



Facultad  
de  
Ciencias

# PROYECCIÓN SOBRE UN CONJUNTO CONVEXO Y CERRADO

(Projection onto a closed convex set)

Trabajo de Fin de Grado  
para acceder al

**GRADO EN MATEMÁTICAS**

Autor:

**Jaime Terán Ramos**

Director:

**Luis Alberto Fernández Fernández**

Junio - 2025

## Resumen

En el presente trabajo se abordará el estudio de la proyección sobre conjuntos convexos y cerrados de un espacio de Hilbert. Para ello, se comenzará introduciendo los conceptos teóricos fundamentales, así como el Teorema Fundamental de la Proyección, lo que permitirá posteriormente ilustrar diferentes situaciones mediante ejemplos y representaciones gráficas.

A continuación, se analizarán casos particulares en dimensión finita, deduciendo las fórmulas clásicas para calcular distancias entre puntos, rectas y planos; y resolviendo una serie de problemas clásicos de mínimos cuadrados, analizados desde el punto de vista de las proyecciones. Todo ello servirá como base para extender el estudio a contextos de dimensión infinita, siendo de especial relevancia el caso del espacio de funciones medibles de cuadrado integrable.

Posteriormente, se estudiarán las proyecciones sobre la intersección de subespacios cerrados, presentando y demostrando el Método de Von Neumann, acompañado de ejemplos que facilitarán su comprensión y aplicación.

Seguidamente, se introducirán las transformadas de Fourier y se aplicarán al análisis de conjuntos previamente estudiados. Además, se abordarán dos aplicaciones prácticas de especial interés relacionadas con la optimización y la teoría de control óptimo. El primer ejemplo tratará el lanzamiento de un cohete al espacio, donde se analizará la relación entre empuje y peso. El segundo se centrará en la Ecuación del Estado, modelo fundamental utilizado en campos tan variados como la economía, la carrera espacial y el tratamiento de tumores en medicina.

Este recorrido teórico-práctico busca ofrecer una visión completa y estructurada del papel de las proyecciones, destacando su relevancia tanto en la teoría matemática como en diversas aplicaciones prácticas. A lo largo del trabajo se ha puesto de manifiesto cómo estas herramientas permiten abordar y resolver problemas de gran interés, evidenciando su papel fundamental en múltiples contextos.

## Abstract

This project focuses on the study of projection onto convex and closed sets in a Hilbert space, a property that ensures the existence and uniqueness of solutions. To this end, the work begins by introducing the fundamental theoretical concepts related to Hilbert spaces, as well as the Fundamental Theorem of the Projection, which will later allow for the illustration of various situations through examples and graphical representations.

Next, particular cases in finite-dimensional spaces are analyzed, deriving the classical formulas for calculating distances between points, lines, and planes, as well as solving a series of classical least squares problems from the perspective of projections. These results serve as a foundation for extending the study to infinite-dimensional contexts, with special emphasis on the space of square-integrable measurable functions.

Subsequently, projections onto the intersection of closed subspaces are studied, including the presentation and proof of the Von Neumann Method, supported by examples that will facilitate its understanding and application.

Finally, Fourier transforms are introduced and applied to the analysis of sets previously studied. Additionally, two practical applications of particular interest related to optimization and optimal control theory are addressed. The first example involves the launch of a rocket into space, analyzing the relationship between thrust and weight. The second focuses on the State Equation, a fundamental model used in diverse fields such as economics, space exploration, and tumor treatment in medicine.

This theoretical and practical journey aims to provide a comprehensive and structured view of the role of projections, highlighting their importance both in mathematical theory and in various real-world applications. Throughout the project, it is demonstrated how these tools can be used to address and solve problems of great interest, underscoring their fundamental role in multiple contexts.

## Agradecimientos

Tras cuatro años de carrera llenos de aprendizajes, experiencias y momentos inolvidables, quiero dedicar unas palabras a todas aquellas personas que han decidido acompañarme en este camino y que, con su presencia y apoyo, han conseguido que nunca me sintiera solo.

En primer lugar, quiero expresar mi más sincero agradecimiento a mi tutor, Luis Alberto, por su apoyo constante a lo largo de todo este curso. Su exigencia y su análisis crítico, siempre constructivo, han facilitado enormemente este proceso y me han ayudado a crecer académicamente. Me he sentido muy cómodo trabajando con él y, tras conocerle más de cerca, me llevo un referente para toda mi vida.

A mi familia, y en especial a mis padres, por vuestro sacrificio, esfuerzo y amor incondicional. Me habéis apoyado en todo lo que me he propuesto, enseñándome con el ejemplo lo que significa la constancia y el compromiso. Vuestras enseñanzas superan cualquier conocimiento académico y me motivan cada día a ser mejor persona.

A todos los jugadores que han pasado por mis manos estos 4 años: probablemente jamás leáis estas líneas, pero poder enseñaros lo que sé y ver que seguís confiando en mí tiempo después de haber sido vuestro entrenador es una de las mayores satisfacciones personales que he vivido. Cada entrenamiento y cada partido me han permitido desconectar de los problemas y disfrutar de lo que hago. Veros crecer es, sin duda, uno de mis mayores orgullos.

A mis amigos de toda la vida, por ser parte fundamental de esta historia. A pesar de que los estudios, los entrenamientos, y la práctica deportiva no me permitan veros tanto como me gustaría, siempre estáis ahí cuando se os necesita. Cada momento compartido con vosotros vale oro.

Por último, a “Poker”, ese grupo de amigos de la universidad que se ha convertido en una parte vital de mi vida. Esta etapa ha sido mucho más amena gracias a vosotros. Cada día de estudio, cada viaje, cada café y, en definitiva, cada instante a vuestro lado ha hecho que todo este camino merezca la pena. Os habéis convertido en parte de mi familia, y no os cambio por nada.

*Jaime Terán Ramos*



# Índice general

<b>1. Introducción</b>	<b>7</b>
<b>2. Teoría fundamental</b>	<b>9</b>
2.1. Teorema fundamental . . . . .	9
2.2. Dimensión finita . . . . .	15
2.2.1. Problemas mínimos cuadrados . . . . .	17
2.3. Dimensión infinita. Espacio $L^2[0, 1]$ . . . . .	22
<b>3. Proyección sobre intersección de subespacios cerrados</b>	<b>29</b>
3.1. Teorema de Von Neumann . . . . .	29
3.1.1. Ejemplos y resultados . . . . .	31
<b>4. Proyecciones relacionadas con Fourier</b>	<b>35</b>
4.1. Teoría de Fourier . . . . .	35
4.2. Proyección sobre conjuntos de $L^2(\mathbb{R})$ . . . . .	36
4.2.1. Conjuntos sin transformada de Fourier . . . . .	36
4.2.2. Conjuntos con transformada de Fourier . . . . .	41
4.3. Ejemplos prácticos . . . . .	44
4.3.1. Lanzamiento de un cohete . . . . .	44
4.3.2. Problema control óptimo . . . . .	48
<b>A. Preliminares</b>	<b>53</b>
A.1. Espacios Hilbert . . . . .	53



# Capítulo 1

## Introducción

En el marco del análisis funcional y la geometría en espacios de Hilbert, uno de los conceptos más importantes y de mayor número de aplicaciones, tanto teóricas como prácticas, es el de la **proyección sobre conjuntos convexos y cerrados**. Este concepto, que puede resultar sencillo desde el punto de vista de la perspectiva geométrica, tiene numerosas aplicaciones y vital importancia en problemas de optimización, en métodos numéricos y en áreas tan distintas como pueden ser el procesamiento de señales, o incluso en modelos económicos y diversas áreas de la medicina.

El presente trabajo tiene como objetivo principal estudiar en detalle el **teorema fundamental de la proyección sobre conjuntos convexos y cerrados** en espacios de Hilbert. Este teorema garantiza la existencia y unicidad de la proyección, lo que constituye una base para numerosos métodos del análisis moderno, y que nos servirá como punto de partida para el desarrollo de todos nuestros resultados. A partir de este, el trabajo explora los conceptos fundamentales y también algunas extensiones y aplicaciones, ofreciendo una visión general que conecta la parte más abstracta con su interpretación geométrica y con ejemplos prácticos.

En el apéndice, se presenta el marco teórico sobre el cual se fundamenta este trabajo. Se introducen definiciones esenciales, tales como las de **espacio de Hilbert, producto interno, y sucesión de Cauchy** junto con la revisión de propiedades fundamentales como la **desigualdad de Cauchy-Schwarz** y otros teoremas relevantes. Asimismo, se explican los fundamentos de espacios  $L^2(\Omega)$ , ejemplificando con funciones reales y complejas en distintos dominios.

En primer lugar, se presenta y demuestra el **teorema de la proyección sobre un conjunto convexo cerrado**. Además de su formulación, se incluye una interpretación geométrica, que ofrece al lector una intuición visual sobre lo que se está trabajando. Además, se recalca la necesidad de que el conjunto sea convexo y cerrado para garantizar la solución, mostrando ejemplos elegidos para ilustrar por qué estas condiciones son necesarias, presentando casos en los que la proyección no existe o no es única si alguna de ellas se omite. Del mismo modo, se aborda el caso de los **subespacios cerrados**, un caso particular del teorema fundamental, que nos simplificará en gran medida los cálculos y razonamientos.

Continuando en la sección, se dedica un apartado al estudio del caso **finito-dimensional**, en la que se deducen fórmulas clásicas como la distancia punto-recta en el plano o la distancia punto-plano en el espacio. Estas fórmulas, conocidas desde cursos básicos de geometría, se reinterpretan aquí como casos particulares de proyecciones en espacios de Hilbert. A continuación, se resuelven una serie de problemas clásicos de mínimos cuadrados, analizados desde el punto de vista de las proyecciones. En particular, se estudian tres situaciones fundamentales: la mejor aproximación polinómica a un conjunto de datos, la resolución de sistemas sobredeterminados y la aproximación de funciones continuas por polinomios en espacios de Hilbert. Cada uno de estos problemas se presenta tanto mediante métodos clásicos del

cálculo diferencial como a través de formulaciones basadas en proyecciones, mostrando la equivalencia entre ambos enfoques.

Para finalizar esta sección, el trabajo se mueve hacia contextos **infinito-dimensionales**, haciendo especial énfasis en el espacio  $L^2[0, 1]$  ampliamente utilizado en el análisis funcional. Se estudian diversos conjuntos convexos cerrados en este espacio, y se analizan las formas explícitas de las proyecciones correspondientes. El lector podrá ver cómo se calculan estas proyecciones de manera explícita, así como las condiciones necesarias para su validez.

Posteriormente, el trabajo también aborda el problema de la proyección sobre la **intersección de subespacios cerrados**, una herramienta que añade una mayor versatilidad para el cálculo de proyecciones. En este capítulo se analizan condiciones para garantizar la existencia y unicidad de la proyección, y se presenta un método efectivo para calcularla. El método de proyecciones sucesivas, también conocido como **método de von Neumann**, es una técnica iterativa que consiste en aplicar alternativamente las proyecciones sobre cada subespacio hasta alcanzar la solución común.

Se termina con una sección dedicada a las proyecciones relacionadas con la **transformada de Fourier**, así como a la presentación de dos **aplicaciones prácticas** de especial relevancia en los ámbitos de la optimización y la teoría de control óptimo. Se inicia con una introducción teórica a la transformada de Fourier, complementada con el análisis de diversos ejemplos clásicos extraídos de la literatura. Estos ejemplos se analizan en primer lugar sin transformadas y posteriormente con ellas, con el objetivo de poner de manifiesto la relación entre ambos casos.

Finalmente, se examinan dos situaciones prácticas que incluyen problemas reales como el lanzamiento de un cohete, analizando la relación empuje-peso, o la resolución de ecuaciones de estado en modelos económicos o biomédicos. Estos ejemplos muestran el gran potencial de las proyecciones como método de resolución de problemas complejos mediante diversas técnicas.

En conjunto, este recorrido teórico y práctico busca ofrecer una visión amplia del concepto de proyección, abordándolo tanto desde una perspectiva rigurosa y abstracta como desde sus aplicaciones en contextos concretos, incluso en áreas que a primera vista pueden parecer alejadas de las matemáticas. Al finalizar el trabajo, podemos concluir que las proyecciones constituyen una herramienta poderosa y versátil para resolver problemas relevantes en entornos muy diversos. Este campo no se agota con lo aquí expuesto: por ejemplo, el análisis de la Ecuación del Estado abre una línea de estudio que podría extenderse considerablemente, lo que sugiere múltiples posibilidades para continuar profundizando en investigaciones futuras.

# Capítulo 2

## Teoría fundamental

### 2.1. Teorema fundamental

En esta sección, nos centraremos en los conjuntos cerrados y convexos y tomará especial importancia el **Teorema Fundamental de la proyección**, a partir del cual desarrollaremos todo el trabajo. Además, se presentará un **corolario para subespacios cerrados**. Toda esta teoría ha sido obtenida del libro [1]. Por otra parte, aportaremos representación gráfica de todos ellos para su mejor comprensión.

Principalmente trabajaremos con conjuntos convexos y cerrados, lo cual nos garantiza la existencia y unicidad de la proyección. Comenzamos definiendo estos conceptos,

**Definición 2.1.** Sea  $X \subset H$ , siendo  $H$  un espacio de Hilbert. Se dice que  $X$  es un **conjunto convexo** si  $\forall a, b \in X, \forall \lambda \in [0, 1]$ , se tiene que  $\lambda a + (1 - \lambda)b \in X$ .

**Propiedad 2.2.** Sea  $X \subset H$ , siendo  $H$  un espacio de Hilbert.  $X$  es un **conjunto cerrado** si  $\forall \{x_n\}_{n \in \mathbb{N}} \subset X$  tal que  $x_n \xrightarrow{n} x$  en  $H$  se tiene que  $x \in X$ .

Una vez introducidos los conjuntos necesarios, estamos en condiciones de enunciar el Teorema de la Proyección sobre un conjunto convexo y cerrado, el cual constituye el resultado central sobre el que se fundamenta el desarrollo de este trabajo.

**Teorema 2.3. Teorema de la proyección sobre un convexo cerrado.**

Sea  $H$  un espacio de Hilbert y sea  $\mathbb{K} \subset H$  un convexo cerrado no vacío. Entonces para todo  $f \in H$ , existe  $\bar{u} \in \mathbb{K}$  único tal que

$$\|f - \bar{u}\| = \min_{v \in \mathbb{K}} \|f - v\|. \quad (2.1)$$

Además, se caracteriza por la propiedad

$$\begin{cases} \bar{u} \in \mathbb{K} \\ (f - \bar{u}, v - \bar{u}) \leq 0 \quad \forall v \in \mathbb{K}. \end{cases} \quad (2.2)$$

Se escribe  $\bar{u} = P_{\mathbb{K}} f$  y se denomina **proyección de  $f$  sobre  $\mathbb{K}$** .

*Demostración.* 1. Existencia.

Dada  $\{v_n\}_{n \in \mathbb{N}}$  una sucesión minimizante para (2.1), es decir

$$d_n = \|f - v_n\| \longrightarrow d = \inf_{v \in \mathbb{K}} \|f - v\|.$$

Probamos que  $\{v_n\}_{n \in \mathbb{N}}$  es de Cauchy aplicando la identidad del paralelogramo A.5. Tomamos  $a = f - v_n$  y  $b = f - v_m$  y tenemos

$$\left\| \frac{(f - v_n) + (f - v_m)}{2} \right\|^2 + \left\| \frac{(f - v_n) - (f - v_m)}{2} \right\|^2 = \frac{1}{2} (\|f - v_n\|^2 + \|f - v_m\|^2)$$

$$\iff \left\| f - \frac{v_n + v_m}{2} \right\|^2 + \left\| \frac{v_n - v_m}{2} \right\|^2 = \frac{1}{2}(d_n^2 + d_m^2)$$

Sabemos que  $\frac{v_n + v_m}{2} \in \mathbb{K}$  ( $v_n, v_m \in \mathbb{K}$  y  $\mathbb{K}$  convexo). Y por tanto  $\|f - \frac{v_n + v_m}{2}\| \geq d$ . Por lo tanto, sustituyendo y despejando en la igualdad previa tenemos

$$\left\| \frac{v_n - v_m}{2} \right\|^2 \leq \frac{1}{2}(d_n^2 + d_m^2) - d^2$$

y al tomar el límite cuando  $n, m \rightarrow \infty$  se tiene que  $\frac{1}{2}(d_n^2 + d_m^2) - d^2 \rightarrow \frac{1}{2}(2d^2) - d^2 = 0$ .

