_j\langle e_j, e_i\rangle = \alpha_i \quad \text{ para toda } \; i=1,\ldots,k.$$

En consecuencia, $\mathcal{O}$ es linealmente independiente. ■

Una forma más completa de expresar el lema anterior es la siguiente.

Lema 1.38 *Sea $\mathcal{I}$ conjunto no vacío y $\mathcal{O}=\{e_i : i\in\mathcal{I}\}$ un subconjunto ortonormal de H. Si $\mathcal{J}\subset\mathcal{I}$ es finito y $u=\sum_{j\in\mathcal{J}}\alpha_j\, e_j$ con $\alpha_j\in\mathbb{R}$ para toda $j\in\mathcal{J}$ entonces $\langle u,e_j\rangle=\alpha_j$ para toda $j\in\mathcal{J}$ y $\langle u,e_j\rangle = 0$ para toda $j\in\mathcal{I}\smallsetminus\mathcal{J}$. Más aún,*

$$\|u\|^2 = \sum_{j\in\mathcal{J}}|\alpha_j|^2.$$

Demostración: Notemos que $\langle u,e_i\rangle = \left\langle \sum_{j\in\mathcal{J}}\alpha_j e_j, e_i\right\rangle = \sum_{j\in\mathcal{J}}\alpha_j\langle e_j,e_i\rangle = \alpha_i$ para toda $i\in\mathcal{J}$. Por otro lado, si $i\in\mathcal{I}\smallsetminus\mathcal{J}$ entonces $i\neq j$ para toda $j\in\mathcal{J}$ de modo que $\langle e_i,e_j\rangle=0$ para toda $j\in\mathcal{J}$. En consecuencia, $\langle u,e_i\rangle = 0$.

Finalmente,

$$\|u\|^2 = \langle u,u\rangle = \left\langle \sum_{j\in\mathcal{J}}\alpha_j e_j, \sum_{i\in\mathcal{J}}\alpha_i e_i\right\rangle = \sum_{j\in\mathcal{J}}\alpha_j\left(\sum_{i\in\mathcal{J}}\alpha_i\,\langle e_j,e_i\rangle\right) = \sum_{j\in\mathcal{J}}|\alpha_j|^2$$

como afirma el enunciado. ■

Proposición 1.39 *Sea $m \in \mathbb{N}$ arbitrario pero fijo, $\mathcal{O} = \{e_i \,:\, 1 \le i \le m\}$ un subconjunto ortonormal de H y $u \in H$. Entonces el subespacio vectorial generado por $\mathcal{O}$, $\operatorname{lin}(\mathcal{O})$, es cerrado en H y*

$$p_{\operatorname{lin}(\mathcal{O})}(u) = \sum_{i=1}^{m} \langle u, e_i\rangle\, e_i.$$

Demostración: Sea $v \in \overline{\operatorname{lin}(\mathcal{O})}$. Existe una sucesión (v_k) tal que $v_k \in \operatorname{lin}(\mathcal{O})$ para todo $k \in \mathbb{N}$ y $v_k \to v$ en H. Así pues, para cada $k \in \mathbb{N}$, existen $\alpha_1^k, \ldots, \alpha_m^k$ reales tales que $v_k = \sum_{i=1}^m \alpha_i^k\, e_i$. Como $v_k \to v$ entonces (v_k) es de Cauchy en H. Dado que para cada $j, k \in \mathbb{N}$ se tiene que $v_j - v_k = \sum_{i=1}^m (\alpha_i^j - \alpha_i^k)\, e_i$, aplicando el Lema 1.48 obtenemos:

$$|\alpha_i^j - \alpha_i^k|^2 \le \sum_{i=1}^{m} |\alpha_i^j - \alpha_i^k|^2 = \|v_j - v_k\|^2 \;\text{ para todo }\; i = 1, \ldots, m.$$

En consecuencia, (α_i^k) es una sucesión de Cauchy en $\mathbb{R}$ y, por tanto, converge para todo $i = 1, \ldots, m$. Sea $\alpha_i := \operatorname{lím}_{k\to\infty} \alpha_i^k$ para todo $i = 1, \ldots, m$ y definamos $\tilde{v} = \sum_{i=1}^m \alpha_i\, e_i \in \operatorname{lin}(\mathcal{O})$. Entonces,

$$\operatorname*{lím}_{k\to\infty} \|v_k - \tilde{v}\| \le \operatorname*{lím}_{k\to\infty} \left\| \sum_{i=1}^{m} (\alpha_i^k - \alpha_i)\, e_i \right\| \le \sum_{i=1}^{m} \operatorname*{lím}_{k\to\infty} |\alpha_i^k - \alpha_i| \|e_i\| = 0$$

por lo que necesariamente $\tilde{v} = v$ y esto prueba que $\operatorname{lin}(\mathcal{O})$ es cerrado en H.

Sea $v := \sum_{i=1}^m \langle u, e_i\rangle\, e_i \in \operatorname{lin}(\mathcal{O})$. Entonces, para todo $j = 1, \ldots, m$ se tiene que:

$$\langle u - v, e_j\rangle = \langle u, e_j\rangle - \left\langle \sum_{i=1}^{m} \langle u, e_i\rangle\, e_i, e_j \right\rangle = \langle u, e_j\rangle - \sum_{i=1}^{m} \langle u, e_i\rangle \langle e_i, e_j\rangle = \langle u, e_j\rangle - \langle u, e_j\rangle = 0$$

es decir, $u - v \in (\operatorname{lin}(\mathcal{O}))^\perp$. El Teorema 1.25 asegura que $v = p_{\operatorname{lin}(\mathcal{O})}(u)$. ■

Corolario 1.40 *Sea $m \in \mathbb{N}$ arbitrario pero fijo, $\mathcal{O} = \{e_i \,:\, 1 \le i \le m\}$ un subconjunto ortonormal de H y $u \in H$. Si $v := \sum_{i=1}^m \langle u, e_i\rangle\, e_i$ entonces:*

$$\|u\|^2 = \|u - v\|^2 + \sum_{i=1}^{m} |\langle u, e_i\rangle|^2. \tag{1.19}$$

En particular, se cumple la ***desigualdad de Bessel:***

$$\|u\|^2 \ge \sum_{i=1}^{m} |\langle u, e_i\rangle|^2. \tag{1.20}$$

Demostración: Sea $V := \operatorname{lin}(\mathcal{O})$. Por la Proposición 1.39 se tiene que $v = p_V(u)$ y, en consecuencia $u - v \in V^{\perp}$. Es decir, $u - v \perp v$ y aplicando el Teorema de Pitágoras (ver Teorema 1.17) junto con el Lema 1.38 obtenemos que:

$$\|u\|^2 = \|u-v\|^2 + \|v\|^2 = \|u-v\|^2 + \sum_{i=1}^{m} |\langle u, e_i\rangle|^2$$

como afirma el enunciado. ■

Teorema 1.41 *Sea $\mathcal{X} = \{v_k : k \in \mathbb{N}\}$ un subconjunto ortogonal de H. Entonces, la serie $\sum_{k=1}^{\infty} v_k$ converge en H si y sólo si $\sum_{k=1}^{\infty} \|v_k\|^2$ converge en $\mathbb{R}$, en cuyo caso se tiene que:*

$$\left\|\sum_{k=1}^{n} u_k\right\|^2 = \sum_{k=1}^{n} \|u_k\|^2. \tag{1.21}$$

Demostración: Sea $w_n := \sum_{k=1}^{n} v_k$ la sucesión de sumas parciales en H. Notemos que, para $m > n$ se tiene que:

$$\|w_m - w_n\|^2 = \langle w_m - w_n, w_m - w_n\rangle = \left\langle \sum_{k=n+1}^{m} v_k, \sum_{k=n+1}^{m} v_k \right\rangle = \sum_{k=n+1}^{m} \|v_k\|^2.$$

Sea $\varepsilon > 0$. Si la sucesión (u_n) de sumas parciales $u_n := \sum_{k=1}^{n} \|v_k\|^2$ converge en $\mathbb{R}$ entonces es de Cauchy en $\mathbb{R}$. Existe $k_0 \in \mathbb{N}$ tal que:

$$|u_n - u_{k_0}| = \left|\sum_{k=1}^{n} \|v_k\|^2 - \sum_{k=1}^{k_0} \|v_k\|^2\right| = \sum_{k=k_0+1}^{n} \|v_k\|^2 < \varepsilon^2, \qquad \forall, u > k_0.$$

En consecuencia,

$$\|w_n - w_{k_0}\|^2 = \sum_{k=k_0+1}^{n} \|v_k\|^2 < \varepsilon^2, \qquad \forall n > k_0.$$

Esto prueba que la sucesión (w_n) es de Cauchy en H y, por tanto, converge en H.

Inversamente, si (w_n) converge en H entonces es de Cauchy en H. Dada $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que:

$$\|w_n - w_{k_0}\| < \sqrt{\varepsilon}, \qquad \forall n > k_0.$$

Por tanto,

$$|u_n - u_{k_0}| = \sum_{k=k_0+1}^{n} \|v_k\|^2 = \|w_n - w_{k_0}\|^2 < \varepsilon \qquad \forall n > k_0.$$

Es decir, (u_n) es de Cauchy en $\mathbb{R}$ y esto implica que (u_n) converge en $\mathbb{R}$.

Finalmente, usando la continuidad de la norma y las propiedades del producto escalar obtenemos:

$$\left\| \sum_{k=1}^{\infty} v_k \right\|^2 = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} v_k \right\|^2 = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} v_k, \sum_{k=1}^{n} v_k \right\rangle = \lim_{n\to\infty} \sum_{k=1}^{n} \|v_k\|^2 = \sum_{k=1}^{\infty} \|v_k\|^2$$

como afirma el enunciado. ■

Teorema 1.42 *Sea* $\mathcal{O} = \{e_k \,:\, k \in \mathbb{N}\}$ *un subconjunto ortonormal de* H*. Entonces,*

$$u = \sum_{k=1}^{\infty} \langle u, e_k \rangle\, e_k \qquad \forall\, u \in \operatorname{lin}(\mathcal{O}).$$

Demostración: Sea $u \in \operatorname{lin}(\mathcal{O})$. Para cada $k \in \mathbb{N}$ definamos $\alpha_k := \langle u, e_k \rangle$ y sea $v = \sum_{k=1}^{\infty} \alpha_k\, e_k$. Por la desigualdad de Bessel (ver Corolario 1.40) se tiene que:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \|\alpha_k\, e_k\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |\langle u, e_k \rangle| \leq \lim_{n\to\infty} \|u\|^2.$$

Es decir, la serie $\sum_{k=1}^{\infty} \|\alpha_k\, e_k\|^2$ converge en $\mathbb{R}$ y, el Teorema 1.41 asegura que la serie $\sum_{k=1}^{\infty} \alpha_k\, e_k$ converge en H. Por tanto, $v \in \operatorname{lin}(\mathcal{O})$ (ver Proposición 1.39). En consecuencia, usando la continuidad del producto escalar obtenemos:

$$\langle v, e_j \rangle = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} \alpha_k\, e_k, e_j \right\rangle = \langle u, e_j \rangle$$

por lo que $v-u \perp e_j$ para todo $j \in \mathbb{N}$ y, por tanto, $v-u \in (\operatorname{lin}(\mathcal{O}))^{\perp}$. Dado que $v-u \in \operatorname{lin}(\mathcal{O})$ concluimos que $v-u = 0_H$ y esto prueba el resultado. ■

Definición 1.43 *Un subconjunto* $\mathcal{B}$ *de* H *se llama una* ***base de Hilbert*** *si es ortonormal y* $\operatorname{lin}(\mathcal{B})$ *es denso en* H*, es decir,* $H = \overline{\operatorname{lin}(\mathcal{B})}$*.*

En general, el conjunto $\operatorname{lin}(\mathcal{B})$ puede no coincidir con el espacio de Hilbert H como lo muestra el siguiente ejemplo.

Ejemplo 1.44 *Sea* $\bar{e}_k := (e_{k,j})$ *la sucesión cuyos términos son* $e_{k,k} = 1$ *y* $e_{k,j} = 0$ *si* $k \neq j$*. Entonces* $\mathcal{B} := \{\bar{e}_k \,:\, k \in \mathbb{N}\}$ *es una base de Hilbert de* $\ell_2(\mathbb{R})$ *(ver* Ejemplo 1.11*).*

Demostración: Es claro que $\mathcal{B}$ es un subconjunto ortonormal de $\ell_2(\mathbb{R})$. Dado que los elementos de $\mathrm{lin}(\mathcal{B})$ son combinaciones lineales finitas de $\mathcal{B}$ entonces

$$\mathrm{lin}(\mathcal{B}) = \{(x_k) \in \ell_2(\mathbb{R}) \,:\, \text{existe } k_0 \in \mathbb{N} \text{ tal que } x_k = 0 \ \forall k \geq k_0\}.$$

En consecuencia $\mathrm{lin}(\mathcal{B}) \neq \ell_2(\mathbb{R})$. Finalmente, sea $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$. Si denotamos por $\bar{x}_k$ a la sucesión cuyos términos son $\bar{x}_{k,j} := x_j$ si $j \leq k$ y $\bar{x}_{k,j} := 0$ si $j > k$ entonces $\bar{x}_k \in \mathrm{lin}(\mathcal{B})$ para todo $k \in \mathbb{N}$ y $\bar{x}_k \to \bar{x}$ en $\ell_2(\mathbb{R})$. ∎

A continuación demostraremos que para todo espacio de Hilbert separable existe una base de Hilbert. Antes, requerimos el siguiente lema.

Lema 1.45 (Ortonormalización de Gram-Schmidt) *Sea $\mathcal{X} = \{v_k \,:\, k \in \mathbb{N}\}$ un subconjunto linealmente independiente de H, es decir, para todo $m \in \mathbb{N}$ y para $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ se cumple que:*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_m = 0_H \quad \Longleftrightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0.$$

Entonces existe un subconjunto ortonormal $\mathcal{O}$ de H tal que $\mathrm{lin}(\mathcal{O}) = \mathrm{lin}(\mathcal{X})$.

Demostración: Fijemos $k \in \mathbb{N}$ y sea $\mathcal{X}_k := \{v_1, \ldots, v_k\}$ subconjunto linealmente independiente de H. Procedamos por inducción sobre k.

Paso Base: Si $k = 1$ entonces $\mathcal{X}_1 = \{v_1\}$ con $v_1 \neq 0_H$. Definiendo $e_1 := \frac{v_1}{\|v\|_1}$ es claro que $\mathcal{O}_1 := \{e_1\}$ es ortonormal y $\mathrm{lin}(\mathcal{O}_1) = \mathrm{lin}(\mathcal{X}_1)$.

Paso Inductivo: Supongamos que para $\mathcal{X}_{k-1} = \{v_1, \ldots, v_{k-1}\}$ existe $\mathcal{O}_{k-1} = \{e_1, \ldots, e_{k-1}\}$ subconjunto ortonormal de H tal que $\mathrm{lin}(\mathcal{O}_{k-1}) = \mathrm{lin}(\mathcal{X}_{k-1})$. Sea $\mathcal{X}_k = \{v_k\} \cup \mathcal{X}_{k-1}$ linealmente independiente.

Definamos $w_k := v_k - \sum_{i=1}^{k-1} \langle v_k, e_j \rangle e_j$. Dado que $\mathcal{X}_k$ es linealmente idependiente entonces $w_k \neq 0_H$. Así pues, definiendo $e_k := \frac{w_k}{\|w_k\|}$ es claro que $\|e_k\| = 1$ y para $j \neq k$ con $j = 1, \ldots, k-1$ se tiene que:

$$\begin{aligned}
\langle e_k, e_j \rangle &= \frac{1}{\|w_k\|} \langle w_k, e_j \rangle = \frac{1}{\|w_k\|} \left\langle v_k - \sum_{i=1}^{k-1} \langle v_i, e_i \rangle e_i, e_j \right\rangle \\
&= \frac{1}{\|w_k\|} \left[\langle v_k, e_j \rangle - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle \langle e_i, e_j \rangle \right] \\
&= \frac{1}{\|w_k\|} \left[\langle v_k, e_j \rangle - \langle v_k, e_j \rangle \langle e_j, e_j \rangle \right] = \frac{1}{\|w_k\|} \left[\langle v_k, e_j \rangle - \langle v_k, e_j \rangle \right] = 0.
\end{aligned}$$

Por tanto, $\mathcal{O}_k = \{e_1, \ldots, e_{k-1}, e_k\}$ es un subconjunto ortonormal de H.

Por otro lado, como por hipótesis $\mathrm{lin}(\mathcal{O}_{k-1}) = \mathrm{lin}(\mathcal{X}_{k-1})$, si $v \in \mathrm{lin}(\mathcal{O}_k)$ entonces $v = \sum_{j=1}^{k} \alpha_j\, e_j$ para $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Así,

$$v = \sum_{j=1}^{k} \alpha_j\, e_j = \sum_{j=1}^{k-1} \alpha_j\, e_j + \alpha_k\, e_k = \sum_{j=1}^{k-1} \alpha_j\, e_j - \frac{\alpha_k}{\|w_k\|} \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j + \frac{\alpha_k}{\|w_k\|}\, v_k.$$

lo que implica que $v \in \mathrm{lin}(\mathcal{B}_k)$. Inversamente, si $u \in \mathrm{lin}(\mathcal{X}_k)$ entonces $u = \sum_{j=1}^{k} \beta_j\, v_j$ para $\beta_1, \ldots, \beta_k \in \mathbb{R}$. Por tanto,

$$\begin{aligned} u &= \sum_{j=1}^{k} \beta_j\, v_j = \sum_{j=1}^{k-1} \beta_j\, v_j + \beta_k\, v_k = \sum_{j=1}^{k-1} \beta_j\, e_j - \beta_k\, w_k + \beta_k \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j \\ &= \sum_{j=1}^{k-1} \beta_j\, e_j - \beta_k \|w_k\|\, e_k + \beta_k \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j, \ \in \mathrm{lin}(\mathcal{O}_k). \end{aligned}$$

Esto concluye la demostración. ■

Teorema 1.46 *Si H es un espacio de Hilbert separable, entonces existe $\mathcal{B}$ una base de Hilbert de H.*

Demostración: Dado que H es separable, existe una sucesión (V_k) creciente de subespacios vectoriales de H tal que $H = \overline{\cup_{k=1}^{\infty} V_k}$ y $\dim(V_k) = k$ para todo $k \in \mathbb{N}$ (estamos suponiendo que $V_k \neq \{0_H\}$ para todo $k \in \mathbb{N}$). En consecuencia, para cada $k \in \mathbb{N}$ existe $\mathcal{B}_k = \{v_1, \ldots, v_k\}$ subconjunto linealmente independiente en H tal que $\mathrm{lin}(\mathcal{B}_k) = V_k$.

Sea $x_1 := v_1 \in \mathcal{B}_1$. Sea $x_2 \in V_2 \setminus \mathrm{lin}(\mathcal{B}_1)$ tal que $\{x_1, x_2\}$ es linealmente independiente. Continuando de este modo obtenemos, para cada $k \in \mathbb{N}$, un $x_k \in V_k \setminus \mathrm{lin}(\mathcal{B}_{k-1})$ tal que $\{x_1, \ldots, x_k\}$ es linealmente independiente. Por tanto, $\mathcal{X} := \{x_k \,:\, k \in \mathbb{N}\}$ es un subconjunto linealmente independiente de H y por el lema de ortonormalización de Gram-Schmidt (ver Lema 1.46) existe $\mathcal{O}$ un subconjunto ortonormal de H tal que $\mathrm{lin}(\mathcal{O}) = \mathrm{lin}(\mathcal{X})$.

Si $v \in \cup_{k=1}^{\infty} V_k$ entonces existe $k_0 \in \mathbb{N}$ tal que $x \in \mathrm{lin}(\mathcal{B}_{k_0}) \subset \mathrm{lin}(\mathcal{X})$. Así pues, $\cup_{k=1}^{\infty} V_k \subset \mathrm{lin}(\mathcal{O})$ y dado que $\cup_{k=1}^{\infty} V_k$ es denso en H concluimos que $\mathrm{lin}(\mathcal{O})$ es denso en H. Por tanto, $\mathcal{O}$ es una base de Hilbert de H. ■

Una forma de caracterizar a una base numerable de Hilbert es la siguiente.

Teorema 1.47 *Sea H un espacio de Hilbert separable y $\mathcal{B} := \{e_k \,:\, k \in \mathbb{N}\}$ un subconjunto ortonormal de H. Las siguientes propiedades son equivalentes:*

(i) *$\mathcal{B}$ es una base de Hilbert.*

(ii) *Para todo $u \in H$ se tiene que:*

$$u = \sum_{k=1}^{\infty} \langle u, e_k \rangle\, e_k. \tag{1.22}$$

(iii) Para todo $u \in H$ se cumple la ***fórmula de Parseval****:*

$$\|u\|^2 = \sum_{k=1}^{\infty} |\langle u, e_k\rangle|^2. \tag{1.23}$$

Demostración: $(i) \Rightarrow (ii)$: Sea $u \in H$ y $\varepsilon > 0$. Dado que $\mathcal{B}$ es base entonces $\mathrm{lin}(\mathcal{B})$ es denso en H. Es decir, existe $v \in \mathrm{lin}(\mathcal{B})$ tal que $\|u - v\| < \varepsilon$.

Por otro lado, la Proposición 1.39 asegura que $p_{\mathrm{lin}(\mathcal{B})}(u) = \sum_{k=1}^{\infty} \langle u, e_k\rangle\, e_k$ de modo que

$$\|u - p_{\mathrm{lin}(\mathcal{B})}(u)\| = \inf_{w \in \mathrm{lin}(\mathcal{B})} \|u - w\|.$$

En consecuencia, $\|u - p_{\mathrm{lin}(\mathcal{B})}(u)\| \leq \|u - v\| < \varepsilon$ y esto prueba el resultado.

$(ii) \Rightarrow (iii)$: Sea $u \in H$. Para cada $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que

$$\left\| u - \sum_{k=1}^{n} \langle u, e_k\rangle\, e_k \right\| < \sqrt{\varepsilon} \qquad \forall\, n \geq k_0.$$

Sea $\mathcal{B}_m := \{e_k \,:\, 1 \leq k \leq m\}$ para cada $m \in \mathbb{N}$. Es claro que $\mathrm{lin}(\mathcal{B}_m) \subset \mathrm{lin}(\mathcal{B}_{m+1}) \subset \mathrm{lin}(\mathcal{B})$ para cada $m \in \mathbb{N}$. Así pues, sea $n \geq k_0$ entonces $v_n := \sum_{k=1}^{n} \langle u, e_k\rangle\, e_k \in \mathrm{lin}(\mathcal{B}_n)$. Aplicando el Corolario 1.40 concluimos que

$$\|u\|^2 - \sum_{k=1}^{n} |\langle u, e_k\rangle|^2 = \|u - v_n\|^2 < \varepsilon \qquad \forall\, n \geq k_0.$$

$(iii) \Rightarrow (i)$: Sea $u \in H$ y $\varepsilon > 0$. Por hipótesis existe $k_0 \in \mathbb{N}$ tal que

$$\|u\|^2 - \sum_{k=1}^{n} |\langle u, e_k\rangle|^2 < \sqrt{\varepsilon} \qquad \forall\, n \geq k_0.$$

Definiendo $v_n := \sum_{k=1}^{n} \langle u, e_k\rangle\, e_k \in \mathrm{lin}(\mathcal{B}_n)$ es claro que $v_n \in \mathrm{lin}(\mathcal{B})$ y se tiene que:

$$\|u - v_n\|^2 = \|u\|^2 - \sum_{k=1}^{n} \langle u, e_k\rangle\, e_k < \sqrt{\varepsilon} \qquad \forall\, n \geq k_0.$$

Por tanto, $\mathrm{lin}(\mathcal{B})$ es denso en H lo que implica que $\mathcal{B}$ es una base de Hilbert. ∎

Ahora probaremos que todo espacio de Hilbert separable es isométricamente isomorfo al espacio $\ell_2(\mathbb{R})$.

Teorema 1.48 *Sea H un espacio de Hilbert separable y $\mathcal{B} := \{e_k : k \in \mathbb{N}\}$ una base de Hilbert de H. La función $\Theta : H \to \ell_2(\mathbb{R})$ dada por:*

$$\Theta(u) := (\langle u, e_k \rangle) \tag{1.24}$$

es un isomorfismo. Más aún, $\|\Theta(u)\|_{\ell_2(\mathbb{R})} = \|u\|$ para todo $u \in H$.

Demostración: Es inmediato de las propiedades del producto escalar que Θ es lineal. Por la fórmula de Parseval (ver Teorema 1.47) tenemos que:

$$\|\Theta(u)\|_{\ell_2(\mathbb{R})} = \left(\sum_{k=1}^{\infty} |\langle u, e_k \rangle| \right)^{1/2} = \|u\| \qquad \forall\, u \in H$$

lo que implica que Θ es continua en H y una isometría.

Dado que Θ es una isometría, entonces $\Theta(H)$ es un subespacio cerrado de $\ell_2(\mathbb{R})$. Sea $\bar{\delta}_k := (\delta_{k,j})$ la sucesión en $\ell_2(\mathbb{R})$ cuyos términos son $\delta_{k,k} = 1$ y $\delta_{k,j} = 0$ si $k \neq j$. Entonces, $\Theta(e_j) = (\langle e_j, e_k \rangle) = \bar{\delta}_j$ para todo $j \in \mathbb{N}$. Sea $\bar{x} \in \ell_2(\mathbb{R})$. El Ejemplo 1.44 asegura que para $\varepsilon > 0$ dada existe $k_0 \in \mathbb{N}$ tal que $\|\bar{x} - \bar{\delta}_{k_0}\| = \|\bar{x} - \Theta(e_{k_0})\| < \varepsilon$. Es decir, $\Theta(H)$ es subespacio cerrado denso de $\ell_2(\mathbb{R})$ por lo que necesariamente $\Theta(H) = \ell_2(\mathbb{R})$. Esto prueba que Θ es suprayectiva y, por tanto, un isomorfismo pues es además una isometría.

Finalmente, para cualesquiera $i, j \in \mathbb{N}$ se tiene que:

$$\langle \Theta(e_j), \Theta(e_i) \rangle_{\ell_2(\mathbb{R})} = \langle \bar{\delta}_j, \bar{\delta}_i \rangle_{\ell_2(\mathbb{R})} = \begin{cases} 0 & \text{si } j \neq 1, \\ 1 & \text{si } j = 1 \end{cases} = \langle e_j, e_i \rangle.$$

.

Dado que Θ es lineal, entonces $\langle \Theta(u), \Theta(v) \rangle$ para cualesquiera $u, v \in \text{lin}(\mathcal{B})$. Y, como Θ es continua entonces $\langle \Theta(u), \Theta(v) \rangle$ para cualesquiera $u, v \in \overline{\text{lin}(\mathcal{B})} = H$. Es decir, la función Θ preserva el producto escalar. ∎

Notemos que, si H es un espacio de Hilbert separable y $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ es una base de Hilbert de H, entonces para cualesquiera $u, v \in H$ el teorema anterior asegura que:

$$\langle u, v \rangle = \langle \Theta(u), \Theta(v) \rangle_{\ell_2(\mathbb{R})} = \sum_{k=1}^{\infty} \langle u, e_k \rangle \langle v, e_k \rangle. \tag{1.25}$$

Finalizamos esta sección con una definición importante que se involucra en otras ramas del Análisis Matemático tales como el Análisis de Fourier.

Definición 1.49 *Sea H un espacio de Hilbert separable y $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ es una base de Hilbert de H. Para cada $u \in H$, la serie:*

$$u = \sum_{k=1}^{\infty} \langle u, e_k \rangle\, e_k \tag{1.26}$$

*se llama **serie de Fourier de** u **relativa a** $\mathcal{B}$. Los elementos de la sucesión $(\langle u, e_k \rangle)$ en $\ell_2(\mathbb{R})$ se llaman **coeficientes de Fourier de** u **relativos a** $\mathcal{B}$.*

Finalizamos este capítulo con un resultado análogo al teorema de representación de Fréchet-Riesz para una forma bilineal $a : H \times H \to \mathbb{R}$ continua.

Definición 1.50 *Sea V un espacio vectorial sobre $\mathbb{R}$. Una función $a : V \times V \to \mathbb{R}$ se llama* ***forma bilineal*** *si verifica las siguientes propiedades:*

(a) $a(v_1 + v_2, w) = a(v_1, w) + a(v_2, w)$ *para cualesquiera* $v_1, v_2, w \in V$.

(b) $a(v, w_1 + w_2) = a(v, w_1) + a(v, w_2)$ *para cualesquiera* $v, w_1, w_2 \in V$.

(c) $a(\lambda v, w)$ *para cualesquiera* $v, w \in V$ *y* $\lambda \in \mathbb{R}$.

(d) $a(v, \mu\, w)$ *para cualesquiera* $v, w \in V$ *y* $\mu \in \mathbb{R}$.

Decimos que a es ***continua*** *si existe $c \in \mathbb{R}$ tal que $|a(v, w)| \leq c\|v\|\,\|w\|$ para cualesquiera $v, w \in V$.*

Por ejemplo, un producto escalar $\langle \cdot, \cdot \rangle$ es una forma bilineal.

Teorema 1.51 (de representación de Fréchet-Riesz para formas bilineales) *Sea $a : H \times H \to \mathbb{R}$ una forma bilineal continua. Entonces, existe una única función $\Psi : H \to H$ lineal y continua tal que:*

$$a(u, v) = \langle \Psi(u), v \rangle \qquad \forall\, u, v \in H. \tag{1.27}$$

Demostración: Fijemos $u \in H$. Definamos la función $\widetilde{a} : H \to \mathbb{R}$ como $\widetilde{a}(v) := a(u, v)$ la cual es claramente lineal y continua. El Teorema 1.33 asegura que existe un único $w \in H$ (que depende de u) tal que:

$$\widetilde{a}(v) = \langle w, v \rangle \qquad \forall\, v \in H. \tag{1.28}$$

Definiendo $\Psi : H \to H$ como $\Psi(u) := w$ se tiene de (1.28) que:

$$\widetilde{a}(v) = a(u, v) = \langle \Psi(u), v \rangle \qquad \forall\, u, v \in H.$$

Además, para cualesquiera $u_1, u_2 \in H$ y $\lambda, \mu \in \mathbb{R}$ se tiene que:

$$\begin{aligned}
\langle \Psi(\lambda u_1 + \mu u_2), v \rangle &= a(\lambda u_1 + \mu u_2, v) \\
&= \lambda\, a(u_1, v) + \mu\, a(u_2, v) \\
&= \lambda \langle \Psi(u_1), v \rangle + \mu \langle \Psi(u_2), v \rangle \\
&= \langle \lambda \Psi(u_1) + \mu \Psi(u_2), v \rangle
\end{aligned}$$

para todo $v \in H$. En consecuencia, $\Psi(\lambda u_1 + \mu u_2) = \lambda \Psi(u_1) + \mu \Psi(u_2)$ lo que implica que Ψ es lineal. Aplicando la continuidad de $a : H \times H \to \mathbb{R}$ se tiene que

$$\|\Psi(u)\|_H^2 = |\langle \Psi(u), \Psi(u) \rangle| = |a(u, \Psi(u))| \leq c\, \|u\|_H\, \|\Psi(u)\|_H \qquad \forall\, u \in H$$

lo que implica que Ψ es continua en H.

Supongamos que existe $\widetilde{\Psi} : H \to H$ tal que cumple (1.27). Entonces, $\langle \Psi(u) - \widetilde{\Psi}(u), v \rangle = 0$ para todo $v \in H$ y, por tanto, $\Psi(u) - \widetilde{\Psi}(u) = 0_H$ para todo $u \in H$ lo que establece la unicidad.

Esto concluye la demostración. $\blacksquare$

Capítulo 2

Mensurabilidad en espacios de Hilbert

Sea $H = (H, \langle \cdot, \cdot \rangle_H, \| \cdot \|_H)$ un espacio de Hilbert sobre $\mathbb{R}$. La norma de este espacio induce una topología a través de los subconjuntos abiertos y, por tanto, es posible definir la σ-álgebra de Borel en H denotada por $\mathcal{B}(H)$. De este modo, consideramos a $(H, \mathcal{B}(H))$ un espacio medible. Así pues, si $(\Omega, \mathcal{S}, \mu)$ es un espacio de medida arbitrario nos es posible considerar funciones de la forma $f : \Omega \to H$. En este capítulo introduciremos nuevas nociones de mensurabilidad de una función llamadas *mensurabilidad fuerte* y *mensurabilidad débil*, que, a diferencia de la mensurabilidad usual, tienen la propiedad de trabajar con criterios de convergencia de sucesiones de funciones. La mensurabilidad fuerte se abarcará desde dos criterios, con respecto de un espacio medible y con respecto de un espacio de medida. Estudiaremos, a través de los teoremas de mensurabilidad de Pettis, los criterios bajo los cuales estos dos conceptos son equivalentes.

El concepto de mensurabilidad fuerte con respecto de un espacio de medida es muy útil pues nos permitirá extender el concepto de integral para funciones que toman valores en un espacio de Banach. En 1933, el matemático Salomon Bochner introdujo un concepto de integral para funciones definidas en un espacio de medida y con valores en un espacio normado. Este nuevo concepto de integral lleva por nombre integral de Bochner. La *integral de Bochner* es una generalización de la integral de Lebesgue, estudiada en el capítulo anterior, en el contexto de los espacios de normados y en este trabajo en particular en espacios de Banach. Algunas propiedades de la integral de Lebesgue tales como la linealidad se extienden naturalmente a la integral de Bochner al ser definida en un espacio de Banach que tiene la estructura algebraica de un espacio vectorial. El punto de partida para definir la integral de Bochner de una función fuertemente medible será análogo al caso de la integral de Lebesgue, comenzando en aquellas funciones que toman solamente una cantidad finita de elementos en un espacio de Hilbert H. El hecho de que H sea de Banach permitirá usar el criterio de convergencia de Cauchy para considerar el límite de integrales de estas funciones.

El criterio de integrabilidad de Lebesgue-Bochner nos permitirá caracterizar aquellas funciones que son Bochner-integrables a través del estudio de sus funciones normas y que se reducen a los criterios establecidos con respecto de las funciones Lebesgue-integrables. Este nos permitirá además dar un cota, en norma, del valor de la integral de Bochner que en muchos casos resulta sumamente complejo calcular.

Finalmente, extendemos el concepto de los espacios de Lebesgue ahora para clases de equivalencia de funciones $f : \Omega \to H$ que resultan ser fuertemente medibles, a tales espacios los llamaremos *espacios de Lebesgue-Bochner*. Demostraremos que los espacios de Lebesgue-Bochner $L^p(\Omega; H)$ son de Banach. Estos conceptos nos permitirán dar una noción de esperanza para variables aleatorias con valores en un espacio de Banach como veremos en próximos capítulos.

La teoría dada en este capítulo está basada principalmente en [1], [2], [6], [24], [25], [50], [79] y [87].

2.1. Funciones medibles en espacios de Hilbert

Sea $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ y consideremos a su espacio dual $H^* := \{\psi : H \to \mathbb{R} \,:\, \psi \text{ es lineal y continua}\}$.

2.1.1. Mensurabilidad fuerte

Sea $(\Omega, \mathcal{S})$ un espacio medible. Dado que H es un espacio normado, entonces es posible considerar la σ-álgebra de Borel $\mathcal{B}(H)$ de H. En este capítulo diremos que una función $\phi : \Omega \to H$ es **$\mathcal{S}$-medible** si $\phi^{-1}(H) \in \mathcal{S}$ para todo $B \in \mathcal{B}(H)$.

Para no generar confusión, a las funciones $\varphi : \Omega \to \mathbb{R}$ medibles relativas a $\mathcal{S}$ y $\mathcal{B}(\mathbb{R})$ las llamaremos simplemente **funciones medibles**.

Observemos que, si $\phi : H \to \mathbb{R}$ es continua entonces $\phi^{-1}(I) \in \mathcal{B}(H)$ para todo subconjunto abierto I de $\mathbb{R}$. Por tanto, $\phi : H \to \mathbb{R}$ es medible.

Además, si $\varphi : \Omega \to H$ es $\mathcal{S}$-medible y $\phi : H \to \mathbb{R}$ es medible, entonces $\phi \circ \varphi : \Omega \to \mathbb{R}$ es medible. En efecto, si $I \in \mathcal{B}(\mathbb{R})$ entonces:

$$(\phi \circ \varphi)^{-1}(I) = \varphi^{-1}(\phi^{-1}(I)) \in \mathcal{S}$$

pues $\phi^{-1}(I) \in \mathcal{B}(H)$.

En esta sección damos distintas definiciones con respecto de la medibilidad de una función $f : \Omega \to H$ en un espacio medible y estudiamos las relaciones que existen entre ellas.

Definición 2.1 *Una función $s : \Omega \to H$ es **$\mathcal{S}$-simple** si existen elementos distintos $v_1, \ldots, v_N \in H$, los subconjuntos disjuntos a pares $A_1 := s^{-1}(\{v_1\}), \ldots, A_N := s^{-1}(\{v_N\})$*

pertenecen a $\mathcal{S}$ y son tales que:

$$s(x) := \sum_{j=1}^{N} v_j \, 1_{A_j}(x) \tag{2.1}$$

en donde 1_{A_j} denota la función indicadora del subconjunto A_j.

Definición 2.2 *Una función $f : \Omega \to H$ es **fuertemente $\mathcal{S}$-medible** si, existe una sucesión de funciones $\mathcal{S}$-simples $s_k : \Omega \to H$ tal que (s_k) converge puntualmente a f en Ω.*

Definición 2.3 *Una función $f : \Omega \to H$ es **débilmente $\mathcal{S}$-medible** si, para cada $\psi \in H^*$, la función $\psi \circ f : \Omega \to \mathbb{R}$ es medible.*

El siguiente resultado, debido a James Pettis, brinda algunas condiciones para relacionar la $\mathcal{S}$-medibilidad fuerte y débil. Antes de enunciarlo requerimos de la siguiente definición.

Definición 2.4 *Sea S un conjunto no vacío. Una función $f : S \to H$ es **separablemente valuada** si, existe un subespacio vectorial cerrado y separable W de H tal que $f(z) \in W$ para cada $z \in S$.*

Sea $V = (V, \|\cdot\|_V)$ un espacio de Banach sobre $\mathbb{R}$. El teorema de extensión de Hahn-Banach permite concluir que, para cada $v \in V$

$$\|v\|_V = \sup\left\{|T(v)| \,:\, T \in V^* \;\text{ y }\; \|T\|_{V^*} \leq 1\right\}. \tag{2.2}$$

Si U un subconjunto no vacío de V y Θ un subespacio vectorial de V^*, decimos que Θ es **normado para** U si, para cada $u \in U$ se tiene que:

$$\|u\|_V = \sup\left\{|T(u)| \,:\, T \in \Theta \;\text{ y }\; \|T\|_{V^*} \leq 1\right\}. \tag{2.3}$$

En el caso en que Θ sea normado para V diremos simplemente que Θ es **normado**. De este modo, si W es un subespacio vectorial separable de V y Θ es normado para W, entonces existe una sucesión (ϑ_k) de elementos de Θ tal que $\{\vartheta_k \,:\, k \in \mathbb{N}\}$ es normado para W y $\|\vartheta_k\|_{V^*} = 1$ para cada $k \in \mathbb{N}$ (ver [87]).

Teorema 2.5 (de mensurabilidad de Pettis, primera versión) *Sean $(\Omega, \mathcal{S})$ un espacio medible, Θ un subespacio vectorial de H^* que es normado para H y $f : \Omega \to H$ una función. Las siguientes afirmaciones son equivalentes:*

(i) *$f : \Omega \to H$ es fuertemente $\mathcal{S}$-medible.*

(ii) *$f : \Omega \to H$ es separablemente valuada y débilmente $\mathcal{S}$-medible.*

(iii) $f : \Omega \to H$ *es separablemente valuada y* $\vartheta \circ f : \Omega \to \mathbb{R}$ *es medible para cada* $\vartheta \in \Theta$.

Demostración: $(i) \Rightarrow (ii)$: Si $f : \Omega \to H$ es fuertemente $\mathcal{S}$-medible, entonces existe una sucesión (s_k) de funciones $\mathcal{S}$-simples tal que $s_k \to f$ puntualmente en Ω.

Sea $\psi \in H^*$. Es claro que $\psi : H \to \mathbb{R}$ es medible pues es continua. Además, $\psi(s_k) \to \psi(f)$ puntualmente en Ω en donde $\psi \circ s_k : \Omega \to \mathbb{R}$ es medible para todo $k \in \mathbb{N}$. Por tanto, $\psi \circ f : \Omega \to \mathbb{R}$ es medible pues es límite puntual de una sucesión de funciones medibles.

En consecuencia, f es débilmente $\mathcal{S}$-medible.

Es claro por definición de función $\mathcal{S}$-simple que $s_k(\Omega)$ es finito para cada $k \in \mathbb{N}$. Así pues, $\mathfrak{C} = \cup_{k=1}^{\infty} s_k(\Omega)$ es, a lo más, numerable. Definiendo $W := \overline{\text{lin}(\mathfrak{C})}$ es claro que W es un subespacio vectorial cerrado y separable de H. Además, como $s_k(x) \to f(x)$ para cada $x \in \Omega$, entonces $f(x) \in \overline{\mathfrak{C}} \subset W$ para cada $x \in \Omega$ lo que prueba que f es separablemente valuada.

$(ii) \Rightarrow (iii)$: Es inmediato del hecho que Θ es subespacio vectorial de H^*.

$(iii) \Rightarrow (i)$: Dado que f es separablemente valuada entonces existe un subespacio vectorial cerrado y separable W de H tal que $f(x) \in W$ para cada $x \in \Omega$. Es claro que Θ es normado para W de modo que existe una sucesión (ϑ_k) en Θ tal que $\|\vartheta_k\|_{H^*} = 1$ y tal que es normada para W. Es decir,

$$\|f(x)\|_H = \sup_{k \in \mathbb{N}} |\vartheta_k(f(x))| \qquad \forall\, x \in \Omega.$$

Para cada $w \in W$ arbitrario pero fijo, definamos $\phi_w : \Omega \to \mathbb{R}$ como:

$$\phi_w(x) := \|f(x) - w\|_H = \sup_{k \in \mathbb{N}} |\vartheta_k(f(x) - w)|. \tag{2.4}$$

Por hipótesis, concluimos que ϕ_w es una función medible.

Sea $\{w_k \;:\; k \in \mathbb{N}\}$ un subconjunto denso de W. Para cada $w \in W$ y $k \in \mathbb{N}$, sea $k^* = k^*(k, x)$ el menor entero en $\{1, \ldots, k\}$ tal que:

$$\|w - w_{k^*}\|_H = \min_{1 \le j \le k} \|w - w_j\|_H.$$

Definamos la sucesión de funciones $\zeta_k : H \to \{w_1, \ldots, w_k\}$ como:

$$\zeta_k(v) := \begin{cases} w_{k^*} & \text{si } v \in W, \\ 0 & \text{si } v \notin W. \end{cases}$$

Es inmediato de la densidad de $\{w_k \;:\; k \in \mathbb{N}\}$ en W que $\zeta_k(w) \to w$ para cada $w \in W$. Definimos, por tanto, la sucesión de funciones $s_k : \Omega \to H$ como:

$$s_k(x) := (\zeta_k \circ f)(x).$$

Es claro que s_k toma sólo una cantidad finita de valores en H para cada $k \in \mathbb{N}$ y $s_k(x) = \zeta_k(f(x)) \to f(x)$ para cada $x \in \Omega$. Así pues, probemos que s_k es $\mathcal{S}$-medible.

Fijemos $k \in \mathbb{N}$. Entonces, para cada $1 \leq \ell \leq k$ se tiene que:

$$\begin{aligned}\{x \in \Omega \,:\, s_k(x) = w_\ell\} &= \{x \in \Omega \,:\, \|f(x) - w_\ell\|_H = \min_{1 \leq j \leq k} \|f(x) - w_j\|_H\} \cap \\ &\bigcap_{m=1}^{\ell-1} \{x \in \Omega \,:\, \|f(x) - w_m\|_H > \|f(x) - w_\ell\|_H\} \\ &= \{x \in \Omega \,:\, \phi_{w_\ell}(x) = \min_{1 \leq j \leq k} \|f(x) - w_j\|_H\} \cap \\ &\bigcap_{m=1}^{\ell-1} \{x \in \Omega \,:\, \phi_{w_m}(x) > \phi_{w_\ell}(x)\} \in \mathcal{S}\end{aligned}$$

pues ϕ_{w_m} es medible para cada $m = 1, \ldots, \ell$.

Por tanto, f es fuertemente $\mathcal{S}$-medible. ■

Observemos que en la demostración del Teorema 2.5 probamos que la imagen de una función $f : \Omega \to H$ que es límite puntual de una sucesión de funciones $\mathcal{S}$-simples queda contenida en un subespacio vectorial cerrado y separable de H (ver [2]).

Corolario 2.6 *Si $f_k : \Omega \to H$ es fuertemente $\mathcal{S}$-medible para todo $k \in \mathbb{N}$ y (f_k) converge puntualmente a f en Ω, entonces $f : \Omega \to H$ es fuertemente $\mathcal{S}$-medible.*

Demostración: El Teorema 2.5 asegura que f_k es separablemente valuada para todo $k \in \mathbb{N}$. Así pues, para todo $k \in \mathbb{N}$, existe W_k subconjunto numerable de H tal que $f_k(x) \in \overline{\text{lin}(W_k)}$. Definiendo $W := \cup_{k=1}^{\infty} W_k$ es claro que es un subconjunto numerable de V y que $f(x) \in \overline{\overline{W}}$ para todo $x \in \Omega$ pues $f_k \to f$ puntualmente en Ω. El conjunto $\overline{\text{lin}(W)}$ es un subespacio vectorial cerrado y separable de H pues W es total en él y $f(x) \in \overline{\overline{W}} \subset \overline{\text{lin}(W)}$ para todo $x \in \Omega$. Es decir, f es separablmente valuada.

Por otra parte, sea $\psi \in H^*$. Dado que f_k es débilmente $\mathcal{S}$-medible entonces $\psi \circ f_k : H \to \mathbb{R}$ es medible para todo $k \in \mathbb{N}$. Como $f_k(x) \to f(x)$ para todo $x \in \Omega$ y $\psi : H \to \mathbb{R}$ es continua, entonces $\psi(f_k(x)) \to \psi(f(x))$ para todo $x \in \Omega$. Así pues, $\psi \circ f : \Omega \to \mathbb{R}$ es medible por ser el límite puntual de funciones medibles. En consecuencia, f es débilmente $\mathcal{S}$-medible.

Por tanto, f es fuertemente $\mathcal{S}$-medible. ■

Corolario 2.7 *Sean H y K dos espacios de Hilbert, $f : \Omega \to H$ una función fuertemente $\mathcal{S}$-medible y $\phi : H \to K$ una función continua. Entonces, $\phi \circ f : \Omega \to K$ es fuertemente $\mathcal{S}$-medible*

Demostración: Sea $s_k : \Omega \to H$ sucesión de funciones $\mathcal{S}$-simples tales que $s_k \to f$ puntualmente en Ω. Como $\phi : H \to K$ es continua entonces $\phi \circ s_k \to \phi \circ f$ puntualmente en Ω. El Corolario 2.6 asegura que $\phi \circ f$ es fuertemente $\mathcal{S}$-medible. ■

Lema 2.8 *Si $f_k : \Omega \to H$ es $\mathcal{S}$-medible para todo $k \in \mathbb{N}$ y (f_k) converge puntualmente a f en Ω, entonces $f : \Omega \to H$ es $\mathcal{S}$-medible.*

Demostración: Sea K un subconjunto cerrado no vacío de H y consideremos $\rho(v) := \inf_{z\in K} \|v - z\|_H$ para cada $v \in H$.

Definamos, para cada $j \in \mathbb{N}$, el subconjunto $U_j := \{v \in H : \rho(v) < \frac{1}{j}\}$. Observemos que, U_j es abierto en H para todo $j \in \mathbb{N}$. Además, como K es cerrado en H, entonces $K = \{v \in H : \rho(v) = 0\} = \cap_{j=1}^{\infty} U_j$.

Supongamos que $x \in f^{-1}(K)$, es decir, $f(x) \in K$. Como $f_k(x) \to f(x)$, entonces para cada $j \in \mathbb{N}$, existe $\ell = \ell(x, j) \in \mathbb{N}$ tal que:

$$\|f_k(x) - f(x)\|_H < \tfrac{1}{j} \qquad \forall\, k \geq \ell.$$

Dado que $f(x) \in K$, entonces:

$$\inf_{z\in K} \|f_k(x) - z\|_H \leq \|f_k(x) - f(x)\|_H < \tfrac{1}{j} \qquad \forall\, k \geq \ell.$$

En consecuencia, $x \in \cap_{j=1}^{\infty} \cup_{\ell=1}^{\infty} \cap_{k=\ell}^{\infty} f_k^{-1}(U_j)$.

Inversamente: si $x \in \cap_{j=1}^{\infty} \cup_{\ell=1}^{\infty} \cap_{k=\ell}^{\infty} f_k^{-1}(U_j)$, entonces para cada $j \in \mathbb{N}$ dada, existe $\ell \in \mathbb{N}$ tal que $f_k(x) \in U_j$ para todo $k \geq \ell$. Así pues, $\lim_{k\to\infty} f_k(x) = f(x) \in \overline{U_j}$ para todo $j \in \mathbb{N}$, es decir, $f(x) \in \cap_{k=1}^{\infty} \overline{U_j}$. Observemos que, $\overline{U_{j+1}} \subset U_j$ para todo $j \in \mathbb{N}$. Entonces, $f(x) \in \cap_{j=1}^{\infty} U_j = K$ lo que implica que $x \in f^{-1}(K)$.

Por tanto,

$$f^{-1}(K) = \bigcap_{j=1}^{\infty} \bigcup_{\ell=1}^{\infty} \bigcap_{k=\ell}^{\infty} f_k^{-1}(U_j) \in \mathcal{S}$$

lo que prueba que f es $\mathcal{S}$-medible. ■

Observemos que en la demostración del Lema 2.8 usamos únicamente propiedades de la distancia dada por la norma en V. Así, si en la hipótesis de este consideramos una sucesión de funciones $\mathcal{S}$-medibles $f_k : \Omega \to X$ en donde X es un espacio métrico y $\mathcal{B}(X)$ es la σ-álgebra de Borel en él, el resultado sigue siendo válido. Invitamos al lector a consultar, por ejemplo, [2] para una prueba detallada de esta afirmación.

Lema 2.9 *Supongamos que H es un espacio de Hilbert separable y totalmente acotado. Entonces, $f : \Omega \to H$ es $\mathcal{S}$-medible si y sólo si es el límite uniforme de una sucesión de funciones $\mathcal{S}$-simples.*

Demostración: $\Leftarrow)$: Es el Lema 2.8 pues si (f_k) converge uniformemente a f, entonces (f_k) converge puntualmente a f.

$\Rightarrow$) : Supongamos que $f : \Omega \to H$ es $\mathcal{S}$-medible y sea $\varepsilon > 0$ dada. Como H es totalmente acotado, existen $v_1, \ldots, v_N \in H$ (que dependen de ε) tales que:

$$H = B_H(v_1, \varepsilon) \cup \cdots \cup B_H(v_N, \varepsilon).$$

Definamos $A_1 := B_H(v_1, \varepsilon)$ y $A_j := B_H(v_j, \varepsilon) \setminus \cup_{k=1}^{j-1} B_H(v_k, \varepsilon)$ para todo $1 < j \leq N$. Es claro que $A_j \in \mathcal{B}(H)$ para todo $j = 1, \ldots, N$ y son mutumaente disjuntos tales que $\cup_{j=1}^{N} A_j = H$. Así pues, $\Omega = \cup_{j=1}^{N} f^{-1}(A_j)$ en donde $f^{-1}(A_j) \in \mathcal{S}$ para todo $j = 1, \ldots, N$ y son mutuamente disjuntos.

Definiendo $s_\varepsilon : \Omega \to \mathbb{R}$ como:

$$s_\varepsilon(x) := \sum_{j=1}^{N} v_j \, 1_{f^{-1}(A_j)}(x),$$

es claro que s_ε es una función $\mathcal{S}$-simple. Además, para $x \in \Omega$, existe $j_0 \in \{1, \ldots, N\}$ tal que $x \in f^{-1}(A_{j_0})$ y, por tanto,

$$\|s_\varepsilon(x) - f(x)\|_H = \|v_{j_0} - f(x)\|_H < \varepsilon.$$

Esto prueba que f puede verse como el límite uniforme de funciones $\mathcal{S}$-simples. ■

Observación 2.10 *Sea $H = (H, \|\cdot\|_H)$ un espacio de Banach separable y sea $f : \Omega \to H$ función $\mathcal{S}$-medible. El* Teorema de metrización de Uryshon *asegura que existe una norma $\|\cdot\|_H^*$ equivalente a $\|\cdot\|_H$ y tal que $(H, \|\cdot\|_H^*)$ es separable y totalmente acotado. Así, por el* Lema 2.9 *existe una sucesión (f_k) de funciones $\mathcal{S}$-simples tales que:*

$$\lim_{k\to\infty} \|f_k(x) - f(x)\|_H^* = 0 \quad \forall\, x \in \Omega.$$

Existe además $\kappa > 0$ tal que $\kappa\|v\|_H \leq \|v\|_H^*$ para todo $v \in H$. Por tanto,

$$\|f_k(x) - f(x)\|_H \leq \frac{1}{\kappa} \|f_k(x) - f(x)\|_H^* \quad \forall\, x \in \Omega \quad \forall\, k \in \mathbb{N},$$

y esto prueba que $f_k(x) \to f(x)$ en $(H, \|\cdot\|_H)$ para todo $x \in \Omega$. Es decir, f es $\mathcal{S}$-medible si y sólo si es el límite puntual de una sucesión de funciones $\mathcal{S}$-simples.

Teorema 2.11 *Sea $f : \Omega \to H$ una función. Las siguientes afirmaciones son equivalentes:*

(i) f es fuertemente $\mathcal{S}$-medible.

(ii) f es separablemente valuada y $\mathcal{S}$-medible.

Demostración: $(i) \Rightarrow (ii)$: El Teorema 2.5 asegura que f es separablemente valuada. Sea U un subconjunto abierto de H y (s_k) sucesión de funciones $\mathcal{S}$-simples tal que $s_k \to f$ puntualmente en Ω. Los Lemas 2.8 y 2.9 junto con la Observación 2.10 implican que f es $\mathcal{S}$-medible.

$(ii) \Rightarrow (i)$: Por hipótesis f es separablemente valuada. Ahora bien, como $f : \Omega \to H$ es $\mathcal{S}$-medible, entonces $\psi \circ f : \Omega \to \mathbb{R}$ es medible para toda $\psi \in H^*$. Es decir, f es débilmente $\mathcal{S}$-medible. El Teorema 2.5 asegura que f es fuertemente $\mathcal{S}$-medible. ■

Observación 2.12 *Si H es un espacio de Hilbert separable, entonces toda función $f : \Omega \to H$ es separablemente valuada. Así pues, el* Teorema 2.8 *implica que, f es $\mathcal{S}$-medible si y sólo si f es fuertemente $\mathcal{S}$-medible.*

2.1.2. Mensurabilidad fuerte con respecto de una medida

En esta sección extendemos los conceptos de la mensurabilidad de una función $f : \Omega \to H$ con respecto de un espacio de medida y estudiamos las relaciones que existen entre estos nuevos conceptos y los dados en el apartado anterior.

Sea $(\Omega, \mathcal{S}, \mu)$ un espacio de medida.

Definición 2.13 *Una función $s : \Omega \to H$ es **μ-simple** si existen elementos distintos $v_1, \ldots, v_N \in H$, los subconjuntos $A_1 := s^{-1}(\{v_1\}), \ldots, A_N := s^{-1}(\{v_N\})$ disjuntos a pares pertenecen a $\mathcal{S}$, son todos de medida finita y son tales que:*

$$s(x) := \sum_{j=1}^{N} v_j \, 1_{A_j}(x) \tag{2.5}$$

en donde 1_{A_j} denota la función indicadora del subconjunto A_j.

Notemos que esta noción es más fuerte que la de función $\mathcal{S}$-simple, es decir, si $s : \Omega \to H$ es una función μ-simple, entonces $s : \Omega \to H$ es $\mathcal{S}$-simple. El recíproco no es cierto en general, a menos que el espacio sea de medida finita.

Definición 2.14 *Una función $f : \Omega \to H$ es **fuertemente μ-medible** si, existe una sucesión de funciones μ-simples $s_k : \Omega \to H$ y existe un subconjunto $N \in \mathcal{S}$ con $\mu(N) = 0$ tal que (s_k) converge puntualmente a f en $\Omega \setminus N$.*

Definición 2.15 *Una función $f : \Omega \to H$ es **μ-separablemente valuada** si, existen un subespacio vectorial cerrado y separable W de H y $N \in \mathcal{S}$ con $\mu(N) = 0$ tales que $f(x) \in W$ para cada $x \in \Omega \setminus N$.*

Teorema 2.16 (de Mensurabilidad de Pettis, segunda versión) *Sean* $(\Omega, \mathcal{S}, \mu)$ *un espacio de medida* σ*-finita y* $f : \Omega \to H$ *una función. Las siguientes afirmaciones son equivalentes:*

(i) $f : \Omega \to H$ *es fuertemente* μ*-medible.*

(ii) $f : \Omega \to H$ *es* μ*-separablemente valuada y débilmente* $\mathcal{S}$*-medible.*

Demostración: La demostración es, en esencia, similar a la dada en el Teorema 2.5. Invitamos al lector a consultar, por ejemplo, [50] para los detalles. ■

Si el espacio de Hilbert H es separable, el Teorema 2.16 implica que una función $f : \Omega \to H$ es fuertemente μ-medible si y sólo si f es débilmente $\mathcal{S}$-medible.

Definición 2.17 *Para* $f : \Omega \to H$ *dada, definimos la* ***función norma de*** f*, denotada por* $\|f\|_H : \Omega \to \mathbb{R}$*, como:*

$$\|f\|_H(x) := (\|\cdot\|_H \circ f)(x) = \|f(x)\|_H.$$

Proposición 2.18 *Si* $f : \Omega \to H$ *es fuertemente* μ*-medible, entonces la función norma de* f *es fuertemente* μ*-medible.*

Demostración: Dado que f es fuertemente μ-medible, entonces existen una sucesión (s_k) de funciones μ-simples y $N \in \mathcal{S}$ con $\mu(N)$ tales que:

$$\lim_{k \to \infty} \|s_k(x) - f(x)\|_H = 0 \quad \forall\, x \in \Omega \setminus N.$$

Aplicando propiedades de norma obtenemos que:

$$|\|s_k(x)\|_H - \|f(x)\|_H| \leq \|s_k(x) - f(x)\|_H \quad \forall\, x \in \Omega \setminus N.$$

Tomando el limite cuando $k \to \infty$ concluimos que $\|s_k(x)\|_H \to \|f(x)\|_H$ para todo $x \in \Omega \setminus N$. ■

En lo que resta examinaremos, bajo que condiciones, las nociones de medibilidad fuerte de una función dadas en la Definición 2.2 y Definición 2.14 son equivalentes.

Teorema 2.19 (Mensurabilidad de Pettis) *Sea* $(\Omega, \mathcal{S}, \mu)$ *espacio de medida* σ*-finita y* $f : \Omega \to H$ *una función. Entonces,* f *es fuertemente* μ*-medible si y sólo si*

(a) f *es* $\mathcal{S}$*-medible.*

(b) f *es* μ*-separablemente valuada.*

Demostración: $\Rightarrow$) : Supongamos que f es fuertemente μ-medible. Existe una sucesión (s_k) de funciones μ-simples y $N \in \mathcal{S}$ con $\mu(N)$ tales que:

$$s_k(x) \to f(x) \qquad \forall\, x \in \Omega \setminus N.$$

Así pues, existe un subespacio vectorial cerrado y separable W de H tal que (ver Teorema 2.5) $f(\Omega \setminus N) = \{f(x) \in H \,:\, x \in \Omega \setminus N\} \subset W$. En consecuencia, f es μ-separablemente valuada.

Por otra parte, como $f : \Omega \setminus N \to W$ es el límite puntual de una sucesión de funciones μ-simples y, por tanto, $\mathcal{S}$-simples, el Lema 2.8 implica que f es $\mathcal{S}$-medible.

$\Leftarrow$) : Existen $N \in \mathcal{S}$ con $\mu(N) = 0$ y W subespacio vectorial cerrado y separable de H tal que $f(\Omega \setminus N) = \{f(x) \in V \,:\, x \in \Omega \setminus N\} \subset W$.

El Lema 2.9 y la Observación 2.10 implican que $f : \Omega \setminus N \to W$ es el límite puntual de una sucesión de funciones $s_k : \Omega \setminus N \to W$ que son $\mathcal{S}$-simples.

Dado que $(\Omega, \mathcal{S}, \mu)$ es σ-finito, existe una sucesión creciente (Ω_k) de elementos de $\mathcal{S}$ tal que $\Omega = \cup_{k=1}^{\infty} \Omega_k$ y $\mu(\Omega_k) < \infty$ para todo $k \in \mathbb{N}$. Definamos, por tanto, la sucesión de funciones $\tilde{s}_k : \Omega \setminus N \to W$ como sigue:

$$\tilde{s}_k(x) := s_k(x) \cdot 1_{\Omega_k}(x).$$

Sea $k \in \mathbb{N}$. Como s_k es $\mathcal{S}$-simple, entonces existen $v_1^k, \ldots, v_N^k \in W$ distintos y subconjuntos $A_1^k, \ldots, A_N^k \in \mathcal{S}$ mutuamente disjuntos tales que:

$$s_k(x) = \sum_{j=1}^{N} v_j^k \, 1_{A_j^k}(x).$$

Por tanto,

$$\tilde{s}_k(x) = s_k(x) \cdot 1_{\Omega_k}(x) = \sum_{j=1}^{N} v_j^k \, 1_{A_j^k \cap \Omega_k}(x).$$

en donde $\mu(A_j^k \cap \Omega_k) \leq \mu(\Omega_k) < \infty$ para todo $j = 1, \ldots, N$.

En consecuencia, $\tilde{s}_k : \Omega \setminus N \to W$ es una función μ-simple para todo $k \in \mathbb{N}$ y se sigue que:

$$\lim_{k\to\infty} \tilde{s}_k(x) = \lim_{k\to\infty} s_k(x) \cdot \lim_{k\to\infty} 1_{\Omega_k} = \lim_{k\to\infty} s_k(x) = f_k(x) \quad \forall\, x \in \Omega \setminus N.$$

Esto prueba que f es fuertemente μ-medible. ■

Notemos que en la demostración anterior, la hipótesis de que el espacio de medida sea σ-finito no es necesaria para el hecho de que toda función fuertemente μ-medible es $\mathcal{S}$-medible.

Repasando la demostración del Teorema 2.19, vemos que hemos probado lo siguiente.

Proposición 2.20 *Si $(\Omega, \mathcal{S}, \mu)$ es un espacio de medida σ-finita, entonces toda función $f : \Omega \to H$ fuertemente $\mathcal{S}$-medible es fuertemente μ-medible.*

Demostración: Como f es fuertemente $\mathcal{S}$-medible, existe una sucesión (s_k) de funciones $\mathcal{S}$-simples tales que $s_k(x) \to f_k(x)$.

Dado que $(\Omega, \mathcal{S}, \mu)$ es σ-finito, existe una sucesión creciente (Ω_k) en $\mathcal{S}$ tal que $\Omega = \cup_{k=1}^{\infty} \Omega_k$ y $\mu(\Omega_k) < \infty$ para todo $k \in \mathbb{N}$. Siguiendo la demostración del Teorema 2.19, concluimos que la sucesión de funciones $\tilde{s}_k : \Omega \to H$ dada por $s_k := s_k \cdot 1_{\Omega_k}$ es una sucesión de funciones μ-simples que converge puntualmente a f en Ω.

Es decir, f es fuertemente μ-medible. ∎

El siguiente resultado muestra que el recíproco es válido salvo en un conjunto de medida cero. Es decir, una función fuertemente μ-medible es igual a una función fuertemente $\mathcal{S}$-medible casi donde quiera relativo a μ.

Teorema 2.21 *Sea $(\Omega, \mathcal{S}, \mu)$ es un espacio de medida σ-finita y $f : \Omega \to H$ una función. Las siguientes afirmaciones son equivalentes:*

(i) f es fuertemente μ-medible.

(ii) Existe una función $\tilde{f} : \Omega \to H$ fuertemente $\mathcal{S}$-medible tal que $f = \tilde{f}$ c.d.rel.μ.

Demostración: $(i) \Rightarrow (ii)$: Dado que f es fuertemente μ-medible, existe una sucesión (s_k) de funciones μ-simples y un subconjunto $N \in \mathcal{S}$ con $\mu(N) = 0$ tales que $s_k(x) \to f(x)$ para todo $x \in \Omega \setminus N$.

Definamos la sucesión de funciones $\tilde{s}_k : \Omega \to H$ como:

$$\tilde{s}_k(x) := \begin{cases} s_k(x) & \text{si } x \in \Omega \setminus N \\ 0 & \text{si } x \notin \Omega \setminus N \end{cases}$$

Es claro que $\tilde{s}_k : \Omega \to H$ es una función $\mathcal{S}$-simple para todo $k \in \mathbb{N}$ y $(\tilde{s}_k)$ converge puntualmente a la función $\tilde{f} : \Omega \to H$ dada por:

$$\tilde{f}(x) := \begin{cases} f(x) & \text{si } x \in \Omega \setminus N \\ 0 & \text{si } x \notin \Omega \setminus N \end{cases}$$

De este modo, $\tilde{f} : \Omega \to H$ es fuertemente $\mathcal{S}$-medible y $\tilde{f}(x) = f(x)$ para todo $x \in \Omega \setminus N$.

$(ii) \Rightarrow (i)$: Sea $\tilde{f} : \Omega \to H$ una función fuertemente $\mathcal{S}$-medible y $N \in \mathcal{S}$ con $\mu(N) = 0$ tal que $\tilde{f}(x) = f(x)$ para todo $x \in \Omega \setminus N$.

Luego, sea (ζ_k) sucesión de funciones $\mathcal{S}$-simples tales que $\zeta_k(x) \to \tilde{f}(x)$ para todo $x \in \Omega$. Así, es claro que $\zeta_k(x) \to f(x)$ para todo $x \in \Omega \setminus N$.

Consideremos (Ω_k) una sucesión creciente en $\mathcal{S}$ tal que $\Omega = \cup_{k=1}^{\infty}\Omega_k$ y $\mu(\Omega_k) < \infty$ para todo $k \in \mathbb{N}$. Definiendo $s_k : \Omega \to V$ como $s_k(x) := \zeta_k(x) \cdot 1_{\Omega_k}(x)$ se tiene que (s_k) es una sucesión de funciones μ-simples tal que $s_k(x) \to f(x)$ para todo $x \in \Omega \setminus N$. Es decir, f es fuertemente μ-medible. ■

Sea $V = (V, \|\cdot\|_V)$ un espacio de Banach sobre $\mathbb{R}$. Si U un subconjunto no vacío de V y Θ un subespacio vectorial de V^*. Decimos que Θ **separa puntos** de U si, para cualesquiera $u, v \in U$ tales que $u \neq v$ entonces existe $\vartheta \in \Theta$ tal que $\vartheta(x) \neq \vartheta(y)$. De este modo, si W un subespacio vectorial separable de V y Θ un subespacio vectorial de V^* que separa puntos de W, entonces existe una sucesión (ϑ_k) de elementos de Θ que separa puntos de W (ver [87]).

Proposición 2.22 *Sea Θ un subespacio vectorial de H^* tal que Θ separa puntos de H. Si $f, g : \Omega \to H$ son funciones fuertemente μ-medibles tales que $\vartheta(f) = \vartheta(g)$ c.d.rel.μ para todo $\vartheta \in \Theta$, entonces $f = g$ c.d.rel.μ.*

Demostración: Existen $N_f, N_g \in \mathcal{S}$ con $\mu(N_f) = \mu(N_g) = 0$ y W_f, W_g subconjuntos numerables H tales que $f(x) \in \overline{\text{lin}(W_f)}$ para todo $x \in \Omega \setminus N_f$ y $g(x) \in \overline{\text{lin}(W_g)}$ para todo $x \in \Omega \setminus N_g$.

Considerando $N_0 := N_f \cup N_g$ y $W := \overline{\text{lin}(W_f \cup W_g)}$ es claro que $f(x), g(x) \in W$ para todo $x \in \Omega \setminus N$ con $\mu(N) = 0$ y W subespacio vectorial cerrado y separable de H.

Existe una sucesión (ϑ_k) de elementos de Θ que separa puntos de W. De este modo, para $k \in \mathbb{N}$ dada, existe $N_k \in \mathcal{S}$ con $\mu(N_k) = 0$ tal que $\vartheta_k(f) = \vartheta_k(g)$ para todo $x \in \Omega \setminus N_k$.

Así pues, definiendo $N := \cup_{k=0}^{\infty} N_k$ es claro que $f(x) = g(x)$ para todo $x \in \Omega \setminus N$ y $\mu(N) = 0$. ■

Ejemplo 2.23 *Sea $(\Omega, \mathcal{S})$ espacio medible y definamos $\mu : \mathcal{S} \to \mathbb{R} \cup \{\infty\}$ como sigue:*

$$\mu(A) := \begin{cases} 0 & si\ A = \varnothing, \\ \infty & si\ A \neq \varnothing. \end{cases}$$

Es claro que $(\Omega, \mathcal{S}, \mu)$ no es de medida σ-finita. Definamos $f : \Omega \to \mathbb{R}$ como $f(x) := 1_\Omega(x)$. Es inmediato que f es $\mathcal{S}$-medible pues $\Omega \in \mathcal{S}$.

Definamos $s_k : \Omega \to \mathbb{R}$ como $s_k := f$ para todo $k \in \mathbb{N}$. Así, (s_k) es una sucesión de funciones $\mathcal{S}$-simples tales que $s_k(x) \to f(x)$ para todo $x \in \Omega$. En consecuencia, f es fuertemente $\mathcal{S}$-medible.

Es claro también que s_k no es una función μ-simple pues $\mu(\Omega) = \infty$. De este modo, la única función μ-simple es $1_\varnothing$ pues $\varnothing$ es el único elemento de $\mathcal{S}$ de medida finita. Por tanto, no existe sucesión de funciones μ-simples tales que converjan puntualmente casi dondequiera a f. Es decir, f no es fuertemente μ-medible.

Sea $L : \mathbb{R} \to \mathbb{R}$ lineal y, por tanto, continua. Es claro que $L \circ f : \Omega \to \mathbb{R}$ es $\mathcal{S}$-medible por ser composición de funciones $\mathcal{S}$-medibles. Esto implica que f es débilmente $\mathcal{S}$-medible.

El ejemplo anterior muestra que los Teoremas 2.16 y 2.19 no son ciertos si se ignora la hipótesis de que el espacio de medida sea σ-finito. Además, prueba que no es cierto que toda función $\mathcal{S}$-simple es μ-simple.

Terminamos esta sección con un ejemplo importante.

Ejemplo 2.24 *Sea* $\Omega = (0,1)$*,* $\mathcal{S} = \mathcal{B}(0,1)$ *y* $\mu = \lambda$ *la medida de Lebesgue en* $\mathcal{B}(0,1)$*. Definamos* $f : (0,1) \to L^2(0,1)$ *como* $f(x) := 1_{(0,x)}$*.*

Notemos que $L^2(0,1)$ *es un espacio de Hilbert separable. Sea* $\psi \in (L^2(0,1))^*$*. El teorema de representación de Fréchet-Riesz afirma que, existe una única función* $\phi \in L^2(0,1)$ *tal que:*

$$\psi(f(x)) = \int_{(0,1)} \phi(y) 1_{(0,x)}(y) \, d\lambda(y).$$

Dado que $(0,1)$ *es de medida finita, entonces* $L^2(0,1) \subset L^1(0,1)$*. Por tanto,* $\psi \circ f : (0,1) \to \mathbb{R}$ *está bien definida.*

Sea $x_0 \in (0,1)$ *y* $\varepsilon > 0$*. Para* $\delta := \frac{\varepsilon}{\|\phi\|_{L^2(0,1)}+1} > 0$ *se cumple que:*

$$\begin{aligned} |\psi(f(x)) - \psi(f(x_0))| &\leq \int_{(0,1)} |\phi(y)| \left|1_{(0,x)}(y) - 1_{(0,x_0)}(y)\right| \, d\lambda(y) \\ &\leq \|\phi\|_{L^2(0,1)} \|1_{(0,x)} - 1_{(0,x_0)}\|_{L^2(0,1)} \\ &\leq \|\phi\|_{L^2(0,1)} \, \lambda((0,x) \triangle (0,x_0)) \\ &\leq \|\phi\|_{L^2(0,1)} \, |x - x_0| \\ &< \varepsilon \end{aligned}$$

si $|x - x_0| < \delta$*.*

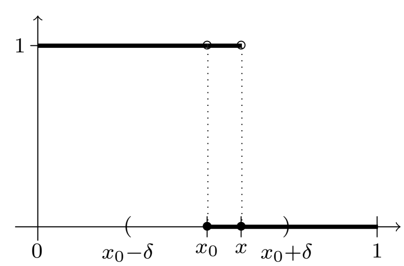


Figura 2.1: Gráfica de la función $f(x) := 1_{(0,x)}$

En consecuencia, $\psi \circ f : (0,1) \to \mathbb{R}$ *es continua y, por tanto, medible. El teorema de mensurabilidad de Pettis (ver* Teorema 2.16*) afirma que* $f : (0,1) \to L^2(0,1)$ *es fuertemente* λ*-medible.*

2.2. La integral de Bochner

En esta sección definimos la integral de Bochner que, en esencia, generaliza a la integral de Lebesgue pues en este caso, trabajamos con funciones que toman valores en un espacio de Hilbert sobre $\mathbb{R}$. Tal como ocurrió con la integral de Lebesgue, comenzamos la construcción de la integral de Bochner a través de la noción de integral para funciones simples. Establecemos además, sus propiedades fundamentales así como algunos teoremas de convergencia.

Sea $H = (H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$, $H^* := (H^*, \|\cdot\|_{H^*})$ su espacio dual y $(\Omega, \mathcal{S}, \mu)$ un espacio de medida.

Definición 2.25 *Sea $s : \Omega \to H$ una función $\mathcal{S}$-simple descrita de la siguiente forma:*

$$s(x) = \sum_{j=1}^{N} v_j \, 1_{A_j}(x).$$

*Decimos que s es **Bochner-integrable en Ω con respecto de la medida** μ (o simplemente **Bochner-integrable**) si s es distinta de 0_H en los conjuntos de medida finita. Es decir, si $v_j = 0_H$ cuando $\mu(A_j) = \infty$ para $j = 1, \ldots, N$.*

*En cuyo caso, definimos la **integral de Bochner de s en Ω con respecto de la medida** μ, denotada por $\mathfrak{B}\int_\Omega s(x)\,d\mu(x)$ (o simplemente por $\mathfrak{B}\int_\Omega s\,d\mu$), como el valor en H dado por:*

$$\mathfrak{B}\int_\Omega s(x)\,d\mu(x) = \mathfrak{B}\int_\Omega s\,d\mu = \sum_{j=1}^{N} v_j\,\mu(A_j). \tag{2.6}$$

Es claro que todas las funciones μ-simples son Bochner-integrables. En un espacio de medida finita, toda función $\mathcal{S}$-simple es Bochner integrable.

Por ejemplo, si s es la función constante con valor 0_H, entonces s es $\mathcal{S}$-simple pues $s(x) := 0_H \cdot 1_\Omega(x)$. De este modo, en el caso más general, si $\mu(\Omega) = \infty$ entonces s es Bochner-integrable y $\mathfrak{B}\int_\Omega s\,d\mu = 0_H$.

Observemos que, en el caso en que $H = \mathbb{R}$, la integral de Bochner de una función $\mathcal{S}$-simple coincide con la integral de Lebesgue.

Proposición 2.26 *Sean $s, t : \Omega \to H$ funciones $\mathcal{S}$-simples y $\alpha \in \mathbb{R}$. Si s y t son Bochner-integrables, entonces $\alpha\, s$ y $s + t$ son Bochner-integrables. En cuyo caso,*

$$\mathfrak{B}\int_\Omega (\alpha\, s)\,d\mu = \alpha\mathfrak{B}\int_\Omega s\,d\mu \quad y \quad \mathfrak{B}\int_\Omega (s+t)\,d\mu = \mathfrak{B}\int_\Omega s\,d\mu + \mathfrak{B}\int_\Omega t\,d\mu.$$

Demostración: Supongamos que s y t están representadas de la siguiente forma:

$$s(x) = \sum_{j=1}^{N} u_j \, 1_{A_j}(x) \quad \text{y} \quad t(x) = \sum_{k=1}^{M} v_k \, 1_{B_k}(x).$$

Si $\alpha = 0$ entonces $\alpha \, s(x) = 0_H$ para todo $x \in \Omega$ y es claro que $\alpha \, s$ es Bochner-integrable. Supongamos que $\alpha \neq 0$. Entonces, $\alpha \, u_1, \ldots, \alpha \, u_N \in H$ son todos distintos y se tiene que:

$$\alpha \, s(x) = \sum_{j=1}^{N} \alpha \, u_j \, 1_{A_j}(x).$$

Es decir, $\alpha \, s$ es una función $\mathcal{S}$-simple. Luego, es claro que $\alpha \, s$ es Bochner-integrable pues $\alpha \, u_j \neq 0_H$ siempre que $\mu(A_j) < \infty$. En consecuencia,

$$\mathfrak{B} \int_{\Omega} (\alpha s) \, d\mu = \sum_{j=1}^{N} \alpha \, u_j \, \mu(A_j) = \alpha \sum_{j=1}^{N} u_j \, \mu(A_j) = \alpha \mathfrak{B} \int_{\Omega} s \, d\mu.$$

Por otra parte, la función $s + t$ admite la siguiente representación:

$$s + t = \sum_{j=1}^{N} \sum_{k=1}^{M} (u_j + v_k) 1_{A_j \cap B_k}$$

Sean $w_1, \ldots, w_p$ los distintos valores en H del conjunto $\{u_j + v_k \, : \, j = 1, \ldots, N \text{ y } k = 1, \ldots, M\}$. Para cada $\ell = 1, \ldots, p$ definamos el conjunto:

$$C_\ell = \cup \left\{ A_j \cap B_k \, : \, u_j + v_k = w_\ell \right\}.$$

Entonces, la función $s + t$ queda descrita de la siguiente:

$$s + t = \sum_{\ell=1}^{p} w_\ell \, 1_{C_\ell}.$$

donde los $w_1, \ldots, w_p$ son todos distintos, $C_1, \ldots, C_p \in \mathcal{S}$ y son mutuamente disjuntos. Es decir, $s + t$ es una función $\mathcal{S}$-simple.