Por lo tanto,  $\lim_{n, m \rightarrow \infty} \|v_n - v_m\| \rightarrow 0$  y  $\{v_n\}$  es de Cauchy, luego existe un  $\bar{u} \in \mathbb{K}$  (por ser cerrado) tal que  $d = \|f - \bar{u}\|$ .

2. Equivalencia entre (2.1) y (2.2).

$\implies$ ) Sea  $\bar{u} \in \mathbb{K}$  verificando (2.1) y sea  $v \in \mathbb{K}$ . Como  $\bar{u}, v \in \mathbb{K}$ , entonces  $w = (1 - \lambda)\bar{u} + \lambda v \in \mathbb{K} \forall \lambda \in [0, 1]$  y, por lo tanto

$$\|f - \bar{u}\| \leq \|f - [(1 - \lambda)\bar{u} + \lambda v]\| = \|(f - \bar{u}) - \lambda(v - \bar{u})\| \iff \|f - \bar{u}\|^2 \leq \|(f - \bar{u}) - \lambda(v - \bar{u})\|^2$$

$$\iff \|f - \bar{u}\|^2 \leq \|f - \bar{u}\|^2 - 2\lambda(f - \bar{u}, v - \bar{u}) + \lambda^2 \|v - \bar{u}\|^2 \stackrel{\lambda \geq 0}{\iff} 2(f - \bar{u}, v - \bar{u}) \leq \lambda \|v - \bar{u}\|^2$$

y cuando  $\lambda \rightarrow 0$ , se verifica (2.2).

$\impliedby$ ) Inversamente, si  $\bar{u}$  verifica (2.2) entonces,  $2(f - \bar{u}, v - \bar{u}) \leq 0$  por hipótesis y  $\|\bar{u} - v\|^2 \geq 0$ . Sumando y restando  $\bar{u}$  tenemos

$$\|f - v\|^2 = \|f - \bar{u} + \bar{u} - v\|^2.$$

Sea  $a = f - \bar{u}$  y  $b = \bar{u} - v$

$$\|f - v\|^2 = \|f - \bar{u} + \bar{u} - v\|^2 = \|f - \bar{u}\|^2 - 2(f - \bar{u}, \bar{u} - v) + \|\bar{u} - v\|^2$$

por lo tanto

$$\|f - \bar{u}\|^2 - \|f - v\|^2 = 2(f - \bar{u}, v - \bar{u}) - \|\bar{u} - v\|^2 \leq 0 \quad \forall v \in \mathbb{K}$$

con lo que  $\|f - \bar{u}\|^2 \leq \|f - v\|^2 \quad \forall v \in \mathbb{K}$ , y se cumple (2.1).

3. Unicidad.

Sean  $\bar{u}_1$  y  $\bar{u}_2$  verificando (2.2). Se tiene

$$\begin{aligned} (f - \bar{u}_1, v - \bar{u}_1) &\leq 0 & \forall v \in \mathbb{K} \\ (f - \bar{u}_2, v - \bar{u}_2) &\leq 0 & \forall v \in \mathbb{K} \end{aligned}$$

sustituyendo  $v = \bar{u}_2$  y  $v = \bar{u}_1$  en las expresiones previas, tenemos

$$(f - \bar{u}_1, \bar{u}_2 - \bar{u}_1) \leq 0 \text{ y } (f - \bar{u}_2, \bar{u}_1 - \bar{u}_2) \leq 0 \implies -(f - \bar{u}_2, \bar{u}_2 - \bar{u}_1) \leq 0$$

y sumando ambas desigualdades obtenemos

$$(f - \bar{u}_1, \bar{u}_2 - \bar{u}_1) - (f - \bar{u}_2, \bar{u}_2 - \bar{u}_1) \leq 0 \implies (f - \bar{u}_1 - f + \bar{u}_2, \bar{u}_2 - \bar{u}_1) \leq 0 \implies$$

$$(\bar{u}_2 - \bar{u}_1, \bar{u}_2 - \bar{u}_1) \leq 0 \implies \|\bar{u}_2 - \bar{u}_1\|^2 \leq 0 \implies \bar{u}_2 = \bar{u}_1.$$

□

**Observación 2.4.** *El Teorema 2.3 es válido tanto en espacios de Hilbert reales como complejos. La demostración es correcta para ambos casos.*

**Interpretación geométrica:**

Veamos ahora la interpretación geométrica y visual del teorema. Tomamos  $\mathbb{K} \subset \mathbb{R}^2$ , siendo  $\mathbb{K}$  el rombo de la Figura 2.1. Realizando el producto escalar obtenemos

$$(f - \bar{u}, v - \bar{u}) = \underbrace{\|f - \bar{u}\|}_{>0} \cdot \underbrace{\|v - \bar{u}\|}_{>0} \cdot \cos(\theta) \leq 0$$

donde  $\theta$  es el ángulo entre ambos vectores, como podemos ver en la Figura 2.1, que es recto u obtuso, siendo  $\theta \in [\frac{\pi}{2}, \frac{3\pi}{2}]$ , por lo que su coseno es siempre negativo o cero, haciendo que se cumpla la propiedad.

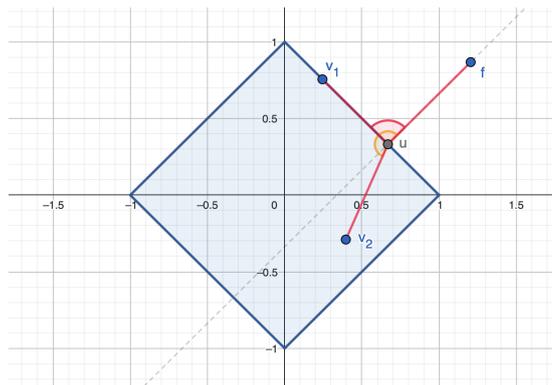


Figura 2.1: Proyección sobre un conjunto convexo y cerrado en  $\mathbb{R}^2$ .

Veamos ahora que  $\bar{u}$  es el único punto que cumple esa propiedad. Tomamos un  $\bar{u}$  distinto de la proyección, como en la Figura 2.2. En este caso, el ángulo  $\theta < \frac{\pi}{2}$ , por lo que  $\cos(\theta) > 0$  lo que hace que no se cumpla la propiedad, luego no puede ser la proyección.

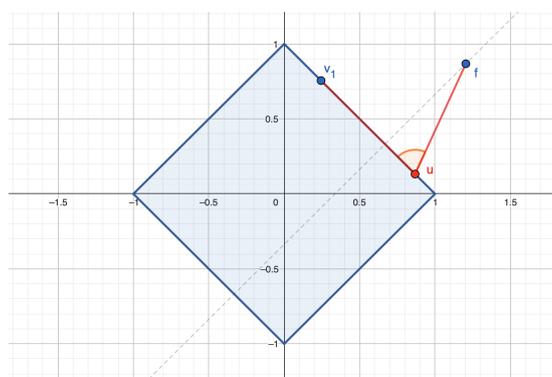


Figura 2.2: Ejemplo proyección errónea.

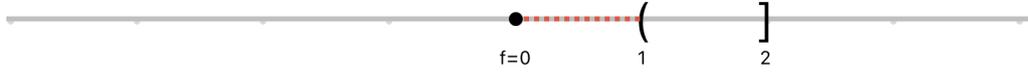
A continuación, analizaremos la importancia de que los conjuntos sobre los que se proyecta sean **cerrados y convexos**. Por medio de dos ejemplos, mostraremos que, si estas condiciones no se cumplen, la existencia o unicidad de la proyección puede no estar garantizada.

**Ejemplo 2.5. Ejemplos de proyecciones sobre conjuntos no cerrados y/o no convexos.**

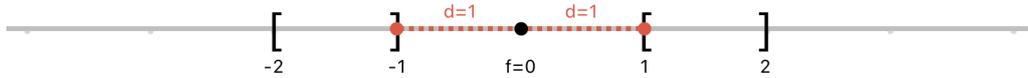
1. Sea  $H = \mathbb{R}$ ,  $\mathbb{K} = (1, 2]$ . Sea  $f = 0$ .  $\mathbb{K}$  es convexo, pero no cerrado. Observamos que la distancia entre  $f$  y el conjunto  $\mathbb{K}$  es

$$d(f, \mathbb{K}) = \inf_{x \in \mathbb{K}} |f - x|$$

que visualmente es 1. Sin embargo, este ínfimo no se alcanza en ningún punto de  $K$  (ya que el intervalo es abierto por la izquierda), luego no existe la proyección.



2. Sea  $H = \mathbb{R}$ ,  $\mathbb{K} = [-2, -1] \cup [1, 2]$ . Sea  $f = 0$ .  $\mathbb{K}$  es cerrado, pero no es convexo. En este caso, la solución no es única, ya que los puntos a mínima distancia son  $\{-1, 1\}$  que se encuentran a distancia 1 del punto.



Con estos dos sencillos ejemplos vemos que, si el conjunto no es cerrado o no es convexo, puede no existir solución, o no ser única. Sin embargo, que el conjunto sea convexo y cerrado es condición suficiente para la existencia de solución única, como hemos probado en el Teorema 2.3.

A continuación, analizaremos el caso particular en que el conjunto sobre el que se proyecta es un subespacio afín cerrado. Esta situación presenta propiedades adicionales que permiten simplificar notablemente el cálculo explícito de las proyecciones. Dentro de este corolario, se incluyen los subespacios vectoriales cerrados, simplemente tomando  $u_0 = 0$ .

**Corolario 2.6.** Sea  $H$  un espacio de Hilbert. Sea  $\mathbb{K} = u_0 + S \subset H$  un subespacio afín cerrado, donde  $S$  es un subespacio vectorial cerrado. Entonces, existe un único punto  $\bar{u} = P_{u_0+S}(f) \in u_0 + S$  tal que

$$\|f - \bar{u}\| = \min_{v \in u_0+S} \|f - v\|, \quad \text{para todo } f \in H.$$

Este punto se caracteriza por la propiedad

$$\begin{cases} \bar{u} \in u_0 + S \\ (f - \bar{u}, v) = 0 \quad \forall v \in S. \end{cases}$$

En el caso particular en el que  $u_0 = 0$ , se recupera el resultado clásico de proyección ortogonal sobre subespacios vectoriales cerrados:

$$\bar{u} = P_S f \in S, \quad \text{con } (f - \bar{u}, v) = 0 \quad \forall v \in S.$$

Además, el operador de proyección  $P_S$  es lineal y continuo.

*Demostración.*  $\implies$  Tenemos de (2.2) que  $(f - \bar{u}, w - \bar{u}) \leq 0 \quad \forall w \in \mathbb{K}$ . Por lo tanto, si  $\mathbb{K} = u_0 + S$  tenemos

$$(f - \bar{u}, u_0 + v - \bar{u}) \leq 0 \quad \forall v \in S$$

y por tanto

$$\begin{aligned} (f - \bar{u}, u_0 + tv - \bar{u}) &\leq 0 \quad \forall v \in S \quad \forall t \in \mathbb{R} \\ \iff (f, tv) - (\bar{u}, tv) + (f, u_0) - (\bar{u}, u_0) + \|\bar{u}\|^2 - (f, \bar{u}) &\leq 0 \quad \forall v \in S \quad \forall t \in \mathbb{R} \\ \iff t[(f, v) - (\bar{u}, v)] + (f, u_0) - (\bar{u}, u_0) + \|\bar{u}\|^2 - (f, \bar{u}) &\leq 0 \quad \forall v \in S \quad \forall t \in \mathbb{R} \end{aligned}$$

que es una expresión de la forma  $tA + B \quad \forall t \in \mathbb{R}$ . Como  $t$  puede tomar cualquier valor de  $\mathbb{R}$ , si  $A < 0$  tomando  $t \rightarrow -\infty$  la expresión crecería sin cota, violando la desigualdad; y si  $A > 0$  tomando  $t \rightarrow +\infty$  ocurriría lo mismo, con lo que  $A$  tiene que ser igual a cero y tenemos

$$(f, v) - (\bar{u}, v) = (f - \bar{u}, v) = 0 \quad \forall v \in S.$$

$\Leftarrow$ ) Si  $\bar{u}$  verifica  $(f - \bar{u}, v) = 0 \quad \forall v \in S$ .

Como  $\bar{u} \in \mathbb{K} = u_0 + S$ , entonces  $\bar{u} = u_0 + m$  con  $m \in S$ , y como  $\forall v \in S$  se tiene  $(f - \bar{u}, v) = 0$ , entonces particularmente  $(f - \bar{u}, m) = 0$  y por lo tanto

$$\begin{aligned} (f - \bar{u}, v) - (f - \bar{u}, m) &= 0 \Rightarrow (f - \bar{u}, v - m) = 0 \\ \Rightarrow (f - \bar{u}, u_0 + v - u_0 - m) &= 0 \Rightarrow (f - \bar{u}, w - \bar{u}) = 0 \quad \forall w \in \mathbb{K} \end{aligned}$$

y cumple (2.2).

Por último, veamos que  $P_S(f)$  es un operador lineal. Para ello, deben cumplirse las dos siguientes propiedades:

1. Aditividad:  $P_S(f_1 + f_2) = P_S f_1 + P_S f_2 \quad \forall f_1, f_2 \in H$
2. Homogeneidad:  $P_S(\lambda f) = \lambda P_S f \quad \forall f_1 \in H \quad \lambda \in \mathbb{R}$

1. Sea  $\bar{u}_1 = P_S f_1$ ,  $\bar{u}_2 = P_S f_2$ . Entonces sabemos que  $(f_1 - \bar{u}_1, v) = 0$  y  $(f_2 - \bar{u}_2, v) = 0 \quad \forall v \in S$ . Llamamos  $\bar{u} = P_S(f_1 + f_2)$ . Entonces,

$$(f_1 + f_2 - \bar{u}, v) = 0 = (f_1 - \bar{u}_1, v) + (f_2 - \bar{u}_2, v) \quad \forall v \in S$$

Por tanto,

$$P_S(f_1 + f_2) = \bar{u}_1 + \bar{u}_2 = P_S f_1 + P_S f_2.$$

2. Sea  $\bar{u} = P_S f$  y  $\lambda \in \mathbb{R}$ . Suponemos que  $u' = \lambda \bar{u}$  y por tanto,

$$(\lambda f - u', v) = \lambda(f - \bar{u}, v) = 0 \quad \forall \lambda \in \mathbb{R}.$$

Esto muestra que  $\lambda \bar{u}$  también satisface la condición de ortogonalidad con respecto a  $\lambda f$ , y por la unicidad de la proyección sobre  $S$ , se concluye que  $P_S(\lambda f) = \lambda P_S f \quad \forall f \in H, \lambda \in \mathbb{R}$ .

Con ello concluimos que es operador lineal. □

**Interpretación geométrica:**

Hagamos nuevamente una interpretación geométrica del Corolario 2.6. En el caso de los subespacios vectoriales cerrados ( $u_0 = 0$ ), presentamos un ejemplo considerando  $S$  como la recta  $y = x$ ,  $S \subset \mathbb{R}^2$ . Realizando el producto escalar, obtenemos

$$(f - \bar{u}, v) = \underbrace{\|f - \bar{u}\|}_{>0} \cdot \underbrace{\|v\|}_{>0} \cdot \underbrace{\cos(\theta)}_{=0} = 0 \quad \forall v \in S.$$

En este caso,  $\theta$  siempre es un ángulo recto, por lo que su coseno será siempre cero y se cumple la propiedad del Corolario 2.6.

Tomando cualquier otro punto distinto de la proyección, el ángulo deja de ser recto, como podemos ver en la Figura 2.4, siendo el coseno distinto de cero. Por tanto  $\bar{u}$  es el único punto que cumple esta propiedad.

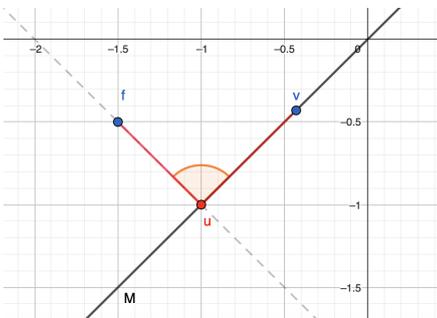


Figura 2.3: Proyección sobre un subespacio cerrado en  $\mathbb{R}^2$ .

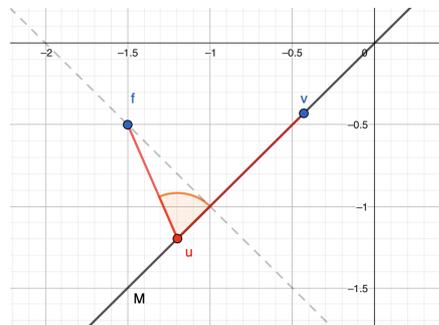


Figura 2.4: Ejemplo proyección errónea.

Para la explicación gráfica de los subespacios afines cerrados, explicamos la forma de calcularlos a partir de los subespacios vectoriales, la cual nos mostrará de forma más visual la conexión y similitud de ambos casos, únicamente realizando una traslación.

Sea  $\mathbb{K} \subset H$  un subespacio **afín cerrado no vacío**. Entonces, por definición, existe un punto  $u_0 \in H$  y un **subespacio vectorial cerrado**  $S \subset H$  tal que  $\mathbb{K} = u_0 + S$ . Dado un elemento  $f \in H$ , para calcular su proyección ortogonal sobre el subespacio afín  $\mathbb{K}$ , se procede de la siguiente manera:

1. Se traslada el problema al subespacio  $S$ , considerando el vector

$$f' = f - u_0.$$

2. Se calcula la proyección ortogonal de  $f'$  sobre el subespacio  $S$ , utilizando el corolario correspondiente a subespacios vectoriales cerrados 2.6. Es decir, se obtiene el único  $w \in S$  tal que

$$(f' - w, v) = 0 \quad \forall v \in S.$$

Esto es,  $w = P_S(f - u_0)$ , donde  $P_S$  es la proyección ortogonal sobre  $S$ .

3. Finalmente, se traslada el resultado de vuelta al subespacio afín  $\mathbb{K}$ , sumando el vector  $u_0$ , y se obtiene

$$\bar{u} = P_{\mathbb{K}}f = u_0 + P_S(f - u_0).$$

De esta forma, la proyección ortogonal sobre un subespacio afín cerrado se reduce al cálculo de una proyección sobre un subespacio vectorial cerrado, junto con un cambio de origen adecuado. Por ejemplo, podemos considerar,

$$\mathbb{K} = \{(x, x + 2); x \in \mathbb{R}\} = (0, 2) + \{(x, x); x \in \mathbb{R}\} = u_0 + S$$

y calcular la proyección sobre el subespacio vectorial  $S$  para luego trasladarlo de nuevo al subespacio afín  $\mathbb{K}$ . A continuación, se presenta el ejemplo de forma gráfica.

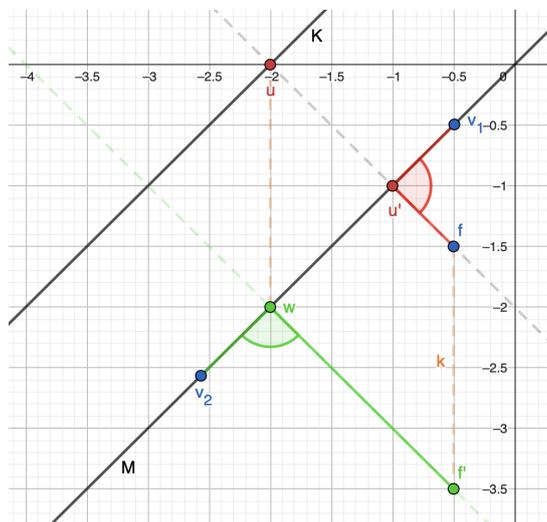


Figura 2.5: Proyección sobre un subespacio afín en  $\mathbb{R}^2$ .

**Proposición 2.7.** En las hipótesis del Teorema de la Proyección sobre un convexo cerrado 2.3, se verifica

$$\|P_{\mathbb{K}}f_1 - P_{\mathbb{K}}f_2\| \leq \|f_1 - f_2\| \quad \forall f_1, f_2 \in H. \quad (2.3)$$

es decir,  $P_{\mathbb{K}}$  es Lipschitz continua con constante 1.

*Demostración.* Tomando  $u_1 = P_{\mathbb{K}}f_1$  y  $u_2 = P_{\mathbb{K}}f_2$  obtenemos

$$\begin{aligned} (f_1 - u_1, v - u_1) &\leq 0 \quad \forall v \in \mathbb{K} \\ (f_2 - u_2, v - u_2) &\leq 0 \quad \forall v \in \mathbb{K} \end{aligned}$$

Sustituyendo  $v = u_2$  y  $v = u_1$  respectivamente en las desigualdades anteriores y sumando ambas, obtenemos

$$\begin{aligned} (f_1 - u_1, u_2 - u_1) + (f_2 - u_2, u_1 - u_2) &\leq 0 \implies (f_1 - u_1, u_2 - u_1) - (f_2 - u_2, u_2 - u_1) \leq 0 \implies \\ ((f_1 - f_2) - (u_1 - u_2), u_2 - u_1) &\leq 0 \implies ((f_1 - f_2) - (u_1 - u_2), u_1 - u_2) \geq 0 \end{aligned}$$

y expandiendo el producto escalar

$$(f_1 - f_2, u_1 - u_2) - \|u_1 - u_2\|^2 \geq 0 \implies \|u_1 - u_2\|^2 \leq (f_1 - f_2, u_1 - u_2).$$

Aplicando la desigualdad de Cauchy-Schwarz A.6,

$$(f_1 - f_2, u_1 - u_2) \leq \|f_1 - f_2\| \|u_1 - u_2\|$$

con lo que obtenemos

$$\|u_1 - u_2\|^2 \leq \|f_1 - f_2\| \|u_1 - u_2\|$$

y por consiguiente  $\|u_1 - u_2\| \leq \|f_1 - f_2\|$ .

□

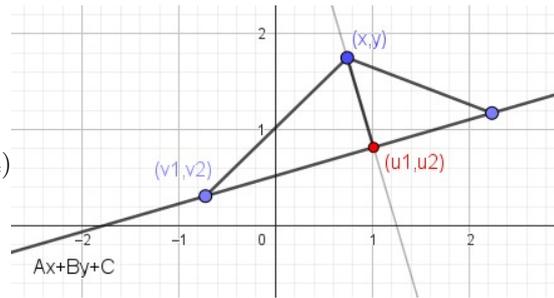
## 2.2. Dimensión finita

En esta sección, deduciremos las fórmulas clásicas de las distancias punto-recta en el plano y recta-plano en el espacio que se estudian en el Bachillerato. Estas fórmulas también pueden demostrarse de manera más sencilla calculando el punto de intersección entre la recta (o el plano) y la recta perpendicular a ella que pasa por el punto dado; posteriormente, se calcula la distancia entre ambos puntos. Sin embargo, en este caso se presenta una demostración basada en el Teorema de la Proyección. El trabajo se ha realizado de forma autónoma, sin uso de bibliografía, a partir del Corolario 2.6.

**Observación 2.8.** *Deducción de las fórmulas de las distancias.*

1. *Distancia punto-recta en el plano.*

$$d(p, r) = \frac{|Ap_1 + Bp_2 + C|}{\sqrt{A^2 + B^2}} \quad (2.4)$$



*Demostración.* Sea  $H = \mathbb{R}^2$ . Dada una recta en forma general  $Ax + By + C = 0$ , definimos

$$\mathbb{K} = \left\{ \left( x, \frac{-C - Ax}{B} \right) : x \in \mathbb{R} \right\} = \left( 0, \frac{-C}{B} \right) + \left\{ \left( x, \frac{-Ax}{B} \right) : x \in \mathbb{R} \right\} = u_0 + S$$

donde  $S$  es un subespacio vectorial. Aplicando Corolario 2.6, si denotamos a  $\hat{u} = P_S(p - u_0)$ , obtenemos

$$\sum_{i=1}^2 (p_i - u_{0i} - \hat{u}_i)v_i = 0 \quad \forall v \in S$$

$$(p_1 - \hat{u}_1)v_1 + \left(p_2 + \frac{C}{B} + \frac{A}{B}\hat{u}_1\right)\left(-\frac{A}{B}v_1\right) = v_1\left(p_1 - \hat{u}_1 - \frac{A}{B}p_2 - \frac{AC}{B^2} - \frac{A^2}{B^2}\hat{u}_1\right) = 0 \quad \forall v_1 \in \mathbb{R}$$

Como  $v_1 \in \mathbb{R}$ , el coeficiente que lo acompaña tiene que ser igual a 0 para que se cumpla la igualdad, por lo tanto

$$p_1 - \hat{u}_1 - \frac{A}{B}p_2 - \frac{AC}{B^2} - \frac{A^2}{B^2}\hat{u}_1 = 0$$

$$\hat{u}_1 = \frac{p_1 - \frac{A}{B}p_2 - \frac{AC}{B^2}}{1 + \frac{A^2}{B^2}} = \frac{B^2p_1 - ABp_2 - AC}{A^2 + B^2}$$

Calculamos  $\hat{u}_2$

$$\hat{u}_2 = \frac{-A\hat{u}_1}{B} = \frac{-AB^2p_1 + A^2Bp_2 + A^2C}{B(A^2 + B^2)}$$

y ahora calculamos  $\bar{u}$

$$\bar{u}_1 = \hat{u}_1 \quad \text{y} \quad \bar{u}_2 = \hat{u}_2 - \frac{C}{B} = \frac{-ABp_1 + A^2p_2 - BC}{A^2 + B^2}$$

y una vez tenemos  $\bar{u}$  la distancia euclídea es

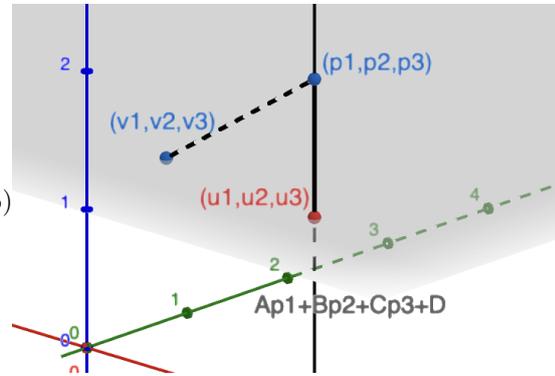
$$d((p_1, p_2)(\bar{u}_1, \bar{u}_2)) = \sqrt{(p_1 - \bar{u}_1)^2 + (p_2 - \bar{u}_2)^2} = \sqrt{\left(\frac{A^2p_1 + ABp_2 + AC}{A^2 + B^2}\right)^2 + \left(\frac{ABp_1 + B^2p_2 + BC}{A^2 + B^2}\right)^2}$$

$$= \frac{1}{\sqrt{A^2 + B^2}} \sqrt{(Ap_1^2 + Bp_2^2 + C)^2} = \frac{|Ap_1 + Bp_2 + C|}{\sqrt{A^2 + B^2}}$$

□

## 2. Distancia punto-plano en el espacio.

$$d(p, \pi) = \frac{|Ap_1 + Bp_2 + Cp_3 + D|}{\sqrt{A^2 + B^2 + C^2}} \quad (2.5)$$



La demostración de este caso se hace de forma análoga al anterior, pero por falta de espacio no la incluimos en el documento.

A continuación, introducimos el espacio  $l_2(t_1, t_2, \dots, t_n)$  el cual representa un espacio euclídeo de dimensión finita, estructurado en torno a un conjunto de puntos dados  $t_1, t_2, \dots, t_n$ . Veremos además que es isomorfo a  $\mathbb{R}^n$ , lo que nos ayudará a interpretar geoméricamente algunos resultados.

**Definición 2.9.** Dado un conjunto finito de puntos  $S = \{t_1, t_2, \dots, t_n\}$ , definimos  $l_2(S)$  como el conjunto de funciones  $f : S \rightarrow \mathbb{R}$  con el producto interno

$$(f, g) = \sum_{i=1}^n f(t_i)g(t_i).$$

Su norma está dada por

$$\|f\| = \sqrt{\sum_{i=1}^n |f(t_i)|^2}$$

Observamos que el espacio  $l_2(S)$  es isomorfo  $\mathbb{R}^n$ . Demostremos esta correspondencia.

Definimos la aplicación biyectiva

$$\Phi : l_2(S) \rightarrow \mathbb{R}^n, \quad f \mapsto (f(t_1), f(t_2), \dots, f(t_n))$$

es decir, que a cada función  $f \in l_2(S)$  le asigna el vector en  $\mathbb{R}^n$  cuyas componentes son los valores de  $f$  evaluados en cada punto del conjunto  $S = \{t_1, t_2, \dots, t_n\}$ . Veamos que cumple las siguientes propiedades.

**1. Linealidad:** Sean  $f, g \in l_2(S)$  y  $\alpha \in \mathbb{R}$ . Entonces

$$\begin{aligned} \Phi(\alpha f + g) &= ((\alpha f + g)(t_1), \dots, (\alpha f + g)(t_n)) = \\ &= (\alpha f(t_1) + g(t_1), \dots, \alpha f(t_n) + g(t_n)) = \alpha \Phi(f) + \Phi(g) \end{aligned}$$

y, por lo tanto,  $\Phi$  es lineal.

**2. Inyectividad:** Supongamos que  $\Phi(f) = \Phi(g)$ . Entonces

$$f(t_i) = g(t_i) \quad \text{para todo } i = 1, \dots, n.$$

Como las funciones están definidas únicamente en  $S = \{t_1, \dots, t_n\}$ , se concluye que  $f = g$ . Por lo tanto,  $\Phi$  es inyectiva.

**3. Sobreyectividad:** Sea  $(a_1, \dots, a_n) \in \mathbb{R}^n$ . Definimos una función  $f : S \rightarrow \mathbb{R}$  por  $f(t_i) := a_i$ . Entonces,  $f \in l_2(S)$ , y se cumple

$$\Phi(f) = (f(t_1), \dots, f(t_n)) = (a_1, \dots, a_n),$$

por lo tanto,  $\Phi$  es sobreyectiva.

**4. Preservación del producto interno:** Sean  $f, g \in l_2(S)$ . Entonces

$$(f, g) = \sum_{i=1}^n f(t_i)g(t_i) = (\Phi(f), \Phi(g))_{\mathbb{R}^n}$$

Como  $\Phi$  es lineal, biyectiva y preserva el producto interno, concluimos que es un isomorfismo de espacios.

### 2.2.1. Problemas mínimos cuadrados

En este apartado, estudiaremos la mejor aproximación en un espacio de Hilbert mediante elementos de un subespacio de dimensión finita. Estudiaremos tres problemas mediante métodos clásicos y mediante proyecciones. Los problemas y la teoría asociadas a esa sección han sido obtenidos de [2].

**Corolario 2.10. Ecuaciones normales.**

Sea  $\{x_1, x_2, \dots, x_n\}$  una base para el subespacio  $n$ -dimensional  $\mathbb{K}$  de  $X$  un espacio de Hilbert. Entonces, para cada  $x \in X$

$$P_{\mathbb{K}}(x) = \sum_{i=1}^n \alpha_i x_i \tag{2.6}$$

donde los  $\alpha_i$  son la única solución de las ecuaciones normales

$$\sum_{i=1}^n \alpha_i (x_i, x_j) = (x, x_j) \quad j = 1, 2, \dots, n.$$

En particular, si  $\{x_1, x_2, \dots, x_n\}$  es una base ortonormal de  $\mathbb{K}$ , entonces

$$P_{\mathbb{K}}(x) = \sum_{i=1}^n (x, x_i) x_i \quad \text{para todo } x \in X.$$

*Demostración.* En primer lugar, como  $\mathbb{K}$  es un subespacio de dimensión finita, por el Teorema A.7 sabemos que es un subespacio cerrado. Dado  $x \in X$  e  $y_0 \in \mathbb{K}$ , sabemos que  $y_0 = \sum_{i=1}^n \alpha_i x_i$  para algunos  $\alpha_i$ . Por el Corolario 2.6, tenemos que  $y_0 = P_{\mathbb{K}}(x)$  es equivalente a  $(x - y_0, y) = 0$  para cada  $y \in \mathbb{K}$  y se cumple

$$(x - y_0, x_j) = 0 \quad \text{para } j = 1, 2, \dots, n.$$

Por el proceso de ortogonalización de Gram-Schmidt sabemos que toda base se puede transformar en una base ortonormal, por lo tanto podemos considerar  $\{x_1, x_2, \dots, x_n\}$  como una base ortonormal, y tenemos que  $(x_i, x_j) = \delta_{ij}$  donde

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Si sustituimos  $(x_i, x_j) = \delta_{ij}$  en (2.6), obtenemos  $\alpha_j = (x, x_j) \forall j = 1, 2, \dots, n$ , y por lo tanto, sustituyendo de nuevo en (2.6), llegamos a que la proyección ortogonal es

$$P_{\mathbb{K}}(x) = \sum_{i=1}^n (x, x_i) x_i \quad \text{para todo } x \in X.$$

□

Veamos un ejemplo en que las funciones base están dadas por  $x_j = \sin(\frac{j\pi x}{L})$ .