Si $\mu(C_\ell) < \infty$ para todo $\ell = 1, \ldots, p$ entonces es claro que $s + t$ es una función Bochner-integrable. Supongamos entonces que existe $\ell_0 \in \{1, \ldots, p\}$ tal que $\mu(C_{\ell_0}) = \infty$. Así pues, existen $j_0 \in \{1, \ldots, N\}$ y $k_0 \in \{1, \ldots, M\}$ tales que $\mu(A_{j_0} \cap B_{k_0}) = \infty$ y, por tanto, $\mu(A_{j_0}) = \infty = \mu(B_{k_0})$. Dado que s y t son Bochner-integrables, entonces $u_{j_0} = v_{k_0} = 0_H$ y, en consecuencia, $w_{\ell_0} = u_{j_0} + v_{k_0} = 0_H$ lo que prueba que $s + t$ es Bochner-integrable.

Ahora bien, denotemos por $\sum_{(\ell)}$ a la suma que se extiende sobre las parejas (j,k) tales que $u_j + v_k = w_\ell$. Así pues, es claro que $\mu(C_\ell) = \sum_{(\ell)} \mu(A_j \cap B_k)$ y por tanto:

$$\begin{aligned}\mathfrak{B}\int_\Omega (s+t)\,d\mu &= \sum_{\ell=1}^{p} w_\ell \mu(C_\ell) = \sum_{\ell=1}^{p}\sum_{(\ell)} w_\ell \mu(A_j \cap B_k) = \sum_{\ell=1}^{p}\sum_{(\ell)} (u_j + v_k)\mu(A_j \cap B_k)\\ &= \sum_{j=1}^{N}\sum_{k=1}^{M}(u_j+v_k)\mu(A_j\cap B_k)\\ &= \sum_{j=1}^{N} u_j \sum_{k=1}^{M}\mu(A_j\cap B_k) + \sum_{k=1}^{M} v_k \sum_{j=1}^{N}\mu(A_j\cap B_k).\end{aligned}$$

Dado que $\Omega = \cup_{j=1}^{N} A_j = \cup_{k=1}^{M} B_k$, entonces $A_j = \cup_{k=1}^{M}(A_j \cap B_k)$ y $B_k = \cup_{j=1}^{N}(A_j \cap B_k)$ para cada $j = 1, \ldots, N$ y $k = 1, \ldots, M$. Aplicando la aditividad de μ obtenemos que:

$$\begin{aligned}\mathfrak{B}\int_\Omega (s+t)\,d\mu &= \sum_{j=1}^{N} u_j \sum_{k=1}^{M}\mu(A_j\cap B_k) + \sum_{k=1}^{M} v_k \sum_{j=1}^{N}\mu(A_j\cap B_k)\\ &= \sum_{j=1}^{N} u_j\mu(A_j) + \sum_{k=1}^{M} v_k\mu(B_k) = \mathfrak{B}\int_\Omega s\,d\mu + \mathfrak{B}\int_\Omega t\,d\mu,\end{aligned}$$

como afirma el enunciado. ■

Sean $A \in \mathcal{S}$ y $s : \Omega \to H$ una función $\mathcal{S}$-simple que es Bochner-integrable. Si $s = \sum_{j=1}^{N} v_j\, 1_{A_j}$ entonces:

$$(s \cdot 1_A)(x) = \sum_{j=1}^{N} v_j\, 1_{A_j \cap A}(x).$$

Así, si $\mu(A_j \cap A) = \infty$ para algún $j = 1, \ldots, N$, entonces $\mu(A_j) = \infty$ y se tiene que $v_j = 0_H$. Por tanto, $s \cdot 1_A$ es Bochner-integrable.

Lo anterior nos permite extender la definición de la integral de Bochner de una función $\mathcal{S}$-simple sobre cualquier subconjunto medible de Ω.

Definición 2.27 *Sean $s : \Omega \to H$ una función $\mathcal{S}$-simple y Bochner-integrable y $A \in \mathcal{S}$. Definimos la **integral de Bochner de** s **en** A **con respecto de la medida** μ (o simplemente la **integral de Bochner de** s **en** A) como el valor en H dado por:*

$$\mathfrak{B}\int_A s(x)\,d\mu(x) := \mathfrak{B}\int_A s\,d\mu = \mathfrak{B}\int_\Omega (s \cdot 1_A)\,d\mu = \sum_{j=1}^{N} v_j\, \mu(A_j \cap A).$$

Sea $s : \Omega \to H$ una función $\mathcal{S}$-simple descrita de la siguiente forma:

$$s(x) := \sum_{j=1}^{N} v_j \, 1_{A_j}(x).$$

Para $x \in \Omega$ dada, existe un único $j \in \{1, \ldots, N\}$ tal que $x = A_j$. Por tanto, $s(x) = v_j$ y se tiene que $\|s(x)\|_H = \|v_j\|_H$.

Así pues, la función norma $\|s\|_H : \Omega \to \mathbb{R}$ de s está dada por:

$$\|s\|_H(x) = \|s(x)\|_H = \sum_{j=1}^{N} \|v_j\|_H \, 1_{A_j}(x)$$

en donde es claro que los números reales $\|v_1\|_H, \ldots, \|v_N\|_H$ son todos distintos. En consecuencia, $\|s\|_H$ es una función simple no negativa.

Además, si $\mu(A_j) = \infty$ para algún $j \in \{1, \ldots, N\}$ entonces $v_j = 0_H$ y, por tanto, $\|v_j\|_H = 0$. Es decir, $\|s\|_H$ es Lebesgue-integrable si s es Bochner-integrable. Más aún,

$$\left\| \mathfrak{B} \int_{\Omega} s \, d\mu \right\|_H = \left\| \sum_{j=1}^{N} v_j \, \mu(A_j) \right\|_H \leq \sum_{j=1}^{N} \|v_j\|_H \, \mu(A_j) = \int_{\Omega} \|s\|_H \, d\mu. \tag{2.7}$$

Damos ahora la definición de la integral de Bochner de una función fuertemente μ-medible. Antes, notemos lo siguiente:

Sean $f, g : \Omega \to H$ funciones fuertemente μ-medibles y $\alpha \in \mathbb{R}$. Existen sucesiones (s_k) y (t_k) de funciones μ-simples y subconjuntos de Ω medida cero N_1 y N_2 tales que:

$$\lim_{k \to \infty} s_k(x) = f(x) \quad \forall\, x \in \Omega \setminus N_1$$

$$\lim_{k \to \infty} t_k(x) = g(x) \quad \forall\, x \in \Omega \setminus N_2$$

Definiendo el conjunto $N := N_1 \cup N_2$ de medida cero y la sucesión de funciones $\xi_k : \Omega \to H$ como:

$$\xi_k(x) := (s_k + t_k)(x) \cdot 1_{\Omega \setminus N}(x)$$

es claro que ξ_k es μ-simple para todo $k \in \mathbb{N}$ y $\lim_{k \to \infty} \xi_k(x) = f(x) + g(x)$ para todo $x \in \Omega \setminus N$. En consecuencia, $f + g$ es fuertemente μ-medible.

Del mismo modo, la función $\alpha\, s_k : \Omega \to H$ es μ-simple para todo $k \in \mathbb{N}$ y se tiene que $\lim_{k \to \infty} \alpha\, s_k(x) = \alpha f(x)$ para todo $x \in \Omega \setminus N$. Es decir, $\alpha\, f$ es fuertemente μ-medible.

Finalmente, si $f(x) = 0_H$ para todo $x \in \Omega$ entonces es claro que f es $\mathcal{S}$-medible pues para $B \in \mathcal{B}(H)$ se tiene que $f^{-1}(B) = \Omega \in \mathcal{S}$ si $0_H \in B$ ó $f^{-1}(B) = \varnothing \in \mathcal{S}$ si $0_H \notin B$. Además, $f(\Omega) = \{0_H\}$ que claramente es un subespacio vectorial cerrado y separable de

H. El teorema de medibilidad de Pettis (ver Teorema 2.16) asegura que f es fuertemente μ-medible.

Denotemos por $\mathfrak{M}(\Omega; H)$ al conjunto de todas las funciones $f : \Omega \to H$ que son fuertemente μ-medibles. Es decir:

$$\mathfrak{M}(\Omega; H) := \{f : \Omega \to H \,:\, f \text{ es fuertemente } \mu - \text{medible}\}.$$

Lo anterior asegura que este es un conjunto con las operaciones:

$$(f+g)(x) := f(x) + g(x) \quad \text{y} \quad (\alpha f)(x) := \alpha\, f(x) \quad x \in \Omega$$

para $f, g \in \mathfrak{M}(\Omega; H)$ y $\alpha \in \mathbb{R}$ es un espacio vectorial sobre $\mathbb{R}$.

Definición 2.28 *Sea $f : \Omega \to H$ una función fuertemente μ-medible. Decimos que f es* ***Bochner-integrable en Ω con respecto de la medida*** *μ (o simplemente* ***Bochner-integrable****) si, existe una sucesión de funciones $f_k : \Omega \to V$ $\mathcal{S}$-simples tales que f_k es Bochner-integrable para todo $k \in \mathbb{N}$ y tal que:*

(a) $\lim_{k\to\infty} f_k(x) = f(x)$ *p.c.t.* $x \in \Omega$.
(b) $\|f_k - f\|_H : \Omega \to \mathbb{R}$ *es Lebesgue-integrable para todo* $k \in \mathbb{N}$.
(c) $\lim_{k\to\infty} \int_\Omega \|f_k - f\|_H \, d\mu = 0$.

Sea $f : \Omega \to H$ una función fuertemente μ-medible y supongamos que existe una sucesión (f_k) de funciones $\mathcal{S}$-simples y Bochner-integrables que satisfacen las propiedades *(a)*, *(b)* y *(c)* de la Definición 2.28.

Dado que $f_k - f : \Omega \to H$ es fuertemente μ-medible, entonces $\|f_k - f\|_H : \Omega \to \mathbb{R}$ es medible no negativa para todo $k \in \mathbb{N}$ (ver Proposición 2.18). Además, por el inciso *(b)* se tiene que, para $\varepsilon > 0$ dada, existe $k_0 \in \mathbb{N}$ tal que:

$$\int_\Omega \|f_k - f\|_H \, d\mu < \frac{\varepsilon}{2} \qquad \forall\, k \geq k_0.$$

Consideremos la sucesión de elementos de H dada por $\varsigma_k := \mathfrak{B}\int_\Omega f_k \, d\mu$ pues f_k es Bochner-integrable para todo $k \in \mathbb{N}$. De la Proposición 2.26 y la identidad (2.7) se tiene que:

$$\begin{aligned}
\left\| \mathfrak{B}\int_\Omega f_k \, d\mu - \mathfrak{B}\int_\Omega f_j \, d\mu \right\|_H &= \left\| \mathfrak{B}\int_\Omega (f_k - f_j) \, d\mu \right\|_H \\
&\leq \int_\Omega \|f_k - f_j\|_H \, d\mu \\
&\leq \int_\Omega \|f_k - f\|_H \, d\mu + \int_\Omega \|f - f_j\|_H \, d\mu \\
&< \varepsilon
\end{aligned}$$

para todo $k, j \geq k_0$.

Es decir, (ς_k) es de Cauchy en H y, por tanto, existe $\lim_{k\to\infty} \varsigma_k \in H$. Esto nos permitirá definir la integral de Bochner en Ω de una función fuertemente μ-medible y Bochner-integrable.

Definición 2.29 *Sea $f : \Omega \to H$ una función fuertemente μ-medible y Bochner-integrable en Ω con respecto de la medida μ. Definimos la* ***integral de Bochner de f sobre Ω con respecto de la medida*** *μ, que denotaremos por $\mathfrak{B}\int_\Omega f(x)\, d\mu(x)$ (o simplemente por $\mathfrak{B}\int_\Omega f\, d\mu$) como el valor en V dado por:*

$$\mathfrak{B}\int_\Omega f(x)\, d\mu(x) = \mathfrak{B}\int_\Omega f\, d\mu := \lim_{k\to\infty} \mathfrak{B}\int_\Omega f_k\, d\mu \tag{2.8}$$

en donde (f_k) es una sucesión de funciones $\mathcal{S}$-simples y Bochner-integrables que satisfacen las propiedades (a), (b) y (c) de la Definición 2.28.

Sea $f : \Omega \to H$ una función fuertemente μ-medible y Bochner-integrable en Ω. Consideremos (f_k) y $(\tilde{f}_k)$ dos sucesiones de funciones $\mathcal{S}$-simple y Bochner-integrables que satisfacen las propiedades *(a)*, *(b)* y *(c)* de la Definición 2.28. Por lo anterior, las sucesiones en H dadas por

$$\varsigma_k := \mathfrak{B}\int_\Omega f_k\, d\mu \qquad \text{y} \qquad \tilde{\varsigma}_k := \mathfrak{B}\int_\Omega \tilde{f}_k\, d\mu$$

son de Cauchy en H y, por tanto, convergen en H. Nuevamente, la Proposición 2.26 y la identidad (2.7) implican que:

$$\begin{aligned}
\left\| \mathfrak{B}\int_\Omega f_k\, d\mu - \mathfrak{B}\int_\Omega \tilde{f}_k\, d\mu \right\|_H &= \left\| \mathfrak{B}\int_\Omega (f_k - \tilde{f}_k)\, d\mu \right\|_H \\
&\leq \int_\Omega \|f_k - \tilde{f}_k\|_H\, d\mu \\
&\leq \int_\Omega \|f_k - f\|_H\, d\mu + \int_\Omega \|f - \tilde{f}_k\|_H\, d\mu.
\end{aligned}$$

Tomando el límite cuando $k \to \infty$ en la desigualdad anterior concluimos que:

$$\left\| \mathfrak{B}\int_\Omega f_k\, d\mu - \mathfrak{B}\int_\Omega \tilde{f}_k\, d\mu \right\|_H \to 0$$

y, por tanto,

$$\lim_{k\to\infty} \mathfrak{B}\int_\Omega f_k\, d\mu = \lim_{k\to\infty} \mathfrak{B}\int_\Omega \tilde{f}_k\, d\mu.$$

Es decir, la integral de Bochner en Ω de $f : \Omega \to H$ no depende de la sucesión aproximadora de funciones $\mathcal{S}$-simples y Bochner-integrables.

Denotemos por:

$$\mathfrak{B}(\Omega; H) := \{f \in \mathfrak{M}(\Omega; H) \,:\, f \text{ es Bochner-integrable}\}.$$

El siguiente resultado afirma que $\mathfrak{B}(\Omega; H)$ es un espacio vectorial y que la integral de Bochner es una función lineal.

Teorema 2.30 (linealidad de la integral de Bochner) *Si f_1, f_2 son Bochner-integrables y $\alpha_1, \alpha_2 \in \mathbb{R}$, entonces $\alpha_1 f_1 + \alpha_2 f_2 \in \mathfrak{B}(\Omega; H)$ y*

$$\mathfrak{B}\int_\Omega (\alpha_1 f_1 + \alpha_2 f_2)\, d\mu = \alpha_1 \mathfrak{B}\int_\Omega f_1\, d\mu + \alpha_2 \mathfrak{B}\int_\Omega f_2\, d\mu.$$

Demostración: Existen sucesiones (s_k) y (t_k) de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $s_k(x) \to f_1(x)$ p.c.t.$x \in \Omega$, $t_k(x) \to f_2(x)$ p.c.t.$x \in \Omega$, $\|s_k - f_1\|_V$, $\|t_k - f_2\|_H$ son Lebesgue-integrables para todo $k \in \mathbb{N}$ y

$$\lim_{k\to\infty} \|s_k - f_1\|_H = 0 \quad \text{y} \quad \lim_{k\to\infty} \|t_k - f_2\|_H = 0.$$

No es difícil notar, en virtud de la Proposición 2.26, que $\alpha_1 s_k + \alpha_2 t_k : \Omega \to H$ es una función $\mathcal{S}$-simple y Bochner-integrable para todo $k \in \mathbb{N}$. Así, la sucesión de funciones dada por $\varsigma_k(x) := \alpha_1 s_k(x) + \alpha_2 t_k(x)$ es tal que $\varsigma_k(x) \to \alpha_1 f_1 + \alpha_2 f_2$ p.c.t. $x \in \Omega$.

Además, para todo $k \in \mathbb{N}$ se tiene que:

$$\begin{aligned}\int_\Omega \|\varsigma_k - (\alpha_1 f_1 + \alpha_2 f_2)\|_H\, d\mu &= \int_\Omega \|\alpha_1 (s_k - f_1) + \alpha_2 (t_k - f_2)\|_H\, d\mu \\ &\leq |\alpha_1| \int_\Omega \|s_k - f_1\|_H\, d\mu + |\alpha_2| \int_\Omega \|t_k - f_2\|_H\, d\mu\end{aligned}.$$

Se tiene que $\|(\alpha_1 s_k + \alpha_2 t_k) - (\alpha_1 f_1 + \alpha_2 f_2)\|_H$ es Lebesgue-integrable para todo $k \in \mathbb{N}$. Luego, tomando el límite cuando $k \to \infty$ en la desigualdad anterior obtenemos que:

$$\lim_{k\to\infty} \int_\Omega \|\varsigma_k - (\alpha_1 f_1 + \alpha_2 f_2)\|_H\, d\mu = 0.$$

Por tanto, $\alpha_1 f_1 + \alpha_2 f_2$ es Bochner-integrable en Ω.

Aplicando la Proposición 2.26 concluimos que:

$$\begin{aligned}\mathfrak{B}\int_\Omega (\alpha_1 f_1 + \alpha_2 f_2)\, d\mu &= \lim_{k\to\infty} \mathfrak{B}\int_\Omega (\alpha_1 s_k + \alpha_2 t_k)\, d\mu \\ &= \alpha_1 \lim_{k\to\infty} \mathfrak{B}\int_\Omega s_k\, d\mu + \alpha_2 \lim_{k\to\infty} \mathfrak{B}\int_\Omega t_k\, d\mu \\ &= \alpha_1\, \mathfrak{B}\int_\Omega f_1\, d\mu + \alpha_2\, \mathfrak{B}\int_\Omega f_2\, d\mu\end{aligned}$$

como afirma el enunciado. ■

La discusión previa a la Definición 2.27 nos permite afirmar que, para $f \in \mathfrak{B}(\Omega; H)$ y $A \in \mathcal{S}$ dados, la función $(f \cdot 1_A)$ es Bochner-integrable en Ω. Por tanto, definimos **la integral de Bochner de f en A con respecto de la medida μ** como:

$$\mathfrak{B}\int_A f(x)\, d\mu(x) = \mathfrak{B}\int_A f\, d\mu := \mathfrak{B}\int_\Omega (f \cdot 1_A)\, d\mu.$$

La idea de encontrar una sucesión de funciones $\mathcal{S}$-simples y Bochner-integrables que converjan puntualmente casi donde quiera a una función $f : \Omega \to H$ fuertemente μ-medible y, que además cumplan las propiedades *(b)* y *(c)* de la Definición 2.28 para probar que f es Bochner-integrable resulta tediosa y difícil de aplicar. Sin embargo, el siguiente resultado establece las condiciones necesarias y suficientes para que una función fuertemente μ-medible sea Bochner-integrable y, que nos permitirá usar la teoría desarrollada para la integral de Lebesgue.

Teorema 2.31 (Criterio de integrabilidad Lebesgue-Bochner) *Sea $(\Omega, \mathcal{S}, \mu)$ un espacio de medida Y sea $f : \Omega \to H$ una función fuertemente μ-medible. Entonces, $f : \Omega \to H$ es Bochner-integrable si y sólo si su función norma $\|f\|_H : \Omega \to \mathbb{R}$ es Lebesgue-integrable.*

Demostración: $\Rightarrow$) : Sea (f_k) sucesión de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $f_k(x) \to f(x)$ p.c.t. $x \in \Omega$, $\|f_k - f\|_H$ sea Lebesgue-integrable para todo $k \in \mathbb{N}$ y $\lim_{k\to\infty} \int_\Omega \|f_k - f\|_H\, d\mu = 0$.

Para $\varepsilon := 1 > 0$, existe $k_1 \in \mathbb{N}$ tal que:

$$\int_\Omega \|f_k - f\|_H\, d\mu < 1 \qquad \forall k \geq k_1.$$

Como f_{k_1} es $\mathcal{S}$-simple y Bochner-integrable, entonces $\|f_{k_1}\|_H$ es Lebesgue-integrable. Es decir, existe $c \in \mathbb{R}$ tal que $\int_\Omega \|f_{k_1}\|_H\, d\mu < c$.

Por tanto,

$$\int_\Omega \|f\|_H\, d\mu \leq \int_\Omega \|f_{k_1}\|_H\, d\mu + \int_\Omega \|f - f_{k_1}\|_H\, d\mu < c + 1$$

y esto prueba que $\|f\|_H$ es Lebesgue-integrable.

$\Leftarrow$) : Dado que $f : \Omega \to H$ es fuertemente μ-medible, existe una sucesión (ξ_k) de funciones μ-simples tales que:

$$\lim_{k\to\infty} \|\xi_k(x) - f(x)\|_H = 0 \qquad p.c.t.\, x \in \Omega. \tag{2.9}$$

Definamos $A_k := \{x \in \Omega \,:\, \|\xi_k(x)\|_H \leq \frac{3}{2}\|f(x)\|_H\}$ para todo $k \in \mathbb{N}$. Es claro que $A_k \in \mathcal{S}$ pues $\|\xi_k\|_H$ es una función medible para todo $k \in \mathbb{N}$.

Definiendo $f_k : \Omega \to H$ como $f_k(x) := \xi_k(x) \cdot 1_{A_k}(x)$ es claro que f_k es μ-simple y, por tanto, Bochner-integrable para todo $k \in \mathbb{N}$. Además, $\|f_k(x)\|_H \leq \frac{3}{2}\,\|f(x)\|_H$ para todo $x \in \Omega$ y para todo $k \in \mathbb{N}$. Supongamos que $\|f(x)\|_H = 0$. Entonces, $f(x) = 0_H$ y, por tanto, $\|f_k(x)\|_H = 0$ para todo $k \in \mathbb{N}$. Así, $f_k(x) = 0_H$ para todo $k \in \mathbb{N}$ y esto implica que $f_k(x) \to f(x)$ p.c.t. $x \in \Omega$. Por otra parte en (2.9), si $\|f(x)\|_H > 0$, entonces para $\varepsilon := \frac{\|f(x)\|_H}{2} > 0$ existe $k_0 \in \mathbb{N}$ tal que:

$$\big|\|\xi_k(x)\|_H - \|f(x)\|_H\big| \leq \|\xi_k(x) - f(x)\|_H < \frac{\|f(x)\|_H}{2} \qquad \forall k \geq k_0.$$

En consecuencia,

$$\|\xi_k(x)\|_H \leq \frac{3}{2}\|f(x)\|_H \qquad \forall k \geq k_0,$$

lo que implica que $\xi_k(x) = f_k(x)$ para todo $k \geq k_0$. Es decir, $f_k(x) \to f(x)$ p.c.t.$x \in \Omega$.

Finalmente, como $\|f\|_H$ es Lebesgue-integrable y $\|f_k(x)\|_H \leq \frac{3}{2}\,\|f(x)\|_H$ para todo $x \in \Omega$ y para todo $k \in \mathbb{N}$, del teorema de la convergencia dominada de Lebesgue concluimos que:

$$\lim_{k\to\infty} \int_\Omega \|f_k - f\|_H \, d\mu = 0.$$

En consecuencia, f es Bochner-integrable. ■

Sea $f : \Omega \to H$ Bochner-integrable. Repasando la demostración del teorema anterior obtenemos que existe una sucesión (f_k) de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $\|f_k\|_V \leq \frac{3}{2}\|f\|_H$ para todo $k \in \mathbb{N}$ y tales que:

$$\mathfrak{B}\int_\Omega f \, d\mu = \lim_{k\to\infty} \mathfrak{B}\int_\Omega f_k \, d\mu.$$

Dado que $f_k : \Omega \to H$ es $\mathcal{S}$-simple y Bochner-integrable para todo $k \in \mathbb{N}$, entonces $\|f_k\|_H : \Omega \to H$ es Lebesgue-integrable y por (2.7) se tiene que:

$$\left\|\mathfrak{B}\int_\Omega f_k \, d\mu\right\|_H \leq \int_\Omega \|f_k\|_H \, d\mu \qquad \forall\, k \in \mathbb{N}.$$

Aplicando la continuidad de norma y el teorema de la convergencia dominada de Lebesgue concluimos que:

$$\left\|\mathfrak{B}\int_\Omega f \, d\mu\right\|_H = \lim_{k\to\infty} \left\|\mathfrak{B}\int_\Omega f_k \, d\mu\right\|_H \leq \lim_{k\to\infty} \int_\Omega \|f_k\|_H \, d\mu = \int_\Omega \|f\|_H \, d\mu. \tag{2.10}$$

La desigualdad (2.10) es llamada **desigualdad de Jensen**.

Ejemplo 2.32 *Sea* $((0,1), \mathcal{B}(0,1), \lambda)$ *espacio de medida. Consideremos la función* $f : (0,1) \to L^2(0,1)$ *del* Ejemplo 2.24 *que es fuertemente* λ*-medible.*

Observemos que, para cada $x \in (0,1)$ *se tiene que:*

$$\|f(x)\|_{L^2(0,1)} = \left(\int_{(0,1)} 1_{(0,x)}(y)\, d\lambda(y) \right)^{1/2} = \lambda((0,x))^{1/2} = x^{1/2}.$$

Es claro que $\|f(x)\|_{L^2(0,1)}$ *es continua, Riemann-integrable y, por tanto, Lebesgue-integrable en* $(0,1)$*. En consecuencia,* $f : (0,1) \to L^2(0,1)$ *es Bochner-integrable. De* (2.10) *se tiene que:*

$$\left\| \mathfrak{B} \int_{(0,1)} f(x)\, d\lambda(x) \right\|_{L^2(0,1)} \leq \int_{(0,1)} \|f(x)\|_{L^2(0,1)}\, d\lambda(x) = \tfrac{2}{3}.$$

Sea $f : \Omega \to H$ una función Bochner-integrable. Siguiendo la demostración del Teorema 2.30 concluimos que existe una sucesión (f_k) de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $f_k(x) \to f(x)$ p.c.t. $x \in \Omega$, $\|f_k(x)\|_H \leq \frac{3}{2}\|f(x)\|_H$ para todo $x \in \Omega$ y $\lim_{k\to\infty} \int_\Omega \|f_k - f\|_H \, d\mu = 0$.

Escribiendo a la función $f_k : \Omega \to H$ como:

$$f_k(x) := \sum_{j=1}^{N_k} v_j^k \, 1_{A_j^k}(x)$$

se tiene que:

$$\mathfrak{B} \int_\Omega f\, d\mu = \lim_{k\to\infty} \mathfrak{B} \int_\Omega f_k\, d\mu = \lim_{k\to\infty} \sum_{j=1}^{N_k} v_j^k \mu(A_j^k). \tag{2.11}$$

Sea $\psi : H \to \mathbb{R}$ lineal y continua. El teorema de medibilidad de Pettis asegura que las funciones $\psi \circ f_k, \psi \circ f : \Omega \to \mathbb{R}$ son medibles. Entonces,

$$|\psi(f_k(x))| \leq \|\psi\|_{H^*} \, \|f_k(x)\|_H \leq \tfrac{3}{2} \, \|\psi\|_{H^*} \, \|f(x)\|_H \qquad \forall\, x \in \Omega \quad \forall\, k \in \mathbb{N}$$

y, como $\|f\|_H : \Omega \to \mathbb{R}$ es Lebesgue-integrable por el Teorema 2.30 entonces $\psi \circ f_k : \Omega \to \mathbb{R}$ es Lebesgue-integrable para todo $k \in \mathbb{N}$.

Así, la continuidad de la función $\psi : H \to \mathbb{R}$ implica que $\psi(f_k(x)) \to \psi(f(x))$ p.c.t. $x \in \Omega$. El teorema de la convergencia dominada de Lebesgue asegura que $\psi \circ f$ es Lebesgue-integrable y

$$\int_\Omega \psi(f)\, d\mu = \lim_{k\to\infty} \int_\Omega \psi(f_k)\, d\mu. \tag{2.12}$$

Por tanto, aplicando la linealidad y continuidad de ψ en (2.11) y (2.12), concluimos que:

$$\begin{aligned}\psi\left(\mathfrak{B}\int_\Omega f\,d\mu\right) &= \psi\left(\lim_{k\to\infty}\mathfrak{B}\int_\Omega f_k\,d\mu\right) = \lim_{k\to\infty}\psi\left(\mathfrak{B}\int_\Omega f_k\,d\mu\right)\\ &= \lim_{k\to\infty}\psi\left(\sum_{j=1}^{N_k} v_j^k\mu(A_j^k)\right) = \lim_{k\to\infty}\sum_{j=1}^{N_k}\psi(v_j^k)\,\mu(A_j^k)\\ &= \lim_{k\to\infty}\int_\Omega \psi(f_k)\,d\mu = \int_\Omega \psi(f)\,d\mu.\end{aligned}$$

Notemos que hemos probado el siguiente resultado.

Teorema 2.33 *Sea $f:\Omega\to H$ una función Bochner-integrable. Entonces, para cada $\psi\in H^*$ se tiene que:*

$$\psi\left(\mathfrak{B}\int_\Omega f\,d\mu\right) = \int_\Omega \psi(f)\,d\mu. \tag{2.13}$$

Existe un análogo al teorema de la convergencia dominada para la integral de Bochner y, que suele conocerse como el teorema de la convergencia dominada vectorial.

Teorema 2.34 (de la convergencia dominada de Lebesgue-Bochner) *Sean $f, f_k : \Omega\to H$ funciones fuertemente μ-medibles con las siguientes propiedades:*

(i) f_k *es Bochner-integrable,*

(ii) $f_k(x)\to f(x)$ *p.c.t.* $x\in\Omega$,

(iii) existe una función Lebesgue-integrable $g:\Omega\to\mathbb{R}$ *tal que, para cada* $k\in\mathbb{N}$, *se cumple que* $\|f_k(x)\|_H\le g(x)$ *p.c.t.* $x\in\Omega$.

Entonces f es Bochner-integrable y

$$\mathfrak{B}\int_\Omega f\,d\mu = \lim_{k\to\infty}\mathfrak{B}\int_\Omega f_k\,d\mu.$$

Demostración: Dado que para cada $k\in\mathbb{N}$, se cumple que $\|f_k(x)\|_H\le g(x)$ p.c.t. $x\in\Omega$ entonces $\|f(x)\|_H\le g(x)$ p.c.t. $x\in\Omega$.

Es claro que $\|f_k-f\|_H:\Omega\to\mathbb{R}$ es una función $\mathcal{S}$-medible para todo $k\in\mathbb{N}$ (ver Proposición 2.18). Observemos que, $\|f_k(x)-f(x)\|_H\le\|f_k(x)\|_H+\|f(x)\|_H\le 2g(x)$ p.c.t. $x\in\Omega$ y para todo $k\in\mathbb{N}$. Por tanto, $\|f_k-f\|_H$ es Lebesgue-integrable para todo $k\in\mathbb{N}$.

Aplicando el teorema de la convergencia dominada de Lebesgue concluimos que:

$$\lim_{k\to\infty}\int_\Omega\|f_k-f\|_H\,d\mu = 0.$$

Sea $k \in \mathbb{N}$. Dado que $f_k : \Omega \to H$ es Bochner-integrable, existe una sucesión (s_j^k) de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $\lim_{j\to\infty} s_j^k(x) = f_k(x)$ p.c.t. $x \in \Omega$ y tales que:

$$\lim_{j\to\infty} \int_\Omega \|f_k - s_j^k\|_H \, d\mu = 0.$$

Así, para $\varepsilon := \frac{1}{k} > 0$ existe $j_k \in \mathbb{N}$ tal que:

$$\int_\Omega \|f_k - s_j^k\|_H \, d\mu \le \tfrac{1}{k} \qquad \forall\, j \ge j_k.$$

Definiendo la sucesión de funciones $\xi_k : \Omega \to H$ como $\xi_k(x) := s_{j_k}^k(x)$ se tiene que:

$$\int_\Omega \|f - \xi_k\|_H \, d\mu \le \int_\Omega \|f - f_k\|_H \, d\mu + \int_\Omega \|f_k - \xi_k\|_H \, d\mu \le \int_\Omega \|f - f_k\|_H \, d\mu + \tfrac{1}{k}$$

y, por tanto, $\lim_{k\to\infty} \int_\Omega \|f - \xi_k\|_H \, d\mu = 0$.

En consecuencia, f es Bochner-integrable y se tiene que:

$$\mathfrak{B}\int_\Omega f \, d\mu = \lim_{k\to\infty} \mathfrak{B}\int_\Omega f_k \, d\mu$$

como afirma el enunciado. ■

2.3. Los espacios $L^p(\Omega\,;\,H)$

Sean $(\Omega, \mathcal{S}, \mu)$ un espacio de medida finita y $H = (H, \langle\cdot,\cdot\rangle, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$.

En el conjunto de todas las funciones fuertemente μ-medibles

$$\mathfrak{M}(\Omega; H) := \{f : \Omega \to H \;:\; f \text{ es fuertemente } \mu\text{-medible}\}.$$

consideramos la relación de equivalencia dada por

$$f \sim_\mu g \qquad \Longleftrightarrow \qquad f(x) = g(x) \quad (c.d.rel.\mu). \tag{2.14}$$

Al conjunto de clases de equivalencia lo denotamos por:

$$M(\Omega; H) := \mathfrak{M}(\Omega; H) \,/\, \sim_\mu . \tag{2.15}$$

Observemos que, si $f_i = g_i$ c.d. en Ω, $i = 1, 2$, entonces $\alpha f_1 + \beta f_2 = \alpha g_1 + \beta g_2$ c.d. en Ω para cualesquiera $\alpha, \beta \in \mathbb{R}$. Así, la estructura de espacio vectorial de $\mathfrak{M}(\Omega; H)$ induce una estructura de espacio vectorial en $M(\Omega; H)$.

Denotaremos a la clase de equivalencia de f simplemente por $f : \Omega \to H$. Si una función está definida c.d. en Ω, la consideraremos definida en todo Ω dándole el valor 0_H en los puntos en los que no está definida.

Definición 2.35 *Si $p \in [1,\infty)$ definimos:*

$$L^p(\Omega,\mathcal{S},\mu;H) := \{f \in M(\Omega;H) \,:\, \|f\|_H \in L^p(\Omega,\mathcal{S},\mu)\}\,.$$

Para $f \in L^p(\Omega,\mathcal{S},\mu;H)$ denotamos por

$$\|f\|_{L^p(\Omega;H)} := \left(\int_\Omega \|f\|_H^p \, d\mu\right)^{1/p}. \tag{2.16}$$

A partir de este momento y siempre que no se preste a confusión escribiremos $L^p(\Omega;H)$ en lugar de $L^p(\Omega,\mathcal{S},\mu;H)$. Algunas de las propiedades de los espacios $L^p(\Omega;H)$ tienen su análogo con los espacios clásicos de Lebesgue $L^p(\Omega)$ y que tienen que ver con el criterio de integrabilidad de Lebesgue-Bochner.

Desigualdad de Hölder-Riesz: Sean $p,q \in (1,\infty)$ tales $\frac{1}{p}+\frac{1}{q}=1$. Consideremos $f \in L^p(\Omega;H)$, $g \in L^q(\Omega;H)$ y, por tanto, $\|f\|_H \in L^p(\Omega)$ y $\|g\|_H \in L^q(\Omega)$.

Aplicando la desigualdad de Hölder-Riesz concluimos que $\|f\|_H\|g\|_H \in L^1(\Omega)$ y

$$\int_\Omega \|f\|_H \, \|g\|_H \, d\mu \le \left(\int_\Omega \|f\|_H^p \, d\mu\right)^{1/p} \left(\int_\Omega \|g\|_H^q \, d\mu\right)^{1/q}.$$

En consecuencia:

$$\int_\Omega \|f\|_H \, \|g\|_H \, d\mu \le \|f\|_{L^p(\Omega;H)}\|g\|_{L^q(\Omega;H)}. \tag{2.17}$$

Observemos que, la expresión fg puede no tener sentido en el espacio de Hilbert H a menos que éste sea una $\mathbb{R}$-álgebra con unidad.

Desigualdad de Minkowski: Sea $p \in [1,\infty)$ y sean $f,g \in L^p(\Omega;H)$. Entonces, $\|f\|_H, \|g\|_H \in L^p(\Omega)$ y aplicando la desigualdad de Minkowski obtenemos que $\|f+g\|_H \in L^p(\Omega)$ y

$$\left(\int_\Omega \|f+g\|_H^p \, d\mu\right)^{1/p} \le \left(\int_\Omega \|f\|_H^p \, d\mu\right)^{1/p} + \left(\int_\Omega \|g\|_H^p \, d\mu\right)^{1/p}.$$

Por tanto, $f+g \in L^p(\Omega;H)$ y se tiene que:

$$\|f+g\|_{L^p(\Omega;H)} \le \|f\|_{L^p(\Omega;H)} + \|g\|_{L^p(\Omega;H)}. \tag{2.18}$$

Proposición 2.36 $L^p(\Omega;H) = (L^p(\Omega;H), \|\cdot\|_{L^p(\Omega;H)})$ *es un espacio normado para todo $p \in [1,\infty)$.*

Demostración: Sea $p \in [1,\infty)$. Si $f = 0_H$ entonces $\|f\|_H = 0 \in L^p(\Omega)$ y, en consecuencia $\|f\|_H^p = 0$. Es decir, $f \in L^p(\Omega; H)$ y $\|f\|_{L^p(\Omega;H)} = 0$. Inversamente, si $\|f\|_{L^p(\Omega;H)} = 0$ entonces $\|f\|_H = 0$ y, por tanto, $f = 0_H$.

Si f es fuertemente μ-medible y $\alpha \in \mathbb{R}$, entonces $\alpha\, f$ es fuertemente μ-medible. Además, observemos que $\|\alpha\, f\|_H^p = |\alpha|^p \|f\|_H^p$ y se sigue $\|\alpha\, f\|_{L^p(\Omega;H)} = |\alpha|\|f\|_{L^p(\Omega;H)}$.

Esta última afirmación junto con la identidad (2.18) aseguran que $L^p(\Omega; H)$ es un espacio vectorial y que $\|\cdot\|_{L^p(\Omega;H)}$ es una norma. ■

Probaremos a continuación que $L^p(\Omega; H)$ es un espacio de Banach. Para ello requerimos el siguiente lema.

Teorema 2.37 *$L^p(\Omega; H) = (L^p(\Omega; H), \|\cdot\|_{L^p(\Omega;H)})$ es un espacio de Banach para todo $p \in [1,\infty)$.*

Demostración: Sea $p \in [1,\infty)$ y (f_k) una sucesión en $L^p(\Omega; H)$ tal que:

$$\sum_{k=1}^{\infty} \|f_k\|_{L^p(\Omega;H)} < \infty.$$

Definamos $g_m, g : \Omega \to \mathbb{R}$ como:

$$g_m := \sum_{k=1}^{m} \|f_k\|_H \qquad \text{y} \qquad g := \sum_{k=1}^{\infty} \|f_k\|_H = \sup_{m\in\mathbb{N}} g_m.$$

Observemos que $\|\|f_k\|_H\|_{L^p(\Omega)}^p = \int_\Omega \|f_k\|_H^p \, d\mu = \|f_k\|_{L^p(\Omega;H)}^p$ para todo $k \in \mathbb{N}$.

En consecuencia:

$$\begin{aligned}
\int_\Omega |g_m|^p \, d\mu = \int_\Omega \left(\sum_{k=1}^{m} \|f_k\|_H\right)^p d\mu = \left\|\sum_{k=1}^{m} \|f_k\|_H\right\|_{L^p(\Omega)}^p &\leq \left(\sum_{k=1}^{m} \|\|f_k\|_H\|_{L^p(\Omega)}\right)^p \\
&\leq \left(\sum_{k=1}^{m} \|f_k\|_{L^p(\Omega;H)}\right)^p \\
&\leq \left(\sum_{k=1}^{\infty} \|f_k\|_{L^p(\Omega;H)}\right)^p \\
&< \infty.
\end{aligned}$$

Por el teorema de la convergencia monótona, la función $|g|^p = \sup_{m\in\mathbb{N}} |g_m|^p$ es Lebesgue-integrable. Existe $N \in \mathcal{S}$ con $\mu(N) = 0$ tal que:

$$g(x) := \sum_{k=1}^{\infty} \|f_k(x)\|_V < \infty \qquad \forall\, x \in \Omega \smallsetminus N.$$

Como H es de Hilbert, el Criterio de Weierstrass implica que la serie $\sum_{k=1}^{\infty} f_k(x)$ converge en H para cada $x \in \Omega \setminus N$. Definimos, por tanto, $f : \Omega \to H$ como:

$$f(x) := \begin{cases} \sum_{k=1}^{\infty} f_k(x) & \text{si } \ x \in \Omega \setminus N, \\ 0_H & \text{si } \ x \in N. \end{cases}$$

Dado que:

$$\left(f(x) - \sum_{k=1}^{m} f_k(x) \right) = \sum_{k=m+1}^{\infty} f_k(x)$$

entonces:

$$\left\| \sum_{k=m+1}^{\infty} f_k(x) \right\|_H^p \to 0 \quad \text{ y } \quad \left\| \sum_{k=m+1}^{\infty} f_k(x) \right\|_H^p \leq |g(x)|^p \quad \text{p.c.t. } x \in \Omega.$$

Por el teorema de la convergencia dominada de Lebesgue $\|f\|_H \in L^p(\Omega)$ y

$$\lim_{m\to\infty} \int_\Omega \left\| f(x) - \sum_{k=1}^{m} f_k(x) \right\|_H^p d\mu = \lim_{m\to\infty} \int_\Omega \left\| \sum_{k=m+1}^{\infty} f_k(x) \right\|_H^p d\mu = 0.$$

Es decir, $\sum_{k=1}^{\infty} f_k$ converge a f en $L^p(\Omega; H)$. Por tanto, $L^p(\Omega; H)$ es de Banach. ■

Proposición 2.38 *Sean $(\Omega, \mathcal{S}, \mu)$ un espacio de medida finita y $H = (H, \langle \cdot, \cdot \rangle_H)$ un espacio de Hilbert. En el espacio $L^2(\Omega; H)$ definimos:*

$$\langle f, g \rangle_{L^2(\Omega;H)} := \int_\Omega \langle f, g \rangle_H \, d\mu.$$

Entonces, $\langle f, g \rangle_{L^2(\Omega;H)}$ es un producto escalar en $L^2(\Omega; H)$. Más aún, $L^2(\Omega; H)$ es un espacio de Hilbert.

Demostración: Sean $f_1, f_2, g \in L^2(\Omega; H)$ y $a, b \in \mathbb{R}$. De las propiedades de producto escalar de $\langle \cdot, \cdot \rangle_H$ y de la integral de Lebesgue se tiene que:

$$\begin{aligned} \langle af_1 + bf_2, g \rangle_{L^2(\Omega;H)} &= \int_\Omega \langle af_1 + bf_2, g \rangle_H \, d\mu \\ &= a \int_\Omega \langle f_1, g \rangle_H \, d\mu + b \int_\Omega \langle f_2, g \rangle_H \, d\mu \\ &= a \langle f_1, g \rangle_{L^2(\Omega;H)} + b \langle f_2, g \rangle_{L^2(\Omega;H)}. \end{aligned}$$

$$\begin{aligned} \langle f_1, f_2 \rangle_{L^2(\Omega;H)} &= \int_\Omega \langle f_1, f_2 \rangle_H \, d\mu \\ &= \int_\Omega \langle f_2, f_1 \rangle_H \, d\mu \\ &= \langle f_1, f_2 \rangle_{L^2(\Omega;H)}. \end{aligned}$$

Esto prueba (PE1) y (PE2). Las propiedades (PE3) Y (PE4) son inmediatas.

Por otra parte, como H es un espacio de Hilbert entonces, $\langle v, v\rangle_H = \|v\|_H^2$ para cada $v \in H$ y, por tanto,

$$\langle f, f\rangle_{L^2(\Omega;H)} = \int_\Omega \langle f, f\rangle_H \, d\mu = \int_\Omega \|f\|_H^2 \, d\mu \qquad \forall\, f \in L^2(\Omega; H).$$

Es decir, la norma inducida por $\langle\cdot,\cdot\rangle_{L^2(\Omega;H)}$ es la norma $\|\cdot\|_{L^2(\Omega;H)}$ en $L^2(\Omega; H)$ que sabemos, por el Teorema 2.37, es completa. ■

Denotemos por $\mathcal{S}^p(\Omega; H)$ al conjunto de todas las funciones $\mathcal{S}$-simples que están en $L^p(\Omega; H)$. Es decir,

$$\mathcal{S}^p(\Omega; H) := \{s : \Omega \to H \,:\, s \text{ es } \mathcal{S} - \text{simple y } s \in L^p(\Omega; H)\}\,.$$

Sea $f \in L^p(\Omega; H)$. Siguiendo la demostración del Teorema 2.30 concluimos que existe una sucesión (s_k) de funciones $\mathcal{S}$-simples y Bochner-integrables tales que $s_k \to f$ puntualmente en Ω y $\|f_k(x)\|_H \leq \frac{3}{2}\|f(x)\|_H$ para todo $x \in \Omega$. De este modo, $f_k \in L^p(\Omega; H)$ para todo $k \in \mathbb{N}$ y por el teorema de la convergencia dominada de Lebesgue concluimos que:

$$\lim_{k\to\infty} \int_\Omega \|f_k - f\|_H^p \, d\mu = 0.$$

Es decir, $\mathcal{S}^p(\Omega; H)$ es denso en $L^p(\Omega; H)$.

Para $\psi_1, \ldots, \psi_N \in L^p(\Omega)$ y $v_1, \ldots, v_N \in H$ dados definimos la función $\sum_{j=1}^N \psi_j \otimes v_j : \Omega \to H$ como sigue:

$$(\sum_{j=1}^N \psi_j \otimes v_j)(x) := \sum_{j=1}^N \psi_j(x)\, v_j. \tag{2.19}$$

Es claro que $(\sum_{j=1}^N \psi_j \otimes v_j) \in L^p(\Omega; H)$ pues:

$$\left\|(\sum_{j=1}^N \psi_j \otimes v_j)(x)\right\|_H^p \leq \sum_{j=1}^N |\psi_j(x)|^p \|v_j\|_H^p.$$

Denotaremos por $L^p(\Omega) \otimes H$ al subespacio vectorial de $L^p(\Omega; H)$ dado por todas las funciones de la forma (2.19).

Observemos que, si $s \in \mathcal{S}^p(\Omega; H)$ entonces:

$$s(x) = \sum_{i=1}^M v_i\, 1_{A_i}(x)$$

en donde $A_1, \ldots, A_M \in \mathcal{S}$ son todos disjuntos y podemos suponer que $\mu(A_i) < \infty$ para todo $i = 1, \ldots, M$. En consecuencia, $1_{A_i} \in L^p(\Omega)$ para todo $i = 1, \ldots, M$ y, por tanto, $s \in L^p(\Omega) \otimes H$. Es decir, $\mathcal{S}^p(\Omega; H) \subset L^p(\Omega) \otimes H$ lo que implica que $L^p(\Omega) \otimes H$ es también denso en $L^p(\Omega; H)$.

Sean $(\Omega, \mathcal{S}, \mu)$ y $(\widetilde{\Omega}, \widetilde{\mathcal{S}}, \widetilde{\mu})$ dos espacios de medida finita y consideremos $T : L^1(\Omega) \to L^1(\widetilde{\Omega})$ una función lineal y continua. Definamos $T_0 : \mathcal{S}^1(\Omega; H) \to L^1(\widetilde{\Omega}; H)$ como:

$$T_0(s) := \sum_{j=1}^{N} v_j \, T(1_{A_j}) \tag{2.20}$$

en donde $s = \sum_{j=1}^{N} v_j \, 1_{A_j}$. Esta función es claramente lineal. Observa que $\|T(1_{A_j})\|_{L^1(\widetilde{\Omega})} \leq \|T\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))} \|1_{A_j}\|_{L^1(\Omega)}$ para todo $j = 1, \ldots, N$. En consecuencia:

$$\begin{aligned}
\|T_0(s)\|_{L^1(\widetilde{\Omega}; H)} &\leq \sum_{j=1}^{N} \|T(1_{A_j})\|_{L^1(\widetilde{\Omega})} \, \|v_j\|_H \\
&\leq \|T\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))} \sum_{j=1}^{N} \|v_j\|_H \|1_{A_j}\|_{L^1(\Omega)} \\
&\leq \|T\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))} \sum_{j=1}^{N} \|v_j\|_H \, \mu(A_j) \\
&= \|T\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))} \|s\|_{L^1(\Omega; H)}
\end{aligned}$$

lo que implica que T_0 es continua en $\mathcal{S}^1(\Omega; H)$.

Dado que $\mathcal{S}^1(\Omega; H)$ es denso en $L^1(\Omega; H)$ entonces la función lineal y continua $T_0 : \mathcal{S}^1(\Omega; H) \to L^1(\widetilde{\Omega}; H)$ admite una única extensión lineal y continua $\widetilde{T} : L^1(\Omega; H) \to L^1(\widetilde{\Omega}; H)$ tal que:

$$\|\widetilde{T}\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))} \leq \|T\|_{\mathcal{L}_c(L^1(\Omega), L^1(\widetilde{\Omega}))}. \tag{2.21}$$

Sea $v \in H$ y consideremos $\psi \in L^1(\Omega)$.

Caso 1: Si $\psi = \sum_{i=1}^{M} \alpha_i \, 1_{B_i}$ con $\alpha_i \in \mathbb{R}$ y $B_i \in \mathcal{S}$ disjuntos tal que $\mu(B_i) < \infty$ para todo $i = 1, \ldots, M$ entonces:

$$\psi \otimes v = \sum_{i=1}^{M} \alpha_i \, v \, 1_{B_i}.$$

Es decir, $\psi \otimes v \in \mathcal{S}^1(\Omega; H)$.

Por lo tanto,

$$\tilde{T}(\psi \otimes v) = T_0(\psi \otimes v) = \sum_{i=1}^{M} \alpha_i \, v \, T(1_{B_i}) \, v = T(\psi) \, v.$$

CASO 2: Consideremos una sucesión (ψ_k) de funciones simples en $L^1(\Omega)$ tales que $\psi_k \to \psi$ en $L^1(\Omega)$.

Entonces,

$$\psi \otimes v = \lim_{k\to\infty} (\psi_k \otimes v).$$

Usando la continuidad de $\widetilde{T}$ y T, del caso anterior concluimos que:

$$\widetilde{T}(\psi \otimes v) = \lim_{k\to\infty} \widetilde{T}(\psi_k \otimes v) = \lim_{k\to\infty} (T(\psi_k)\, v) = T(\psi)\, v.$$

De los casos anteriores concluimos que para toda $\psi \in L^1(\Omega)$ y todo $v \in H$ se tiene que:

$$\widetilde{T}(\psi \otimes v) = T(\psi)\, v. \tag{2.22}$$

Hemos demostrado el siguiente resultado.

Proposición 2.39 *Sean $(\Omega, \mathcal{S}, \mu)$ y $(\widetilde{\Omega}, \widetilde{\mathcal{S}}, \widetilde{\mu})$ dos espacios de medida finita. Para cada $T : L^1(\Omega) \to L^1(\widetilde{\Omega})$ función lineal y continua, existe una única función lineal y continua $\widetilde{T} : L^1(\Omega; H) \to L^1(\widetilde{\Omega}; H)$ tal que:*

$$\widetilde{T}(\psi \otimes v) = T(\psi)\, v \qquad \forall\, \psi \in L^1(\Omega), \quad \forall\, v \in H.$$

Más aún, $\|T\|_{\mathcal{L}_c(L^1(\Omega),L^1(\tilde{\Omega}))} = \|\widetilde{T}\|_{\mathcal{L}_c(L^1(\Omega;H),L^1(\tilde{\Omega};H))}$.

Sea $p \in [1, \infty)$ arbitrario pero fijo. Consideremos $T : L^p(\Omega) \to L^p(\Omega)$ una función lineal y continua. Definimos la función $T \otimes id_V : L^P(\Omega) \otimes H \to L^p(\Omega) \otimes H$ como:

$$(T \otimes id_V)(\psi \otimes v) := T(\psi) \otimes v$$

Esta función es claramente lineal. Del mismo modo que en la Proposición 2.39 nos preguntamos cuando es posible extender a esta función a una función lineal y continua en $L^p(\Omega; H)$. De manera general, no es posible a menos que la función T sea un operador positivo, es decir, $T(\psi) \geq 0$ si $\psi \geq 0$.

Supongamos que T es un operador positivo y sea $f \in L^p(\Omega) \otimes H$.

Consideremos que $f = \sum_{j=1}^{N} 1_{A_j} \otimes v_j$ con $A_1, \ldots, A_N \in \mathcal{S}$ mutuamente disjuntos y $v_1, \ldots, v_N \in H$ distintos. Dado que T es positivo, entonces $|T(1_{A_j})| = T(1_{A_j})$ para todo

$j = 1, \ldots, N$. Por tanto:

$$\begin{aligned}
\|(T \otimes id_H)(f)\|_{L^p(\Omega;H)}^p &= \int_\Omega \left\| \sum_{j=1}^N T(1_{A_j}) \otimes v_j \right\|_H^p d\mu \\
&\leq \int_\Omega \left| \sum_{j=1}^N |T(1_{A_j})| \|v_j\|_H \right|^p d\mu \\
&= \int_\Omega \left| \sum_{j=1}^N T(1_{A_j}) \|v_j\|_H \right|^p d\mu \\
&= \int_\Omega \left| T\left(\sum_{j=1}^N 1_{A_j} \|v_j\|_H \right) \right|^p d\mu \\
&\leq \|T\|_{\mathcal{L}_c(L^p(\Omega),L^p(\Omega))}^p \int_A \left| \sum_{j=1}^N 1_{A_j} \|v_j\|_H \right|^p d\mu \\
&\leq \|T\|_{\mathcal{L}_c(L^p(\Omega),L^p(\Omega))}^p \left\| \sum_{j=1}^N 1_{A_j} \otimes v_j \right\|_{L^p(\Omega;H)}^p .
\end{aligned}$$

Dado que $L^p(\Omega) \otimes H$ es denso en $L^p(\Omega; H)$ entonces $T \otimes id_H$ tiene una única extensión lineal y continua $\widetilde{T \otimes id_H} : L^p(\Omega; H) \to L^p(\Omega; H)$ tal que:

$$\left\| \widetilde{T \otimes id_H} \right\|_{\mathcal{L}_c(L^p(\Omega;H))} \leq \|T\|_{\mathcal{L}_c(L^p(\Omega))}$$

De manera general, tenemos el siguiente resultado.

Proposición 2.40 *Supongamos que H es separable y sean $1 \leq p, q < \infty$. Para $T : L^p(\Omega) \to L^q(\Omega)$ una función lineal y continua, la función*

$$T \otimes id_H : L^p(\Omega) \otimes H \to L^q(\Omega) \otimes H$$

es el único operador lineal tal que:

$$(T \otimes id_H)(\psi \otimes v) = T(\psi) \otimes v \qquad \forall\, \psi \in L^p(\Omega), \quad \forall\, v \in H. \tag{2.23}$$

Además, si T es positivo (e.d. $T(\psi) \geq 0$ si $\psi \geq 0$) entonces $T \otimes id_H$ se extiende a una función lineal y continua $\widetilde{T \otimes id_H} : L^p(\Omega; H) \to L^q(\Omega; H)$ que tiene norma igual a T.

Demostración: Consulta, por ejemplo, [61]. ■

Si $(\Omega, \mathcal{S}, \mu)$ es de medida finita, para cualesquiera $1 \leq p < q < \infty$, se tiene que $L^q(\Omega; H) \subset L^p(\Omega; H)$.

Observa lo siguiente: Sea $p \in [1,\infty)$ arbitrario pero fijo y (f_k) una sucesión en $L^p(\Omega;H)$ tal que $f_k \to f$ en $L^p(\Omega;H)$. Es decir,

$$\int_\Omega \|f_k - f\|_H^p \, d\mu \to 0.$$

Entonces,

$$|\|f_k\|_H - \|f\|_H|^p \leq \|f_k - f\|_H^p$$

de modo que:

$$\int_\Omega |\|f_k\|_H - \|f\|_H|^p \, d\mu \leq \int_\Omega \|f_k - f\|_H^p \, d\mu \to 0$$

lo que implica que $\|f_k\|_H \to \|f\|_H$ en $L^p(\Omega)$.

Así pues, existe una subsucesión $(\|f_{k_j}\|_H)$ de $(\|f_k\|_H)$ y una función $g \in L^p(\Omega)$ tal que $\|f_{k_j}(x)\|_H \to \|f(x)\|_H$ p.c.t. $x \in \Omega$ y $\|f_{k_j}(x)\|_H \leq g(x)$ p.c.t. $x \in \Omega$.

Desigualdad de Jensen: Sea $(\Omega, \mathcal{S}, \mu)$ un espacio de medida finita y H un espacio de Hilbert real. Sea $f : \Omega \to H$ una función Bochner-integrable y $\phi : \mathbb{R} \to \mathbb{R}$ una función convexa tal que $\phi(\|f\|_H) : \Omega \to \mathbb{R}$ es Lebesgue-integrable. Entonces:

$$\phi\left(\int_\Omega \|f\|_H \, d\mu\right) \leq \int_\Omega \phi(\|f\|_H) \, d\mu. \tag{2.24}$$

Dado que ϕ es convexa, es posible hallar (a_k) y (b_k) sucesiones de números reales tales que $\phi(t) = \sup_{k\geq 1}(a_k t + b_k)$ para todo $k \in \mathbb{N}$ (ver [71]). Entonces,

$$\begin{aligned}\int_\Omega \phi(\|f\|_H) \, d\mu &\geq \int_\Omega (a_k\|f\|_H + b_k) \, d\mu \\ &\geq a_k \int_\Omega \|f\|_H \, d\mu + b_k \qquad \forall\, k \in \mathbb{N}.\end{aligned}$$

Tomando el supremo sobre $k \in \mathbb{N}$ en la desigualdad anterior obtenemos la desigualdad (2.24).

De este modo, para $p \in (1,\infty)$, si $(f_k) \to f$ en $L^p(\Omega;V)$ entonces de la desigualdad de Jensen y la convexidad de la función $\phi(t) = |t|^p$ se tiene que:

$$\left(\int_\Omega \|f_k - f\|_H \, d\mu\right)^p \leq \int_\Omega \|f_k - f\|_H^p \, d\mu \to 0.$$

Es decir, $(f_k) \to f$ en $L^1(\Omega;H)$.

El siguiente resultado nos brinda un criterio sobre la separabilidad del espacio $L^p(\Omega;H)$.

Definición 2.41 *Sea $X = (X, d)$ un espacio métrico y $\mathcal{B}(X)$ la σ-álgebra de Borel de X. Decimos que una medida $\mathfrak{v} : \mathcal{B}(X) \to \mathbb{R} \cup \{\infty\}$ es una* ***medida de Radon*** *si satisface las siguientes propiedades:*

(a) Para cualquier subconjunto compacto K de X se tiene que $\mathfrak{v}(K) < \infty$.

(b) Para cualquier subconjunto abierto O de X se tiene que:

$$\mathfrak{v}(O) = \sup \{\mathfrak{v}(K) \,:\, K \subset O \;\; y \;\; K \;\; \textit{es compacto}\} .$$

(c) Para cualquier subconjunto A de X se tiene que:

$$\mathfrak{v}(A) = \operatorname{ínf} \{\mathfrak{v}(B) \,:\, A \subset B \;\; y \;\; B \;\; \textit{es abierto}\} .$$

Por ejemplo, la medida de Lebesgue λ en $(0,1)$ es una medida de Radon.

Invitamos al lector a consultar, por ejemplo, [11] y [18] para un mejor estudio de estas medidas.

Teorema 2.42 *Si X es un espacio métrico, μ es una medida de Radon finita y H es un espacio de Hilbert separable, entonces $L^p(X; H)$ es separable para cada $p \in [1, \infty)$.*

En este caso omitimos la demostración pues esta queda fuera de los conceptos esenciales de este trabajo. Invitamos al lector a consultar [24; 50] para algunas versiones de la demostración.

Del teorema anterior podemos concluir que el espacio $L^p((0,1); L^p(0,1))$ con la medida de Lebesgue en $(0,1)$ es separable para todo $p \in (1, \infty)$.

Capítulo 3

Operadores compactos en espacios de Hilbert

El *teorema de descomposición espectral*, también llamado solamente teorema espectral, en espacios vectoriales de dimensión finita sobre $\mathbb{R}$ establece las condiciones bajos las cuales una función lineal A puede diagonalizarse, es decir, cuándo existe una base ortonormal de vectores propios de A tales que la matriz de A asociada a dicha base puede verse como una matriz diagonal. Recordemos que un vector propio de A es un elemento del espacio vectorial, en este caso de dimensión finita, distinto del vector cero tal que $A(v) = \lambda v$ para λ algún número real.

El objetivo de este capítulo es extender este resultado para espacios de Hilbert separables de dimensión infinita. Para lograrlo, comenzamos con el estudio de los operadores compactos entre dos espacios de Hilbert. Un operador compacto es una función lineal $T : H \to K$ entre dos espacios de Hilbert tal que la imagen de cualquier subconjunto acotado de H bajo T es un subconjunto relativamente compacto en K. Demostramos que cualquier función lineal y continua de rango finito es un operador compacto, lo cual nos permite entender que la clase de operadores compactos es una generalización de la clase de operadores de rango finito en espacios vectoriales de dimensión infinita. De hecho, cuando H es igual a K, cualquier operador compacto es el límite de una sucesión de operadores de rango finito.

Por tanto, estudiar operadores compactos entre espacios de Banach nos obliga a proporcionar resultados que nos permitan caracterizar a los subconjuntos compactos en ellos. Uno de estos resultado es el *teorema de Riesz* que caracteriza a los espacios de Banach de dimensión infinita a través de su bola unitaria cerrada.

El teorema espectral en espacios de dimensión finita establece que para que una función lineal pueda diagonalizarse esta debe ser un operador hermítico, o equivalentemente, que la matriz asociada a A sea simétrica. En espacios de Hilbert de dimensión infinita este concepto se generaliza a través de los operadores autoadjuntos. De este modo, damos una construcción explícita del operador adjunto de Hilbert de una función lineal y continua además de estudiar sus propiedades básicas. Daremos aquí algunos ejemplos interesantes. Probamos también el teorema de la alternativa de Fredholm en espacios de Hilbert que nos brinda información acerca del kernel e imagen de operadores compactos y sus adjuntos con respecto de la función identidad. Así, con todos estos resultados fundamentales, desarrollamos la teoría espectral de los operadores compactos y autoadjuntos en espacios de Hilbert separables de dimensión infinita.

Para el desarrollo de este capítulo nos basamos principalmente en [13], [31], [32], [33] [47], [48], [63], [65] y [75].

3.1. Conceptos y propiedades básicas

Sean H y K dos espacios de Hilbert sobre $\mathbb{R}$ con normas $\|\cdot\|_H$ y $\|\cdot\|_K$ respectivamente. Por simplicidad denotaremos por B_H a la bola cerrada en H con centro en 0_H y radio 1.

Definición 3.1 *Sea $T : H \to K$ una función lineal. Decimos que T es un **operador compacto** si el subconjunto $T(B_H)$ es relativamente compacto en K. Es decir, si la cerradura del conjunto $T(B_H)$ es compacta en K.*

Observa que si $T : H \to K$ es un operador compacto entonces $\overline{T(B_H)}$ es compacto en K y, por tanto, es acotado en K. Es decir, existen $w \in K$ y $\varepsilon > 0$ tal que $T(B_H) \subset \overline{T(B_H)} \subset B_K(w, \varepsilon)$. Esto implica que $T(B_H)$ es acotado en K lo que demuestra que T es continuo en H.

Denotaremos por $\mathcal{K}(H, K)$ al conjunto de todos los operadores $T : H \to K$ lineales y continuos que son compactos. A continuación damos una caracterización para los operadores compactos la cual resultará más sencilla en algunos resultados importantes de este trabajo. Supondremos que el lector está familiarizado con las propiedades de los espacios métricos compactos.

Teorema 3.2 *Sea $T \in \mathcal{L}_c(H, K)$. Entonces, T es compacto si y sólo si para cualquier sucesión (v_k) acotada en H, la sucesión $(T(v_k))$ tiene una subsucesión convergente en K.*

Demostración: $\Rightarrow)$: Sea (v_k) sucesión acotada en H. Existe $c \in \mathbb{R}$ con $c > 0$ tal que $\|v_k\|_H \leq c$ para todo $k \in \mathbb{N}$. Por tanto, la sucesión $(w_k) := (\frac{v_k}{c})$ en H es tal que $w_k := \frac{v_k}{c} \in B_H$ para todo $k \in \mathbb{N}$. Así pues, $(T(w_k))$ es una sucesión en K tal que $T(w_k) \in T(B_H) \subset \overline{T(B_H)}$ para todo $k \in \mathbb{N}$. Existe una subsucesión $(T(w_{k_j}))$ de $(T(w_k))$ que converge en K a un elemento w de $\overline{T(B_H)}$. En consecuencia, usando la continuidad de la norma y que T es lineal obtenemos:

$$\lim_{j\to\infty} \|T(v_{k_j}) - cw\|_K = \lim_{j\to\infty} |c| \left\| T\left(\frac{v_{k_j}}{c}\right) - w \right\|_K = 0,$$

lo que demuestra que $(T(v_{k_j}))$ converge en K.

$\Leftarrow$) : Sea (w_k) una sucesión en $\overline{T(B_H)}$. Entonces, existe $v_k \in B_H$ tal que $T(v_k) = w_k$ para cada $k \in \mathbb{N}$. De este modo, como (v_k) está acotada en H entonces existe $(T(v_{k_j}))$ subsucesión de $(T(v_k))$ tal que $T(v_{k_j}) \to w$ en K para algún $w \in K$. En consecuencia, (w_{k_j}) es una subsucesión de (w_k) de elementos de $\overline{T(B_H)}$ tal que $w_{k_j} \to w$ en K. Por tanto, $w \in \overline{T(B_H)}$ lo que demuestra que $\overline{T(B_H)}$ es compacto en K. ■

Probaremos a continuación que el conjunto de operadores compactos $\mathcal{K}(H,K)$ es un espacio vectorial sobre $\mathbb{R}$.

Proposición 3.3 *Si $T, S \in \mathcal{K}(H,K)$ y $\alpha, \beta \in \mathbb{R}$ entonces $\alpha T + \beta S \in \mathcal{K}(H,K)$. Además, si denotamos por 0 a la función constante tal que $0(v) = 0_K$ para todo $v \in H$ entonces $0 \in \mathcal{K}(H,K)$.*

Demostración: Sea (v_k) una sucesión acotada en H. Existe (v_{k_j}) una subsucesión de (v_k) tal que $(T(v_{k_j}))$ converge en K. Del mismo modo, como (v_{k_j}) es acotada entonces existe una subsucesión $(v_{k_{j_\ell}})$ tal que $(S(v_{k_{j_\ell}}))$ converge en K. Sea $w := \lim_{j\to\infty} T(v_{k_j})$ y $\tilde{w} := \lim_{\ell\to\infty} S(v_{k_{j_\ell}})$.

Por tanto,

$$\lim_{\ell\to\infty} \left[\alpha\, T(v_{k_{j_\ell}}) + \beta\, S(v_{k_{j_\ell}})\right] = \alpha \lim_{\ell\to\infty} T(v_{k_{j_\ell}}) + \beta \lim_{\ell\to\infty} S(v_{k_{j_\ell}}) = \alpha\, w + \beta\, \tilde{w} \in W$$

y el Teorema 3.2 asegura que $\alpha T + \beta S \in \mathcal{K}(V,W)$.

Por otro lado, $0(v_k) = 0_K$ para todo $k \in \mathbb{N}$ por lo que $0(v_k) \to 0_K \in WK$ y, en consecuencia, $0 \in \mathcal{K}(H,K)$. ■

Teorema 3.4 *$\mathcal{K}(H,K)$ es un subespacio vectorial cerrado de $\mathcal{L}_c(H,K)$.*

Demostración: Sea $T \in \overline{\mathcal{K}(H,K)}$. Existe una sucesión (T_k) tal que $T_k \in \mathcal{K}(H,K)$ para todo $k \in \mathbb{N}$ y $T_k \to T$ en $\mathcal{L}_c(H,K)$. Sea (v_k) una sucesión acotada en H y sea $c \in \mathbb{R}$ con $c > 0$ tal que $\|v_k\|_H \le c$ para todo $k \in \mathbb{N}$. Existe una subsucesión (v_k^1) de (v_k) tal que $(T_1(v_k^1))$ converge en K. La sucesión (v_k^1) también está acotada y, por tanto, existe una subsucesión (v_k^2) de (v_k^1) tal que $(T_2(v_k^2))$ converge en K. Continuando de este modo obtenemos, para cada $j \in \mathbb{N}$, una subsucesión (v_k^j) de (v_k^{j-1}) tal que $(T_j(v_k^j))$ converge en K cuando $k \to \infty$. Definamos $u_k := v_k^k$ para todo $k \in \mathbb{N}$. La sucesión (u_k) es una subsucesión de (v_k). Probaremos que $(T(u_k))$ converge en K. Como K es de Banach basta probar que la sucesión $(T(u_k))$ es de Cauchy en K.

Sea $\varepsilon > 0$. Como $T_k \to T$ en $\mathcal{L}_c(H,K)$ existe $k_0 \in \mathbb{N}$ tal que:

$$\|T - T_k\|_{\mathcal{L}_c(H,K)} < \frac{\varepsilon}{3c} \qquad \forall k \ge k_0.$$

Dado que $(T_{k_0}(u_k))$ converge en K entonces es de Cauchy en K. Existe $k_1 \in \mathbb{N}$ tal que

$$\|T_{k_0}(u_k) - T_{k_0}(u_j)\|_K < \frac{\varepsilon}{3} \qquad \forall k, j \ge k_1.$$

Definiendo $k_* := \text{máx}\{k_0, k_1\}$ obtenemos:

$$\begin{aligned}\|T(u_k) - T(u_j)\|_K &\leq \|T(u_k) - T_{k_0}(u_k)\|_K + \|T_{k_0}(u_k) - T_{k_0}(u_j)\|_K + \|T_{k_0}(u_j) - T(u_j)\|_K \\ &\leq 2\|T - T_k\|_{\mathcal{L}_c(H,K)}\|u_k\|_H + \|T_{k_0}(u_k) - T_{k_0}(u_j)\|_K \\ &< 2c\,\frac{\varepsilon}{3c} + \frac{\varepsilon}{3} = \varepsilon \qquad \forall k \geq k_*.\end{aligned}$$

En consecuencia, $(T(u_k))$ es de Cauchy en K. El Teorema 3.2 asegura que $T \in \mathcal{K}(H,K)$ y esto prueba que $\mathcal{K}(H,K)$ es cerrado en $\mathcal{L}_c(H,K)$. ■

Proposición 3.5 *Sea $Z = (Z, \|\cdot\|_Z)$ espacio de Hilbert sobre $\mathbb{R}$ y sean $T \in \mathcal{L}_c(H,K)$ y $S \in \mathcal{L}_c(K,Z)$. Si $S \in \mathcal{K}(K,Z)$ entonces $S \circ T \in \mathcal{K}(H,Z)$.*

Demostración: Sea (v_k) una sucesión acotada en H y sea $c \in \mathbb{R}$ con $c > 0$ tal que $\|v_k\|_H \leq c$. Como $T : H \to K$ es lineal y continua, entonces

$$\|T(v_k)\|_H \leq \|T\|_{\mathcal{L}_c(H,K)}\|v_k\|_H \leq \|T\|_{\mathcal{L}_c(H,K)}\, c \qquad \forall k \in \mathbb{N}$$

lo que implica que la sucesión $(T(v_k))$ es acotada en K. Dado que $S \in \mathcal{K}(K,Z)$ entonces existe una subsucesión $(T(v_{k_j}))$ de $(T(v_k))$ tal que $(S(T(v_{k_j})))$ converge en Z. Por tanto, $S \circ T \in \mathcal{K}(H,Z)$. ■

Definición 3.6 *Sea $T : H \to K$ función lineal. Decimos que T* ***es de rango finito*** *si la dimensión del subespacio vectorial* $\operatorname{im} T := \{T(v) : v \in H\}$ *de K es finita.*

Notemos que si $T \in \mathcal{L}_c(H,K)$ es de rango finito, digamos que $\dim \operatorname{im} T = N$, entonces al subespacio vectorial $\operatorname{im} T$ lo podemos identificar naturalmente con $\mathbb{R}^N$ de modo que si (v_k) es una sucesión acotada en H, de la Proposición 3.6 se tiene que $(T(v_k))$ es una sucesión acotada de elementos de $\operatorname{im} T$. El teorema de Bolzano-Weierstrass asegura que $T(v_k)$ contiene una subsucesión convergente. En consecuencia, $T \in \mathcal{K}(H,K)$.

Hemos demostrado el siguiente resultado.

Proposición 3.7 *Sea $T \in \mathcal{L}_c(H,K)$. Si T es de rango finito, entonces $T \in \mathcal{K}(H,K)$.*

De los resultados anteriores podemos concluir lo siguiente.

Corolario 3.8 *Sea (T_k) una sucesión en $\mathcal{L}_c(H,K)$ tal que T_k es de rango finito para cada $k \in \mathbb{N}$. Si existe $T \in \mathcal{L}_c(H,K)$ tal que $\|T_k - T\|_{\mathcal{L}_c(H,K)} \to 0$, entonces $T \in \mathcal{K}(H,K)$.*

Demostración: La Proposición 3.7 asegura que $T_k \in \mathcal{K}(H,K)$ para cada $k \in \mathbb{N}$. Aplicando el Teorema 3.4 concluimos que $T \in \mathcal{K}(H,K)$ como afirma el enunciado. ■

Damos a continuación una caracterización de los operadores compactos.

Teorema 3.9 *Sea $T \in \mathcal{L}_c(H,K)$, entonces:*

(a) T es un operador compacto si y sólo si para cada subconjunto acotado A de H, el subconjunto $\overline{T(A)}$ es compacto en K.

(b) Si T es un operador compacto, entonces los subconjuntos $\operatorname{im} T$ y $\overline{\operatorname{im} T}$ son separables en K.

Demostración: (a) $\Rightarrow)$: Sea $A \subset H$ acotado y (w_k) sucesión en $\overline{T(A)}$. Para cada $k \in \mathbb{N}$, existe (v_j^k) sucesión en A tal que $\lim_{j\to\infty} \|T(v_j^k) - w_k\|_K = 0$. Así pues, para $k = 1$ existe $j_1 \in \mathbb{N}$ tal que $\|T(v_{j_1}^1) - w_1\|_K \leq 1$. Luego, para $k = 2$ elegimos $j_2 \in \mathbb{N}$ con $j_2 > j_1$ tal que $\|T(v_{j_2}^2) - w_2\|_K \leq \frac{1}{2}$. Continuando de este modo obtenemos, para cada $k \in \mathbb{N}$, un $j_k \in \mathbb{N}$ con $j_k > j_{k-1}$ tal que $\|T(v_{j_k}^k) - w_k\|_K \leq \frac{1}{k}$. Definamos $v_k := v_{j_k}^k$ para cada $k \in \mathbb{N}$. Entonces (v_k) es una sucesión de elementos de A de modo que (v_k) es acotada en H. Como T es un operador compacto, existe (v_{k_ℓ}) subsucesión de (v_k) tal que $T(v_{k_\ell}) \to w$ en K cuando $\ell \to \infty$. Se sigue entonces que $v \in \overline{T(A)}$ y obtenemos que:

$$\|w_{k_\ell} - w\|_K \leq \|w_{k_\ell} - T(v_{k_\ell})\|_K + \|T(v_{k_\ell}) - w\|_K \leq \frac{1}{k_\ell} + \|T(v_{k_\ell}) - w\|_K \qquad \forall\, \ell \in \mathbb{N}.$$

En consecuencia, $w_{k_\ell} \to w$ cuando $\ell \to \infty$. Es decir, (w_k) tiene una subcesión convergente lo que implica que $\overline{T(A)}$ es compacto en K.