**Ejemplo 2.11.** Sea  $H = L^2[0, \pi]$  y sea  $\mathbb{K} = \{f : f = a_1 \sin(x) + a_2 \sin(2x) + a_3 \sin(3x); a_1, a_2, a_3 \in \mathbb{R}\}$ . Sea  $f \in H$ , calcular la proyección de  $f$  sobre  $\mathbb{K}$ .

Claramente,  $\mathbb{K}$  es un subespacio cerrado. De nuevo, como  $\mathbb{K}$  es un subespacio de dimensión finita, por el Teorema A.7, sabemos que es cerrado.

Como  $\mathbb{K}$  es un subespacio cerrado, podemos emplear el Corolario 2.6

$$(f - P_{\mathbb{K}}(f), \sin(kx)) = 0 \quad \text{para cada } k = 1, 2, 3$$

por ser el conjunto  $\{\sin(x), \sin(2x), \sin(3x)\}$  una base ortogonal de  $\mathbb{K}$ , es decir

$$\int_0^\pi \left( f(x) - \sum_{i=1}^3 a_i \sin(ix) \right) \sin(kx) dx = 0 \implies \int_0^\pi f(x) \sin(kx) dx - \sum_{i=1}^3 a_i \int_0^\pi \sin(ix) \sin(kx) dx = 0.$$

Sabemos por la ortogonalidad de la función seno en  $[0, \pi]$  que se cumple

$$\int_0^\pi \sin(ix) \sin(kx) dx = \begin{cases} 0 & \text{si } i \neq k \\ \frac{\pi}{2} & \text{si } i = k \end{cases}$$

Por lo tanto

$$\int_0^\pi f(x) \sin(kx) dx - \sum_{i=1}^3 a_i \int_0^\pi \sin(ix) \sin(kx) dx = 0 \implies a_k \frac{\pi}{2} = \int_0^\pi f(x) \sin(kx) dx$$

y despejando  $a_k$  tenemos

$$a_k = \frac{2}{\pi} \int_0^\pi f(x) \sin(kx) dx.$$

**Problema 1: Mejor aproximación polinómica a un conjunto de datos.**

Sea  $\{(t_j, y_j) : j = 1, 2, \dots, m\}$  un conjunto de datos. Para cualquier entero  $n < m$  queremos encontrar el polinomio  $p(x) = \sum_{i=0}^n \alpha_i x^i$ , de grado como mucho  $n$  tal que minimice la expresión  $\sum_{k=1}^m [y_k - p(t_k)]^2$ .

**Resolución por técnicas clásicas del Cálculo Diferencial.**

Queremos minimizar el error cuadrático total dado por

$$E(\alpha_0, \alpha_1, \dots, \alpha_n) = \sum_{j=1}^m [y_j - p(t_j)]^2 = \sum_{j=1}^m \left[ y_j - \sum_{i=0}^n \alpha_i t_j^i \right]^2$$

Para determinar los valores óptimos de  $\alpha_0, \alpha_1, \dots, \alpha_n$  que minimizan la función de error  $E(\alpha_0, \alpha_1, \dots, \alpha_n)$ , observamos que se trata de una función cuadrática, por lo tanto, es una función convexa. Por las condiciones de optimalidad, sabemos que el mínimo global se alcanza en el punto donde el gradiente de  $E$  se anula, es decir, donde las derivadas parciales son cero. Calculamos las derivadas parciales

$$\frac{\partial E}{\partial \alpha_k} = -2 \sum_{j=1}^m t_j^k \left( y_j - \sum_{i=0}^n \alpha_i t_j^i \right).$$

e igualando a cero obtenemos el sistema

$$\sum_{i=0}^n \alpha_i \left( \sum_{j=1}^m t_j^{i+k} \right) = \sum_{j=1}^m y_j t_j^k, \quad \text{para } k = 0, 1, \dots, n.$$

que es un sistema lineal de  $n + 1$  ecuaciones con  $n + 1$  incógnitas. Veamos el siguiente ejemplo.

**Ejemplo 2.12.** *Dados los puntos (1,2) (2,3) y (3,5), veamos cuál es la mejor aproximación al modelo lineal  $p(x) = \alpha_1 x + \alpha_0$ . Planteamos la función de error*

$$E(\alpha_1, \alpha_0) = \sum_{i=1}^3 (y_i - (\alpha_1 t_i + \alpha_0))^2 = (2 - (\alpha_1 \cdot 1 + \alpha_0))^2 + (3 - (\alpha_1 \cdot 2 + \alpha_0))^2 + (5 - (\alpha_1 \cdot 3 + \alpha_0))^2.$$

Calculamos las derivadas parciales respecto de  $\alpha_0$  y  $\alpha_1$ .

$$\frac{\partial E}{\partial \alpha_1} = -2 \sum_{i=1}^3 t_i (y_i - (\alpha_1 t_i + \alpha_0)) = 2\alpha_1 \sum_{i=1}^3 t_i^2 + 2\alpha_0 \sum_{i=1}^3 t_i - 2 \sum_{i=1}^3 t_i y_i = 0$$

$$\frac{\partial E}{\partial \alpha_0} = -2 \sum_{i=1}^3 (y_i - (\alpha_1 t_i + \alpha_0)) = 2\alpha_1 \sum_{i=1}^3 t_i + 6\alpha_0 - 2 \sum_{i=1}^3 y_i = 0$$

y sustituyendo llegamos al sistema

$$\begin{cases} 14\alpha_1 + 6\alpha_0 = 23 \\ 6\alpha_1 + 3\alpha_0 = 10 \end{cases}$$

De donde se obtiene  $\alpha_1 = \frac{3}{2}$  y  $\alpha_0 = \frac{1}{3}$ . Luego la recta es  $p(x) = \frac{3}{2}x + \frac{1}{3}$ .

**Resolución por proyecciones.**

Vemos que este problema se puede convertir en el cálculo de una proyección. Dado el conjunto de datos  $\{(t_j, y_j) : j = 1, 2, \dots, m\}$  queremos encontrar el polinomio  $p(x) = \sum_{i=0}^n \alpha_i x^i$  que minimice el error cuadrático entre  $p(t)$  y los puntos dados. Dicho polinomio puede interpretarse como la proyección ortogonal de  $y$  en el subespacio  $S$  de los polinomios de grado  $n$ .

Sea  $H = l_2(t_1, t_2, \dots, t_m)$ . Consideramos  $\mathbb{K} = \{f \in H : \exists p \in S \text{ tal que } f(t_j) = p(t_j), \forall j = 1, 2, \dots, m\} \subset H$ . Como  $S$  es un subespacio de dimensión finita, y la evaluación de los puntos es una operación lineal,  $\mathbb{K}$  es un subespacio de dimensión finita y por el Teorema A.7 sabemos que es un subespacio cerrado. Por tanto, si consideramos la base de  $\mathbb{K}$  generada por las funciones  $\phi_j(t) = t^j$  evaluadas en los puntos  $t_1, t_2, \dots, t_m$ ; podemos aplicar el Corolario 2.6 y obtenemos

$$(y - p, \phi_j) = 0 \quad \text{para cada } j = 0, 1, \dots, n.$$

Sustituyendo  $p = \sum_{i=0}^n \alpha_i \phi_i$  obtenemos

$$\left(y - \sum_{i=0}^n \alpha_i \phi_i, \phi_j\right) = 0 \quad \text{para cada } j = 0, 1, \dots, n$$

lo que equivale a

$$\sum_{i=0}^n \alpha_i (\phi_i, \phi_j) = (y, \phi_j) \quad \text{para cada } j = 0, 1, \dots, n$$

y desarrollando los productos internos llegamos a

$$\sum_{i=0}^n \alpha_i \sum_{k=1}^m t_k^{i+j} = \sum_{k=1}^m y_k t_k^j \quad \text{para } j = 0, 1, \dots, n$$

que es el mismo sistema que el obtenido con el método clásico.

### Problema 2: Resolución de una sistema de ecuaciones sobredeterminado.

Consideramos el sistema de  $m$  ecuaciones con  $n$  incógnitas, donde  $m > n$ ,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

que podemos expresar matricialmente como  $Ax = b$ . En general, este sistema no tiene solución, tiene más ecuaciones que incógnitas. El objetivo será buscar el vector  $x$  que minimice el error cuadrático, es decir, que minimice la expresión

$$E(x) = \sum_{i=1}^m \left( \sum_{j=1}^n a_{ij}x_j - b_i \right)^2 \quad (2.7)$$

### Resolución por técnicas clásicas del Cálculo Diferencial.

Para minimizar la función  $E(x) = \|Ax - b\|^2$ , observamos que se trata de una función cuadrática y, por lo tanto, convexa. El mínimo se alcanza en el punto donde el gradiente de  $E$  se anula. Calculamos dicho gradiente:

$$E(x) = x^T A^T A x - 2b^T A x + b^T b$$

$$\nabla E(x) = 2A^T A x - 2A^T b = 2A^T (Ax - b).$$

Igualando el gradiente a cero obtenemos

$$A^T A x = A^T b$$

que nos genera un sistema de ecuaciones cuadrado, el cual ya es un problema standard. La solución de dicho sistema minimiza el error cuadrático dado en la expresión (2.7). Veamos un ejemplo.

**Ejemplo 2.13.** Dado el sistema

$$\begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + x_2 = 1 \\ 3x_1 + 4x_2 = 7 \end{cases} \implies A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \end{pmatrix} \quad A^T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix}$$

Por lo tanto

$$A^T Ax = A^T b \implies \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix} \implies \begin{pmatrix} 14 & 16 \\ 16 & 21 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 26 \\ 35 \end{pmatrix}$$

cuya solución es  $x_1 \approx -0,36$  e  $x_2 \approx 1,94$ , que evidentemente no cumple las ecuaciones exactamente, pero sí minimiza su error cuadrático.

### Resolución por proyecciones.

De nuevo, podemos convertir este problema en el cálculo de una proyección.

Sea  $H = l_2(x_1, x_2, \dots, x_m)$  y  $\mathbb{K} = \{f_y \in H; y = Ax, x \in \mathbb{R}^n\} \subset H$ , donde  $f_y \in H$  tal que  $f_y(t_i) = y_i$  y  $A$  es una matriz real  $m \times n$ . Podemos considerar el problema como encontrar la proyección de  $b$  en  $\mathbb{K}$ . Para ello, escogemos cualquier  $x_0$  tal que  $Ax_0 = y_0$ . Como  $\mathbb{K}$  es un subespacio de dimensión finita, por el Teorema 2.6 sabemos que es un subespacio cerrado y, por el Corolario 2.6, tenemos que  $y_0 = P_{\mathbb{K}}(b)$  si y solo si  $(b - y_0, Ax) = 0$  para todo  $x \in \mathbb{R}^n$ . Trabajando con la matriz traspuesta de  $A$  tenemos

$$(y, Ax) = \sum_{i=1}^m y_i \left( \sum_{j=1}^n a_{ij} x_j \right) = \sum_{j=1}^n \left( \sum_{i=1}^m a_{ij} y_i \right) x_j = (A^T y, x).$$

Sabemos que  $y_0 = P_{\mathbb{K}}(b)$  es equivalente a  $(A^T(b - y_0), x) = 0$  para todo  $x \in \mathbb{R}^n$ , que es equivalente a  $A^T(b - y_0) = 0$ , es decir,  $A^T b = A^T y_0$ . Entonces  $y_0 \in \mathbb{K}$  y  $y_0 = Ax_0$  para algún  $x_0 \in \mathbb{R}^n$ . Por lo tanto,  $x_0$  es solución si y solo si

$$A^T Ax_0 = A^T b.$$

$A^T A$  es simétrica y semidefinida positiva luego el sistema siempre es compatible. Si la matriz  $A^T A$  no es singular, el sistema tiene una única solución. En el caso de que fuera singular, tendríamos más de una posible solución. En el Teorema Fundamental 2.3, habíamos demostrado que la proyección es única, y es cierto. Sin embargo, aquí no se afirma que hay más de un  $y_0$ , es decir, más de un punto de  $\mathbb{K}$  que esté a distancia mínima, sino que puede haber más de un punto  $x_0 \in \mathbb{R}^n$  tal que  $Ax_0 = y_0$ .

### Problema 3: Mejor aproximación polinómica a una función.

Sea  $y$  una función continua en el intervalo  $[a, b]$ , queremos encontrar la función polinómica  $p(x) = \sum_{i=0}^n \alpha_i x^i$  de grado menor o igual a  $n$  que minimice la expresión

$$E(\alpha_1, \alpha_2, \dots, \alpha_n) = \int_a^b [y(x) - p(x)]^2 dx.$$

### Resolución por técnicas clásicas del Calculo Diferencial.

Queremos que el polinomio  $p$  aproxime la función  $y$  con el menor error posible en  $L^2[a, b]$ . Para ello, se minimiza la integral del error cuadrático entre  $y$  y  $p$

$$\int_a^b \left[ y(x) - \sum_{i=0}^n \alpha_i x^i \right]^2 dx = \int_a^b \left[ y(x)^2 - 2y(x) \sum_{i=0}^n \alpha_i x^i + \left( \sum_{i=0}^n \alpha_i x^i \right)^2 \right] dx.$$

Derivando esta expresión respecto de cada  $\alpha_j$  e igualando a cero, obtenemos

$$\sum_{i=0}^n \alpha_i \int_a^b x^{i+j} dx = \int_a^b y(x) x^j dx \quad \text{para cada } j = 0, 1, \dots, n.$$

Esto nos da un sistema de ecuaciones normales, que al resolverlo nos proporciona los coeficientes  $\alpha_i$  de  $p$ . Veamos de nuevo un ejemplo.

**Ejemplo 2.14.** Supongamos que queremos aproximar  $y(x) = \sin(x)$  en el intervalo  $[0, 2]$  por un polinomio  $p(x) = \alpha_1 x + \alpha_0$ . Planteamos la ecuación del error

$$E(\alpha_1, \alpha_2) = \int_0^2 [\sin(x) - (\alpha_1 x + \alpha_0)]^2 dx =$$

Realizamos las derivadas parciales respecto a  $\alpha_1$  y  $\alpha_0$  e igualamos a cero

$$\begin{cases} \int_0^2 \sin(x) dx = \alpha_0 \int_0^2 1 dx + \alpha_1 \int_0^2 x dx & \implies 2\alpha_0 + 2\alpha_1 = 1 - \cos(2) \\ \int_0^2 x \sin(x) dx = \alpha_0 \int_0^2 x dx + \alpha_1 \int_0^2 x^2 dx & \implies 2\alpha_0 + \frac{8}{3}\alpha_1 = \sin(2) - 2\cos(2) \end{cases}$$

y resolviendo el sistema obtenemos  $\alpha_1 \approx 0,49$  y  $\alpha_0 \approx 0,22$ , siendo el polinomio  $p(x) = 0,49x + 0,22$ .

### Resolución por proyecciones.

Sea  $X = L^2[a, b]$  y  $S$  el subespacio de los polinomios de grado menor o igual a  $n$ . Interpretamos el problema como la proyección de  $y$  sobre  $S$ . Como hemos visto anteriormente,  $S$  es un subespacio cerrado. De nuevo, para que  $p$  sea la proyección, por el Corolario 2.6, tenemos que  $y - p$  debe cumplir

$$(y - p, x^j) = 0 \quad \text{para cada } j = 0, 1, \dots, n$$

es decir

$$\int_a^b [y(x) - \sum_{i=0}^n \alpha_i x^i] x^j dx = 0 \quad \text{para cada } j = 0, 1, \dots, n$$

llegando al mismo sistema de ecuaciones normales que anteriormente.

## 2.3. Dimensión infinita. Espacio $L^2[0, 1]$ .

En esta sección, trabajaremos una serie de ejemplos en el espacio  $L^2[0, 1]$ .  $\mathbb{K}_1$ ,  $\mathbb{K}_2$  y  $\mathbb{K}_3$  son ejemplos clásicos, mientras que  $\mathbb{K}_4$  ha sido estudiado, por ejemplo, en [4].

**Ejemplo 2.15.** Consideramos el espacio de Hilbert de funciones reales definido en A.8

$$L^2[0, 1] = \left\{ f : [0, 1] \longrightarrow \mathbb{R} \text{ medible} : \int_0^1 |f(x)|^2 dx < \infty \right\}$$

con el producto escalar

$$(f, g)_{L^2[0,1]} = \int_0^1 f(x)g(x) dx.$$

Consideramos los siguientes conjuntos:

$$\mathbb{K}_1 = \{ f \in L^2[0, 1] : a \leq f(x) \leq b \text{ c.t.p. } x \in [0, 1] \} \quad a, b \in \mathbb{R}.$$

$$\mathbb{K}_2 = \left\{ f \in L^2[0, 1] : \|f\|_{L^2[0,1]} \leq r \right\} \quad r > 0.$$

$$\mathbb{K}_3 = \left\{ f \in L^2[0, 1] : \int_0^1 f(x)p(x)dx \leq r \right\} \quad r \in \mathbb{R}, p \in L^2[0, 1], p > 0.$$

$$\mathbb{K}_4 = \mathbb{K}_1 \cap \mathbb{K}_3.$$

Veamos que  $\mathbb{K}_1$ ,  $\mathbb{K}_2$ ,  $\mathbb{K}_3$  y  $\mathbb{K}_4$  son subconjuntos convexos y cerrados de  $L^2[0, 1]$ .

Para probar que dichos subconjuntos son convexos, probaremos que dados  $f_1, f_2$  pertenecientes al conjunto, entonces  $\lambda f_1 + (1 - \lambda)f_2$  también pertenece al conjunto para todo  $\lambda$  en  $[0, 1]$ .

Dados  $f_1, f_2 \in \mathbb{K}_1$  sabemos que  $a \leq f_1(x), f_2(x) \leq b$  entonces, tenemos que

$$a = \lambda a + (1 - \lambda)a \leq \lambda f_1(x) + (1 - \lambda)f_2(x) \leq \lambda b + (1 - \lambda)b = b \quad \forall \lambda \in [0, 1]$$

Sean  $f_1, f_2 \in \mathbb{K}_2$  sabemos que  $\|f_1\|_{L^2[0,1]} \leq r$  y  $\|f_2\|_{L^2[0,1]} \leq r$ . Sabemos que  $x^2$  es una función convexa, por lo tanto

$$(\lambda f_1(x) + (1 - \lambda)f_2(x))^2 \leq \lambda f_1(x)^2 + (1 - \lambda)f_2(x)^2.$$

Por tanto, integrando a ambos lados, obtenemos

$$\begin{aligned} \int_0^1 (\lambda f_1(x) + (1 - \lambda)f_2(x))^2 dx &\leq \int_0^1 \lambda f_1(x)^2 + (1 - \lambda)f_2(x)^2 dx = \\ &= \lambda \int_0^1 f_1(x)^2 dx + (1 - \lambda) \int_0^1 f_2(x)^2 dx \leq \lambda r + (1 - \lambda)r = r. \end{aligned}$$

Dados  $f_1, f_2 \in \mathbb{K}_3$ , sabemos que  $\int_0^1 f_1(x)p(x)dx \leq r$  y  $\int_0^1 f_2(x)p(x)dx \leq r$ . Por lo tanto,

$$\begin{aligned} \int_0^1 (\lambda f_1(x) + (1 - \lambda)f_2(x))p(x)dx &= \int_0^1 (\lambda f_1(x)p(x) + (1 - \lambda)f_2(x)p(x))dx = \\ &= \int_0^1 \lambda f_1(x)p(x) dx + \int_0^1 (1 - \lambda)f_2(x)p(x) dx = \lambda \int_0^1 f_1(x)p(x) dx + (1 - \lambda) \int_0^1 f_2(x)p(x) dx \leq \\ &\leq \lambda r + (1 - \lambda)r = r. \end{aligned}$$

Además, como la intersección de convexos es convexo,  $\mathbb{K}_4$  también es convexo. Con todo ello, concluimos que los cuatro conjuntos son convexos. Veamos ahora que son cerrados.

$\mathbb{K}_1$  es cerrado en  $L^2[0, 1]$  si toda sucesión de funciones  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_1$  que converge a  $f \in L^2[0, 1]$  satisface que  $f \in \mathbb{K}_1$ . Si  $f_n \rightarrow f$ , entonces

$$\lim_{n \rightarrow \infty} \|f_n - f\| = 0$$

es decir,

$$\int_0^1 |f_n(x) - f(x)|^2 dx \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Dado que  $f_n \in \mathbb{K}_1$  para cada  $n$  entonces

$$a \leq f_n(x) \leq b \quad \text{c.t.p } x \in [0, 1].$$

Empleando el Teorema A.15, sabemos que existe una subsucesión  $\{f_{n_k}\}$  tal que  $f_{n_k}(x) \rightarrow f(x)$  c.t.p.  $x \in [0, 1]$ . Entonces, también tenemos que para casi todo  $x \in [0, 1]$

$$a \leq f_{n_k}(x) \leq b.$$

Por lo tanto, pasando al límite se deduce que  $a \leq f(x) \leq b$  c.t.p.  $x \in [0, 1]$ . Por lo tanto, el conjunto  $\mathbb{K}_1$  es cerrado en  $L^2[0, 1]$ .

Para  $\mathbb{K}_2$  procedemos de forma similar. Por la definición de  $\mathbb{K}_2$  sabemos que

$$0 \leq \|f_n\|_{L^2[0,1]} = \sqrt{\int_0^1 |f_n(x)|^2 dx} \leq r.$$

Como  $f_n$  converge a  $f$  entonces tenemos que

$$\|f\|_{L^2[0,1]} = \lim_{n \rightarrow \infty} (\|f_n\|_{L^2[0,1]}) \leq r$$

y concluimos que  $\mathbb{K}_2$  es cerrado.

Por su parte, sea  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_3$  una sucesión tal que  $f_n \rightarrow f$  en  $L^2[0, 1]$ , consideramos el funcional

$$\Phi(f) := \int_0^1 f(x)p(x) dx = (f, p)_{L^2}$$

que es lineal y continuo, ya que  $p \in L^2[0, 1]$  y, por la desigualdad de Cauchy-Schwarz A.6,

$$|\Phi(f)| = |(f, p)| \leq \|f\|_{L^2} \cdot \|p\|_{L^2}.$$

Entonces, por la continuidad de  $\Phi$ ,

$$\Phi(f_n) \rightarrow \Phi(f) \text{ es decir, } \int_0^1 f_n(x)p(x) dx \rightarrow \int_0^1 f(x)p(x) dx.$$

Como  $f_n \in \mathbb{K}_3$ , se cumple que

$$\int_0^1 f_n(x)p(x) dx \leq r \quad \forall n \in \mathbb{N}$$

pasando al límite

$$\int_0^1 f(x)p(x) dx = \lim_{n \rightarrow \infty} \int_0^1 f_n(x)p(x) dx \leq r$$

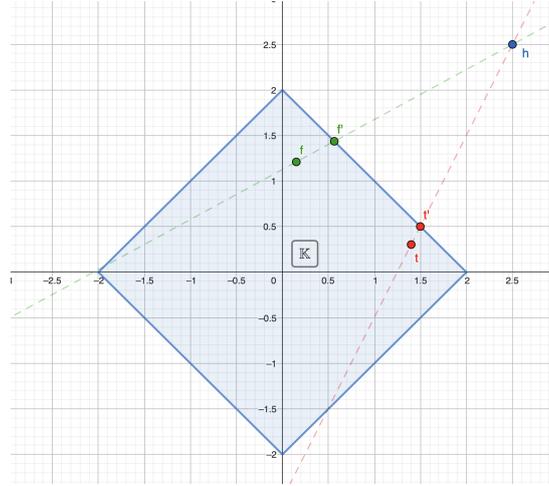
lo que implica que  $f \in \mathbb{K}_3$ , y por lo tanto,  $\mathbb{K}_3$  es cerrado en  $L^2[0, 1]$ .

Por último, como la intersección de cerrados es cerrado, sabemos que  $\mathbb{K}_4$  también es cerrado.

**Observación 2.16.** Sea  $\mathbb{K}$  el conjunto sobre el que queremos proyectar. Si  $h \in \mathbb{K}$ , es evidente que su proyección sobre  $\mathbb{K}$  es el propio  $h$  ya que

$$(f - \bar{u}, v - \bar{u}) = (h - h, v - h) = (0, v - h) = 0 \quad \forall v \in \mathbb{K}$$

es decir, la proyección de un elemento perteneciente al conjunto es él mismo. Por otro lado, si  $h \notin \mathbb{K}$ , podemos intuir que  $\bar{u}$  se va a encontrar en la "frontera" del conjunto ya que, dado cualquier punto que no pertenezca a la frontera, podríamos encontrar uno más cercano a  $h$  que pertenezca a  $\mathbb{K}$ . Veámoslo con el siguiente ejemplo en  $\mathbb{R}^2$ .

Figura 2.6: Proyección de un elemento que no pertenece a  $\mathbb{K}$ .

Por lo tanto, cuando  $h \notin \mathbb{K}$  presupondremos que la proyección se encuentra en la frontera, para luego comprobarlo rigurosamente.

Comprobemos ahora que las proyecciones sobre  $\mathbb{K}_1$  y  $\mathbb{K}_2$  son las siguientes:

$$P_{\mathbb{K}_1}(h(x)) = \begin{cases} a & \text{donde } h(x) < a \\ h(x) & \text{donde } a \leq h(x) \leq b = P_{[a,b]}(h(x)) \\ b & \text{donde } h(x) > b \end{cases}$$

$$P_{\mathbb{K}_2}(h) = \begin{cases} h & \text{si } \|h\|_{L^2[0,1]} \leq r \\ \frac{rh}{\|h\|_{L^2[0,1]}} & \text{si } \|h\|_{L^2[0,1]} > r \end{cases}$$

Comenzamos con  $\mathbb{K}_2$ . Si  $h \in \mathbb{K}_2$ , por la Observación 2.16,  $\bar{u}(x) = h(x)$ . Si  $h \notin \mathbb{K}_2$  entonces  $\|h\|_{L^2[0,1]} > r$  y  $\bar{u}(x) = \frac{rh(x)}{\|h\|_{L^2[0,1]}} \in \mathbb{K}_2$ .

$$\int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \int_0^1 \left( h(x) - \frac{rh(x)}{\|h\|_{L^2[0,1]}} \right) \left( v(x) - \frac{rh(x)}{\|h\|_{L^2[0,1]}} \right) dx \leq 0 \quad \forall v \in \mathbb{K}_2$$

$$\iff \int_0^1 (\|h\|_{L^2[0,1]} h(x) - rh(x)) (\|h\|_{L^2[0,1]} v(x) - rh(x)) dx \leq 0$$

$$\iff (\|h\|_{L^2[0,1]} - r) \int_0^1 h(x) (\|h\|_{L^2[0,1]} v(x) - rh(x)) dx \leq 0$$

y como  $\|h\|_{L^2[0,1]} - r > 0$

$$\iff \|h\|_{L^2[0,1]} \int_0^1 h(x)v(x) dx \leq r \int_0^1 h^2(x) dx = r \|h\|_{L^2[0,1]}^2 \iff \int_0^1 h(x)v(x) dx \leq r \|h\|_{L^2[0,1]} \quad \forall v \in \mathbb{K}_2.$$

Aplicando la desigualdad de Hölder, tenemos

$$\int_0^1 h(x)v(x) dx \leq \left| \int_0^1 h(x)v(x) dx \right| \leq \int_0^1 |h(x)| |v(x)| dx \leq \|h\|_{L^2[0,1]} \|v\|_{L^2[0,1]} \leq r \|h\|_{L^2[0,1]}$$

y por lo tanto se cumple.

Por otro lado, para  $\mathbb{K}_1$ , dividimos el intervalo  $[0, 1]$  de la siguiente manera,

$$[0, 1] = \Omega_0 \cup \Omega_+ \cup \Omega_- \quad (2.8)$$

donde

$$\Omega_0 = \{x \in [0, 1] : h(x) \in [a, b]\} \quad \Omega_+ = \{x \in [0, 1] : h(x) > b\} \quad \Omega_- = \{x \in [0, 1] : h(x) < a\}$$

que son conjuntos medibles por ser  $h$  medible. Por lo tanto, vemos que

$$\begin{aligned} & \int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \\ &= \int_{\Omega_0} (h(x) - h(x))(v(x) - h(x)) + \int_{\Omega_+} (h(x) - b)(v(x) - b) + \int_{\Omega_-} (h(x) - a)(v(x) - a) \leq 0 \quad \forall v \in \mathbb{K}_1. \end{aligned}$$

Con todo ello, concluimos que las proyecciones son correctas.

Veamos ahora cuáles son las proyecciones sobre  $\mathbb{K}_3$  y  $\mathbb{K}_4$ . Comenzamos con  $\mathbb{K}_3$ .

Si  $h \notin \mathbb{K}_3$  entonces  $\int_0^1 h(x)p(x)dx > r$  y por la Observación 2.16, la proyección parece que va a satisfacer  $\int_0^1 \bar{u}(x)p(x)dx = r$ . La función proyectada podría ser una corrección lineal de  $h$  en la dirección de  $p$  dada por  $\bar{u}(x) = h(x) - \gamma p(x)$ . Sustituyendo, obtenemos

$$\int_0^1 (h(x) - \gamma p(x))p(x) dx = r \Leftrightarrow \int_0^1 h(x)p(x) dx - \gamma \int_0^1 p(x)^2 dx = r$$

y despejando  $\gamma$  obtenemos

$$\gamma = \frac{\int_0^1 h(x)p(x) dx - r}{\|p\|_{L^2[0,1]}^2} > 0 \quad (2.9)$$

Por tanto, la proyección podría ser

$$\bar{u}(x) = h(x) - \gamma p(x).$$

Veamos que esto es cierto

$$\begin{aligned} & \int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \int_0^1 (h(x) - (h(x) - \gamma p(x)))(v(x) - (h(x) - \gamma p(x))) dx = \\ &= \int_0^1 \gamma p(x)(v(x) - h(x) + \gamma p(x)) dx = \int_0^1 \gamma p(x)v(x) dx - \int_0^1 \gamma p(x)h(x) dx + \int_0^1 \gamma^2 p(x)^2 dx \\ &= \gamma \left( \int_0^1 p(x)v(x) dx - \int_0^1 p(x)h(x) dx + \gamma \int_0^1 p^2(x) dx \right) \leq 0 \\ &\stackrel{\gamma \geq 0}{\Leftrightarrow} \int_0^1 p(x)v(x) dx - \int_0^1 p(x)h(x) dx + \int_0^1 p(x)h(x) dx - r \leq 0 \end{aligned}$$

ya que, por definición de  $\gamma$  (2.9), se tiene que  $\gamma \int_0^1 p^2(x) dx = \gamma \|p\|^2 = \int_0^1 p(x)h(x) dx - r$ . Por tanto, tenemos

$$\int_0^1 p(x)v(x) dx - r \leq 0 \quad \forall v \in \mathbb{K}_3$$

que se cumple por ser la definición de  $\mathbb{K}_3$ . Como la proyección es la única que cumple esta propiedad, concluimos que la proyección calculada es correcta.

Por último, veamos cuál es la proyección sobre  $\mathbb{K}_4$ .

$$\mathbb{K}_4 = \mathbb{K}_1 \cap \mathbb{K}_3 = \{h \in L^2[0, 1] : a \leq h(x) \leq b \text{ c.t.p. } x \in [0, 1]; \int_0^1 h(x)p(x) dx \leq r\}$$

En primer lugar,  $\mathbb{K}_4$  podría ser vacío si la intersección de  $\mathbb{K}_1$  y  $\mathbb{K}_2$  es vacía. En este caso, no existe proyección. Supongamos ahora que  $\mathbb{K}_4 \neq \emptyset$ , lo que significa que  $\int_0^1 ap(x) dx \leq \int_0^1 h(x)p(x) dx \leq r$ . Para cada  $h \in L^2[0, 1]$  tenemos los siguientes dos posibles casos.