$\Leftarrow)$: Sea (v_k) una sucesión acotada en H. Entonces, el subconjunto $\mathfrak{B} = \{v_k : k \in \mathbb{N}\}$ de V es acotado y, por hipótesis, $\overline{T(\mathfrak{B})}$ es compacto en K. Dado que $T(v_k) \in T(\mathfrak{B}) \subset \overline{T(\mathfrak{B})}$ para cada $k \in \mathbb{N}$, entonces la sucesión $(T(v_k))$ tiene una subsucesión convergente lo que demuestra que T es compacto.

(b) Para $k \in \mathbb{N}$ consideremos la bola abierta $B_H(0_H, k)$ en H. Es claro que $B_H(0_H, k)$ es un conjunto acotado en H para cada $k \in \mathbb{N}$. Para cada $k \in \mathbb{N}$, denotemos por $\mathfrak{B}_k := T(B_H(0_H, k))$. Dado que T es un operador compacto, del inciso *(a)* concluimos que $\overline{\mathfrak{B}_k}$ es compacto en K para cada $k \in \mathbb{N}$ y, por tanto, que $\overline{\mathfrak{B}_k}$ es separable en K para cada $k \in \mathbb{N}$. Luego, como $\mathfrak{B}_k \subset \overline{\mathfrak{B}_k}$ entonces $\mathfrak{B}_k$ es separable en K para cada $k \in \mathbb{N}$.

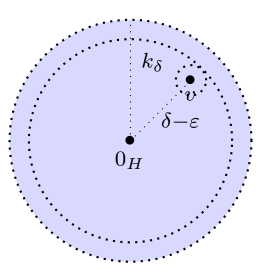


Figura 3.1: $B_H(v,\varepsilon) \subset B_H(0_H,\delta) \subset B_H(0_H,k_\delta)$

Observemos que, si $v \in H$ entonces existe $\varepsilon > 0$ tal que $B_H(v,\varepsilon) \subset H$. Definiendo $\delta := \|v\|_H + \varepsilon > 0$ se sigue que $B_H(v,\varepsilon) \subset B_H(0_H,\delta)$. En efecto, si $u \in B_H(v,\varepsilon)$ entonces

$\|u\|_H - \|v\|_H \leq \|u-v\|_H < \varepsilon$ y, por tanto, $u \in B_H(0_H, \delta)$. Escojamos $k_\delta \in \mathbb{N}$ tal que $\delta < k_\delta$. Así pues, $B_H(v, \varepsilon) \subset B_H(0_H, \delta) \subset B_H(0_H, k_\delta)$ lo que demuestra que $v \in \cup_{k=1}^{\infty} B_H(0_H, k)$ y, por tanto, $H = \cup_{k=1}^{\infty} B_H(0_H, k)$.

En consecuencia, $\operatorname{im} T = \cup_{k=1}^{\infty} \mathfrak{B}_k$.

Sea $k \in \mathbb{N}$. Dado que $\mathfrak{B}_k$ es separable, existe $\mathfrak{D}_k$ subconjunto denso numerable de $\mathfrak{B}_k$. Definiendo $\mathfrak{D} := \cup_{k=1}^{\infty} \mathfrak{D}_k$ es claro que es un subconjunto numerable de $\operatorname{im} T$. Luego, para cada $j \in \mathbb{N}$ se sigue que $\mathfrak{D}_j \subset \mathfrak{D}$ de modo que $\overline{\mathfrak{D}_j} = \mathfrak{B}_j \subset \overline{\mathfrak{D}}$ y, en consecuencia, $\overline{\mathfrak{D}} = \operatorname{im} T$ lo que significa que $\mathfrak{D}$ es denso en $\operatorname{im} T$. Por tanto, $\operatorname{im} T$ es separable. Luego, $\mathfrak{D}$ es también un subconjunto denso de $\overline{\operatorname{im} T}$ de modo que $\overline{\operatorname{im} T}$ es separable.

Esto concluye la demostración. ■

Veamos algunos ejemplos de operadores compactos.

Ejemplo 3.10 *La identidad $id : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ no es un operador compacto.*

Demostración: Es suficiente probar que la bola cerrada $\bar{B}_{\ell_2(\mathbb{R})}(0,1) := \{(x_k) \in \ell_2(\mathbb{R}) : \sum_{k=1}^{\infty} x_k^2 \leq 1\}$ no es compacta en $\ell_2(\mathbb{R})$.

Sea $\bar{e}_k := (e_{k,j})$ la sucesión cuyos términos son $e_{k,k} = 1$ y $e_{k,j} = 0$ si $k \neq j$. Es claro que $\bar{e}_k \in \bar{B}_{\ell_2(\mathbb{R})}(0,1)$ para cada $k \in \mathbb{N}$. Notemos que:

$$\|\bar{e}_k - \bar{e}_j\|_2 = \sqrt{2} \qquad \forall k \neq j.$$

Argumentando por contradicción, supongamos que existe una subsucesión $(\bar{e}_{k_j})$ que converge a e en $\ell_2(\mathbb{R})$. Entonces, para $\varepsilon_0 := \frac{\sqrt{2}}{2}$ existe $j_0 \in \mathbb{N}$ tal que:

$$\|\bar{e}_{k_j} - e\|_2 < \varepsilon_0 \qquad \forall\, j \geq j_0.$$

En consecuencia,

$$\|\bar{e}_{k_j} - \bar{e}_{k_i}\|_2 \leq \|\bar{e}_{k_j} - e\|_2 + \|e - \bar{e}_{k_i}\|_2 < \tfrac{\sqrt{2}}{2} + \tfrac{\sqrt{2}}{2} = \sqrt{2} \qquad \forall\, i, j \geq j_0$$

lo cual contradice nuestra suposición. Por tanto, $(\bar{e}_k)$ no contiene ninguna subsucesión convergente por lo que $\bar{B}_{\ell_2(\mathbb{R})}(0,1)$ no es compacto. ■

Ejemplo 3.11 *La función $T : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ dada por $T(\bar{x}) := \left(\frac{x_j}{j}\right)$ con $\bar{x} = (x_j)$ es un operador compacto.*

Demostración: La función T es claramente lineal. Notemos que para todo $\bar{x} = (x_j) \in \ell_2$

$$\|T(\bar{x})\|_2 = \left\| \left(\frac{x_j}{j} \right) \right\|_2 = \left(\sum_{j=1}^{\infty} \frac{x_j^2}{j^2} \right)^{1/2} \leq \left(\sum_{j=1}^{\infty} x_j^2 \right)^{1/2} = \|\bar{x}\|_2$$

lo que implica que T es continuo.

Definamos la sucesión de funciones $T_k : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ como $T_k(\bar{x}) := (y_{k,j})$ en donde $y_{k,j} := \frac{x}{k}$ si $j \leq k$ y $y_{k,j} := 0$ si $k < j$ con $\bar{x} = (x_j)$. Es inmediato que T_k es lineal y continua para cada $k \in \mathbb{N}$.

Fijemos $k \in \mathbb{N}$ y consideremos el subespacio vectorial $\operatorname{im} T_k$ de $\ell_2(\mathbb{R})$. Sea $\bar{y} = (y_j) \in \operatorname{im} T_k$ entonces existe $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ tal que $y_j = \frac{x_j}{j}$ si $j \leq k$ y $y_j = 0$ si $k < j$. Definamos $\iota : \operatorname{im} T_k \to \mathbb{R}^k$ como $\iota(\bar{y}) := (x_1, x_2, \ldots, x_k)$. Es claro que ι es lineal e inyectiva. Sea $(x_1, \ldots, x_k) \in \mathbb{R}^k$. Consideremos la sucesión $\bar{\varsigma} = (\varsigma_j)$ dada por $\varsigma_j := jx_j$ si $j \leq k$ y $\varsigma_j := 0$ si $k < j$. Es claro que $\bar{\varsigma} \in \ell_2(\mathbb{R})$ y $T_k(\bar{\varsigma})$ es la sucesión igual a x_j si $j \leq k$ y 0 si $k < j$. Por tanto, $\iota(T_k(\bar{\varsigma})) = (x_1, \ldots, x_k)$ lo que demuestra que ι es suprayectiva. En consecuencia, $\dim \operatorname{im} T_k = k < \infty$ por lo que T_k es un operador de rango finito. La Proposición 3.7 asegura que T_k es compacto.

Así pues, (T_k) es una sucesión de operadores compactos. Observemos que

$$\|T_k(\bar{x}) - T(\bar{x})\|_2^2 = \sum_{j=k+1}^{\infty} \frac{x_j^2}{j^2} \leq \frac{1}{(k+1)^2} \sum_{j=k+1} x_j^2 \leq \frac{\|\bar{x}\|_2^2}{(k+1)^2} \qquad \forall\, \bar{x} = (x_j) \in \ell_2(\mathbb{R})$$

y, en consecuencia $\|T_k - T\|_{\mathcal{L}_c(\ell_2(\mathbb{R}),\ell_2(\mathbb{R}))} \leq \frac{1}{k+1}$ para cada $k \in \mathbb{N}$. El Teorema 3.4 asegura que T es compacto. $\blacksquare$

Sea S un conjunto y V espacio de Banach sobre $\mathbb{R}$ con norma $\|\cdot\|_V$. El conjunto

$$\mathcal{B}(S,V) := \{f : S \to V \ : \ f \ \text{ es acotada}\}$$

es un espacio de Banach con la norma $\|f\|_\infty := \sup_{z \in S} \|f(z)\|_V$. De hecho, una sucesión (f_k) en $\mathcal{B}(S,V)$ converge a f en $\mathcal{B}(S,V)$ si y sólo si $f_k \to f$ uniformemente en S.

Notemos que, si $T : H \to K$ es una función lineal y continua, entonces T es acotada en H de modo que $T \in \mathcal{B}(H,K)$. Es decir, $\mathcal{L}_c(H,K)$ es un subespacio vectorial del espacio de funciones acotadas $\mathcal{B}(H,K)$. El Teorema 3.4 asegura entonces que $\mathcal{K}(H,K)$ es un subespacio vectorial de $\mathcal{B}(H,K)$. Así pues, podemos concluir que el límite uniforme de operadores compactos es compacto.

Veamos en el siguiente ejemplo que, el límite puntual de operadores compactos no es compacto en general.

Ejemplo 3.12 *Sea $T_k : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ y $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ la función dada por $T_k(\bar{x}) := x_k$ si $j \leq k$ y $T(\bar{x}) := 0$ si $k < j$. Es claro que T_k es lineal y continuo. Además, siguiendo un procedimiento análoga al ejemplo anterior podemos concluir que* $\dim \operatorname{im} T_k = k < \infty$ *por lo que T_k es un operador compacto para cada $k \in \mathbb{N}$.*

La sucesión (T_k) converge puntualmente a la función identidad id pero del Ejemplo 3.11 *se sigue que id no es compacto en $\ell_2(\mathbb{R})$.*

3.2. Operadores compactos-autoadjuntos

Sea H un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|_H$.

Observemos lo siguiente: Sea $T : H \to H$ una función lineal y continua. Para $v \in H$ arbitrario pero fijo consideremos la función $T_v : H \to \mathbb{R}$ dada por $T_v(u) := \langle T(u), v \rangle$. La Proposición 1.31 asegura que $T_v : H \to \mathbb{R}$ es lineal y continua. De este modo, el teorema de representación de Fréchet-Riesz asegura que existe un único $\tau_v \in H$ tal que:

$$T_v(u) = \langle u, \tau_v \rangle \qquad \forall u \in H,$$

es decir,

$$\langle T(u), v \rangle = \langle u, \tau_v \rangle \qquad \forall u \in H.$$

Definamos la función $T_H^* : H \to H$ como:

$$T_H^*(v) := \tau_v \tag{3.1}$$

Es claro del teorema de representación de Fréchet-Riesz que T_H^* es la única función que satisface que:

$$\langle T(u), v \rangle = \langle u, T_H^*(v) \rangle \qquad \forall\, u, v \in H.$$

Ahora, consideremos $v, w \in H$ y $\alpha, \beta \in \mathbb{R}$. Entonces:

$$\begin{aligned}
\langle u, T_H^*(\alpha\, v + \beta\, w) \rangle &= \langle T(u), \alpha\, v + \beta\, w \rangle = \langle T(u), \alpha v \rangle + \langle T(u), \beta\, w \rangle \\
&= \alpha \langle T(u), v \rangle + \beta \langle T(u), w \rangle = \alpha \langle u, T_H^*(v) \rangle + \beta \langle u, T_H^*(w) \rangle \\
&= \langle u, \alpha T_H^*(v) \rangle + \langle u, \beta T_H^*(w) \rangle \\
&= \langle u, \alpha T_H^*(v) + \beta T_H^*(w) \rangle
\end{aligned}$$

por lo que $T_H^*(\alpha\, v + \beta\, w) = \alpha T_H^*(v) + \beta T_H^*(w)$. Es decir, T_H^* es lineal.

Luego, sea $v \in H$. Si $T_H^*(v) = 0_H$ entonces $\|T_H^*(v)\|_H = 0$ y, por tanto, $\|T_H^*(v)\|_H \leq c\, \|v\|_H$ para cada $c \in \mathbb{R}$ con $c \geq 0$. Supongamos entonces que $T_H^*(v) \neq 0_H$. Dado que T es lineal y continua, entonces:

$$\|T_H^*(v)\|_H^2 = |\langle T_H^*(v), T_H^*(v) \rangle| = |\langle T(T_H^*(v)), v \rangle| \leq \|T\|_{\mathcal{L}_c(H,H)} \|T_H^*(v)\|_H \|v\|_H.$$

En consecuencia, $\|T_H^*(v)\|_H \leq \|T\|_{\mathcal{L}_c(H,H)} \|v\|_H$ para todo $v \in H$ lo que implica que T_H^* es continua.

Definición 3.13 *Sea $T : H \to H$ una función lineal y continua. La función $T_H^* : H \to H$ que cumple que*

$$\langle T(u), v \rangle = \langle u, T_H^*(v) \rangle \qquad \forall\, u, v \in H$$

se llama ***operador adjunto de Hilbert de T.***

Por ejemplo, si denotamos por id_H a la función identidad en H, dado que $id_H(u) = u$ para todo $u \in H$, entonces $\langle id(u), v\rangle = \langle u, v\rangle = \langle u, id_H^*(v)\rangle$ para todo $u, v \in H$. Por tanto, $id_H^*(v) = v$ para todo $v \in H$ y concluimos que $id_H = id_H^*$.

Veamos algunas propiedades de los operadores adjuntos de Hilbert.

Teorema 3.14 *Sean $T, S : H \to H$ funciones lineal y continuas y $\alpha \in \mathbb{R}$. Si $T_H^*, S_H^* : H \to H$ son los operadores adjuntos de Hilbert de T y S respectivamente, entonces su cumple lo siguiente:*

(a) $T_H^* : H \to H$ *es lineal y continua y* $\|T^*\|_{\mathcal{L}_c(H,H)} = \|T\|_{\mathcal{L}_c(H,H)}$.
(b) $(S+T)_H^* = S_H^* + T_H^*$.
(c) $(\alpha T)_H^* = \alpha\, T_H^*$.
(d) $(T_H^*)_H^* = T$.
(e) $\ker T = (\operatorname{im} T_H^*)^\perp$.
(f) $\overline{\operatorname{im} T} = (\ker T_H^*)^\perp$.
(g) $(S \circ T)_H^* = T_H^* \circ S_H^*$.
(h) *Si* $T : H \to H$ *es un isomorfismo de Banach, entonces* $T_H^* : H \to H$ *es un isomorfismo de Banach y se cumple que* $(T_H^*)^{-1} = (T^{-1})_H^*$.

Demostración: (a) Ya vimos que T_H^* es lineal y continua. Observemos lo siguiente: Sean $u, v \in H$ con $\|u\|_H = \|v\|_H = 1$. De la desigualdad de Cauchy-Schwarz se sigue que:

$$|\langle T(u), v\rangle| \leq \|T(u)\|_H \|v\|_H,$$

y, dado que T es lineal y continua concluimos que:

$$|\langle T(u), v\rangle| \leq \|T(u)\|_H \|v\|_H \leq \|T\|_{\mathcal{L}_c(H,H)} \|u\|_H \|v\|_H = \|T\|_{\mathcal{L}_c(H,H)}.$$

Así pues, $\sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\} \leq \|T\|_{\mathcal{L}_c(H,H)}$.

Si $T(u) = 0_H$ es inmediato que $\|T(u)\|_H \leq \sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\}$. Supongamos entonces que $T(u) \neq 0_H$. Como $\|T(u)\|_H^2 = \langle T(u), T(u)\rangle$ entonces:

$$\|T(u)\|_H = \left\langle T(u), \frac{T(u)}{\|T(u)\|_H} \right\rangle \leq \sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\}$$

y, en consecuencia $\|T\|_{\mathcal{L}_c(H,H)} \leq \sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\}$.

Por tanto, $\|T\|_{\mathcal{L}_c(H,H)} = \sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\}$.

De lo anterior obtenemos que:

$$\begin{aligned}
\|T\|_{\mathcal{L}_c(H,H)} &= \sup\{|\langle T(u), v\rangle| \,:\, \|u\|_H = \|v\|_H = 1\} \\
&= \sup\{|\langle u, T_H^*(v)\rangle| \,:\, \|u\|_H = \|v\|_H = 1\} \\
&= \sup\{|\langle T_H^*(v), u\rangle| \,:\, \|u\|_H = \|v\|_H = 1\} \\
&= \|T_H^*\|_{\mathcal{L}_c(H,H)}
\end{aligned}$$

como afirma el enunciado.

(b) Sean $u, v \in H$, entonces:

$$\begin{aligned}\langle u, (S+T)^*_H(v)\rangle &= \langle (S+T)(u), v\rangle = \langle S(u)+T(u), v\rangle = \langle S(u), v\rangle + \langle T(u), v\rangle \\ &= \langle u, S^*_H(v)\rangle + \langle u, T^*_H(v)\rangle = \langle u, S^*_H(v)+T^*_H(v)\rangle = \langle u, (S^*_H+T^*_H)(v)\rangle\end{aligned}$$

y, en consecuencia $(S+T)^*_H(v) = (S^*_H + T^*_H)(v)$.

(c) Sean $u, v \in H$, entonces:

$$\begin{aligned}\langle u, (\alpha\, T)^*_H(v)\rangle &= \langle (\alpha T)(u), v\rangle = \langle \alpha\, T(u), v\rangle = \alpha\langle T(u), v\rangle \\ &= \alpha\langle u, T^*_H(v)\rangle = \langle u, \alpha\, T^*_H(v)\rangle = \langle u, (\alpha\, T^*_H)(v)\rangle\end{aligned}$$

y, por tanto, $(\alpha T)^*_H(v) = \alpha\, T^*_H(v)$.

(d) Sean $u, v \in H$, entonces:

$$\langle (T^*_H)^*_H(u), v\rangle = \langle v, (T^*_H)^*_H(u)\rangle = \langle T^*_H(u), v\rangle = \langle v, T^*_H(u)\rangle = \langle T(u), v\rangle$$

por lo tanto $(T^*_H)^*_H(u) = T(u)$.

(e) Sean $u \in \ker T$ y $w \in \operatorname{im} T^*_H$. Existe $v \in H$ tal que $T^*_H(v) = w$. Notemos que:

$$\langle u, w\rangle = \langle u, T^*_H(v)\rangle = \langle T(u), v\rangle = \langle 0_H, v\rangle = 0$$

por lo que $u \in (\operatorname{im} T^*_H)^\perp$. Inversamente, si $u \in (\operatorname{im} T^*_H)^\perp$ entonces $\langle u, T^*_H(v)\rangle = 0$ para cada $v \in H$. En consecuencia, $\langle T(u), v\rangle = 0$ para cada $v \in H$ y, por tanto, $T(u) = 0_H$. Es decir, $u \in \ker T$.

(f) De los incisos *(d)* y *(e)* podemos concluir que:

$$\ker T^*_H = (\operatorname{im} T)^\perp.$$

Aplicando la Proposición 1.20 y el Corolario 1.29 obtenemos que:

$$\overline{\operatorname{im} T} = \left(\left(\overline{\operatorname{im} T}\right)^\perp\right)^\perp = \left((\operatorname{im} T)^\perp\right)^\perp = (\ker T^*_H)^\perp$$

como afirma el enunciado.

(g) Para cualesquiera $u, v \in H$ se tiene que:

$$\langle u, (S\circ T)^*_H(v)\rangle = \langle (S\circ T)(u), v\rangle = \langle S(T(u)), v\rangle = \langle T(u), S^*_H(v)\rangle = \langle u, T^*_H(S^*_H(v))\rangle\,.$$

En consecuencia, $(S\circ T)^*_H(v) = (T^*_H \circ S^*_H)(v)$ para cada $v \in H$ y esto prueba el resultado.

(h) Dado que $T \in \mathcal{L}_c(H, H)$ es un isomorfismo de Banach, existe $T^{-1} \in \mathcal{L}_c(H, H)$ tal que:

$$T^{-1}\circ T = T^{-1}\circ T = id_H.$$

Así pues, del inciso *(g)* se sigue que:

$$id_H = id_H^* = \left(T^{-1} \circ T\right)_H^* = T_H^* \circ (T^{-1})_H^*,$$
$$id_H = id_H^* = \left(T \circ T^{-1}\right)_H^* = (T^{-1})_H^* \circ T_H^*.$$

Es decir, $(T^{-1})_H^*$ es un inverso izquiero y derecho de T_H^* lo que implica que T_H^* es biyectivo y $(T_H^*)^{-1} = (T^{-1})_H^*$. Luego, como $T^{-1} \in \mathcal{L}_c(H, H)$ entonces por el inciso *(a)* se tiene que $(T^{-1})_H^* \in \mathcal{L}_c(H, H)$ lo que prueba que T^* es un isomorfismo de Banach. ■

Por simplicidad y siempre que no se preste a confusión escribiremos $T^* : H \to H$ para referirnos al operador adjunto de Hilbert de una función lineal y continua $T : H \to H$. Del mismo modo, diremos únicamente operador adjunto de T.

En la siguiente observación extendemos el concepto del operador adjunto de Hilbert.

Observación 3.15 *Sean H y K espacios de Hilbert sobre $\mathbb{R}$ con productos escalares $\langle \cdot, \cdot \rangle_H$ y $\langle \cdot, \cdot \rangle_K$ respectivamente. Consideremos una función lineal y continua $T : H \to K$.*

Para $z \in K$ arbitrario pero fijo, definimos $T_z : H \to \mathbb{R}$ como $T_z(u) := \langle T(u), z \rangle_K$. Esta función es claramente lineal y continua por la Proposición 1.41. *El teorema de representación de Fréchet-Riesz asegura que existe un único $z^* \in H$ tal que $\langle T(u), z \rangle_K = \langle u, z^* \rangle_H$ para todo $u \in H$. Esto nos permite definir el operador $T^* : K \to H$ como:*

$$T^*(z) := z^*$$

que satisface que $\langle T(u), z \rangle_K = \langle u, T^*(z) \rangle_H$ para cualesquiera $u \in H$ y $z \in K$. La función T^* definida se conoce como operador adjunto de Hilbert de T.

Las propiedades dadas en el Teorema 3.14 *permanecen válidas, salvo algunas modificaciones para $T^* : K \to H$.*

Definición 3.16 *Una función lineal y continua $T : H \to H$ se llama* ***operador autoadjunto*** *si $T = T^*$. Es decir, si*

$$\langle T(u), v \rangle = \langle u, T(v) \rangle \qquad \forall\, u, v \in H.$$

Por ejemplo, si $0(u) := 0_H$ para todo $u \in H$ entonces $\langle 0(u), v \rangle = 0$ para todo $u, v \in H$. Así pues, $\langle u, 0^*(v) \rangle = 0$ para todo $u, v \in H$ de modo que $0^*(v) = 0$ para todo $v \in H$. En consecuencia $0 = 0^*$.

Consideremos a la función $T : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ dada por $T(\bar{x}) := \left(\frac{x_j}{j}\right)$ con $\bar{x} = (x_j)$ vista en el Ejemplo 3.12. Sean $\bar{x} = (x_j), \bar{y} = (y_j) \in \ell_2(\mathbb{R})$, entonces:

$$\langle T(\bar{x}), \bar{y} \rangle_{\ell_2(\mathbb{R})} = \sum_{j=1}^{\infty} \frac{x_j}{j}\, y_j = \sum_{j=1}^{\infty} x_j\, \frac{y_j}{j} = \langle \bar{x}, T(\bar{y}) \rangle_{\ell_2(\mathbb{R})}.$$

Por tanto, T es un operador compacto autoadjunto.

Notemos que si $T : H \to H$ es una función lineal y continua, entonces el conjunto $\overline{\operatorname{im} T}$ es un subespacio vectorial cerrado de H y por el teorema del complemento ortogonal (ver Teorema 1.28) concluimos que $H = \overline{\operatorname{im} T} \oplus \left(\overline{\operatorname{im} T}\right)^{\perp}$. Entonces, si T es un operador autoadjunto, de los inicisos *(d)* y *(e)* del Teorema 3.14 se tiene que $\left(\overline{\operatorname{im} T}\right)^{\perp} = \ker T$ y, por tanto, $H = \overline{\operatorname{im} T} \oplus \ker T$.

Veamos algunas propiedades de los operadores autoadjuntos.

Proposición 3.17 *Sean $S, T \in \mathcal{L}_c(H, H)$ tales que S y T son operadores autoadjuntos. Entonces, $S \circ T$ es autoadjunto si y sólo si $S \circ T = T \circ S$.*

Demostración: El Teorema 3.14 asegura que $(S \circ T)^* = T^* \circ S^*$. Por tanto, $S \circ T = (S \circ T)^* = T^* \circ S^* = T \circ S$ si y sólo si $T \circ S = S \circ T$ como afirma el enunciado. ■

Observemos que, si $S, T : H \to H$ son operadores autoadjuntos y $\alpha \in \mathbb{R}$, del Teorema 3.14 se sigue que $(T + S)^* = T^* + S^* = T + S$ y $(\alpha T)^* = \alpha T^* = \alpha T$. Además, la función constante igual a 0_H en H es un operador autoadjunto. De este modo, el conjunto de todos los operadores autoadjuntos en H espacio de Hilbert es un subespacio vectorial de $\mathcal{L}_c(H, H)$. A continuación vemos que, de hecho, es un subespacio cerrado de $\mathcal{L}_c(H, H)$.

Teorema 3.18 *Sea $T_k : H \to H$ una sucesión de funciones lineales y continuas tal que T_k converge a T en $\mathcal{L}_c(H, H)$, entonces T_k^* converge a T^* en $\mathcal{L}_c(H, H)$.*

Demostración: Sea $\varepsilon > 0$. Existe $k_0 \in \mathbb{N}$ tal que:

$$\|T_k - T\|_{\mathcal{L}_c(H,H)} < \varepsilon \qquad \forall\, k \geq k_0.$$

El Teorema 3.14 asegura que:

$$\|T_k^* - T^*\|_{\mathcal{L}_c(H,H)} = \|(T_k - T)^*\|_{\mathcal{L}_c(H,H)} = \|T_k - T\|_{\mathcal{L}_c(H,H)}$$

y, en consecuencia:

$$\|T_k^* - T^*\|_{\mathcal{L}_c(H,H)} < \varepsilon \qquad \forall\, k \geq k_0$$

lo que demuestra que $T_k^* \to T^*$ en $\mathcal{L}_c(H, H)$. ■

Teorema 3.19 *Sea $T_k : H \to H$ una sucesión de funciones lineales y continuas tales que T_k es autoadjunto para cada $k \in \mathbb{N}$. Supongamos que T_k converge a T en $\mathcal{L}_c(H, H)$, entonces $T : H \to H$ es un operador autoadjunto.*

Demostración: Del mismo modo que en el resultado anterior, el Teorema 3.14 asegura que:
$$\|T_k^* - T^*\|_{\mathcal{L}_c(H,H)} = \|(T_k - T)^*\|_{\mathcal{L}_c(H,H)} = \|T_k - T\|_{\mathcal{L}_c(H,H)}.$$

Por tanto,
$$\|T - T^*\|_{\mathcal{L}_c(H,H)} \leq \|T - T_k\|_{\mathcal{L}_c(H,H)} + \|T_k - T_k^*\|_{\mathcal{L}_c(H,H)} + \|T_k^* - T^*\|_{\mathcal{L}_c(H,H)}$$
usando el teorema anterior y el hecho de que $T_k = T_k^*$ para cada $k \in \mathbb{N}$, obtenemos:
$$\|T - T^*\|_{\mathcal{L}_c(H,H)} \leq 2\|T - T_k\|_{\mathcal{L}_c(H,H)}. \tag{3.2}$$

Tomando el límite cuando $k \to \infty$ en (3.2) concluimos que $\|T - T^*\|_{\mathcal{L}_c(H,H)} = 0$ y, por tanto, $T = T^*$. ■

A continuación demostraremos que cualquier operador compacto entre un espacio de Banach y un espacio de Hilbert es el límite de una sucesión de operadores de rango finito.

Teorema 3.20 *Sea V un espacio de Hilbert sobre $\mathbb{R}$ y $T : V \to H$ una función lineal y continua. Si $T : V \to H$ es compacto, entonces existe una sucesión $T_k : V \to H$ de operadores de rango finito tal que $T_k \to T$ en $\mathcal{L}_c(V, H)$.*

Demostración: Es claro que $\overline{\operatorname{im} T}$ es un subespacio vectorial cerrado de H y, por tanto, es un espacio de Hilbert. Dado que T es compacto, por la Proposición 3.10 se tiene que $\overline{\operatorname{im} T}$ es separable. Por tanto, existe $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ una base de Hilbert de $\overline{\operatorname{im} T}$ (ver Teorema 1.47).

Definiendo el conjunto $V_k := \operatorname{lin}(\{e_1, \ldots, e_k\})$ es claro que V_k es un subespacio vectorial cerrado en $\overline{\operatorname{im} T}$ para cada $k \in \mathbb{N}$. Consideremos $p_{V_k} : \overline{\operatorname{im} T} \to V_k$ la proyección de $\overline{\operatorname{im} T}$ sobre V_k y definamos $T_k : V \to V_k$ como $T_k := p_{V_k} \circ T$. Esta función es claramente lineal y continua. Además, como $\operatorname{im} T_k \subset V_k$ y V_k es de dimensión finita k, entonces $\operatorname{im} T_k$ es de dimensión finita y, por tanto, T_k es un operador de rango finito.

Demostraremos que $\|T_k - T\|_{\mathcal{L}_c(V,H)} \to 0$. Argumentando por contradicción, supongamos que existe $\varepsilon_0 > 0$ tal que:
$$\|T_k - T\|_{\mathcal{L}_c(V,H)} \geq \varepsilon_0 \qquad \forall k \in \mathbb{N}. \tag{3.3}$$

Así pues, existe (x_k) una sucesión en V tal que $\|x_k\|_V = 1$ para cada $k \in \mathbb{N}$ y tal que:
$$\|T_k(x_k) - T(x_k)\|_H \geq \frac{\varepsilon_0}{2} \qquad \forall k \in \mathbb{N}. \tag{3.4}$$

Como T es compacto y (x_k) es acotada en V, existe (x_{k_j}) subsucesión de (x_k) tal que $(T(x_{k_j}))$ converge en H. Sea $w := \lim_{j\to\infty} T(x_{k_j}) \in \overline{\operatorname{im} T}$. Observemos que:
$$\begin{aligned}(T_j - T)(x_{k_j}) &= T_j(x_{k_j}) - T(x_{k_j}) = p_{V_j}(T(x_{k_j})) - T(x_{k_j}) = (p_{V_j} - id_H)(T(x_{k_j}))\\ &= (p_{V_j} - id_H)(T(x_{k_j})) + (p_{V_j} - id_H)(w) - (p_{V_j} - id_H)(w)\\ &= (p_{V_j} - id_H)(w) + (p_{V_j} - id_H)(T(x_{k_j}) - w) \qquad \forall j \in \mathbb{N}.\end{aligned}$$

Usando el Teorema 1.42 obtenemos:

$$(p_{V_j} - id_H)(w) = p_{V_j}(w) - w = \sum_{n=1}^{j} \langle w, e_n \rangle \, e_n - \sum_{m=1}^{\infty} \langle w, e_m \rangle \, e_m = \sum_{n=j+1}^{\infty} \langle w, e_n \rangle \, e_n.$$

Por lo tanto:

$$\begin{aligned}
\|T_j(x_{k_j}) - T(x_{k_j})\|_H &= \big\|(p_{V_j} - id_H)(w) + (p_{V_j} - id_H)(T(x_{k_j}) - w)\big\|_H \\
&= \big\|(p_{V_j} - id_H)(w)\big\|_H + \big\|p_{V_j}(T(x_{k_j}) - w)\big\|_H + \big\|T(x_{k_j}) - w\big\|_H \\
&\leq \left(\sum_{n=j+1}^{\infty} |\langle w, e_n \rangle|^2 \right)^{1/2} + 2\,\big\|T(x_{k_j}) - w\big\|_H \qquad \forall j \in \mathbb{N}
\end{aligned}$$

y, en consecuencia:

$$0 < \frac{\varepsilon_0}{2} \leq \|T_j(x_{k_j}) - T(x_{k_j})\|_H \leq 0$$

lo cual es imposible. Esto concluye la demostración. ■

Lema 3.21 *Sea $T : H \to H$ función lineal y continua. Entonces, T es de rango finito si y sólo si T^* es de rango finito.*

Demostración: Supongamos que $\dim \operatorname{im} T < \infty$. Dado que T^* es lineal y continua, entonces por el Lema 1.32 se tiene que $\ker T^*$ es un subespacio vectorial cerrado de H. El teorema del complemento ortogonal (ver Teorema 1.33) asegura que $H = \ker T^* \oplus (\ker T^*)^{\perp}$ y, por el Teorema 3.14, $(\ker T^*)^{\perp} = \overline{\operatorname{im} T} = \operatorname{im} T$ pues es de dimensión finita. Por tanto, $H = \ker T^* \oplus \operatorname{im} T$. Así, para $u \in H$ dada existen únicos $v \in \ker T^*$ y $w \in \operatorname{im} T$ tales que $u = v + w$. Entonces, $T^*(u) = T^*(v + w) = T^*(v) + T^*(w) = T^*(w)$ de modo que $\operatorname{im} T^* \subset T^*(\operatorname{im} T)$ y, por tanto, $\dim \operatorname{im} T^* \leq \dim \operatorname{im} T < \infty$.

Inversamente: si $\dim \operatorname{im} T^* < \infty$, el resultado se sigue directamente de aplicar la discusión anterior a T^* y usando el hecho de que $(T^*)^* = T$. ■

Teorema 3.22 (de Schauder en espacios de Hilbert) *Sea $T : H \to H$ función lineal y continua. Entonces, T es compacto si y sólo si T^* es compacto.*

Demostración: Supongamos que T es compacto. El Teorema 3.20 asegura que existe una sucesión de operadores de rango finito $T_k : H \to H$ tal que $T_k \to T$ en $\mathcal{L}_c(H, H)$. Luego, $T_k^* : H \to H$ es de rango finito por el Lema 3.21 y $T_k^* \to T^*$ en $\mathcal{L}_c(H, H)$ por el Teorema 3.18. El Corolario 3.9 asegura que T^* es compacto.

Inversamente: si T^* es compacto entonces de la discusión anterior se sigue que $(T^*)^*$ es compacto. El Teorema 3.14 implica que $(T^*)^* = T$ y, por tanto, que T es compacto. ■

Una versión más general del teorema de Schauder para un operador compacto entre dos espacios de Banach puede encontrarse, por ejemplo, en [13] y [75].

3.3. Los teoremas de Riesz y Fredholm

En esta sección estudiamos algunos resultados importantes con respecto a la bola unitaria cerrada en espacios de dimensión finita e infinita. Además, establecemos algunas propiedades de los operadores compactos.

Lema 3.23 (de Riesz) *Sea $V = (V, \|\cdot\|_V)$ un espacio normado. Si W es un subespacio vectorial cerrado de V y $W \neq V$ entonces, para cada $\delta \in (0,1)$, existe $v_\delta \in V$ con $\|v_\delta\|_V = 1$ tal que:*

$$\|v_\delta - w\|_V \geq \delta \qquad \forall\, w \in W.$$

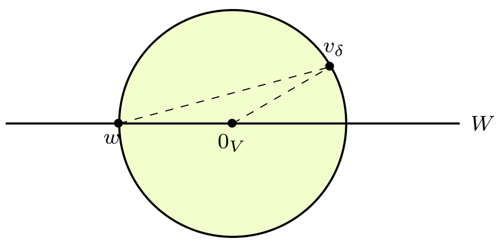


Figura 3.2: Lema de Riesz

Demostración: Sea $v_0 \in V \setminus W$ fijo. Dado que W es cerrado, entonces $\|v_0 - w\|_V > 0$ para cada $w \in W$. Definamos $\rho_0 := \inf_{w \in W} \|v_0 - w\|_V > 0$ y sea $\delta \in (0,1)$. Escojamos $w_\delta \in W$ tal que

$$\rho_0 \leq \|v_0 - w_\delta\|_V < \frac{\rho_0}{\delta}.$$

Definiendo $v_\delta := \frac{v_0 - w_\delta}{\|v_0 - w_\delta\|_V} \in V$ es claro que $\|v_\delta\|_V = 1$. Por tanto, para cada $w \in W$ se tiene que:

$$\|v_\delta - w\|_V = \left\| \frac{v_0 - w_\delta}{\|v_0 - w_\delta\|_V} - w \right\|_V = \left\| \frac{v_0 - w_\delta - w\|v_0 - w_\delta\|_V}{\|v_0 - w_\delta\|_V} \right\|_V \geq \frac{\rho_0}{\|v_0 - w_\delta\|_V} \geq \delta$$

y esto prueba el resultado. ■

Teorema 3.24 (de Riesz) *Sea $V = (V, \|\cdot\|_V)$ un espacio normado y $\mathbb{S}_V := \{v \in V : \|v\|_V = 1\}$ la esfera unitaria. Si $\mathbb{S}_V$ es compacto, entonces* $\dim V < \infty$.

Demostración: Fijemos $\delta \in (0,1)$. Es claro que $\mathfrak{C} := \{B_V(z, \frac{\delta}{2}) : z \in \mathbb{S}_V\}$ es una cubierta abierta de $\mathbb{S}_V$. Como $\mathbb{S}_V$ es compacto, existen $z_1, \ldots, z_m \in \mathbb{S}_V$ tales que:

$$\mathbb{S}_V \subset B_V(z_1, \tfrac{\delta}{2}) \cup \cdots \cup B_V(z_m, \tfrac{\delta}{2}). \tag{3.5}$$

Definamos $W := \text{lin}\,(\{z_1, \ldots, z_m\})$. Entonces W es un subespacio vectorial de dimensión finita que es además cerrado en V. Supongamos que $W \neq V$. El lema de Riesz (ver Lema 3.23) asegura que existe $v_\delta \in \mathbb{S}_V$ tal que $\|v_\delta - w\|_V \geq \delta$ para cada $w \in W$. Por tanto, para cada $j = 1, \ldots, m$ se tiene que $\|v_\delta - z_j\|_V \geq \delta$ lo cual es imposible pues contradice (3.5). En consecuencia, $V = W$ y se sigue que V es de dimensión finita. ∎

A continuación probaremos que un espacio vectorial normado tiene dimensión finita si su bola unitaria cerrada es compacta. El procedimiento que seguiremos puede aplicarse, salvo algunas modificaciones, como una prueba alternativa al teorema de Riesz.

Teorema 3.25 *Sea* $V = (V, \|\cdot\|_V)$ *un espacio normado y* $B_V := \bar{B}_V(0_V, 1) = \{v \in V : \|v\|_V \leq 1\}$ *la bola unitaria cerrada. Si* $\bar{B}_V(0_V, 1)$ *es compacta, entonces* $\dim V < \infty$.

Demostración: Supongamos que $\dim V = \infty$ y fijemos $\delta \in (0, 1)$.

Sea $v_1 \in B_V \setminus \{0_H\}$. Es claro que $W_1 := \text{lin}(\{v_1\})$ es un subespacio vectorial cerrado de dimensión finita y, por tanto, $W_1 \neq V$. El lema de Riesz (ver Lema 3.23) asegura que existe $v_2 \in V$ tal que $\|v_2\|_V = 1$ y $\|v_2 - v_1\|_V \geq \delta$. El conjunto $W_2 := \text{lin}\,(\{v_1, v_2\})$ es también un subespacio vectorial cerrado de dimensión finita y, por tanto, $W_2 \neq V$. El lema de Riesz asegura que existe $v_3 \in B_V$ tal que $\|v_3 - v_j\|_V \geq \delta$ para cada $j = 1, 2$. Continuando de este modo obtenemos, para cada $k \in \mathbb{N}$, un subespacio vectorial cerrado y dimensión finita $W_k := \text{lin}(\{v_1, \ldots, v_k\})$ tal que $W_k \neq V$ y, por el lema de Riesz existe $v_{k+1} \in B_V$ tal que $\|v_{k+1} - v_j\|_V \geq \delta$ para cada $j < k + 1$. Así pues, definimos la sucesión (v_k) de elementos de B_V. De este modo,

$$\|v_j - v_k\|_V \geq \delta \qquad \forall\, j \neq k.$$

En consecuencia, cualquier subsucesión de (v_k) no es de Cauchy y, por tanto, cualquier subsucesión de (v_k) no converge en B_V. Por tanto, B_V no es compacta en V y esto prueba el resultado. ∎

Observemos que si $V = (V, \|\cdot\|_V)$ es un espacio normado de dimensión finita N entonces existe una equivalencia con $\mathbb{R}^N = (\mathbb{R}^N, \|\cdot\|_1)$. Es decir, existe una función $\phi : V \to \mathbb{R}^N$ que es lineal, Lipschitz continua y biyectiva y su inversa $\phi^{-1} : \mathbb{R}^N \to V$ es lineal y Lipschitz continua. De este modo, $\bar{B}_V(0_V, 1)$ es equivalente al conjunto $\{x \in \mathbb{R}^N : \|x\|_1 \leq 1\}$ que es compacto en $\mathbb{R}^N$. Así pues, $\bar{B}_V(0_V, 1)$ es compacta en V. Acabamos de probar que el hecho de que un espacio normado sea de dimensión finita es condición necesaria y suficiente para ver que la bola unitaria cerrada es compacta. Para los detalles de esta prueba invitamos al lector a consultar por ejemplo [13], [17], [73] y [75].

Notemos además que si H es un espacio de Hilbert de dimensión infinita, en el Teorema 3.25 hemos demostrado que existe una sucesión (v_k) de elementos de $\bar{B}_H(0_H, 1)$ que no contiene ninguna subsucesión convergente. El Teorema 3.3 asegura que la función identidad en H no es un operador compacto.

Sea H es un espacio de Hilbert de dimensión infinita y $T : H \to H$ un operador compacto. Supongamos que T es un isomorfismo de Banach, es decir, existe $T^{-1} : H \to H$ función lineal y continua tal que $T^{-1} \circ T = T \circ T^{-1} = id_H$. Dado que T es un operador compacto y $T^{-1} \in \mathcal{L}_c(H, H)$, de la Proposición 3.6 concluimos que $T^{-1} \circ T$ es un operador compacto. Es decir, id_H es un operador compacto lo cual es imposible por lo comentado en el párrafo anterior. Por tanto, un operador compacto en un espacio de dimensión infinita no puede ser una función biyectiva.

Notación 3.26 *Sea S un conjunto no vacío y $\phi : S \to S$ una función. Denotamos por $\phi^k : S \to S$ a la composición:*

$$\phi^k = \underbrace{\phi \circ \cdots \circ \phi}_{k \text{ veces}} \quad \text{si } k \in \mathbb{N}, \qquad \phi^0 = id_S,$$

en donde $id_S : S \to S$ es la función identidad.

Teorema 3.27 (Alternativa de Fredholm en espacios de Hilbert) *Sea H un espacio de Hilbert con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|_H$. Sean $T : H \to H$ una función lineal y continua, y $T^* : H \to H$ su operador adjunto. Si T es un operador compacto, entonces:*

(a) $\ker(id_H - T)$ *es de dimensión finita.*

(b) $\operatorname{im}(id_H - T)$ *es cerrado. Más aún,* $\operatorname{im}(id_H - T) = (\ker(id_H - T^*))^{\perp}$.

(c) $\ker(id_H - T) = \{0_H\}$ *si y sólo si* $\operatorname{im}(id_H - T) = H$.

(d) $\dim\ker(id_H - T) = \dim\ker(id_H - T^*)$.

Demostración: (a) Sea B_H la bola unitaria cerrada de H y $V := \ker(id_H - T)$. Dado que T es lineal y continua, el Lema 1.32 asegura que V es un subespacio vectorial cerrado de H. Consideremos B_V su bola unitaria cerrada. Si $v \in B_V$, entonces $v \in V$ y $\|v\|_H \leq 1$. Notemos que, $(id_H - T)(v) = v - T(v) = 0_H$ y, por lo tanto, $T(v) = v$. En consecuencia, $B_V \subset T(B_H) \subset \overline{T(B_H)}$. Como T es un operador compacto entonces B_V es compacta. Aplicando el Teorema 3.25 se sigue que $\dim V < \infty$.

(b) Sea $v \in \overline{\operatorname{im}(id_H - T)}$. Existe una sucesión (v_k) tal que $v_k \in \operatorname{im}(id_H - T)$ para cada $k \in \mathbb{N}$ y $v_k \to v$ en H. Así pues, existe $u_k \in H$ tal que $v_k = u_k - T(u_k)$ para cada $k \in \mathbb{N}$.

Consideremos el subespacio vectorial cerrado $W := \ker(id_H - T)$ en H y denotemos por $\rho_k := \inf_{w \in W} \|u_k - w\|_H$ para cada $k \in \mathbb{N}$. Dado que W es de dimensión finita por el inciso anterior, existe $w_k \in W$ tal que $\rho_k = \|u_k - w_k\|_H$ para cada $k \in \mathbb{N}$. Es decir, $w_k = T(w_k)$ para cada $k \in \mathbb{N}$. Por tanto,

$$v_k = u_k - T(u_k) = (u_k - w_k) - T(u_k - w_k) \qquad \forall\, k \in \mathbb{N}. \tag{3.6}$$

Supongamos que la sucesión (ρ_k) no es acotada en $\mathbb{R}$. Entonces, existe (ρ_{k_j}) subsucesión de (ρ_k) tal que $\rho_{k_j} \to \infty$ cuando $j \to \infty$. Definamos $\widetilde{w}_k := \frac{u_k - w_k}{\|u_k - w_k\|_H}$ para cada $k \in \mathbb{N}$ y, por tanto, de (3.6) concluimos que

$$\widetilde{w}_k - T(\widetilde{w}_k) = \frac{1}{\|u_k - w_k\|_H}\left[(u_k - w_k) - T(u_k - w_k)\right] = \frac{v_k}{\|u_k - w_k\|_H} = \frac{v_k}{\rho_k} \qquad \forall\, k \in \mathbb{N}.$$

De este modo, obtenemos que $\widetilde{w}_{k_j} - T(\widetilde{w}_{k_j}) \to 0_H$ cuando $j \to \infty$. O, equivalentemente, $\widetilde{w}_{k_j} \to T(\widetilde{w}_{k_j})$ cuando $j \to \infty$. Así pues, $(\widetilde{w}_{k_j})$ es una sucesión acotada en V y, como T es compacto, existe una subsucesión $(\widetilde{w}_{k_{j_\ell}})$ de $(\widetilde{w}_{k_j})$ tal que $T(\widetilde{w}_{k_{j_\ell}})$ converge a $\widetilde{w} \in H$ cuando $\ell \to \infty$. De la unicidad del límite concluimos que $\widetilde{w}_{k_{j_\ell}} \to \widetilde{w}$ cuando $\ell \to \infty$ y, por tanto, $\widetilde{w} \in W$ pues es cerrado. En consecuencia, $\operatorname{ínf}_{w\in W} \|\widetilde{w}_{k_{j_\ell}} - w\|_H \to 0$ cuando $\ell \to \infty$.

Por otro lado,

$$\begin{aligned}\operatorname*{ínf}_{w\in W} \|\widetilde{w}_k - w\|_H &= \operatorname*{ínf}_{w\in W} \left\| \frac{u_k - w_k}{\|u_k - w_k\|_H} - w \right\|_H = \frac{\operatorname{ínf}_{w\in W} \|u_k - (w_k + w\|u_k - w_k\|_H)\|_H}{\|u_k - w_k\|_H} \\ &= \frac{\operatorname{ínf}_{w\in W} \|u_k - w\|_H}{\|u_k - w_k\|_H} = \frac{\|u_k - w_k\|_H}{\|u_k - w_k\|_H} = 1 \qquad \forall\, k \in \mathbb{N},\end{aligned}$$

lo cual contradice que $\operatorname{ínf}_{w\in W} \|\widetilde{w}_{k_{j_\ell}} - w\|_H \to 0$ cuando $\ell \to \infty$. De este modo, concluimos que la sucesión (ρ_k) es acotada en $\mathbb{R}$. Como T es compacto, entonces existe una subsucesión $(u_{k_j} - w_{k_j})$ de $(u_k - w_k)$ tal que $T(u_{k_j} - w_{k_j})$ converge a x en H. De (3.6) se sigue que $u_{k_j} - w_{k_j} \to v + x$ cuando $j \to \infty$. Definiendo $u := v + x \in H$ es claro que $u - T(u) = v$ y, por tanto, $v \in \operatorname{im}(id_H - T)$.

(c) $\Rightarrow$) : Consideremos $V_1 := \operatorname{im}(id_H - T)$ subespacio vectorial cerrado de H y argumentando por contradicción supongamos que $V_1 \neq H$. Es claro que V_1 es de Banach. Si $w \in T(V_1)$, entonces existe $v \in V_1$ tal que $T(v) = w$. Dado que $v \in V_1$ entonces $v = u - T(u)$ para algún $u \in H$. Así pues, $w = T(v) = T(u) - T(T(u))$ para algún $u \in H$ y por lo tanto $w \in V_1$. Es decir, $T(V_1) \subset V_1$. Por tanto, $T_{|V_1} : V_1 \to H$ es un operador compacto de modo que el subespacio vectorial $V_2 := (id_H - T)(V_1)$ de V_1 es cerrado por el inciso anterior. Más aún, $V_2 \neq V_1$ pues $id_H - T$ es inyectiva por hipótesis.

Continuando de este modo definimos, para cada $k \in \mathbb{N}$, al subespacio vectorial cerrado $V_k := (id_H - T)^k(H)$. Así pues, obtenemos (V_k) una sucesión decreciente de subespacios vectoriales cerrados de H. Fijemos $\delta \in (0,1)$. Por el lema de Riez (ver Lema 3.23), para cada $k \in \mathbb{N}$, existe $v_\delta^k \in V_k$ con $\|v_\delta^k\|_H = 1$ tal que $\|v_\delta^k - w\|_H \geq \delta$ para todo $w \in V_{k+1}$. Definamos pues la sucesión $v_k := v_\delta^k$ en H que es acotada.

Notemos que, para cualesquiera $j, k \in \mathbb{N}$ se tiene:

$$T(v_k) - T(v_j) = -(v_k - T(v_k)) + (v_j - T(v_j)) + (v_k - v_j). \tag{3.7}$$

De este modo, si $j < k$ entonces $V_{k+1} \subset V_k \subset V_{j+1} \subset V_j$ y, por tanto,

$$-(v_k - T(v_k)) + (v_j - T(v_j)) + v_k \in V_{j+1}$$

y, se sigue de (3.7) que $\|T(v_k) - T(v_j)\|_H \geq \operatorname{ínf}_{w\in V_{j+1}} \|v_j - w\| \geq \delta$ lo cual es imposible pues T es un operador compacto y (v_k) es una sucesión acotada en H. Por tanto, $V_1 = H$ como afirma el enunciado.

$\Leftarrow$) : Supongamos que $\operatorname{im}(id_H - T) = H$. Dado que T es un operador compacto, entonces $T^* : H \to H$ es un operador compacto y, así $id_H - T^*$ es un operador compacto (ver Teoremas

3.5 y 3.22). El Teorema 3.14 asegura que $\ker(id_H - T^*) = (\mathrm{im}\,(id_H - T))^\perp = \{0_H\}$. Así pues, de la implicación anterior podemos concluir que $\mathrm{im}\,(id_H - T^*) = H$ y, nuevamente por el Teorema 3.14 tenemos que $\mathrm{im}\,(id_H - T^*)^\perp = \ker(id_H - T) = \{0_H\}$. Esto concluye la prueba.

(d) Sea $d = \dim\ker(id_H - T)$ y $d^* = \dim\ker(id_H - T^*)$. Primero veamos que $d^* \leq d$. Argumentando por contradicción, supongamos que $d < d^*$ y denotemos por $V := \ker(id_H - T)$. Entonces, V es un subespacio vectorial de dimensión finita de H y, por tanto, cerrado.

Por otro lado, como $\mathrm{im}\,(id_H - T) = (\ker(id_H - T^*))^\perp$ es un subespacio vectorial cerrado de H, si denotamos por $W := \mathrm{im}\,(id_H - T)$ entonces su complemento ortogonal en H tiene dimensión finita igual a d^*. De este modo, como $d < d^*$ entonces existe una función lineal $\Lambda : V \to W^\perp$ que es inyectiva y no suprayectiva. Si $p_V : H \to V$ es la proyección ortogonal de H sobre V, entonces $\Lambda \circ p_V : H \to W^\perp$ es un operador compacto pues es de rango finito (ver Proposición 3.7) y, en consecuencia, $S := T + \Lambda \circ p_V$ es un operador compacto. Sea $u \in \ker(id_H - S)$, entonces $u - S(u) = 0_H$. Es decir, $u - T(u) - (\Lambda \circ p_V)(u) = 0_H$ y, por tanto, $u - T(u) = 0_H$ y $\Lambda(p_V(u)) = 0_V$. Esto implica que $u \in V$ y así $\Lambda(u) = 0_V$ por lo que necesariamente $u = 0_H$. En consecuencia, $\ker(id_H - S) = \{0_H\}$ y se sigue del iniciso *(c)* que $\mathrm{im}\,(id_H - S) = H$ lo cual es imposible. En efecto, dado que Λ no es suprayectiva, existe $w \in W^\perp$ tal que $w \notin \mathrm{im}\,\Lambda$. Es decir, no existe $u \in H$ tal que $u - S(u) = w$ y, por tanto, $d^* \leq d$.

Inversamente: aplicando el procedimiento anterior al operador compacto T^* obtenemos que:

$$\dim\ker(id_H - (T^*)^*) \leq \dim\ker(id_H - T^*)$$

y, como $\ker(id_H - T) \subset \ker(id_H - (T^*)^*)$ dado que $T = (T^*)^*$ entonces $d \leq d^*$. Por lo tanto, $d = d^*$ como afirma el enunciado. ■

Observemos que en la prueba anterior, salvo el inciso *(d)*, no usamos propiedades especiales del espacio de Hilbert H más que aquellas que tiene la norma por lo que es posible sustituir a H por cualquier espacio de Banach y la prueba es exactamente la misma. Para ello, invitamos al lector a consultar, por ejemplo, [13] para una versión más general del teorema de Fredholm al considerar un operador compacto $T : V \to V$ con V un espacio de Banach. La prueba dada en [13] del iniciso *(d)* considera una extensión del concepto de complemento ortogonal de un subespacio vectorial cerrado de un espacio de Banach arbitrario.

3.4. Teoría espectral de operadores compactos

Sea $H = (H, \langle\cdot,\cdot\rangle, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$. Consideremos $T : H \to H$ una función lineal y continua. Para cada $\lambda \in \mathbb{R}$, definimos a la función $T_\lambda : H \to H$ como:

$$T_\lambda(v) := (T - \lambda\, id_H)(v)$$

donde $id_H : H \to H$ es la función identidad en H. Es claro que T_λ es una función lineal y continua.

Definición 3.28 *Sea $T : H \to H$ una función lineal y continua. Definimos el* ***conjunto resolvente de*** *T como:*

$$\rho(T) := \{\lambda \in \mathbb{R} \,:\, T_\lambda = (T - \lambda\, id_H) \;\textit{ es biyectiva de } H \textit{ en } H\}\,.$$

Definimos el ***espectro de*** *T que denotamos por $\sigma(T)$ como el complemento del conjunto resolvente de T. Es decir, $\sigma(T) := \mathbb{R} \setminus \rho(T)$.*

Sea $T : H \to H$ lineal y continua. Observemos los siguiente: $\lambda \in \sigma(T)$ si y sólo si $T_\lambda : V \to V$ no es biyectiva si y sólo si T_λ no es inyectiva o T_λ no es suprayectiva. Es decir,

(i) $\ker T_\lambda \neq \{0_H\}$ o

(ii) $\operatorname{im} T_\lambda \neq V$.

Definición 3.29 *Sea $T : H \to H$ una función lineal y continua y $\lambda \in \mathbb{R}$. Decimos que λ* ***es un valor propio de*** *T si $\ker T_\lambda \neq \{0_H\}$.*

Es decir, $\lambda \in \mathbb{R}$ es un valor propio de T si existe $v \in H$ con $v \neq 0_H$ tal que $T(v) = \lambda\, v$. El vector v que cumple dicha condición se llama ***vector propio asociado al valor propio*** λ.

Al conjunto de todos los valores propios de T lo denotamos por $\sigma_p(T)$ y es claro que $\sigma_p(T) \subset \sigma(T)$. Notemos que, si $\dim V < \infty$ entonces $\sigma_p(T) = \sigma(T)$ pues toda función lineal $L : H \to H$ es inyectiva si y sólo si es suprayectiva. En dimensión infinita, la relación $\sigma_p(T) \subset \sigma(T)$, en general, puede ser estricta como lo muestra el siguiente ejemplo.

Ejemplo 3.30 *Sea $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ y definamos $T : \ell_2(R) \to \ell_2(\mathbb{R})$ como $T(\bar{x}) := 0$ si $j = 1$ y $T(\bar{x}) := x_{j-1}$ si $j > 1$ la cual es claramente T es lineal y continua. Entonces, $\sigma_p(T) \subsetneq \sigma(T)$.*

Demostración: Sea $\lambda \in \mathbb{R}$ y consideremos la función $T_\lambda := T - \lambda\, id$ en donde id denota la función identidad en $\ell_2(\mathbb{R})$. Observemos que $T_\lambda(\bar{x}) = -\lambda\, x_1$ si $j = 1$ y $T_\lambda(\bar{x}) = x_{j-1} - \lambda\, x_j$ si $j > 1$. Consideremos los siguientes casos:

Caso 1. Si $\lambda = 0$ entonces $T_\lambda = T$. Así pues, si $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ es tal que $T(\bar{x}) = \bar{0}$ entonces $x_{j-1} = 0$ para todo $j > 1$ y, en consecuencia, $x_j = 0$ para todo $j \in \mathbb{N}$ de modo que $\ker T_\lambda = \{\bar{0}\}$.

Caso 2. Supongamos $\lambda \neq 0$. Entonces, si $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ es tal que $T(\bar{x}) = \bar{0}$ entonces $-\lambda\, x_1 = 0$ y $x_{j-1} - \lambda\, x_j = 0$ para todo $j > 1$. Como $-\lambda\, x_1 = 0$ entonces $x_1 = 0$. Luego, como $x_1 - \lambda\, x_2 = 0$ entonces $-\lambda\, x_2 = 0$ y, por tanto, $x_2 = 0$. Continuando de este modo se tiene que, para cada $m \in \mathbb{N}$, $x_{m-1} = 0$. Así, de la expresión $x_{m-1} - \lambda\, x_m = 0$ concluimos que $x_m = 0$. Por tanto, $x_j = 0$ para todo $j \in \mathbb{N}$ lo que implica que $\bar{x} = \bar{0}$, es decir, $\ker T_\lambda = \{\bar{0}\}$.

De ambos casos concluimos que $\sigma_p(T) = \varnothing$.

Ahora, sea $\lambda \in [-1,1]$ y consideremos la función T_λ definida arriba. Sea $\bar{y} = (y_j)$ la sucesión de números reales dada por $y_1 := -1$ y $y_j := 0$ para todo $j > 1$. Es claro que $\bar{y} \in \ell_2(\mathbb{R})$. Es inmediato que si $\lambda = 0$ entonces la función $T_\lambda = T$ no es suprayectiva pues no existe sucesión $\bar{x} = (x_j)$ tal que $T(\bar{x}) = \bar{y}$. Sea pues $\lambda \neq 0$ y supongamos que existe sucesión $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ tal que $T_\lambda(\bar{x}) = \bar{y}$. Entonces, $-\lambda\, x_1 = -1$ y $x_{j-1} - \lambda\, x_j = 0$ para todo $j > 1$. Entonces, $x_1 = \frac{1}{\lambda}$ y, por tanto, $\frac{1}{\lambda} - \lambda\, x_2 = \lambda(\frac{1}{\lambda^2} - x_2) = 0$ implica que $x_2 = \frac{1}{\lambda^2}$. Continuando de este modo, podemos concluir que $x_j = \frac{1}{\lambda^j}$ para todo $j \in \mathbb{N}$ lo cual es imposible pues la serie $\sum_{j=1}^{\infty} \frac{1}{\lambda^{2j}}$ no converge en $\mathbb{R}$ ya que $\frac{1}{|\lambda|} \geq 1$. En consecuencia, T_λ no es suprayectiva y, por tanto, T_λ no es biyectiva. Así pues, $[-1,1] \subset \sigma(T)$. ■

En general, puede darse el caso que exista un espacio de Hilbert H sobre $\mathbb{R}$ y una función lineal y continua $T : H \to H$ tal que $\sigma_p(T) = \sigma(T) = \varnothing$ independientemente de la dimensión de H como lo muestran los siguientes ejemplos.

Ejemplo 3.31 *Sea $\bar{x} = (x_j)$ y definamos $T : \ell_2(R) \to \ell_2(\mathbb{R})$ lineal y continua como $T(\bar{x}) := -x_{j+1}$ si j es impar y $T(\bar{x}) := x_{j-1}$ si j es par. Entonces, $\sigma_p(T) = \sigma(T) = \varnothing$.*

Demostración: Sea $\lambda \in \mathbb{R}$ y consideremos la función $T_\lambda := T - \lambda\, id$ en donde id denota la función identidad en $\ell_2(\mathbb{R})$. Observemos que $T_\lambda(\bar{x}) = -x_{j+1} - \lambda\, x_j$ si j es impar y $T_\lambda(\bar{x}) = x_{j-1} - \lambda\, x_j$ si j es par.

Sea $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ tal que $T_\lambda(\bar{x}) = \bar{0}$. Es inmediato que si $\lambda = 0$ entonces $\ker T_\lambda = \{\bar{0}\}$. Así pues, supongamos que $\lambda \neq 0$ y consideremos los siguientes casos:

Caso 1. Si j es impar, entonces de la descripción de $T_\lambda(\bar{x})$ tenemos el siguiente sistema de ecuaciones:

$$\begin{cases} -x_{j+1} - \lambda\, x_j &= 0 \qquad (1.1) \\ x_j - \lambda\, x_{j+1} &= 0 \qquad (1.2) \end{cases}$$

De (1.1) obtenemos que $x_{j+1} = -\lambda\, x_j$ y sustituyendo en (1.2) se sigue que $x_j + \lambda^2\, x_j = (1+\lambda^2)x_j = 0$. Por tanto, $x_j = 0$ y en consecuencia $x_{j+1} = 0$.

Caso 2. Si j es par, entonces obtenemos el sistema:

$$\begin{cases} -x_j - \lambda\, x_{j-1} &= 0 \qquad (2.1) \\ x_{j-1} - \lambda\, x_j &= 0 \qquad (2.2) \end{cases}$$

y resolviendo de forma análoga al caso anterior obtenemos que $x_j = 0 = x_{j-1}$.

Por tanto, $x_j = 0$ para todo $j \in \mathbb{N}$ lo que implica que $\ker T_\lambda = \{\bar{0}\}$. Es decir, T_λ es inyectiva para todo $\lambda \in \mathbb{R}$.

Sea $\bar{y} = (y_j) \in \ell_2(\mathbb{R})$ y $\lambda \in \mathbb{R}$.

Caso 1. Si $\lambda = 0$, entonces $T_\lambda = T$. Definamos la sucesión $\bar{x} = (x_j)$ como $x_j := y_{j+1}$ si j es impar y $x_j := -y_{j-1}$ si j es par. Así pues, $T_\lambda(\bar{x}) = -x_{j+1} = y_j$ si j es impar y $T_\lambda(\bar{x}) = x_{j-1} = y_j$ si j es par. Es decir, $T_\lambda(\bar{x}) = \bar{y}$

Caso 2. Supongamos $\lambda \neq 0$ y definamos la sucesión $\bar{x} = (x_j)$ como sigue

$$\bar{x}_j := \begin{cases} \frac{1}{1+\lambda^2} y_{j+1} - \frac{\lambda}{1+\lambda^2} y_j & \text{si } j \text{ es impar,} \\ -\frac{1}{1+\lambda^2} y_{j-1} - \frac{\lambda}{1+\lambda^2}\, y_j & \text{si } j \text{ es par.} \end{cases}$$

Es claro que $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$. Observemos además, que $T_\lambda(\bar{x}) = -x_{j+1} - \lambda\, x_j = -\left(-\frac{1}{1+\lambda^2}\, y_j - \frac{\lambda}{1+\lambda^2}\, y_{j+1}\right) - \lambda \left(\frac{1}{1+\lambda^2}\, y_{j+1} - \frac{\lambda}{1+\lambda^2}\, y_j\right) = y_j$ si j es impar y $T_\lambda(\bar{x}) = -x_{j+1} - \lambda\, x_j = \left(\frac{1}{1+\lambda^2}\, y_j - \frac{\lambda}{1+\lambda^2}\, y_{j-1}\right) - \lambda \left(-\frac{1}{1+\lambda^2}\, y_{j-1} - \frac{\lambda}{1+\lambda^2}\, y_j\right) = y_j$ si j es par. Es decir, $T_\lambda(\bar{x}) = \bar{y}$.

En consecuencia, T_λ es suprayectiva para todo $\lambda \in \mathbb{R}$.

Por tanto, $\sigma_p(T) = \sigma(T) = \varnothing$. ∎

Ejemplo 3.32 *Sea* $\mathbb{R}^2 = (\mathbb{R}^2, \|\cdot\|_2)$ *con la norma Euclidiana y consideremos* $\vartheta : \mathbb{R}^2 \to \mathbb{R}^2$ *la rotación por* $\frac{\pi}{2}$ *en torno al origen. Es decir,*

$$\begin{aligned} \vartheta(x_1, x_2) &= (x_1 \cos(\tfrac{\pi}{2}) - x_2 \sin(\tfrac{\pi}{2}), x_1 \sin(\tfrac{\pi}{2}) + x_2 \cos(\tfrac{\pi}{2})) \\ &= (-x_2, x_1). \end{aligned}$$

Entonces, $\sigma_p(\vartheta) = \sigma(\vartheta) = \varnothing$*.*

Demostración: Es claro que ϑ es una función lineal en $\mathbb{R}^2$ y, por tanto, continua.

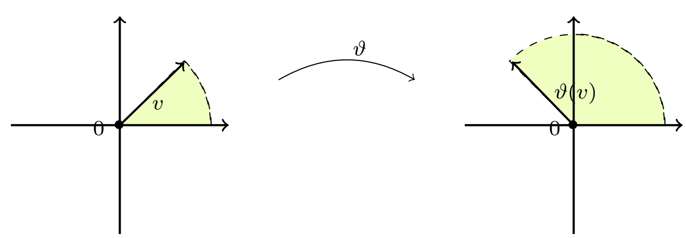


Figura 3.3: Rotación por $\frac{\pi}{2}$ en torno al origen

Sea $(x_1, x_2) \in \mathbb{R}^2$ tal que $\vartheta(x_1, x_2) = (0, 0)$. Entonces, $(-x_2, x_1) = (0, 0)$ y, por lo tanto, $x_1 = x_2 = 0$ lo que implica que $\ker \vartheta = \{(0, 0)\}$.

Ahora bien, sea $\lambda \in \mathbb{R}$ con $\lambda \neq 0$ y $\vartheta_\lambda(x_1, x_2) = (\vartheta - \lambda\, id)(x_1, x_2) = (-x_2 - \lambda x_1, x_1 - \lambda x_2)$. Si $\vartheta_\lambda(x_1, x_2) = (0, 0)$ entonces,

$$\begin{aligned} -x_2 - \lambda x_1 &= 0, \\ x_1 - \lambda x_2 &= 0. \end{aligned}$$

De la primera ecuación obtenemos que $x_2 = -\lambda\, x_1$ y, sustituyendo en la segunda se tiene que $(1 + \lambda^2) x_1 = 0$ de donde concluimos que $x_1 = 0$ y, por tanto, $x_2 = 0$. Es decir, $(x_1, x_2) = (0, 0)$ lo que prueba que $\ker \vartheta_\lambda = \{(0, 0)\}$.

En consecuencia, $\sigma_p(\vartheta) = \sigma(\vartheta) = \varnothing$. ∎

Sea $T : H \to H$ función lineal y continua. En los siguientes resultados estudiaremos para cuáles valores reales λ, la función $T_\lambda := (T - \lambda\, id_H)$ es biyectiva de H a H. De los ejemplos anteriores, supondremos a partir de este momento que las funciones en estudio son tales que el conjunto de sus valores propios es distinto del vacío.

Proposición 3.33 *Sea $T : H \to H$ una función lineal continua y $\lambda \in \mathbb{R}$ tal que $\lambda \in \sigma_p(T)$. Entonces, $|\lambda| \leq \|T\|_{\mathcal{L}_c(H,H)}$.*

Demostración: Dado que $\lambda \in \sigma_p(T)$, entonces existe $v \in H \setminus \{0_H\}$ tal que $T_\lambda(v) = 0_H$. Es decir, $T(v) = \lambda\, v$. Definiendo $\widetilde{v} := \frac{v}{\|v\|_H} \in H$ es claro que $\|\widetilde{v}\|_H = 1$ y $T(\widetilde{v}) = \lambda\, \widetilde{v}$. En consecuencia,

$$|\lambda| = |\lambda|\, \|\widetilde{v}\|_H = \|\lambda\, \widetilde{v}\|_H = \|T(\widetilde{v})\|_H \leq \|T\|_{\mathcal{L}_c(H,H)}$$

como afirma el enunciado. ■

Es decir, si $\lambda \in \mathbb{R}$ es tal que $|\lambda| > \|T\|_{\mathcal{L}_c(H,H)}$ entonces $\ker T_\lambda = \{0_H\}$ lo que significa que T_λ es inyectiva. Veremos que, de hecho, la función T_λ es biyectiva. Para ello requerimos del teorema de punto fijo de Banach.

Teorema 3.34 (de punto fijo de Banach) *Sea $X = (X, d)$ un espacio métrico completo, no vacío, y sea $\phi : X \to X$ una contracción. Entonces se cumple lo siguiente:*

(a) ϕ tiene un único punto fijo x^*, es decir, existe un único $x^* \in X$ tal que $\phi(x^*) = x^*$.

(b) Para cualquier $x_0 \in X$, la sucesión $(\phi^k(x_0))$ converge a x^* en X, y se cumple que:

$$d(\phi^k(x_0), x^*) \leq \frac{\alpha^k}{1 - \alpha}\, d(\phi(x_0), x_0), \tag{3.8}$$

donde $\alpha \in (0, 1)$ satisface (3.8).

Sea $T : H \to H$ función lineal y continua y $v \in H$ arbitrario pero fijo. Sea λ un número real y escribimos la siguiente ecuación:

$$T(u) - \lambda\, u = v. \tag{3.9}$$

Nos preguntamos, ¿para qué valores de λ la ecuación (3.9) tiene solución única?. De acuerdo a los conceptos dados en el teorema de punto fijo de Banach, queremos expresar este problema como un problema de punto fijo. Es decir, queremos encontrar una función $\phi_\lambda : H \to H$ cuyos puntos fijos sean las soluciones de la ecuación (3.9). Para ello, procederemos del siguiente modo:

Si $\lambda \neq 0$, entonces la ecuación (3.9) se escribe como:

$$u = \frac{1}{\lambda}\left(T(u) - v\right).$$

Esto nos permite definir la función $\phi_\lambda : H \to H$ como sigue:

$$\phi_\lambda(u) = \frac{1}{\lambda}\left(T(u) - v\right).$$

Las soluciones de (3.9) son justamente los puntos fijos de ϕ_λ.

Para poder aplicar el teorema de punto fijo de Banach requerimos que H sea completo y, que para ciertos valores de λ, la función ϕ_λ sea una contracción. La primera condición se sigue inmediatamente pues H es de Banach. Ahora, observemos lo siguiente: Sean $x, y \in H$, entonces:

$$\|\phi_\lambda(x) - \phi_\lambda(y)\|_H = \left|\tfrac{1}{\lambda}\right| \|T(x-y)\|_H \leq \left|\tfrac{1}{\lambda}\right| \|T\|_{\mathcal{L}_c(H,H)} \|x-y\|_H.$$

Así pues, definiendo $\alpha := \left|\frac{1}{\lambda}\right| \|T\|_{\mathcal{L}_c(H,H)}$ se tiene entonces que $\alpha \in (0,1)$ si y sólo si $|\lambda| > \|T\|_{\mathcal{L}_c(H,H)}$. En consecuencia, ϕ_λ es una contracción si $|\lambda| > \|T\|_{\mathcal{L}_c(V,V)}$. Dado que H es completo, el Teorema 3.34 asegura que existe un único $u^* \in H$ tal que:

$$u^* = \phi_\lambda(u^*) = \frac{1}{\lambda}\left(T(u^*) - v\right).$$

Es decir, T_λ es suprayectiva si $|\lambda| > \|T\|_{\mathcal{L}_c(H,H)}$.

Hemos demostrado lo siguiente:

Teorema 3.35 *Sea $T : H \to H$ función lineal y continua. Si $|\lambda| > \|T\|_{\mathcal{L}_c(H,H)}$ entonces, para cada $v \in H$, la ecuación* (3.9) *tiene una única solución. Es decir, T_λ es suprayectiva.*

Proposición 3.36 *Sea $T : H \to H$ función lineal y continua, entonces $\sigma(T) \subset \mathbb{R}$ es acotado. Más aún,*

$$\sigma(T) \subset \left[-\|T\|_{\mathcal{L}_c(H,H)}, \|T\|_{\mathcal{L}_c(V,V)}\right]$$

Demostración: De la Proposición 3.33 y el Teorema 3.35 tenemos que, si $\lambda \in \mathbb{R}$ es tal que T_λ no es biyectiva de H en H, entonces $|\lambda| \leq \|T\|_{\mathcal{L}_c(H,H)}$. En consecuencia, $\sigma(T) \subset \left[-\|T\|_{\mathcal{L}_c(H,H)}, \|T\|_{\mathcal{L}_c(H,H)}\right]$ como afirma el enunciado. ■

A continuación probamos que el conjunto espectral de una función $T : H \to H$ lineal y continua es cerrado en $\mathbb{R}$ para lo cual, usaremos de nuevo el teorema de punto fijo de Banach.

Sea $\lambda_0 \in \rho(T)$. Deseamos encontrar un $\varepsilon > 0$ tal que $(\lambda_0 - \varepsilon, \lambda_0 + \varepsilon) \subset \rho(T)$. Procedemos de la siguiente forma:

Sea $\varepsilon > 0$ y $\lambda \in \mathbb{R}$ tal que $|\lambda - \lambda_0| < \varepsilon$. Para $v \in H$ arbitrario pero fijo consideremos de nuevo la ecuación (3.9).

Notemos que, la ecuación (3.9) puedes escribirse como:

$$T(u) - \lambda_0\, u = v + (\lambda - \lambda_0)\, u. \tag{3.10}$$

Dado que $\lambda_0 \in \rho(T)$ entonces la función $T_{\lambda_0} := (T - \lambda_0\, id_H)$ es biyectiva. Es decir, existe $T_{\lambda_0}^{-1} : H \to H$ tal que $T_{\lambda_0}^{-1} \circ T_{\lambda_0} = T_{\lambda_0} \circ T_{\lambda_0}^{-1} = id_H$. De este modo, podemos escribir a la ecuación (3.10) como:

$$u = T_{\lambda_0}^{-1}(v + (\lambda - \lambda_0)u). \tag{3.11}$$

Definamos la función $\varphi : H \to H$ como:

$$\varphi(u) := T_{\lambda_0}^{-1}(v + (\lambda - \lambda_0)u).$$

Entonces las soluciones de (3.11) son justamente los puntos fijos de φ.

Usaremos el teorema de punto fijo de Banach para lo cual, necesitamos que φ sea una contracción. Así pues, encontraremos el valor de ε para el cual la función φ cumple con ser una contracción. Sean $x, y \in V$, entonces:

$$\|\varphi(x) - \varphi(y)\|_H = |\lambda - \lambda_0| \|T_{\lambda_0}^{-1}(x - y)\|_H \leq |\lambda - \lambda_0| \|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)} \|x - y\|_H.$$

Definiendo $\alpha := |\lambda - \lambda_0| \|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}$ se tiene que $\alpha \in (0,1)$ si y sólo si $|\lambda - \lambda_0| < \frac{1}{\|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}}$. Es decir, φ es una contracción si $\varepsilon := \frac{1}{\|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}}$. Dado que H es completo, el Teorema 3.34 asegura que existe un único $u^* \in H$ tal que:

$$u^* = \varphi_\lambda(u^*) = T_{\lambda_0}^{-1}(v + (\lambda - \lambda_0)u^*).$$

Notemos que si $v = 0_H$ entonces la ecuación en estudio se reduce a $T(u) = \lambda\, u$ y $u = 0_H$ es una solución. De lo anterior podemos asegura que esta es única si $|\lambda - \lambda_0| < \frac{1}{\|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}}$.

Por tanto, si $\lambda \in \mathbb{R}$ es tal que $|\lambda - \lambda_0| < \frac{1}{\|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}}$ entonces T_λ es biyectiva de H en H.