1. Si  $\int_0^1 P_{[a,b]}(h(x))p(x) dx \leq r$ , entonces  $P_{\mathbb{K}_4}(h) = P_{[a,b]}(h)$ .
2. Si  $\int_0^1 P_{[a,b]}(h(x))p(x) dx > r$ , por la Observación 2.16, parece que debería existir  $\gamma > 0$  tal que

$$P_{\mathbb{K}_4}(h) = P_{[a,b]}(h - \gamma p) \quad \text{y} \quad \int_0^1 P_{\mathbb{K}_4}(h(x))p(x) dx = r.$$

Denotamos  $\bar{u} = P_{\mathbb{K}_4}(h)$  y sabemos que tiene que cumplir

$$\int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx \leq 0 \quad \forall v \in \mathbb{K}_4. \quad (2.10)$$

Consideramos la función real

$$g(s) = \int_0^1 P_{[a,b]}(h(x) - sp(x))p(x) dx - r$$

que es una función Lipschitz continua no creciente. Veámoslo. Queremos ver que existe  $L > 0$  tal que, para todo  $s_1, s_2 \in \mathbb{R}$

$$|g(s_1) - g(s_2)| \leq L |s_1 - s_2|.$$

Por tanto,

$$\begin{aligned} |g(s_1) - g(s_2)| &= \left| \int_0^1 (P_{[a,b]}(h(x) - s_1p(x)) - r - (P_{[a,b]}(h(x) - s_2p(x)))p(x) - r) dx \right| \leq \\ &\leq \int_0^1 |P_{[a,b]}(h(x) - s_1p(x)) - P_{[a,b]}(h(x) - s_2p(x))| |p(x)| dx. \end{aligned}$$

Como  $P_{[a,b]}$  es función Lipschitz (por ser proyección sobre un intervalo cerrado) se tiene

$$|P_{[a,b]}(u) - P_{[a,b]}(v)| \leq |u - v| \quad \forall u, v \in \mathbb{R}.$$

Aplicando esto al integrando

$$|P_{[a,b]}(h(x) - s_1p(x)) - P_{[a,b]}(h(x) - s_2p(x))| \leq |s_1 - s_2| \cdot |p(x)|$$

por tanto

$$|g(s_1) - g(s_2)| \leq |s_1 - s_2| \cdot \int_0^1 |p(x)|^2 dx = |s_1 - s_2| \cdot \|p\|_{L^2}^2.$$

y concluimos que es Lipschitz continua con constante  $L = \|p\|_{L^2}^2$ .

Continuando con la prueba, si  $g(0) \leq 0$ , que corresponde con el caso 1, vemos que  $\bar{u} = P_{[a,b]}(h)$  porque  $P_{[a,b]}(h) \in \mathbb{K}_4$ . Veamos que se cumple (2.10). Para ello, descompondremos el intervalo  $[0, 1]$  al igual que en (2.8). Entonces obtenemos

$$\int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx =$$

$$\int_{\Omega_0} (h(x)-h(x))(v(x)-h(x)) dx + \int_{\Omega_+} (h(x)-b)(v(x)-b) dx + \int_{\Omega_-} (h(x)-a)(v(x)-a) dx \leq 0 \quad \forall v \in \mathbb{K}_4.$$

Cuando  $g(0) > 0$ , que corresponde con el caso 2, podemos demostrar, aplicando el Teorema de la Convergencia Dominada A.15, que  $P_{[a,b]}(h - np) \rightarrow a$  cuando  $n \rightarrow \infty$ . Veámoslo.

$$P_{[a,b]}(h(x) - np(x)) = \begin{cases} a & \text{si } h(x) - np(x) < a \\ h(x) - np(x) & \text{si } h(x) - np(x) \in [a, b] \\ b & \text{si } h(x) - np(x) > b \end{cases}$$

observamos que, como  $p(x) > 0$ , entonces  $h(x) - np(x) \rightarrow -\infty$  en c.t.p. cuando  $n \rightarrow \infty$ . Dado que  $P_{[a,b]}$  proyecta cualquier valor por debajo de  $a$  al valor  $a$ , se tiene que

$$P_{[a,b]}(h(x) - np(x)) \rightarrow a \quad \text{en c.t.p. cuando } n \rightarrow \infty$$

y como  $p \in L^2[0, 1]$ , podemos aplicar el Teorema de la Convergencia Dominada A.13 y concluimos que

$$\int_0^1 P_{[a,b]}(h(x) - np(x))p(x) dx \rightarrow a \int_0^1 p(x) dx \quad \text{cuando } n \rightarrow \infty.$$

Por la definición de  $\mathbb{K}_4$ , como hemos visto anteriormente,  $a \int_0^1 p(x) dx < r$  (suponemos la desigualdad estricta), entonces

$$g(n) \rightarrow a \int_0^1 p(x) dx - r < 0 \quad \text{cuando } n \rightarrow \infty.$$

Como además

$$g(0) = \int_0^1 P_{[a,b]}(h(x))p(x) dx - r > 0$$

por hipótesis, y  $g$  es continua, por el Teorema de Bolzano existe  $\gamma > 0$  tal que  $g(\gamma) = 0$ . Es decir,

$$\int_0^1 P_{[a,b]}(h(x) - \gamma p(x))p(x) dx = r.$$

Por tanto, es claro que  $P_{[a,b]}(h - \gamma p) \in \mathbb{K}_4$ ,  $\bar{u}(x) = P_{[a,b]}(h - \gamma p)$  y de nuevo vemos que se cumple la desigualdad (2.10). En este caso, consideramos la descomposición  $[0, 1] = \bar{\Omega}_0 \cup \bar{\Omega}_+ \cup \bar{\Omega}_-$  donde

$$\begin{aligned} \bar{\Omega}_0 &= \{x \in [0, 1] : h(x) - \gamma p(x) \in [a, b]\} \\ \bar{\Omega}_+ &= \{x \in [0, 1] : h(x) - \gamma p(x) > b\} \\ \bar{\Omega}_- &= \{x \in [0, 1] : h(x) - \gamma p(x) < a\} \end{aligned}$$

Entonces

$$\begin{aligned} & \int_0^1 (h(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \\ & \int_{\bar{\Omega}_0} \gamma p(x)(v(x) - h(x) + \gamma p(x)) dx + \int_{\bar{\Omega}_-} (h(x) - a)(v(x) - a) dx + \int_{\bar{\Omega}_+} (h(x) - b)(v(x) - b) dx \leq \\ & \leq \int_{\bar{\Omega}_0} \gamma p(x)(v(x) - h(x) + \gamma p(x)) dx + \int_{\bar{\Omega}_-} \gamma p(x)(v(x) - a) dx + \int_{\bar{\Omega}_+} \gamma p(x)(v(x) - b) dx = \\ & = \gamma \int_0^1 p(x)(v(x) - \bar{u}(x)) dx = \gamma \left( \int_0^1 p(x)v(x) dx - r \right) \leq 0 \quad \forall v \in \mathbb{K}_4 \end{aligned}$$

con lo que queda comprobado.

## Capítulo 3

# Proyección sobre intersección de subespacios cerrados

### 3.1. Teorema de Von Neumann

Veamos ahora qué ocurre con la proyección sobre la intersección de dos subespacios cerrados. Este caso fue analizado por el matemático húngaro-estadounidense John von Neumann (1903-1957), del que toma nombre el algoritmo de proyecciones sucesivas, que servirá como método de cálculo. La teoría de esta sección se ha obtenido de [6].

**Teorema 3.1.** (*John von Neumann*)

Sean  $S_1$  y  $S_2$  dos subespacios cerrados de  $H$  un espacio de Hilbert, y sean  $P_1 : H \rightarrow S_1$  y  $P_2 : H \rightarrow S_2$  las proyecciones de  $H$  sobre  $S_1$  y  $S_2$  respectivamente. Sea  $x_0$  un punto arbitrario de  $H$ ,  $\forall n \in \{0, 1, 2, \dots\}$  se define la siguiente sucesión,

$$x_{2n+1} = P_1 x_{2n}; \quad x_{2n+2} = P_2 x_{2n+1}. \quad (3.1)$$

entonces, dicha sucesión converge en norma a  $Px_0$ , donde  $P : H \rightarrow S$  es la proyección de  $H$  en la intersección  $S = S_1 \cap S_2$ .

Para la demostración de este teorema, veamos previamente una serie de conceptos.

**Lema 3.2.** Sea  $\{x_n\}_{n \in \mathbb{N}}$  la sucesión definida en (3.1). Entonces

(a)  $(x_n, s) = (x_0, s) \quad \forall s \in S \text{ y } n \in \mathbb{N}$ .

(b)  $Px_n = Px_0 \quad \forall n \in \mathbb{N}$ .

(c) Si  $\{x_n\}_{n \in \mathbb{N}}$  converge débilmente a  $x_\infty$ , entonces  $x_\infty = Px_0$ .

*Demostración.* a) Razonamos por inducción.

$$(x_1, s) = (P_1 x_0, s) = (x_0, s).$$

Sabemos que  $(P_1 x_0, s) = (x_0, s)$  por el Corolario 2.6 ya que  $S_1$  y  $S_2$  son subespacios (mismo argumento en el resto de la demostración). Supongamos que se cumple para un cierto  $n \in \mathbb{N}$ , veamos qué pasa con  $n + 1$ .

Si  $n$  es par

$$(x_{n+1}, s) = (P_1 x_n, s) = (x_n, s).$$

Si  $n$  es impar

$$(x_{n+1}, s) = (P_2 x_n, s) = (x_n, s).$$

Y por lo tanto, por hipótesis de inducción tenemos  $(x_{n+1}, s) = (x_n, s) = (x_0, s)$ .

b) Usando a) tenemos que para cualquier  $s \in S$  y  $n \in \mathbb{N}$

$$(x_n - Px_0, s) = (x_n, s) - (Px_0, s) = (x_0, s) - (x_0, s) = 0.$$

c) Partiendo de b) sabemos que  $Px_\infty = Px_0$ . Como  $x_\infty \in S$ , entonces  $Px_\infty = x_\infty = Px_0$ .  $\square$

Volviendo a la sucesión  $\{x_n\}$  definida por (3.1) tenemos que

$$(x_n - x_{n+1}, x_{n+1}) = 0$$

y aplicando el Teorema de Pitágoras para  $u = x_n - x_{n+1}$  y  $v = x_{n+1}$  obtenemos

$$\|x_n\|^2 = \|x_n - x_{n+1}\|^2 + \|x_{n+1}\|^2. \quad (3.2)$$

Por otro lado, dados  $k, l$  dos números enteros, denotamos por  $m = m(k, l) = \lfloor (k + l + 1)/2 \rfloor$  la parte entera de  $(k + l + 1)/2$ .

**Lema 3.3.** *Sea  $\{x_n\}_{n \in \mathbb{N}}$  la sucesión definida en (3.1), sean  $k, l$  dos números enteros y  $m = m(k, l)$ . Entonces*

$$\|x_k - x_l\|^2 = \|x_k\|^2 + \|x_l\|^2 - 2\|x_m\|^2. \quad (3.3)$$

*Demostración.* Razonamos por inducción en  $n = k - l$ . Si  $n = 0$  ( $k = l$ )

$$0 = \|x_k - x_k\|^2 = \|x_k\|^2 + \|x_k\|^2 - 2\|x_m\|^2 = 2\|x_k\|^2 - 2\|x_k\|^2 = 0.$$

Si  $n = 1$ , supongamos que  $k = l + 1$ , de modo que

$$m = \left\lfloor \frac{k + l + 1}{2} \right\rfloor = \left\lfloor \frac{2l + 2}{2} \right\rfloor = l + 1 = k.$$

Aplicando 3.2 obtenemos

$$\|x_l\|^2 = \|x_l - x_{l+1}\|^2 + \|x_{l+1}\|^2 = \|x_k - x_l\|^2 + \|x_k\|^2$$

y despejando

$$\|x_k - x_l\|^2 = \|x_l\|^2 - \|x_k\|^2$$

llegando a 3.3.

Supongamos que se cumple para un cierto  $n \in \mathbb{N}$ , veamos que se cumple para  $k - l = n + 1$ .

Si  $n$  es par, entonces  $x_{k+1}$  y  $x_l$  pertenecen al mismo subespacio y por tanto  $(x_k - x_{k+1}, x_{k+1} - x_l) = 0$ . Además,  $m(k + 1, l) = m(k, l) = (k + l + 1)/2$ . Por tanto,

$$\|x_k - x_l\|^2 = \|x_k - x_{k+1} + x_{k+1} - x_l\|^2 = \|x_k - x_{k+1}\|^2 + \|x_{k+1} - x_l\|^2 =$$

y usando (3.2) para el primer cuadrado y la hipótesis de inducción para el segundo obtenemos

$$= \|x_k\|^2 - \|x_{k+1}\|^2 + \|x_{k+1}\|^2 - 2\|x_{m(k+1, l)}\|^2 + \|x_l\|^2 = \|x_k\|^2 - 2\|x_{m(k+1, l)}\|^2 + \|x_l\|^2.$$

Si  $n$  es impar, entonces los puntos  $x_k$  y  $x_l$  pertenecen al mismo subespacio y por tanto  $(x_{l-1} - x_l, x_k - x_l) = 0$  y  $m(k, l - 1) = m(k, l) = (k + l)/2$  y de nuevo por (3.2) y la hipótesis de inducción, obtenemos

$$\|x_k - x_{l-1}\|^2 = \|x_k - x_l + x_l - x_{l-1}\|^2 = \|x_k - x_l\|^2 + \|x_l - x_{l-1}\|^2$$

donde

$$\begin{aligned} \|x_k - x_l\|^2 &= \|x_k - x_{l-1}\|^2 - \|x_l - x_{l-1}\|^2 = \|x_k\|^2 + \|x_{l-1}\|^2 - 2\|x_{m(k, l-1)}\|^2 - (\|x_{l-1}\|^2 - \|x_l\|^2) = \\ &= \|x_k\|^2 + \|x_l\|^2 - 2\|x_{m(k, l-1)}\|^2 \end{aligned}$$

con lo que queda demostrado el lema.  $\square$

Ahora ya tenemos todo lo necesario para probar el Teorema 3.1.

*Demostración.* Teorema 3.1.

Dado que la sucesión  $\{\|x_n\|\}_{n \in \mathbb{N}}$  es decreciente,  $k \leq m(k, l) \leq l$  y partiendo de la igualdad (3.3) tenemos

$$\|x_k\| \geq \|x_m\| \geq \|x_l\|$$

por lo que

$$\|x_k\|^2 \geq \|x_m\|^2 \geq \|x_l\|^2.$$

Por tanto

$$\|x_k - x_l\|^2 = \|x_k\|^2 + \|x_l\|^2 - 2\|x_m\|^2 \leq \|x_k\|^2 - \|x_l\|^2$$

ya que

$$\|x_l\|^2 - 2\|x_m\|^2 \leq \|x_l\|^2 - 2\|x_l\|^2 = -\|x_l\|^2$$

Por lo tanto, la sucesión es de Cauchy y converge a  $Px_0$  como resultado del Lema 3.2 (c).  $\square$

Dicho teorema que hemos visto para dos subespacios, se puede extender a  $m$  subespacios, donde  $m \in \mathbb{N}$ , haciendo alternativamente la proyección sobre cada uno de los  $m$  subespacios. En este caso, el algoritmo sería

$$x_{mn+1} = P_1 x_{mn}; \quad x_{mn+2} = P_2 x_{mn+1}; \quad \dots \quad x_{mn} = P_m x_{mn-1}.$$

### 3.1.1. Ejemplos y resultados

**Ejemplo 3.4.** Por otro lado, veamos un ejemplo donde el orden en el que se hacen las proyecciones no influye en el resultado.

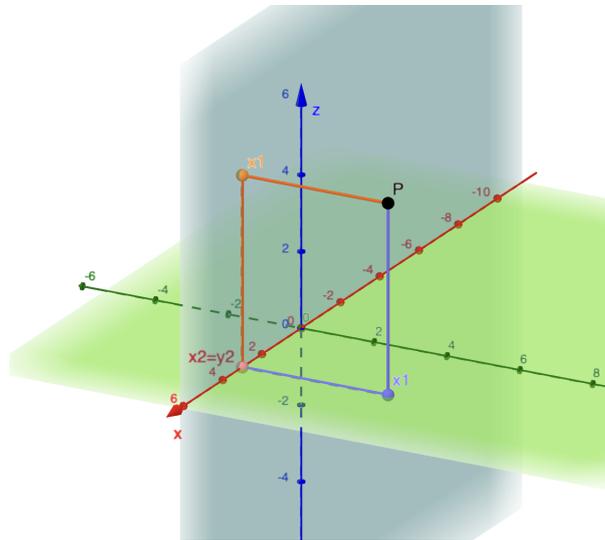


Figura 3.1: Ejemplo intersección empezando por  $S_1$  y  $S_2$ .

Sea  $H = \mathbb{R}^3$ , consideramos los planos  $y = 0$  y  $z = 0$ , es decir, los subespacios

$$S_1 = \{(x, y, 0) : x, y \in \mathbb{R}\} \text{ y } S_2 = \{(x, 0, z) : x, z \in \mathbb{R}\}$$

y el punto  $x_0 = (3, 4, 5)$ . Veamos que indistintamente del orden en el que realicemos las proyecciones, llegamos a la misma solución.

Comenzamos realizando la proyección sobre  $S_1$ . Aplicando el Teorema (2.3) tenemos que  $x_1 = (3, 4, 0)$ , como podemos observar visualmente. Realizamos ahora la proyección sobre  $S_2$  y obtenemos que  $x_2 = (3, 0, 0)$  que es un punto de la intersección de ambos planos, por lo que hemos llegado a la solución.

Veamos ahora el caso contrario. Comenzamos haciendo la proyección sobre  $S_2$  y obtenemos  $y_1 = (3, 0, 5)$ . Realizando ahora la proyección sobre  $S_1$  obtenemos  $y_2 = (3, 0, 0)$  que vuelve a ser un punto de la intersección, por lo que es la solución buscada.

Por lo tanto, llegamos en ambos casos a la misma solución, aunque los puntos intermedios sean distintos.

Dicho resultado se puede extender de nuevo al problema sobre  $m$  subespacios. Veamos el siguiente lema, cuyo enunciado y demostración se encuentran en [6].

**Lema 3.5.** Sean  $\{S_j : j = 1, 2, \dots, m\}$   $m$  subespacios cerrados de  $H$  y sea  $P_j : H \rightarrow S_j$  las proyecciones de  $H$  sobre  $S_j$ ,  $j = 1, 2, \dots, m$ . Sea  $r : \{1, 2, \dots\} \rightarrow \{1, 2, \dots, m\}$  una aplicación sobreyectiva que asume cada uno de sus valores infinitas veces. Sea  $x_0$  un punto arbitrario de  $H$ , se define la sucesión  $\{x_n\}_{n \in \mathbb{N}}$  como

$$x_{n+1} = P_{r(n+1)}x_n. \quad (3.4)$$

Dicha sucesión converge a la proyección sobre la intersección  $S = \cap\{S_j : j = 1, 2, \dots, m\}$ .

Por último, veamos la necesidad en el algoritmo de von Neumann de que los conjuntos sobre los que proyectamos sean subespacios cerrados. En el caso de conjuntos convexos cerrados, el algoritmo nos lleva a un punto de la intersección de los dos conjuntos (suponiendo que esta no es vacía), pero no necesariamente a la proyección sobre dicha intersección. Veámoslo en el siguiente ejemplo.

**Ejemplo 3.6.** Sean  $S_1 = \{(x, y) \in \mathbb{R}^2 : x \geq 0\}$  y  $S_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ . Sea  $p = (-1, 1)$ .

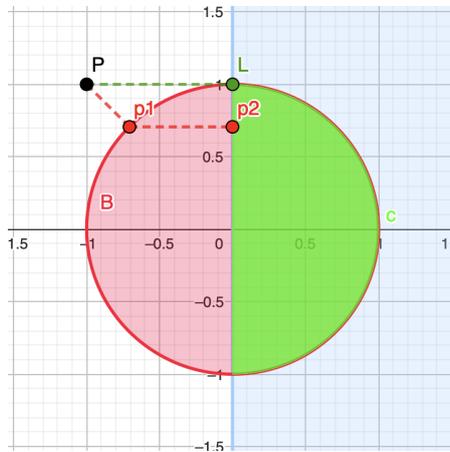


Figura 3.2: Proyección sobre la intersección de dos conjuntos convexos cerrados.

Observamos que  $S_2$  es el disco de centro el origen y radio 1, que no es un subespacio. Aplicamos ahora el algoritmo de von Neumann. Comenzamos proyectando sobre  $S_2$ .

$$P_{S_2}(p) = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = p_1$$

A continuación proyectamos sobre  $S_1$ .

$$P_{S_1}(p_1) = \left(0, \frac{1}{\sqrt{2}}\right) = p_2$$

Y  $p_2$  es un punto de la intersección de  $S_1$  y  $S_2$ , por tanto, el algoritmo ha llegado a la solución. Claramente, este punto no es la proyección de  $p$  sobre la intersección de  $S_1$  y  $S_2$ ; como podemos ver gráficamente, el punto correcto es el  $(0, 1)$ . De este modo, vemos que el algoritmo no siempre funciona cuando los conjuntos no son subespacios.

**Observación 3.7.** En el algoritmo de von Neumann, la hipótesis de que la intersección  $A \cap B$  de los subespacios cerrados es no vacía es esencial. Si dicha intersección es vacía, la sucesión generada por las proyecciones alternadas no converge, y en general, oscila o converge a un elemento de mínima distancia entre los conjuntos, pero no corresponde a una proyección sobre una intersección (ya que esta no existe). Por tanto, la validez del algoritmo como método de proyección ortogonal requiere que  $A \cap B \neq \emptyset$ .

Para concluir esta sección, veamos un ejemplo sencillo del algoritmo de von Neumann.

**Ejemplo 3.8.** Sean

$$S_1 = \{f \in L^2[0, 1] : \int_0^1 f(x) dx = 0\}; \quad S_2 = \langle \sin(\pi x), \sin(2\pi x), \sin(3\pi x) \rangle$$

queremos calcular la proyección de  $f(x) = x$  sobre  $S = S_1 \cap S_2$ , aplicando el algoritmo de Von Neumann.

Comenzamos calculando la proyección sobre  $S_1$ . En primer lugar,  $S_1$  es claramente un subespacio cerrado. Para que la proyección tenga integral nula tendremos que

$$x_1 = P_{S_1} f = f - \int_0^1 f(x) dx$$

Calculamos

$$\int_0^1 x dx = \frac{1}{2} \implies x_1 = x - \frac{1}{2}$$

A continuación, continuando con el algoritmo, proyectaremos  $x_1$  sobre  $S_2$ . Como podemos observar, esta proyección es equivalente al Ejemplo 2.11 simplemente adaptando el dominio. Por lo tanto tenemos

$$\int_0^1 \sin(i\pi x) \sin(k\pi x) dx = \begin{cases} 0 & \text{si } i \neq k \\ \frac{1}{2} & \text{si } i = k \end{cases}$$

y los coeficientes serán

$$a_k = 2 \int_0^1 f(x) \sin(k\pi x) dx = 2 \int_0^1 \left(x - \frac{1}{2}\right) \sin(k\pi x) dx \quad \text{para } k = 1, 2, 3$$

e integrando en cada caso obtenemos que  $a_1 = 0$ ,  $a_2 = -\frac{1}{\pi}$ ,  $a_3 = 0$ . Entonces la proyección es

$$x_2 = P_{S_2} x_1 = -\frac{1}{\pi} \sin(2\pi x)$$

Veamos si pertenece a  $S_1$  realizando la integral.

$$\int_0^1 -\frac{1}{\pi} \sin(2\pi x) dx = 0$$

luego pertenece a  $S = S_1 \cap S_2$  y hemos llegado a la proyección buscada.



## Capítulo 4

# Proyecciones relacionadas con Fourier

En este capítulo se analiza el papel de las proyecciones en contextos donde intervienen transformadas de Fourier, y se aplican a problemas prácticos de control y optimización. Se estudian distintos tipos de conjuntos definidos en términos de propiedades de sus transformadas, y se presentan resultados clásicos de la literatura.

Finalmente, se exponen aplicaciones concretas de las proyecciones en modelos físicos y de ingeniería, que muestran la potencia de esta herramienta en espacios funcionales.

### 4.1. Teoría de Fourier

En esta sección, se presenta teoría y resultados obtenidos de [5], adaptados convenientemente a la notación empleada en este trabajo. Asimismo, se ha optado por utilizar la definición simétrica de la transformada de Fourier, en la que tanto la transformada directa como su inversa incluyen un factor constante  $\frac{1}{\sqrt{2\pi}}$ .

En el Ejemplo 2.11, se considera el caso particular de las ecuaciones normales en las que las funciones base están dadas por  $x_j = \sin(\frac{j\pi x}{L})$ . Este planteamiento conduce a las **series de Fourier de senos**, es decir, la proyección de cualquier función de  $H$  sobre el conjunto  $\mathbb{K}$  (en este caso senos hasta  $n = 3$ , pero podría extenderse a cualquier  $n \in \mathbb{N}$ ) que coincide con la aproximación de la función  $f$  en serie de Fourier de senos en sus primeros términos.

**Definición 4.1.** Dada una función  $f : \mathbb{R} \rightarrow \mathbb{R}$ , la **transformada de Fourier** de  $f$  es otra función  $\hat{f}$  que para cada  $\xi \in \mathbb{R}$  viene dada por

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx$$

en caso de que exista.

**Definición 4.2.** Dada una función  $g(\xi)$  definida para cada  $\xi \in \mathbb{R}$ , la **transformada inversa de Fourier** de  $g$  es otra función  $\mathcal{F}^{-1}(g)$  que para cada  $x \in \mathbb{R}$  viene dada por

$$\mathcal{F}^{-1}(g)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(\xi) e^{i\xi x} d\xi.$$

**Aplicación a funciones ortogonales en  $L^2[0, \pi]$ .**

El espacio  $L^2[0, \pi]$  es un espacio de Hilbert cuya base ortonormal puede estar formada por funciones como los exponenciales complejos  $e^{2\pi i \xi x}$  o por funciones seno o coseno. Usando la serie de Fourier en senos, sabemos que se verifica la identidad de Parseval,

$$\|f\|_{L^2[0, \pi]}^2 = \int_0^\pi \left| \sum_{n=1}^{\infty} a_n \sin(nx) \right|^2 dx = \sum_{n=1}^{\infty} |a_n|^2.$$

Esta igualdad es una versión del Teorema de Pitágoras en un espacio infinito-dimensional, donde en lugar de una suma finita de vectores ortogonales, tenemos una suma infinita (una serie) de funciones ortogonales. En el caso de la transformada de Fourier, el resultado equivalente se conoce como Teorema de Plancherel, que introducimos a continuación (ver [3]).

**Teorema 4.3. (Teorema de Plancherel)**

Dada una función  $f \in L^2(\mathbb{R})$ , su transformada de Fourier también pertenece a  $L^2(\mathbb{R})$  y se cumple

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |\hat{f}(\xi)|^2 d\xi$$

o lo que es lo mismo

$$\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})}$$

es decir, que la transformada de Fourier es una transformación que conserva la norma en  $L^2(\mathbb{R})$ , es decir, es una **isometría**.

**4.2. Proyección sobre conjuntos de  $L^2(\mathbb{R})$** **4.2.1. Conjuntos sin transformada de Fourier**

Comenzamos calculando las proyecciones sobre una serie de conjuntos similares a los ya tratados anteriormente, que nos ayudarán a entender mejor los conjuntos posteriores, en los que introduciremos las transformadas de Fourier. Todos los conjuntos que se estudian en esta sección se presentan en el artículo [7], de los cuales demostraremos que cumplen las condiciones de existencia de solución y calcularemos sus proyecciones.

Comenzamos con el siguiente conjunto.

$$\mathbb{K}_1 = \{f \in L^2(\mathbb{R}) : f(x) = 0 \text{ c.t.p. } x \notin [-a, a]\}$$

Este conjunto es claramente un subespacio, veamos que es cerrado. Sea  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_1$  una sucesión tal que  $f_n \rightarrow f$  en  $L^2(\mathbb{R})$ . Es decir,  $\|f_n - f\|_{L^2} \rightarrow 0$  cuando  $n \rightarrow \infty$ . Queremos demostrar que  $f \in \mathbb{K}_1$ , es decir, que  $f(x) = 0$  c.t.p.  $x \notin [-a, a]$ . Por la convergencia de  $f_n \rightarrow f$  en  $L^2(\mathbb{R})$ , tenemos que

$$\int_{\mathbb{R}} |f_n(x) - f(x)|^2 dx \rightarrow 0 \quad \text{cuando } n \rightarrow \infty.$$

Esto implica, en particular, que

$$\int_{\mathbb{R} \setminus [-a, a]} |f_n(x) - f(x)|^2 dx \rightarrow 0 \quad \text{cuando } n \rightarrow \infty$$

ya que  $\mathbb{R} \setminus [-a, a] \subset \mathbb{R}$ . Dado que  $f_n(x) = 0$  c.t.p. en  $x \notin [-a, a]$ , se tiene

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R} \setminus [-a, a]} |f_n(x) - f(x)|^2 dx = \int_{\mathbb{R} \setminus [-a, a]} |f(x)|^2 dx = 0.$$

Por lo tanto,  $|f(x)|^2 = 0$  c.t.p. en  $x \notin [-a, a]$ , lo que implica que  $\mathbb{K}_1$  es cerrado. Dada  $f \in L^2(\mathbb{R})$ , intuitivamente la proyección es

$$P_{\mathbb{K}_1}(f)(x) = f(x)\chi_{[-a,a]}(x).$$

Comprobémoslo. Empleando el Corolario 2.6 se tiene que cumplir

$$(f - P_{\mathbb{K}_1}(f), v) = 0 \quad \forall v \in \mathbb{K}_1$$

Calculándolo

$$\int_{\mathbb{R}} (f(x) - f(x)\chi_{[-a,a]}(x))v(x) dx = \int_{-\infty}^{-a} f(x) \cdot 0 dx + \int_{-a}^a 0 \cdot v dx + \int_a^{+\infty} f(x) \cdot 0 dx = 0 \quad \forall v \in \mathbb{K}_1$$

Por lo tanto, se cumple.

Veamos ahora el siguiente conjunto.

$$\mathbb{K}_2 = \{f \in L^2(\mathbb{R}) : f(x) = g(x) \text{ c.t.p. } x \in [-a, a]\}$$

Veamos que el conjunto es convexo y cerrado.

En primer lugar, dados  $f_1, f_2 \in \mathbb{K}_2$ ,  $\lambda \in [0, 1]$  vemos que en  $[-a, a]$ ,  $\lambda f_1 + (1 - \lambda)f_2 = \lambda g + (1 - \lambda)g = g$ , por lo tanto es convexo.

Veamos que es cerrado. Sea  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_2$  tal que  $f_n \rightarrow f$  en  $L^2(\mathbb{R})$ , queremos ver que  $f \in \mathbb{K}_2$ . Por definición de  $\mathbb{K}_2$  tenemos que  $\forall n \in \mathbb{N}$ ,  $f_n = g$  en  $[-a, a]$ . Por el Teorema A.15, existe una subsucesión  $\{f_{n_k}\}$  que converge a  $f$  en casi todo punto. Como para cada  $k$ ,  $f_{n_k}(x) = g(x)$  en  $x \in [-a, a]$ , pasando al límite tenemos que  $f \in \mathbb{K}_2$  y concluimos que es cerrado.

Veamos que la proyección es la siguiente.

$$P_{\mathbb{K}_2}(f)(x) = g(x)\chi_{[-a,a]}(x) + [1 - \chi_{[-a,a]}(x)]f(x).$$

Aplicamos de nuevo el Corolario 2.6

$$(f - P_{\mathbb{K}_2}(f), v) = 0 \quad \forall v \in \mathbb{K}_2$$

$$\begin{aligned} (f - (g\chi_{[-a,a]} + [1 - \chi_{[-a,a]}]f), v) &= 0 \quad v \in \mathbb{K}_2 \Leftrightarrow \\ \Leftrightarrow ((f - g)\chi_{[-a,a]}, v) &= 0 \quad v \in \mathbb{K}_2 \end{aligned}$$

y como  $g\chi_{[-a,a]} = f\chi_{[-a,a]}$ , obtenemos  $(0, v) = 0 \quad \forall v \in \mathbb{K}_2$  y queda demostrado.