Teorema 3.37 *Sea $T : H \to H$ una función lineal y continua. Entonces, el conjunto $\rho(T)$ es abierto en $\mathbb{R}$.*

Demostración: Sea $\lambda_0 \in \rho(T)$. Definiendo $\varepsilon := \frac{1}{\|T_{\lambda_0}^{-1}\|_{\mathcal{L}_c(H,H)}} > 0$ es claro de la discusión anterior que $(\lambda_0 - \varepsilon, \lambda_0 + \varepsilon) \subset \rho(T)$. Es decir, $\rho(T)$ es un subconjunto abierto de $\mathbb{R}$. ■

Corolario 3.38 *Sea $T : H \to H$ una función lineal y continua. Entonces, el conjunto $\sigma(T)$ es compacto en $\mathbb{R}$.*

Demostración: Por la Proposición 3.36 se tiene que $\sigma(T)$ es acotado y, por el Teorema 3.37, el conjunto $\sigma(T) = \mathbb{R} \setminus \rho(T)$ es cerrado en $\mathbb{R}$. El teorema de Heine-Borel asegura que $\sigma(T) \subset \mathbb{R}$ es compacto. ■

De los resultados anteriores obtenemos que todo número real λ que esté en el conjunto espectral de una función lineal y continua T se encuentra entre los valores $-\|T\|_{\mathcal{L}_c(H,H)}$ y $\|T\|_{\mathcal{L}_c(H,H)}$. Es decir, aún cuando encontrar un valor explícito λ en el espectro de T resulte en un proceso sumamente complejo, calcular la norma del operador T nos brinda un valor aproximado. Veamos un ejemplo.

Ejemplo 3.39 *En $\ell_2(\mathbb{R})$ consideremos la función $T : \ell_2(R) \to \ell_2(\mathbb{R})$ dada por $T(\bar{x}) := 0$ si $j = 1$ y $T(\bar{x}) := x_{j-1}$ si $j > 1$ (ver* Ejemplo 3.30*).*

Es claro que $\|T(\bar{x})\|_{\ell_2(\mathbb{R})} = \|\bar{x}\|_{\ell_2(\mathbb{R})}$ para todo $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$. Por tanto,

$$\|T\|_{\mathcal{L}_c(\ell_2(\mathbb{R}),\ell_2(\mathbb{R}))} = \sup\left\{\|T(\bar{x})\|_{\ell_2(\mathbb{R})} \,:\, \|\bar{x}\|_{\ell_2(\mathbb{R})} = 1\right\} = 1.$$

De la Proposición 3.36 *concluimos que:*

$$\sigma(T) \subset [-1, 1].$$

En consecuencia, $\sigma(T) = [-1, 1]$ pues en el Ejemplo 3.30 *se probó que $[-1, 1] \subset \sigma(T)$.*

Veamos ahora algunas propiedades del espectro y el conjunto de valores propios de un operador compacto definido en un espacio de Hilbert de dimensión infinita.

Teorema 3.40 *Supongamos que $\dim H = \infty$ y sea $T : H \to H$ un operador compacto. Entonces,*

(a) $0 \in \sigma(T)$.

(b) $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$.

Demostración: (a) Argumentando por contradicción, supongamos que $0 \notin \sigma(T)$. Entonces $0 \in \rho(T)$ lo que implica que T es biyectiva de H en H lo cual es imposible pues T es un operador compacto. En consecuencia, $0 \in \sigma(T)$.

(b) Sea $\lambda \in \sigma(T) \setminus \{0\}$ y supongamos que $\lambda \notin \sigma_p(T)$. Entonces, $\ker T_\lambda = \{0_H\}$. Así, por el Teorema 3.27 se sigue que $\operatorname{im} T_\lambda = H$ y, por tanto, T_λ es biyectiva de H en H. En consecuencia, $\lambda \in \rho(T)$ lo cual es imposible pues $\lambda \in \sigma(T)$. Por tanto, $\lambda \in \sigma_p(T) \setminus \{0\}$.

Inversamente: dado que $\sigma_p(T) \subset \sigma(T)$ entonces $\sigma_p(T) \setminus \{0\} \subset \sigma(T) \setminus \{0\}$. Esto prueba el resultado. ∎

Lema 3.41 *Supongamos que $\dim H = \infty$. Sea $T : H \to H$ un operador compacto y (λ_k) una sucesión de números reales distintos tales que $\lambda_k \to \lambda$ y $\lambda_k \in \sigma(T) \setminus \{0\}$ para cada $k \in \mathbb{N}$. Entonces, $\lambda = 0$.*

Demostración: Del Teorema 3.40 se sigue que $\lambda_k \in \sigma_p(T) \setminus \{0\}$ para cada $k \in \mathbb{N}$. Argumentando por contradicción, supongamos que $\lambda \neq 0$. Entonces, existe $\varepsilon_0 > 0$ tal que $|\lambda_k| \geq \varepsilon_0$ para todo $k \in \mathbb{N}$.

Para cada $k \in \mathbb{N}$, existe $e_k \in H \setminus \{0_H\}$ tal que $T_{\lambda_k}(e_k) = 0_V$. Observemos lo siguiente: Como $e_1 \neq 0_V$ entonces $\{e_1\}$ es un subconjunto de H linealmente independiente. Supongamos que, para $k \in \mathbb{N}$ arbitrario pero fijo, el conjunto $\{e_1, \ldots, e_{k-1}\}$ es linealmente independiente y que $e_k = \sum_{i=1}^{k-1} \alpha_i\, e_i$ para algunos $\alpha_1, \ldots, \alpha_{k-1} \in \mathbb{R}$. Entonces,

$$T(e_k) = \sum_{i=1}^{k-1} \alpha_i\, T(e_i) = \sum_{i=1}^{k-1} \alpha_i\, \lambda_i\, e_i \quad \text{y} \quad \lambda_k\, e_k = \lambda_k \sum_{i=1}^{k-1} \alpha_i\, e_i = \sum_{i=1}^{k-1} \alpha_i\, \lambda_k\, e_i$$

y, como $T(e_k) = \lambda_k\, e_k$ se sigue que:

$$T(e_k) - \lambda_k\, e_k = \sum_{i=1}^{k-1} \alpha_i(\lambda_i - \lambda_k)\, e_i = 0_V,$$

y, en consecuencia, $\alpha_i(\lambda_i - \lambda_k) = 0$ para todo $i = 1, \ldots, k-1$. Dado que $\lambda_i \neq \lambda_k$ para todo $i = 1, \ldots, k-1$ entonces $\alpha_i = 0$ para todo $i = 1, \ldots, k-1$ lo cual es imposible. Por tanto, el conjunto $\{e_1, \ldots, e_k\}$ es linealmente independiente.

Definiendo, para cada $k \in \mathbb{N}$, el conjunto $E_k := \mathrm{lin}(\{e_1, \ldots, e_k\})$ se sigue que E_k es un subespacio vectorial de H de dimensión finita, y por tanto, cerrado. Más aún, $E_k \subset E_{k+1}$ y $E_k \neq E_{k+1}$ para todo $k \in \mathbb{N}$ por lo demostrado en el párrafo anterior.

Sea $k \in \mathbb{N}$ y $v \in E_k$. Entonces, $v = \sum_{i=1}^{k} \alpha_i\, e_i$ para algunos $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Así pues,

$$T_{\lambda_k}(v) = T(v) - \lambda_k\, u = \sum_{i=1}^{k} \alpha_i(\lambda_i - \lambda_k)\, e_i = \sum_{i=1}^{k-1} \alpha_i(\lambda_i - \lambda_k)\, e_i \in E_{k-1}.$$

Es decir, $T_{\lambda_k}(E_k) \subset E_{k-1}$ para todo $k \in \mathbb{N}$.

Fijemos $\delta \in (0,1)$. Por el lema de Riez, para cada $k \in \mathbb{N}$ con $k \geq 2$, existe $v_\delta^k \in E_k$ con $\|v_\delta^k\|_H = 1$ tal que $\|v_\delta^k - w\|_V \geq \delta$ para todo $w \in E_{k-1}$. Definamos pues la sucesión $v_k := v_\delta^k$ en H que es acotada.

Sean $2 \leq k < j$, entonces:

$$\begin{aligned}
\|T(v_j) - T(v_k)\|_H = \|T_{\lambda_j}(v_j) + \lambda_j\, v_j - T(v_k)\|_H &= \|\lambda_j\, v_j - (T(v_k) - T_{\lambda_j}(v_j))\|_H \\
&= |\lambda_j| \|v_j - \tfrac{1}{\lambda_j}(T(v_k) - T_{\lambda_j}(v_j))\|_H \\
&\geq \varepsilon_0\, \delta
\end{aligned}$$

pues $T_{\lambda_j}(v_j) - T(v_k) \in E_{j-1}$ ya que $E_{k+1} \subset E_k \subset E_{j-1} \subset E_j$. En consecuencia, la sucesión $(T(v_k))$ no contiene ninguna subsucesión convergente en H. Esto contradice el hecho de que T es un operador compacto. Por lo tanto, $\lambda = 0$ como afirma el enunciado. ∎

Teorema 3.42 *Supongamos que* $\dim H = \infty$ *y sea* $T : H \to H$ *un operador compacto. Entonces, se tiene uno de los siguientes casos:*

(a) $\sigma(T) = \{0_H\}$.
(b) $\sigma(T) \setminus \{0\}$ *es finito.*
(c) Existe una sucesión (x_k) *de números reales, todos distintos, tal que* $x_k \to 0$ *y* $\sigma(T) \setminus \{0\} = \{x_k \,:\, k \in \mathbb{N}\}$.

Demostración: Consideremos, para cada $k \in \mathbb{N}$, el conjunto

$$\mathfrak{C}_k := \sigma(T) \cap \left\{\lambda \in \mathbb{R} \,:\, |\lambda| \geq \tfrac{1}{k}\right\}.$$

Entonces, para cada $k \in \mathbb{N}$, el conjunto $\mathfrak{C}_k$ o bien es vacío o bien es finito. En efecto, si $\mathfrak{C}_k = \varnothing$ para cada $k \in \mathbb{N}$, entonces $\lambda \in \sigma(T)$ es tal que $|\lambda| < \frac{1}{k}$ para cada $k \in \mathbb{N}$ por lo que $\lambda = 0$. Ahora, supongamos que $\mathfrak{C}_k$ no es finito. Entonces, sea (λ_j) una sucesión de elementos distintos de $\sigma(T)$ tal que $|\lambda_j| \geq \frac{1}{k}$ para todo $j \in \mathbb{N}$. Como $\sigma(T)$ es compacto, entonces existe una subsucesión (λ_{j_m}) de (λ_j) tal que $\lambda_{j_m} \to \lambda$ con $\lambda \in \mathfrak{C}_k$. Entonces, $\lambda \neq 0$ pues $|\lambda| \geq \frac{1}{k}$ lo cual no es posible por el Lema 3.41.

Por tanto, si $\sigma(T) \setminus \{0\}$ contiene una cantidad numerable de puntos distintos, ■

3.4.1. Teoría espectral de operadores compactos-autoadjuntos

Sea H un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\|\cdot\|_H$.

Consideremos el siguiente ejemplo.

Ejemplo 3.43 *Sean* $a, b \in \mathbb{R}$ *con* $0 < a < b$, *arbitrarios pero fijos. Consideremos el espacio de medida* $([a,b], \mathcal{B}([a,b]), \mu)$ *con* μ *la medida de Lebesgue en* $\mathcal{B}([a,b])$. *La función* $T : L^2([a,b]) \to L^2([a,b])$ *dada por:*

$$(Tu)(x) := x\, u(x).$$

está bien definida pues es el producto de funciones de cuadrado integrable. Es claro además que T *es lineal. Luego, para cada* $u \in L^2([a,b])$

$$\|Tu\|^2_{L^2([a,b])} = \int_{[a,b]} |x|^2 |u(x)|^2 \, d\mu \leq \text{máx}\{a^2, b^2\} \int_{[a,b]} |u(x)|^2 \, d\mu = \text{máx}\{a^2, b^2\} \|u\|^2_{L^2([a,b])}$$

lo que implica que T *es continua.*

Observemos que:

$$\begin{aligned}\langle T(u), v\rangle_{L^2([a,b])} &= \int_{[a,b]} (Tu)(x)\, v(x)\, d\mu = \int_{[a,b]} x\, u(x)\, v(x)\, d\mu \\ &= \int_{[a,b]} u(x)\, (Tv)(x)\, d\mu = \langle u, T(v)\rangle_{L^2([a,b])} \qquad \forall\, u, v \in L^2([a,b]).\end{aligned}$$

Por tanto, T es un operador autoadjunto.

Sea $\lambda \in \mathbb{R}$ y sea $u \in L^2([a,b])$ tal que $(T_\lambda u)(x) = 0$ para cada $x \in [a,b]$. Consideremos los siguientes casos:

CASO 1. *Si $\lambda = 0$, entonces $(T_\lambda u)(x) = (Tu)(x) = x\,u(x) = 0$ para cada $x \in [a,b]$. Por tanto, $u(x) = 0$ para cada $x \in [a,b]$ lo que implica que* $\ker T_\lambda = \{0\}$.

CASO 2. *Supongamos que $\lambda \neq 0$. Entonces, $(T_\lambda u)(x) = (Tu)(x) - \lambda u(x) = (x - \lambda)\,u(x)$ para cada $x \in [a,b]$ y, por tanto, $u(x) = 0$ para cada $x \in [a,b]$. Es decir,* $\ker T_\lambda = \{0\}$.

El ejemplo anterior nos dice que el conjunto de valores propios de una función lineal y continua en un espacio de Hilbert puede ser vacío, aún cuando esta función sea un operador autoadjunto o un operador compacto. En esta sección estudiaremos las condiciones que debe cumplir una función lineal y continua para que al menos exista un valor propio. Daremos también algunas propiedades de este conjunto.

Proposición 3.44 *Sea $T : H \to H$ un operador autoadjunto. Entonces, los vectores propios correspondientes a valores propios distintos, si existen, son ortogonales.*

Demostración: Sean $\lambda_1, \lambda_2 \in \mathbb{R}$ con $\lambda_1 \neq \lambda_2$ tales que $\lambda_1, \lambda_2 \in \sigma_p(T)$ y sean $v_1, v_2 \in V \setminus \{0_V\}$ los vectores propios asociados a los valores propios λ_1 y λ_2 respectivamente.

Observemos que:

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle T(v_1), v_2 \rangle = \langle v_1, T(v_2) \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

y, dado que $\lambda_1 \neq \lambda_2$ entonces $\langle v_1, v_2 \rangle = 0$ como afirma el enunciado. ■

El siguiente resultado garantiza la existencia de al menos un valor propio cuando la función lineal y continua es un operador compacto y autoadjunto. Más aún, nos dice cuál podría ser este valor.

Teorema 3.45 *Sea $T : H \to H$ una función lineal y continua. Si T es un operador compacto y autoadjunto entonces, al menos uno de los valores reales, $-\|T\|_{\mathcal{L}_c(H,H)}$ o $\|T\|_{\mathcal{L}_c(H,H)}$, es un valor propio de T.*

Demostración: Si $T(u) = 0_H$ para cada $u \in H$, entonces $\|T\|_{\mathcal{L}_c(H,H)} = 0$ y es claro que $\ker T \neq \{0_V\}$. Así pues, supongamos que T no es la función constante con valor 0_H en H.

Del Teorema 3.14 se tiene que $\|T\|_{\mathcal{L}_c(H,H)} = \sup \{ |\langle T(u), u \rangle| : \|u\|_H = 1 \}$. Por tanto, existe una sucesión (u_k) en H tal que $\|u_k\|_H = 1$ para cada $k \in \mathbb{N}$ y tal que $\|T\|_{\mathcal{L}_c(H,H)} = \lim_{k\to\infty} |\langle T(u_k), u_k \rangle|$. Definamos $x_k := \langle T(u_k), u_k \rangle \in \mathbb{R}$ para cada $k \in \mathbb{N}$. La sucesión (x_k) es acotada en $\mathbb{R}$ de modo que, existe (x_{k_j}) una subsucesión de (x_k) tal que $x_{k_j} \to \lambda$ cuando $j \to \infty$ en donde $\lambda \in \mathbb{R} \setminus \{0\}$ es tal que $|\lambda| = \|T\|_{\mathcal{L}_c(H,H)}$.

Notemos lo siguiente:

$$\begin{aligned}\|T(u_{k_j}) - \lambda\, u_{k_j}\|_H^2 &= \|T(u_{k_j})\|_H^2 + \lambda^2\, \|u_{k_j}\|_H^2 - \lambda\, \langle\, T(u_{k_j}), u_{k_j}\,\rangle - \lambda\, \langle\, u_{k_j}, T(u_{k_j})\,\rangle \\ &= \|T(u_{k_j})\|_H^2 + \lambda^2 - 2\lambda\, \langle\, T(u_{k_j}), u_{k_j}\,\rangle \\ &\leq 2\lambda^2 - 2\lambda\, \langle\, T(u_{k_j}), u_{k_j}\,\rangle \qquad\qquad \forall\, j \in \mathbb{N}.\end{aligned}$$

En consecuencia, $\|T(u_{k_j}) - \lambda\, u_{k_j}\|_H \to 0$ cuando $j \to \infty$.

Por otro lado, como T es un operador compacto y (u_{k_j}) es una sucesión acotada en H, entonces existe $(u_{k_{j_\ell}})$ subsucesión de (u_{k_j}) tal que $T(u_{k_{j_\ell}})$ converge a algún v en H cuando $\ell \to \infty$. Denotemos por $v_\ell := u_{k_{j_\ell}}$ para todo $\ell \in \mathbb{N}$. Entonces,

$$\|\lambda\, v_\ell - v\|_H \leq \|\lambda\, v_\ell - T(v_\ell)\|_H + \|T(v_\ell) - v\|_H \qquad \forall\, \ell \in \mathbb{N}$$

de modo que $v_\ell \to \frac{v}{\lambda}$ en H cuando $\ell \to \infty$. De la continuidad de T concluimos que $v = \lim_{\ell\to\infty} T(v_\ell) = \frac{1}{\lambda}\, T(v)$. Es decir, $T(v) = \lambda\, v$. Notemos además que $v \neq 0_H$ pues $\|v\|_H = \lim_{\ell\to\infty} \|\lambda\, v_\ell\|_H = \|T\|_{\mathcal{L}c(H,H)} \neq 0$. Por tanto, $\lambda \in \sigma_p(T)$ y esto concluye la demostración. ∎

El siguiente resultado da una propiedad interesante del espectro de un operador compacto autoadjunto. Esta vez, omitimos su demostración pues dentro de ella se consideran algunos conceptos que quedan, en principio, fuera del objetivo central de este trabajo.

Teorema 3.46 *Sea $T : H \to H$ una función lineal y continua. Supongamos que T es un operador autoadjunto y definamos:*

$$m := \inf\{\langle T(u), u\rangle \,:\, \|u\|_H = 1\} \qquad y \qquad M := \sup\{\langle u, T(u)\rangle \,:\, \|u\|_H = 1\}.$$

Entonces, $\sigma(T) \subset [m, M]$, $m \in \sigma(T)$ y $M \in \sigma(T)$. Más aún, $\|T\|_{\mathcal{L}_c(H,H)} = \max\{|m|, |M|\}$.

Demostración: Consulta, por ejemplo, [13]. ∎

Corolario 3.47 *Sea $T : H \to H$ una función lineal y continua. Si T es un operador autoadjunto tal que $\sigma(T) = \{0\}$, entonces $T(u) = 0_H$ para todo $u \in H$.*

Demostración: Sean m y M como en el Teorema 3.46. Es decir:

$$m := \inf\{\langle T(u), u\rangle \,:\, \|u\|_H = 1\} \qquad \text{y} \qquad M := \sup\{\langle u, T(u)\rangle \,:\, \|u\|_H = 1\}.$$

Entonces, $m, M \in \sigma(T)$ por el Teorema 3.46 y, por hipótesis concluimos que $m = M = 0$. Así pues, $\|T\|_{\mathcal{L}_c(H,H)} = \max\{|m|, |M|\} = 0$ lo que implica que T es la función constante igual a 0_H en H como afirma el enunciado. ∎

3.5. El teorema espectral de operadores compactos-autoadjuntos

En esta sección demostramos una versión del teorema espectral para operadores compactos autoadjuntos en espacios de Hilbert separables. Esta versión del teorema espectral nos garantiza la existencia de una base de Hilbert compuesta por vectores propios para un operador compacto y autoadjunto. Recordemos que el espectro de un operador compacto autoadjunto consta de una cantidad finita o, a lo más numerable, de números reales cuyo único punto de acumulación es el 0 (ver Teorema 3.42).

Del mismo modo que en la sección anterior, el conjunto H denota un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|_H$.

Lema 3.48 *Sea $T : H \to H$ una función lineal y continua y V un subespacio vectorial cerrado de H. Si T es un operador compacto y autoadjunto tal que $T(V) \subset V$ entonces se tiene lo siguiente:*

(a) $T(V^\perp) \subset V^\perp$.

(b) $T_{|_{V^\perp}} : V^\perp \to V^\perp$ *es también un operador compacto y autoadjunto. Además,*

$$\|T_{|_{V^\perp}}\|_{\mathcal{L}_c(V^\perp, V^\perp)} \leq \|T\|_{\mathcal{L}_c(H,H)}$$

Demostración: (a) Sea $w \in T(V^\perp)$. Existe $v \in V^\perp$ tal que $T(v) = w$. Dado que $v \in V^\perp$ entonces $\langle v, u \rangle = 0$ para todo $u \in V$. Como T es autoadjunto se tiene que:

$$\langle w, u \rangle = \langle T(v), u \rangle = \langle v, T(u) \rangle = 0 \qquad \forall\, u \in V$$

pues $T(u) \in V$ por hipótesis. Por tanto, $w \in V^\perp$.

(b) Primero veamos que $T_{|_{V^\perp}}$ es autoadjunto:

Sean $v, w \in V^\perp$. Dado que $T_{|_{V^\perp}}(v) = T(v)$ y $T_{|_{V^\perp}}(w) = T(w)$ y, T es un operador adjunto, entonces:

$$\left\langle T_{|_{V^\perp}}(v), w \right\rangle = \langle T(v), w \rangle = \langle v, T(w) \rangle = \left\langle v, T_{|_{V^\perp}}(w) \right\rangle.$$

Ahora veamos que $T_{|_{V^\perp}}$ es compacto:

Sea (w_k) una sucesión acotada en $V^\perp$. Es inmediato que (w_k) es también acotada en H por lo que existe $(T(w_{k_j}))$ una subsucesión de $(T(w_k))$ que converge en H. Dado que $T(w_k) = T_{|_{V^\perp}}(w_k)$ para todo $k \in \mathbb{N}$, concluimos que $(T_{|_{V^\perp}}(w_{k_j}))$ es una subsucesión de $(T_{|_{V^\perp}}(w_k))$ que converge en H a algún elemento de $V^\perp$.

Luego, es inmediato que $\|T_{|_{V^\perp}}\|_{\mathcal{L}_c(V^\perp,V^\perp)} \leq \|T\|_{\mathcal{L}_c(H,H)}$.

Esto concluye la demostración. ■

Lema 3.49 *Sea $T : H \to H$ una función lineal y continua. Si T es un operador compacto entonces, para cada $\lambda \in \mathbb{R} \setminus \{0\}$ el subespacio vectorial* $\ker T_\lambda$ *es cerrado y de dimensión finita en H.*

Demostración: Sea $\lambda \in \mathbb{R} \setminus \{0\}$. Es inmediato de la continuidad de T que $\ker T_\lambda$ es cerrado. Consideremos $V := \ker T_\lambda$ y B_V su bola unitaria cerrada. Sea (v_k) una sucesión de elementos de B_V. Es claro que (v_k) es acotada en H de modo que existe una subsucesión $(T_\lambda(v_{k_j}))$ de $(T_\lambda(v_k))$ que converge en H a algún $v \in V$. Notemos que, $T_\lambda(v_k) = 0$ para cada $k \in \mathbb{N}$, entonces, $T(v_k) = \lambda\, v_k$ para cada $k \in \mathbb{N}$ y, en consecuencia, $v_k = \frac{1}{\lambda}\, T(v_k)$ para cada $k \in \mathbb{N}$. Por tanto,

$$\frac{1}{\lambda}\, v = \lim_{j\to\infty} \frac{1}{\lambda}\, T(v_{k_j}) = \lim_{j\to\infty} v_{k_j}.$$

Es decir, (v_k) tiene una subsucesión convergente en B_V y, por tanto, B_V es compacto en V. El teorema de Riesz (ver Teorema 3.25) asegura que $\dim V < \infty$. ■

Notemos del lema anterior que, si $\lambda = 0$, de manera general se tiene que $0 \leq \dim \ker T \leq \infty$ pues si tomamos, por ejemplo, $T(u) = 0_H$ para todo $u \in H$ entonces $\ker T = H$.

Además, la condición de que T sea un operador compacto es necesaria pues si, por ejemplo, $\dim H = \infty$ y $T = id_H$ entonces $\ker T_1 = \ker (T - id_H) = H$.

Lema 3.50 *Sea $T : H \to H$ una función lineal y continua. Si T es un operador compacto entonces, para cada $\lambda \in \mathbb{R} \setminus \{0\}$ el subespacio vectorial* $\operatorname{im} T_\lambda$ *es cerrado en H.*

Demostración: Sea $\lambda \in \mathbb{R} \setminus \{0\}$ y consideremos el subespacio vectorial $W := \operatorname{im} T_\lambda$ de H. Sea $w \in \overline{W}$, entonces existe una sucesión (w_k) en H tal que $\lim_{k\to\infty} T_\lambda(w_k) = w$ en H.

Por el Lema 3.49 el subespacio vectorial $V := \ker T_\lambda$ es cerrado y, por tanto, del teorema del complemento ortogonal (ver Teorema 1.28), concluimos que $H = V \oplus V^\perp$. Así pues, para cada $k \in \mathbb{N}$, existen únicos $u_k \in V$ y $w_k \in V^\perp$ tales que $w_k = u_k + v_k$. Consideremos la sucesión (v_k) y veamos que es acotada en H.

Argumentando por contradicción, supongamos que (v_k) no es acotada. Entonces, existe (v_{k_j}) subsucesión de (v_k) tal que $\|v_{k_j}\|_H \to \infty$ cuando $j \to \infty$. Es posible suponer que $v_{k_j} \neq 0_H$ para cada $j \in \mathbb{N}$ de modo que $\widetilde{v}_j := \frac{v_{k_j}}{\|v_{k_j}\|_H} \in V^\perp$ y $\|\widetilde{v}_j\|_H = 1$ para cada $j \in \mathbb{N}$. Observemos que:

$$T_\lambda(\widetilde{v}_j) = \frac{1}{\|v_{k_j}\|_H}\, T_\lambda(v_{k_j}) = \frac{1}{\|v_{k_j}\|_H}\, T_\lambda(w_{k_j} - u_{k_j}) = \frac{1}{\|v_{k_j}\|_H}\, T_\lambda(w_{k_j}) \qquad \forall\, j \in \mathbb{N}$$

de modo que $\lim_{j\to\infty} T_\lambda(\widetilde{v}_j) = 0_H$ pues $\lim_{j\to\infty} T_\lambda(w_{k_j}) = w$. Dado que T es un operador compacto, entonces existe $(T(\widetilde{v}_{j_\ell}))$ subsucesión de $(T(\widetilde{v}_j))$ que converge, a saber, a $\widetilde{v}$ en H. Así pues,

$$\lim_{\ell\to\infty} \widetilde{v}_{j_\ell} = \lim_{\ell\to\infty} \frac{1}{\lambda}\, T(\widetilde{v}_{j_\ell}) = \frac{1}{\lambda}\, \tilde{v}.$$

Observemos que $T_\lambda(\frac{1}{\lambda}\,\widetilde{v}) = \frac{1}{\lambda}\,T_\lambda(\widetilde{v}) = \frac{1}{\lambda}\,\text{lím}_{\ell\to\infty} T_\lambda(\widetilde{v}_{j_\ell}) = 0_H$. Por tanto, $\frac{1}{\lambda}\,\widetilde{v} \in V$ y es claro que $\|\frac{1}{\lambda}\,\widetilde{v}\|_H = 1$.

Por otro lado, como $\widetilde{v}_{j_\ell} \in V^\perp$ para cada $\ell \in \mathbb{N}$, entonces por el teorema de Pitágoras (ver Teorema 1.17) concluimos que:

$$\|\widetilde{v}_{j_\ell} - \tfrac{1}{\lambda}\widetilde{v}\|_H^2 = \|\widetilde{v}_{j_\ell}\|_H^2 + \|\tfrac{1}{\lambda}\widetilde{v}\|_H^2 = 2 \qquad \forall \ell \in \mathbb{N}$$

y esto contradice que $\widetilde{v}_{j_\ell} \to \frac{1}{\lambda}\widetilde{v}$. Por tanto, (v_k) es acotada en H.

Así, como T es un operador compacto, existe $(T(v_{k_j}))$ subsucesión de $(T(v_k))$ que converge, a saber, a $\widetilde{w}$ en H. Notemos que:

$$v_k = \frac{1}{\lambda}\,(\lambda\, v_k - T(v_k) + T(v_k)) = \frac{1}{\lambda}(T(v_k) - T_\lambda(v_k)) = \frac{1}{\lambda}(T(v_k) - w_k)) \qquad \forall\, k \in \mathbb{N}.$$

Por tanto,

$$\lim_{j\to\infty} v_{k_j} = \lim_{j\to\infty} \frac{1}{\lambda}(T(v_{k_j}) - w_{k_j})) = \frac{1}{\lambda.}(\tilde{w} - w) := v \in H$$

En consecuencia,

$$w = \lim_{j\to\infty} T_\lambda(w_{k_j}) = \lim_{j\to\infty} T_\lambda(v_{k_j}) = T_\lambda(v)$$

lo que demuestra que $w \in W$. Es decir, W es cerrado en H. ■

Lema 3.51 *Supongamos que H es separable y sea $T : H \to H$ una función lineal y continua. Si T es un operador compacto y autoadjunto, entonces existe $\mathcal{I} \subset \mathbb{N}$ no vacío y una sucesión $\{\lambda_k \,:\, k \in \mathcal{I}\}$ de números reales tales que $\lambda_k \in \sigma_p(T) \setminus \{0\}$, $|\lambda_{k+1}| \le |\lambda_k|$ para cada $k \in \mathcal{I}$ y un correspondiente conjunto ortonormal $\mathcal{O} = \{e_k \,:\, k \in \mathcal{I}\}$ de H tal que:*

(i) $T(e_i) = \lambda_i\, e_i$ *para cada* $i \in \mathcal{I}$.

(ii) $\ker T = (\text{lin}(\{e_i \,:\, i \in \mathcal{I}\}))^\perp$

(iii) Si $\mathcal{I} = \mathbb{N}$ *entonces* $\text{lím}_{k\to\infty} \lambda_k = 0$.

Demostración: Si $T(u) = 0_H$ para cada $u \in H$ entonces $\ker T = H$. Dado que $\{0_H\} = \text{lin}(\varnothing)$ entonces definiendo $\mathcal{I} := \varnothing$ el resultado se satisface trivialmente. Supongamos entonces que T no es la función constante con valor 0_H en H. De acuerdo al Teorema 3.44 el conjunto de valores propios de T es finito o numerable. De este modo, consideremos los siguientes casos:

Caso 1. Si el cardinal del conjunto $\sigma_p(T) \setminus \{0\}$ es finito, a saber, $N \in \mathbb{N}$. Definamos $\mathcal{I} := \{1, \ldots, N\}$.

Aplicando el Teorema 3.45 al operador T, sabemos que existe un vector propio e_1 de T con valor propio $\lambda_1 \in \mathbb{R}$ tal que $|\lambda_1| = \|T_1\|_{\mathcal{L}_c(H,H)}$. Dado que $T(u) \neq 0$ para todo $u \in H$ entonces $\lambda_1 \neq 0$ y, por tanto, podemos suponer que $\|e_1\|_H = 1$. Sea $H_1 := \text{lin}(\{e_1\})^\perp$.

Entonces por el Lema 3.51, la restricción $T\,|_{H_1}$ es un operador compacto y autoadjunto tal que $\|T_{|_{H_1}}\|_{\mathcal{L}_c(H_1,H_1)} \leq \|T\|_{\mathcal{L}_c(H,H)} = |\lambda_1|$. Si $T\,|_{H_1}$ es el operador constante igual a 0_H en H_1 entonces el resultado queda probado, si no es así, aplicando nuevamente el Teorema 3.45 al operador $T\,|_{H_1}$ obtenemos e_2 vector propio de $T\,|_{H_1}$ con valor propio $\lambda_2 \in \mathbb{R} \setminus \{0\}$ tales que $|\lambda_2| = \|T_{|_{H_1}}\|_{\mathcal{L}_c(H_1,H_1)} \leq |\lambda_1|$ y $\|e_2\|_H = 1$. Continuando de este modo obtenemos, para $k-1 \in \mathbb{N}$, una sucesión de vectores propios $e_1, \ldots, e_{k-1} \in H$ ortogonales entre sí con correspondientes valores propios reales $\lambda_1, \ldots, \lambda_{k-1}$ tales que $|\lambda_1| \geq \cdots |\lambda_{k-1}| > 0$ y, tal que si definimos $H_{k-1} := \operatorname{lin}(\{e_1, \ldots, e_{k-1}\})^{\perp}$ entonces $T\,|_{H_{k-1}}$ es un operador compacto y autoadjunto tal que $\|T_{|_{H_{k-1}}}\|_{\mathcal{L}_c(H_{k-1},H_{k-1})} \leq |\lambda_{k-1}|$. Si $T\,|_{H_{k-1}}$ es el operador constante igual a 0_H en H_{k-1} entonces el resultado queda probado, si no es así, aplicamos el Teorema 3.45 al operador $T\,|_{H_{k-1}}$ para obtener e_k vector propio de $T\,|_{H_{k-1}}$ con valor propio $\lambda_k \in \mathbb{R}\setminus\{0\}$ tales que $0 < |\lambda_k| = \|T_{|_{H_{k-1}}}\|_{\mathcal{L}_c(H_{k-1},H_{k-1})} \leq |\lambda_{k-1}|$ y $\|e_k\|_H = 1$. Dado que $e_k \in H_{k-1}$ entonces $\langle e_k, e_j \rangle = 0$ para todo $j = 1, \ldots, k-1$ de modo que $e_1, \ldots, e_k$ son vectores propios ortogonales entre sí. Sea $H_k := \operatorname{lin}(\{e_1, \ldots, e_k\})^{\perp}$. Entonces por el Lema 3.48, la restricción $T\,|_{H_k}$ es un operador compacto y autoadjunto. Dado que $H_k \subset H_{k-1}$ entonces $\|T_{|_{H_k}}\|_{\mathcal{L}_c(H_k,H_k)} \leq \|T\|_{\mathcal{L}_c(H_{k-1},H_{k-1})} = |\lambda_k|$. Y, repetimos el proceso anterior hasta obtener que $T\,|_{H_N}(u) = 0_H$ para todo $u \in H_N$. Así, $\operatorname{lin}(\{e_1, \ldots, e_N\})^{\perp} = H_N \subset \ker T$. Inversamente, si $u \in \ker T$ entonces:

$$\lambda_j \langle e_j, v \rangle = \langle \lambda_j\, e_j, u \rangle = \langle\, T(e_j), u \rangle = \langle e_j, T(u) \rangle = 0 \quad \text{para todo } j = 1, \ldots, N \qquad (3.12)$$

de modo que $u \in \operatorname{lin}(\{e_1, \ldots, e_N\})^{\perp} = H_N$ y, por tanto, $H_N = \ker T$. Esto prueba los incisos *(i)* y *(ii)*.

Caso 2. Si el cardinal del conjunto $\sigma_p(T) \setminus \{0\}$ es numerable entonces consideramos $\mathcal{I} = \mathbb{N}$. En este caso, asumimos que el proceso del caso anterior no se detiene y, por tanto, obtenemos una sucesión (λ_k) de números reales tales que $\lambda_k \in \sigma_p(T)$ y $0 < |\lambda_{k+1}| \leq |\lambda_k|$ para cada $k \in \mathbb{N}$. El Lema 3.41 asegura que $\lambda_k \to 0$ y esto prueba *(iii)*. Sea $\{e_k : k \in \mathbb{N}\}$ el conjunto de vectores propios asociados a los valores propios $\{\lambda_k : k \in \mathbb{N}\}$, es decir, $T(e_k) = \lambda_k\, e_k$ para todo $k \in \mathbb{N}$. La Proposición 3.44 asegura que $\{e_k : k \in \mathbb{N}\}$ es ortonormal y el inciso *(i)* queda probado.

Notemos que, siguiendo la misma idea de (3.12) obtenemos que si $u \in \ker T$ entonces $u \perp e_k$ para todo $k \in \mathbb{N}$ y, por tanto, $u \in (\operatorname{lin}(\{e_k : k \in \mathbb{N}\}))^{\perp}$. Inversamente, si $u \in (\operatorname{lin}(\{e_k : k \in \mathbb{N}\}))^{\perp}$ entonces $u \in \operatorname{lin}(\{e_1, \ldots, e_k\})^{\perp}$ para cada $k \in \mathbb{N}$ y, dado que $\|T_{|_{H_k}}\|_{\mathcal{L}_c(H_k,H_k)} \leq |\lambda_k|$ entonces $\|T(u)\|_H \leq |\lambda_k|\, \|u\|_H$ para todo $k \in \mathbb{N}$. Como $\lambda_k \to 0$ concluimos que $T(u) = 0_H$ y, por tanto, $u \in \ker T$. Esto prueba el inciso *(ii)*. ■

Teorema 3.52 (espectral de operadores compactos-autoadjuntos) *Supongamos que H es separable y sea $T : H \to H$ una función lineal y continua. Si T es un operador compacto y autoadjunto, entonces existe una clase $\mathcal{B} = \{e_k \in H : k \in \mathbb{N}\}$ tal que e_k es un vector propio de T para cada $k \in \mathbb{N}$ y $\mathcal{B}$ es una base de Hilbert de H.*

Demostración: Sea (λ_k) sucesión de números reales tal que $\lambda_k \in \sigma_p(T)\setminus\{0\}$ para cada $k \in \mathbb{N}$ (ver Lema 3.51) y consideremos $\lambda_0 := 0$. Definamos los siguientes subespacios vectoriales de H

$$H_0 := \ker T \qquad \text{y} \qquad H_k := \ker T_{\lambda_k} \qquad \forall\, k \in \mathbb{N}.$$

El Lema 3.49 asegura que $0 \leq \dim H_0 \leq \infty$ y $0 < \dim H_k < \infty$ para todo $k \in \mathbb{N}$. Notemos que, para cualesquiera $u \in H_k$ y $v \in H_j$ con $k \neq j$ se tiene que:

$$\lambda_j \langle u, v\rangle = \langle u, T(v)\rangle = \langle T(u), v\rangle = \lambda_k \langle u, v\rangle$$

y, por lo tanto, $\langle u, v\rangle = 0$ pues $\lambda_j \neq \lambda_k$.

Sea $V := \operatorname{lin}(\cup_{k=0}^{\infty} H_k) \subset H$. Es claro que $T(V) \subset V$ y, por el Lema 3.48 se sigue que $T(V^{\perp}) \subset V^{\perp}$ y que $T|_{V^{\perp}} : V^{\perp} \to V^{\perp}$ es un operador compacto y autoadjunto.

Luego, $\sigma_p \left(T|_{V^{\perp}} \right) = \{0\}$. En efecto, supongamos que existe $\lambda \in \sigma_p(T|_{V^{\perp}})$ tal que $\lambda \neq 0$. De este modo, existe $u \in V^{\perp} \setminus \{0_H\}$ tal que $T(u) = \lambda\, u$ y, en consecuencia, existe $k_0 \in \mathbb{N}$ tal que $\lambda = \lambda_{k_0}$. Por tanto, $u \in H_{k_0} \subset V$ lo que implica que $u \in V \cap V^{\perp}$ y, en consecuencia, $u = 0_H$ lo cual contradice nuestra suposición. El Corolario 3.47 implica que $T(v) = 0_H$ para todo $v \in V^{\perp}$ y, por tanto, $V^{\perp} \subset H_0$. Notemos además que $H_0 \subset V$ por lo que $V^{\perp} \subset V$ y esto implica que $V^{\perp} = \{0_H\}$. Aplicando el Corolario 1.29 concluimos que $\overline{V} = (V^{\perp})^{\perp} = H$ y esto prueba que V es denso en H.

El Teorema 1.46 asegura que existe $\mathcal{B}_0$ una base de Hilbert de H_0 y, dado que H_k es de dimensión finita para cada $k \in \mathbb{N}$, entonces es inmediato que existe $\mathcal{B}_k$ base de H_k para cada $k \in \mathbb{N}$. Definiendo $\mathcal{B} := \cup_{k=0}^{\infty} \mathcal{B}_k \subset H$ se sigue el resultado. ■

Por simplicidad damos la siguiente definición.

Definición 3.53 *Sea $T : H \to H$ una función lineal y continua y $\mathcal{I} \subset \mathbb{N}$ no vacío. Un subconjunto ortonormal $\{e_i \,:\, i \in \mathcal{I}\}$ en H de vectores propios de T con sus correspondientes valores propios no nulos $\{\lambda_i \,:\, i \in \mathcal{I}\}$ se llama **sistema básico de vectores propios y valores propios de** T si:*

$$T(u) = \sum_{i \in \mathcal{I}} \lambda_i \langle u, e_i\rangle\, e_i \qquad \forall\, u \in H.$$

Recordemos que el conjunto de valores propios no nulos de un operador compacto es finito o a lo más numerable con único punto de acumulación igual a 0. Teniendo esto en cuenta, damos el siguiente resultado pues el teorema espectral garantiza la existencia de un sistema básico de vectores propios de un operador compacto y autoadjunto.

Proposición 3.54 *Supongamos que H es separable. Sea $T : H \to H$ una función lineal y continua y $\mathcal{I} \subset \mathbb{N}$ no vacío arbitrario pero fijo. Si T es un operador compacto y autoadjunto y $\mathcal{B} = \{e_i \,:\, i \in \mathcal{I}\}$ y $\mathcal{L} = \{\lambda_i \,:\, i \in \mathcal{I}\}$ sistema básico de vectores propios y valores propios de T, entonces se cumple lo siguiente:*

(a) $\mathcal{B}$ *es una base de Hilbert de* $\overline{\operatorname{im} T}$ *y para cada* $u \in H$ *se cumple que:*

$$u = p_{\ker T}(u) + \sum_{i \in \mathcal{I}} \langle u, e_i \rangle \, e_i.$$

(b) H *tiene una base de Hilbert que consiste de un sistema básico de vectores propios si y sólo si* $\ker T = \{0_H\}$.

(c) Si $\lambda \in \mathbb{R} \setminus \{0\}$ *es un valor propio de* T*, entonces existe* $j_0 \in \mathcal{I}$ *tal que* $\lambda = \lambda_{j_0}$.

Demostración: (a) Dado que $e_j = \frac{1}{\lambda_j} T(e_j)$ para todo $j \in \mathcal{I}$ entonces $e_j \in \operatorname{im} T$ para todo $j \in \mathcal{I}$ y, por tanto, $\operatorname{lin}(\mathcal{B}) \subset \operatorname{im} T$.

Por otro lado, para cada $u \in H$ se tiene que:

$$T(u) = \sum_{i \in \mathcal{I}} \lambda_i \, \langle u, e_i \rangle \, e_i$$

lo que implica que $\operatorname{im} T \subset \overline{\operatorname{lin}(\mathcal{B})}$. En consecuencia, $\overline{\operatorname{im} T} = \overline{\operatorname{lin}(\mathcal{B})}$.

Dado que $\overline{\operatorname{im} T}$ es un subespacio vectorial cerrado de H entonces por el teorema del complemento ortogonal (ver Teorema 1.28) concluimos que $H = \overline{\operatorname{im} T} \oplus \left(\overline{\operatorname{im} T}\right)^{\perp}$. El Teorema 3.14 asegura que $\left(\overline{\operatorname{im} T}\right)^{\perp} = \ker T$ y, por tanto, $H = \overline{\operatorname{im} T} \oplus \ker T$ pues T es compacto y autoadjunto. De este modo, para cada $u \in H$ se tiene que $u = p_{\ker T}(u) + (u - p_{\ker T}(u))$ en donde $u - p_{\ker T}(u) \in \overline{\operatorname{im} T}$. Luego, por el Corolario 1.29 se sigue que $u - p_{\ker T}(u) = \sum_{i \in \mathcal{I}} \langle u, e_i \rangle \, e_i$ y, por lo tanto,

$$u = p_{\ker T}(u) + \sum_{i \in \mathcal{I}} \langle u, e_i \rangle \, e_i$$

como afirma el enunciado.

(b) $\Rightarrow$) : Sea $\mathcal{B} = \{e_i \,:\, i \in \mathcal{I}\}$ base de Hilbert de H constituida de vectores propios de T. Como $\operatorname{lin}(\mathcal{B})$ es denso en H, entonces $(\operatorname{lin}(\mathcal{B}))^{\perp} = \{0_H\}$ y el Lema 3.51 asegura que $\ker T = \{0_H\}$.

$\Leftarrow$:) Es inmediato del inciso anterior y el hecho de que $H = \overline{\operatorname{im} T} \oplus \ker T$ por ser T un operador autadjunto.

(c) Dado que $\lambda \in \mathbb{R} \setminus \{0\}$ es un valor propio de T entonces existe $v \in H \setminus \{0_H\}$ tal que $T(v) = \lambda \, u$. Si $\lambda \neq \lambda_j$ para todo $j \in \mathcal{I}$ entonces $\langle v, e_j \rangle = 0$ para todo $j \in \mathcal{I}$ (ver Proposición 3.44). En consecuencia, $\lambda \, u = T(v) = \sum_{i \in \mathcal{I}} \lambda_i \langle v, e_i \rangle \, e_i = 0_H$ lo cual es imposible pues $\lambda \neq 0$ y $v \neq 0_H$. Por tanto, existe $j_0 \in \mathcal{I}$ tal que $\lambda = \lambda_{j_0}$. ■

El siguiente resultado establece las condiciones para obtener el recíproco del teorema espectral de operador compactos-autoadjuntos.

Teorema 3.55 *Supongamos que H un espacio de Hilbert separable y sea $\mathcal{I} \subset N$ no vacío. Sea $\mathcal{O} = \{e_i : i \in \mathcal{I}\}$ un subconjunto ortonormal de H y $\mathcal{L} = \{\lambda_i : i \in \mathcal{I}\}$ un subconjunto de $\mathbb{R} \setminus \{0\}$ tal que $\lambda_{j+1} \leq \lambda_j$ para todo $j \in \mathcal{I}$ y, si $\mathcal{I} = \mathbb{N}$, entonces $\lim_{j\to\infty} \lambda_j = 0$. Definamos $T : H \to H$ como:*

$$T(u) = \sum_{i\in\mathcal{I}} \lambda_i \langle u, e_i \rangle \, e_i.$$

Entonces, T es un operador compacto y autoadjunto.

Demostración: Es inmediato que T es lineal. Por la desigualdad de Bessel, se tiene que $\|T(u)\|_H^2 = \sum_{i\in\mathcal{I}} \lambda_i^2 |\langle u, e_i\rangle|^2 \leq \lambda_1^2 \sum_{i\in\mathcal{I}} \lambda_i^2 |\langle u, e_i\rangle|^2 \leq \lambda_1^2 \|u\|_H^2$ para cada $u \in H$. Por tanto, T es continua. Veamos que T es un operador compacto y autoadjunto para lo cual consideremos los siguientes casos.

Caso 1. $\mathcal{I}$ es finito.

Supongamos que $\mathcal{I} = \{1, \ldots, m\}$ para $m \in \mathbb{N}$, entonces:

$$T(u) = \sum_{i=1}^{m} \lambda_i \langle u, e_i \rangle \, e_i \qquad \forall\, u \in H$$

lo que significa que T es un operador de rango finito (ver Definición 3.7) y, por tanto, un operador compacto (ver Proposición 3.8). Además, para $u, v \in H$ se tiene que:

$$\begin{aligned}\langle T(u), v\rangle &= \left\langle \sum_{i=1}^{m} \lambda_i \langle u, e_i\rangle \, e_i, v \right\rangle = \sum_{i=1}^{m} \lambda_i \langle u, e_i\rangle \, \langle e_i, v\rangle = \sum_{i=1}^{m} \lambda_i \langle v, e_i\rangle \, \langle e_i, u\rangle \\ &= \left\langle \sum_{i=1}^{m} \lambda_i \langle v, e_i\rangle \, e_i, u \right\rangle = \langle T(v), u\rangle = \langle u, T(v)\rangle\end{aligned}$$

lo que implica que T es un operador autoadjunto.

Caso 2. $\mathcal{I} = \mathbb{N}$.

Consideremos, para cada $m \in \mathbb{N}$, la función $T_m : H \to H$ dada por $T_m(u) = \sum_{i=1}^{m} \lambda_i \langle u, e_i\rangle \, e_i$. Del caso anterior concluimos que T_m es un operador compacto y autoadjunto para cada $m \in \mathbb{N}$.

Notemos que:

$$\|T_m(u) - T(u)\|_H^2 = \sum_{i=m+1}^{\infty} \lambda_i^2 |\langle u, e_i\rangle|^2 \leq \lambda_{m+1}^2 \|u\|_H^2 \qquad \forall\, u \in H, \qquad \forall m \in \mathbb{N}.$$

En consecuencia, $\|T_m - T\|_{\mathcal{L}_c(H,H)} \leq |\lambda_{m+1}|$ para cada $m \in \mathbb{N}$ y, por tanto, $\|T_m - T\|_{\mathcal{L}_c(H,H)} \to 0$ pues $\lambda_m \to 0$. De los Teoremas 3.4 y 3.19 concluimos que T es un operador compacto y autodadjunto. ■

Finalizamos esta sección con un ejemplo.

Ejemplo 3.56 *Consideremos la función $T : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ dada por $T(\bar{x}) := \left(\frac{x_j}{j}\right)$. Sabemos que T es un operador compacto y autoadjunto.*

Fijemos $k \in \mathbb{N}$. Definamos la sucesión $\bar{x} = (x_j) \in \ell_2(\mathbb{R})$ como $x_j := 0$ si $j \neq k$ y $x_j := 1$ si $j = k$. Entonces, es claro que $T_{\frac{1}{k}}(\bar{x}) = T(\bar{x}) - \frac{1}{k}\bar{x} = \bar{0}$ por lo que $\ker T_{\frac{1}{k}} \neq \{\bar{0}\}$. Es decir, $\frac{1}{k}$ es un valor propio de T.

En consecuencia, $\sigma_p(T) = \{\frac{1}{k} \;:\; k \in \mathbb{N}\}$. De lo anterior concluimos que la sucesión $\bar{x}_k = (x_j)$ dada por $x_j := 0$ si $j \neq k$ y $x_j := 1$ si $j = k$ es un vector propio de T asociado al valor propio $\frac{1}{k}$.

Sea $\bar{y} = (y_j) \in \ell_2(\mathbb{R})$, entonces

$$\langle \bar{y}, \bar{x}_k \rangle_{\ell_2(\mathbb{R})} = \sum_{j=1}^{\infty} y_j\, x_j = y_k \qquad \forall k \in \mathbb{N}$$

y, en consecuencia, la descomposición espectral de T está dada por:

$$T(\bar{y}) = \sum_{k=1}^{\infty} \langle \bar{y}, \bar{x}_k \rangle_{\ell_2(\mathbb{R})} \frac{\bar{x}_k}{k} = \sum_{k=1}^{\infty} \frac{y_k}{k}\, \bar{x}_k.$$

3.5.1. El teorema espectral de operadores compactos

Sea H un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|_H$.

En la sección anterior mostramos que dada una función lineal y continua $T : H \to H$ esta puede no tener ningún valor propio aún cuando T resulte ser un operador compacto o un operador autoadjunto. Sin embargo, cuando T es tanto un operador compacto como autoadjunto existe al menos un valor propio cuyo valor dependende de la norma de T en $\mathcal{L}_c(H, H)$. Entonces, si tenemos que T es únicamente un operador compacto en esta sección daremos una solución parcial en donde le asociaremos valores reales cuya propiedad sea, en esencia, similar a un valor propio.

Definición 3.57 *Sea $T : H \to H$ función lineal y continua. Decimos que T es un* ***operador positivo*** *si $\langle T(u), u \rangle \geq 0$ para todo $u \in H$.*

Proposición 3.58 *Supongamos que H es separable. Sea $T : H \to H$ función lineal y continua que es un operador compacto y autoadjunto. Entonces, T es positivo si y sólo si todos sus valores propios son no negativos.*

Demostración: Supongamos que T es un operador positivo y sea $\lambda \in \sigma_p(T)$. Entonces, existe $v \in H \setminus \{0_H\}$ tal que $T(v) = \lambda\, v$. Así pues,

$$0 \leq \langle T(v), v \rangle = \langle \lambda\, v, v \rangle = \lambda\, \langle v, v \rangle \leq \lambda\, \|v\|_H^2$$

por lo que necesariamente $\lambda \geq 0$.

Inversamente: El Teorema 3.52 asegura que existe $\mathcal{B} = \{e_k \in H \,:\, k \in \mathbb{N}\}$ una base de Hilbert de H conformada por vectores propios de T y tal que:

$$T(u) = \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle \, e_k \qquad \forall\, u \in H.$$

En consecuencia,

$$\langle T(u), u \rangle = \left\langle \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle \, e_k, u \right\rangle = \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle \, \langle e_k, u \rangle = \sum_{k=1}^{\infty} \lambda_k \, |\langle u, e_k \rangle|^2 \geq 0 \qquad \forall\, u \in H$$

lo que implica que T es un operador positivo. ■

Teorema 3.59 *Supongamos que H es separable y sea $T : H \to H$ función lineal y continua. Si T es un operador compacto y autoadjunto, entonces existe una función $S : H \to H$ tal que S es un operador compacto y autoadjunto y, $S \circ S = T$.*

Demostración: El Teorema 3.52 asegura que existe $\mathcal{B} = \{e_k \,:\, k \in \mathbb{N}\}$ una base de Hilbert de H conformada por vectores propios de T y tal que:

$$T(u) = \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle \, e_k \qquad \forall\, u \in H.$$

en donde $\lambda_k \geq 0$ pues T es positivo (ver Proposición 3.58). Definamos $S : H \to H$ como:

$$S(u) := \sum_{k=1}^{\infty} \sqrt{\lambda_k} \, \langle u, e_k \rangle \, e_k \qquad \forall\, u \in H.$$

Entonces, para cada $u \in H$ se tiene que:

$$\begin{aligned}
(S(S(u))) = S\left(\sum_{k=1}^{\infty} \sqrt{\lambda_k} \, \langle u, e_k \rangle \, e_k \right) &= \sum_{j=1}^{\infty} \sqrt{\lambda_j} \, \left\langle \left(\sum_{k=1}^{\infty} \sqrt{\lambda_k} \, \langle u, e_k \rangle \, e_k \right), e_j \right\rangle \, e_j \\
&= \sum_{j=1}^{\infty} \sqrt{\lambda_j} \, \left\langle \sqrt{\lambda_j} \, u, e_j \right\rangle \, e_j \\
&= \sum_{j=1}^{\infty} \lambda_j \, \langle u, e_j \rangle \, e_j \\
&= T(u).
\end{aligned}$$

El Teorema 3.55 asegura que S es un operador compacto y autoadjunto. Luego, por la Proposición 3.54 concluimos que un valor propio de S es el 0 ó algún $\sqrt{\lambda_k}$ para $k \in \mathbb{N}$. De este modo, S es un operador positivo por la Proposición 3.58. ■

Teorema 3.60 *Supongamos que H es separable y sea $T : H \to H$ un operador compacto. Entonces, existe (λ_k) una sucesión decreciente de números reales no negativos y $\{e_k : k \in \mathbb{N}\}$, $\{\xi_k : k \in \mathbb{N}\}$ conjuntos ortonormales de H tales que:*

$$T(u) = \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle \xi_k \qquad \forall\, u \in H.$$

Demostración: El resultado se satisface trivialmente si T es el operador constante con valor 0_H. Así pues, supongamos que $T \neq 0_H$. Consideremos $T^* : H \to H$ el operador adjunto de T y definamos $\Lambda := T^* \circ T : H \to H$ que es claramente un operador compacto (ver Proposición 3.5). Se tiene que $\Lambda^* = (T^* \circ T)^* = T^* \circ T = \Lambda$ por el Teorema 3.14. Es decir, Λ es un operador autoadjunto. Observemos que, $\langle \Lambda(u), u \rangle = \langle T(u), T(u) \rangle = \|T(u)\|_H^2 \geq 0$ para cada $u \in H$ lo que implica que Λ es un operador positivo.

Como Λ es un operador compacto y autoadjunto, el Teorema 3.25 asegura que existe $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ una base de Hilbert de H constituida de vectores propios de Λ y (μ_k) una sucesión decreciente y positiva de valores propios de Λ tal que:

$$\Lambda(u) = \sum_{k=1}^{\infty} \mu_k \langle u, e_k \rangle\, e_k \qquad \forall\, u \in H.$$

Definiendo $\xi_k := \frac{1}{\sqrt{\mu_k}} T(e_k) \in H$ para todo $k \in \mathbb{N}$ se tiene que:

$$\langle \xi_k, \xi_j \rangle = \frac{1}{\sqrt{\mu_k}} \frac{1}{\sqrt{\mu_j}} \langle T(e_k), T(e_j) \rangle = \frac{1}{\sqrt{\mu_k}} \frac{1}{\sqrt{\mu_j}} \langle \Lambda(e_k), e_j \rangle = \frac{\mu_k}{\sqrt{\mu_k}} \frac{1}{\sqrt{\mu_j}} \langle e_k, e_j \rangle$$

por lo que $\{\xi_k : k \in \mathbb{N}\}$ es un conjunto ortonormal de H. La Proposición 3.54 asegura que:

$$u = p_{\ker \Lambda}(u) + \sum_{k=1}^{\infty} \langle u, e_k \rangle\, e_k \qquad \forall\, u \in H. \tag{3.13}$$

Es claro que $\ker T \subset \ker \Lambda$. Inversamente, si $u \in \ker \Lambda$ entonces $\Lambda(u) = T^*(T(u)) = 0_H$. Así, $0 = \langle \Lambda(u), u \rangle = \langle T(u), T(u) \rangle = \|T(u)\|_H^2$ de donde se sigue que $T(u) = 0_H$ y, por tanto, $u \in \ker T$. De este modo, $p_{\ker \Lambda} = p_{\ker T}$.

Aplicando la linealidad y continuidad de T en (3.13) concluimos que:

$$T(u) = T(p_{\ker \Lambda}(u)) + \sum_{k=1}^{\infty} \langle u, e_k \rangle\, T(e_k) = \sum_{k=1}^{\infty} \sqrt{\mu_k} \langle u, e_k \rangle\, \xi_k \qquad \forall\, u \in H.$$

El resultado se sigue definiendo la sucesión (δ_k) en $\mathbb{R}$ como $\delta_k := \sqrt{\mu_k}$. ■

Capítulo 4

Medidas Gaussianas en espacios de Hilbert

En los cursos básicos de probabilidad se define a una *medida Gaussiana* como una medida de probabilidad en un espacio euclidiano de dimensión finita, $\mathbb{R}^N$ y es de gran importancia pues está directamente relacionada con la distribución normal y algunos resultados fundamentales en estadística, por ejemplo, el teorema del límite central.

En un espacio de Hilbert H sobre $\mathbb{R}$ es posible considerar $\mathcal{B}(H)$ la σ-álgebra generada por la topología inducida por la norma y, que conocemos como la σ-álgebra de Borel de H. Esto nos permite considerar medidas $\mu : \mathcal{B}(H) \to \mathbb{R}$ y, en particular, medidas de probabilidad.

El objetivo de este capítulo es extender el concepto de medida Gaussiana para espacios de Hilbert separables y de dimensión infinita. Profundizaremos en los resultados que se aprenden en un curso básico de teoría de la probabilidad y los extenderemos a espacios distintos del euclidiano a través de los resultados que genera el producto escalar y la norma definidas en H como lo son el teorema de representación de Fréchet-Riesz, la teoría espectral de operadores compactos y las bases de Hilbert.

Este capítulo está basado enteramente en [9], [20], [21], [23] y [86].

4.1. Definiciones y propiedades básicas

Sea H un espacio de Hilbert sobre $\mathbb{R}$ con producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|_H$ que es separable. Denotaremos por $\mathcal{B}(H)$ la σ-álgebra de Borel de H.

Denotaremos por $\mathbb{P}(H)$ al conjunto de todas las medidas de probabilidad definidas en el espacio medible $(H, \mathcal{B}(H))$.

Definición 4.1 *Sea $\mu \in \mathbb{P}(H)$ y $k \in \mathbb{N}$. Decimos que μ tiene k-**ésimo momento finito** si:*

$$\int_H \|u\|_H^k \, d\mu(u) < \infty.$$

Por ejemplo, sea $\delta_{0_H} \in \mathbb{P}(H)$ la medida unitaria concentrada en 0_H. Es decir, $\delta_{0_H} : \mathcal{B}(H) \to \{0,1\}$ está dada por $\delta_{0_H}(A) := 1$ si $0_H \in A$ ó $\delta_{0_H}(A) := 0$ si $0_H \notin A$. Entonces, para todo $k \in \mathbb{N}$ se tiene que:

$$\int_H \|u\|_H^k \, d\delta_{0_H}(u) = \|0_H\|_H^k = 0 < \infty.$$

Es inmediato que si μ tiene k-ésimo momento finito con $k > 1$ entonces μ tiene primer momento finito.

Sea $\mu \in \mathbb{P}(H)$ con primer momento finito y consideremos la función $F : H \to \mathbb{R}$ dada por:

$$F(u) := \int_H \langle u, v \rangle \, d\mu(v). \tag{4.1}$$

Es claro que F es lineal. Aplicando la desigualdad de Cauchy-Schwarz obtenemos:

$$|F(u)| \leq \int_H |\langle u, v \rangle| \, d\mu(v) \leq \|u\|_H \int_H \|v\|_H \, d\mu(v) \qquad \forall\, u \in H$$

lo que implica que F es continua en H. Por el teorema de representación de Fréchet-Riesz existe un único $m \in H$ tal que:

$$\langle u, m \rangle = F(u) = \int_H \langle u, v \rangle \, d\mu(v) \qquad \forall\, u \in H. \tag{4.2}$$

En este caso, m se conoce como la media de μ.

Definición 4.2 *Sea $\mu \in \mathbb{P}(H)$ con primer momento finito. La **media de** μ es el único elemento m de H que cumple que:*

$$\langle u, m \rangle = \int_H \langle u, v \rangle \, d\mu(v) \qquad \forall\, u \in H.$$

Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito. Notemos que, de (4.1), se sigue que μ tiene primer momento finito y, por tanto, la media m de μ está bien definida. Consideremos la función $G : H \times H \to \mathbb{R}$ como:

$$G(u, v) := \int_H \langle u, w - m \rangle \langle v, w - m \rangle \, d\mu(w). \tag{4.3}$$

La función G es una forma bilineal continua pues para cualesquiera $u, v \in H$ se tiene que:

$$|G(u,v)| \leq \int_H |\langle u, w-m\rangle| \, |\langle v, w-m\rangle| \, d\mu(w) \leq \|u\|_H \|v\|_H \int_H \|w-m\|_H^2 \, d\mu(w). \tag{4.4}$$

Por el Teorema 1.51 existe una única función $Q \in \mathcal{L}_c(H,H)$ tal que:

$$\langle Q(u), v\rangle = \int_H \langle u, w-m\rangle \, \langle v, w-m\rangle \, d\mu(w) \qquad \forall\, u, v \in H. \tag{4.5}$$

En este caso, Q se conoce como la covarianza de μ.

Definición 4.3 *Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito. La* ***covarianza de*** *μ es el único elemento Q de $\mathcal{L}_c(H,H)$ que cumple que:*

$$\langle Q(u), v\rangle = \int_H \langle u, w-m\rangle \, \langle v, w-m\rangle \, d\mu(w) \qquad \forall\, u, v \in H$$

en donde m denota la media de μ.

Veamos unas propiedades de este operador.

Proposición 4.4 *Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito, $Q \in \mathcal{L}_c(H,H)$ su convarianza y m su media. Entonces:*

(i) Q es un operador autoadjunto y positivo.

(ii) Si $\mathcal{B} = \{e_k \,:\, k \in \mathbb{N}\}$ es una base de Hilbert entonces:

$$\sum_{k=1}^{\infty} \langle Q(e_k), e_k\rangle = \int_H \|u-m\|_H^2 \, d\mu(u) < \infty \tag{4.6}$$

(iii) Q es un operador compacto.

Demostración: (i) Sean $u, v \in H$. Entonces $\langle Q(u), u\rangle = \int_H \langle u, w-m\rangle^2 \, d\mu(w) \geq 0$ lo que prueba que Q es positivo. Luego, $\langle Q(u), v\rangle = \int_H \langle u, w-m\rangle \, \langle v, w-m\rangle \, d\mu(w) = \int_H \langle v, w-m\rangle \, \langle u, w-m\rangle \, d\mu(w) = \langle u, Q(v)\rangle$ lo que implica que Q es un operador autoadjunto.

(ii) De la identidad (4.5) obtenemos que:

$$\langle Q(e_k), e_k\rangle = \int_H |\langle e_k, w-m\rangle|^2 \, d\mu(w) \qquad \forall\, k \in \mathbb{N}.$$

Sumando sobre $k \in \mathbb{N}$ y usando el teorema de la convergencia monótona en la identidad anterior obtenemos que:

$$\sum_{k=1}^{\infty} \langle Q(e_k), e_k \rangle = \sum_{k=1}^{\infty} \int_H |\langle e_k, w - m \rangle|^2 \, d\mu(w) = \int_H \sum_{k=1}^{\infty} |\langle e_k, w - m \rangle|^2 \, d\mu(w).$$

Aplicando la fórmula de Parseval concluimos que:

$$\sum_{k=1}^{\infty} \langle Q(e_k), e_k \rangle = \int_H \sum_{k=1}^{\infty} |\langle e_k, w - m \rangle|^2 \, d\mu(w) = \int_H \|w - m\|_H^2 \, d\mu(w).$$

(iii) Definamos la sucesión de funciones $Q_k : H \to H$ como:

$$Q_k(u) := \sum_{j=1}^{k} \langle Q(u), e_j \rangle \, e_j.$$

Sea $k \in \mathbb{N}$. Es inmediato que Q_k es lineal y continua. Observemos que $\operatorname{im} Q_k \subset \operatorname{lin}(\{e_1, \ldots, e_k\})$ para cada $k \in \mathbb{N}$ de modo que la función Q_k es un operador de rango finito y, por tanto, compacto (ver Proposición 3.8). Por el inciso *(i)* tenemos que Q es un operador autoadjunto y positivo de modo que existe una única función $Q^{1/2} : H \to H$ lineal y continua que es un operador autoadjunto y positivo tal que $Q^{1/2} \circ Q^{1/2} = Q$ (ver Teorema 3.59). Así pues,

$$\begin{aligned}
\|Q_k(u) - Q(u)\|_H^2 &= \sum_{j=k+1}^{\infty} |\langle Q(u), e_j \rangle|^2 \\
&= \sum_{j=k+1}^{\infty} |\langle Q^{1/2}(u), Q^{1/2}(e_j) \rangle|^2 \\
&\leq \|Q\|_{\mathcal{L}_c(H,H)} \, \|u\|_H^2 \sum_{j=k+1}^{\infty} \|Q^{1/2}(e_j)\|_H^2 \\
&= \|Q\|_{\mathcal{L}_c(H,H)} \, \|u\|_H^2 \sum_{j=k+1}^{\infty} \langle Q(e_j), e_j \rangle \qquad \forall\, u \in H
\end{aligned}$$

y, en consecuencia:

$$\|Q_k - Q\|_{\mathcal{L}_c(H,H)} \leq \|Q\|_{\mathcal{L}_c(H,H)} \sum_{j=k+1}^{\infty} \langle Q(e_j), e_j \rangle.$$

Tomando el límite cuando $k \to \infty$ en la identidad anterior concluimos que $\|Q_k - Q\|_{\mathcal{L}_c(H,H)} \to 0$ por el inciso *(ii)* y el criterio de Cauchy para series. El Teorema 3.5 asegura que Q es un operador compacto. ∎

Definición 4.5 *Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito, $Q \in \mathcal{L}_c(H,H)$ su convarianza y $\mathcal{B} = \{e_k \,:\, k \in \mathbb{N}\}$ una base de Hilbert de H. Definimos la* ***traza de*** *Q como el número real dado por:*

$$\operatorname{Tr} Q := \sum_{k=1}^{\infty} \langle Q(e_k), e_k \rangle = \int_H \|u - m\|_H^2 \, d\mu(u).$$

Observa que la definición de la traza de Q no depende de la base de Hilbert elegida.

Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito y $Q \in \mathcal{L}_c(H,H)$ su convarianza. Dado que Q es un operador compacto y autoadjunto, por el teorema espectral (ver Teorema 3.55) existe base de Hilbert $\{\xi_k \,:\, k \in \mathbb{N}\}$ y una sucesión de números reales no negativos tales que:

$$Q(\xi_k) = \lambda_k \, \xi_k \qquad \forall\, k \in \mathbb{N}. \tag{4.7}$$

Por tanto, de (4.6) concluimos que:

$$\operatorname{Tr} Q = \sum_{k=1}^{\infty} \langle Q(\xi_k), \xi_k \rangle = \sum_{k=1}^{\infty} \lambda_k \langle \xi_k, \xi_k \rangle = \sum_{k=1}^{\infty} \lambda_k < \infty. \tag{4.8}$$

En lo que resta y, por simplicidad, denotaremos por $\mathcal{L}_c(H) := \mathcal{L}_c(H,H)$ al espacio de las funciones lineales y continuas $f : H \to H$. Al subconjunto de éste constituido por las funciones que son operadores autoadjuntos y positivos lo denotaremos por $\mathcal{L}_c^+(H)$. Finalmente, al subconjunto de $\mathcal{L}_c^+(H)$ constituido por los operadores compactos que tienen traza finita lo denotaremos por $\mathcal{L}_1^+(H)$.

Proposición 4.6 *Sea $\mu \in \mathbb{P}(H)$ con segundo momento finito, $Q \in \mathcal{L}_c(H)$ su convarianza y m su media. Entonces:*

$$\int_H \|u\|_H^2 \, d\mu(u) = \operatorname{Tr} Q + \|m\|_H^2.$$

Demostración: Por cálculos directos se tiene que:

$$\|u - m\|_H^2 = \|u\|_H^2 - 2\langle u, m \rangle + \|m\|_H^2 \qquad \forall\, u \in H.$$

Así pues,

$$\int_H \|u\|_H^2 \, d\mu(u) = \int_H \|u - m\|_H^2 \, d\mu(u) + 2\int_H \langle u, m \rangle \, d\mu(u) - \int_H \|m\|_H^2 \, d\mu(u).$$

Dado que m es la media de μ, entonces $\langle m, m \rangle = \int_H \langle u, m \rangle \, d\mu(u)$ de modo que:

$$2\int_H \langle u, m \rangle \, d\mu(u) - \int_H \|m\|_H^2 \, d\mu(u) = 2\|m\|_H^2 - \|m\|_H^2.$$

En consecuencia,

$$\int_H \|u\|_H^2 \, d\mu(u) = \int_H \|u - m\|_H^2 \, d\mu(u) + \|m\|_H^2 = \operatorname{Tr} Q + \|m\|_H^2$$

como afirma el enunciado. ■

Definición 4.7 *Sea $\mu \in \mathbb{P}(H)$. Definimos la* ***transformada de Fourier de*** *μ como:*

$$\widehat{\mu}(h) := \int_H \exp\{i\langle h, u\rangle\} \, d\mu(u) \qquad \forall\, h \in H.$$

Es posible demostrar que la transformada de Fourier de μ es única.

Teorema 4.8 *Si $\mu, \nu \in \mathbb{P}(H)$ son tales que $\widehat{\mu}(h) = \widehat{\nu}(h)$ para todo $h \in H$ entonces, $\mu = \nu$.*

Demostración: Consulta, por ejemplo, [86; 87]. ■

4.2. Medidas Gaussianas

Definición 4.9 *Una medida $\mu \in \mathbb{P}(H)$ se llama una* ***medida Gaussiana*** *si existen $m \in H$ y $Q \in \mathcal{L}_1^+(H)$ tales que:*

$$\widehat{\mu}(h) = \exp\{i\langle m, h\rangle\} \exp\{-\tfrac{1}{2}\langle Q(h), h\rangle\} \qquad \forall\, h \in H.$$

En este caso, escribimos a μ como $\mu = N_{m,Q}$ y si $m = 0_H$ entonces $\mu = N_Q$.

Si $\ker Q = \{0_H\}$ decimos que la medida μ es ***no degenerada****.*

En las siguientes secciones veremos algunos ejemplos importantes de medidas Gaussianas en distintos espacios de Hilbert conocidos. El objetivo de esta sección es demostrar que en todo espacio de Hilbert separable H existe una única medida Gaussiana $N_{m,Q}$ para $m \in H$ y $Q \in \mathcal{L}_1^+(H)$ dados.

4.2.1. Medidas Gaussianas en $\mathbb{R}$

Consideremos el espacio de Hilbert $\mathbb{R}$ con el producto escalar definido en el Ejemplo 1.2.

Proposición 4.10 *Sea $\mathbb{R}$ con el producto escalar $\langle x, y\rangle_{\mathbb{R}} = xy$ para todos $x, y \in \mathbb{R}$.*

(i) Sea $m \in \mathbb{R}$ y consideremos δ_m la medida unitaria concentrada en m. Entonces, δ_m es una medida Gaussiana con $\delta_m = N_{m,0}$. Más aún, m es la media de δ_m y la función constante 0 es la covarianza de δ_m.

(ii) Sea $\lambda > 0$ y $m \in \mathbb{R}$. Consideremos $\mu : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ tal que μ cumple que:

$$d\mu(x) := \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{\frac{-(x-m)^2}{2\lambda}\right\} dx. \tag{4.9}$$

Entonces, $\mu = N_{m,\lambda}$. Más aún, m es la media y λ es la covarianza de μ.

Demostración: (i) Es inmediato que δ_m tiene k-ésimo momento finito para todo $k \in \mathbb{N}$ pues:

$$\int_{\mathbb{R}} |x|^k \, d\delta_m(x) = |m|^k < \infty.$$

La transformada de Fourier de δ_m está dada por:

$$\widehat{\delta}_m(h) = \int_{\mathbb{R}} \exp\{i\langle h, x\rangle_{\mathbb{R}}\} \, d\delta_m(x) = \exp\{i\langle h, m\rangle_{\mathbb{R}}\} \qquad \forall\, h \in \mathbb{R}$$

de donde se sigue que $\delta_m = N_{m,0}$. Luego, es claro que m es su media pues para todo $x \in \mathbb{R}$ se tiene que:

$$\int_{\mathbb{R}} x\, y \, d\delta_m(y) = x\, m = \langle x, m\rangle_{\mathbb{R}}.$$

Del mismo modo, para cualesquiera $x, y \in \mathbb{R}$ se tiene que:

$$\int_{\mathbb{R}} xy(z-m)^2 \, d\delta_m(z) = xy(0-0)^2 = 0 = \langle x, 0\rangle_{\mathbb{R}} = \langle 0, y\rangle_{\mathbb{R}}$$

y, por lo tanto, 0 es la covarianza de δ_m.

(ii) Es claro que la función μ que cumple (4.9) es una medida de probabilidad pues:

$$\mu(\mathbb{R}) = \frac{1}{\sqrt{2\pi\lambda}} \int_{-\infty}^{\infty} \exp\left\{\frac{-(x-m)^2}{2\lambda}\right\} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{x^2}{2}\right\} dx = 1.$$

Ahora bien,

$$\begin{aligned}
\widehat{\mu}(h) &= \int_{\mathbb{R}} \exp\left\{i\langle h, x\rangle_{\mathbb{R}}\right\} \, d\mu(x) \\
&= \int_{-\infty}^{\infty} \exp\{ihx\} \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{\frac{-(x-m)^2}{2\lambda}\right\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{\frac{-(x^2 - 2x(m - ih\lambda) + m^2)}{2\lambda}\right\} dx \\
&= \exp\left\{\frac{-m^2 + (m - ih\lambda)^2}{2\lambda}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{\frac{-(x - (m - ih\lambda))^2}{2\lambda}\right\} dx \\
&= \exp\left\{\frac{-m^2 + (m - ih\lambda)^2}{2\lambda}\right\} = \exp\{ihm\} \exp\{-\tfrac{1}{2}h^2\lambda\} \qquad \forall\, h \in \mathbb{R}
\end{aligned}$$

lo que implica que $\mu = N_{m,\lambda}$. Finalmente probaremos que:

$$\mathcal{I}_1 = \int_{\mathbb{R}} x\, d\mu(x) = m \quad \text{y} \quad \mathcal{I}_2 = \int_{\mathbb{R}} (x-m)^2\, d\mu(x) = \lambda.$$

Haciendo el cambio de variable $\xi := \frac{x-m}{\sqrt{\lambda}}$ las integrales anteriores se reducen a lo siguiente:

$$\mathcal{I}_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sqrt{\lambda}\xi + m) \exp\left\{\frac{-\xi^2}{2}\right\} d\xi \quad \text{y} \quad \mathcal{I}_2 = \frac{\lambda}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi^2 \exp\left\{\frac{-\xi^2}{2}\right\} d\xi.$$

Resolvamos primero $\int_{-\infty}^{\infty} \xi \exp\left\{\frac{-\xi^2}{2}\right\} d\xi$.

$$\begin{aligned}
\int_{-\infty}^{\infty} \xi \exp\left\{\frac{-\xi^2}{2}\right\} d\xi &= \int_{-\infty}^{0} \xi \exp\left\{\frac{-\xi^2}{2}\right\} d\xi + \int_{0}^{\infty} \xi \exp\left\{\frac{-\xi^2}{2}\right\} d\xi \\
&= \left[-\exp\left\{-\frac{\xi^2}{2}\right\}\right]_{-\infty}^{0} + \left[-\exp\left\{-\frac{\xi^2}{2}\right\}\right]_{0}^{\infty} \\
&= -1 + 1 = 0.
\end{aligned}$$

En consecuencia,

$$\begin{aligned}
\mathcal{I}_1 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sqrt{\lambda}\xi + m) \exp\left\{\frac{-\xi^2}{2}\right\} d\xi \\
&= \frac{\sqrt{\lambda}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi \exp\left\{\frac{-\xi^2}{2}\right\} d\xi + \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-\xi^2}{2}\right\} d\xi \\
&= \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-\xi^2}{2}\right\} d\xi \\
&= m.
\end{aligned}$$

Por otra parte, haciendo integración por partes obtenemos que:

$$\begin{aligned}
\int_{-\infty}^{\infty} \xi^2 \exp\left\{\frac{-\xi^2}{2}\right\} d\xi &= \int_{-\infty}^{0} \xi \left[\xi \exp\left\{\frac{-\xi^2}{2}\right\}\right] d\xi + \int_{0}^{\infty} \xi \left[\xi \exp\left\{\frac{-\xi^2}{2}\right\}\right] d\xi \\
&= \left[-\exp\left\{-\frac{\xi^2}{2}\right\}\right]_{-\infty}^{0} + \int_{-\infty}^{0} \exp\left\{-\frac{1}{2}\xi^2\right\} d\xi \\
&\quad + \left[-\xi \exp\left\{-\frac{1}{2}\xi^2\right\}\right]_{0}^{\infty} + \int_{0}^{\infty} \exp\left\{-\frac{1}{2}\xi^2\right\} d\xi \\
&= \int_{-\infty}^{0} \exp\left\{-\frac{1}{2}\xi^2\right\} d\xi + \int_{0}^{\infty} \exp\left\{-\frac{1}{2}\xi^2\right\} d\xi \\
&= \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\xi^2\right\} d\xi \\
&= \sqrt{2\pi}.
\end{aligned}$$

Así pues,

$$\mathcal{I}_2 = \frac{\lambda}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi^2 \exp\left\{\frac{-\xi^2}{2}\right\} d\xi = \lambda.$$

Esto concluye la demostración. ■

Un resultado útil es el siguiente pues nos permitirá concluir que para $\lambda > 0$, la medida Gaussiana N_λ tiene j-ésimo momento finito para todo $j \in \mathbb{N}$.

Por simplicidad supongamos que $\lambda = 1$ y consideremos la medida N_λ dada en (4.9). La Proposición 4.10 asegura que:

$$\int_{\mathbb{R}} x^1 \, dN_\lambda(x) = 0.$$

Fijemos $j \in \mathbb{N}$ con $j > 1$. Queremos encontrar una expresión para la integral

$$\mathcal{I} := \int_{\mathbb{R}} x^j \, dN_\lambda(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^j \exp\{-\tfrac{x^2}{2}\} \, dx. \tag{4.10}$$

Observa que $\mathcal{I}$ puede escribirse como:

$$\mathcal{I} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{j-1} (x \exp\{-\tfrac{x^2}{2}\}) \, dx. \tag{4.11}$$

Paso 1: Definamos a las funciones $f_1(x) := x^{j-1}$ y $dg_1(x) := x \exp\{-\frac{x^2}{2}\}$. Haciendo integración por partes se tiene que:

$$\begin{aligned}
\mathcal{I} &= \frac{1}{\sqrt{2\pi}} \left(\left[-x^{j-1} \exp\{\tfrac{-x^2}{2}\} \right]_{-\infty}^{\infty} + (j-1) \int_{-\infty}^{\infty} x^{j-2} \exp\{-\tfrac{x^2}{2}\} \, dx \right) \\
&= \frac{j-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{j-2} \exp\{-\tfrac{x^2}{2}\} \, dx \\
&= (j-1) \int_{\mathbb{R}} x^{j-2} \, dN_\lambda(x)
\end{aligned}$$

Paso 2: De manera similar al paso anterior, definamos $f_2(x) = x^{j-3}$ y $dg_2(x) := x \exp\{-\frac{x^2}{2}\}$. En consecuencia,

$$\begin{aligned}
\int_{\mathbb{R}} x^{j-2} \, dN_\lambda(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{j-2} \exp\{-\tfrac{x^2}{2}\} \, dx. \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{j-3} (x \exp\{-\tfrac{x^2}{2}\}) \, dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[-x^{j-3} \exp\{\tfrac{-x^2}{2}\} \right]_{-\infty}^{\infty} + (j-3) \int_{-\infty}^{\infty} x^{j-4} \exp\{-\tfrac{x^2}{2}\} \, dx \right) \\
&= \frac{(j-3)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{j-4} \exp\{-\tfrac{x^2}{2}\} \, dx \\
&= (j-3) \int_{\mathbb{R}} x^{j-4} \, dN_\lambda(x)
\end{aligned}$$

Sea $m \in \mathbb{N}$ tal que $2m \leq j \leq 2m+1$. Entonces, continuando como los pasos anteriores concluimos, a través de un argumento inductivo, que, para cada $n \in \{0, 1, \ldots, m-1\}$:

$$\int_{\mathbb{R}} x^{j-2n} \, dN_\lambda(x) = (j - (2n+1)) \int_{\mathbb{R}} x^{j-2(n+1)} \, dN_\lambda(x) \tag{4.12}$$

Dado que $\int_{\mathbb{R}} x^0 \, dN_\lambda(x) = 1$ podemos concluir por un argumento recursivo que:

$$\mathcal{I} = (j-1)(j-3)(j-5)\cdots(3)(1). \tag{4.13}$$

Así pues, si $j = 2m+1$ entonces $\mathcal{I} = 0$ por lo que supondremos que $j = 2m$. En consecuencia,

$$\begin{aligned} \mathcal{I} = (j-1)(j-3)(j-5)\cdots(3)(1) &= \frac{(j)(j-1)(j-2)(j-3)(j-4)\cdots(4)(3)(2)(1)}{(j)(j-2)(j-4)\cdots(4)(2)} \\ &= \frac{j!}{j!!} = \frac{j!!(j-1)!!}{j!!} = (j-1)!! = (2m-1)!! = \frac{(2m)!}{2^m m!} \end{aligned}$$

en donde la expresión $j!!$ denota a la función doble factorial (ver [81]).

Podemos ahora demostrar la siguientes proposición.

Proposición 4.11 *Sea $\lambda > 0$ y N_λ la medida Gaussiana dada en* (4.9)*. Entonces, para cada $j \in \mathbb{N}$ se tiene que:*

$$\int_{\mathbb{R}} x^{2j} \, dN_\lambda(x) = \frac{(2j)!}{2^j j!} \lambda^j. \tag{4.14}$$

Demostración: Sea $j \in \mathbb{N}$. Haciendo el cambio de variable $\xi := \frac{x}{\sqrt{\lambda}}$ en (4.14) se tiene de la discusión anterior que:

$$\begin{aligned} \int_{\mathbb{R}} x^{2j} \, dN_\lambda(x) &= \frac{1}{\sqrt{2\pi\lambda}} \int_{-\infty}^{\infty} x^{2j} \exp\{-\tfrac{x^2}{2\lambda}\} \, dx \\ &= \frac{\lambda^j}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi^{2j} \exp\{-\tfrac{\xi^2}{2}\} \, d\xi \\ &= \lambda^j \frac{(2j)!}{2^j j!} \end{aligned}$$

como afirma el enunciado. ■

4.2.2. Medidas Gaussianas en $\mathbb{R}^N$

Sea $N \in \mathbb{N}$ con $N > 1$ fijo. Consideremos el espacio de Hilbert $\mathbb{R}^N$ con el producto escalar definido en el Ejemplo 1.2.

Sean $m \in \mathbb{R}^N$ y $Q \in \mathcal{L}_1^+(\mathbb{R}^N)$. Dado que Q es un autoadjunto, el Teorema 3.55 asegura que existe una base $\{e_1, \ldots, e_N\}$ de $\mathbb{R}^N$ y números reales no negativos $\lambda_1, \ldots, \lambda_N$ tales que:

$$Q(e_j) = \lambda_j \, e_j \qquad \forall \, j = 1, \ldots, N. \tag{4.15}$$

Definamos

$$m_j := \langle m, e_j \rangle_{\mathbb{R}^N} \qquad \forall \, j = 1, \ldots, N. \tag{4.16}$$

Sea $(\mathbb{R}, \mathcal{B}(\mathbb{R}), N_{m_j,\lambda_j})$ el espacio de medida dado en la Proposición 4.10 para cada $j = 1, \ldots, N$. Es decir,

$$dN_{m_j,\lambda_j}(x_j) = \frac{1}{\sqrt{2\pi\lambda_j}} \exp\left\{ \frac{-(x_j - m_j)^2}{2\lambda_j} \right\} dx_j \,, \qquad \forall \, j = 1, \ldots, N.$$

Así pues, consideremos el espacio de medida producto en $\mathbb{R}^N$ con la σ-álgebra $\mathcal{B}(\mathbb{R}^N) = \mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})$ y la medida producto:

$$\mu := \bigotimes_{j=1}^{N} N_{m_j,\lambda_j}. \tag{4.17}$$

Sea $j \in \{1, \ldots, N\}$. Dado que N_{m_j,λ_j} es una medida Gaussiana en $\mathbb{R}$ por la Proposición 4.10 entonces:

$$\widehat{N_{m_j,\lambda_j}}(h_j) = \int_{\mathbb{R}} \exp\{i \, x_j \, h_j\} \, dN_{m_j,\lambda_j}(x_j) = \exp\{i \, m_j \, h_j\} \exp\{-\tfrac{1}{2}\lambda_j \, h_j^2\} \tag{4.18}$$

para todo $h_j \in \mathbb{R}$. Más aún, m_j y λ_j son la media y covarianza de N_{m_j,λ_j} lo que implica que m_j es el único elemento de $\mathbb{R}$ que cumple que:

$$m_j \, x_j = \int_{\mathbb{R}} x_j \, y_j \, dN_{m_j,\lambda_j}(y_j) \qquad \forall \, x_j \in \mathbb{R} \tag{4.19}$$

y, λ_j es el único elemento de $\mathbb{R}$ que cumple que:

$$\lambda_j \, x_j \, y_j = \int_{\mathbb{R}} (x_j z_j - x_j m_j)(y_j z_j - y_j m_j) \, dN_{m_j,\lambda_j}(z_j) \qquad \forall \, x_j, y_j \in \mathbb{R}. \tag{4.20}$$

Teniendo estas propiedades en cuenta demostraremos, a través del teorema de Tonelli-Fubini que la medida dada en (4.17) es una medida Gaussiana en $\mathbb{R}^N$ con media igual a $m \in \mathbb{R}^N$ y covarianza $Q : \mathbb{R}^N \to \mathbb{R}^N$.

Proposición 4.12 *La medida $\mu : \mathcal{B}(\mathbb{R}^N) \to \mathbb{R}$ definida en (4.17) es una medida Gaussiana $N_{m,Q}$. Más aún, m es la media de μ y el operador $Q \in \mathcal{L}_1^+(\mathbb{R}^N)$ es su convarianza.*

Demostración: Para cada $h = (h_1, \ldots, h_N) \in \mathbb{R}^N$, aplicando el teorema de Tonelli, se tiene que:

$$\begin{aligned}
\widehat{\mu}(h) &= \int_{\mathbb{R}^N} \exp\{i\langle x, h\rangle_{\mathbb{R}^N}\}\, d\mu(x) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\int_{\mathbb{R}} \exp\left\{ i \sum_{j=1}^{N} x_j\, h_j \right\}\, dN_{m_1,\lambda_1}(x_1) \right) \cdots \right)\, dN_{m_N,\lambda_N}(x_N) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\int_{\mathbb{R}} \prod_{j=1}^{N} \exp\left\{ i\, x_j\, h_j \right\}\, dN_{m_1,\lambda_1}(x_1) \right) \cdots \right)\, dN_{m_N,\lambda_N}(x_N) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\prod_{j=2}^{N} \exp\left\{ i\, x_j\, h_j \right\} \int_{\mathbb{R}} \exp\left\{ i\, x_1\, h_1 \right\}\, dN_{m_1,\lambda_1}(x_1) \right) \cdots \right)\, dN_{m_N,\lambda_N}(x_N) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\prod_{j=2}^{N} \exp\left\{ i\, x_j\, h_j \right\} \exp\{i\, m_1\, h_1\} \exp\{-\tfrac{1}{2}\lambda_1\, h_1^2\} \right) \cdots \right)\, dN_{m_N,\lambda_N}(x_N) \\
&= \prod_{j=1}^{N-1} \exp\{i\, m_j\, h_j\} \exp\{-\tfrac{1}{2}\lambda_j\, h_j^2\} \int_{\mathbb{R}} \exp\left\{ i\, x_N\, h_N \right\}\, dN_{m_N,\lambda_N}(x_N) \\
&= \prod_{j=1}^{N} \exp\{i\, m_j\, h_j\} \exp\{-\tfrac{1}{2}\lambda_j\, h_j^2\} \\
&= \exp\left\{ \sum_{j=1}^{N} i\, m_j\, h_j \right\} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{N} \lambda_j\, h_j^2 \right\} \\
&= \exp\left\{ i\langle m, h\rangle_{\mathbb{R}^N} - \tfrac{1}{2}\langle Q(h), h\rangle_{\mathbb{R}^N} \right\}
\end{aligned}$$

lo que implica que $\mu = N_{m,Q}$.

Ahora, sea $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$. Usando de manera recursiva la identidad (4.19) obtenemos que:

$$\begin{aligned}
\int_{\mathbb{R}^N} \langle x, y\rangle_{\mathbb{R}^N}\, d\mu(y) &= \int_{\mathbb{R}} \left(\cdots \left(\int_{\mathbb{R}} \sum_{j=1}^{N} x_j\, y_j\, dN_{m_1,\lambda_1}(y_1) \right) \cdots \right)\, dN_{m_N,\lambda_N}(y_N) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\sum_{j=1}^{N} \int_{\mathbb{R}} x_j\, y_j\, dN_{m_1,\lambda_1}(y_1) \right) \cdots \right)\, dN_{m_N,\lambda_N}(y_N) \\
&= \int_{\mathbb{R}} \left(\cdots \left(\int_{\mathbb{R}} \left(m_1 x_1 + \sum_{j=2}^{N} x_j y_j \right)\, dN_{m_2,\lambda_2}(y_2) \right) \cdots \right)\, dN_{m_N,\lambda_N}(y_N) \\
&= \int_{\mathbb{R}} \left(\sum_{j=1}^{N-1} m_j x_j + x_N y_N \right)\, dN_{m_N,\lambda_N}(y_N) = \sum_{j=1}^{N} m_j\, x_j = \langle m, x\rangle_{\mathbb{R}^N}.
\end{aligned}$$

En consecuencia, $m \in \mathbb{R}^N$ es la media de μ.

Finalmente, sean $x = (x_1, \ldots, x_N), y = (y_1, \ldots, y_N) \in \mathbb{R}^N$. Entonces, para cada $z = (z_1, \ldots, z_N) \in \mathbb{R}^N$ se tiene que:

$$
\begin{aligned}
&\langle x, z-m\rangle_{\mathbb{R}^N} \langle y, z-m\rangle_{\mathbb{R}^N} = \left(\sum_{j=1}^{N} (x_j z_j - x_j m_j)\right) \left(\sum_{j=1}^{N} (y_j z_j - y_j m_j)\right) \\
&= (x_1 z_1 - x_1 m_1) \left(\sum_{j=1}^{N} (y_j z_j - y_j m_j)\right) + \ldots + (x_N z_N - x_N m_N) \left(\sum_{j=1}^{N} (y_j z_j - y_j m_j)\right) \\
&= (x_1 z_1 - x_1 m_1)(y_1 z_1 - y_1 m_1) + (x_1 z_1 - x_1 m_1) \left(\sum_{j=2}^{N} (y_j z_j - y_j m_j)\right) + \\
&+ (x_2 z_2 - x_2 m_2)(y_2 z_2 - y_2 m_2) + (x_2 z_2 - x_2 m_2) \left(\sum_{\substack{j=1 \\ j \neq 2}}^{N} (y_j z_j - y_j m_j)\right) + \ldots \\
&+ (x_N z_N - x_N m_N)(y_N z_N - y_N m_N) + (x_N z_N - x_N m_N) \left(\sum_{j=1}^{N-1} (y_j z_j - y_j m_j)\right).
\end{aligned}
$$

Fijemos $i \in \{1, \ldots, N\}$. Observemos que, aplicando las identidades (4.18) y (4.20) se tiene que:

$$
\begin{aligned}
&\int_{\mathbb{R}} \left[(x_i z_i - x_i m_i)(y_i z_i - y_i m_i) + (x_i z_i - x_i m_i) \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} (y_j z_j - y_j m_j)\right)\right] dN_{m_i, \lambda_i}(z_i) \\
&= \int_{\mathbb{R}} (x_i z_i - x_i m_i)(y_i z_i - y_i m_i)\, dN_{m_i, \lambda_i}(z_i) \;\; + \\
&\qquad + \sum_{\substack{j=1 \\ j \neq i}}^{N} (y_j z_j - y_j m_j) \int_{\mathbb{R}} (x_i z_i - x_i m_i)\, dN_{m_i, \lambda_i}(z_i) \\
&= \lambda_i\, x_i\, y_i \;+\; \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} (y_j z_j - y_j m_j)\right) (x_i m_i - x_i m_i) = \lambda_i\, x_i\, y_i.
\end{aligned}
$$

Por lo tanto, aplicando el procedimiento de manera recursiva, para cada $i \in \{1, \ldots, N\}$, obtenemos:

$$\begin{aligned}
&\int_{\mathbb{R}^N} \langle x, z-m\rangle_{\mathbb{R}^N} \langle y, z-m\rangle_{\mathbb{R}^N}\, d\mu(z)\\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left(\sum_{j=1}^{N} (x_j z_j - x_j m_j)\right) \left(\sum_{j=1}^{N} (y_j z_j - y_j m_j)\right) dN_{m_1,\lambda_1}(y_1) \cdots\ dN_{m_N,\lambda_N}(y_N)\\
&= \sum_{j=1}^{N} \lambda_j\, x_j\, y_j = \langle Q(x), y\rangle_{\mathbb{R}^N}.
\end{aligned}$$

Esto concluye la demostración. ■

4.2.3. Medidas Gaussianas en espacios de Hilbert

En esta sección supondremos que el lector está familiarizado con los conceptos básicos de la medida producto para una cantidad numerable de espacios de probabilidad. Invitamos a consultar, por ejemplo, [37] en donde se reúnen los resultados elementales de este tema.

Sea $H = (H, \langle\cdot\rangle, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable de dimensión infinita.

Sean $m \in H$ y $Q \in \mathcal{L}_1^+(H)$. El Teorema 3.55 asegura que existe $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ una base de Hilbert de H y una sucesión de números reales no negativos (λ_k) tales que $Q(e_k) = \lambda_k\, e_k$ para todo $k \in \mathbb{N}$.

El Teorema 1.49 asegura que la función $\Theta : H \to \ell_2(\mathbb{R})$ dada por:

$$\Theta(u) := (\langle u, e_k\rangle) := u_k \tag{4.21}$$

es un isomorfismo isométrico. En esta sección identificaremos a H con el espacio $\ell_2(\mathbb{R})$ y escribimos $m = (m_k)$ y $Q(u) = (\lambda_k\, u_k)$ para todo $u \in H$.

Sea $\mathbb{R}_k := (\mathbb{R}, \mathcal{B}(\mathbb{R}), N_{m_k,\lambda_k})$ el espacio de medida dado en la Proposición 4.10 para todo $k \in \mathbb{N}$. Definamos entonces $\mathbb{R}^{\mathbb{N}}$ como el producto sobre $\mathbb{N}$ del conjunto $\mathbb{R}$ y consideremos en él la topología producto. Es decir:

$$\mathbb{R}^{\mathbb{N}} = \{f : \mathbb{N} \to \cup_{k=1}^{\infty} \mathbb{R}\ :\ f(k) \in \mathbb{R}\quad \forall\, k \in \mathbb{N}\},$$

por lo que afirmamos que $\mathbb{R}^{\mathbb{N}}$ es el conjunto de todas las sucesiones de números reales.

Como $\mathbb{R}_k$ es un espacio de probabilidad entonces es posible considerar la medida producto sobre $\mathbb{N}$ de estos espacios siguiendo la idea de la sección anterior. Por tanto, sea $\mu : \mathcal{B}(\mathbb{R}^{\mathbb{N}}) \to \mathbb{R}$ dada por:

$$\mu := \bigotimes_{k=1}^{\infty} N_{m_k,\lambda_k}. \tag{4.22}$$

Para cada $j \in \mathbb{N}$, se define la proyección en la j-ésima coordenada como la función $\pi_j : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ dada por $\pi_j((x_k)) := x_j$. Es inmediato que π_j es una función suprayectiva. Más aún, la σ-álgebra producto $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) = \bigotimes_{k=1}^{\infty} \mathcal{B}(\mathbb{R})$ hace a todas las proyecciones funciones medibles.

Observemos que μ está definida en $\mathbb{R}^{\mathbb{N}}$ lo que nos sugiere restringirla al espacio $\ell_2(\mathbb{R})$ para lo cual es necesario probar que $\ell_2(\mathbb{R})$ es un conjunto de Borel de $\mathbb{R}^{\mathbb{N}}$. En efecto, sea $\rho : \mathbb{R}^{\mathbb{N}} \to \mathbb{R} \cup \{\infty\}$ dada por $\rho((x_k)) := \sum_{j=1}^{\infty} |\pi_j((x_k))|^2$. Es claro que ρ es medible en $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ pues $|\pi_j((x_k))|^2$ es medible para cada $j \in \mathbb{R}$. Luego, es inmediato que $\ell_2(\mathbb{R}) = \rho^{-1}([0, \infty))$ por lo que concluimos que $\ell_2(\mathbb{R}) \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$.

Finalmente, el espacio $\ell_2(\mathbb{R})$ tiene medida igual a 1. En efecto, por el teorema de la convergencia monótona se tiene que:

$$\begin{aligned}
\int_{\mathbb{R}^{\mathbb{N}}} \|(x_k)\|_{\ell_2(\mathbb{R})}^2 \, d\mu((x_k)) &= \int_{\mathbb{R}^{\mathbb{N}}} \sum_{k=1}^{\infty} |x_k|^2 \, d\mu((x_k)) = \sum_{k=1}^{\infty} \int_{\mathbb{R}} |x_k|^2 \, dN_{m_k, \lambda_k}(x_k) \\
&= \sum_{k=1}^{\infty} (\lambda_k + m_k^2) = \operatorname{Tr} Q + \|m\|_H^2.
\end{aligned}$$

En consecuencia,

$$\mu\left(\left\{(x_k) \in \mathbb{R}^{\mathbb{N}} \, : \, \sum_{k=1}^{\infty} |x_k|^2 < \infty\right\}\right) = 1.$$

Podemos, por tanto, completar el objetivo de esta sección.

Teorema 4.13 *La restricción de la medida μ definida en* (4.22) *en el espacio medible* $(\ell_2(\mathbb{R}), \mathcal{B}(\ell_2(\mathbb{R})))$ *es una medida Gaussiana $N_{m,Q}$.*

Demostración: Definamos $p_k : H \to H$ como $p_k(u) := \sum_{j=1}^{k} \langle u, e_j \rangle \, e_j$ para cada $k \in \mathbb{N}$. El Teorema 1.48 asegura que $\lim_{k\to\infty} p_k(u) = u$ para todo $u \in H$.

Así pues, para cada $h \in H$, por el teorema de la convergencia dominada de Lebesgue y la definición de la medida producto μ se tiene que:

$$\begin{aligned}
\int_H \exp\{i\langle u, h\rangle\} \, d\mu(u) &= \lim_{k\to\infty} \int_H \exp\{i\langle p_k(u), h\rangle\} \, d\mu(u) \\
&= \lim_{k\to\infty} \prod_{j=1}^{k} \int_{\mathbb{R}} \exp\{iu_j \, h_j\} \, dN_{m_j, \lambda_j}(u_j) \\
&= \lim_{k\to\infty} \prod_{j=1}^{k} \exp\{im_j \, h_j\} \exp\left\{-\tfrac{1}{2}\lambda_j h_j^2\right\} \\
&= \lim_{k\to\infty} \exp\{i\langle p_k(m), h\rangle\} \exp\{-\tfrac{1}{2}\langle p_k(Q(h)), h\rangle\} \\
&= \exp\{i\langle m, h\rangle\} \exp\{-\tfrac{1}{2}\langle Q(h), h\rangle\}
\end{aligned}$$

lo que prueba que $\mu = N_{m,Q}$.

Observa que

$$\int_H \|u\|_H \, d\mu(u) = \operatorname{Tr} Q + \|m\|_H^2 < \infty$$

lo que implica que μ tiene segundo momento finito.

Aplicando de nuevo el teorema de la convergencia monótona obtenemos:

$$\int_H \langle u, v\rangle \, d\mu(v) = \sum_{k=1}^{\infty} \int_{\mathbb{R}} u_k v_k \, dN_{m_k,\lambda_k}(v_k) = \sum_{k=1}^{\infty} u_k \, m_k = \langle u, m\rangle \quad \forall\, u \in H$$

lo que implica que m es la media de μ. Del mismo modo,

$$\begin{aligned}\int_H \langle u, w-m\rangle\langle v, w-m\rangle \, d\mu(w) &= \sum_{k=1}^{\infty} \int_{\mathbb{R}} (u_k w_k - u_k m_k)(v_k w_k - v_k m_k) \, dN_{m_k,\lambda_k}(w_k) \\ &= \sum_{k=1}^{\infty} \lambda_k \, u_k \, v_k \\ &= \langle Q(u), v\rangle \qquad \forall\, u, v \in H\end{aligned}$$

lo que establece que Q es la covarianza de μ. ∎

Capítulo 5

Variables aleatorias en espacios de Hilbert

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. En los cursos básicos de teoría de la probabilidad se estudian a las funciones medibles $X : \Omega \to \mathbb{R}$ que usualmente son llamadas variables aleatorias. De manera intuitiva, una variable aleatoria real puede entenderse como un valor númerico que está generado por la ocurrencia de un fenómeno aleatorio y del cual interesa conocer la probabilidad de la misma. Así pues, las variables aleatorias reales inducen una medida de probabilidad en el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ conocida como distribución de la variable aleatoria y la cual resuelve, en muchas ocasiones, el problema planteado.

Como es bien sabido, el Análisis Matemático profundiza en algunos conceptos y resultados establecidos para el conjunto de los números reales $\mathbb{R}$ extendiéndolos muy especialmente a espacios de funciones y, en teoría de la probabilidad no hay excepción. Veamos el siguiente problema:

Sea $\Omega = (0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1])$ y $P = \lambda$ la medida de Lebesgue en $(0, 1)$. Definamos $X : \Omega \to L^2(0, 1)$ como sigue $X(\omega) = 1_{(0,\omega)}$. *¿Cuál es la probabilidad de que la norma-L^2 de X sea menor que 1?*

Precisemos esta pregunta. En el Capítulo 2 extendimos el concepto de mensurabilidad para funciones que toman valores en un espacio de Hilbert real. De este modo, la noción de variable aleatoria puede generalizarse para funciones que toman valores en cualquier espacio de Hilbert real H. Es decir, $\xi : \Omega \to H$ es una *variable aleatoria H-valuada* si ξ es una función fuertemente P-medible. Dado que ξ es fuertemente P-medible, el Teorema 2.21 nos permite definir la *distribución de* ξ como la medida de probabilidad sobre $(H, \mathcal{B}(H))$ dada por $P(\xi^{-1}(B))$ para cada $B \in \mathcal{B}(H)$.

En este trabajo, restringiremos el concepto de variable aleatoria al caso de los espacios de Hilbert y, en particular, a *espacios de Hilbert separables*. Considerar la separabilidad en los espacios de Hilbert nos permitirá, a través del teorema de mensurabilidad de Pettis, trabajar con equivalencias en el concepto de variable aleatoria pues, por ejemplo, toda función fuertemente P-medible en un espacio de Banach separable es $\mathcal{F}$-medible y viceversa. Considerar la separabilidad en un espacio de Hilbert H nos permitirá además extender la noción de *variables aleatorias Gaussianas* cuyo concepto surge de una implicación directa del capítulo anterior.

En este capítulo introducimos diversos tipos de convergencia de una sucesión de variables aleatorias H-valuadas y se estudia su relación entre ellas. Extendemos conceptos clásicos de teoría de la probabilidad como lo es la *esperanza condicional* y la *independencia* para variables aleatorias H-valuadas. Demostramos el *teorema de cambio de variable* en una versión más general para espacios de medida finita y del cual se deducirán distintos resultados importantes.

La construcción de este capítulo se basa principalmente en [9], [20], [21], [23], [38], [61], [62], [64], [83], [86] y [87].

5.1. Variables Aleatorias

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable. Consideraremos en el espacio de Hilbert H a $\mathcal{B}(H)$ la σ-álgebra de Borel generada por la topología inducida por la norma $\|\cdot\|_H$. A los elementos de $\mathcal{B}(H)$ los llamaremos subconjuntos Borel-medibles.

Definición 5.1 *Sea $X : \Omega \to H$ una función. Decimos que X es una* ***variable aleatoria H-valuada*** *en $(\Omega, \mathcal{F}, P)$ si, X es una función fuertemente P-medible.*

Sea $X : \Omega \to H$ una función. El Teorema 2.19 asegura que $X : \Omega \to H$ es una variable aleatoria H-valuada si y sólo si X es $\mathcal{F}$-medible, es decir, $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(H)$, pues H es separable.

Definición 5.2 *Sea $X : \Omega \to H$ una variable aleatoria H-valuada. Si X es Bochner-integrable, definimos la* ***esperanza de X*** *como:*

$$E(X) := \mathfrak{B}\int_\Omega X\,dP.$$

Dado un subconjunto B de H Borel-medible y $X : \Omega \to H$ una variable aleatoria H-valuada, denotaremos por

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}$$

a la **imagen inversa de B bajo X**. La observación anterior asegura que $\{X \in B\} \in \mathcal{F}$.

Definición 5.3 *Sea $X : \Omega \to H$ una variable aleatoria H-valuada. La* ***distribución de la variable aleatoria*** *X es la medida de probabilidad $X_\sharp P$ en $(H, \mathcal{B}(H))$ definida por:*

$$X_\sharp P(B) := P(\{X \in B\}) \qquad \forall\, B \in \mathcal{B}(H). \tag{5.1}$$

Decimos que X, Y variables aleatorias H-valuadas son ***idénticamente distribuidas*** *si tienen la misma distribución.*

Definición 5.4 *Sea $\mu : \mathcal{B}(H) \to \mathbb{R}$ una medida finita. Decimos que μ es una* ***medida de Radon*** *si, para cada $B \in \mathcal{B}(H)$ y $\varepsilon > 0$ existe un subconjunto compacto K de B tal que $\mu(B \setminus K) < \varepsilon$.*

Decimos que μ es ***tight*** *si, dada $\varepsilon > 0$ existe un subconjunto compacto K de H tal que $\mu(H \setminus K) < \varepsilon$.*

En la definición anterior, debe entenderse a cada subconjunto $B \in \mathcal{B}(H)$, como subespacio métrico de H contiene a un subconjunto compacto K con tal propiedad para μ. A continuación probamos que toda medida de probabilidad en H es una medida *tight*.

Proposición 5.5 *Sea $\mu : \mathcal{B}(H) \to \mathbb{R}$ una medida de probabilidad. Entonces, para cada $\varepsilon > 0$ existe un subconjunto compacto K de H tal que $\mu(K) \geq 1 - \varepsilon$.*

Demostración: Sea $\{e_k : k \in \mathbb{N}\}$ subconjunto denso de H y sea $\varepsilon > 0$. Fijemos $j \in \mathbb{N}$. Para cada $u \in H$ existe $k_0 \in \mathbb{N}$ tal que $\|u - e_{k_0}\|_H < \frac{1}{j}$.

En consecuencia, $H = \cup_{j=1}^{\infty} B_H(e_j, \frac{1}{j}) = \cup_{j=1}^{\infty} \bar{B}_H(e_j, \frac{1}{j})$.

Definiendo la sucesión $F_k := \cup_{j=1}^{k} \bar{B}_H(e_j, \frac{1}{j})$ de subconjuntos de $\mathcal{B}(H)$ es claro que (F_k) es no decreciente y $\cup_{k=1}^{\infty} F_k = \cup_{j=1}^{\infty} \bar{B}_H(e_j, \frac{1}{j}) = H$. En consecuencia, $\lim_{k\to\infty} \mu(F_k) = \mu(H) = 1$. De este modo, para cada $j \in \mathbb{N}$ existe $k_j \in \mathbb{N}$ tal que:

$$1 - \mu(F_{k_j}) \leq \frac{\varepsilon}{2^j}.$$

O equivalentemente:

$$\mu(F_{k_j}) \geq 1 - \frac{\varepsilon}{2^j}.$$

Definiendo $K := \cap_{j=1}^{\infty} F_{k_j}$ es claro que K es cerrado, totalmente acotado y $K \in \mathcal{B}(H)$. En consecuencia, K es compacto pues H es completo. De las propiedades de medida de probabilidad concluimos que:

$$\mu(K) = 1 - \mu(H \setminus K) = 1 - \mu\left(\cup_{j=1}^{\infty}(H \setminus F_{k_j})\right) \geq 1 - \sum_{j=1}^{\infty} \mu(H \setminus F_{k_j}) \geq 1 - \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} \geq 1 - \varepsilon$$

y esto demuestra el enunciado. ■

Una consecuencia inmediata de la proposición anterior es la siguiente.

Corolario 5.6 *Si $X : \Omega \to H$ una variable aleatoria H-valuada, entonces para cada $\varepsilon > 0$ existe un subconjunto compacto K de H tal que $P(\{X \notin K\}) < \varepsilon$.*

Demostración: Es inmediato de aplicar la Proposición 5.5 a la distribución $X_\sharp P : \mathcal{B}(H) \to \mathbb{R}$ de X que es una medida de probabilidad. ∎

Definición 5.7 *Una familia $\mathcal{H}$ de variables aleatorias H-valuadas en $(\Omega, \mathcal{F}, P)$ es **uniformemente tight** si, dada $\varepsilon > 0$ existe un subconjunto compacto K de H (que depende de ε) tal que, para toda $X \in \mathcal{H}$,*

$$P(\{X \notin K\}) < \varepsilon.$$

O equivalentemente:

$$P(\{X \in K\}) \geq 1 - \varepsilon.$$

Proposición 5.8 *Si $\mathcal{H}$ es uniformemente tight entonces $\mathcal{H} - \mathcal{H} := \{X_1 - X_2 \,:\, X_1, X_2 \in \mathcal{H}\}$ es uniformemente tight.*

Demostración: Sea $\varepsilon > 0$ dada. Elijamos K subconjunto compacto de H tal que $P(\{X \notin K\}) < \frac{\varepsilon}{2}$ para toda $X \in \mathcal{H}$.

Definamos $f : X \times X \to X$ como $f(u, v) := u - v$. Esta función, es claramente lineal. Observa que, $\|u - v\|_H \leq \|u\|_H + \|v\|_H$ para cualesquiera $u, v \in H$. Por tanto,

$$\|f(u, v)\|_H \leq \|(u, v)\|_{H \times H, 1} \qquad \forall\, (u, v) \in H \times H$$

lo que implica que f es continua en $H \times H$. Como $K \times K$ es un subconjunto compacto en $H \times H$ y f es continua, entonces $f(K \times K)$ es compacto en H.

Es inmediato que $f(K \times K) = \{u - v \,:\, u, v \in K\}$. De este modo, si $w \in \Omega$ es tal que $X_1(\omega), X_2(\omega) \in K$, entonces $X_1(\omega) - X_2(\omega) \in f(K \times K)$ y, por lo tanto:

$$P(\{X_1 \in K\} \cap \{X_2 \in K\}) \leq P(\{X_1 - X_2 \in f(K \times K)\})$$

O equivalentemente:

$$P(\{X_1 - X_2 \notin f(K \times K\}) \leq P(\{X_1 \notin K\}) + P(\{X_2 \notin K\}) < \varepsilon.$$

Esto prueba que $\mathcal{H} - \mathcal{H}$ es uniformemente *tight*. ∎

El siguiente resultado es de gran importancia pues nos permite relacionar la integral de Lebesgue de una variable aleatoria real con respecto de la distribución de una variable aleatoria H-valuada.

Teorema 5.9 (de cambio de variable) *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $X : \Omega \to H$ una variable aleatoria H-valuada. Entonces, para cualquier función $\varphi : H \to \mathbb{R}$ acotada y Borel-medible se tiene que:*

$$\int_\Omega (\varphi \circ X)(\omega)\, dP(\omega) = \int_H \varphi(x)\, dX_\sharp P(x).$$

Demostración: Consideremos los siguientes casos:

CASO 1. Si $\varphi = 1_A$ para algún $A \in \mathcal{B}(H)$, entonces $(\varphi \circ X)(\omega) = 1$ si $X(\omega) \in A$ y $(\varphi \circ X)(\omega) = 0$ si $X(\omega) \notin A$. Es decir, $(\varphi \circ X) = 1_{X^{-1}(A)}$.

En consecuencia,

$$\int_H \varphi(x)\, dX_\sharp P(x) = X_\sharp P(A) = P(X^{-1}(A)) = \int_\Omega (\varphi \circ X)(\omega)\, dP(\omega).$$

CASO 2. φ es una función acotada simple no negativa con descripción canónica igual a $\sum_{i=1}^N \alpha_i\, 1_{A_i}$. Aplicando el caso anterior concluimos que $\int_\Omega (1_{A_i} \circ X)(\omega)\, dP(\omega) = \int_Y 1_{A_i}(x)\, dX_\sharp P(x)$ para cada $i = 1, \ldots, N$.

Por tanto, de la linealidad de la integral obtenemos:

$$\begin{aligned}\int_H \varphi(x)\, dX_\sharp P(x) &= \sum_{i=1}^N \alpha_i \int_H 1_{A_i}(x)\, dX_\sharp P(x) = \sum_{i=1}^N \alpha_i \int_\Omega (1_{A_i} \circ X)(\omega)\, dP(\omega) \\ &= \int_\Omega \sum_{i=1}^N \alpha_i (1_{A_i} \circ X)(\omega)\, dP(\omega) = \int_\Omega (\varphi \circ X)(\omega)\, dP(\omega)\end{aligned}$$

que es la identidad deseada.

CASO 3. φ es una función Borel-medible no negativa y acotada. El lema de aproximación de funciones medibles afirma que existe (s_k) sucesión no decreciente y no negativa de funciones simples tal que $s_k \to \varphi$ uniformemente en Y. Por el caso anterior y el teorema de la convergencia monótona se tiene:

$$\int_H \varphi(x)\, dX_\sharp P(x) = \lim_{k\to\infty} \int_H s_k(x)\, dX_\sharp P(x) = \lim_{k\to\infty} \int_\Omega (s_k \circ X)(\omega)\, dP(\omega)$$

donde claramente $(s_k \circ X)$ es una sucesión no decreciente de funciones simples no negativas tal que $s_k \circ X \to \varphi \circ X$ en Ω. Así, nuevamente del teorema de la convergencia monótona concluimos que:

$$\lim_{k\to\infty} \int_\Omega (s_k \circ X)(\omega)\, dP(\omega) = \int_\Omega (\varphi \circ X)(\omega)\, dP(\omega)$$

lo que demuestra la igualdad deseada.

CASO 4. Sea φ cualquier función acotada y Borel-medible. Consideremos φ^+ y φ^- : $H \to \mathbb{R}$ tales que $\varphi = \varphi^+ - \varphi^-$. Es claro que φ^+ y φ^- son funciones Borel-medibles acotadas y no negativas. Así pues, del caso anterior y la linealidad de la integral se tiene que:

$$\begin{aligned}\int_H \varphi(x)\,dX_\sharp P(x) &= \int_H \varphi^+(x)\,dX_\sharp P(x) - \int_H \varphi^-(x)\,dX_\sharp P(x)\\ &= \int_\Omega (\varphi^+ \circ X)(\omega)\,dP(\omega) - \int_\Omega (\varphi^- \circ X)(\omega)\,dP(\omega)\\ &= \int_\Omega ((\varphi^+ - \varphi^-) \circ X)(\omega)\,dP(\omega)\\ &= \int_\Omega (\varphi \circ X)(\omega)\,dP(\omega)\end{aligned}$$

como afirma el enunciado. ■

Una versión del teorema de cambio de variable para cualquier espacio de medida $(\Omega, \mathcal{S}, \mu)$ considera fundamental la hipótesis de que la función medible $\phi : H \to \mathbb{R}$ sea Lebesgue-integrable con respecto de la medida $\mu(X^{-1}(B))$ para $X : \Omega \to H$ una función $\mathcal{S}$-medible fija. En este trabajo hablaremos del teorema de cambio de variable sin hacer distinción de estas versiones, es decir, lo consideraremos válido en los casos cuando la función medible $\phi : H \to \mathbb{R}$ sea acoatda y sea Lebesgue-integrable con respecto de la de distribución de la variable aleatoria $X : \Omega \to H$. La prueba de la seguna versión es completamente análoga a la dada aquí e invitamos al lector a consultar, por ejemplo, [18] y [78].

Definición 5.10 *Sea $X : \Omega \to H$ una variable aleatoria H-valuada. Definimos la* ***transformada de Fourier de*** *X como la transformada de Fourier de su distribución $\widehat{X_\sharp P} : H \to \mathbb{C}$. Es decir:*

$$\widehat{X}(h) := \widehat{X_\sharp P}(h) = \int_H \exp\{i\langle h, u\rangle_H\}\; dX_\sharp P(u).$$

Observación 5.11 *Fijemos $h \in H$ y consideremos la función $\langle h, \cdot\rangle_H : H \to \mathbb{R}$. Es inmediato de la* Proposición 1.6 *que $\langle h, \cdot\rangle_H$ es medible.*

Por otra parte, las funciones $\cos, \sin : \mathbb{R} \to \mathbb{R}$ *son funciones acotadas y medibles. De este modo, las funciones* $\cos(\langle h, \cdot\rangle_H) : H \to \mathbb{R}$ *y* $\sin(\langle h, \cdot\rangle_H) : H \to \mathbb{R}$ *son medibles y acotadas.*

De la fórmula de Euler (ver [56]*) se sigue que:*

$$\exp\{i\langle h, u\rangle_H\} = \cos(\langle h, u\rangle_H) + i\sin(\langle h, u\rangle_H) \qquad \forall\, u \in H.$$

En consecuencia,

$$\int_H \exp\{i\langle h, u\rangle_H\}\; dX_\sharp P(u) = \int_H \cos(\langle h, u\rangle_H)\,dX_\sharp P(u) + i\int_H \sin(\langle h, u\rangle_H)\,dX_\sharp P(u)$$

por lo que la Definición 5.9 *es consistente.*

Ahora bien, del teorema de cambio de variable se sigue que:

$$\int_H \cos(\langle h,u\rangle_H)\,dX_\sharp P(u) = \int_\Omega \cos(\langle X(\omega),u\rangle_H)\,dP(\omega),$$
$$\int_H \sin(\langle h,u\rangle_H)\,dX_\sharp P(u) = \int_\Omega \sin(\langle X(\omega),u\rangle_H)\,dP(\omega).$$

En consecuencia,

$$\widehat{X}(h) = \int_H \exp\{i\langle h,u\rangle_H\}\,dX_\sharp P(u) = \int_\Omega \exp\{i\langle h,X(\omega)\rangle_H\}\,dP(\omega) = E(i\langle h,X\rangle_H) \quad \forall\, h\in H.$$

Observemos que si X_1, X_2 son variables aleatorias H-valuadas tales que $\widehat{X}_1(h) = \widehat{X}_2(h)$ para todo $h\in H$ entonces X_1 y X_2 son idénticamente distribuidas por el Teorema 4.8.

El siguiente es un resultado útil y clásico de cálculo.

Teorema 5.12 (Fórmula de integración por partes) *Sea $\vartheta:\Omega\to\mathbb{R}$ una variable aleatoria no negativa y sea $p\in[1,\infty)$. Entonces,*

$$E(\vartheta^p) = \int_\Omega \vartheta^p\,dP = \int_0^\infty p\,\xi^{p-1}\,P(\vartheta>\xi)\,d\xi. \tag{5.2}$$

Demostración: Definamos $\gamma:\Omega\times[0,\infty)\to\mathbb{R}$ como:

$$\gamma(\omega,\xi) := \begin{cases} 1 & \text{si } \vartheta(\omega)>\xi, \\ 0 & \text{si } \vartheta(\omega)\le\xi \end{cases}$$

en donde consideramos $\mathcal{B}([0,\infty))$ la σ-álgebra de Borel y λ la medida de Lebesgue en $[0,\infty)$. De este modo, es inmediato que la función γ es medible en $\mathcal{F}\otimes\mathcal{B}([0,\infty))$. La función $\eta:[0,\infty)\to[0,\infty)$ dada por $\eta(\xi):=p\xi^{p-1}$ es continua y, por tanto, medible.

Aplicando el teorema de Tonelli obtenemos que:

$$\begin{aligned}
\int_0^\infty p\,\xi^{p-1}\,P(\vartheta>\xi)\,d\xi &= \int_0^\infty p\,\xi^{p-1}\left(\int_\Omega 1_{\{\vartheta(\omega)>\xi\}}(\omega)\,dP(\omega)\right)d\xi \\
&= \int_0^\infty \left(\int_\Omega p\,\xi^{p-1}\,\gamma(\omega,\xi)\,dP(\omega)\right)d\xi \\
&= \int_\Omega \left(\int_0^{\vartheta(\omega)} p\,\xi^{p-1}\,d\xi\right)dP(\omega) \\
&= \int_\Omega (\vartheta(\omega))^p\,dP(\omega) \\
&= E(\vartheta^p),
\end{aligned}$$

lo cual prueba el teorema. ■

5.1.1. Vectores aleatorios en espacios de Hilbert

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

El Ejemplo 1.12 asegura que $H^2 := H \times H$ es un espacio de Hilbert con el producto escalar definido:

$$\langle (u_1, u_2), (v_1, v_2)\rangle_{H^2} := \langle u_1, v_1\rangle_H + \langle u_2, v_2\rangle_H$$

y que induce la norma producto $\|(u,v)\|^2_{H^2} = \|u\|^2_H + \|v\|^2_H$.

Como H es separable, sea D subconjunto denso y numerable de H. Defiendo $D^2 := D \times D$ es inmediato que D^2 es un subconjunto numerable de H^2. Veamos que es separable en H^2. Sea $(u,v) \in H^2$ y $\varepsilon > 0$ dada. Existen $\xi, \eta \in D$ tales que:

$$\|u - \xi\|^2_H < \frac{\varepsilon^2}{2} \qquad \text{y} \qquad \|v - \eta\|^2_H < \frac{\varepsilon^2}{2}.$$

De este modo,

$$\|(u,v) - (\xi, \eta)\|^2_{H^2} = \|u - \xi\|^2_H + \|v - \eta\|^2_H < \varepsilon^2$$

lo que implica que $\overline{D^2} = H^2$.

En consecuencia, $H^2 = (H^2, \langle\cdot,\cdot\rangle_{H^2}, \|\cdot\|_{H^2})$ es un espacio de Hilbert separable sobre $\mathbb{R}$.

Consideremos ahora la σ-álgebra de Borel en H^2 inducida por la norma producto $\|(u,v)\|^2_{H^2} = \|u\|^2_H + \|v\|^2_H$. La denotaremos por $\mathcal{B}(H^2)$.

Por otra parte, sea $\mathcal{B}(H) \otimes \mathcal{B}(H) = \mathcal{S}(\mathcal{B}(H) \times \mathcal{B}(H))$ la σ-álgebra producto en H^2.

Definimos $\pi_i : H^2 \to H$ dada por $\pi_i(u_1, u_2) = u_i$ la proyección en la i-ésima coordenada $(i = 1, 2)$. Dada $\varepsilon > 0$ y $(u_1^0, u_2^0) \in H^2$ se tiene que:

$$\|\pi_i(u_1^0, v_2^0) - \pi_i(u_1, u_2)\|^2_H = \|u_i^0 - u_i\|^2_H \leq \|(u_1^0, u_2^0) - (u_1, u_2)\|^2_{H^2} < \varepsilon^2$$

si $\|(u_0, v_0) - (u_1, u_2)\|_{H^2} < \varepsilon$. Es decir, π_i es continua en H^2 y, por tanto, medible en $(H^2, \mathcal{B}(H^2))$ $(i = 1, 2)$.

Sea $B := B_1 \times B_2 \subset H^2$ con $B_1, B_2 \in \mathcal{B}(H)$. Observa que $\pi_i^{-1}(B_i) \in \mathcal{B}(H^2)$ para $i = 1, 2$. De este modo $\pi_1^{-1}(B_1) \cap \pi_2^{-1}(B_2) = B_1 \times B_2 = B \in \mathcal{B}(H^2)$ por lo que $\mathcal{B}(H) \times \mathcal{B}(H) \subset \mathcal{B}(H^2)$ y, en consecuencia $\mathcal{S}(\mathcal{B}(H) \times \mathcal{B}(H)) = \mathcal{B}(H) \otimes \mathcal{B}(H) \subset \mathcal{B}(H^2)$.

Inversamente: Dado que la norma $\|(u,v)\|_{H^2,\infty} = \text{máx}\{\|u\|_H, \|v\|_H\}$ genera a la topología producto en H^2 y es equivalente a la norma $\|\cdot\|_{H^2}$ entonces $\mathcal{B}(H^2) \subset \mathcal{S}(\mathcal{B}(H) \times \mathcal{B}(H)) = \mathcal{B}(H) \otimes \mathcal{B}(H)$.

Argumentando por inducción, es posible generalizar lo anterior para una cantidad finita N. Es decir, H^N con el producto escalar

$$\langle (u_1, \ldots, u_N), (v_1, \ldots, v_N)\rangle_{H^N} := \sum_{j=1}^{N} \langle u_j, v_j\rangle_H$$

que induce la norma $\|(u_1,\ldots,u_N)\|_{H^N}^2 = \sum_{j=1}^N \|u_j\|_H^2$ es un espacio de Hilbert separable tal que $\mathcal{B}(H^N)$ coincide con la σ-álgebra producto $\otimes_{j=1}^N \mathcal{B}(H)$.

En esta sección nos interesa estudiar funciones del tipo $X := (X_1,\ldots,X_N) : \Omega \to H^N$.

Proposición 5.13 *Sean $(\Omega,\mathcal{F},P)$ un espacio de probabilidad y $X_i : \Omega \to H$ una variable aleatoria H-valuada para cada $i = 1,\ldots,N$. Definamos $X : \Omega \to H^N$ como:*

$$X(\omega) := (X_1(\omega),\ldots,X_N(\omega)).$$

Entonces, X es una variable aleatoria H^N valuada.

Demostración: Como H^N es separable, es suficiente probar que X es una función $\mathcal{F}$-medible (ver Teorema 2.19). Definamos la colección $\mathscr{B} := \left\{B \in \mathcal{B}(H^N) : X^{-1}(B) \in \mathcal{F}\right\}$. Afirmamos que $\mathscr{B}$ es una σ-álgebra de subconjuntos de H^N tal que $\mathscr{B} \subset \mathcal{B}(H^N)$.

Sea $B_i \in \mathcal{B}(H)$ para cada $i = 1,\ldots,N$ y definamos $B := B_1 \times \cdots \times B_N$ subconjunto de H^N. Es claro que $X^{-1}(B) = \cap_{i=1}^N X_i^{-1}(B_i)$ en donde $X_i^{-1}(B_i) \in \mathcal{F}$ para cada $i = 1,\ldots,N$. En consecuencia, $X^{-1}(B) \in \mathscr{B}$ por lo que:

$$\mathcal{B}(H) \times \cdots \times \mathcal{B}(H) \subset \mathscr{B} \subset \mathcal{B}(H).$$

Así pues,

$$\mathcal{S}\left(\mathcal{B}(H) \times \cdots \times \mathcal{B}(H)\right) \subset \mathscr{B} \subset \mathcal{B}(H^N)$$

en donde $\mathcal{S}\left(\mathcal{B}(H) \times \cdots \times \mathcal{B}(H)\right) = \mathcal{B}(H^N)$. Esto concluye la demostración. ■

Observa que si $X = (X_1,\ldots,X_N) : \Omega \to H^N$ es una función $\mathcal{F}$-medible entonces $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(H^N)$. Así pues, si $B_1 \in \mathcal{B}(H)$ es claro que $B := B_1 \times H \times \cdots \times H \in \mathcal{B}(H^N)$ y que $X^{-1}(B) = X_1^{-1}(B_1)$. Por tanto, $X_1^{-1}(B_1) \in \mathcal{F}$ lo que prueba que $X_1 : \Omega \to H$ es una función $\mathcal{F}$-medible. De manera análoga se prueba para las demás coordenadas que X_j es una función $\mathcal{F}$-medible $(j = 1,\ldots,N)$.

Lo anterior motiva la siguiente definición.

Definición 5.14 *Una función $X = (X_1,\ldots,X_N) : \Omega \to H^N$ es un* ***vector aleatorio*** *si y sólo si X_j es una variable aleatoria H-valuada para todo $j = 1,\ldots,N$.*

Observación 5.15 *Sean $X_1, X_2 : \Omega \to H$ variables aleatorias y consideremos el vector aleatorio $X = (X_1,X_2) : \Omega \to H$. Definamos en $(H^2,\mathcal{B}(H^2))$ la medida producto de las distribuciones de X_1 y X_2 que denotaremos por $X_{1,\sharp}P \otimes X_{2,\sharp}P$.*

Como en la Observación 5.11*, fijemos $h_1, h_2 \in H$ y usando la fórmula de Euler se tiene que:*

$$\exp\{i\langle h_j,u\rangle_H\} = \cos(\langle h_j,u\rangle_H) + i\sin(\langle h_j,u\rangle_H) \qquad \forall\, u \in H \qquad (j = 1,2).$$

En consecuencia,

$$\int_H \exp\{i\langle h_j, u\rangle_H\}\, dX_{j,\sharp}P(u) = \int_H \cos(\langle h_j, u\rangle_H)\, dX_{j,\sharp}P(u) + i\int_H \sin(\langle h_j, u\rangle_H)\, dX_{j,\sharp}P(u)$$

Por cálculos directos tenemos que, para cualesquiera $u, v \in H$

$$\begin{aligned}\exp\{i\langle h_1, u\rangle_H\}\exp\{i\langle h_2, v\rangle_H\} = (&\cos\langle h_1, u\rangle_H \cos\langle h_2, v\rangle_H - \sin\langle h_1, u\rangle_H \sin\langle h_2, v\rangle_H) + \\ &i(\cos\langle h_1, u\rangle_H \sin\langle h_2, v\rangle_H + \sin\langle h_1, u\rangle_H \cos\langle h_2, v\rangle_H).\end{aligned}$$

Aplicando el teorema de Fubini, de acuerdo a la Observación 5.11 *se tiene que:*

$$\begin{aligned}\int_{H^2} \cos\langle h_1, u_1\rangle_H \cos\langle h_2, u_2\rangle_H\, d(X_{1,\sharp}P\otimes X_{2,\sharp}P)(u_1,u_2) = &\left(\int_H \cos\langle h_1, u_1\rangle_H\, dX_{1,\sharp}P(u_1)\right)\\ &\left(\int_H \cos\langle h_2, u_2\rangle_H\, dX_{2,\sharp}P(u_2)\right),\end{aligned}$$

$$\begin{aligned}\int_{H^2} \sin\langle h_1, u_1\rangle_H \sin\langle h_2, v\rangle_H\, d(X_{1,\sharp}P\otimes X_{2,\sharp}P)(u_1,u_2) = &\left(\int_H \sin\langle h_1, u_1\rangle_H\, dX_{1,\sharp}P(u_1)\right)\\ &\left(\int_H \sin\langle h_2, u_2\rangle_H\, dX_{2,\sharp}P(u_2)\right),\end{aligned}$$

$$\begin{aligned}\int_{H^2} \cos\langle h_1, u\rangle_H \sin\langle h_2, v\rangle_H\, d(X_{1,\sharp}P\otimes X_{2,\sharp}P)(u,v) = &\left(\int_H \cos\langle h_1, u_1\rangle_H\, dX_{1,\sharp}P(u_1)\right)\\ &\left(\int_H \sin\langle h_2, u_2\rangle_H\, dX_{2,\sharp}P(u_2)\right),\end{aligned}$$

$$\begin{aligned}\int_{H^2} \sin\langle h_1, u_1\rangle_H \cos\langle h_2, u_2\rangle_H\, d(X_{1,\sharp}P\otimes X_{2,\sharp}P)(u_1,u_2) = &\left(\int_H \sin\langle h_2, u_1\rangle_H\, dX_{1,\sharp}P(u_1)\right)\\ &\left(\int_H \cos\langle h_2, u_2\rangle_H\, dX_{2,\sharp}P(u_2)\right).\end{aligned}$$

Utilizando la linealidad de la integral de Lebesgue y los cálculos anteriores podemos concluir que:

$$\int_{H^2} \prod_{j=1}^{2} \exp\{i\langle h_j, u_j\rangle_H\}\, d(X_{1,\sharp}P\otimes X_{2,\sharp}P)(u_1,u_2) = \prod_{j=1}^{2}\int_H \exp\{i\langle h_j, u_j\rangle_H\}\, dX_{j,\sharp}P(u_j).$$

Lo anterior puede generalizarse por inducción para una cantidad finita N*. Haremos uso de esta observación en las secciones restantes.*

5.1.2. Modos de convergencia

En esta sección introducimos diversos tipos de convergencia de una sucesión de variables aleatorias H-valuadas y estudiamos la relación que existe entre ellos.

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle \cdot, \cdot \rangle_H, \| \cdot \|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 5.16 *Sea $X_k : \Omega \to H$ una sucesión de variables aleatorias H-valuadas. Decimos que:*

(a) (X_k) ***converge casi dondequiera*** *a $X : \Omega \to H$ una variable aleatoria H-valuada si, existe $N \in \mathcal{F}$ con $P(N) = 0$ tal que:*

$$\lim_{k\to\infty} X_k(\omega) \to X(\omega) \qquad \forall\, \omega \in \Omega \smallsetminus N.$$

(b) (X_k) ***converge casi uniformemente*** *a $X : \Omega \to H$ una variable aleatoria H-valuada si, dada $\delta > 0$ existe $F \in \mathcal{F}$ con $P(F) < \delta$ y tal que (X_k) converge uniformemente a X en $\Omega \smallsetminus F$.*

(c) (X_k) ***converge en probabilidad*** *a $X : \Omega \to H$ una variable aleatoria H-valuada si, dada $\varepsilon > 0$*

$$\lim_{k\to\infty} P\left(\{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \varepsilon\}\right) = 0.$$

(d) (X_k) ***converge en media*** *p para $p \in [1, \infty)$ a $X \in L^p(\Omega\,;\, H)$ si $\lim_{k\to\infty} \|X_k - X\|_{L^p(\Omega;H)} = 0$.*

El siguiente resultado es bien conocido y útil en la teoría de la probabilidad clásica.

Lema 5.17 (Desigualdad de P. Tchebyshev) *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\varphi : [0, +\infty) \to [0, +\infty)$ una función no-decreciente y mayor que cero en $(0, +\infty)$. Si $f : \Omega \to \mathbb{R}$ es una función medible tal que $\varphi \circ |f| \in L^1(\Omega)$ entonces para toda $\varepsilon > 0$*

$$P(\{\omega \in \Omega \,:\, |f(\omega)| \geq \varepsilon\}) \leq \frac{1}{\varphi(\varepsilon)} \int_\Omega \varphi \circ |f| \, dP$$

Demostración: Sea $\varepsilon > 0$ y $A = \{\omega \in \Omega \,:\, |f(\omega)| \geq \varepsilon\} \subset \Omega$. Dado que φ es no decreciente y positiva en $(0, +\infty)$ entonces:

$$(\varphi \circ |f|) \cdot 1_\Omega \geq (\varphi \circ |f|) \cdot 1_A \geq \varphi(\varepsilon) \cdot 1_A$$

y, como, $\Omega = \Omega \cup (\Omega \smallsetminus A)$ de la monotonía de la integral se tiene que:

$$\int_\Omega \varphi \circ |f| \, dP = \int_\Omega \varphi \circ |f| \, dP + \int_{\Omega \smallsetminus A} \varphi \circ |f| \, dP \geq \int_A \varphi \circ |f| \, dP \geq \int_A \varphi(\varepsilon) \, dP = \varphi(\varepsilon) \cdot P(A).$$

Por lo tanto:

$$P(\{\omega \in \Omega \,:\, |f(\omega)| \geq \varepsilon\}) \leq \frac{1}{\varphi(\varepsilon)} \int_\Omega \varphi \circ |f| \, dP,$$

como afirma el enunciado. ■

Es una consecuencia directa de la desigualdad de Tchebyshev que, si $f \in L^p(\Omega)$ para algún $p \in [1, \infty)$ entonces, para toda $\varepsilon > 0$ se tiene que:

$$P(\{\omega \in \Omega \,:\, |f(\omega)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon^p} \int_\Omega |f|^p \, dP. \tag{5.3}$$

Esto nos permite concluir el siguiente resultado.

Teorema 5.18 *Sea $p \in [1, \infty)$ fijo. Si (X_k) es una sucesión de variables aleatorias H-valuadas tales que convergen en media p a $X \in L^p(\Omega; H)$ entonces $X_k \to X$ en probabilidad.*

Demostración: Sea $\varepsilon > 0$. Para $k \in \mathbb{N}$, definimos el subconjunto medible $A_k(\varepsilon) := \{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \varepsilon\}$. De la desigualdad de Tchebyshev (ver (5.3)) se tiene que:

$$P(A_k(\varepsilon)) \leq \frac{1}{\varepsilon^p} \int_\Omega \|X_k - X\|_H^p \, dP = \frac{1}{\varepsilon^p} \|X_k - X\|_{L^p(\Omega;H)}^p \qquad \forall\, k \in \mathbb{N}.$$

Tomando el límite cuando $k \to \infty$ concluimos que (X_k) converge en probabilidad a X como afirma el enunciado. ■

El recíproco no es cierto en general como lo muestra el siguiente ejemplo.

Ejemplo 5.19 *Sea $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$ y $\mu = \lambda$ la medida de Lebesgue en S. Sea $p \in [1, \infty)$ fija y la sucesión de funciones $X_k := k 1_{[0, \frac{1}{k}]}$. Entonces X_k converge a 0 en probabilidad pero no converge en media p.*

Demostración: Fijemos $\varepsilon > 0$ y sea $k \in \mathbb{N}$. Consideremos los siguientes casos:

Caso 1. Si $\varepsilon > k$ entonces el conjunto $\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}$ es vacío por lo que necesariamente:

$$\lambda\left(\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}\right) = 0.$$

Caso 2. Si $\varepsilon \leq k$ entonces:

$$\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\} = \left[0, \frac{1}{k}\right]$$

por lo que:

$$\lambda\left(\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}\right) = \lambda\left(\left[0, \frac{1}{k}\right]\right) = \frac{1}{k}.$$

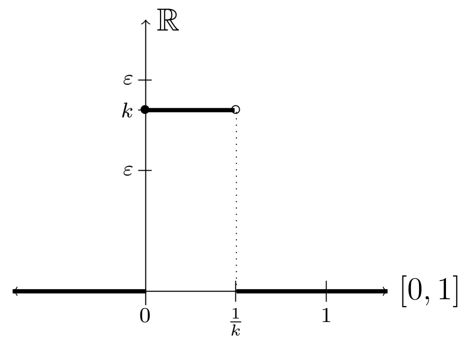


Figura 5.1: *Función* X_k

Tomando el límite cuando $k \to \infty$ concluimos que:

$$\lim_{k\to\infty} \lambda\left(\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}\right) = \lambda\left(\left[0, \frac{1}{k}\right]\right) = \lim_{k\to\infty} \frac{1}{k} = 0.$$

Esto prueba que $X_k \to 0$ en probabilidad.

Ahora, supongamos que (X_k) converge en media p. Entonces, (X_k) es de Cauchy en $(L^p([0,1]), \|\cdot\|_{L^p([0,1])})$. Sin embargo, para todo $k \in \mathbb{N}$, se tiene que:

$$\|X_{2k} - X_k\|^p_{L^p([0,1])} = \int_{[0,1]} |X_{2k} - X_k|^p \, d\lambda$$

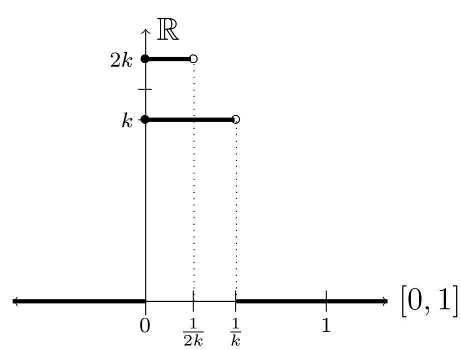


Figura 5.2: $|X_{2k} - X_k|$

en donde:

$$|X_{2k} - X_k|^p(\omega) = \begin{cases} |2k-k|^p & \text{si } \omega \in [0, \frac{1}{2k}), \\ |-k|^p & \text{si } \omega \in [\frac{1}{2k}, \frac{1}{k}), \\ 0 & \text{si } \omega \in [\frac{1}{k}, 1). \end{cases}$$

Por lo tanto:

$$\|X_{2k} - X_k\|^p_{L^p([0,1])} = \int_{[0,1]} |X_{2k} - X_k|^p \, d\mu = |k|^p \left[\lambda([0, \tfrac{1}{2k})) + \lambda([\tfrac{1}{2k}, \tfrac{1}{k}))\right] = \frac{|k|^p}{k} = k^{p-1}$$

para todo $k \in \mathbb{N}$. Tomando $\varepsilon_0 \in (0,1)$ es claro que:

$$\|X_{2k} - X_k\|^p_{L^p([0,1])} \geq \varepsilon_0 \qquad \forall k \in \mathbb{N}.$$

lo que contradice nuestra suposición. ■

Veamos un ejemplo interesante.

Ejemplo 5.20 *Sea $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}([0,1])$ y $\mu = \lambda$ la medida de Lebesgue en $\mathcal{B}([0,1])$. Definamos la siguiente sucesión de intervalos:*

$$\begin{array}{lllll}
L_1 & I_1 := [0,1) & & & \\
L_2 & I_2 := [0,1/2) & I_3 := [1/2,1) & & \\
L_3 & I_4 := [0,1/4) & I_5 := [1/4,1/2) & I_6 := [1/2,3/4) & I_7 := [3/4,1) \\
\vdots & & & & \\
L_k & I_{2^{k-1}} := [0,1/2^{k-1}) & \ldots\ldots & I_{2^k-1} := [1-1/2^{k-1},1) &
\end{array}$$

Definimos la sucesión de funciones:

$$\begin{array}{lllll}
L_1 & X_1 := 1_{I_1} & & & \\
L_2 & X_2 := 1_{I_2} & X_3 := 1_{I_3} & & \\
L_3 & X_4 := 1_{I_4} & X_5 := 1_{I_5} & X_6 := 1_{I_6} & X_7 := 1_{I_7} \\
\vdots & & & & \\
L_k & X_{2^{k-1}} := 1_{I_{2^{k-1}}} & \ldots\ldots & X_{2^k-1} := 1_{I_{2^k-1}} &
\end{array}$$

Veamos que $X_k \to X$ en probabilidad en donde $X = 0$.

Sea $\varepsilon \in (1,\infty)$. En este caso tenemos que:

$$\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\} = \varnothing \quad \forall k \in \mathbb{N}.$$

Por lo tanto:

$$\lambda\left(\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}\right) = 0 \quad \forall k \in \mathbb{N}.$$

Ahora, sea $\varepsilon \in (0,1]$ y $k \in L_j$ con $j \in \mathbb{N}$. Entonces $2^{j-1} \leq k \leq 2^j - 1$ y por lo tanto $\lambda(I_k) = \frac{1}{2^{j-1}}$ para todo $k \in L_j$.

Así pues:

$$\lambda\left(\{\omega \in \Omega \,:\, |X_k(\omega)| \geq \varepsilon\}\right) = \tfrac{1}{2^{j-1}} \leq \tfrac{2}{k}$$

pues:

$$2^{j-1} \leq k \leq 2^j - 1 = 22^{j-1} - 1 \leq 22^{j-1}$$

por lo que necesariamente:

$$\frac{1}{2^{j-1}} \leq \frac{2}{k} \leq \frac{2}{2^{j-1}}.$$

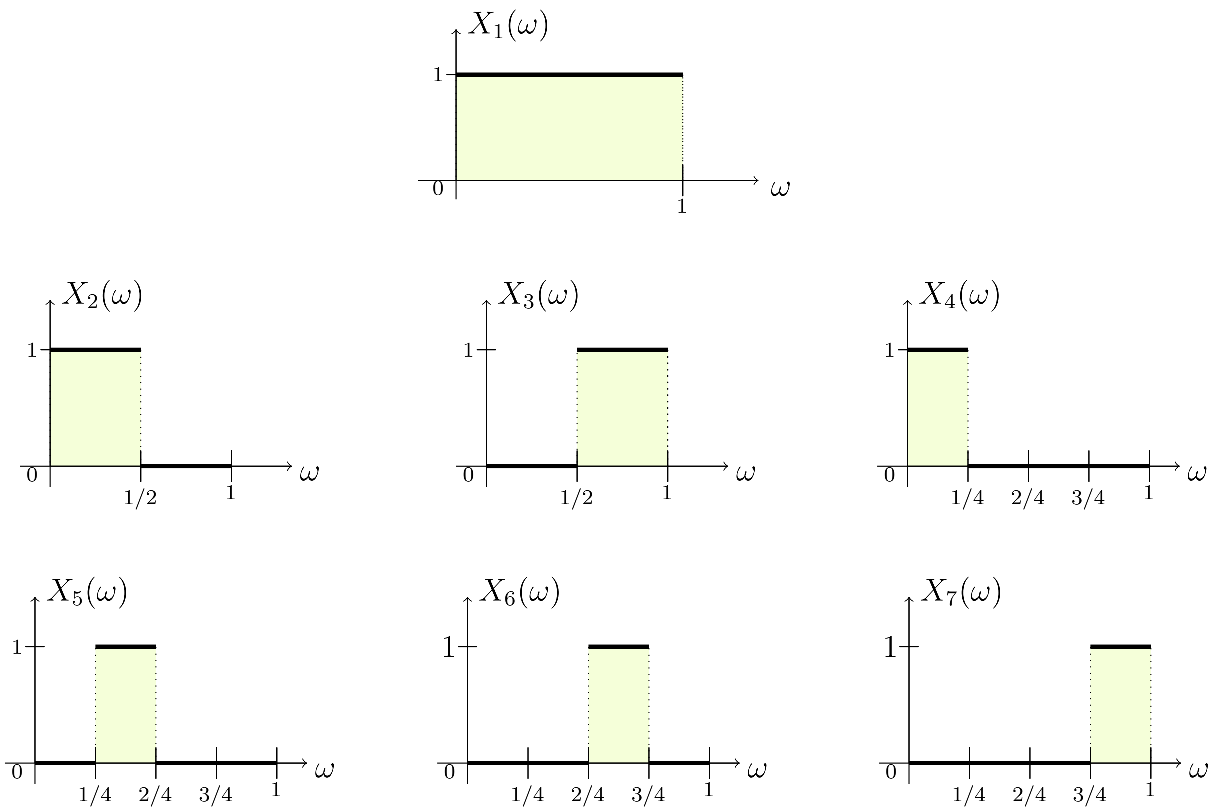
Tomando el límite cuando $k \to \infty$ tenemos que:

$$0 \leq \lim_{k\to\infty} \lambda\left(\{\omega \in \Omega\,:\, |X_k(\omega)| \geq \varepsilon\}\right) \leq \lim_{k\to\infty} \tfrac{2}{k} = 0.$$

En consecuencia, $X_k \to 0$ en probabilidad.

Notemos, sin embargo, que $X_k(\omega) \nrightarrow 0$ para todo $\omega \in [0,1)$ pues $\limsup_{k\to\infty} X_k(\omega) = 1$ y $\liminf_{k\to\infty} X_k(\omega) = 0$ para todo $\omega \in [0,1)$.

Las gráficas de las primeras siete funciones son de la siguiente forma:



Del ejemplo anterior afirmamos que la convergencia en probabilidad no implica convergencia casi dondequiera.

El siguiente resultado debido a F. Riesz y H. Weyl establece un criterio de relación entre una sucesión (X_k) de variables aleatorias H-valuadas que converge en probabilidad con respecto del criterio de convergencia casi dondequiera.

Teorema 5.21 (Riesz-Weyl) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y (X_k) una sucesión de variables aleatorias H-valuadas que converge en probabilidad a $X : \Omega \to H$ variable aleatoria H-valuada. Entonces, existe (X_{k_j}) una subsucesión de (X_k) tal que $X_{k_j} \to X$ casi dondequiera.*

Demostración: Para un mejor desarrollo de la demostración, la subdividimos en dos pasos.

PASO 1: Como (X_k) converge a X en probabilidad, entonces para $\varepsilon_j := \frac{1}{2^j} > 0$ existe $k_j \in \mathbb{N}$ tal que:

$$P\left(\left\{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \frac{1}{2^j}\right\}\right) < \frac{1}{2^j} \quad \forall k, i \geq k_j.$$

Por la condición de Cauchy podemos suponer que $k_1 < k_2 < k_3 < \ldots < k_j < k_{j+1} < \ldots$ de modo que:

$$P\left(\left\{\omega \in \Omega \,:\, \|X_{k_j}(\omega) - X(\omega)\|_H \geq \frac{1}{2^j}\right\}\right) < \frac{1}{2^j}.$$

Para lo anterior, uno procede de la siguiente manera: Elegimos k_1 tal que:

$$P\left(\left\{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \frac{1}{2}\right\}\right) < \frac{1}{2} \quad \forall k, i \geq k_1.$$

Después, elegimos $k_2 > k_1$ tal que:

$$P\left(\left\{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \frac{1}{2^2}\right\}\right) < \frac{1}{2^2} \quad \forall k, i \geq k_2.$$

Continuando de este modo obtenemos, para cada $j \in \mathbb{N}$ un $k_j > k_{j-1}$ tal que

$$P\left(\left\{\omega \in \Omega \,:\, \|X_k(\omega) - X(\omega)\|_H \geq \frac{1}{2^j}\right\}\right) < \frac{1}{2^j} \quad \forall k, i \geq k_j.$$

PASO 2: Para cada $j \in \mathbb{N}$ definamos el conjunto medible:

$$E_j := \left\{\omega \in \Omega \,:\, \|X_{k_j}(\omega) - X(\omega)\|_H \geq \frac{1}{2^j}\right\}.$$

Del PASO 1 tenemos que $P(E_j) < \frac{1}{2^j}$ y, por tanto, $\sum_{j=1}^{\infty} P(E_j) \leq \sum_{j=1}^{\infty} \frac{1}{2^j} = 1 < \infty$. El Lema de Borel-Cantelli afirma que:

$$P\left(\limsup_{j\to\infty} E_j\right) = 0.$$

Definimos el conjunto medible F_p como:

$$F_p := \bigcup_{j=p}^{\infty} E_j.$$

Notemos que:

$$\limsup_{p\to\infty} E_p = \bigcap_{p=1}^{\infty}\bigcup_{j=p}^{\infty} E_j = \bigcap_{p=1}^{\infty} F_p.$$

Sea $\omega \in \Omega \setminus F_p$ arbitraria, entonces $\omega \in \Omega \setminus E_j$ para todo $j \geq p$ y, por tanto,

$$\|X_{k_j}(\omega) - X(\omega)\|_H < \frac{1}{2^j} \qquad \forall\, j \geq p.$$

Esto prueba que $X_{k_j}(\omega) \to X(\omega)$ en $\Omega \setminus F_p$ y como

$$\Omega \setminus \limsup_{p\to\infty} E_p = \bigcup_{p=1}^{\infty}(\Omega \setminus F_p),$$

entonces $X_{k_j} \to X(\omega)$ en $\Omega \setminus \limsup_{p\to\infty} E_p$ y esto concluye la demostración. ■

En general, sabemos que la convergencia puntual de una sucesión de funciones no implica la convergencia uniforme. Sin embargo, el siguiente resultado debido a D. Egorov, establece un criterio bajo el cual una sucesión de variables aleatorias que converge casi dondequiera converge también casi uniformemente.

Teorema 5.22 (D. Egorov) *Sea (X_k) una sucesión de variables aleatorias H-valuadas tales que convergen casi dondequiera a $X : \Omega \to H$ variable aleatoria. Entonces, $X_k \to X$ casi uniformemente.*

Demostración: Para $\varepsilon > 0$ y $j \in \mathbb{N}$ dadas definimos el conjunto medible:

$$C_j(\varepsilon) = \{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H < \varepsilon\}.$$

Sea A el conjunto dado por:

$$A := \{\omega \in \Omega \,:\, X_k(\omega) \to X(\omega)\}.$$

Por hipótesis $P(\Omega \setminus A) = 0$ y, por propiedades de medida, tenemos que $0 = P(\Omega \setminus A) = P(\Omega) \setminus P(A)$. En consecuencia, $P(\Omega) = P(A)$.

Sea $\omega \in A$, entonces existe $k_0 \in \mathbb{N}$ tal que $\|X_k(\omega) - X(\omega)\|_H < \varepsilon$ para todo $k \geq k_0$. Es decir, $\omega \in C_k(\varepsilon)$ para todo $k \geq k_0$ lo que implica que:

$$\omega \in \bigcap_{k=k_0}^{\infty} C_k(\varepsilon) \quad \text{para algún } k_0 \in \mathbb{N}.$$

Por lo tanto:

$$A \subset \bigcup_{k=1}^{\infty}\bigcap_{j=k}^{\infty} C_j(\varepsilon) = \liminf_{k\to\infty} C_k(\varepsilon).$$

De las propiedades de medida tenemos que:

$$P(\Omega) = P(A) \leq P\left(\liminf_{k\to\infty} C_k(\varepsilon)\right) \leq \lim_{k\uparrow\infty} P\left(\bigcap_{j=k}^{\infty} C_j(\varepsilon)\right) \leq P(\Omega)$$

de donde obtenemos:

$$P(\Omega) = \lim_{k\uparrow\infty} \mu\left(\bigcap_{j=k}^{\infty} C_j(\varepsilon)\right).$$

Entonces:

$$0 = \lim_{k\uparrow\infty}\left[P(\Omega) - P\left(\bigcap_{j=k}^{\infty} C_j(\varepsilon)\right)\right] = \lim_{k\uparrow\infty}\left[P\left(\bigcup_{j=k}^{\infty}(\Omega \setminus C_j(\varepsilon))\right)\right]$$

en donde:

$$\bigcup_{j=k}^{\infty}(\Omega \setminus C_j(\varepsilon)) = \bigcup_{j=k}^{\infty}\left\{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H \geq \varepsilon\right\}.$$

Así pues:

$$\lim_{k\uparrow\infty} P\left(\bigcup_{j=k}^{\infty}\{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H \geq \varepsilon\}\right) = 0.$$

Tomando $\varepsilon = \frac{1}{m}$ para $m \in \mathbb{N}$ y $\delta > 0$ arbitraria es posible hallar $k_m \in \mathbb{N}$ tal que:

$$P\left(\bigcup_{j=k_m}^{\infty}\left\{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H \geq \frac{1}{m}\right\}\right) < \frac{\delta}{2^m}.$$

Sea

$$F = \bigcup_{m=1}^{\infty}\bigcup_{j=k_m}^{\infty}\left\{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H \geq \frac{1}{m}\right\}$$

entonces:

$$P(F) \leq \sum_{m=1}^{\infty} P\left(\bigcup_{j=k_m}^{\infty}\left\{\omega \in \Omega \,:\, \|X_j(\omega) - X(\omega)\|_H \geq \frac{1}{m}\right\}\right) < \sum_{m=1}^{\infty}\frac{\delta}{2^m} = \delta.$$

Así pues, $\omega \notin F$ entonces $\|X_j(\omega) - X(\omega)\|_H < \frac{1}{m}$ para todo $j \geq k_m$ para todo $m \in \mathbb{N}$. Esto demuestra que $X_k \to X$ uniformemente en $\Omega \setminus F$. ■

5.2. Independencia

Definición 5.23 (Independencia) *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Decimos que:*

(i) Los eventos $A_1, \dots, A_N \in \mathcal{F}$ ($N \in \mathbb{N}$) son independientes si, para cada subconjunto $\{1, \dots, n\}$ de $\{1, \dots, N\}$ y cada permutación $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ se tiene que:

$$P(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)}) = \prod_{j=1}^{n} P(A_{\sigma(j)}).$$

(ii) $X_1, \dots, X_N$ variables aleatorias H-valuadas ($N \in \mathbb{N}$) son independientes si para cualesquiera $A_1, \dots, A_N \in \mathcal{B}(H)$ los eventos $\{X_1 \in A_1\}, \dots, \{X_N \in A_N\}$ son independientes.

(iii) Una sucesión $X_k : \Omega \to Y$ de variables aleatorias H-valuadas es independiente si cualquier subcolección finita de (X_k) es independiente.

(iv) Sean $\mathcal{G}_1, \dots, \mathcal{G}_N$ sub σ-álgebras de $\mathcal{F}$ ($N \in \mathbb{N}$). Decimos que son independientes si para cualesquiera $A_1 \in \mathcal{G}_1, \dots, A_N \in \mathcal{G}_N$ los eventos $A_1, \dots, A_N$ son independientes.

Dadas $X_1, \dots, X_N : \Omega \to H$ variables aleatorias H-valuadas, la distribución de la variable aleatoria $X = (X_1, \dots, X_N) : \Omega \to H^N$ dada en la Proposición 5.10, que denotamos como $X_\sharp P : \mathcal{B}(H^N) \to \mathbb{R}$, está dada por:

$$X_\sharp P(B) := P\{(X_1, \dots, X_N) \in B\} \quad \forall\, B \in \mathcal{B}(H^N).$$

Por otra parte, consideremos la medida producto de las distribuciones de las variables aleatorias $X_1, \dots, X_N : \Omega \to H$. Es decir,

$$\bigotimes_{j=1}^{N} X_{j,\sharp} P.$$

El teorema de la medida producto asegura que $\bigotimes_{j=1}^{N} X_{\sharp,j} P$ es la única medida de probabilidad tal que:

$$\bigotimes_{j=1}^{N} X_{j,\sharp} P(B_1 \times \cdots \times B_N) = \prod_{j=1}^{N} X_{j,\sharp} P(B_j)$$

para todo $B_1 \times \cdots \times B_N \in \mathcal{B}(H) \times \cdots \times \mathcal{B}(H)$.

Supongamos que $X_1, \dots, X_N : \Omega \to H$ son independientes. Consideremos $B_j \in \mathcal{B}(H)$ para cada $j = 1, \dots, N$ y definamos $B := B_1 \times \cdots B_N \in \mathcal{B}(H) \times \cdots \times \mathcal{B}(H) \subset \mathcal{B}(H^N)$.

Como $X_1, \dots, X_N$ son independientes, entonces:

$$P(\{X_1 \in B_1\} \cap \cdots \cap \{X_N \in B_N\}) = \prod_{j=1}^{N} P(\{X_j \in B_j\}).$$

Dado que $P(\{X \in B\}) = P(\{X_1 \in B_1\} \cap \cdots \cap \{X_N \in B_N\})$ (ver Proposición 5.13) entonces:

$$X_\sharp P(B) = \prod_{j=1}^{N} X_{j,\sharp} P(B_j) = \bigotimes_{j=1}^{N} X_{j,\sharp} P(B).$$

El teorema de unicidad de medidas implica que $X_\sharp P = \bigotimes_{j=1}^{N} X_{j,\sharp} P$.

Inversamente: Si $X_\sharp P = \bigotimes_{j=1}^{N} X_{j,\sharp} P$ entonces para $B = B_1 \times \cdots B_N \in \mathcal{B}(H^N)$ se tiene que:

$$X_\sharp P(B) = P(\{X_1 \in B_1\} \cap \cdots \cap \{X_N \in B_N\}) = \bigotimes_{j=1}^{N} X_{j,\sharp} P(B) = \prod_{j=1}^{N} P(\{X_j \in B_j\})$$

Es decir, $X_1, \ldots, X_N$ son independientes.

Hemos demostrado el siguiente resultado.

Proposición 5.24 *Sean $X_1, \ldots, X_N : \Omega \to H$ variables aleatorias H-valuadas y sea $X := (X_1, \ldots, X_N) : \Omega \to H^N$. Las siguientes son equivalentes:*

(a) $X_1, \ldots, X_N$ son independientes.

(b) $X_\sharp P = \bigotimes_{j=1}^{N} X_{j,\sharp} P$.

Este resultado nos permitirá demostrar la siguientes proposición.

Proposición 5.25 *Sean $X_1, \ldots, X_N : \Omega \to H$ variables aleatorias H-valuadas y sea $X := (X_1, \ldots, X_N) : \Omega \to H^N$. Las siguientes son equivalentes:*

(a) $X_1, \ldots, X_N$ son independientes.

(b) $\widehat{X}(h_1, \ldots, h_N) = \prod_{j=1}^{N} \widehat{X}_j(h_j)$ para todo $(h_1, \ldots, h_N) \in H^N$.

Demostración: Probaremos únicamente el caso cuando $N = 2$ pues el caso general se deduce fácilmente de éste.

$(a) \Rightarrow (b)$: Sean $h_1, h_2 \in H$. Aplicando la Proposición 5.24 y el teorema de Fubini concluimos que:

$$\begin{aligned}
\widehat{X}(h_1, h_2) &= \int_{H^2} \exp\left\{i\langle (h_1, h_2), (u, v)\rangle_{H^2}\right\} \, dX_\sharp P(u, v) \\
&= \int_{H^2} \exp\left\{i\left(\langle h_1, u\rangle_H + \langle h_2, v\rangle_H\right)\right\} \, dX_\sharp P(u, v) \\
&= \int_{H^2} \exp\left\{i\langle h_1, u\rangle_H\right\} \exp\left\{i\langle h_2, v\rangle_H\right\} \, dX_\sharp P(u, v) \\
&= \int_H \left(\int_H \exp\left\{i\langle h_1, u\rangle_H\right\} \exp\left\{i\langle h_2, v\rangle_H\right\} \, dX_{1,\sharp}(u) \right) \, dX_{2,\sharp}(v) \\
&= \left(\int_H \exp\left\{i\langle h_1, u\rangle_H\right\} \, dX_{1,\sharp}(u) \right) \left(\int_H \exp\left\{i\langle h_2, v\rangle_H\right\} \, dX_{2,\sharp}(v) \right) \\
&= \widehat{X}_1(h_1) \, \widehat{X}_2(h_2).
\end{aligned}$$

$(b) \Rightarrow (a)$: La transformada de Fourier de la medida producto $X_{\sharp,1}P \otimes X_{\sharp,2}P : \mathcal{B}(H^2) \to \mathbb{R}$ está dada por:

$$\begin{aligned}
X_{1,\sharp}\widehat{P \otimes X}_{2,\sharp}P(h_1, h_2) &= \int_{H^2} \exp\left\{i\langle (h_1, h_2), (u, v)\rangle_{H^2}\right\} \, d(X_{1,\sharp}P \otimes X_{2,\sharp}P)(u, v) \\
&= \int_{H^2} \exp\left\{i\left(\langle h_1, u\rangle_H + \langle h_2, v\rangle_H\right)\right\} \, d(X_{1,\sharp}P \otimes X_{2,\sharp}P)(u, v),
\end{aligned}$$

para cualquier $(h_1, h_2) \in H^2$.

Aplicando el teorema de Fubini obtenemos:

$$\begin{aligned}
&\int_{H^2} \exp\left\{i\left(\langle h_1, u\rangle_H + \langle h_2, v\rangle_H\right)\right\} \, d(X_{1,\sharp}P \otimes X_{2,\sharp}P)(u, v) = \\
&= \int_H \left(\int_H \exp\left\{i\langle h_1, u\rangle_H\right\} \exp\left\{i\langle h_2, v\rangle_H\right\} \, dX_{1,\sharp}(u) \right) \, dX_{2,\sharp}(v) \\
&= \left(\int_H \exp\left\{i\langle h_1, u\rangle_H\right\} \, dX_{1,\sharp}(u) \right) \left(\int_H \exp\left\{i\langle h_2, v\rangle_H\right\} \, dX_{2,\sharp}(v) \right) \\
&= \widehat{X_{1,\sharp}P}(h_1) \, \widehat{X_{2,\sharp}P}(h_2).
\end{aligned}$$

De lo anterior y el teorema de cambio de variable concluimos que:

$$X_{1,\sharp}\widehat{P \otimes X}_{2,\sharp}P(h_1, h_2) = \widehat{X}_1(h_1) \, \widehat{X}_2(h_2).$$

El Teorema 4.8 asegura que $X_\sharp P = \bigotimes_{j=1}^{N} X_{j,\sharp}P$ lo que implica que X_1 y X_2 son independientes por la Proposición 5.24. ■

Usando el teorema de cambio de variable podemos concluir que la transformada de Fourier de la variable aleatoria $X = (X_1, \ldots, X_N) : \Omega \to H^N$ está dada por:

$$\widehat{X}(h_1, \ldots, h_N) = \int_\Omega \exp\left\{ i \sum_{j=1}^N \langle h_j, X_j(\omega)\rangle_H \right\} dP(\omega) = E\left(i \sum_{j=1}^N \langle h_j, X_j\rangle_H \right)$$

para cualquier $(h_1, \ldots, h_N) \in H^N$.

Proposición 5.26 *Consideremos H y K espacios de Hilbert. Sean $X, Y : \Omega \to H$ variables aleatorias H-valuadas e independientes y sean $\varphi, \phi : H \to K$ funciones medibles. Entonces las variables aleatorias $\varphi(X), \phi(Y) : \Omega \to K$ son independientes.*

Demostración: Si A y B son dos cualesquiera conjuntos de Borel en K, entonces:

$$\begin{aligned} P\left(\{\varphi(X) \in A\} \cap \{\phi(Y) \in B\}\right) &= P\left(\{X \in \varphi^{-1}(A)\} \cap \{Y \in \phi^{-1}(B)\}\right) \\ &= P\left(\{X \in \varphi^{-1}(A)\}\right) P\left(\{Y \in \phi^{-1}(B)\}\right) \\ &= P\left(\{\varphi(X) \in A\}\right) P\left(\{\phi(Y) \in B\}\right) \end{aligned}$$

lo que prueba que $\varphi(X)$ y $\phi(Y)$ son independientes. ■

Proposición 5.27 *Sean $X_1, \ldots, X_N : \Omega \to H$ variables aleatorias H-valuadas e independientes. Entonces, $E(\varphi_1(X_1) \cdots \varphi_N(X_N)) = E(\varphi_1(X_1)) \cdots E(\varphi_N(X_N))$ para cualesquiera $\varphi_1, \ldots, \varphi_N : H \to \mathbb{R}$ funciones acotadas y medibles.*

Demostración: Del teorema de cambio de variable se tiene que:

$$E(\varphi_j(X_j)) = \int_\Omega (\varphi_j \circ X_j)(\omega_j)\, dP(\omega_j) = \int_H \varphi_j(x_j)\, dX_{j,\sharp}P(x_j) \quad \forall\, j = 1, \ldots, N.$$

Consideremos $X := (X_1, \ldots, X_N) : \Omega \to H^N$ variable aleatoria H^N-valuada y definamos $\varphi : H^N \to \mathbb{R}$ como:

$$\varphi(u_1, \ldots, u_N) := \prod_{j=1}^N \varphi_j(u_j).$$

Es claro que φ es acotada y medible. Del teorema de cambio de variable se tiene que:

$$E(\varphi \circ X) = \int_\Omega (\varphi \circ X)(\omega)\, dP(\omega) = \int_{H^N} \varphi(u_1, \ldots, u_N)\, dX_\sharp P(u_1, \ldots, u_N).$$

Como $X_1, \ldots, X_N$ son independientes entonces $X_\sharp P = \bigotimes_{j=1}^N X_{\sharp,j} P$ por la Proposición 5.19. Aplicando el teorema de Fubini concluimos que:

$$\begin{aligned}
\int_{H^N} \varphi(u_1, \ldots, u_N)\, dX_\sharp P(u_1, \ldots, u_N) &= \int_H \cdots \left(\int_H \prod_{j=1}^N \varphi_j(u_j) dX_{1,\sharp} P(u_1) \right) \cdots\, dX_{N,\sharp} P(u_N) \\
&= \prod_{j=1}^N \int_H \varphi_j(u_j) dX_{j,\sharp} P(u_j) \\
&= \prod_{j=1}^N E(\varphi_j(X_j)).
\end{aligned}$$

En consecuencia,

$$E\left(\prod_{j=1}^N \varphi_j(X_j) \right) = E(\varphi \circ X) = \prod_{j=1}^N E(\varphi_j(X_j)),$$

como afirma el enunciado. ■

Proposición 5.28 *Sean (X_k) y (Y_k) sucesiones de variables aleatorias H-valuadas tales que $X_k \to X$ y $Y_k \to Y$ en probabilidad. Si X_j es independiente de Y_k para toda $k \in \mathbb{N}$ y $j \in \mathbb{N}$, entonces X y Y son independientes.*

Demostración: Sean $\xi_j := X_{k_j}$ y $\eta_j := Y_{k_j}$ subsucesiones de (X_k) y (Y_k) respectivamente tales que $\xi_j \to X$ c.d. y $\eta_j \to Y$ c.d. (ver Teorema 5.21).

Definamos $Z_j : \Omega \to H \times H$ como $Z_j(\omega) := (\xi_j(\omega), \eta_j(\omega))$ y $Z : \Omega \to H \times H$ como $Z(\omega) := (X(\omega), Y(\omega))$. Es claro que $Z_j \to Z$ c.d.

Del teorema de la convergencia dominada de Lebesgue y la Proposición 5.24 se tiene que:

$$\begin{aligned}
\widehat{Z}(h_1, h_2) &= E(\exp\{i\left(\langle h_1, X\rangle_H + \langle h_1, Y\rangle_H\right)\}) \\
&= \int_\Omega \exp\left\{i\langle h_1, X(\omega)\rangle_H\right\} \exp\left\{i\langle h_2, Y(\omega)\rangle_H\right\}\, dP(\omega) \\
&= \lim_{j\to\infty} \int_\Omega \exp\left\{i\langle h_1, \xi_j(\omega)\rangle_H\right\} \exp\left\{i\langle h_2, \eta_j(\omega)\rangle_H\right\}\, dP(\omega) \\
&= \lim_{j\to\infty} E(\exp\{i\langle h_1, \xi_j\rangle_H)\ \lim_{j\to\infty} E(\exp\{i\langle h_2, \eta_j\rangle_H) \\
&= E(\exp\{i\langle h_1, X\rangle_H)\, E(\exp\{i\langle h_1, Y\rangle_H) \\
&= \widehat{X}(h_1)\widehat{Y}(h_2).
\end{aligned}$$

El resultado se sigue de la Proposición 5.24. ■

Definición 5.29 *Sea $X : \Omega \to H$ una variable aleatoria H-valuada. Decimos que X es* ***simétrica*** *si, X y $-X$ son idénticamente distribuidas. Es decir, $X_\sharp P = -X_\sharp P$.*

Sean X, Y variables aleatorias H-valuadas independientes y tal que X es simétrica. La Proposición 5.26 implica que $-X$ y Y son independientes. Aplicando la Proposición 5.24 se sigue que:

$$\begin{aligned}\widehat{X+Y}(u) &= E(i\langle u, X+Y\rangle_H) = E(i\langle u, X\rangle_H + i\langle u, Y\rangle_H)\\ &= E(i\langle u, X\rangle_H)\,E(i\langle u, Y\rangle_H) = \widehat{X}(u)\widehat{Y}(u) = \widehat{-X}(u)\widehat{Y}(u) = \widehat{-X+Y}(u).\end{aligned}$$

El Teorema 4.8 asegura que $X+Y$ y $-X+Y$ son idénticamente distribuidas.

Proposición 5.30 *Sean X, Y variables aleatorias H-valuadas independientes y tal que X es simétrica. Entonces, para todo $p \in [1, \infty)$ se tiene que:*

$$E(\|X\|_H^p) \leq E(\|X+Y\|_H^p)$$

Demostración: Dado que $X+Y$ y $-X+Y$ son idénticamente distribuidas, entonces:

$$\begin{aligned}(E(\|X\|_H^p))^{1/p} &= \frac{1}{2}\,(E(\|(X+Y)+(X-Y)\|_H^p))^{1/p}\\ &\leq \frac{1}{2}\,(E(\|X+Y\|_H^p))^{1/p} + \frac{1}{2}\,(E(\|X-Y\|_H^p))^{1/p}\\ &= (E(\|X+Y\|_H^p))^{1/p}\end{aligned}$$

como afirma el enunciado. ■

5.3. Esperanza condicional

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{G} \subset \mathcal{F}$ una sub σ-álgebra.

En el Ejemplo 1.27 demostramos que, dada $\vartheta : \Omega \to \mathbb{R}$ una variable aleatoria real tal que $\|\vartheta\|_{L^2(\Omega)} < \infty$, existe una única variable aleatoria $\xi : \Omega \to \mathbb{R}$ llamada **esperanza condicional de ϑ con respecto de la σ-álgebra** $\mathcal{G}$ tal que ξ cumple que:

(i) ξ es $\mathcal{G}$-medible.

(ii) Para cada $A \in \mathcal{G}$,

$$\int_A \xi(\omega)\, dP(\omega) = \int_A \vartheta(\omega)\, dP(\omega).$$

A la esperanza condicional de ϑ con respecto de la σ-álgebra $\mathcal{G}$ la denotaremos por $E(\vartheta \,|\, \mathcal{G})$. El primer objetivo de esta sección es extender la noción de esperanza condicional con respecto de la σ-álgebra $\mathcal{G}$ a variables aleatorias reales en $L^1(\Omega)$. Para ello, requerimos de los siguientes lemas.

En esta sección, por simplicidad en algunas demostraciones escribiremos $E(\vartheta)$ en lugar de $\int_\Omega \vartheta(\omega)\, dP(\omega)$ para variables aleatorias reales que son Lebesgue-integrables.

Lema 5.31 *(a) Para cada $\vartheta \in L^2(\Omega)$ se tiene que:*

$$E(|E(\vartheta \,|\, \mathcal{G})|) \leq E(|\vartheta|). \tag{5.4}$$

(b) Si $\vartheta \in L^2(\Omega)$ es tal que $\vartheta \geq 0$, entonces $E(\vartheta \,|\, \mathcal{G}) \geq 0$.

(c) Si $\vartheta, \eta \in L^2(\Omega)$ son tales que $\vartheta \geq \eta$, entonces $E(\vartheta \,|\, \mathcal{G}) \geq E(\eta \,|\, \mathcal{G})$.

Demostración: (a) Sea $\vartheta \in L^2(\Omega)$ y consideremos la descomposición de ϑ dada por $\vartheta = \vartheta^+ - \vartheta^-$ en donde es claro que $\vartheta^+, \vartheta^- \in L^2(\Omega)$.

Entonces,

$$\begin{aligned} |E(\vartheta \,|\, \mathcal{G})| = \left|E(\vartheta^+ \,|\, \mathcal{G}) - E(\vartheta^- \,|\, \mathcal{G})\right| &\leq |E(\vartheta^+ \,|\, \mathcal{G})| + |E(\vartheta^- \,|\, \mathcal{G})| \\ &= E(\vartheta^+ \,|\, \mathcal{G}) + E(\vartheta^- \,|\, \mathcal{G}) = E(|\vartheta| \,|\, \mathcal{G}), \end{aligned}$$

y, por lo tanto, del inciso *(ii)* del Ejemplo 1.27 concluimos que

$$E(|E(\vartheta \,|\, \mathcal{G})|) \leq E(E(|\vartheta| \,|\, \mathcal{G})) = E(|\vartheta|).$$

(b) Es inmediato de las propiedades de la integral de Lebesgue y la esperanza condicional en $L^2(\Omega)$.

(c) Se sigue del inciso anterior al considerar que $\vartheta - \eta \geq 0$ y la linealidad de la integral de Lebesgue. ■

Lema 5.32 *$L^p(\Omega)$ es denso en $L^1(\Omega)$ para cada $p \in (1, \infty)$.*

Demostración: Fijemos $p \in (1, \infty)$. Sea $\vartheta \in L^1(\Omega)$ y consideremos el conjunto medible $A_k := \{\omega \in \Omega \,:\, |\vartheta(\omega)| < k\}$ para todo $k \in \mathbb{N}$. Definamos la sucesión de variables aleatorias reales $\vartheta_k : \Omega \to \mathbb{R}$ como $\vartheta_k(\omega) := (\vartheta \cdot 1_{A_k})(\omega)$. Observa que, $\|\vartheta_k\|^p_{L^p(\Omega)} = \int_{A_k} |\vartheta(\omega)|^p \, dP(\omega) \leq k^p$ para todo $k \in \mathbb{N}$ lo que implica que (ϑ_k) es una sucesión de elementos de $L^p(\Omega)$.

En consecuencia,

$$\begin{aligned} \int_\Omega |\vartheta(\omega) - \vartheta_k(\omega)| \, dP(\omega) &= \int_{A_k} |\vartheta(\omega) - \vartheta_k(\omega)| \, dP(\omega) + \int_{\Omega \setminus A_k} |\vartheta(\omega) - \vartheta_k(\omega)| \, dP(\omega) \\ &= \int_{\Omega \setminus A_k} |\vartheta(\omega)| \, dP(\omega) \\ &= \int_\Omega |\vartheta(\omega) \cdot 1_{|\vartheta(\omega)| \geq k}| \, dP(\omega) := \||\vartheta| \cdot 1_{|\vartheta| \geq k}\|_{L^1(\Omega)} \to 0 \end{aligned}$$

por el teorema de la convergencia dominada de Lebesgue pues $|\vartheta| \cdot 1_{|\vartheta| \geq k} \to 0$ c.d. (ver [17]). Por tanto, $\vartheta_k \to \vartheta$ en $L^1(\Omega)$. Esta última afirmación junto con la relación $L^p(\Omega) \to L^1(\Omega)$ aseguran que $L^p(\Omega)$ es denso en $L^1(\Omega)$. ■

De este modo, la definición de esperanza condicional con respecto de la σ-álgebra $\mathcal{G}$ se extiende a variables aleatorias reales que son Lebesgue-integrables como lo muestra el siguiente teorema.

Teorema 5.33 *Sea $\vartheta \in L^1(\Omega)$. Existe una única variable aleatoria $\xi : \Omega \to \mathbb{R}$ en $L^1(\Omega)$ tal que ξ cumple que:*

(a) ξ es $\mathcal{G}$-medible.

(b) Para cada $A \in \mathcal{G}$,

$$\int_A \xi(\omega)\, dP(\omega) = \int_A \vartheta(\omega)\, dP(\omega).$$

Demostración: Supongamos que $\vartheta \geq 0$. Para $k \in \mathbb{N}$, definamos $\vartheta_k : \Omega \to \mathbb{R}$ como $\vartheta_k := \text{mín}\{\vartheta, k\}$ la cual es claramente una función medible y acotada en $(\Omega, \mathcal{F})$. En consecuencia, $\vartheta_k \in L^2(\Omega)$ para cada $k \in \mathbb{N}$. Así, (ϑ_k) es una sucesión en $L^2(\Omega)$ que converge puntualmente a ϑ. Del Ejemplo 1.27 existe $E(\vartheta_k \,|\, \mathcal{G}) : \Omega \to \mathbb{R}$ tal que sastiface las propiedades *(a)* y *(b)* del enunciado para cada ϑ_k y para cada $k \in \mathbb{N}$. El Lema 5.31 asegura que $(E(\vartheta_k \,|\, \mathcal{G}))$ es una sucesión no decreciente y no negativa de elementos de $L^2(\Omega)$. Definiendo $\xi := \text{lím}_{k\to\infty} E(\vartheta_k \,|\, \mathcal{G})$ es claro que ξ es $\mathcal{G}$-medible por ser el límite puntual de una sucesión de funciones $\mathcal{G}$-medibles. Además, el Lema 5.32 asegura que $(E(\vartheta_k \,|\, \mathcal{G}))$ es una sucesión de funciones Lebesgue-integrables.

Aplicando el teorema de la convergencia monótona concluimos que $\xi \in L^1(\omega)$ y que

$$\int_A \vartheta(\omega)\, dP(\omega) = \lim_{k\to\infty} \int_A \vartheta_k(\omega)\, dP(\omega) = \lim_{k\to\infty} \int_A E(\vartheta_k \,|\, \mathcal{G})\, dP(\omega) = \int_A \xi(\omega)\, dP(\omega)$$

para todo $A \in \mathcal{G}$.

El caso general se reduce a aplicar lo anterior a las funciones ϑ^+, ϑ^-. ■

Para $\vartheta \in L^1(\Omega)$ dada, denotaremos por $E(\vartheta \,|\, \mathcal{G})$ a la esperanza condicional de ϑ con respecto de la σ-álgebra $\mathcal{G}$.

Veamos algunas propiedades de esta función.

Proposición 5.34 *Sean ϑ, η variables aleatorias reales en $L^1(\Omega)$ y $a, b \in \mathbb{R}$. Entonces:*

(a) Si $\vartheta \geq 0$ entonces $E(\vartheta \,|\, \mathcal{G}) \geq 0$.

(b) $E(E(\vartheta \,|\, \mathcal{G})) = E(\vartheta)$.

(c) $E(a\vartheta + b\eta \,|\, \mathcal{G}) = aE(\vartheta \,|\, \mathcal{G}) + bE(\eta \,|\, \mathcal{G})$.

(d) Si $\vartheta \leq \eta$ *entonces* $E(\vartheta \,|\, \mathcal{G}) \leq E(\eta \,|\, \mathcal{G})$*.*

(e) Si ϑ *es* $\mathcal{G}$*-medible, entonces* $E(\vartheta \,|\, \mathcal{G}) = \vartheta$ *c.d.*

Demostración: (a) Es inmediato de la propiedad *(b)* del Teorema 5.33 y las propiedades de la integral de Lebesgue.

(b) Es inmediato del inciso *(b)* del Teorema 5.33 al considerar $A = \Omega \in \mathcal{G}$.

(c) Es claro que $aE(\vartheta \,|\, \mathcal{G}) + bE(\eta \,|\, \mathcal{G})$ es una función $\mathcal{G}$-medible. Además, para cada $A \in \mathcal{G}$ se tiene que:

$$\begin{aligned}
\int_A aE(\vartheta \,|\, \mathcal{G}) + bE(\eta \,|\, \mathcal{G})\, dP &= a \int_A E(\vartheta \,|\, \mathcal{G})\, dP + b \int_\Omega E(\eta \,|\, \mathcal{G})\, dP \\
&= a \int_\Omega \vartheta\, dP + b \int_\Omega \eta\, dP \\
&= \int_\Omega a\vartheta + b\eta\, dP
\end{aligned}$$

lo que implica que $E(a\vartheta + b\eta \,|\, \mathcal{G}) = aE(\vartheta \,|\, \mathcal{G}) + bE(\eta \,|\, \mathcal{G})$.

(d) Es consecuencia directa de los incisos *(a)* y *(b)* al considerar la variable $\eta - \vartheta \geq 0$.

(e) Como ϑ es $\mathcal{G}$-medible, es claro que satisface las condiciones del Teorema 5.33. De la unicidad se obtiene la igualdad casi dondequiera. ■

Una desigualdad importante es la siguiente.

Proposición 5.35 (Desigualdad de Jensen condicional) *Sea* $\phi : \mathbb{R} \to \mathbb{R}$ *una función convexa. Entonces, para cada* $\vartheta \in L^1(\Omega)$ *tal que* $\phi \circ \vartheta \in L^1(\Omega)$ *se tiene que:*

$$\phi \circ E(\vartheta \,|\, \mathcal{G}) \leq E((\phi \circ \vartheta) \,|\, \mathcal{G}), \quad \textit{c.d.} \tag{5.5}$$

Demostración: Si $a, b \in \mathbb{R}$ son tales que $at + b \leq \phi(t)$ para cada $t \in \mathbb{R}$, de la Proposición 5.34 se sigue que:

$$aE(\vartheta \,|\, \mathcal{G}) + b = E(a\vartheta + b \,|\, \mathcal{G}) \leq E(\phi \circ \vartheta \,|\, \mathcal{G}) \quad \textit{c.d.}$$

Dado que ϕ es convexa, es posible hallar sucesiones (a_k) y (b_k) de números reales tales que $\phi(t) = \sup_{k \geq 1}(a_k t + b_k)$ para todo $k \in \mathbb{N}$. En consecuencia,

$$\phi \circ E(\vartheta \,|\, \mathcal{G}) = \sup_{k \geq 1} E(a_k \vartheta \,|\, \mathcal{G}) + b_k \leq E(\phi \circ \vartheta) \,|\, \mathcal{G}), \quad \textit{c.d.}$$

como afirma el enunciado. ■

El objetivo final de esta sección es extender el concepto de esperanza condicional para variables aleatorias que toman valores en un espacio de Hilbert separable.

Teorema 5.36 *Sea $H = (H, \langle\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable y $X : \Omega \to H$ una variable aleatoria que es Bochner-integrable. Entonces, existe una única, salvo en un evento de probabilidad cero, variable aleatoria $\xi : \Omega \to H$ Bochner-integrable tal que:*

(a) ξ es $\mathcal{G}$-medible.

(b) Para cada $A \in \mathcal{G}$,

$$\mathfrak{B}\int_A \xi(\omega)\,dP(\omega) = \mathfrak{B}\int_A X(\omega)\,dP(\omega).$$

Demostración: Probamos primero la unicidad y luego la existencia.

Unicidad: Supongamos que existen $\xi, \widetilde{\xi} : \Omega \to H$ variables aleatorias que satisfacen *(a)* y *(b)* y tales que $P(\{\omega \in \Omega \,:\, \xi(\omega) \neq \tilde{\xi}(\omega)\}) > 0$.

Dado que H es separable, existe $\{e_k \,:\, k \in \mathbb{N}\}$ subconjunto denso. Así, para $\omega \in \Omega$ tal que $\xi(\omega) - \widetilde{\xi}(\omega) \neq 0_H$, existe $e_j \in H \setminus \{0_H\}$ tal que:

$$\left\|\xi(\omega) - \widetilde{\xi}(\omega) - e_j\right\|_H < \frac{\|e_j\|_H}{3}.$$

En consecuencia, $P\left(\left\{\omega \in \Omega \,:\, \left\|\xi(\omega) - \widetilde{\xi}(\omega) - e_j\right\|_H < \frac{\|e_j\|_H}{3}\right\}\right) > 0.$

El Teorema de Hahn-Banach asegura que existe $\psi : H \to \mathbb{R}$ función lineal y continua tal que $\psi(e_j) = \|e_j\|_H$ y $\|\psi\|_{\mathcal{L}_c(H,\mathbb{R})} = 1$. Es inmediato que $\psi \circ \xi$ y $\psi \circ \widetilde{\xi}$ son variables aleatorias reales tales que:

$$|\psi(\xi(\omega))| \leq \|\psi\|_{\mathcal{L}_c(H,\mathbb{R})}\,\|\xi(\omega)\|_H = \|\xi(\omega)\|_H,$$
$$|\psi(\widetilde{\xi}(\omega))| \leq \|\psi\|_{\mathcal{L}_c(H,\mathbb{R})}\,\|\widetilde{\xi}(\omega)\|_H = \|\widetilde{\xi}(\omega)\|_H.$$

Es decir, $\psi \circ \xi$ y $\psi \circ \widetilde{\xi}$ son Lebesgue-integrables. Además, el Teorema 2.33 asegura que $\psi(\xi) = E(\psi(X)\,|\,\mathcal{G}) = \psi(\widetilde{\xi})$ c.d.

Por otra parte, sea $\omega \in \Omega$ tal que:

$$\left\|\xi(\omega) - \widetilde{\xi}(\omega) - e_j\right\|_H < \frac{\|e_j\|_H}{3}.$$

Observa que de la linealidad y continuidad de ψ se tiene que:

$$\psi(e_j) - \psi(\xi(\omega) - \widetilde{\xi}(\omega)) \leq |\psi(\xi(\omega) - \widetilde{\xi}(\omega) - e_j)| \leq \|\psi\|_{\mathcal{L}_c(H,\mathbb{R})}\left\|\xi(\omega) - \widetilde{\xi}(\omega) - e_j\right\|_H < \frac{\|e_j\|_H}{3}$$

lo que implica que $\psi(\xi(\omega) - \widetilde{\xi}(\omega)) > \frac{2}{3}\|e_j\|_H$ y, por tanto,

$$P\left\{\left(\omega \in \Omega \,:\, \psi(\xi(\omega) - \widetilde{\xi}(\omega)) > \frac{2}{3}\|e_j\|_H\right)\right\} > 0$$

lo cual contradice que $\psi(\xi) = \psi(\widetilde{\xi})$ c.d.

En consecuencia, $P(\{\omega \in \Omega : \xi(\omega) \neq \tilde{\xi}(\omega)\}) = 0$ lo que significa que $\xi = \widetilde{\xi}$ c.d.

Existencia: Dado que X es Bochner-integrable, consideremos los siguientes dos casos.

Caso 1. X es una función P-simple con descripción $X = \sum_{j=1}^{N} u_j 1_{A_j}$ y definamos $\xi : \Omega \to H$ como:

$$\xi := \sum_{j=1}^{N} u_j P(A_j \,|\, \mathcal{G})$$

en donde $P(A_j \,|\, \mathcal{G}) = E(1_{A_j} \,|\, \mathcal{G}) : \Omega \to \mathbb{R}$ es $\mathcal{G}$-medible pues $1_{A_j} : \Omega \to \mathbb{R}$ es Lebesgue-integrable para todo $j = 1, \ldots, N$ (ver Teorema 5.33). Observa que, la Proposición 5.35 implica que:

$$E\left(\|\xi\|_H\right) \leq \sum_{j=1}^{N} E\left(\|u_j\|_H |P(A_j \,|\, \mathcal{G})|\right) \leq \sum_{j=1}^{N} \|u_j\|_H \,|P(A_j)| \leq E(\|X\|_H) < \infty. \tag{5.6}$$

El teorema de integrabilidad de Lebesgue-Bochner asegura que $\xi : \Omega \to H$ es Bochner-integrable. Además, se tiene que:

$$\mathfrak{B}\int_\Omega \xi(\omega)\, dP(\omega) = \sum_{j=1}^{N} u_j\, E(P(A_j \,|\, \mathcal{G})) = \sum_{j=1}^{N} u_j\, E(A_j) = \sum_{j=1}^{N} u_j\, P(A_j) = \mathfrak{B}\int_\Omega X(\omega)\, dP(\omega)$$

lo que implica que:

$$\mathfrak{B}\int_A \xi(\omega)\, dP(\omega) = \mathfrak{B}\int_A X(\omega)\, dP(\omega) \qquad \forall A \in \mathcal{G}.$$

Caso 2. Consideremos (X_k) una sucesión de funciones P-simples tales que $X_k \to X$ c.d. y tales que $\mathfrak{B}\int_\Omega X(\omega)\, dP(\omega) = \lim_{k\to\infty} \mathfrak{B}\int_\Omega X_k(\omega)\, dP(\Omega)$

Del caso anterior afirmamos que para cada $k \in \mathbb{N}$ existe $\xi_k : \Omega \to H$ tal que ξ_k es Bochner-integrable y

$$\mathfrak{B}\int_\Omega \xi_k(\omega)\, dP(\omega) = \mathfrak{B}\int_\Omega X_k(\omega)\, dP(\omega).$$

De (5.6) sigue que:

$$E(\|\xi_k - \xi_j\|_H) \leq E(\|X_k - X_j\|_H) \qquad \forall, k, j \in \mathbb{N}.$$

En consecuencia, (ξ_k) es de Cauchy en $L^1(\Omega; H)$ por lo que existe $\xi \in L^1(\Omega; H)$ tal que $\xi_k \to \xi$ en $L^1(\Omega; H)$. El teorema de Riesz-Weyl asegura que existe una subsucesión (ξ_{k_j}) de (ξ_k) tal que $\xi_{k_j} \to \xi$ c.d. Dado que

$$E(\|\xi_{k_j} - \xi_{k_\ell}\|_H) \leq E(\|X_{k_j} - X_{k_\ell}\|_H) \qquad \forall j, \ell \in \mathbb{N}$$

tomando el límite cunado $\ell \to \infty$ obtenemos que:

$$E(\|\xi_{k_j} - \xi\|_H) \leq E(\|X_{k_j} - X\|_H) \to 0.$$

En consecuencia, ξ es Bochner-integrable y por lo tanto:

$$\mathfrak{B}\int_\Omega \xi(\omega)\, dP(\omega) = \lim_{j\to\infty} \mathfrak{B}\int_\Omega \xi_{k_j}(\omega)\, dP(\omega).$$

De este modo, para cualquier $A \in \mathcal{G}$ se tiene que:

$$\mathfrak{B}\int_A X(\omega)\, dP(\omega) = \lim_{j\to\infty} \mathfrak{B}\int_A X_{k_j}(\omega)\, dP(\omega) = \lim_{j\to\infty} \mathfrak{B}\int_A \xi_{k_j}(\omega)\, dP(\omega) = \mathfrak{B}\int_A \xi(\omega)\, dP(\omega)$$

lo que prueba el resultado. ■

El teorema anterior motiva la siguiente definición.

Definición 5.37 (Esperanza condicional) *Sea $X : \Omega \to H$ una variable aleatoria Bochner-integrable y sea $\mathcal{G}$ una sub σ-álgebra de $\mathcal{F}$. La **esperanza condicional de** X **dada** $\mathcal{G}$ es la única c.d. variable aleatoria denotada por $E(X \,|\, \mathcal{G}) : \Omega \to H$ que cumples las siguientes tres propiedades:*

(i) $E(X \,|\, \mathcal{G})$ *es* $\mathcal{G}$*-medible.*
(ii) $E(X \,|\, \mathcal{G})$ *es Bochner-integrable*
(iii) *Para cualquier evento* A *en* $\mathcal{G}$

$$\mathfrak{B}\int_A X\, dP = \mathfrak{B}\int_A E(X \,|\, \mathcal{G})\, dP. \tag{5.7}$$

Una construcción usual de la esperanza condicional en la bibliografía es la siguiente: Definamos $J : L^1(\Omega) \to L^1(\Omega)$ como:

$$J(\vartheta) := E(\vartheta \,|\, \mathcal{G}).$$

El Teorema 5.33 asegura que J está bien definida. Luego, de la Proposición 5.34 se sigue que J es una función lineal y positiva. Observa que por la desigualdad de Jensen condicional $|E(\vartheta \,|\, \mathcal{G})| \leq E(|\vartheta| \,|\, \mathcal{G})$ para todo $\vartheta \in L^1(\Omega)$. Por tanto, usando el Teorema 5.33 obtenemos:

$$\|J(\vartheta)\|_{L^1(\Omega)} \leq \|\vartheta\|_{L^1(\Omega)} \qquad \forall\, \vartheta \in L^1(\Omega),$$

lo que implica que J es continua.

La Proposición 2.40 asegura que la función $(J \otimes id_H) : L^1(\Omega) \otimes H \to L^1(\Omega) \otimes H$ se extiende a una única función lineal y continua $\widetilde{J \otimes id_H} : L^1(\Omega; H) \to L^1(\Omega; H)$. Resulta pues que $\widetilde{J \otimes id_H}$ es la esperanza condicional con respecto de la σ-álgebra $\mathcal{G}$ definida para variables aleatorias H-valuadas que son Bochner-integrables.

Terminamos esta sección con algunas propiedades importantes de la esperanza condicional.

Proposición 5.38 *Sean $X, Y : \Omega \to H$ variables aleatorias Bochner-integrables. Entonces:*

(a) $E(E(X \,|\, \mathcal{G})) = E(X)$

(b) $E(aX + bY \,|\, \mathcal{G}) = aE(X \,|\, \mathcal{G}) + bE(Y \,|\, \mathcal{G})$ *c.d.*

(c) Si X es $\mathcal{G}$-medible, entonces $X = E(X \,|\, \mathcal{G})$ c.d.

(d) Si $\mathcal{H} \subset \mathcal{G}$ es una sub σ-álgebra entonces $E((E(X \,|\, \mathcal{G})) \,|\, \mathcal{H}) = E(X \,|\, \mathcal{H})$ c.d.

(e) Si X es independiente de $\mathcal{G}$, es decir, X es independiente de 1_A para cada $A \in \mathcal{G}$, entonces $E(X \,|\, \mathcal{G}) = E(X)$.

Demostración: (a) Es inmediato del inciso *(b)* del Teorema 5.36 para $A = \Omega$ en $\mathcal{G}$.

(b) La demostración es similar a la dada en el inciso *(c)* de la Proposición 5.33 considerando la linealidad de la integral de Bochner (ver Teorema 2.30).

(c) Si X es $\mathcal{G}$-medible, entonces X satisface las propiedades de la Definición 5.36. De la unicidad se sigue que $X = E(X \,|\, \mathcal{G})$ c.d.

(d) Es inmediato que $E((E(X \,|\, \mathcal{G})) \,|\, \mathcal{H})$ es $\mathcal{H}$-medible y Bochner-integrable. Además, para cada $B \in \mathcal{H} \subset \mathcal{G}$ se tiene que:

$$\mathfrak{B}\int_B E((E(X \,|\, \mathcal{G})) \,|\, \mathcal{H})\, dP = \mathfrak{B}\int_B E(X \,|\, \mathcal{G})\, dP = \mathfrak{B}\int_B X\, dP.$$

De la unicidad se sigue que $E((E(X \,|\, \mathcal{G})) \,|\, \mathcal{H}) = E(X \,|\, \mathcal{H})$ c.d.

(e) Sea $A \in \mathcal{G}$. Aplicando la Proposición 5.27 se tiene que:

$$\mathfrak{B}\int_A X\, dP = E(1_A \cdot X) = E(1_A) \cdot E(X) = \mathfrak{B}\int_A E(X)\, dP$$

lo que prueba el resultado. ∎

Proposición 5.39 *Sea $p \in [1, \infty)$ y X una variable aleatoria H-valuada tal que $X \in L^p(\Omega; H)$. Entonces, $\|E(X \,|\, \mathcal{G})\|_H^p \leq E(\|X\|_H^p \,|\, \mathcal{G})$ c.d. En particular, $E(X \,|\, \mathcal{G}) \in L^p(\Omega; H)$ y se tiene que:*

$$\|E(X \,|\, \mathcal{G})\|_{L^p(\Omega;H)} \leq \|X\|_{L^p(\Omega;H)}.$$

Demostración: Dado que $X \in L^p(\Omega; H)$ entonces $\|X\|_H \in L^p(\Omega)$ y el Teorema **??** implica que $\|X\|_H \in L^1(\Omega)$. Por tanto, $X \in L^1(\Omega; H)$ y la esperanza condicional con respecto de la σ-álgebra $\mathcal{G}$ para la variable aleatoria X existe.

Aplicando la desigualdad de Jensen condicional a la función convexa $\phi(x) = |x|^p$ obtenemos que $\|E(X \,|\, \mathcal{G})\|_H^p \leq E(\|X\|_H^p \,|\, \mathcal{G})$ c.d. y esto implica además que $E(X \,|\, \mathcal{G}) \in L^p(\Omega; H)$. Más aún,

$$\int_\Omega \|E(X \,|\, \mathcal{G})\|_H^p\, dP \leq \int_\Omega E(\|X\|_H^p \,|\, \mathcal{G})\, dP = \int_\Omega \|X\|_H^p\, dP$$

lo que prueba el teorema. ■

Con el fin de simplificar algunas pruebas y resultados usaremos la siguiente notación que es clásica en la bibliografía.

Notación 5.40 *Sea $X : \Omega \to H$ una variable aleatoria Bochner-integrable.*

(i) Si $\mathcal{C}$ es una colección de subconjuntos de Ω entonces $\mathcal{S}(\mathcal{C})$ denota a la σ-álgebra generada por $\mathcal{C}$. En este contexto, denotamos $E(X \,|\, \mathcal{C}) := E(X \,|\, \mathcal{S}(\mathcal{C}))$.

(ii) Si $Y : \Omega \to H$ es una variable aleatoria, entonces $\sigma(Y)$ denota a la σ-álgebra $\{Y^{-1}(B) : B \in \mathcal{B}(H)\}$ que es la σ-álgebra más pequeña que hace Borel-medible a Y. Denotamos $E(X \,|\, Y) := E(X \,|\, \sigma(Y))$.

(iii) Para cada $A \in \mathcal{F}$ denotamos por $P(A \,|\, \mathcal{G}) := E(1_A \,|\, \mathcal{G})$ para cualquier $\mathcal{G}$ sub σ-álgebra de $\mathcal{F}$.

5.4. Variables aleatorias Gaussianas

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sean $H = (H, \langle\cdot\rangle_H, \|\cdot\|_H)$ y $K = (K, \langle\cdot\rangle_K, \|\cdot\|_K)$ espacios de Hilbert separables sobre $\mathbb{R}$.

Definición 5.41 *Un **espacio de Hilbert Gaussiano** es un espacio de Hilbert H separable con una medida de probabilidad Gaussiana $\mu = N_{m,Q} : \mathcal{B}(H) \to \mathbb{R}$ dada. Lo denotaremos por (H, μ).*

Definición 5.42 *Una variable aleatoria X H-valuada se llama **variable aleatoria Gaussiana** si su distribución es una medida Gaussiana.*

*Si $X_\sharp P$ es una medida Gaussiana con media m y covarianza Q entonces decimos que X es una **variable aleatoria Gaussiana** $N_{m,Q}$.*

Sea $X : \Omega \to H$ una variable aleatoria. Observemos que, $X_\sharp P$ es una medida Gaussiana $N_{m,Q}$ si y sólo si:

$$\widehat{X_\sharp P}(h) = \exp\{i\langle m, h\rangle_H\} \exp\{-\tfrac{1}{2}\langle Q(h), h\rangle_H\} \qquad \forall\, h \in H.$$

Del teorema de cambio de variable (ver Teorema 5.8) se sigue que:

$$\widehat{X_\sharp P}(h) = \int_H \exp\{i\langle h, u\rangle_H\}\, dX_\sharp P(u) = \int_\Omega \exp\{i\langle h, X(\omega)\rangle_H\}\, dP(\omega) \qquad \forall\, h \in H.$$

Por tanto, $X_\sharp P$ es una medida Gaussiana $N_{m,Q}$ si y sólo si:

$$\widehat{X_\sharp P}(h) = \int_\Omega \exp\{i\langle h, X(\omega)\rangle_H\}\, dP(\omega) = \exp\{i\langle m, h\rangle_H\} \exp\{-\tfrac{1}{2}\langle Q(h), h\rangle_H\} \qquad \forall\, h \in H.$$

Veamos un ejemplo.

Proposición 5.43 *Sean* $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad,* $X : \Omega \to H$ *una variable aleatoria Gaussiana* $N_{m,Q}$ *y sea* $u \in H$ *arbitrario pero fijo. Definamos la función* $R : \Omega \to \mathbb{R}$ *como:*

$$R(\omega) := \langle X(\omega), u \rangle_H. \tag{5.8}$$

Entonces, R *es una variable aleatoria Gaussiana* $N_{\langle m,u \rangle_H \, \langle Q(u),u \rangle_H}$.

Demostración: Es inmediato que R es una variable aleatoria pues $R(\omega) = (\langle \cdot, u \rangle_H \circ X)(\omega)$ para cada $\omega \in \Omega$ y el producto escalar es una función Borel-medible pues es una función continua (ver Proposición 1.6).

Por otra parte, para cada $x \in \mathbb{R}$ se tiene que:

$$\widehat{R_\sharp P}(x) = \int_{\mathbb{R}} \exp\{ixy\} \, dR_\sharp P(y).$$

El teorema de cambio de variable (ver Teorema 5.9) asegura que para cada $x \in \mathbb{R}$

$$\begin{aligned}
\int_{\mathbb{R}} \exp\{ixy\} \, dR_\sharp P(y) &= \int_\Omega \exp\{ixR(\omega)\} \, dP(\omega) \\
&= \int_\Omega \exp\{ix\langle X(\omega), u \rangle_H\} \, dP(\omega) \\
&= \int_\Omega \exp\{i\langle X(\omega), xu \rangle_H\} \, dP(\omega).
\end{aligned}$$

Dado que $X : \Omega \to H$ es una variable aleatoria Gaussiana entonces para cada $x \in \mathbb{R}$ se tiene que:

$$\widehat{X_\sharp P}(xu) = \int_\Omega \exp\{i\langle X(\omega), xu \rangle_H\} \, dP(\omega) = \exp\{i\langle m, xu \rangle_H\} \exp\{-\frac{1}{2}\langle Q(xu), xu \rangle_H\}.$$

En consecuencia,

$$\widehat{R_\sharp P}(x) = \exp\{ix\langle m, u \rangle_H\} \exp\{-\tfrac{1}{2}x^2\langle Q(u), u \rangle_H\} \qquad \forall x \in \mathbb{R}$$

lo que implica que R es una variable aleatoria Gaussiana $N_{\langle m,u \rangle_H \, \langle Q(u),u \rangle_H}$. ■

Proposición 5.44 *Sea* $X_j : \Omega \to \mathbb{R}$ *una variable aleatoria Gaussiana* N_{m_j,λ_j} *para cada* $j = 1, \ldots, N$ *y supongamos que* $X_1, \ldots, X_N$ *son independientes. Entonces,* $X := (X_1, \ldots, X_N) : \Omega \to \mathbb{R}^N$ *es una variable aleatoria Gaussiana con media* $(m_1, \ldots, m_N)$ *y covarianza* $\operatorname{diag}(\lambda_1, \ldots, \lambda_N)$.

Demostración: Dado que X_j es una variable aleatoria Gaussiana N_{m_j,λ_j} para cada $j = 1, \ldots, N$, entonces

$$\widehat{X}_j(x) = \widehat{X_{j\sharp} P}(x) = \exp\{i\, x\, m_j\} \exp\{-\tfrac{1}{2}\, x\, \lambda_j\} \qquad \forall\, x \in \mathbb{R}, \quad \forall\, j = 1, \ldots, N.$$

Como $X_1, \ldots, X_N$ son independientes, aplicando la Proposición 5.25 y lo anterior obtenemos que:

$$\begin{aligned}\widehat{X}(x_1, \ldots, x_N) &= \prod_{j=1}^{N} \widehat{X}_j(x_j) = \prod_{j=1}^{N} \exp\{i\, x_j\, m_j\} \prod_{j=1}^{N} \exp\{-\tfrac{1}{2}\, x_j\, \lambda_j\} \\ &= \exp\left\{ i \sum_{j=1}^{N} x_j\, m_j \right\} \exp\left\{ i \sum_{j=1}^{N} x_j\, \lambda_j \right\}\end{aligned}$$

para cualesquiera $(x_1, \ldots, x_N) \in \mathbb{R}^N$.

Por lo tanto,

$$\begin{aligned}\widehat{X_\sharp P}(x_1, \ldots, x_N) &= \int_\Omega \exp\left\{ i \sum_{j=1}^{N} x_j\, X_j(\omega) \right\} dP(\omega) \\ &= \widehat{X}(x_1, \ldots, x_N) \\ &= \exp\left\{ i \sum_{j=1}^{N} x_j\, m_j \right\} \exp\left\{ i \sum_{j=1}^{N} x_j\, \lambda_j \right\}\end{aligned}$$

para cualesquiera $(x_1, \ldots, x_N) \in \mathbb{R}^N$. Esto concluye la demostración. ■

Definición 5.45 *Sea $L : H \to K$ una función lineal y continua. Consideremos $b \in K$ arbitrario pero fijo y definamos la función $T : H \to K$ como:*

$$T(u) := L(u) + b. \tag{5.9}$$

*La función T se llama una **transformación afín de** H **en** K.*

Notemos que toda función lineal y continua es una transformación afín pues $L(u) = L(u) + 0_K$ para todo $u \in H$. A continuación, demostraremos que toda transformación afín de H en K es una variable aleatoria Gaussiana.

Proposición 5.46 *Sean $H = (H, \mu)$ un espacio de Hilbert Gaussiano, $L : H \to K$ una función lineal y continua y $b \in K$ arbitrario pero fijo. Entonces, la transformación afín $T(u) := L(u) + b$ es una variable aleatoria Gaussiana K-valuada con $N_{L(m)+b, L\circ Q\circ L^*}$.*

Demostración: Es inmediato que T es una variable aleatoria pues es una función continua.

Por otro lado, por el teorema de cambio de variable (ver Teorema 5.9) se tiene que:

$$\begin{aligned}
\widehat{T_\sharp\mu}(k) &= \int_K \exp\{i\langle k, v\rangle_K\}\, dT_\sharp\mu(v) = \int_H \exp\{i\langle k, T(u)\rangle_K\}\, d\mu(u)\\
&= \int_H \exp\{i\langle k, L(u)+b\rangle_K\}\, d\mu(u) = \int_H \exp\{i\langle k, L(u)\rangle_K + i\langle k, b\rangle_K\}\, d\mu(u)\\
&= \exp\{i\langle k, b\rangle_K\} \int_H \exp\{i\langle k, L(u)\rangle_K\}\, d\mu(u)\\
&= \exp\{i\langle k, b\rangle_K\} \int_H \exp\{i\langle L^*(k), u\rangle_H\}\, d\mu(u)\\
&= \exp\{i\langle k, b\rangle_K\}\, \hat{\mu}(L^*(k))\\
&= \exp\{i\langle k, b\rangle_K\}\, \exp\{i\langle m, L^*(k)\rangle_H\} \exp\{-\tfrac{1}{2}\langle Q(L^*(k)), L^*(k)\rangle_H\}\\
&= \exp\{i\langle k, b\rangle_K\}\, \exp\{i\langle L(m), k\rangle_K\} \exp\{-\tfrac{1}{2}\langle L(Q(L^*(k))), k\rangle_K\}\\
&= \exp\{i\langle L(m)+b, k\rangle_K\}\, \exp\{-\tfrac{1}{2}\langle L(Q(L^*(k))), k\rangle_K\} \qquad \forall\, k\in K.
\end{aligned}$$

En consecuencia, $T_\sharp\mu$ es una medida Gaussiana con media $L(m)+b$ y covarianza $L\circ Q\circ L^*$ como afirma el enunciado. ■

Corolario 5.47 *Sea $\mu = N_Q$ una medida Gaussiana en H y sean $u_1,\ldots,u_N \in H$. Definamos la función $A : H \to \mathbb{R}^N$ como sigue:*

$$A(u) := (\langle u, u_1\rangle_H, \ldots, \langle u, u_N\rangle_H)\,. \tag{5.10}$$

Entonces, A es una variable aleatoria Gaussiana $\mathbb{R}^N$-valuada N_{Q_A} con

$$(Q_A)_{ij} = \langle Q(z_i), z_j\rangle_H, \qquad i,j = 1,\ldots,N.$$

Demostración: Es inmediato que A es una función lineal. Observemos que, para cada $u\in H$, se tiene que:

$$\|A(u)\|_2^2 = \sum_{j=1}^N |\langle u, u_j\rangle_H|^2 \le \|u\|_H^2 \,\|\sum_{j=1}^N \|u_j\|_H^2$$

lo que implica que A es una función continua.

Sean $u\in H$ y $x = (x_1,\ldots,x_N)\in\mathbb{R}^N$, entonces:

$$\begin{aligned}
\langle A(u), x\rangle_{\mathbb{R}^N} &= \langle u, u_1\rangle_H\, x_1 + \ldots + \langle u, u_N\rangle_H\, x_N\\
&= \langle u, u_1x_1\rangle_H + \ldots + \langle u, u_Nx_N\rangle_H\\
&= \langle u, u_1x_1 + \ldots + u_Nx_N\rangle_H\,.
\end{aligned}$$

Por tanto, el operador transpuesto $A^* : \mathbb{R}^N \to H$ de A está dado por:

$$A^*(x) = \sum_{j=1}^{N} u_j \, x_N \qquad \forall \, x = (x_1, \ldots, x_N) \in \mathbb{R}^N.$$

En consecuencia,

$$A(Q(A^*(x))) = A\left(\sum_{j=1}^{N} x_j \, Q(u_j)\right) = \sum_{j=1}^{N} x_j \, (\langle Q(z_j), z_1\rangle_H, \ldots, \langle Q(z_j), z_N\rangle_H)$$

y el resultado se sigue de la Proposición 5.46. ∎

El siguiente resultado establece que el límite en media 2 (o media cuadrática) de una sucesión de variables aleatorias Gaussianas es una variable aleatoria Gaussiana.

Notemos primero lo siguiente: Sea (X_k) una sucesión de variables aleatorias H-valuadas tales que $X_k \to X$ en $L^2(\Omega; H)$. Entonces $\|X_k - X\|_H \in L^2(\Omega)$ y, por tanto, $\|X_k - X\|_H \in L^1(\Omega)$ y se cumple que:

$$\int_\Omega \|X_k(\omega) - X(\omega)\|_H \, dP(\omega) \leq \left(\int_\Omega \|X_k(\omega) - X(\omega)\|_H^2 \, dP(\omega)\right)^{1/2} \qquad \forall \, k \in \mathbb{N}.$$

En consecuencia, $X_k \to X$ en $L^1(\Omega; H)$

Teorema 5.48 *Sea (X_k) una sucesión de variables aleatorias H-valuadas tales que X_k es Gaussiana N_{m_k, Q_k} para todo $k \in \mathbb{N}$. Si $X_k \to X$ en $L^2(\Omega; H)$, $X_k \to X$ c.d. (e incluso solamente una subuscesión), m es la media y $Q \in \mathcal{L}_1^+(H)$ es la covarianza de la distribución de X, entonces:*

(i) $\lim_{k\to\infty} \langle m_k, u\rangle_H = \langle m, u\rangle_H$ *para todo* $u \in H$.

(ii) $\lim_{k\to\infty} \langle Q_k(u), u\rangle_H = \langle Q(u), u\rangle_H$ *para todo* $u \in H$.

(iii) X *es una variable aleatoria Gaussiana* $N_{m,Q}$.

Demostración: Fijemos $u \in H$. De la desigualdad de Cauchy-Schwarz (ver Proposición 1.4) se sigue que:

$$|\langle X_k(\omega), u\rangle_H - \langle X(\omega), u\rangle_H| \leq |\langle X_k(\omega) - X(\omega), u\rangle_H| \leq \|X_k(\omega) - X(\omega)\|_H \|u\|_H.$$

En consecuencia, $\langle X_k(\omega), u\rangle_H \to \langle X(\omega), u\rangle_H$ c.d. De la observación anterior se tiene que $\langle X_k(\omega), u\rangle_H \in L^1(\Omega)$ para todo $k \in \mathbb{N}$. Aplicando el teorema de la convergencia dominada de Lebesgue concluimos que:

$$\lim_{k\to\infty} \int_\Omega \langle X_k(\omega), u\rangle_H \, dP(\omega) = \int_\Omega \langle X(\omega), u\rangle_H \, dP(\omega).$$

(i) y *(ii)* Sea $u \in H$. De la Definición 4.6 y el teorema de cambio de variable se tiene que:

$$\begin{aligned}\lim_{k\to\infty}\langle m_k, u\rangle_H = \lim_{k\to\infty}\int_H \langle v, u\rangle_H \, dX_{k,\sharp}P(v) &= \lim_{k\to\infty}\int_\Omega \langle X_k(\omega), u\rangle_H \, dP(\omega)\\ &= \int_\Omega \langle X(\omega), u\rangle_H \, dP(\omega)\\ &= \int_H \langle v, u\rangle_H \, dX_\sharp P(v)\\ &= \langle m, u\rangle_H.\end{aligned}$$

Del mismo modo, la Definición 4.3 y el teorema de cambio de variable implican que:

$$\begin{aligned}\lim_{k\to\infty}\langle Q_k(u), u\rangle_H &= \lim_{k\to\infty}\int_H \langle u, v-m_k\rangle_H\langle u, v-m_k\rangle_H \, dX_{k,\sharp}P(v)\\ &= \lim_{k\to\infty}\int_\Omega \langle u, X_k(\omega)-m_k\rangle_H\langle u, X_k(\omega)-m_k\rangle_H \, dP(\omega)\\ &= \int_\Omega \langle u, X(\omega)-m\rangle_H\langle u, X(\omega)-m\rangle_H \, dP(\omega)\\ &= \int_H \langle u, v-m\rangle_H\langle u, v-m\rangle_H \, dX_\sharp P(v)\\ &= \langle Q(u), u\rangle_H.\end{aligned}$$

(iii) Aplicando los incisos anteriores, el teorema de la convergencia dominada de Lebesgue, la continuidad de $\exp\{ix\} : \mathbb{R} \to \mathbb{C}$ y del producto escalar obtenemos que:

$$\begin{aligned}\int_H \exp\{i\langle h, u\rangle_H\} \, dX_\sharp P(u) &= \int_\Omega \exp\{i\langle h, X(\omega)\rangle_H\} \, dP(\omega)\\ &= \lim_{k\to\infty}\int_\Omega \exp\{i\langle h, X_k(\omega)\rangle_H\} \, dP(\omega)\\ &= \lim_{k\to\infty}\exp\left\{i\langle m_k, h\rangle_H\right\}\exp\left\{-\tfrac{1}{2}\langle Q_k(h), h\rangle_H\right\}\\ &= \exp\left\{i\langle m, h\rangle_H\right\}\exp\left\{-\tfrac{1}{2}\langle Q(h), h\rangle_H\right\}.\end{aligned}$$

En consecuencia, X es Gaussiana y $X_\sharp P = N_{m,Q}$. ■

Un resultado bien conocido es el siguiente.

Proposición 5.49 *Sea $X_j : \Omega \to H$ una variable aleatoria H-valuada Gaussiana N_{m_j,Q_j} para $j = 1, \dots, N$. Si $X_1, \dots, X_N$ son independientes, entonces $S := \sum_{j=1}^N X_j : \Omega \to H$ es una variable H-valuada Gaussiana con media $m := \sum_{j=1}^N m_j$ y covarianza $Q := \sum_{j=1}^N Q_j$.*

Demostración: Para $h \in H$ dada, se sigue del teorema de cambio de variable y la Proposición 5.27 que:

$$\begin{aligned}
\widehat{S}(h) &= \int_H \exp\left\{i\langle h, u\rangle_H\right\} \, dS_\sharp P(u) \\
&= \int_\Omega \exp\left\{i\langle h, S(\omega)\rangle_H\right\} \, dP(\omega) \\
&= \int_\Omega \exp\left\{ i \left\langle h, \sum_{j=1}^N X_j(\omega) \right\rangle_H \right\} dP(\omega) \\
&= \prod_{j=1}^N \int_\Omega \exp\left\{i\langle h, X_j(\omega)\rangle_H\right\} \, dP(\omega) \\
&= \prod_{j=1}^N \int_H \exp\left\{i\langle h, u\rangle_H\right\} \, dX_{j,\sharp}P(u) \\
&= \prod_{j=1}^N \exp\left\{i\langle m_j, h\rangle_H\right\} \exp\left\{-\tfrac{1}{2}\langle Q_j(h), h\rangle_H\right\} \\
&= \exp\left\{ i \left\langle \sum_{j=1}^N m_j, h \right\rangle_H \right\} \exp\left\{ -\tfrac{1}{2} \left\langle \sum_{j=1}^N Q_j(h), h \right\rangle_H \right\} \\
&= \exp\left\{i\langle m, h\rangle_H\right\} \exp\left\{-\tfrac{1}{2}\langle Q(h), h\rangle_H\right\}.
\end{aligned}$$

En consecuencia, $S_\sharp P = N_{m,Q}$ como afirma el enunciado. ■

Sea $X : \Omega \to H$ es una variable aleatoria H-valuada y supongamos X es Gaussiana N_Q. Definamos la variable aleatoria $-X : \Omega \to H$ como $(-X)(\omega) = -X(\omega)$. Observa que, del teorema de cambio de variable se tiene:

$$\begin{aligned}
\widehat{-X}(h) = \int_H \exp\{i\langle h, u\rangle_H\} \, d-X_\sharp P(u) &= \int_\Omega \exp\{i\langle h, -X(\omega)\rangle_H\} \, dP(\omega) \\
&= \int_\Omega \exp\{-i\langle h, X(\omega)\rangle_H\} \, dP(\omega) \\
&= \int_H \exp\{-i\langle h, u\rangle_H\} \, dX_\sharp P(u) \\
&= \int_H \exp\{i\langle h, -u\rangle_H\} \, dX_\sharp P(-u) \\
&= \widehat{X}(h).
\end{aligned}$$

En consecuencia, X es simétrica.

Probamos a continuación que la composición de una variable aleatoria Gaussiana con un función lineal y continua es una variable aleatoria Gaussiana.

Proposición 5.50 *Sea $X : \Omega \to H$ una variable aleatoria H-valuada y $L : H \to K$ una función lineal y continua. Si X es una variable aleatoria Gaussiana $N_{m,Q}$ entonces $L \circ X : \Omega \to K$ es una variable aleatoria K-valuada y Gaussiana $N_{L(m),L\circ Q\circ L^*}$.*

Demostración: Es claro que $L\circ X$ es una variable aleatoria por ser composición de variables aleatorias. Del teorema de cambio de variable se tiene que:

$$\begin{aligned}
\widehat{L \circ X}(z) &= \int_K \exp\{\langle z, v\rangle_K\}\, d(L \circ X)_\sharp P(v) = \int_\Omega \exp\{\langle z, L(X(\omega))\rangle_K\}\, dP(\omega) \\
&= \int_\Omega \exp\{\langle L^*(z), X(\omega)\rangle_H\}\, dP(\omega) \\
&= \int_\Omega \exp\{\langle L^*(z), u\rangle_H\}\, dX_\sharp P(u) \\
&= \exp\left\{\langle m, L^*(z)\rangle_H\right\} \exp\left\{-\tfrac{1}{2}\langle Q(L^*(z)), L^*(z)\rangle_H\right\} \\
&= \exp\left\{\langle L(m), z\rangle_K\right\} \exp\left\{-\tfrac{1}{2}\langle L(Q(L^*(z))), z\rangle_K\right\}
\end{aligned}$$

como afirma el enunciado. ■

5.4.1. Variables aleatorias lineales en un espacio de Hilbert

Sean $H = (H, \langle\cdot\rangle, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable y $\mu = N_Q$ una medida Gaussiana definida en H.

Para cada $w \in H$ consideremos la función $T_w : H \to \mathbb{R}$ dada por:

$$T_w(u) := \langle w, u\rangle.$$

Es claro que $T_w \in \mathcal{L}_c(H, \mathbb{R})$. La Proposición 5.43 asegura que T_w es una variable aleatoria Gaussiana $N_{\langle Q(w),w\rangle}$.

Por lo tanto:

$$\widehat{T_w}(x) = \exp\{-\tfrac{1}{2}\, x^2\, \langle Q(w), w\rangle\} \qquad \forall\, x \in \mathbb{R}. \tag{5.11}$$

Del mismo modo, para $w_1, \ldots, w_N \in H$ dados, el Corolario 5.47 asegura que la función:

$$T_{w_1,\ldots,w_N}(u) := (T_{w_1}(u), \ldots, T_{w_N}(u))$$

es una variable aleatoria Gaussiana $\mathbb{R}^N$-valuada con media $0_{\mathbb{R}^N}$ y covarianza tal que:

$$(Q_{w_1,\ldots,w_N})_{j,k} = \langle Q(w_j), w_k\rangle, \qquad \forall\, j, k = 1, \ldots, N.$$

Es decir,

$$(Q_{w_1,\ldots,w_N}) = \begin{pmatrix} \langle Q(w_1), w_1\rangle & \langle Q(w_1), w_2\rangle & \ldots & \langle Q(w_1), w_N\rangle \\ \langle Q(w_2), w_1\rangle & \langle Q(w_2), w_2\rangle & \ldots & \langle Q(w_2), w_N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle Q(w_N), w_1\rangle & \langle Q(w_N), w_2\rangle & \ldots & \langle Q(w_N), w_N\rangle \end{pmatrix}.$$

Así pues, para cualquier $\bar{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ se tiene que:

$$\begin{aligned}(Q_{w_1,\ldots,w_N})(\bar{x}) &= \begin{pmatrix} \langle Q(w_1), w_1\rangle & \langle Q(w_1), w_2\rangle & \ldots & \langle Q(w_1), w_N\rangle \\ \langle Q(w_2), w_1\rangle & \langle Q(w_2), w_2\rangle & \ldots & \langle Q(w_2), w_N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle Q(w_N), w_1\rangle & \langle Q(w_N), w_2\rangle & \ldots & \langle Q(w_N), w_N\rangle \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \\ &= \begin{pmatrix} \langle Q(w_1), w_1\rangle x_1 + \langle Q(w_1), w_2\rangle x_2 + \ldots + \langle Q(w_1), w_N\rangle x_N \\ \langle Q(w_2), w_1\rangle x_1 + \langle Q(w_2), w_2\rangle x_2 + \ldots + \langle Q(w_2), w_N\rangle x_N \\ \vdots \\ \langle Q(w_N), w_1\rangle x_1 + \langle Q(w_N), w_2\rangle x_2 + \ldots + \langle Q(w_N), w_N\rangle x_N \end{pmatrix}.\end{aligned}$$

Por lo tanto, para cualquier $\bar{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ tenemos que:

$$\begin{aligned}\widehat{T_{w_1,\ldots,w_N}}(\bar{x}) &= \exp\left\{-\tfrac{1}{2}\langle Q_{w_1,\ldots,w_N}(\bar{x}), \bar{x}\rangle_{\mathbb{R}^N}\right\} \\ &= \exp\left\{-\frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle Q(w_j), w_k\rangle x_k\, x_j\right\}.\end{aligned}$$

De las observaciones anteriores podemos concluir el siguiente resultado.

Proposición 5.51 *Sean $w_1, \ldots, w_N \in H$ dados. Entonces, $T_{w_1}, \ldots, T_{w_N}$ son independientes si y sólo si la matriz $(Q_{w_1,\ldots,w_N})$ es diagonal, es decir, si y sólo si*

$$\langle Q(w_j), w_k\rangle = 0, \qquad \forall\, j, k = 1, \ldots, N \text{ con } j \neq k.$$

Demostración: $\Rightarrow)$: Supongamos que $T_{w_1}, \ldots, T_{w_N}$ son independientes.

La Proposición 5.25 asegura que:

$$\begin{aligned}\widehat{T_{w_1,\ldots,w_N}}(\bar{x}) = \prod_{j=1}^{N}\widehat{T_{w_j}}(x_j) &= \prod_{j=1}^{N}\exp\left\{-\tfrac{1}{2}\, x_j^2\, \langle Q(w_j), w_j\rangle\right\} \\ &= \exp\left\{-\frac{1}{2}\sum_{j=1}^{N} x_j^2\, \langle Q(w_j), w_j\rangle\right\}\end{aligned}$$

para cualquier $\bar{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$.

Por otra parte,

$$\widehat{T_{w_1,\ldots,w_N}}(\bar{x}) = \exp\left\{-\frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle Q(w_j), w_k\rangle x_k\, x_j\right\}$$

para cualquier $\bar{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ de donde concluimos que:

$$\langle Q(w_j), w_k\rangle = 0, \qquad \forall\, j,k = 1,\ldots,N \text{ con } j \neq k.$$

$\Leftarrow)$: Supongamos que:

$$\langle Q(w_j), w_k\rangle = 0, \qquad \forall\, j,k = 1,\ldots,N \text{ con } j \neq k.$$

Entonces:

$$\begin{aligned}\widehat{T_{w_1,\ldots,w_N}}(\bar{x}) &= \exp\left\{-\frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle Q(w_j), w_k\rangle x_k\, x_j\right\}\\ &= \exp\left\{-\frac{1}{2}\sum_{j=1}^{N} x_j^2\,\langle Q(w_j), w_j\rangle\right\}\\ &= \prod_{j=1}^{N}\widehat{T_{w_j}}(x_j).\end{aligned}$$

La Proposición 5.25 implica que $T_{w_1}, \ldots, T_{w_N}$ son independientes y esto concluye la demostración. ■

Ejemplo 5.52 *Sea $\mu = N_Q$ una medida Gaussiana en H y sea $\mathcal{B} = \{e_k \,:\, k \in \mathbb{N}\}$ una base de vectores propios de Q. Definamos la sucesión de variables aleatorias $X_k : H \to \mathbb{R}$ como $X_k(u) := \langle u, e_k\rangle$. De la* Proposición 5.51 *se sigue que (X_k) es independiente.*

5.5. El teorema de Fernique

El objetivo de esta sección es extender a espacios de Hilbert de dimensión infinita resultado usual en teoría de la probabilidad clásica sobre el hecho de que una variable aleatoria Gaussiana tiene colas exponenciales. Este resultado fue probado en 1970 por el matemático Xavier Fernique. Para demostrar el teorema de Fernique requerimos de los siguientes lemas.

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot\rangle, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Lema 5.53 *Sean $X, Y : \Omega \to H$ variables aleatorias H-valuadas. Supongamos que X y Y son independientes e idénticamente distribuidas variables aleatorias Gaussianas N_Q. Entonces, las variables aleatorias H-valuadas $U := (X+Y)/\sqrt{2}$ y $V := (X-Y)/\sqrt{2}$ son independientes y tienen la misma distribución que X y Y.*

Demostración: Consideremos $\mu := X_\sharp P = Y_\sharp P$ la distribución de X y Y. Dado que X y Y son independientes, la Proposición 5.10 asegura que:

$$
\begin{aligned}
\widehat{U_\sharp P}(h) &= \int_H \exp\left\{i\langle h, u\rangle\right\}\, dU_\sharp P(u) = \int_\Omega \exp\left\{i\langle h, U(\omega)\rangle\right\}\, dP(\omega) \\
&= \int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, X(\omega)+Y(\omega)\rangle\right\}\, dP(\omega) \\
&= \int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, X(\omega)\rangle\right\} \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, Y(\omega)\rangle\right\}\, dP(\omega) \\
&= \left(\int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, X(\omega)\rangle\right\}\, dP(\omega)\right)\left(\int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, Y(\omega)\rangle\right\}\, dP(\omega)\right) \\
&= \left(\int_H \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, u\rangle\right\}\, d\mu(u)\right)\left(\int_H \exp\left\{\tfrac{i}{\sqrt{2}}\ \langle h, u\rangle\right\}\, d\mu(u)\right) \\
&= \widehat{\mu}\left(\tfrac{h}{\sqrt{2}}\right)\widehat{\mu}\left(\tfrac{h}{\sqrt{2}}\right) \\
&= \exp\left\{-\tfrac{1}{4}\langle Q(h), h\rangle\right\}\exp\left\{-\tfrac{1}{4}\langle Q(h), h\rangle\right\} \\
&= \exp\left\{-\tfrac{1}{2}\langle Q(h), h\rangle\right\}.
\end{aligned}
$$

En consecuencia, U es una variable aleatoria Gaussiana N_Q. De manera análoga se prueba que V es Gaussiana N_Q.

Observemos que:

$$
\begin{aligned}
\widehat{X}\left(\tfrac{u+v}{\sqrt{2}}\right)\widehat{Y}\left(\tfrac{u-v}{\sqrt{2}}\right) &= \exp\left\{-\tfrac{1}{2}\left\langle Q\left(\tfrac{u+v}{\sqrt{2}}\right), \left(\tfrac{u+v}{\sqrt{2}}\right)\right\rangle\right\}\exp\left\{-\tfrac{1}{2}\left\langle Q\left(\tfrac{u-v}{\sqrt{2}}\right), \left(\tfrac{u-v}{\sqrt{2}}\right)\right\rangle\right\} \\
&= \exp\left\{-\tfrac{1}{4}\left\langle Q\left(u\right)+Q(v), u+v\right\rangle\right\}\exp\left\{-\tfrac{1}{4}\left\langle Q\left(u\right)-Q(v), u-v\right\rangle\right\} \\
&= \exp\left\{-\tfrac{2}{4}\left\langle Q\left(u\right), v\right\rangle\right\}\exp\left\{-\tfrac{2}{4}\left\langle Q\left(v\right), v\right\rangle\right\} \\
&= \exp\left\{-\tfrac{1}{2}\left\langle Q\left(u\right), v\right\rangle\right\}\exp\left\{-\tfrac{1}{2}\left\langle Q\left(v\right), v\right\rangle\right\} \\
&= \widehat{U}(u)\widehat{V}(v)
\end{aligned}
$$

para cualesquiera $u, v \in H$.

De este modo, aplicando la Proposición 5.25 obtenemos que:

$$\begin{aligned}
\widehat{(U,V)}(u,v) &= \int_{H^2} \exp\{i\langle u,x\rangle + i\langle v,y\rangle\}\, d(U,V)_\sharp P(x,y) \\
&= \int_\Omega \exp\{i\langle u,U(\omega)\rangle + i\langle v,V(\omega)\rangle\}\, dP(\omega) \\
&= \int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\langle u,X(\omega)+Y(\omega)\rangle + \tfrac{i}{\sqrt{2}}\langle v,X(\omega)-Y(\omega)\rangle\right\} dP(\omega)
\end{aligned}$$

$$\begin{aligned}
&= \int_\Omega \exp\left\{\tfrac{i}{\sqrt{2}}\langle u+v,X(\omega)\rangle + \tfrac{i}{\sqrt{2}}\langle u-v,Y(\omega)\rangle\right\} dP(\omega) \\
&= \int_\Omega \exp\left\{i\left\langle \tfrac{u+v}{\sqrt{2}},X(\omega)\right\rangle + i\left\langle \tfrac{u-v}{\sqrt{2}},Y(\omega)\right\rangle\right\} dP(\omega) \\
&= \int_{H^2} \exp\left\{i\left\langle \tfrac{u+v}{\sqrt{2}},x\right\rangle + i\left\langle \tfrac{u-v}{\sqrt{2}},y\right\rangle\right\} d(X,Y)_\sharp P(x,y) \\
&= \widehat{(X,Y)}\left(\tfrac{u+v}{\sqrt{2}},\tfrac{u-v}{\sqrt{2}}\right) \\
&= \widehat{X}\left(\tfrac{u+v}{\sqrt{2}}\right)\widehat{Y}\left(\tfrac{u-v}{\sqrt{2}}\right) \\
&= \widehat{U}(u)\widehat{V}(v)
\end{aligned}$$

lo que implica que U y V son independientes. ■

Lema 5.54 *Sea $X:\Omega\to H$ una variable aleatoria Gaussiana N_Q. Existe una variable aleatoria H-valuada que es una copia independiente de X.*

Demostración: Consideremos el espacio de probabilidad producto $(\Omega\times\Omega,\mathcal{F}\otimes\mathcal{F},P\otimes P)$ y definamos $\chi(x,y):=X(x)$ y $\widetilde{\chi}(x,y):=X(y)$. Por simplicidad denotaremos por P^2 a la medida producto $P\otimes P$. Veamos que χ y $\widetilde{\chi}$ son independientes.

Si $Y:\Omega\times\Omega\to H^2$ es el vector aleatorio $Y(x,y):=(\chi(x,y),\widetilde{\chi}(x,y))$ entonces para cualesquiera $B_1,B_2\in\mathcal{B}(H)$ se tiene que:

$$\begin{aligned}
Y_\sharp P^2(B_1\times B_2) &= P^2(\{Y\in B_1\times B_2\}) = P^2(\{\chi\in B_1\}\cap\{\widetilde{\chi}\in B_2\}) \\
&= P^2\left([\{X\in B_1\}\times\Omega]\cap[\Omega\times\{X\in B_2\}]\right) \\
&= P^2\left([\{X\in B_1\}\cap\Omega]\times[\Omega\cap\{X\in B_2\}]\right) \\
&= P(\{X\in B_1\})P(\{X\in B_2\}).
\end{aligned}$$

Por otra parte,

$$\begin{aligned}
(\chi_\sharp P^2\otimes\widetilde{\chi}_\sharp P^2)(B_1\times B_2) &= \chi_\sharp P^2(B_1)\widetilde{\chi}_\sharp P^2(B_2) \\
&= P^2(\{\chi\in B_1\})P^2(\{\widetilde{\chi}\in B_2\}) \\
&= P^2(\{X\in B_1\}\times\Omega)P^2(\Omega\times\{X\in B_2\}) \\
&= P(\{X\in B_1\})P(\{X\in B_2\})
\end{aligned}$$

para cualesquiera $B_1, B_2 \in \mathcal{B}(H)$.

El Teorema 4.8 implica que $Y_\sharp P^2 = \chi_\sharp P^2 \otimes \widetilde{\chi}_\sharp P^2$ y en consecuencia χ y $\widetilde{\chi}$ son independientes por el Teorema 5.24. ■

Teorema 5.55 (Fernique) *Sea $X : \Omega \to H$ una variable aleatoria Gaussiana N_Q. Entonces, existe una constante $\gamma > 0$ tal que:*

$$E\left(\exp\{\gamma \, \|X\|_H^2\}\right) < \infty.$$

Demostración: Sea $\widetilde{X}$ una copia independiente de X. Fijemos $t \geq s \geq 0$. Del Lema 5.53 se tiene que:

$$\begin{aligned} P(\{\|X\|_H \leq s\}) \cdot P(\{\|\widetilde{X}\|_H > t\}) &= P\left(\left\{\left\|\frac{X - \widetilde{X}}{\sqrt{2}}\right\|_H \leq s\right\}\right) \cdot P\left(\left\{\left\|\frac{X + \widetilde{X}}{\sqrt{2}}\right\|_H > t\right\}\right) \\ &\leq P\left(\left\{\left|\frac{\|X\|_H - \|\widetilde{X}\|_H}{\sqrt{2}}\right| \leq s\right\} \cap \left\{\frac{\|X\|_H + \|\widetilde{X}\|_H}{\sqrt{2}} > t\right\}\right) \\ &\leq P\left(\left\{\|X\|_H > \frac{t-s}{\sqrt{2}}\right\} \cap \left\{\|\widetilde{X}\|_H > \frac{t-s}{\sqrt{2}}\right\}\right) \\ &= P\left(\left\{\|X\|_H > \frac{t-s}{\sqrt{2}}\right\}\right) \cdot P\left(\left\{\|\widetilde{X}\|_H > \frac{t-s}{\sqrt{2}}\right\}\right), \end{aligned}$$

pues el conjunto

$$\left\{(\xi, \eta) \in \mathbb{R}_+^2 \,:\, |\xi - \eta| \leq \sqrt{2}s \;\text{ y }\; \xi + \eta > t\sqrt{2}\right\}$$

está contenido en el conjunto

$$\left\{(\xi, \eta) \in \mathbb{R}_+^2 \,:\, \xi \geq \frac{t-s}{\sqrt{2}} \;\text{ y }\; \eta > \frac{t-s}{\sqrt{2}}\right\}.$$

Dado que X y $\widetilde{X}$ tienen la misma distribución, entonces:

$$P(\{\|X\|_H \leq s\}) \cdot P(\{\|\widetilde{X}\|_H > t\}) \leq P\left(\left\{\|X\|_H > \frac{t-s}{\sqrt{2}}\right\}\right)^2. \tag{5.12}$$

Elijamos $r > 0$ tal que $P(\{\|X\|_H \leq r\}) \geq \frac{2}{3}$. Definimos $t_0 := r$ y $t_j := r + \sqrt{2}\, t_{j-1}$ para cada $j \in \mathbb{N}$.

Argumentando por inducción puede probarse que $t_j := r\frac{\sqrt{2}^{j+1}-1}{\sqrt{2}-1}$ para cada $j \in \mathbb{N} \cup \{0\}$. En consecuencia, $t_j \leq r\sqrt{2}^{j+4}$ para cada $j \in \mathbb{N} \cup \{0\}$.

Definamos:

$$\alpha_j := \frac{P(\{\|X\|_H > t_j\})}{P(\{\|X\|_H \leq r\})} \qquad \forall\, j \in \mathbb{N} \cup \{0\}. \tag{5.13}$$

Observa que $\alpha_0 \leq \frac{(1-\frac{2}{3})}{\frac{2}{3}} = \frac{1}{2}$. De (5.12) con $s = r$ y $t = t_{j+1}$ obtenemos que:

$$\alpha_{j+1} \leq \left(\frac{P\left(\{\|X\|_H > r + \sqrt{2}t_j\}\right)}{P(\{\|X\|_H \leq r\})} \right)^2 = \alpha_j^2 \qquad \forall\, j \in \mathbb{N} \cup \{0\}.$$

En consecuencia, $\alpha_j \leq \alpha_0^{2^j} \leq 2^{-2^j}$ y se sigue que:

$$P(\{\|X\|_H > t_j\}) = \alpha_j \, P(\{\|X\|_H \leq r\}) \leq 2^{-2^j}$$

De lo anterior, para $\gamma > 0$ se tiene que:

$$\begin{aligned} E\left(\exp\{\gamma \, \|X\|_H^2\}\right) &\leq P(\{\|X\|_H \leq t_0\}) \cdot \exp\{\gamma t_0^2\} \\ &\quad + \sum_{j=0}^{\infty} P(\{t_j < \|X\|_H \leq t_{j+1}\}) \cdot \exp\{\gamma t_{j+1}^2\} \\ &\leq \exp\{\gamma r^2\} + \sum_{j=0}^{\infty} 2^{-2^j} \exp\{\gamma r^2 2^{j+5}\} \\ &= \exp\{\gamma r^2\} + \sum_{j=0}^{\infty} \exp\{2^j(-\log(2) + \gamma r^2 2^5)\} \end{aligned}$$

en donde esta suma converge si $\gamma > 0$ es suficientemente pequeño. ■

Los detalles sobre una aproximación sobre la constante $\gamma > 0$ elegida en el teorema de Fernique pueden consultarse con más detalle en [21].

5.6. Solución al problema inicial

Iniciamos este capítulo con la siguiente pregunta: Dados $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}([0,1])$ y $P = \lambda$ la medida de Lebesgue en $(0,1)$ y definiendo $X : \Omega \to L^2(0,1)$ como $X(\omega) = 1_{(0,\omega)}$, *¿cuál es la probabilidad de que la norma-L^2 de X sea menor que 1?*

El primer punto a resolver es que X sea una variable aleatoria con valores en el espacio de Hilbert separable $L^2(0,1)$ pero esto lo afirma el Ejemplo 2.24 de modo que la pregunta inicial tiene sentido.

La pregunta planteada está considerando que tenemos algunas nociones básicas de teoría de la probabilidad clásica y para resolverla procedemos de la siguiente forma:

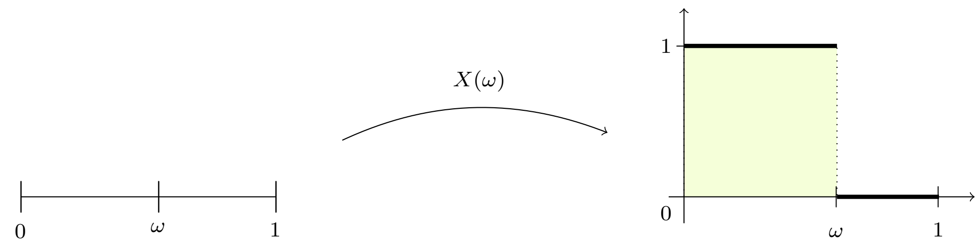


Figura 5.3: Representación de la variable aleatoria X

Denotaremos por 0 a la clase de la función constante con valor 0 en $(0,1)$. Así pues, consideremos la bola abierta unitaria en $L^2(0,1)$

$$B_{L^2}(0,1) := \left\{u \in L^2(0,1) \,:\, \|u\|_{L^2(0,1)} < 1\right\}$$

que es claramente un subconjunto en la σ-álgebra $\mathcal{B}(L^2(0,1))$.

Por tanto, la pregunta inicial es equivalente a calcular la probabilidad de que X esté en la bola abierta unitaria en $L^2(0,1)$ y lo haremos a través de su distribución. Es decir, calcular:

$$P(\{X \in B_{L^2}(0,1)\}).$$

Observemos que, para $\omega \in (0,1)$ se tiene que $\int_{(0,1)} |1_{(0,\omega)}|^2 \, d\lambda = \lambda(0,\omega) = \omega < 1$ lo cual implica que $\{X \in B_{L^2}(0,1)\} = \{\omega \in (0,1) \,:\, 1_{(0,\omega)} \in B_{L^2}(0,1)\} = (0,1)$.

En consecuencia, la probabilidad de que la norma-L^2 de X sea menor que 1 es 1.

En realidad podemos generalizar un poco la pregunta considerando $B_{L^2(0,1)}(0,\varepsilon)$ para cualquier $\varepsilon \in (0,1)$. Deseamos ahora calcular $P(\{X \in B_{L^2}(0,\varepsilon)\})$ para $\varepsilon \in (0,1)$ dada. Observemos que esta es, de hecho, la pregunta con la que empezamos este trabajo.

Si $\omega \in (0,1)$ es tal que $\|X(\omega)\|_{L^2(0,1)} < \varepsilon$ entonces $\int_{(0,1)} 1_{(0,\omega)} \, d\lambda < \varepsilon^2$ y se sigue directamente que $\omega < \varepsilon^2$. Es decir, $\omega \in (0,\varepsilon^2)$.

Inversamente: Sea $\omega \in (0,\varepsilon^2)$. Entonces $(\int_{(0,1)} 1_{(0,\omega)} \, d\lambda)^{1/2} = (\lambda(0,\omega))^{1/2} = \omega^{1/2} < \varepsilon$ por lo que $X(\omega) \in B_{L^2}(0,\varepsilon)$.

Por lo tanto,

$$P(\{X \in B_{L^2}(0,\varepsilon)\}) = P((0,\varepsilon^2)) = \varepsilon^2.$$

Estas observaciones nos serán de utilidad en el próximo capítulo.

Capítulo 6

Procesos estocásticos en espacios de Hilbert

Después de que en el capítulo anterior extendimos el concepto de variable aleatoria para funciones que toman valores en un espacio de Hilbert separable, el siguiente paso, y guardando una estrecha relación con los cursos básicos de probabilidad, es estudiar colecciones de variables aleatorias que estén parametrizadas por un conjunto $\mathcal{I}$. Estas colecciones reciben el nombre de procesos estocásticos y tienen una enorme relevancia en análisis e importantes aplicaciones en la física, la biología, las finanzas y otras disciplinas pues suelen caracterizar la forma en como evoluciona un fenómeno aleatorio a través del tiempo.

En este capítulo extendemos el concepto de *proceso estocástico* cuando tenemos una colección de variables aleatorias valuadas en un espacio de Hilbert separable H. Un proceso estocástico puede entenderse como una función de dos variables $X : \Omega \times \mathcal{I} \to H$ lo nos permitirá dar distintas definiciones que lo caractericen y, sobre todo, a sus trayectorias.

El teorema de la prueba de Kolmogorov garantiza la existencia de un proceso estocástico H-valuado cuyas trayectorias son funciones α-Hölder continuas y, por tanto, continuas.

En general, las variables aleatorias que conforman a un proceso estocástico suelen estar relacionadas de distintas formas la cuales les brindan ciertas características. Una de estas características la desarrollaremos en lo que se conoce como *procesos Gaussianos*.

Finalizamos con una muy breve introducción de la teoría de *martingalas* en espacios de Hilbert pues esta permite, por ejemplo, construir el concepto de integral estocástica de Itô en espacios de Hilbert (ver [54; 62]).

Este capítulo se desarrolló siguiendo [21], [23], [38], [53], [61], [62] y [87].

6.1. Procesos estocásticos

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle \cdot, \cdot \rangle_H, \| \cdot \|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 6.1 *Un **proceso estocástico H-valuado** es una colección de variables aleatorias H-valuadas $\{X_t : t \in \mathcal{I}\}$ parametrizada por un conjunto no vacío $\mathcal{I}$.*

Lo denotaremos usualmente por $X = \{X_t : t \in \mathcal{I}\}$.

En general, podemos considerar a un proceso estocástico H-valuado como una función de dos variables

$$X : \mathcal{I} \times \Omega \to H$$

tal que a la pareja (t, ω) le asocial el valor en H dado por $X_t(\omega)$. Para cada $\omega \in \Omega$ fijo, la función $t \mapsto X(\cdot, \omega)$ es llamada una *trayectoria* o realización del proceso.

Decimos que un proceso estocástico Y H-valuado es una *versión* de X si

$$P(\{\omega \in \Omega : Y_t(\omega) \neq X_t(\omega)\}) = 0 \qquad \forall\, t \in \mathcal{I}.$$

En lo que resta del capítulo consideraremos a $\mathcal{I}$, salvo que se indique lo contrario, un intervalo en $\mathbb{R}$ de la forma $[0, T]$ para T un número real positivo arbitrario pero fijo o el intervalo de la forma $[0, \infty)$. De este modo, consideraremos al espacio medible $(\mathcal{I}, \mathcal{B}(\mathcal{I}))$ y a menos que se indique lo contrario, a λ la medida de Lebesgue en $\mathcal{I}$.

Observa que si $X : \mathcal{I} \otimes \Omega \to H$ es una función $\mathcal{B}(\mathcal{I}) \otimes \mathcal{F}$-medible entonces para cada $t \in \mathcal{I}$ fijo, la función $X_i : \Omega \to H$ dada por $X_t(\omega) := X(t, \omega)$ es $\mathcal{F}$-medible y, por tanto, X es proceso estocástico H-valuado. Sin embargo, no todo proceso estocástico H-valuado es una función $\mathcal{B}(\mathcal{I}) \otimes \mathcal{F}$-medible.

El objetivo de esta sección, es demostrar que un proceso estocástico H-valuado $\{X_t : t \in [0, T]\}$ que cumple una cierta propiedad tiene una versión que es una función $\mathcal{B}(\mathcal{I}) \otimes \mathcal{F}$-medible. Para ello, requerimos de las siguientes definiciones y resultados.

Definición 6.2 *Sea $X = \{X_t : t \in [0, T]\}$ un proceso estocástico H-valuado. Decimos que X es **estocásticamente continuo en el punto** $t_0 \in [0, T]$ si, para cada $\varepsilon > 0$ y para cada $\delta > 0$, existe $\eta > 0$ tal que:*

$$P\left(\{\|X_t - X_{t_0}\|_H \geq \varepsilon\}\right) \leq \delta \qquad \forall\, t \in [t_0 - \eta, t_0 + \eta] \cap [0, T].$$

*Decimos que X es **estocásticamente continuo** si lo es en todo punto de $[0, T]$.*

Definición 6.3 *Sea $X = \{X_t : t \in [0, T]\}$ un proceso estocástico H-valuado. Decimos que X es **estocásticamente uniformemente continuo en** $[0, T]$ si, para cada $\varepsilon > 0$ y para cada $\delta > 0$, existe $\eta > 0$ tal que:*

$$P\left(\{\|X_t - X_s\|_H \geq \varepsilon\}\right) \leq \delta \qquad \forall\, t, s \in [0, T] \text{ tales que } |t - s| < \eta.$$

Veamos un ejemplo.

Ejemplo 6.4 *Sea* $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}([0,1])$ *y* $P = \lambda$ *la medida de Lebesgue en* $(0,1)$. *Definamos* $X : [0,1] \times (0,1) \to L^2(0,1)$ *como:*

$$X(t,\omega) := \begin{cases} X_t(\omega) = 1_{\varnothing}(\omega) & si\ t = 0, \\ X_t(\omega) = Y(\omega) & si\ t \in (0,1), \\ X_t(\omega) = 1_{(0,1)}(\omega) & si\ t = 1, \end{cases}$$

en donde $Y(\omega) = 1_{(0,\omega)}$. *El* Ejemplo 2.24 *asegura que* X_t *es una variable aleatoria* $L^2(0,1)$-*valuada para toda* $t \in [0,1]$. *En consecuencia,* X *es un proceso estocástico* $L^2(0,1)$-*valuado.*

Podemos entender a este proceso de la siguiente forma:

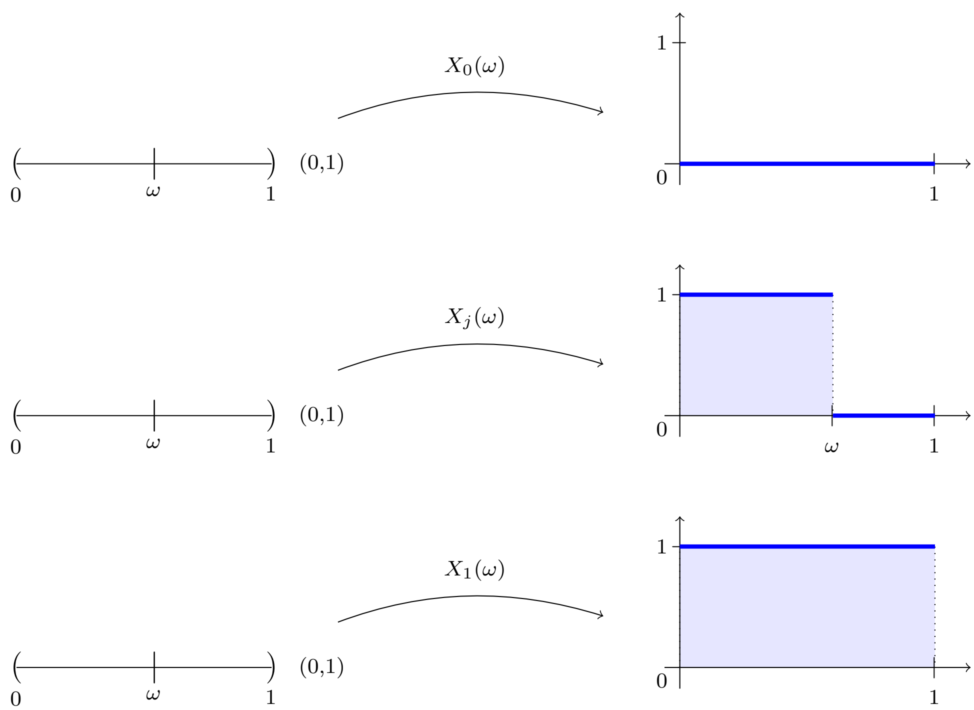


Figura 6.1: Representación del proceso X

Probaremos que X *no es estocásticamente continuo en el punto* $0 \in [0,1]$. *Sea* $\varepsilon_0 = \delta_0 := \frac{1}{2}$. *Afirmamos que para cualquier* $\eta \in (0,1)$ *existe* $t_\eta \in [-\eta,\eta] \cap [0,1]$ *tal que* $P(\{\|X_{t_\eta} - X_0\|_{L^2(0,1)} \geq \varepsilon_0\}) > \delta_0$. *En efecto,* $t_\eta := \frac{\eta}{2}$ *tiene estas propiedades pues:*

$$\begin{aligned} P(\{\|X_{t_\eta} - X_0\|_{L^2(0,1)} \geq \varepsilon_0\}) &= P(\{\|X_{t_\eta}\|_{L^2(0,1)} \geq \varepsilon_0\}) \\ &= 1 - P(\{\|X_{t_\eta}\|_{L^2(0,1)} < \varepsilon_0\}) \\ &= 1 - \varepsilon_0^2 > \delta_0. \end{aligned}$$

En consecuencia, X *no es un proceso estocásticamente continuo en* $[0,1]$.

El siguiente resultado establece una relación entre las definiciones anteriores para un proceso estocástico H-valuado definido en $[0,T]$.

Proposición 6.5 *Un proceso* $X = \{X_t : \Omega \to H\}$ *estocásticamente continuo en* $[0,T]$ *es estocásticamente uniformemente continuo en* $[0,T]$.

Demostración: Sean $\varepsilon, \delta > 0$. Dado que X es estocásticamente continuo en $[0,T]$ entonces, para cada $r \in [0,T]$ existe un intervalo abierto no vacío $\mathcal{I}(r)$ con centro en r tal que:

$$P\left(\left\{\|X_s - X_r\|_H \geq \frac{\varepsilon}{2}\right\}\right) \leq \frac{\delta}{2} \qquad \forall\, s \in \mathcal{I}(r) \cap [0,T].$$

Observa que, para cualesquiera $s,t \in \mathcal{I}(r) \cap [0,T]$, la desigualdad del triángulo de H implica que:

$$\left\{\|X_s - X_r\|_H < \frac{\varepsilon}{2}\right\} \cap \left\{\|X_r - X_t\|_H < \frac{\varepsilon}{2}\right\} \subset \left\{\|X_s - X_t\|_H < \varepsilon\right\}.$$

En consecuencia,

$$P\left(\{\|X_s - X_t\|_H \geq \varepsilon\}\right) \leq P\left(\left\{\|X_s - X_r\|_H \geq \frac{\varepsilon}{2}\right\}\right) + P\left(\left\{\|X_r - X_t\|_H \geq \frac{\varepsilon}{2}\right\}\right) \leq \delta$$

para cualesquiera $s,t \in \mathcal{I}(r) \cap [0,T]$.

Dado que $[0,T]$ es compacto en $\mathbb{R}$, existe una cantidad finita $r_1, \ldots, r_m \in [0,T]$ tal que $[0,T] \subset \mathcal{I}(r_1) \cup \cdots \cup \mathcal{I}(r_m)$ y de lo anterior se deduce fácilmente el resultado. ■

Teorema 6.6 *Sea* $X = \{X_t : \Omega \to H \,:\, t \in [0,T]\}$ *un proceso estocásticamente continuo en* $[0,T]$*. Entonces existe* $\widetilde{X} : [0,T] \times \Omega$ *una función* $\mathcal{B}([0,T]) \otimes \mathcal{F}$*-medible tal que* $\widetilde{X}$ *es una versión de* X*.*

Demostración: Fijemos $k \in \mathbb{N}$. Dado que X es estocásticamente uniformemente continuo en $[0,T]$ por la Proposición 6.5, entonces es posible construir una partición $0 = t_{k,0} < t_{k,1} < \cdots < t_{k,j_k} = T$ tal que para todo $t \in (t_{k,\ell}, t_{k,\ell+1}]$ se tiene que:

$$P\left(\left\{\|X_{t_{k,\ell}} - X_t\|_H \geq \frac{1}{2^k}\right\}\right) \leq \frac{1}{2^k} \qquad \forall\, \ell = 0, \ldots, j_k - 1. \tag{6.1}$$

Definamos $\widetilde{X}_k : [0,T] \times \Omega \to H$ como:

$$\widetilde{X}_k(t,\omega) := \begin{cases} X_0(\omega) & \text{si } t = 0, \\ X_{t_{k,\ell}}(\omega) & \text{si } t \in (t_{k,\ell}, t_{k,\ell+1}] \ \textit{y} \ \ell \leq j_k - 1. \end{cases}$$

Esta función es claramente $\mathcal{B}([0,T])\otimes\mathcal{F}$-medible pues puede representarse como:

$$\widetilde{X}_k(t,\omega)=X_0(\omega)\cdot 1_{\{0\}\times\Omega}(t,\omega)+\sum_{\ell=0}^{j_k-1}X_{t_{k,\ell}}(\omega)\cdot 1_{(t_{k,\ell},t_{k,\ell+1}]\times\Omega}(t,\omega).$$

Consideremos $A:=\{(t,\omega)\in[0,T]\times\Omega\,:\,(\widetilde{X}_k(t,\omega))$ converge en $H\}$. Observa que $(\widetilde{X}_k(t,\omega))$ converge en H si y sólo si $(\widetilde{X}_k(t,\omega))$ es de Cauchy en H si y sólo si, para cada $n\in\mathbb{N}$, existe $k\in\mathbb{N}$ tal que:

$$\|\widetilde{X}_k(t,\omega)-\widetilde{X}_{k+p}(t,\omega)\|_H<\tfrac{1}{n}\qquad\forall\,p\geq 1.$$

Es decir, $A=\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty}\bigcap_{p=1}^{\infty}\left\{(t,\omega)\in[0,T]\times\Omega\,:\,\|\widetilde{X}_k(t,\omega)-\widetilde{X}_{k+p}(t,\omega)\|_H<\frac{1}{n}\right\}$ y esto implica que $A\in\mathcal{B}([0,T])\otimes\mathcal{F}$.

Definimos, por tanto, $\widetilde{X}:[0,T]\times\Omega\to H$ como sigue:

$$\widetilde{X}(t,\omega):=\begin{cases}\lim_{k\to\infty}\widetilde{X}_k(t,\omega) & \text{si } (t,\omega)\in A,\\ 0 & \text{si } (t,\omega)\notin A,\end{cases}$$

la cual es claramente una función $\mathcal{B}([0,T])\otimes\mathcal{F}$-medible.

Ahora, fijemos $t\in[0,T]$ y definamos $E_k:=\left\{\omega\in\Omega\,:\,\|\widetilde{X}_k(t,\omega)-X(t,\omega)\|_H\geq\frac{1}{2^k}\right\}\in\mathcal{F}$ para cada $k\in\mathbb{N}$. De (6.1) se sigue que $P(E_k)\leq\frac{1}{2^k}$ para cada $k\in\mathbb{N}$ y, por tanto, $\sum_{k=1}^{\infty}P(E_k)\leq 1$. El Lema de Borel-Cantelli afirma que:

$$P\left(\limsup_{k\to\infty}E_k\right)=0.$$

Definimos el conjunto medible F_j como:

$$F_j:=\bigcup_{k=j}^{\infty}E_k.$$

Notemos que:

$$\limsup_{j\to\infty}E_j=\bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty}E_k=\bigcap_{j=1}^{\infty}F_j.$$

Sea $\omega\in\Omega\setminus F_j$ arbitraria, entonces $\omega\in\Omega\setminus E_k$ para todo $k\geq j$ y, por tanto,

$$\|\widetilde{X}_k(t,\omega)-X(t,\omega)\|_H<\frac{1}{2^k}\qquad\forall\,k\geq j.$$

Esto prueba que $\widetilde{X}_k(t,\omega)-X(t,\omega)$ en $\Omega\setminus F_j$ y como $\Omega\setminus\limsup_{j\to\infty}E_j=\bigcup_{j=1}^{\infty}(\Omega\setminus F_j)$ entonces $\widetilde{X}_k(t,\omega)-X(t,\omega)$ en $\Omega\setminus\limsup_{j\to\infty}E_j$. En consecuencia $\widetilde{X}(t,\omega)=X(t,\omega)$ c.d. y esto concluye la demostración. ■

Finalicemos esta sección con algunas definiciones importantes.

Definición 6.7 *Sea $X = \{X_t : t \in \mathcal{I}\}$ un proceso estocástico H-valuado. Decimos que X es* ***continuo en media cuadrática en el punto*** *$t_0 \in \mathcal{I}$ si,*

$$\lim_{t \to t_0} E\left(\|X_t - X_{t_0}\|_H^2\right) = 0.$$

Es decir, si dada $\varepsilon > 0$, existe $\delta > 0$ tal que:

$$\|X_t - X_{t_0}\|_{L^2(\Omega;H)} < \varepsilon \qquad \forall\, t \in \mathcal{I} \text{ con } |t - t_0| < \delta.$$

Decimos que X es ***continuo en media cuadrática*** *si lo es en todo punto de $\mathcal{I}$.*

Definición 6.8 *Sea $X = \{X_t : t \in \mathcal{I}\}$ un proceso estocástico H-valuado. Decimos que X es* ***continuo con probabilidad 1*** *(o simplemente* ***continuo****) si, la función $X_\omega : \mathcal{I} \to H$ definida por $X_\omega(t) := X(t, \omega)$ es continua p.c.t. $\omega \in \Omega$. Es decir, sus trayectorias son continuas c.d.*

Ejemplo 6.9 *Sea $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}([0,1])$ y $P = \lambda$ la medida de Lebesgue en $(0,1)$ y consideremos el proceso estocástico $X : [0,1] \times (0,1) \to L^2(0,1)$ dado por:*

$$X(t,\omega) := \begin{cases} X_t(\omega) = 1_{\varnothing}(\omega) & \text{si } t = 0, \\ X_t(\omega) = Y(\omega) & \text{si } t \in (0,1), \\ X_t(\omega) = 1_{(0,1)}(\omega) & \text{si } t = 1, \end{cases}$$

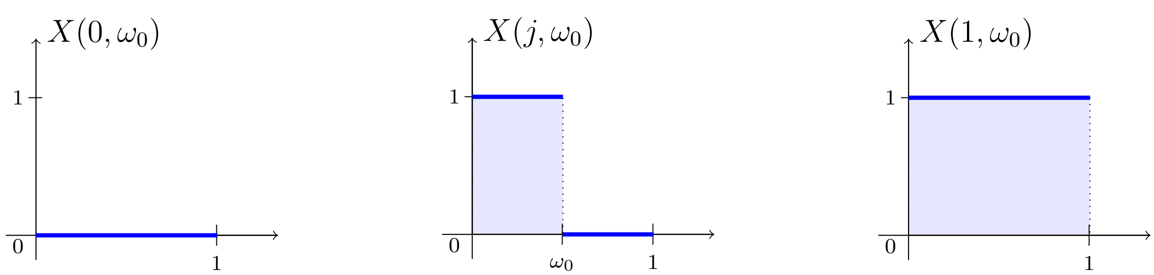
en donde $Y(\omega) = 1_{(0,\omega)}$.

Entonces X no es continuo.

Demostración: Sea $\omega \in \Omega$ arbitraria pero fija y consideremos la función $X_\omega : [0,1] \to L^2(0,1)$ definida por $X_\omega(t) := X(t,\omega)$. Afirmamos que X_ω no es continua en 0 (de hecho no lo es también en 1). Sea $\varepsilon := \omega$. Afirmamos que para cualquier $\delta \in (0,1)$ existe $t_\delta \in [0,1]$ tal que $|t_\delta| < \delta$ y $\|X_\omega(t_\delta)\|_{L^2(0,1)} > \varepsilon$. En efecto, el punto $t_\delta := \frac{\delta}{2}$ tiene estas propiedades.

Por tanto, $X_\omega[0,1] \to L^2(0,1)$ no es continua para cada $\omega \in (0,1)$. ■

Una representación de las trayectorias de X puede ser la siguiente en donde $j \in (0,1)$:



6.2. Prueba de Kolmogorov

En esta sección demostramos un resultado bastante importante e interesante que garantiza, bajo ciertas condiciones, la existencia de versiones continuas para un proceso estocástico.

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 6.10 *Sea $X = (X, d)$ un espacio métrico y $V = (V, \|\cdot\|_V)$ un espacio normado real. Sea $0 < \alpha \leq 1$ una constante. Una función $\phi : X \to V$ es **α-Hölder continua** si existe una constante $c > 0$ tal que*

$$\|\phi(x) - \phi(y)\|_V \leq c\,(d(x,y))^\alpha \qquad \forall\, x, y \in X. \tag{6.2}$$

Definición 6.11 *Sea $X = (X, d)$ un espacio métrico y $V = (V, \|\cdot\|_V)$ un espacio normado real. Si $\phi : X \to V$ es una función α-Hölder continua entonces ϕ es uniformemente continua y, por tanto, continua.*

Demostración: Sea $c > 0$ una constante que cumple (6.2) para ϕ. Entonces, dada $\varepsilon > 0$, para $\delta := (\frac{\varepsilon}{c})^{1/\alpha} > 0$ se cumple que:

$$\|\phi(x) - \phi(y)\|_V < \varepsilon \;\text{ si }\; d(x,y) < \delta \qquad \forall\, x, y \in X.$$

Por tanto, ϕ es uniformemente continua. ■

Definición 6.12 *Sea $X = \{X_t : t \in \mathcal{I}\}$ un proceso estocástico H-valuado y $0 < \alpha leq 1$ una constante. Decimos que X es **α-Hölder continuo con probabilidad 1** (o simplemente **α-Hölder continuo**) si, la función $X_\omega : \mathcal{I} \to H$ definida por $X_\omega(t) := X(t, \omega)$ es α-Hölder continua p.c.t. $\omega \in \Omega$. Es decir, sus trayectorias son α-Hölder continuas c.d.*

Teorema 6.13 (Prueba de Kolmogorov) *Sea $X = \{X_t : t \in [0, T]\}$ un proceso estocástico H-valuado tal que, existen constantes $c > 0$, $1 \geq \varepsilon > 0$ y $\delta > 1$ tales que para cada $t, s \in [0, T]$*

$$E\left(\|X_t - X_s\|_H^\delta\right) \leq c\,|t - s|^{\varepsilon+1}. \tag{6.3}$$

Entonces, existe $\widetilde{X}$ una versión de X tal que $\widetilde{X}$ es α-Hölder continuo para cada $\alpha \in (0, \frac{\varepsilon}{\delta})$. En particular, X tiene una versión continua.

Demostración: Es posible suponer que $T = 1$. Sea $\beta > 0$. Definamos $A := \{\omega \in \Omega : \|X_t(\omega) - X_s(\omega)\|_H \geq \beta\} \in \mathcal{F}$ para cualesquiera $s, t \in [0, T]$. Dado que $\delta > 1$ entonces:

$$\beta^\delta \cdot 1_A \leq \|X_t - X_s\|_H^\delta \cdot 1_A \leq \|X_t - X_s\|_H^\delta \cdot 1_\Omega$$

y de (6.3) sigue que:

$$E(\beta^\delta \cdot 1_A) \leq E\left(\|X_t - X_s\|_H^\delta \cdot 1_A\right) \leq E\left(\|X_t - X_s\|_H^\delta\right) \leq c\,|t-s|^{\varepsilon+1}.$$

En consecuencia:

$$P\left(\{\|X_t - X_s\|_H \geq \beta\}\right) \leq \frac{1}{\beta^\delta}\, c\,|t-s|^{\varepsilon+1} \tag{6.4}$$

para cualesquiera $s, t \in [0, T]$. Esto prueba que X es estocásticamente uniformemente continuo en $[0, T]$.

Sea $0 < \gamma < \frac{\varepsilon}{\delta}$. Fijemos $k \in \mathbb{N}$. De (6.4) se tiene que para cada $j \in \{1, 2, \ldots, 2^k\}$

$$P\left(\left\{\|X_{j2^{-k}} - X_{(j-1)2^{-k}}\|_H \geq 2^{-\gamma\, k}\right\}\right) \leq c\, 2^{-k(1+\varepsilon-\gamma\,\delta)}.$$

Por tanto,

$$\begin{aligned} P\left(\left\{\max_{1\leq j\leq 2^k} \left\{\|X_{j2^{-k}} - X_{(j-1)2^{-k}}\|_H \geq 2^{-\gamma\, k}\right\}\right\}\right) &\leq \sum_{j=1}^{2^k} P\left(\left\{\|X_{j2^{-k}} - X_{(j-1)2^{-k}}\|_H \geq c\, 2^{-\gamma\, k}\right\}\right) \\ &\leq c\, 2^{-k(\varepsilon-\gamma\,\delta)}. \end{aligned}$$

Dado que $\sum_{k=1}^{\infty} 2^{-k(\varepsilon-\gamma\,\delta)} < \infty$, por el lema de Borel-Cantelli, existe un subconjunto $\widetilde{\Omega} \in \mathcal{F}$ y una variable aleatoria $\widetilde{N} : \Omega \to \mathbb{N}$ tal que para cada $\omega \in \widetilde{\Omega}$ y $k \geq \widetilde{N}(\omega)$

$$\max_{1\leq j\leq 2^k} \left\{\|X_{j2^{-k}} - X_{(j-1)2^{-k}}\|_H \geq 2^{-\gamma\, k}\right\} < 2^{-\gamma\, k}. \tag{6.5}$$

Definamos $D_k := \left\{0, \frac{1}{2^k}, \ldots, \frac{2^{k-1}}{2^k}\right\}$ y $D := \cup_{k=1}^{\infty} D_k$. Observa que, para cualquier $x \in D_k$, existe una única representación de él de la forma:

$$x = \sum_{n=1}^{\ell} \varepsilon_n\, 2^{-n}, \quad \text{donde } \varepsilon_n = 0 \text{ o } \varepsilon_n = 1 \quad k < \ell. \tag{6.6}$$

Así pues, sean $\ell > k \geq \widetilde{N}(\omega)$ y $t = j2^{-\ell}$, $s = i2^{-\ell}$, $0 \leq i \leq j < 2^\ell$ tales que $t - s < 2^{-k}$. De (6.6) se sigue que:

$$t - s = j2^{-\ell} - i2^{-\ell} = \sum_{n=k+1}^{\ell} \varepsilon_n\, 2^{-n}, \quad \text{con } \varepsilon_n = 0 \text{ o } \varepsilon_n = 1 \quad n = k+1, \ldots, \ell. \tag{6.7}$$

Por simplicidad, escribimos $X(t) := X_t$. En consecuencia, de (6.7) y (6.5) se tiene que:

$$\begin{aligned}\|X(t) - X(s)\|_H &\leq \left\|X\left(\tfrac{i}{2^\ell}\right) - X\left(\tfrac{i}{2^\ell} + \tfrac{\varepsilon_1}{2^{k+1}}\right)\right\|_H \\ &+ \left\|X\left(\tfrac{i}{2^\ell} + \tfrac{\varepsilon_1}{2^{k+1}}\right) - X\left(\tfrac{i}{2^\ell} + \tfrac{\varepsilon_1}{2^{k+1}} + \tfrac{\varepsilon_2}{2^{k+2}}\right)\right\|_H \\ &+ \cdots + \\ &+ \left\|X\left(\tfrac{i}{2^\ell} + \tfrac{\varepsilon_1}{2^{k+1}} + \cdots + \tfrac{\varepsilon_{\ell-1}}{2^{\ell-1}}\right) - X\left(\tfrac{i}{2^\ell} + \tfrac{\varepsilon_1}{2^{k+1}} + \cdots + \tfrac{\varepsilon_\ell}{2^\ell}\right)\right\|_H \\ &\leq 2^{-\gamma(k+1)} + \cdots + 2^{-\gamma\ell} \\ &\leq 2^{-\gamma(k+1)}(1 - 2^{-\gamma})^{-1}.\end{aligned}$$

Eligiendo $k \in \mathbb{N}$ tal que $2^{-k-1} \leq t - s \leq 2^{-k}$ obtenemos:

$$\|X(t) - X(s)\|_H \leq \frac{(t-s)^\gamma}{1 - 2^{-\gamma}}. \tag{6.8}$$

Por tanto, $X(t,\omega)$ con $t \in D$ y $\omega \in \widetilde{\Omega}$ es una función uniformemente continua en D y tiene una única extensión continua $\widetilde{X}(t,\omega)$ en $[0,T]$. Definimos $\widetilde{X}(t,\omega) = 0$ para $t \in [0,T]$ y $\omega \notin \widetilde{\Omega}$. La continuidad estocástica implica que el proceso $\widetilde{X}(t,\omega)$ es una versión de $X(t,\omega)$ y la desigualdad (6.8) implica que las trayectorias de $\widetilde{X}$ son γ-Hölder continuas como afirma el enunciado. ■

6.3. Procesos Gaussianos

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 6.14 *Sea $X = \{X_t : t \in \mathcal{I}\}$ un proceso estocástico H-valuado. Decimos que X es un* ***proceso Guassiano en*** *H si, dada $N \in \mathbb{N}$ y $t_1, \ldots, t_N \in \mathcal{I}$, la variable aleatoria H^N-valuada $(X_{t_1}, \ldots, X_{t_N})$ es Gaussiana.*

Proposición 6.15 *Sea $X = \{X_t : t \in [0,\infty)\}$ un proceso Guassiano en H. Supongamos que $E(X_t) = 0_H$ para cada $t \geq 0$ y que existen constantes $M > 0$ y $\gamma \in (0,1]$ tales que:*

$$E(\|X_t - X_s\|_H^2) \leq M|t-s|^\gamma \qquad \forall, t, s \in [1,\infty). \tag{6.9}$$

Entonces, X tiene una α-Hölder continua versión, para cada $\alpha \in (0, \frac{\gamma}{2})$.

Demostración: De (6.9) se sigue que:

$$E(\|X_t - X_s\|_H^{2\varepsilon}) \leq M^\varepsilon |t-s|^{(\gamma\varepsilon - 1)+1} \qquad \forall, t, s \in [1,\infty), \quad \forall 0 < \varepsilon < 1$$

El teorema de la prueba de Kolmogorv (6.13) asegura que existe $\widetilde{X}$ una versión de X tal que $\widehat{X}$ es un proceso α_ε Hölder continuo con $\alpha_\varepsilon < \frac{\gamma\varepsilon-1}{2\varepsilon}$ para todo $0 < \varepsilon < 1$.

En consecuencia, X tiene una versión $\widetilde{X}$ que α-Hölder continua para cada $\alpha \in (0, \frac{\gamma}{2})$. ■

Sea $\xi : \Omega \to H$ una variable aleatoria Gaussiana $N_{m,Q}$. Notemos, que por el teorema de cambio de variable,

$$\int_\Omega \|\xi\|_H^2 \, dP(\omega) = \int_H \|u\|_H^2 \, d\xi_\sharp P(u) < \infty$$

pues $d\xi_\sharp P$ es una medida Gaussiana (ver Teorema 4.13). Luego, $\|\xi\|_H$ es Lebesgue-integrable y, por tanto, ξ es Bochner-integrable. Esta observación nos permite dar la siguiente definición.

La **covarianza de** ξ se define como el operador $Q_\xi : H \to H$ dado por:

$$Q_\xi(u) := E(\langle \xi - m, u\rangle_H \, (\xi - m)) = \mathfrak{B} \int_\Omega \langle \xi(\omega) - m, u\rangle_H \, (\xi(\omega) - m) \, dP(\omega) \qquad (6.10)$$

Esta función es claramente lineal y continua. Observa que, para cualesquiera u, v la función $\langle \cdot, v\rangle_H : H \to \mathbb{R}$ dada por $(\langle \cdot, v\rangle_H)(x) = \langle x, v\rangle_H$ es lineal y continua de modo que por el Teorema 2.33 tenemos que:

$$\begin{aligned}
\langle Q_\xi(u), u\rangle_H &= \left\langle \mathfrak{B} \int_\Omega \langle \xi(\omega) - m, u\rangle_H \, (\xi(\omega) - m) \, dP(\omega), v \right\rangle_H \\
&= \int_\Omega \langle \langle \xi(\omega) - m, u\rangle_H \, (\xi(\omega) - m), v\rangle_H \, dP(\omega) \\
&= \int_\Omega \langle \xi(\omega) - m, u\rangle_H \langle \xi(\omega) - m, v\rangle_H \, dP(\omega) \\
&= \int_H \langle w - m, u\rangle_H \langle w - m, v\rangle_H \, d\xi_\sharp P(w).
\end{aligned}$$

La Definición 4.3 asegura que $Q_\xi = Q$.

Por otra parte, dado que ξ es Bochner-integrable, siguiendo una idea similar a la anterior (ver Teorema 2.33), concluimos que:

$$\langle E(\xi), u\rangle_H = E(\langle \xi, u\rangle_H) = \int_\Omega \langle \xi(\omega), u\rangle_H \, dP(\omega) = \int_H \langle v, u\rangle \, d\xi_\sharp P(v).$$

En consecuencia, $E(\xi) = m$ por la Definición 4.2.

Argumentando como antes, podemos dar la siguiente definición.

Dadas $\xi, \eta : \Omega \to H$ variables aleatorias Gaussianas $N_{m,Q}$ y $N_{n,L}$ respectivamente, definimos la correlación de ξ y η como el operador $\rho(\xi, \eta) : H \to H$ dado por:

$$\rho(\xi, \eta)(u) := E(\langle \xi - m, u\rangle_H \, (\eta - n)) = \mathfrak{B} \int_\Omega \langle \xi(\omega) - m, u\rangle_H \, (\eta(\omega) - n) \, dP(\omega). \qquad (6.11)$$

Y, $\rho(\xi,\eta)$ satisface que, para cualesquiera $u, v \in H$

$$\begin{aligned}\langle \rho(\xi,\eta)(u), v\rangle_H &= \langle E(\langle \xi - m, u\rangle_H\,(\eta - n)), v\rangle_H \\ &= E(\langle \xi - m, u\rangle_H\,\langle \eta - n, v\rangle_H) \\ &= E(\langle \xi, u\rangle_H\,\langle \eta, v\rangle_H) - \langle m, u\rangle_H \langle n, v\rangle_H.\end{aligned}$$

Observa, que si ξ y η son independientes entonces $E(\langle \xi, u\rangle_H)E(\langle \eta, v\rangle_H)$ (ver Proposición 5.26). Y, en consecuencia $\langle \rho(\xi,\eta)(u), v\rangle_H = 0$. El recíproco no es cierto en general ver, por ejemplo, [87].

Extendemos este razonamiento para procesos estocásticos H-valuados Gaussianos.

Definición 6.16 *Sea $X = \{X_t : t \in [0,\infty)\}$ un proceso Guassiano en H. Definimos:*

(i) MEDIA: $m_t := E(X_t)$ *para cada* $t \geq 0$.

(ii) COVARIANZA: $Q_t(u) := E(\langle X_t - m_t, u\rangle_H\,(X_t - m_t))$ *para cada* $u \in H$ *y para cada* $t \geq 0$.

(iii) CORRELACIÓN: $B_{t,s}(u) := E(\langle X_t - m_t, u\rangle_H\,(X_s - m_s))$ *para cada* $u \in H$ *y para cada* $t, s \geq 0$.

Definición 6.17 *Sea $X = \{X_t : t \in [0,\infty)\}$ un proceso estocástico H-valuado. Decimos que X* ***es estacionario*** *si*

$$E\left(\exp\left\{i\sum_{j=1}^{N}\langle X_{t_j+r}, h_j\rangle_H\right\}\right) = E\left(\exp\left\{i\sum_{j=1}^{N}\langle X_{t_j}, h_j\rangle_H\right\}\right) \tag{6.12}$$

para cada $N \in \mathbb{N}$ y para cualesquiera $t_1, \ldots, t_N \in [0,\infty)$, $h_1, \ldots, h_N \in H$ y $r \geq 0$.

Es decir, la distribución del vector $(X_{t_1}, \ldots, X_{t_N})$ es la misma que la del vector $(X_{t_1+r}, \ldots, X_{t_N+r})$ para cualquier valor de $r > 0$. En particular, la distribución de X_t es la misma que la de X_{t+r} para cualquier $r > 0$.

Proposición 6.18 *Sea $X = \{X_t : t \in [0,\infty)\}$ un proceso Guassiano en H. Entonces, X es estacionario si y sólo si*

(a) $m_{t+r} = m_t$ *para cualesquiera* $t, r \geq 0$.

(b) $B_{t+r,s+r} = B_{t,r}$ *para cualesquiera* $t, s, r \geq 0$.

Demostración: Sean $N \in \mathbb{N}$, $t_1, \ldots, t_N \in [0,\infty)$, $h_1, \ldots, h_N \in H$ y $r \in [0,\infty)$. Dado que X es Gaussiano, entonces $(X_{t_1}, \ldots, X_{t_N})$ es una variable aleatoria Gaussiana.

Consideremos la función $A : H^N \to \mathbb{R}^N$ dada por:

$$A(u_1, \ldots, u_N) := (\langle u_1, h_1\rangle_H, \ldots, \langle u_N, h_N\rangle_H).$$

Esta función es claramente lineal. Observa que $\|A(u_1,\ldots,u_N)\|^2_{\mathbb{R}^N} \leq \sum_{j=1}^N \|u_j\|^2_H\|h_j\|^2_H$ para todo $(u_1,\ldots,u_N) \in H^2$. En consecuencia, A es continua. La Proposición 5.49 asegura que:

$$(\langle X_{t_1}, h_1\rangle_H, \ldots, \langle X_{t_N}, h_N\rangle_H)$$

es una variable aleatoria Gaussiana $\mathbb{R}^N$-valuada. De hecho,

$$E\left(\exp\left\{i\sum_{j=1}^N \langle X_{t_j}, h_j\rangle_H\right\}\right) = \exp\left\{i\sum_{j=1}^N \langle m_j, h_j\rangle\right\}\exp\left\{-\tfrac{1}{2}\sum_{j=1}^N\sum_{k=1}^N \langle B_{t_k,t_j}h_k, h_j\rangle_H\right\}$$

y el resultado se sigue directamente. ■

6.4. Filtraciones

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 6.19 *Sea $\mathcal{I} = (\mathcal{I}, \leq)$ un conjunto parcialmente ordenado. Una familia $\{\mathcal{F}_i : i \in \mathcal{I}\}$ de σ-álgebras de subconjuntos de Ω es una **filtración** si, $\mathcal{F}_i \subset \mathcal{F}_j$ para cada $i, j \in \mathcal{I}$ tales que $i \leq j$.*

Veamos algunos ejemplo, pero antes demos las siguientes definiciones.

Definición 6.20 *Sea $\mathcal{I} = (\mathcal{I}, \leq)$ un conjunto parcialmente ordenadoy $\{X_i : \Omega \to H : i \in \mathcal{I}\}$ una familia de variables aleatorias H-valuadas. Sea $i \in \mathcal{I}$ arbitrario pero fijo. Entonces, la σ-álgebra generada por $\{X_j : j \leq i\}$, que denotaremos por $\sigma(\{X_j : j \leq i\})$, es la menor σ-álgebra de subconjuntos de Ω tal que X_j es medible (con respecto de esta σ-álgebra) para cada $j \leq i$.*

Extendemos la noción de independencia para una familia arbitraria de variables aleatorias y σ-álgebras.

Definición 6.21 *Sea $\mathcal{I} = (\mathcal{I}, \leq)$ un conjunto parcialmente ordenado y $(X_i) := \{X_i : \Omega \to H : i \in \mathcal{I}\}$ una familia de variables aleatorias H-valuadas. Decimos que (X_i) es independiente si, cualquier colección finita $X_{i_1}, \ldots, X_{i_N}$ de (X_i) es independiente.*

Del mismo modo, una familia $(\mathcal{G}_i) := \{\mathcal{G}_i : i \in \mathcal{I}\}$ de sub-σ-álgebras de $\mathcal{F}$ es independiente si, cualquier colección finita $\mathcal{G}_{i_1}, \ldots, \mathcal{G}_{i_N}$ lo es.

Proposición 6.22 *Sea $\mathcal{I} = (\mathcal{I}, \leq)$ un conjunto parcialmente ordenado y $(X_i) := \{X_i : \Omega \to H : i \in \mathcal{I}\}$ una familia de variables aleatorias H-valuadas. Entonces, (X_i) es independiente si y sólo si la familia de las σ-álgebras generadas por las variables aleatorias, es decir, $\{\sigma(X_i) : i \in \mathcal{I}\}$ es independiente.*

Demostración: $\Rightarrow$) : Fijemos $N \in \mathbb{N}$. Sean $X_{i_1}, \ldots, X_{i_N}$ en (X_i) con $i_n \neq i_m$ para cada $n \neq m$ en $\{1, \ldots, N\}$.

Sea $A_m \in \sigma(X_{i_m})$ para cada $m \in \{1, \ldots, N\}$. Observa que, para cada $m \in \{1, \ldots, N\}$, $A_m = \{\omega \in \Omega : X_{i_m}(\omega) \in B_{i_m}\} = \{X_{i_m} \in (B_{i_m})\}$ para algún $B_{i_m} \in \mathcal{B}(H)$ (ver 5.39).

En consecuencia,

$$P\left(\prod_{m=1}^{N} A_{i_m}\right) = P\left(\prod_{m=1}^{N} \{X_{i_m} \in B_{i_m}\}\right) = \prod_{m=1}^{N} P(\{X_{i_m} \in B_{i_m}\}) = \prod_{m=1}^{N} P(A_{i_m})$$

lo que demuestra que $(\sigma(X_i))$ es independiente.

$\Leftarrow$) : Fijemos $N \in \mathbb{N}$. Sean $X_{i_1}, \ldots, X_{i_N}$ en (X_i), $B_{i_1}, \ldots, B_{i_N}$ subconjuntos en $\mathcal{B}(H)$, con $i_n \neq i_m$ para cada $n \neq m$ en $\{1, \ldots, N\}$. Dado que $\{X_{i_m} \in (B_{i_m})\} \in \sigma(X_{i_m})$ para cada $m = 1, \ldots, N$, la independencia de $(\sigma(X_i))$ implica que:

$$P\left(\{X_{i_1} \in B_{i_1}\} \cap \cdots \cap \{X_{i_N} \in B_{i_N}\}\right) = \prod_{m=1}^{N} P(\{X_{i_m} \in B_{i_m}\})$$

y, por tanto, (X_i) es independiente. Esto demuestra la proposición. ∎

En lo que resta de esta sección, $\mathcal{I}$ denotará un conjunto parcialmente ordenado. Si $X = \{X_i : i \in \mathcal{I}\}$ es un proceso estocástico H-valuado y definimos la σ-álgebra

$$\mathcal{F}_i := \sigma(\{X_j : j \leq i\}) \qquad \forall\, i \in \mathcal{I}$$

entonces $\{\mathcal{F}_i : i \in \mathcal{I}\}$ es claramente una filtración y, usualmente es llamada la **filtración canónica del proceso** X**.**

Definición 6.23 *Sea $X = \{X_i : i \in \mathcal{I}\}$ un proceso estocástico H-valuado. Decimos que X es **adaptado** a una filtración $\{F_i : i \in \mathcal{I}\}$ si, la variable aleatoria $X_i : \Omega \to \mathcal{F}$ es una función $\mathcal{F}_i$-medible para cada $i \in \mathcal{I}$.*

Por ejemplo, $X = \{X_i : i \in \mathcal{I}\}$ es siempre adaptado a su filtración canónica.

6.5. Martingalas

Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle \cdot, \cdot \rangle_H, \| \cdot \|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Esta última sección tiene como objetivo dar un breve introducción del concepto de martingala para procesos estocásticos H-valuados.

Definición 6.24 *Sean $\mathcal{I} = (\mathcal{I}, \leq)$ un orden parcial, $\{X_i : i \in \mathcal{I}\}$ un proceso estocástico H-valuado y $(\mathcal{F}_i) := \{\mathcal{F}_i : i \in \mathcal{I}\}$ una filtración. Decimos que $\{X_i : i \in \mathcal{I}\}$ es una **martingala H-valuada** con respecto de la filtración $(\mathcal{F}_i)$ si cumple las siguientes tres propiedades:*

(MG1) *X_i es Bochner-integrable para cada $i \in \mathcal{I}$.*

(MG2) *Es adaptado a la filtración.*

(MG3) *$E(X_j \,|\, \mathcal{F}_i) = X_i$ c.d. para cualesquiera $i \leq j$ en $\mathcal{I}$.*

En lo que resta de la sección, $\mathcal{I}$ denotará un conjunto parcialmente ordenado. Veamos algunos ejemplos.

Ejemplo 6.25 *Sea $\xi : \Omega \to H$ una variable aleatoria Bochner-integrable fija y $(\mathcal{F}_i) := \{\mathcal{F}_i : i \in \mathcal{I}\}$ una filtración. Definamos $X_i : \Omega \to H$ como:*

$$X_i := E(\xi \,|\, \mathcal{F}_i).$$

Entonces, $(X_i) = \{X_i : i \in \mathcal{I}\}$ es una martingala H-valuada.

Demostración: El Teorema 5.35 asegura que $X_i : \Omega \to H$ es una función $\mathcal{F}_i$-medible y Bochner-integrable para cada $i \in \mathcal{I}$. Esto prueba (MG1) y (MG2).

Finalmente, sean $i \leq j$. Dado que $\mathcal{F}_i \subset \mathcal{F}_j$, la Proposición 5.37 implica que:

$$E(X_j \,|\, \mathcal{F}_i) = E(E(\xi \,|\, \mathcal{F}_j) \,|\, \mathcal{F}_i) = E(\xi \,|\, \mathcal{F}_i) = X_i \quad c.s.$$

lo que prueba (MG3).

Por tanto, (X_i) es una martingala H-valuada. ■

Ejemplo 6.26 *Sea (ξ_k) una sucesión de variables aleatorias H-valuadas e independientes tales que $E(\xi_k) = 0_H$ para cada $k \in \mathbb{N}$ y consideremos $(\mathcal{F}_k)$ la filtración canónica generada por (ξ_k). Definamos $S_n := \sum_{k=1}^{n} \xi_k$ la sucesión de sumas parciales de (ξ_k).*

Entonces, (S_n) es una martinagala con respecto de la filtración $(\mathcal{F}_k)$

Demostración: Las propiedades (MG1) y (MG2) son inmediatas. Observa que, para cualesquiera $n \leq m$, la Proposición 5.37 implica que:

$$E(S_m \,|\, \mathcal{F}_n) = E(\xi_1 + \ldots + \xi_m \,|\, \mathcal{F}_n) = E(S_n \,|\, \mathcal{F}_n) + E(\xi_{n+1} \,|\, \mathcal{F}_n) + \ldots + E(\xi_m \,|\, \mathcal{F}_n) \quad (6.13)$$

Como (ξ_k) es independiente, la Proposición 6.22 asegura que ξ_j es independiente de $\mathcal{F}_n$ para cada $n \leq j$ y, por tanto, $E(\xi_j \,|\, \mathcal{F}_n) = E(\xi_j)$ c.s. para cada $n \leq j$ (ver Proposición 5.37).

Aplicando nuevamente la Proposición 5.37, en (6.13) concluimos que:

$$\begin{aligned} E(S_m \,|\, \mathcal{F}_n) &= E(\xi_1 + \ldots + \xi_m \,|\, \mathcal{F}_n) \\ &= E(S_n \,|\, \mathcal{F}_n) + \sum_{j=n+1}^{m} E(\xi_j \,|\, \mathcal{F}_n) \\ &= E(S_n \,|\, \mathcal{F}_n) + \sum_{j=n+1}^{m} E(\xi_j) \\ &= E(S_n \,|\, \mathcal{F}_n) = S_n \ \ c.s \end{aligned}$$

pues S_n es claramente $\mathcal{F}_n$-medible. En consecuencia, se satisface (MG3) y esto prueba la afirmación. ■

Proposición 6.27 *Sean $X = \{X_i \ :, i \in \mathcal{I}\}$ y $Y = \{Y_i \,:\, i \in \mathcal{I}\}$ dos martingalas con respecto de la misma filtración $\mathcal{F} = \{\mathcal{F}_i \,:\, i \in \mathcal{I}\}$. Entonces, el proceso $\{\alpha X_i + \beta Y_i \,:\, i \in \mathcal{I}\}$ es una martingala con respecto de la filtración $\mathcal{F} = \{\mathcal{F}_i \,:\, i \in \mathcal{I}\}$ para cualesquiera $\alpha, \beta \in \mathbb{R}$.*

Demostración: Sean $\alpha, \beta \in \mathbb{R}$. Es inmediato que $\alpha X_i + \beta Y_i$ es una función $\mathcal{F}_i$-medible para cada $i \in \mathcal{I}$. La Proposición 2.30 asegura que $\alpha X_i + \beta Y_i$ es Bochner-integrable para cada $i \in \mathcal{I}$. Esto prueba (MG1) y (MG2).

Finalmente, sean $i \leq j$ en $\mathcal{I}$. Se sigue directamente de la Proposición 5.38 que

$$E(\alpha X_j + \beta Y_j \,|\, \mathcal{F}_i) = \alpha\, E(X_j \,|\, \mathcal{F}_i) + \beta\, E(Y_j \,|\, \mathcal{F}_i) = \alpha\, X_i + \beta\, Y_i \ \ \text{c.d.}$$

lo que prueba (MG3).

Por tanto, $\{\alpha X_i + \beta Y_i \,:\, i \in \mathcal{I}\}$ es una martingala H-valuada. ■

Observemos lo siguiente: Sea $X = \{X_t \,:\, 0 \leq t \leq T\}$ un proceso estocástico H-valuado tal que X_t es Bochner-integrable para cada $t \in [0, T]$ y consideremos una filtración $\mathcal{F} = \{\mathcal{F}_t \,:\, 0 \leq t \leq T\}$.

Si X es una martinagala H-valuada con respecto de la filtración $\mathcal{F}$ entonces para $u \in H$ arbitraria pero fija definamos $\eta_t : \Omega \to \mathbb{R}$ como:

$$\eta_t(\omega) := \langle X_t(\omega), u \rangle_H \qquad \forall\, t \in [0, T]. \tag{6.14}$$

Es claro que $\eta = \{\eta_t \,:\, 0 \leq t \leq T\}$ es un proceso estocástico real adaptado a la filtración $\mathcal{F}$. Ahora, como X_t es Bochner-integrable, aplicando el Teorema 2.33 se tiene que:

$$E(\eta_t) = E\left(\langle X_t, u \rangle_H\right) = \langle E(X_t), u \rangle_H < \infty \quad \forall\, t \in [0, T].$$

En consecuencia, η_t es Lebesgue-integrable para todo $t \in [0, T]$.

Finalmente, si $0 \leq s \leq t \leq T$ entonces:

$$\mathfrak{B}\int_\Omega E(X_t \,|\, \mathcal{F}_s)\, dP = \mathfrak{B}\int_\Omega X_s\, dP \quad \text{c.d.}$$

Aplicando nuevamente el Teorema 2.33 se tiene que:

$$\begin{aligned}\int_\Omega \langle E(X_t \,|\, \mathcal{F}_s), u\rangle_H\, dP &= \left\langle \mathfrak{B}\int_\Omega E(X_t \,|\, \mathcal{F}_s)\, dP, u\right\rangle_H \\ &= \left\langle \mathfrak{B}\int_\Omega X_s\, dP, u\right\rangle_H \\ &= \int_\Omega \langle X_s, u\rangle_H\, dP \quad \text{c.d.}\end{aligned}$$

por lo que concluimos que el proceso $\eta = \{\eta_t \,:\, 0 \leq t \leq T\}$ es una martingala real con respecto de la filtración $\mathcal{F}$.

Inversamente: Supongamos que $\{\langle X_t, u\rangle_H \,:\, 0 \leq t \leq T\}$ es una martingala real para todo $u \in H$ con respecto de la filtración $\mathcal{F}$. El Teorema de mensurabilidad de Pettis asegura que el proceso $\{X_t \,:\, 0 \leq t \leq T\}$ es adaptado a la filtración $\mathcal{F}$.

Por otra parte, el Teorema 2.33 implica que, para cualesquiera $0 \leq s \leq t \leq T$ se tiene que:

$$\begin{aligned}\left\langle \mathfrak{B}\int_\Omega E(X_t \,|\, \mathcal{F}_s)\, dP, u\right\rangle_H &= \int_\Omega \langle E(X_t \,|\, \mathcal{F}_s), u\rangle_H\, dP \\ &= \int_\Omega \langle X_s, u\rangle_H\, dP \\ &= \left\langle \mathfrak{B}\int_\Omega X_s\, dP, u\right\rangle_H \qquad \forall\, u \in H,\end{aligned}$$

por lo que necesariamente:

$$\mathfrak{B}\int_\Omega E(X_t \,|\, \mathcal{F}_s)\, dP = \mathfrak{B}\int_\Omega X_s\, dP$$

y esto prueba que $\{X_t \,:\, 0 \leq t \leq T\}$ es una martingala H-valuada.

Definición 6.28 *Sean $\mathcal{I} = (\mathcal{I}, \leq)$ un orden parcial, $\{\eta_i \,:\, i \in \mathcal{I}\}$ un proceso estocástico real y $(\mathcal{F}_i) := \{\mathcal{F}_i \,:\, i \in \mathcal{I}\}$ una filtración. Decimos que $\{\eta_i \,:\, i \in \mathcal{I}\}$ es una* ***submartingala*** *con respecto de la filtración $(\mathcal{F}_i)$ si cumple las siguientes tres propiedades:*

(i) *η_i es Lebesgue-integrable para cada $i \in \mathcal{I}$.*

(ii) *Es adaptado a la filtración.*

(iii) *$E(\eta_j \,|\, \mathcal{F}_i) \geq \eta_i$ c.d. para cualesquiera $i \leq j$ en $\mathcal{I}$.*

Por otra parte, decimos que $\{\eta_i \,:\, i \in \mathcal{I}\}$ es una ***supermartingala*** *con respecto de la filtración $(\mathcal{F}_i)$ si cumple las siguientes tres propiedades:*

(i) η_i *es Lebesgue-integrable para cada* $i \in \mathcal{I}$.

(ii) Es adaptado a la filtración.

(iii) $E(\eta_j \,|\, \mathcal{F}_i) \le \eta_i$ *c.d. para cualesquiera* $i \le j$ *en* $\mathcal{I}$.

Proposición 6.29 *Sea $X = \{X_t \,:\, 0 \le t \le T\}$ una martingala H-valuada con respecto de una filtración $\mathcal{F} = \{\mathcal{F}_t \,:\, 0 \le t \le T\}$. Entonces, $\{\|X_t\|_H \,:\, 0 \le t \le T\}$ es una submartingala con respecto de $\mathcal{F} = \{\mathcal{F}_t \,:\, 0 \le t \le T\}$.*

Demostración: Para cualesquiera $0 \le s \le t \le T$ se sigue directamente de la Proposición 5.39 que:

$$\|X_s\|_H = \|E(X_t \,|\, \mathcal{F}_s)\|_H \le E(\|X_t\|_H \,|\, \mathcal{F}_s).$$

Esta última afirmación junto con la Proposición 2.18 y el criterio de integrabilidad de Lebesgue-Bochner (ver Teorema 2.31) implican que $\{\|X_t\|_H \,:\, 0 \le t \le T\}$ es una submartingala. ■

Definición 6.30 *Sea $X = \{X_i \,:\, i \in \mathcal{I}\}$ una martingala H-valuada y $p \in [1, \infty)$. Decimos que X es una* ***L^p-martingala*** *si, $X_i \in L^p(\Omega; H)$ para cada $i \in \mathcal{I}$.*

Es posible concluir de las Proposiciones 5.39 y 6.29 que si $X = \{X_t \,:\, 0 \le t \le T\}$ una L^p-martingala entonces $\{\|X_t\|_H^p \,:\, 0 \le t \le T\}$ es una submartingala para todo $p \in [1, \infty)$.

Fijemos $T > 0$ número real y $\mathcal{F} := \{\mathcal{F}_t \,:\, 0 \le t \le T\}$ una filtración. Denotaremos por $\mathcal{M}_T^2(H)$ al conjunto de todas los procesos estocásticos H-valuados $\{X_t \,:\, 0 \le t \le T\}$ que son adaptados a $\mathcal{F}$, continuos, L^2-martingalas y que satisfacen que $X_0 = 0_H$.

La Proposición 6.27 nos permite concluir que el conjunto $\mathcal{M}_T^2(H)$ es un espacio vectorial sobre $\mathbb{R}$. De hecho, probaremos que es un espacio de Banach para lo cual requerimos los siguientes resultados.

El siguiente teorema, debido a Joseph Doob, establece que el supremo de un colección de L^p-martingalas está también en L^p.

Teorema 6.31 (Desigualdad maximal de Doob) *Sea $\{X_1, \ldots, X_N\}$ una martingala H-valuada con respecto de una filtración $\{\mathcal{F}_1, \ldots, \mathcal{F}_N\}$ y definamos $X_N^* : \Omega \to \mathbb{R}$ como:*

$$X_N^* := \sup_{1 \le k \le N} \|X_k\|_H.$$

Entonces, se tiene que:

(a) Para cada $\gamma > 0$ se tiene que:

$$P(\{X_N^* > \gamma\}) \leq \frac{1}{\gamma}\, E(\|X_N\|_H). \tag{6.15}$$

(b) Si $p \in (1, \infty)$ y $X_N \in L^p(\Omega; H)$, entonces $X_N^* \in L^p(\Omega)$ y

$$\|X_N^*\|_{L^p(\Omega)} \leq \frac{p}{p-1}\, \|X_N\|_{L^p(\Omega;H)} \tag{6.16}$$

Demostración: (a) Fijemos $\gamma > 0$ y definamos $\tau : \Omega \to \{1, \ldots, N\}$ como $\tau := \text{mín}\{1 \leq k \leq N : \|X_K\|_H > \gamma\}$ en donde convenimos $\text{mín}\, \varnothing := N + 1$. Así, $\{X_N^* > \gamma\} = \{\tau \leq N\}$. Observa que en $\{\tau = k\}$ se tiene que $\|X_k\|_H > \gamma$ y, por tanto,

$$\gamma\, P(\{X_N^* > \gamma\}) = \gamma \sum_{k=1}^{N} P(\{\tau = k\}) \leq \sum_{k=1}^{N} E(1_{\{\tau = k\}} \cdot \|X_k\|_H).$$

De la desigualdad de Jensen condicional se tiene que:

$$\|X_k\|_H = \|E(X_N \,|\, \mathcal{F}_k)\|_H \leq E(\|X_N\|_H \,|\, \mathcal{F}_k)$$

por lo que:

$$\begin{aligned}
\gamma\, P(\{X_N^* > \gamma\}) = \gamma \sum_{k=1}^{N} P(\{\tau = k\}) &\leq \sum_{k=1}^{N} E(1_{\{\tau = k\}} \cdot \|X_k\|_H) \\
&\leq \sum_{k=1}^{N} E(1_{\{\tau = k\}} \cdot \|X_N\|_H) \\
&= E(1_{\{\tau \leq N\}} \cdot \|X_N\|_H) \\
&= E(1_{\{X_N^* > \gamma\}} \cdot \|X_N\|_H)
\end{aligned}$$

lo que prueba la desigualdad (6.15).

(b) Si $\|X_N^*\|_{L^p(\Omega)} = 0$ el resultado se satisface trivialmente así es que supongamos que $\|X_N^*\|_{L^p(\Omega)} > 0$. Aplicando el inciso anterior, la fórmula de integración por partes (ver 5.12) y la desigualdad de Hölder obtenemos:

$$\begin{aligned}
\|X_N^*\|_{L^p(\Omega)}^p = \int_\Omega |X_N^*|^p \, dP &= \int_0^\infty p\gamma^{p-1} P(\{X_N^* > \gamma\}) \, d\gamma \\
&\leq \int_0^\infty p\gamma^{p-2} P(\{X_N^* > \gamma\}) \, d\gamma \\
&= E\left(\|X_N\|_H \int_0^{X_N^*} p\gamma^{p-2} \, d\gamma \right) \\
&= \frac{p}{p-1} E(\|X_N\|_H \, |X_N^*|^{p-1}) \\
&\leq \frac{p}{p-1} E(\|X_N\|_H^p) E(|X_N^*|^p)^{(p-1)/p} \\
&= \frac{p}{p-1} \|X_N\|_{L^p(\Omega;H)} \, \|X_N\|_{L^p(\Omega)}^{p-1}.
\end{aligned}$$

Dividendo la desigualdad anterior entre $\|X_N\|_{L^p(\Omega)}^{p-1}$ obtenemos la desigualdad (6.16). ■

La desigualdad máximal de Doob es válida si la martingala está definido sobre un conjunto numerable $\mathcal{I}$. Por otra parte, si la martingala es un proceso continuo, la desigualdad máximal de Doob es válida también cuando $\mathcal{I} = [0, T] \subset \mathbb{R}$ como enunciamos a continuación:

Si $\{X_t \,:\, 0 \leq t \leq T\}$ es una martingala continua H-valuada con respecto de una filtración $\{\mathcal{F}_t \,:\, 0 \leq t \leq T\}$ entonces para todo $p \in [1, \infty)$ se cumple:

$$\left(E\left(\sup_{t\in[0,T]} \|X_t\|_H^p \right) \right)^{1/p} \leq \frac{p}{p-1} \sup_{t\in[0,T]} \left(E(\|X_t\|_H^p) \right)^{1/p}. \tag{6.17}$$

Invitamos al lector a consular, por ejemplo, [62] para los detalles de la prueba.

Para $X \in \mathcal{M}_T^2(H)$ denotamos por:

$$\|X\|_{\mathcal{M}_T^2(H)} := \left(E\left(\sup_{t\in[0,T]} \|X_t\|_H^2 \right) \right)^{1/2}. \tag{6.18}$$

Nota que, como $\|X_t\|_H^2$ es una submartingala real por la Proposición 6.29 entonces aplicando la Proposición 5.39 concluimos que:

$$E(\|X_t\|_H^2) \leq E(\|X_T\|_H^2) \qquad \forall\, t \in [0, T].$$

De la desigualdad máximal de Doob se sigue que:

$$\|X\|_{\mathcal{M}_T^2(H)} \leq 2 \sup_{t\in[0,T]} \left(E(\|X_t\|_H^2) \right)^{1/2} \leq 2E(\|X_T\|_H^2)^{1/2}$$

y, por tanto, $\|X\|_{\mathcal{M}_T^2(H)} \in \mathbb{R}$. De hecho, lo anterior implica que $\|\cdot\|_{\mathcal{M}_T^2(H)}$ es una norma pues probamos que $\sup_{t\in[0,T]} \|X_t\|_H \in L^2(\Omega)$.

Teorema 6.32 *El espacio* $(\mathcal{M}_T^2(H), \|\cdot\|_{\mathcal{M}_T^2(H)})$ *es de Banach.*

Demostración: Sea (X_k) una sucesión de Cauchy en $\mathcal{M}_T^2(H)$. Denotaremos por $X_k = \{X_{k,t} \,:\, 0 \leq t \leq T\}$ al proceso estocástico H-valuado para todo $k \in \mathbb{N}$.

Para $\varepsilon > 0$ dada, existe $k_0 \in \mathbb{N}$ tal que:

$$\|X_k - X_j\|^2_{\mathcal{M}_T^2(H)} := E\left(\sup_{t\in[0,T]} \|X_{k,t} - X_{j,t}\|^2_H\right) < \varepsilon \qquad \forall\, k, j \geq k_0. \tag{6.19}$$

Dado que:

$$\|X_{k,t} - X_{j,t}\|^2_H \leq \sup_{t\in[0,T]} \|X_{k,t} - X_{j,t}\|^2_H \qquad \forall\, t \in [0,T], \tag{6.20}$$

por lo que $(X_{k,t})$ es de Cauchy en $L^2(\Omega; H)$ para todo $t \in [0,T]$. De este modo, existe $X_t \in L^2(\Omega; H)$ tal que $X_{k,t} \to X_t$ en $L^2(\Omega; H)$ para todo $t \in [0,T]$. Por tanto, consideremos el proceso H-valuado $X := \{X_t \,:\, 0 \leq t \leq T\}$ que es claramente adaptado a $\mathcal{F}$ y satisface que $X_0 = 0_H$.

Como $X_{k,t} \to X_t$ en $L^2(\Omega; H)$ para todo $t \in [0,T]$, el Teorema 5.18 implica que $X_{k,t} \to X_k$ en probabilidad para todo $t \in [0,T]$. El teorema de Riesz-Weyl junto con el teorema de Egorov aseguran que existe una subsucesión $(X_{k_j,t})$ de $(X_{k,t})$ tal que $X_{k_j,t} \to X_t$ casi uniformemente para todo $t \in [0,T]$. Dado que $X_{k_j,t}$ es un proceso continuo para todo $j \in \mathbb{N}$ y $X_{k_j,t} \to X_t$ casi uniformemente para todo $t \in [0,T]$ entonces X es un proceso continuo.

Para cualesquiera $0 \leq s \leq t \leq T$ se tiene que:

$$E(X_{k_j,t} \,|\, \mathcal{F}_s) = X_{k_j,s} \quad c.d. \tag{6.21}$$

Observa que $X_{k_j} \to X$ en $L^2(\Omega; H)$ por lo que es posible afirmar que para todo $t \in [0,T]$, existe $g_t \in L^1(\Omega)$ tal que $\|X_{k_j,t}\|_H \leq g_t$ c.d. Aplicando el teorema de la convergencia dominada de Lebesgue-Bochner (ver Teorema 2.34) concluimos que $X_s \in L^1(\Omega; H)$ y que para cualesquiera $A \in \mathcal{F}_s$ en (6.21) se tiene que:

$$\mathfrak{B}\int_A E(X_t \,|\, \mathcal{F}_s)\, dP = \lim_{j\to\infty} \mathfrak{B}\int_A E(X_{k_j,t} \,|\, \mathcal{F}_s)\, dP = \lim_{j\to\infty} \mathfrak{B}\int_A X_{k_j,s}\, dP = \mathfrak{B}\int_A X_s\, dP.$$

En consecuencia, $(X_t \,|\, \mathcal{F}_s) = X_s$ para todo $0 \leq s \leq t \leq T$ lo que establece que $X \in \mathcal{M}_T^2(H)$. Finalmente, de (6.19) y (6.20) concluimos que:

$$\|X_{k,t} - X_t\|^2_H < \varepsilon \qquad \forall\, k \geq k_0, \quad \forall\, t \in [0,T]$$

y, por tanto, que $X_k \to X$ en $\mathcal{M}_T^2(H)$. ■

Capítulo 7

Proceso de Wiener en espacios de Hilbert

En este último capítulo desarrollamos uno de los conceptos más importantes del análisis estocástico por sus múltiples aplicaciones en distintas áreas de las matemáticas, la física, la biología y otras disciplinas como las finanzas e ingeniería. Un *proceso de Wiener* es un proceso estocástico a tiempo continuo cuyas características están determinadas, de alguna forma, a través del estudio de procesos Gaussianos.

Un proceso de Wiener es también llamado movimiento Browniano en honor al botánico escocés Robert Brown que, en 1827, observó a través de un microscopio que las trayectorias formadas por partículas atrapadas en las cavidades dentro de un grano de polen en el agua no podían mecanizarse de la misma forma en distintos intervalos de tiempo. Aunque a lo largo de la historia este fenómeno derivó en muchos estudios relevantes de la física fue hasta 1923 que el matemático Norbert Wiener probó la existencia y unicidad de un proceso con tales características.

En este capítulo establecemos las propiedades más elementales de un proceso de Wiener real y las cuales nos permitirán extender este concepto para procesos definidos en un espacio de Hilbert separable de dimensión infinita. Daremos la construcción explícita de un proceso de Wiener real a través de la teoría dada en los capítulos anteriores y de lo que comúnmente se conoce como *función de ruido blanco*. Una función de ruido blanco permite desarrollar modelos que representan, de mejor manera, a los fenómenos físicos reales pues una de sus principales características es que sus valores no están correlacionados a lo largo del tiempo. Damos aquí un ejemplo concreto.

Los procesos de Wiener en espacios de Hilbert arbitrarios permiten construir conceptos más complejos dentro del análisis estocástico, por ejemplo, la integral estocástica que puede usarse para resolver ecuaciones de evolución estocástica.

El desarrollo de este capítulo se basó enteramente en [3], [16], [20], [21], [40], [53], [54], [55], [64] y [62].

7.1. Función de ruido blanco

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable de dimensión infinita.

Definición 7.1 *Sea $\mathscr{W} : H \to L^2(\Omega)$ una función. Decimos que, $\mathscr{W}$ es un* ***proceso H-isonormal Gaussiano en*** *Ω si cumple las siguientes dos propiedades:*

(i) $\mathscr{W}(h)$ es una variable Gaussiana para cada $h \in H$.

(ii) $E(\mathscr{W}(u)\cdot\mathscr{W}(v)) = \langle u, v\rangle_H$ para cualesquiera $u, v \in H$.

En la literatura, a un proceso H-isonormal Gaussiano se le conoce como *función de ruido blanco*. Veamos que un proceso isonormal H es siempre una función lineal.

Proposición 7.2 *Si $\mathscr{W} : H \to L^2(\Omega)$ es un proceso H-isonormal en Ω entonces, $\mathscr{W}$ es una función lineal.*

Demostración: Sean $u, v \in H$ y $a, b \in \mathbb{R}$. Observemos que:

$$\begin{aligned}
&\int_\Omega [\mathscr{W}(au+bv) - (a\mathscr{W}(u) + b\mathscr{W}(v))]^2\, dP \\
&= \int_\Omega [\mathscr{W}(au+bv)]^2\, dP - 2\int_\Omega \mathscr{W}(au+bv)\cdot[a\mathscr{W}(u)+b\mathscr{W}(v)]\, dP + \int_\Omega [a\mathscr{W}(u)+b\mathscr{W}(v)]^2\, dP \\
&= \int_\Omega \mathscr{W}^2(au+bv)\, dP - 2a\int_\Omega \mathscr{W}(au+bv)\cdot\mathscr{W}(u)\, dP - 2b\int_\Omega \mathscr{W}(au+bv)\cdot\mathscr{W}(v)\, dP \\
&+ a^2\int_\Omega [\mathscr{W}(u)]^2\, dP + 2ab\int_\Omega \mathscr{W}(u)\cdot\mathscr{W}(v)\, dP + b^2\int_\Omega [\mathscr{W}(v)]^2\, dP.
\end{aligned}$$

De lo anterior y el inciso *(ii)* de la definición de proceso isonormal se tiene que:

$$\begin{aligned}
E([\mathscr{W}(au+bv) - (a\mathscr{W}(u)+b\mathscr{W}(v))]^2) &= \langle au+bv, au+bv\rangle_H - 2a\langle au+bv, u\rangle_H \\
&- 2b\langle au+bv, v\rangle_H + a^2\langle u,v\rangle_H + 2ab\langle u,v\rangle_H \\
&+ b^2\langle v,v\rangle_H \\
&= 0.
\end{aligned}$$

En consecuencia, $\mathscr{W}(au+bv) = a\mathscr{W}(u) + b\mathscr{W}(v)$ y esto prueba la afirmación. ■

Notemos que, para cualesquiera $h_1, \ldots, h_N \in H$, de la linealidad de $\mathscr{W}$ se tiene que $\mathscr{W}(h_1) + \ldots + \mathscr{W}(h_N) = \mathscr{W}(h_1 + \ldots + h_N)$ por lo que $(\mathscr{W}(h_1), \ldots, \mathscr{W}(h_N))$ es una variable aleatoria $\mathbb{R}^N$-valuada Gaussiana. De este modo, $\{\mathscr{W}(h) : h \in H\}$ es un proceso real Gaussiano.

Supongamos $H = (H, \mu)$ un espacio de Hilbert Gaussiano con $\mu = N_Q$ tal que $\ker Q = \{0_H\}$. El objetivo de esta sección es construir un proceso isonormal en H.

Dado que $Q : H \to H$ es compacto, existe $\mathcal{B}$ una base de Hilbert de H compuesta por los vectores propios de Q, a saber, $\mathcal{B} = \{e_k \,:\, k \in \mathbb{N}\}$. Consideremos (λ_k) la sucesión de números reales positivos constituida por los valores propios de Q.

Lema 7.3 *El rango $Q(H)$ de Q es un subespacio vectorial propio y denso de H dado por:*

$$Q(H) = \left\{ v \in H \,:\, \sum_{k=1}^{\infty} \frac{1}{\lambda_k^2} \langle v, e_k \rangle_H^2 < \infty \right\}.$$

Demostración: Sea $v \in Q(H)$. Existe $u \in H$ tal que $v = Q(u)$. Observa que $\langle v, e_k \rangle_H = \langle Q(u), e_k \rangle_H = \langle u, Q(e_k) \rangle_H = \lambda_k \rangle u, e_k \rangle_H$ para cada $k \in \mathbb{N}$. En consecuencia (ver Teorema 1.47),

$$u = \sum_{k=1}^{\infty} \langle u, e_k \rangle_H \, e_k = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \langle v, e_k \rangle_H \, e_k$$

y, de la fórmula de Parseval se tiene que:

$$\infty > \|u\|_H^2 = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^2} |\langle v, e_k \rangle_H|^2.$$

Así, es inmediato además que $Q(H)$ es subconjunto propio.

Sea $u_0 \in (Q(H))^{\perp}$. Entonces, $0 = \langle u_0, Q(u) \rangle_H = \langle Q(u_0), u \rangle_H$ para todo $u \in H$ lo que implica que $Q(u_0) = 0_H$ y, por tanto, $u_0 = 0_H$ pues $\ker Q = \{0_H\}$.

Así, por el teorema del complemento ortogonal, $\overline{Q(H)} = H$ lo que prueba que $Q(H)$ es denso en H. ■

Consideremos ahora el operador $Q^{1/2} : H \to H$ tal que $Q^{1/2} \circ Q^{1/2} = Q$ y cuya descripción es la siguiente:

$$Q^{1/2}(u) := \sum_{k=1}^{\infty} \sqrt{\lambda_k} \, \langle u, e_k \rangle_H \, e_k.$$

El Teorema 3.49 asegura que $Q^{1/2}$ es un operador compacto, autoadjunto, positivo y de traza finita. El rango $Q^{1/2}(H)$ es llamado *reproducción del kernel de la medida* $\mu = N_Q$. Argumentando como en Lema 7.3 podemos afirmar que el rango de $Q^{1/2}$ es un subespacio denso y propio de H descrito de la forma:

$$Q^{1/2}(H) = \left\{ v \in H \,:\, \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \langle v, e_k \rangle_H^2 < \infty \right\}.$$

Para $v \in Q^{1/2}(H)$ definimos $W_v : H \to \mathbb{R}$ como $W_v(u) = \langle Q^{-1/2}(v), u\rangle_H$ en donde $Q^{-1/2} : H \to H$ denota al inverso izquierdo de $Q^{1/2}(H)$ que existe pues $\ker Q^{1/2} = \{0_H\}$. Para cada $v \in Q^{1/2}(H)$ se tiene que $\int_H |\langle Q^{-1/2}(v), u\rangle_H|\, d\mu(u) \leq \|Q^{-1/2}(v)\|_H^2 \int_H \|u\|_H^2\, d\mu(u) < \infty$ pues μ tiene segundo momento finito por ser Gaussiana.

En consecuencia, definamos $\mathcal{W} : Q^{1/2}(H) \to L^2(H,\mu)$ como:

$$\mathcal{W}(v) := W_v. \tag{7.1}$$

Esta función es claramente lineal. Sean $v_1, v_2 \in Q^{1/2}(H)$. Existen $u_1, u_2 \in H$ tales que $Q^{1/2}(u_i) = v_i$ para $i = 1, 2$. Así pues,

$$\begin{aligned}
\langle v_1, v_2\rangle_H = \left\langle Q^{1/2}(u_1), Q^{1/2}(u)\right\rangle_H &= \langle u_1, Q^{1/2}(Q^{1/2}(u_2))\rangle_H \\
&= \langle u_1, Q(u_2)\rangle_H \\
&= \int_H \langle u_1, w\rangle_H \langle u_2, w\rangle_H\, d\mu(w) \\
&= \int_H \langle Q^{-1/2}(v_1), w\rangle_H \langle Q^{-1/2}(v_2), w\rangle_H\, d\mu(w) \\
&= \int_H W_{v_1}(w) \cdot W_{v_2}(v)\, d\mu(w).
\end{aligned}$$

En consecuencia, $\mathcal{W}$ es continua en $Q^{1/2}(H)$. Y, dado que $Q^{1/2}(H)$ es denso en H, $\mathcal{W}$ tiene una única extensión lineal y continua $\widetilde{\mathcal{W}} : H \to L^2(H,\mu)$.

En lo que resta del capítulo, haciendo abuso de notación, escribimos para $h \in H$

$$W_h(u) = \langle Q^{-1/2}(h), u\rangle_H,$$

por lo que no haremos distinción entre $\mathcal{W}$ y $\widetilde{\mathcal{W}}$. Observa que hemos probado que, para $u_1, u_2 \in H$, $E(\mathcal{W}(u_1) \cdot \mathcal{W}(u_2)) = \langle u_1, u_2\rangle_H$.

Finalizamos esta sección y con ella el objetivo planteado con las siguientes proposiciones.

Proposición 7.4 *Sea $v \in H$. Entonces $W_v : H \to \mathbb{R}$ es una variable aleatoria Gaussiana con media 0 y covarianza $\|v\|_H^2$.*

Demostración: Sea $\nu_v := (W_v)_\sharp \mu$. Dado que $v \in H$, existe una sucesión $(v_k) = (Q^{1/2}(u_k))$ en $Q^{1/2}(H)$ tal que $v_k \to v$ en H y, por tanto, $W_{v_k}(u) \to W_v(u)$ para cada $u \in H$. Entonces, por el teorema de la convergencia dominada de Lebesgue se tiene que:

$$\begin{aligned}
\int_H \exp\{i\eta\, W_v(u)\}\, d\mu(u) &= \lim_{k\to\infty} \int_H \exp\{i\eta W_{v_k}(u)\}\, d\mu(u) \\
&= \lim_{k\to\infty} \int_H \exp\{i\eta\, \langle Q^{-1/2}(v_k), u\rangle_H\}\, d\mu(u) \\
&= \lim_{k\to\infty} \exp\{-\tfrac{1}{2}\langle \eta Q(Q^{-1/2}(v_k)), \eta Q^{-1/2}(v_k)\rangle_H\}
\end{aligned}$$

$$\begin{aligned}
&= \lim_{k\to\infty} \exp\{-\tfrac{1}{2}\langle \eta Q(u_k), \eta u_k\rangle_H\} \\
&= \lim_{k\to\infty} \exp\{-\tfrac{1}{2}\eta^2\langle Q^{1/2}(u_k), Q^{1/2}(u_k)\rangle_H\} \\
&= \lim_{k\to\infty} \exp\{-\tfrac{1}{2}\eta^2\|v_k\|_H^2\} \\
&= \exp\{-\tfrac{1}{2}\eta^2\|v\|_H^2\} \qquad \forall\, \eta \in \mathbb{R}.
\end{aligned}$$

Aplicando el teorema de cambio de variable (ver Teorema 5.9) concluimos que:

$$\widehat{\nu_v}(\eta) = \int_{\mathbb{R}} \exp\{i\eta\, x\}\, d\nu_v(x) = \int_H \exp\{i\eta\, W_v(u)\}\, d\mu(u) = \exp\{-\tfrac{1}{2}\eta^2\|v\|_H^2\} \quad \forall\, \eta \in \mathbb{R}.$$

lo que prueba la afirmación. ■

Concluimos que $\mathcal{W} : H \to L^2(H, \mu)$ es un proceso isonormal Gaussiano. De hecho, aplicando un procedimiento similar a la prueba del Corolario 5.47 podemos concluir lo siguiente.

Proposición 7.5 *Sean $v_1, \ldots, v_N \in H$ ($N \in \mathbb{N}$). Entonces, $(W_{v_1}, \ldots, W_{v_N})$ es una variable aleatoria Gaussiana $\mathbb{R}^N$-valuada con medida $0_{\mathbb{R}^N}$ y covarianza $Q_{v_1,\ldots,v_N}$ dada por $(Q_{v_1,\ldots,v_N})_{i,j} = \langle v_i, v_j\rangle_H$.*

Corolario 7.6 *Las variables aleatorias reales $W_{v_1}, \ldots, W_{v_N}$ son independientes si y sólo si $v_1, \ldots, v_N$ son mutuamente ortogonales.*

Demostración: Se sigue directamente de la Proposición 5.51. ■

Un dato curioso sobre el rango $Q^{1/2}(H)$ de $Q^{1/2}$ puede consultarse, por ejemplo, en [20; 23] donde se prueba que $\mu(Q^{1/2}(H)) = 0$.

7.2. Proceso de Wiener real

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. En esta sección recordamos la definición de un proceso de Wiener real y algunas de sus propiedades básicas.

Definición 7.7 *Sea $W = \{W_t\,;\, 0 \le t \le T\}$ un proceso estocástico real. Decimos que W es un **proceso de Wiener real de parámetro** $\lambda > 0$ si,*

(W1) *$W_0 : \Omega \to \mathbb{R}$ es tal que $W_0(\omega) = 0$ p.c.t. $\omega \in \Omega$.*

(W2) *Para cualesquiera $0 \le s < t \le T$, la variable aleatoria $W_t - W_s$ tiene distribución $N(0, \lambda(t-s))$.*

(W3) *Para cualquier partición finita $0 < t_1 < \cdots < t_m \le T$ de $[0, T]$, las variables aleatorias:*

$$W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_m} - W_{t_{m-1}}$$

son independientes.

(W4) *W es continuo.*

Si $\lambda = 1$ diremos que W es un ***proceso de Wiener real estándar****.*

Para referirnos a la propiedad (W3) de la definición anterior decimos que el proceso W tiene incrementos independientes. En (W2), a la variable aleatoria $W_t - W_s$ para $s \leq t$ la llamaremos variable de incremento y a la propiedad dada en este inciso la nombraremos incremento Gaussiano.

En lo que resta de la sección consideraremos a W un proceso de Wiener estándar y lo llamaremos únicamente proceso de Wiener real. Veamos algunas propiedades.

Proposición 7.8 *Sea $W : [0,T] \times \Omega \to \mathbb{R}$ un proceso de Wiener. Entonces, W es un proceso Gaussiano. Más aún, si $m \in \mathbb{N}$ y $0 < t_1 < \cdots < t_m \leq T$ es una partición finita de $[0,T]$, entonces la variable aleatoria $\mathbb{R}^m$-valuada $(W_{t_1}, \ldots, W_{t_m})$ es Gaussiana con media $0_{\mathbb{R}^m}$ y operador de covarianza dado por:*

$$Q_{t_1,\ldots,t_m} = \Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^* \tag{7.2}$$

donde:

$$D_{t_1,\ldots,t_m} = \operatorname{diag}(t_1, t_2 - t_1, \ldots, t_m - t_{m-1}), \tag{7.3}$$

$\Theta : \mathbb{R}^m \to \mathbb{R}^m$ es un operador lineal dado por:

$$\Theta(x_1, \ldots, x_m) = (x_1, x_1 + x_2, \ldots, x_1 + \ldots + x_m) \tag{7.4}$$

y Θ^* es el operador adjunto de Θ.

Demostración: Sea $m \in \mathbb{N}$. Consideremos $0 < t_1 < \cdots < t_m \leq T$ partición finita de $[0,T]$ y las variables aleatorias:

$$\xi := (W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}) \quad \text{y} \quad \eta := (W_{t_1}, W_{t_2}, \ldots, W_{t_m}).$$

Dado que W es un proceso de Wiener, las variables aleatorias $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}$ son independientes. Luego, la variable aleatoria W_{t_1} tiene distribución $N(0, t_1)$ y la variable de incremento $W_{t_j} - W_{t_{j-1}}$ tiene distribución $N(0, t_j - t_{j-1})$ para cada $j = 2, \ldots, m$. Por tanto, la Proposición 5.44 asegura que la variable aleatoria ξ es una variable aleatoria Gaussiana con media $0_{\mathbb{R}^m}$ y covarianza

$$Q_\xi = \operatorname{diag}(t_1, t_2 - t_1, \ldots, t_m - t_{m-1}) := D_{t_1,\ldots,t_m}.$$

Consideremos la función $\Theta \in \mathcal{L}_c(\mathbb{R}^m, \mathbb{R}^m)$ dada en (7.4). Observa que $\eta = \Theta(\xi)$ y por la Proposición 5.46 se tiene que η es una variable aleatoria Gaussiana con media $0_{\mathbb{R}^m}$ y covarianza $\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*$ como afirma el enunciado. ■

Proposición 7.9 *Sea $W : [0,T] \times \Omega \to \mathbb{R}$ un proceso de Wiener y $m \in \mathbb{N}$. Entonces, si $A \in \mathcal{B}(\mathbb{R}^m)$ se tiene que la probabilidad:*

$$P(\{(W_{t_1}, \ldots, W_{t_m}) \in A\})$$

es igual a:

$$\frac{1}{\sqrt{(2\pi)^m (t_1)(t_2-t_1)\cdots(t_m-t_{m-1})}} \int_A \exp\left\{-\tfrac{x_1^2}{2t_1} - \tfrac{(x_2-x_1)^2}{(t_2-t_1)} - \cdots - \tfrac{(x_m-x_{m-1})^2}{(t_m-t_{m-1})}\right\} dx_1 \ldots dx_m$$

Demostración: Es inmediato que la representación matricial del operador lineal $\Theta : \mathbb{R}^m \to \mathbb{R}^m$ dado en (7.4) con respecto de la base canónica de $\mathbb{R}^m$ es igual a:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}.$$

En consecuencia, $\det(\Theta) = 1$ y puede calcularse que:

$$\Theta^{-1}(x_1, \ldots, x_m) = (x_1, x_2 - x_1, \ldots, x_m - x_{m-1}).$$

La Proposición 7.8 asegura que la distribución de $(W_{t_1}, \ldots, W_{t_m})$ es $N_{\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*}$. Así pues, $P(\{(W_{t_1}, \ldots, W_{t_m}) \in A\})$ es igual a:

$$\frac{1}{\sqrt{(2\pi)^m \det(\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*)}} \int_A \exp\left\{-\tfrac{1}{2}\langle (\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*)^{-1}(x), x\rangle_{\mathbb{R}^m}\right\} dx_1 \ldots dx_m,$$

en donde $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$.

Observa que, para cada $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$:

$$\begin{aligned} \big\langle (\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*)^{-1}(x), x\big\rangle_{\mathbb{R}^m} &= \big\langle ((\Theta^*)^{-1} \circ D_{t_1,\ldots,t_m}^{-1} \circ \Theta^{-1})(x), x\big\rangle_{\mathbb{R}^m} \\ &= \big\langle (D_{t_1,\ldots,t_m}^{-1} \circ \Theta^{-1})(x), \Theta^{-1}(x)\big\rangle_{\mathbb{R}^m} \\ &= \frac{x_1^2}{t_1} + \frac{(x_2-x_1)^2}{(t_2-t_1)} + \ldots + \frac{(x_m-x_{m-1})^2}{(t_m-t_{m-1})}. \end{aligned}$$

Y, dado que $\det(\Theta) = 1 = \det(\Theta^*)$ entonces:

$$\begin{aligned} \det(\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*) &= \det(\Theta)\det(D_{t_1,\ldots,t_m})\det(\Theta^*) \\ &= \det(D_{t_1,\ldots,t_m}) = (t_1)(t_2-t_1)\cdots(t_m-t_{m-1}). \end{aligned}$$

Por lo tanto,

$$\frac{1}{\sqrt{(2\pi)^m \det(\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*)}} \int_A \exp\left\{-\tfrac{1}{2}\left\langle (\Theta \circ D_{t_1,\ldots,t_m} \circ \Theta^*)^{-1}(x), x\right\rangle_{\mathbb{R}^m}\right\} dx_1 \ldots dx_m$$

es igual a

$$\frac{1}{\sqrt{(2\pi)^m (t_1)(t_2-t_1)\cdots(t_m-t_{m-1})}} \int_A \exp\left\{-\tfrac{x_1^2}{2t_1} - \sum_{j=1}^{m} \tfrac{(x_{j+1}-x_j)^2}{(t_{j+1}-t_j)}\right\} dx_1 \ldots dx_m$$

como afirma el enunciado. ■

Proposición 7.10 *Si $W = \{W_t : 0 \leq t \leq T\}$ es un proceso de Wiener, entonces*

$$E(W_s \cdot W_t) = \text{mín}\{s,t\} \qquad \forall s,t \in [0,T].$$

Demostración: Para cualesquiera $0 \leq s \leq t \leq T$, se tiene que:

$$E((W_s - W_t)^2) = E(W_s^2 - 2W_s \cdot W_t + W_t^2) = E(W_s^2) - 2E(W_s \cdot W_t) + E(W_t^2)$$

y, en consecuencia:

$$2E(W_s \cdot W_t) = E(W_s^2) + E(W_t^2) - E((W_s - W_t)^2) = s + t - (t - s) = 2s = 2\,\text{mín}\{s,t\}$$

como afirma el enunciado. ■

A continuación probamos que un proceso de Wiener tiene una versión α-Hölder continua para $\alpha < \frac{1}{2}$.

Proposición 7.11 *Un proceso de Wiener real $W = \{W_t : 0 \leq t \leq T\}$ tiene una versión α-Hölder continua para cada $\alpha < \frac{1}{2}$.*

Demostración: Para cualesquiera $s,t \in [0,T]$ se tiene que $E(|W_s - W_t|^2) = |s-t|$. Así, aplicando la Proposición 4.11 obtenemos que:

$$E(|W_s - W_t|^{2k}) = \frac{(2k)!}{2^k\, k!}\,|s-t|^k \qquad\qquad \forall\, k \in \mathbb{N}.$$

El resultado se sigue directamente de la prueba de Kolmogorov (ver Teorema 6.13). ■

Proposición 7.12 *Un proceso de Wiener real $W = \{W_t : 0 \leq t \leq T\}$ es una martingala con respecto de la filtración dada por*

$$\mathcal{F}_t^W := \sigma(\{W_s : 0 \leq s \leq t \leq T\}).$$

Demostración: Claramente el proceso es adaptado a la filtración $\{\mathcal{F}_t^W : 0 \le t \le T\}$ y cada variable aleatoria del proceso es integrable. Por otro lado, para cualesquiera $0 \le s \le t \le T$, aplicando la Proposición 6.22 por la propiedad de incrementos independientes, se tiene que:

$$\begin{aligned} E(W_t \,|\, \mathcal{F}_s^W) &= E((W_t - W_s) + W_s \,|\, \mathcal{F}_s^W) \\ &= E(W_t - W_s \,|\, \mathcal{F}_s^W) + E(W_s \,|\, \mathcal{F}_s^W) \\ &= E(W_t - W_s) + W_s \\ &= W_s. \end{aligned}$$

Esto prueba la afirmación. ■

7.2.1. Construcción de un proceso de Wiener

El objetivo de esta sección es probar la existencia de un proceso de Wiener real en un espacio de Hilbert Gaussiano.

Consideremos el espacio de medida $((0,T), \mathcal{B}(0,T), \lambda)$ con λ la medida de Lebesgue en $(0,T)$. De acuerdo al Teorema 4.13 es posible considerar en el espacio de Hilbert separable $L^2(0,T)$ una medida Gaussiana $\mu = N_Q$ no degenerada.

Definamos $W : [0,T] \times L^2(0,T) \to \mathbb{R}$ como:

$$W(t,u) := \mathscr{W}_{1_{[0,t]}}(u) \tag{7.5}$$

en donde $\mathscr{W}$ es la función de ruido blanco construida en la sección anterior (ver (7.1)).

Observa que, si $t = 0$ entonces:

$$\mathscr{W}_{1_{\{0\}}}(u) = \langle 1_{\{0\}}, Q^{-1/2}(u) \rangle_{L^2(0,T)} = \int_{\{0\}} Q^{-1/2}(u)\, d\lambda \quad \forall\, u \in L^2(0,T).$$

Como $\lambda(\{0\}) = 0$, entonces $\int_{\{0\}} Q^{-1/2}(u)\, d\mu = 0$ para todo $u \in L^2(0,T)$. En consecuencia, $W_0 = 0$. Esto prueba la primera propiedad del proceso de Wiener para $\{W_t : L^2(0,T) \to \mathbb{R} : t \in [0,T]\}$. Veamos las demás propiedades.

Incremento Gaussiano: Sean $0 \le s < t \le T$. Aplicando la linealidad de $\mathscr{W}$ en $L^2(0,T)$ obtenemos:

$$\mathscr{W}_{1_{[0,t]}}(u) - \mathscr{W}_{1_{[0,s]}}(u) = \mathscr{W}_{1_{(s,t]}}(u) \qquad \forall\, u \in L^2(0,T).$$

En consecuencia, $\mathscr{W}_{1_{[0,t]}} - \mathscr{W}_{1_{[0,s]}}$ es una variable aleatoria Gaussiana con media 0 y covarianza $\|1_{(s,t]}\|^2_{L^2(0,T)} = (t-s)$.

Incrementos Independientes: Fijemos $N \in \mathbb{N}$ y elijamos $0 \leq t_1 < t_2 < \ldots < t_N \leq T$ una partición del intervalo $[0, T]$. Afirmamos que el conjunto

$$\{1_{[0,t_1]}, 1_{(t_1,t_2)}, \ldots, 1_{(t_{N-1},t_N]}\}$$

es ortogonal en $L^2(0, T)$. En efecto, dado que $[0, t_1] \cap (t_{j-1}, t_j] = \varnothing$ y $(t_{j-1}, t_j] \cap (t_{k-1}, t_k] = \varnothing$ para cualesquiera $1 < j < k \leq N$ entonces:

$$\int_{L^2(0,T)} 1_{(t_{j-1},t_j]} \cdot 1_{(t_{k-1},t_k]} \, d\lambda = 0 \quad \forall\, 1 \leq j < k < N.$$

El Corolario 7.6 junto con la propiedad anterior aseguran que las variables aleatorias

$$\mathscr{W}_{1_{[0,t_1]}}, \mathscr{W}_{1_{[0,t_2]}} - \mathscr{W}_{1_{[0,t_1]}}, \ldots, \mathscr{W}_{1_{[0,t_N]}} - \mathscr{W}_{1_{[0,t_{N-1}]}}$$

son independientes.

Lo único que resta probar es la continuidad del proceso y para ello requerimos del siguiente lema.

Lema 7.13 *Sean $n > 1$, $\alpha \in (\frac{1}{2n}, 1)$ y $f \in L^{2n}(0, T)$. Definamos $F : [0, T] \to \mathbb{R}$ como:*

$$F(t) := \int_0^t (t - \xi)^{\alpha-1} f(\xi)\, d\xi. \tag{7.6}$$

Entonces, $F \in \mathcal{C}^0([0, T], \mathbb{R})$.

Demostración: Aplicando la desigualdad de Hölder-Riesz obtenemos que:

$$|F(t)| \leq \left(\int_0^t (t - \xi)^{(\alpha-1)\,\frac{2n}{2n-1}}\, d\xi\right)^{\frac{2n-1}{2n}} \|f\|_{L^{2n}(0,T)} \qquad \forall\, t \in [0, T]. \tag{7.7}$$

Por tanto, $F \in L^\infty(0, T)$. La desigualdad 7.7 implica que F es continua en 0. Ahora, probaremos que F es continua en $[\frac{t_0}{2}, T]$ para cada $t_0 \in (0, T]$. Sea $0 < \varepsilon < \frac{t_0}{2}$ y definamos:

$$F_\varepsilon(t) = \int_0^{t-\varepsilon} (t - \xi)^{\alpha-1} f(\xi)\, d\xi \qquad t \in [0, T].$$

Esta función es continua en $[\frac{t_0}{2}, T]$. Usando nuevamente la desigualdad de Hölder-Riesz se tiene que:

$$|F(t) - F_\varepsilon(t)| \leq M \left(\frac{2n-1}{2n\alpha - 1}\right)^{\frac{2n}{2n-1}} \varepsilon^{\alpha - \frac{1}{2n}} \|f\|_{L^{2n}(0,T)}.$$

En consecuencia, $\lim_{\varepsilon\to 0} F_\varepsilon(t) = F(t)$ uniformemente en $[\frac{t_0}{2}, T]$ y esto demuestra que F es continua. ■

Teorema 7.14 *La familia de variables aleatorias* $\{W_t : L^2(0,T) \to \mathbb{R} : 0 \leq t \leq T\}$ *definidas en* (7.5) *es un proceso continuo.*

Demostración: Como la demostración es larga, la subdividimos en varios pasos.

PASO 1: Queremos resolver la siguiente integral

$$\int_s^t (t-\sigma)^{\alpha-1}(\sigma-s)^{-\alpha}\, d\sigma \qquad 0 \leq s \leq \sigma \leq t \leq T. \tag{7.8}$$

con $\alpha \in (0,1)$.

Sea $\alpha \in (0,1)$. Haciendo el cambio de variable $\sigma = r(t-s)+s$ en (7.8) se sigue fácilmente que:

$$\int_0^1 (1-r)^{\alpha-1} r^{-\alpha}\, dr = \frac{\pi}{\sin \pi\alpha}.$$

PASO 2: Sea $t \in [0,T]$. Fijemos $\sigma \in [0,t]$ y definamos $g_\sigma(s) := 1_{[0,\sigma]}(s)(\sigma-s)^{-\alpha}$ para $s \in [0,t]$. Entonces,

$$\int_{(0,T)} |g_\delta(s)|^2\, d\lambda(s) = \int_0^\sigma (\sigma-s)^{-2\alpha}\, ds = \frac{\sigma^{1-2\alpha}}{1-2\alpha}$$

lo que establece que $g_\sigma(s) \in L^2(0,T)$ y $\|g_\sigma\|^2_{L^2(0,T)} = \frac{\sigma^{1-2\alpha}}{1-2\alpha}$ en donde $\alpha \in (0,\frac{1}{2})$.

A partir de ahora elegimos a $\alpha \in (0,\frac{1}{2})$. De la identidad (7.8) se sigue directamente que:

$$1_{[0,t]}(s) = \frac{\sin \pi\alpha}{\pi} \int_0^t (t-\sigma)^{\alpha-1} g_\sigma(s)\, d\sigma, \qquad s \in [0,T]. \tag{7.9}$$

PASO 3: Recordando que la función de ruido blanco $\mathcal{W} : L^2(0,T) \to L^2(L^2(0,T),\mu)$ dada por $u \mapsto \mathscr{W}_u$ es continua, obtenemos que:

$$\mathscr{W}_{1_{[0,t]}} = \frac{\sin \pi\alpha}{\pi} \int_0^t (t-\sigma)^{\alpha-1} \mathscr{W}_{g_\sigma}\, d\sigma. \tag{7.10}$$

Dado que $\mathscr{W}_{g_\sigma}$ es una variable aleatoria Gaussiana con distribución $N_{\frac{\sigma^{1-2\alpha}}{1-2\alpha}}$ entonces:

$$\int_{L^2(0,T)} |\mathscr{W}_{g_\sigma}(u)|^{2n}\, d\mu(u) = \frac{(2n)!}{2^n\, n!}(1-2\alpha)^{-n}\sigma^{n(1-2\alpha)} \qquad n > 1. \tag{7.11}$$

Como $\alpha < \frac{1}{2}$, por el teorema de Fubini se tiene que:

$$\int_0^T \left(\int_{L^2(0,T)} |\mathscr{W}_{g_\sigma}(u)|^{2n}\, d\mu(u) \right) d\sigma = \int_{L^2(0,T)} \left(\int_0^T |\mathscr{W}_{g_\sigma}(u)|^{2n}\, d\sigma \right) d\mu(u) < \infty$$

lo que prueba que $\mathscr{W}_{g_\sigma}(u) \in L^{2n}(0,T)$ p.c.t. $u \in L^2(0,T)$.

Aplicando el Lema 7.13 a la identidad (7.10) se sigue que el proceso $\{W_t : L^2(0,T) \to \mathbb{R} : 0 \leq t \leq T\}$ es continuo como afirma el enunciado. ∎

7.3. Procesos de Wiener en espacios de Hilbert

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $H = (H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ un espacio de Hilbert sobre $\mathbb{R}$ separable.

Definición 7.15 *Sea $Q \in \mathcal{L}_1^+(H)$. Un proceso estocástico $\{W_t : \Omega \to H \,:\, 0 \leq t \leq T\}$ es llamado un **Q-proceso de Wiener H-valuado** si, satisface las siguientes propiedades:*

(QW1) $W_0 = 0_H$ *c.d.*

(QW2) W *es un proceso continuo.*

(QW3) *Para cualquier partición finita $0 \leq t_1 < \cdots < t_N \leq T$ de $[0,T]$, las variables aleatorias:*

$$W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_N} - W_{t_{N-1}}$$

son independientes.

(QW4) *La variable aleatoria $W_t - W_s$ tiene distribución Gaussiana $N_{(t-s)Q}$ para cualesquiera $0 \leq s \leq t \leq T$.*

En lo que resta de esta sección y siempre que no se preste a confusión, nos referiremos únicamente por Q-proceso de Wiener.

Fijemos $u \in H$. Consideremos $\{W_t : \Omega \to H \,:\, 0 \leq t \leq T\}$ un Q-proceso de Wiener. Definamos $\vartheta_t : \Omega \to \mathbb{R}$ como:

$$\vartheta_t(\omega) := \langle W_t(\omega), u\rangle_H \qquad \forall\, 0 \leq t \leq T. \tag{7.12}$$

En consecuencia, $\theta = \{\vartheta_t \,:\, 0 \leq t \leq T\}$ es un proceso estocástico real. Observa que, $\vartheta_0(\omega) = \langle W_0(\omega), u\rangle_H = \langle 0_H, u\rangle_H = 0$ p.c.t. $\omega \in \Omega$. El primer objetivo de esta sección es probar que θ es un proceso de Wiener real.

Continuidad: Fijemos $\omega \in \Omega$. La función $W(t,\omega) : [0,T] \to H$ es continua y, como el producto escalar es una función continua entonces $\langle W(t,\omega), u\rangle_H : [0,\infty) \to \mathbb{R}$ es continua. Es decir, $\theta(t,\omega) : [0,T] \to \mathbb{R}$ es una función continua.

Incremento Gaussiano: Sean $0 \leq s < t \leq T$ y consideremos la variable aleatoria de incremento $W_t - W_s$. Dado que $W_t - W_s$ es Gaussiana $N_{(t-s)Q}$, la Proposición 5.43 asegura que la variable aleatoria real $\langle W_t - W_s, u\rangle_H$ es Gaussiana $N_{(t-s)\langle Q(u),u\rangle_H}$.

Usando la linealidad del producto escalar, obtenemos que $\vartheta_t - \vartheta_s$ tiene distribución $N(0, (t-s)\langle Q(u), u\rangle_H)$.

Incrementos Independientes: Sean $0 \leq t_1 < t_2 < \ldots < t_N \leq T$ y consideremos las variables independientes:

$$W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_N} - W_{t_{N-1}}.$$

El producto escalar $\langle \cdot, u \rangle_H : H \to \mathbb{R}$ es claramente una función medible por ser continua. Así pues, la Proposición 5.26 asegura que las variables

$$\langle W_{t_1}, u \rangle_H, \langle W_{t_2} - W_{t_1}, u \rangle_H, \ldots, \langle W_{t_N} - W_{t_{N-1}}, u \rangle_H.$$

son independientes. Nuevamente, usando la linealidad del producto escalar obtenemos que las variables aleatorias:

$$\vartheta_1, \vartheta_{t_2} - \vartheta_{t_1}, \ldots, \vartheta_{t_N} - \vartheta_{t_{N-1}}$$

son independientes.

En consecuencia, $\theta = \{\vartheta_t \,:\, 0 \leq t \leq T\}$ es un proceso de Wiener de parámetro $\langle Q(u), u \rangle_H$.

Es posible suponer que $\|u\|_H = 1$. Si H es de dimensión infinita, nos preguntamos si es posible que exista un operador $Q \in \mathcal{L}_1^+(H)$ tal que $\theta = \{\vartheta_t \,:\, t \geq 0\}$ es un proceso de Wiener estándar. La respuesta es negativa pues, en caso contrario, el operador Q tendría que cumplir que $\langle Q(u), u \rangle_H = 1$ lo cual ocurre si $Q(u) = u$. Es decir, si $Q = id_H$ lo cual es imposible pues la identidad no es un operador compacto por el Teorema de Riesz (ver Teorema 3.54).

Fijemos $Q \in \mathcal{L}_1^+(H)$. Existe $\{e_k \,:\, k \in \mathbb{N}\}$ una base de Hilbert de H constituida por los vectores propios de Q y una sucesión de números reales no negativos (λ_k) tal que $Q(e_k) = \lambda_k \, e_k$ para cada $k \in \mathbb{N}$. El objetivo de esta sección es dar una representación para un Q-proceso de Wiener en H. Para ello, requerimos de los siguientes resultados.

Proposición 7.16 *Sean $X_1, \ldots, X_N$ variables aleatorias H-valuadas tales que $X = (X_1, \ldots, X_N)$ es una variable aleatoria H^N-valuada Gaussiana N_Q. Entonces, $X_1, \ldots, X_N$ son independientes si y sólo si $X_1, \ldots, X_N$ no son correlacionadas, es decir, si $E(\langle X_i, u \rangle_H \cdot \langle X_j, v \rangle_H) = 0$ para cualesquiera $u, v \in H$ y para todo $i \neq j$ con $i, j = 1, \ldots, N$.*

Demostración: Procedemos por casos.

Caso 1. $H = \mathbb{R}$.

Sea $j \in \{1, \ldots, N\}$ fija. Observa que la variable aleatoria $X_j : \Omega \to \mathbb{R}$ es Gaussiana pues, por hipótesis, la variable aleatoria $X = (X_1, \ldots, X_N)$ $\mathbb{R}^N$-valuada es Gaussiana y $X_j = \pi_j \circ X$ en donde $\pi_j : \mathbb{R}^N \to \mathbb{R}$ es la función proyección que es lineal y continua (ver Proposición 5.50). Más aún, X_j es Gaussiana con medida 0 y covarianza igual a $\lambda_j := E(X_j^2)$.

Si $X_1, \ldots, X_N$ son independientes, la Proposición 5.27 asegura que $E(X_i \cdot X_j) = E(X_i) \cdot E(X_j) = 0$ para todo $i \neq j$ con $i, j \in \{1, \ldots, N\}$.

Inversamente: De la observación anterior y el hecho de que $X_1, \ldots, X_N$ no son correlacionadas se tiene que la representación matricial de Q está dada por $\operatorname{diag}(\lambda_1 \ldots, \lambda_N)$. En consecuencia, para cualquier $(h_1, \ldots, h_N) \in \mathbb{R}^N$ se tiene que:

$$\widehat{X}(h_1, \ldots, h_N) = \exp\left\{-\tfrac{1}{2}\sum_{j=1}^{N} \lambda_j\, h_j^2\right\} = \prod_{j=1}^{N} \exp\{-\tfrac{1}{2}\lambda_j\, h_j^2\} = \prod_{j=1}^{N} \widehat{X}_j(h_j).$$

La Proposición 5.25 asegura que $X_1, \ldots, X_N$ son independientes.

Caso 2. H es cualquier espacio de Hilbert separable.

Sean $u_1, \ldots, u_N \in H$ arbitrarios pero fijos y consideremos la función $\varsigma : H^N \to \mathbb{R}^N$ dada por $\varsigma(h_1, \ldots, h_N) := (\langle h_1, u_1\rangle_H, \ldots, \langle h_N, u_N\rangle_H)$ la cual es claramente lineal y continua. La Proposición 5.50 asegura que la variable aleatoria $\mathbb{R}^N$-valuada $\xi = (\langle X_1, u_1\rangle_H, \ldots, \langle X_n, u_N\rangle_H)$ es Gaussiana pues $\xi = \varsigma \circ X$.

Si $X_1, \ldots, X_N$ son independientes entonces $\langle X_1, u_1\rangle_H, \ldots, \langle X_n, u_N\rangle_H$ son independientes por la Proposición 5.26 y del Caso 1 concluimos que $E(\langle X_i, u_i\rangle_H \cdot \langle X_j, u_j\rangle_H) = E(\langle X_i, u_i\rangle_H) \cdot E(\langle X_j, u_j\rangle_H) = 0$ para todo $i \neq j$ con $i, j \in \{1, \ldots, N\}$.

Inversamente: Del Caso 1 se tiene que $\langle X_1, u_1\rangle_H, \ldots, \langle X_n, u_N\rangle_H$ son independientes. En consecuencia,

$$\hat{X}(u_1, \ldots, u_N) = E\left(i\sum_{j=1}^{N} \langle X_j, u_j\rangle_H\right) = \prod_{j=1}^{N} E\left(i\langle X_j, u_j\rangle_H\right) = \prod_{j=1}^{N} \hat{X}_j(u_j).$$

Aplicando la Proposición 5.25 concluimos que $X_1, \ldots, X_N$ son independientes. ∎

Teorema 7.17 (Representación de una variable aleatoria Gaussiana) *Sea $\xi : \Omega \to H$ una variable aleatoria. Entonces, ξ es Gaussiana N_Q si y sólo si*

$$\xi = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \beta_k(t)\, e_k \quad \text{en } L^2(\Omega; H) \tag{7.13}$$

en donde $\{\beta_k : \Omega \to \mathbb{R} : k \in \mathbb{N}\}$ son variables aleatorias Gaussianas $N_{0,1}$ mutuamente independientes.

Demostración: $\Rightarrow$) : El Teorema 1.47 asegura que:

$$\xi(\omega) = \sum_{k=1}^{\infty} \langle \xi(\omega), e_k\rangle_H\, e_k \qquad \forall\, \omega \in \Omega.$$

La Proposición 5.43 asegura que $\langle \xi, e_k\rangle_H$ es una variable aleatoria Gaussiana con media $\langle 0_H, e_k\rangle_H = 0$ y covarianza $\langle Q(e_k), e_k\rangle_H = \lambda_k$ para todo $k \in \mathbb{N}$. De este modo, definimos $\beta_k : \Omega \to \mathbb{R}$ como:

$$\beta_k(\omega) := \begin{cases} \langle \xi(\omega), e_k\rangle_H / \sqrt{\lambda_k} & \text{si } \lambda_k > 0, \\ 0 & \text{si } \lambda_k \leq 0, \end{cases}$$

en donde es claro que β_k es una variable aleatoria Gaussiana $N_{0,1}$ para todo $k \in \mathbb{N}$. Aplicando nuevamente el Teorema 1.47 concluimos que:

$$\xi(\omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \beta_k(\omega)\, e_k \qquad\qquad \forall\, \omega \in \Omega.$$

Fijemos $k \in \mathbb{N}$. Sea $a_j \in \mathbb{R}$ arbitraria para $1 \leq j \leq k$ entonces,

$$\sum_{j=1}^{k} a_j\, \beta_j = \sum_{\substack{j=1 \\ \lambda_j \neq 0}}^{k} \frac{a_j}{\sqrt{\lambda_j}} \langle \xi, e_j \rangle_H = \left\langle \xi, \sum_{\substack{j=1 \\ \lambda_j \neq 0}}^{k} \frac{a_j}{\sqrt{\lambda_j}} e_j \right\rangle_H$$

En consecuencia, $\sum_{j=1}^{k} a_j\, \beta_j$ es una variable aleatoria Gaussiana por la Proposición 5.43. Esta última afirmación junto con el Corolario 5.47 implican que $(\beta_1, \ldots, \beta_k)$ es una variable aleatoria $\mathbb{R}^k$-valuada. Es decir, $\{\beta_k \,:\, k \in \mathbb{N}\}$ es un proceso Gaussiano.

Sean $i, j \in \mathbb{N}$ tales que $i \neq j$ y tales que $\lambda_i \neq 0 \neq \lambda_j$. Entonces,

$$E(\beta_i \cdot \beta_j) = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} E(\langle \xi, e_i \rangle_H \cdot \langle \xi, e_j \rangle_H) = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle Q(e_i), e_j \rangle_H = \frac{\lambda_i}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle e_i, e_j \rangle_H = 0$$

lo que implica que $\{\beta_k \,:\, k \in \mathbb{N}\}$ es independiente por la Proposición 7.16.

Finalmente, para cualesquiera $n > m \geq 1$ se tiene que:

$$E\left(\| \sum_{k=m}^{n} \sqrt{\lambda_k}\, \beta_k\, e_k \|_H^2 \right) = \sum_{k=m}^{n} \lambda_k\, E(\beta_k^2) = \sum_{k=m}^{n} \lambda_k < \infty$$

pues $\sum_{k=1}^{\infty} \lambda_k = \operatorname{Tr} Q < \infty$. Por tanto, la sucesión $(\sum_{k=1}^{n} \sqrt{\lambda_k}\, \beta_k\, e_k)$ converge en $L^2(\Omega; H)$.

$\Leftarrow)$: Fijemos $k \in \mathbb{N}$ y definamos $\tau_k := \sqrt{\lambda_k} \beta_k\, e_k : \Omega \to H$. Observa que la función $L : \mathbb{R} \to H$ dada por $L(x) := \sqrt{\lambda_k}\, e_k\, x$ es lineal y, por tanto, continua pues $\mathbb{R}$ es de dimensión finita. Para cualesquiera $x \in \mathbb{R}$ y $u \in H$ se tiene que:

$$\langle L(x), u \rangle_H = \langle \sqrt{\lambda_k}\, e_k\, x, u \rangle_H = \sqrt{\lambda_k} \langle e_k, u \rangle_H\, x$$

de donde deducimos fácilmente que $L^*(u) = \sqrt{\lambda_k} \langle e_k, u \rangle_H$. Como $\beta_k : \Omega \to \mathbb{R}$ es una variable Gaussiana $N_{0,1}$, la Proposición 5.50 asegura que $L \circ \beta_k = \tau_k$ es una variable aleatoria Gaussiana con media 0_H y covarianza $(L \circ id_{\mathbb{R}} \circ L^*)(u) = \lambda_k \langle e_k, u \rangle_H\, e_k$.

Ahora, definamos $\eta_n := \sum_{k=1}^{n} \sqrt{\lambda_k}\, \beta_k\, e_k$ para todo $n \in \mathbb{N}$. Como $\{\beta_k \,:\, k \in \mathbb{N}\}$ es independiente, entonces η_n es una variable aleatoria Gaussiana con media 0_H y covarianza igual a $\mathcal{Q}_n(u) = \sum_{k=1}^{n} \lambda_k \langle e_k, u \rangle_H\, e_k$ (ver Proposición 5.49). Luego es claro que $\mathcal{Q}_n(u) \to Q(u)$ para $u \in H$.

Estamos suponiendo que $\eta_n \to \xi$ en $L^2(\Omega; H)$ y, por tanto, existe una subsucesión η_{n_j} de η_n tal que $\eta_{n_j} \to \xi$ c.d. La Proposición 5.48 asegura que ξ es una variable aleatoria Gaussiana N_Q. ■

Teorema 7.18 (Representación de un Q-proceso de Wiener) *Sea $W = \{W_t : \Omega \to H : 0 \leq t \leq T\}$ un proceso estocástico H-valuado. Entonces, W es un Q-proceso de Wiener si y sólo si*

$$W_t = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \beta_{k,t}\, e_k, \qquad t \in [0,T] \tag{7.14}$$

en donde $\{\beta_k : [0,T] \times \Omega \to \mathbb{R} : k \in \{n \in \mathbb{N} : \lambda_n > 0\}\}$ son procesos de Wiener reales mutuamente independientes.

Demostración: $\Rightarrow$) : Supongamos que W es un Q-proceso de Wiener. De las propiedades (QW1) y (QW4) se tiene que $W_t : \Omega \to H$ es Gaussiana N_{tQ} para todo $0 \leq t \leq T$. El Teorema 7.17 asegura que:

$$W_t = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \beta_{k,t}\, e_k \tag{7.15}$$

en donde:

$$\beta_{k,t} := \begin{cases} \langle W_t, e_k \rangle_H / \sqrt{\lambda_k} & \text{si } \lambda_k > 0, \\ 0 & \text{si } \lambda_k \leq 0, \end{cases}$$

para todo $0 \leq t \leq T$. Además, $\beta_{k,t} : \Omega \to \mathbb{R}$ es Gaussiana $N_{0,t}$ y $\beta_{k,t}$ son independientes para todo $k \in \mathbb{N}$ y para todo $0 \leq t \leq T$.

Fijemos $k \in \mathbb{N}$. Probaremos que $\beta_k = \{\beta_{k,t} : \Omega \to \mathbb{R} : 0 \leq t \leq T\}$ es un proceso de Wiener real.

La propiedad (W1) es inmediata de (QW1) y la Proposición 1.17.

Sea $0 = t_0 < t_1 < \ldots < t_N = T$ una partición finita de $[0,T]$ y consideremos las variables aleatorias reales:

$$\beta_{k,t_1}, \beta_{k,t_2} - \beta_{k,t_1}, \ldots, \beta_{k,t_N} - \beta_{k,t_{N-1}}.$$

Las variables aleatorias H-valuadas

$$W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_N} - W_{t_{N-1}}$$

son independientes por la Propiedad (QW3).

Observemos que, para cada $1 \leq j \leq N$ se tiene que:

$$\beta_{k,t_j} - \beta_{k,t_{j-1}} := \begin{cases} \frac{1}{\sqrt{\lambda_k}} \langle W_{t_j} - W_{t_{j-1}}, e_k \rangle_H & \text{si } \lambda_k > 0, \\ 0 & \text{si } \lambda_k \leq 0. \end{cases} \tag{7.16}$$

De (7.16) y la Proposición 5.26 concluimos que $\beta_{k,t_1}, \beta_{k,t_2} - \beta_{k,t_1}, \ldots, \beta_{k,t_N} - \beta_{k,t_{N-1}}$ son independientes lo prueba la propiedad (W3).

Análogamente, de (7.16), la propiedad (QW4) y la Proposición 5.43 concluimos que $\beta_{k,t} - \beta_{k,s}$ es Gaussiana N_{t-s} lo que prueba (W2).

Finalmente, de la continuidad de $\langle \cdot, e_k \rangle_H$ y la Propiedad (QW2) se tiene que la aplicación

$$t \mapsto \frac{1}{\sqrt{\lambda_k}} \langle W_t, e_k \rangle_H$$

es continua c.d. Esto prueba (W4).

Por tanto, $\beta_k = \{\beta_{k,t} : \Omega \to \mathbb{R} : 0 \leq t \leq T\}$ es un proceso de Wiener real para todo $k \in \mathbb{N}$.

Probaremos ahora que la familia $\{\beta_k : k \in \mathbb{N}\}$ es independiente. Sean $k_1, \ldots, k_j \in \mathbb{N}$ todos distintos y sea $0 = t_0 < t_1 < \ldots < t_N \leq T$ con $N \in \mathbb{N}$ una partición del intervalo $[0, T]$. Para evitar confusiones escribiremos $\beta_{k_j,t} := \beta_{k_j}(t)$ para $t \in [0, T]$. Usaremos la Proposición 6.22 para demostrar la independencia, es decir, probaremos que las σ-álgebras:

$$\sigma(\beta_{k_1}(t_1), \ldots, \beta_{k_1}(t_N)), \ldots, \sigma(\beta_{k_j}(t_1), \ldots, \beta_{k_j}(t_N))$$

son independientes. Procedemos por inducción sobre N:

Paso Base: Si $N = 1$ entonces las variables $\beta_{k_1}(t_1), \ldots, \beta_{k_j}(t_1)$ son independientes por la discusión anterior y el Teorema 7.17.

Paso Inductivo: Supongamos que:

$$\sigma(\beta_{k_1}(t_1), \ldots, \beta_{k_1}(t_{N-1})), \ldots, \sigma(\beta_{k_j}(t_1), \ldots, \beta_{k_j}(t_{N-1}))$$

son independientes.

Dado que $\beta_{k_i}(t_N) = (\beta_{k_i}(t_N) - \beta_{k_i}(t_{N-1})) + \beta_{k_i}(t_{N-1})$ entonces:

$$\sigma(\beta_{k_i}(t_1), \ldots, \beta_{k_i}(t_{N-1}), \beta_{k_i}(t_N)) = \sigma(\beta_{k_i}(t_1), \ldots, \beta_{k_i}(t_{N-1}), \beta_{k_i}(t_N) - \beta_{k_i}(t_{N-1})),$$

para todo $1 \leq i \leq j$. Observemos que:

$$\beta_{k_i}(t_N) - \beta_{k_i}(t_{N-1}) := \begin{cases} \frac{1}{\sqrt{\lambda_{k_i}}} \langle W_{t_N} - W_{t_{N-1}}, e_{k_i} \rangle_H & \text{si } \ \lambda_{k_i} > 0, \\ 0 & \text{si } \ \lambda_{k_i} \leq 0. \end{cases}$$

para $1 \leq i \leq j$ son ortogonales por pares en $L^2(\Omega)$ y $W_{t_N} - W_{t_{N-1}}$ es Gaussiana por la propiedad (QW4). Por tanto, $\{\beta_{k_i}(t_N) - \beta_{k_i}(t_{N-1}) : 1 \leq i \leq j\}$ es independiente.

Sea $A_{i,\ell} \in \mathcal{B}(\mathbb{R})$ con $1 \leq i \leq j$ y $1 \leq \ell \leq N$. Usando la propiedad (QW3) junto con la Proposición 6.22 afirmamos que $\sigma(W_s : s \leq t_{N-1})$ y $\sigma(W_{t_N} - W_{t_{N-1}})$ son independientes y, por tanto,

$$
\begin{aligned}
&P\left(\bigcap_{i=1}^{j}\{\{\beta_{k_i}(t_1)\in A_{i,1}\}\cap\cdots\cap\{\beta_{k_i}(t_{N-1})\in A_{i,N-1}\}\cap\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\}\right)=\\
&=P\left(\bigcap_{i=1}^{j}\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\cap\bigcap_{i=1}^{j}\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\right)\\
&=P\left(\bigcap_{i=1}^{j}\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\right)\cdot P\left(\bigcap_{i=1}^{j}\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\right)
\end{aligned}
$$

pues $\bigcap_{i=1}^{j}\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\in\sigma(W_s\ :\ s\leq t_{N-1})$ y $\bigcap_{i=1}^{j}\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\in\sigma(W_{t_N}-W_{t_{N-1}})$. Usando hipótesis de inducción y las observaciones previas obtenemos que:

$$
\begin{aligned}
&P\left(\bigcap_{i=1}^{j}\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\right)\cdot P\left(\bigcap_{i=1}^{j}\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\right)=\\
&=\left(\prod_{i=1}^{j}P\left(\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\right)\right)\cdot\left(\prod_{i=1}^{j}P\left(\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\right)\right)\\
&=\prod_{i=1}^{j}P\left(\bigcap_{\ell=1}^{N-1}\{\beta_{k_i}(t_\ell)\in A_{i,\ell}\}\cap\{\beta_{k_i}(t_N)-\beta_{k_i}(t_{N-1})\in A_{i,N}\}\right)
\end{aligned}
$$

y esto prueba el resultado.

$\Leftarrow)$: Definiendo

$$
W_t := W(t)=\sum_{k=1}^{\infty}\sqrt{\lambda_k}\,\beta_k(t)\,e_k,\qquad t\in[0,T]
$$

es claro que $W_t\in L^2(\Omega;H)$ para todo $t\in[0,T]$.

Dado que β_k es un proceso de Wiener para cada $k\in\mathbb{N}$ entonces $\beta_k(0)=0$ c.d. para cada $k\in\mathbb{N}$ y, por tanto, $W(0)=0_H$ c.d. Ahora, sean $0\leq s\leq t\leq T$. Por hipótesis, $\beta_k(t)-\beta_k(s)$ es una variable aleatoria Gaussiana N_{t-s} para cada $k\in\mathbb{N}$ de modo que $W(t)-W(s)$ es Gaussiana $N_{(t-s)Q}$ por el Teorema 7.17. Esto prueba (QW1) y (QW4).

Definamos $W_n:[0,T]\times\Omega\to H$ como $W_n(t,\omega):=\sum_{k=1}^{n}\sqrt{\lambda_k}\,\beta_k(t,\omega)\,e_k$ para todo $n\in\mathbb{N}$. Es claro que W_n es continua c.d. para todo $n\in\mathbb{N}$. Así pues, para cualesquiera $1\leq m<n$,

usando la desigualdad maximal de Doob obtenemos que:

$$\begin{aligned}
E\left(\sup_{t\in[0,T]} \|W_n(t)-W_m(t)\|_H^2\right) &= E\left(\sup_{t\in[0,T]} \sum_{k=m+1}^{n} \lambda_k \beta_k^2(t)\right) \\
&\leq \sum_{k=m+1}^{n} \lambda_k\, E\left(\sup_{t\in[0,T]} \beta_k^2(t)\right) \\
&\leq \sum_{k=m+1}^{n} \lambda_k\, 2^2 \sup_{t\in[0,T]} E\left(\beta_k^2(t)\right) \\
&\leq 2^2\, T \sum_{k=m+1}^{n} \lambda_k < \infty
\end{aligned}$$

lo que establece que $\sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \beta_k(t,\omega)\, e_k$ es continua c.d. lo que prueba (QW2)

Finalmente, sea $0 = t_0 \leq t_1 < t_2 < \ldots < t_N \leq T$ una partición del intervalo $[0,T]$ y consideremos las variables aleatorias:

$$W(t_1), W(t_2)-W(t_1), \cdots, W(t_N)-W(t_{N-1}).$$

El Teorema 7.17 asegura que:

$$W(t_{j+1})-W(t_j) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, e_k\, (\beta_k(t_{j+1})-\beta_k(t_j)), \qquad \forall\, 1 \leq j \leq N.$$

Por hipótesis, las variables $\beta_k(t_1), \beta_k(t_2)-\beta_k(t_1), \cdots, \beta_k(t_N)-\beta_k(t_{N-1})$ son independientes para todo $k \in \mathbb{N}$. Luego, la Proposición 5.26 asegura que las variables $\sqrt{\lambda_k}\, e_k \beta_k(t_1), \sqrt{\lambda_k}\, e_k(\beta_k(t_2)-\beta_k(t_1)), \cdots, \sqrt{\lambda_k}\, e_k(\beta_k(t_N)-\beta_k(t_{N-1}))$ son independientes para todo $k \in \mathbb{N}$. La discusión anterior junto con las Proposiciones 5.28 y 5.49 implican que las variables $W(t_1), W(t_2)-W(t_1), \cdots, W(t_N)-W(t_{N-1})$ son independientes, es decir, se cumple (QW3). Esto concluye la demostración. ■

Bibliografía

[1] Aliprantis, C.D. *Principles of Real Analysis*. Third Edition. San Diego, Academic Press.

[2] Aliprantis, C.D. & Border, K.C. *Infinite Dimensional Analysis. A Hitchhiker's Guide*. Third Edition. Springer-Verlag Berlin Heidelberg, 2006.

[3] Alvarado-Solano, A.E. & Fonseca-Mora, C.A. Stochastic integration in Hilbert spaces with respect to cylindrical martingale-valued measures. *ALEA, Lat. Am. J. Probab. Math. Stat.* **18** (2021), 1267–1295.

[4] Applebaum, D. Martingale-valued measures, Ornstein-Uhlenbeck processes with jumps and operator self-decomposability in Hilbert space. *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX* **1874** Lecture Notes in Math, Springer, Berlin (2006), 171-196.

[5] Athreya, K.B. & Lahiri, S.N. *Measure Theory and Probability Theory*. Springer, United States of America, 2006.

[6] Bartle, R.G. A Modern Theory of Integration. *American Mathematical Society.* **32** Providence, 2001.

[7] Bartle, R.G. *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, United States of America, 1995.

[8] Bartle, R.G. & Sherbert, D.R. *Introduction to Real Analysis*. Fourth Edition. John Wiley & Sons, United States of America, 2011.

[9] Bogachev, V.I. Gaussian Measures. *American Mathematical Society.* **62**. 1961.

[10] Bogachev, V.I. *Measure Theory. Volume I.* Springer, 2007.

[11] Bogachev, V.I. *Measure Theory. Volume II.* Springer, 2007.

[12] Bréhier, C.E. A short introduction to Stochastic PDEs. 2014. hal-00973887v2.

[13] Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2010.

[14] Castillo, E.R. *Espacios Lp.* Universidad Nacional de Colombia, Facultad de Ciencias. Departamento de Matemáticas, Bogotá, 2014.

[15] Chamorro, D. *Espacios de Lebesgue y de Lorentz. Volumen 1.* Unidad de Publicaciones de la Facultad de Ciencias de la Escuela Politécnica Nacional, Ladrón de Guevara E11-253, Quito, Ecuador, 2010.

[16] Chantladze, T. A stochastic differential equation in Hilbert space. *Soobsc. Akad. Nauk Gruzin. SSR.* **33** (1964), 529–534.

[17] Clapp, M. *Análisis Matemático.* Colección Papirhos, Serie Textos, Núm. 2, Instituto de Matemáticas de la UNAM, 3a. Edición, México, 2018.

[18] Cohn, D.L. *Measure Theory.* Second Edition. Birkhäuser, Boston, 2013.

[19] Cuevas, J.B. *Funciones Lipschitz sobre espacios métricos.* Benemérita Universidad Autónoma de Puebla, Puebla, México, 2016.

[20] Da Prato, G. *Introduction to Stochastic Analysis and Malliavin Calculus.* Scuola Normale Superiore Pisa, Pisa, 2014.

[21] Da Prato, G. & Zabczyk, J. Stochastic Equations in Infinite Dimensions. *Encyclopedia of Mathematics and its Applications (44)* Cambridge University Press, Cambridge 1992.

[22] Da Prato, G., Kwapien, S. & Zabczyk, J. Regularity of solutions of linear stochastic equations in Hilbert spaces, *Stochastics.* **23** (1987), 1-23.

[23] Da Prato, G., Kunstmann, C., Lasiecka, I., Lunardi, A., Schnaubelt, R. & Weis, L. *Functional Analytic Methods for Evolution Equations.* Springer Berlin Heidelberg, 2004.

[24] Diestel, J. & Uhl, J.J.Jr. Vector Measures. *American Mathematical Society.* **15** Providence, 1997.

[25] Dinov, I.D. *Bochner Integrals and Vector Measures*, Open Access Master's Report, Michigan Technological University, 1993.

[26] Dudley, R.M. *Real Analysis and Probability.* Cambridge, 2002.

[27] Edgar, G.A. Measurability in a Banach Space. *Indiana University Mathematics Journal.* **26** (1997), 663–677.

[28] Fabian, M., Habala, P., Hájek, P., Montesinos, V., Pelant, J. & Zizler, V. *Functional Analysis and Infinite Dimensional Geometry.* Springer, New York, 2001.

[29] Fernique, X. Fonctions aléatoires gaussiennes, les résultats récents de M. Talagrand. *Astérisque, Séminaire Bourbaki.* **38** (1986), 177-186.

[30] Folland, G.B. *Real Analysis. Modern Theory and Their Applications*. Second Edition. John Wiley & Sons, United States of America, 1999.

[31] Gohberg, I. & Goldberg, S. *Basic operator theory.* Birkhäuser, Boston, 1981.

[32] Gohberg, I. & Krein, M. Introduction to the Theory of linear Nonselfadjoint Operators in Hilbert Space. *American Mathematical Society.* **18**. 1969.

[33] Gosson, M. *Symplectic Methods in Harmonic Analysis and in Mathematical Physics*. Birkhäuser Basel, 2011.

[34] Grabinsky, G. *Teoría de la medida.* Las prensas de ciencias, Universidad Nacional Autónoma de México, Facultad de Ciencias, México, 2016.

[35] Gross, L. Potential theory on Hilbert space. *J. Funct. Anal.* **1** (1967), 123–181.

[36] Haagerup, U., Hoffmann-Jorgensen, J. & Nielsen, N.J. *Probability in Banach Spaces 6. Proceedings of the Sixth International Conference, Sandbjerg, Denmark 1986*. Birkhäuser Boston, MA, 1990.

[37] Halmos, P.R. *Measure Theory.* Springer-Verlag, United States of America, 1914.

[38] Hytönen, T., van Neerven, J., Veraar, M., & Weis, L. *Analysis in Banach Spaces: Volume I: Martingales and Littlewood-Paley Theory.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge; Vol. 63, Springer, 2016.

[39] Issoglio, E. & Riedle, M. Cylindrical fractional Brownian motion in Banach spaces. *Stochastic Processes and their Applications.* **124** (2014), 3507–3534.

[40] Itô, K. *Foundations of Stochastic Differential Equations in Infinite-Dimensional Spaces.* SIAM, Philadelphia, Pennsylvania, 1984.

[41] Itô, K. & Nisio, M. On the convergence of sums of independent Banach space valued random variables. *Osaka J. Math.* **5** (1968), 35-48.

[42] Jordan, P. & Neumann, J.v. On Inner Product in Linear Metric Spaces. *Ann. of Math.* **36** (1935), 719-723.

[43] Kallianpur, G. & Xiong, J. *Stochastic Differential Equations in Infinite Dimensional Spaces.* IMS, Inst. of Math. Statistics, Hayward, CA, 1996.

[44] Kanan, D. & Kannappan, PL. On a characterization of gaussian measures in a Hilbert space. *Ann. Inst. Henri Poincaré (XI).* **4** (1975), 397 -404.

[45] KLISINSKA, A. *Clarkson Type Inequalities and Geollletric Properties of Banach Spaces*. Lulea University of Technology, Department of Mathematics, Lulea, 1999.

[46] KOLMOGOROV, A.N. & FOMIN, S.V. *Elementos de la Teoría de Funciones y del Análisis Funcional*. Second Edition. Moscú: Editorial MIR, 1975.

[47] KÖNIG, H. *Eigenvalue Distribution of Compact Operators*. Birkhäuser, Boston, 1986.

[48] KREYSZIG, E. *Introductory Functional Analysis with Applications*. John Wiley & Sons, United States of America, 1978.

[49] KUBRUSLY, C.V. *Essentials of Measure Theory*. Springer, 2015.

[50] LEONI, G. A First Course in Sobolev Spaces. Second Edition. *American Mathematical Society* **181** Providence, 2017.

[51] LEORATO, S. A refined Jensen's inequality in Hilbert spaces and empirical approximations. *Journal of Multivariate Analysis.* **100** (2009), 1044-1060.

[52] LIMAYE, B.V. *Functional Analysis*. New Age International Publisher 1996.

[53] MAMPORIA, B. On the Wiener processes in a Banach space. *Bull. Georgian Natl. Acad. Sci.* **7** (2013), 5–14.

[54] MAMPORIA, B. Stochastic differential equations in a Banach space driven by the cylindrical Wiener process. *Transactions of A. Razmadze Mathematical Institute.* **171** (2017), 76-89.

[55] MAMPORIA, B. Stochastic integral from operator-valued function by the Wiener process in a Banach space. *Soobsc. Akad. Nauk Gruzin. SSR.* **105** (1982), 29-32.

[56] MARSDEN, J.E. & HOFFMAN, M.J. *Basic Complex Analysis.* Third Edition. W.H. Freeman and Co. 1999.

[57] MEISE, R. & VOGT, D. *Einführung in die Funktionalanalysis.* ViewegTeubner Verlag Wiesbaden, 1992.

[58] MEYER, P.A. Les Processus Stochastiques de 1950 à Nos Jours. *Development of Mathematics 1950–2000.* **5** (2009), 813-848.

[59] PIETSCH, A. *Eigenvalues and s-numbers*. Cambridge studies in advanced mathematics 13, Cambridge University Press, 1987.

[60] PIETSCH, A. *History of Banach Spaces and Linear Operators*. Birkhäuser, Boston, 2007.

[61] PISIER, G. *Martingales in Banach Spaces.* Cambridge University Press, 2016.

[62] Prévôt, C. & Röckner, M. *A Concise Course on Stochastic Partial Differential Equations.* Springer-Verlag, Berlin Heidelberg, 2007.

[63] Reed, M. & Simon, B. *Methods of modern mathematical physics I: Functional Analysis.* Academic Press, California, 1980.

[64] Riedle, M. Cylindrical Wiener Processes. In: Donati-Martin, C., Lejay, A., Rouault, A. (eds) *Séminaire de Probabilités XLIII. Lecture Notes in Mathematics*, vol 2006. Springer, Berlin, Heidelberg.

[65] Riesz, F. & Nagy, B. *Functional Analysis*, Dover, 1990.

[66] Rincón, L.A. *Curso intermedio de probabilidad.* Las prensas de ciencias, Universidad Nacional Autónoma de México, Facultad de Ciencias, México, 2015.

[67] Rincón, L.A. Introducción a la ecuaciones diferenciales estocásticas. *Memorias del Congreso Regional de Probabilidad y Estadística de la UAA,* (2005).

[68] Rincón, L.A. *Introducción a los procesos estocásticos.* Las prensas de ciencias, Universidad Nacional Autónoma de México, Facultad de Ciencias, México, 2012.

[69] Robdera, M.A. On the Riesz Representation Theorem and Integral Operators, *Quaestiones Mathematicae.* (2015), 1-15.

[70] Robdera, M.A. Tensor Integral: A New Comprehensive Approach to the Integration Theory. *British Journal of Mathematics & Computer Science.* **4** (2014), 3236-3244.

[71] Roberts, W.A. & Varberg, D.E. *Convex functions.* Pure and Applied Mathematics 57, Academic Press, New York, 1973.

[72] Rosinski, J. & Suchanecki, Z. On the space of vector-valued functions integrable with respect to the white noise. *Colloq. Math.* **43** (1980), 183–201.

[73] Royden, H.L. *Real Analysis.* Second Edition. The Macmillan Company, New York, 1968.

[74] Royden, H.L. & Fitzpatrick, D.R. *Real Analysis.* Fourth Edition. Pearson Education Asia Limited and China Machine Press, 2010.

[75] Rudin, W. *Functional Analysis.* Second Edition. McGraw-Hill, Singapore, 1991.

[76] Rudin, W. *Principios de Análisis Matemático.* Tercera Edición. McGraw-Hill, México, 1980.

[77] Santana, B.J. *Determinación del dual de algunos espacios clásicos en análisis.* Universidad de Sonora, México, 2006.

[78] Salamon, D.A. *Measure and Integration.* European Mathematical Society, Textbooks in Mathematics, ETH Zürich, 2016.

[79] Sanchéz, J.R. *La Integral de Bochner*. Benemérita Universidad Autónoma de Puebla, Puebla, México, 2017.

[80] Schatten, R. *Norm Ideals of Completely Continuous Operators*. Springer, 1970.

[81] Simpson, D.G. *The Factorial Function, n!*. Department of Physical Sciences and Engineering. Prince George's Community College, 2012.

[82] Steele, J.M. *Stochastic Calculus and Financial Applications.* Springer New York, NY, 2001.

[83] Straatman, M. *Random variables on non-separable Banach spaces.* Mathematical Institute, University of Leiden, 2017.

[84] Thamban, N.M. *Functional Analysisis A First Course.* Second Edition. PHI, Learning Private Limited, Delhi, 2002.

[85] Tindel, S., Tudor, C. & Viens, F. Stochastic evolution equations with fractional Brownian motion. *Probab. Theory Relat. Fields.* **127** (2003), 186–204.

[86] Vakhaniya, N., Tarieladze, V. & Chobanyan, S. *Probability Distributions on Banach Spaces.* D. Reidel Publishing Company, Dordrecht, 1987.

[87] van Neerven J. *Stochastic Evolution Equations, lecture notes*. University of Delft, Delft, 2007.

[88] Vera, S.E. & Vera, M.R *Análisis funcional. Espacios seminormados*. Las prensas de ciencias, Universidad Nacional Autónoma de México, Facultad de Ciencias, México, 2018.

[89] Weaver, N. *Measure Theory and Functional Analysis*. World Scientific Publishing Co. Pte. Ltd. 2013

[90] Xie, Y. Φ'-valued martingale measures and their limit theorems. *Stochastic Anal. Appl.* **3** (2001), 413–432 .

[91] Yosida, K. *Functional Analysis*. Sixth Edition. Springer-Verlag, Berlin Heidelberg, New York, 1980.