Veamos otro conjunto.

$$\mathbb{K}_3 = \{f \in L^2(\mathbb{R}) : f(x) = 0 \text{ c.t.p. } x \notin [-a, a]; f(x) \geq 0 \text{ c.t.p. } x \in [-a, a]\}$$

Veamos que es un conjunto convexo y cerrado. Dados  $f_1, f_2 \in \mathbb{K}_3$  y  $\lambda \in [0, 1]$  tenemos que

$$\begin{aligned} \lambda f_1(x) + (1 - \lambda)f_2(x) &\geq 0 \text{ para } x \in [-a, a] \\ \lambda f_1(x) + (1 - \lambda)f_2(x) &= 0 \text{ para } x \notin [-a, a] \end{aligned}$$

luego la combinación lineal pertenece a  $\mathbb{K}_3$  y es convexo. Veamos que es cerrado. Por la demostración realizada en  $\mathbb{K}_1$ , sabemos que cumple la primera condición. Por otro lado, sea  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_3$ , suponemos

que  $f_n \rightarrow f$ . Por el Teorema A.15 sabemos que existe una subsucesión  $\{f_{n_k}\}$  tal que  $f_{n_k}(x) \rightarrow f(x)$ . Dado que  $f_{n_k}(x) \geq 0$  en c.t.p. en  $[-a, a]$ , el límite puntual de  $f_{n_k}(x)$ , que es  $f(x)$ , también cumple  $f(x) \geq 0$  en  $[-a, a]$ . Luego  $f \in \mathbb{K}_3$  y  $\mathbb{K}_3$  es cerrado.

Veamos que su proyección es

$$P_{\mathbb{K}_3}(f)(x) = \text{máx}(f(x), 0)\chi_{[-a, a]}(x).$$

En este caso, aplicamos el Teorema 2.3. Para ello, dividimos  $\mathbb{R}$  de la siguiente manera.

$$\mathbb{R} = \Omega_0 \cup \Omega_1 \cup \Omega_2$$

donde

$$\Omega_0 = (-\infty, -a) \cup (a, +\infty); \quad \Omega_1 = \{x \in [-a, a] : f(x) \geq 0\}; \quad \Omega_2 = \{x \in [-a, a] : f(x) < 0\}.$$

Por lo tanto, obtenemos

$$\begin{aligned} \int_{\mathbb{R}} (f(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx &= \int_{\Omega_0} (f(x) - 0)(v(x) - 0) dx + \int_{\Omega_1} (f(x) - f(x))(v(x) - f(x)) dx + \\ &+ \int_{\Omega_2} (f(x) - 0)(v(x) - 0) dx = \int_{\Omega_0} f(x)0 dx + \int_{\Omega_2} f(x)v(x) dx \leq 0 \quad \forall v \in \mathbb{K}_3 \end{aligned}$$

y se cumple la condición.

Añadimos una condición adicional a  $\mathbb{K}_3$ .

$$\mathbb{K}_4 = \{f \in L^2(\mathbb{R}) : f(x) = 0 \text{ c.t.p. } x \notin [-a, a]; f(x) \geq 0 \text{ c.t.p. } x \in [-a, a]; \int_{-a}^a f(x) dx = \rho\}$$

donde  $\rho > 0$ , o lo que es lo mismo  $\mathbb{K}_4 = \{f \in \mathbb{K}_3 : \int_{-a}^a f(x) dx = \rho, \rho > 0\}$ . Probemos que el conjunto es convexo. Como  $\mathbb{K}_3$  es convexo, basta probar la última condición. Dados  $f_1, f_2 \in \mathbb{K}_4$  y  $\lambda \in [0, 1]$  tenemos

$$\int_{-a}^a \lambda f_1(x) + (1 - \lambda)f_2(x) dx = \lambda \int_{-a}^a f_1(x) dx + (1 - \lambda) \int_{-a}^a f_2(x) dx = \lambda\rho + (1 - \lambda)\rho = \rho$$

y por tanto  $\mathbb{K}_4$  es convexo. Veamos que es cerrado. Como  $\mathbb{K}_3$  es cerrado, basta comprobar que la condición adicional no afecta a esta propiedad. Dada una sucesión  $\{f_n\} \subset \mathbb{K}_4$  tal que  $f_n \rightarrow f$  en  $L^2(\mathbb{R})$ , sabemos que

$$\int_{-a}^a f_n(x) dx = \rho \quad \forall n \in \mathbb{N}$$

y al tomar el límite, obtenemos

$$\int_{-a}^a f(x) dx = \lim_{n \rightarrow \infty} \int_{-a}^a f_n(x) dx = \rho$$

luego  $f \in \mathbb{K}_4$  y concluimos que  $\mathbb{K}_4$  es cerrado.

Veamos ahora cuál es la proyección. Dada la similitud con  $\mathbb{K}_3$ , podemos pensar que la proyección será la misma, pero añadiendo una constante que nos haga que se cumpla la condición de la integral. Por tanto, intuimos que la proyección será

$$P_{\mathbb{K}_4}(f)(x) = \text{máx}(f(x) + c, 0)\chi_{[-a, a]}(x)$$

donde  $c$  es una constante tal que

$$\int_{-a}^a \text{máx}(f(x) + c, 0) dx = \rho.$$

En primer lugar, veamos que existe tal constante  $c$ . Denotamos por  $A(c) = \int_{-a}^a \max(f(x) + c, 0) dx - \rho$ . Veamos que  $A$  es continua. Para ello, veamos que es Lipschitz-continua, es decir

$$|A(c) - A(\hat{c})| \leq L |c - \hat{c}|.$$

En nuestro caso,

$$\begin{aligned} |A(c) - A(\hat{c})| &= \left| \int_{-a}^a \max(f(x) + c, 0) dx - \rho - \left( \int_{-a}^a \max(f(x) + \hat{c}, 0) dx - \rho \right) \right| \leq \\ &\leq \int_{-a}^a |\max(f(x) + c, 0) - \max(f(x) + \hat{c}, 0)| dx. \end{aligned}$$

Observamos ahora que  $H(c) = \max(c_0 + c, 0)$  es Lipschitz continua, es decir, para cualesquiera  $u, v \in \mathbb{R}$

$$|\max(u, 0) - \max(v, 0)| \leq |u - v|.$$

Aplicando esto para  $u = f(x) + c$  y  $v = f(x) + \hat{c}$  obtenemos

$$|\max(f(x) + c, 0) - \max(f(x) + \hat{c}, 0)| \leq |(f(x) + c) - (f(x) + \hat{c})| = |c - \hat{c}|$$

y por lo tanto

$$\int_{-a}^a |\max(f(x) + c, 0) - \max(f(x) + \hat{c}, 0)| dx \leq \int_{-a}^a |c - \hat{c}| dx = 2a |c - \hat{c}|$$

y se cumple que es Lipschitz-continua con constante  $L = 2a$ , que implica que es continua. Veamos que existe un valor de  $c$  para el cual  $A(c) = 0$ .

1. Si  $A(0) = 0$ , ya está.

2. Si  $A(0) < 0$  (o análogamente  $A(0) > 0$ ).

Si  $f$  está acotada inferiormente,  $k_1 \leq f(x) \forall x \in [-a, a]$ . Tomando  $c + k_1 \geq n$  con  $n \in \mathbb{N}$ ,

$$\int_{-a}^a \max(f(x) + c, 0) dx - \rho = \int_{-a}^a n - \rho dx = 2an - \rho \implies n > \frac{\rho}{2a}$$

En el caso general, basta aproximar  $f$  por una función continua en  $[-a, a]$  y obtenemos

$$\int_{-a}^a \max(f(x) + c, 0) dx \approx \int_{-a}^a \max(\varphi(x) + c, 0) dx \geq 0$$

convirtiéndolo en el caso en el que  $f$  está acotado inferiormente. Por lo tanto, vemos que existe  $c_1$  suficientemente grande tal que  $A(c_1) > 0$ . Aplicando el teorema de Bolzano, como  $A$  es continua,  $A(0) < 0$  y  $A(c_1) > 0$  entonces tiene que existir un  $c$  entre 0 y  $c_1$  tal que  $A(c) = 0$ .

Con todo ello queda probada la existencia de  $c$ .

Una vez demostrado que existe  $c$ , veamos que la proyección es correcta. Consideramos de nuevo el intervalo  $[-a, a] = \Omega_+^c \cup \Omega_-^c$  donde

$$\Omega_+^c = \{x \in [-a, a] : f(x) + c \geq 0\}; \quad \Omega_-^c = \{x \in [-a, a] : f(x) + c < 0\}$$

$$\begin{aligned} \int_{\mathbb{R}} (f(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx &= \int_{-a}^a (f(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \int_{\Omega_+^c} (f(x) - (f(x) + c))(v(x) - (f(x) + c)) dx + \\ &+ \int_{\Omega_-^c} f(x)v(x) dx = \int_{\Omega_+^c} (-c)(v(x) - (f(x) + c)) dx + \int_{\Omega_-^c} f(x)v(x) dx \end{aligned}$$

y sabemos que en  $\Omega_-^c$ ,  $f(x) < -c$ , y como  $\int_{\Omega_+^c} (f(x) + c) dx = \rho$ , obtenemos

$$\begin{aligned} & \int_{\Omega_+^c} (-c)(v(x) - (f(x) + c)) dx + \int_{\Omega_-^c} f(x)v(x) dx < \int_{\Omega_+^c} (-c)(v(x) - (f(x) + c)) dx - \int_{\Omega_-^c} cv(x) dx = \\ & - \int_{\Omega_+^c} cv(x) dx + \int_{\Omega_+^c} c(f(x) + c) dx - \int_{\Omega_-^c} cv(x) dx = -c \int_{-a}^a v(x) dx + c \int_{\Omega_+^c} (f(x) + c) dx = -c\rho + c\rho = 0 \end{aligned}$$

y se cumple la desigualdad.

Por último, consideramos otro conjunto añadiéndole otra condición a  $\mathbb{K}_3$ .

$$\mathbb{K}_5 = \{f \in L^2(\mathbb{R}) : f(x) = 0 \text{ c.t.p. } \notin [-a, a], f(x) \geq 0 \text{ c.t.p. } \in [-a, a], \int_{-a}^a f^2(x) dx \leq \rho^2\}$$

Dicho conjunto también se puede expresar como  $\mathbb{K}_5 = \{f \in \mathbb{K}_3 : \int_{-a}^a f^2(x) dx \leq \rho^2, \rho > 0\}$ . Veamos que es convexo.

Como  $\mathbb{K}_3$  es convexo, basta comprobar la condición adicional. Sean  $f_1, f_2 \in \mathbb{K}_5$  y  $\lambda \in [0, 1]$ , tenemos que

$$(\lambda f_1(x) + (1 - \lambda)f_2(x))^2 \leq \lambda f_1^2(x) + (1 - \lambda)f_2^2(x)$$

e integrando obtenemos

$$\int_{-a}^a (\lambda f_1(x) + (1 - \lambda)f_2(x))^2 dx \leq \lambda \int_{-a}^a f_1^2(x) dx + (1 - \lambda) \int_{-a}^a f_2^2(x) dx \leq \lambda\rho^2 + (1 - \lambda)\rho^2 = \rho^2$$

y  $\mathbb{K}_5$  es convexo. Veamos que es cerrado. Dada una sucesión convergente  $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{K}_5$ , tal que  $f_n \rightarrow f$  en  $L^2(\mathbb{R})$ . Como  $\mathbb{K}_3$  es cerrado y  $f_n \in \mathbb{K}_3$ , entonces  $f \in \mathbb{K}_3$ . Dado que cada  $f_n$  satisface que  $\int_{-a}^a f_n^2(x) dx \leq \rho^2$ , tomando el límite obtenemos

$$\int_{-a}^a f^2(x) dx = \lim_{n \rightarrow \infty} \int_{-a}^a f_n^2(x) dx \leq \rho^2$$

y es cerrado.

Como  $\mathbb{K}_5$  es convexo y cerrado, podemos aplicar el Teorema 2.3

$$\int_{\mathbb{R}} (f(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx \leq 0 \quad \forall v \in \mathbb{K}_5.$$

Veamos que su proyección es

$$\bar{u} = c \text{ máx}(f(x), 0) \chi_{[-a, a]}(x)$$

donde

$$c = \begin{cases} 1 & \text{si } \int_{-a}^a \text{máx}(f(x), 0)^2 dx \leq \rho^2 \\ \frac{\rho}{\sqrt{\int_{-a}^a \text{máx}(f(x), 0)^2 dx}} & \text{si } \int_{-a}^a \text{máx}(f(x), 0)^2 dx > \rho^2 \end{cases}$$

que está bien definida ya que si  $\int_{-a}^a \text{máx}(f(x), 0)^2 dx > \rho^2 > 0$ , entonces el denominador es siempre distinto de 0. Además, podemos ver que  $c \in (0, 1]$  ya que  $\sqrt{\int_{-a}^a \text{máx}(f(x), 0)^2 dx} > \rho$  y  $c > 0$  ya que numerador y denominador son mayores a cero en cualquier caso.

Dividiendo el intervalo  $[-a, a] = \Omega_+ \cup \Omega_-$  donde

$$\Omega_+ = \{x \in [-a, a] : f(x) \geq 0\}; \quad \Omega_- = \{x \in [-a, a] : f(x) < 0\}$$

obtenemos

$$\begin{aligned} (f - \bar{u}, v - \bar{u}) &= \int_{-a}^a (f(x) - \bar{u}(x))(v(x) - \bar{u}(x)) dx = \\ &= \int_{\Omega_+} (f(x) - cf(x))(v(x) - cf(x)) dx + \int_{\Omega_-} (f(x) - 0)(v(x) - 0) dx = \\ &= \int_{\Omega_+} (f(x) - cf(x))(v(x) - cf(x)) dx + \int_{\Omega_-} \underset{<0}{f(x)v(x)} \underset{\geq 0}{dx} \leq (1-c) \int_{\Omega_+} f(x)(v(x) - cf(x)) dx \end{aligned}$$

donde  $1 - c \geq 0$  ya que  $c \in (0, 1]$ . Si  $c = 1$  ya está porque  $1 - c = 0$ . Si  $c \neq 1$ , queremos ver entonces que  $\int_{\Omega_+} f(x)(v(x) - cf(x)) dx \leq 0$  o lo que es lo mismo

$$\int_{\Omega_+} f(x)v(x) dx \leq c \int_{\Omega_+} f^2(x) dx.$$

Aplicando la Desigualdad de Cauchy-Schwarz, Lema A.6

$$\int_{\Omega_+} f(x)v(x) dx \leq \left( \int_{\Omega_+} f^2(x) dx \right)^{\frac{1}{2}} \left( \int_{\Omega_+} v^2(x) dx \right)^{\frac{1}{2}}$$

y como  $\int_{\Omega_+} v^2(x) dx \leq \rho^2$

$$\int_{\Omega_+} f(x)v(x) dx \leq \rho \sqrt{\int_{\Omega_+} f^2(x) dx} = \frac{\rho}{\sqrt{\int_{\Omega_+} f^2(x) dx}} \int_{\Omega_+} f^2(x) dx = c \int_{\Omega_+} f^2(x) dx$$

con lo que se cumple la desigualdad y por tanto la proyección es correcta.

### 4.2.2. Conjuntos con transformada de Fourier

En los conjuntos que se trabajan a partir de ahora introducimos las transformadas directas e inversas de Fourier, para las cuales serán de vital importancia los Teoremas 4.3 y 4.4. Cabe recalcar que cada uno de los conjuntos citados está relacionado con los conjuntos anteriormente trabajados sin introducir las transformadas.

Veamos un teorema que nos facilitará todo el trabajo.

**Teorema 4.4.** *En el espacio  $L^2(\mathbb{R})$  con  $f \in L^2(\mathbb{R})$  y  $\mathbb{K} \subset L^2(\mathbb{R})$  un conjunto convexo y cerrado.*

$$P_{\hat{\mathbb{K}}}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}}(\mathcal{F}(f)))$$

donde  $\hat{\mathbb{K}} = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K}\}$ , es decir, es el conjunto de funciones cuya transformada de Fourier pertenece a  $\mathbb{K}$ .

*Demostración.* Por el Teorema 4.3, sabemos que la transformada de Fourier es una isometría, es decir, conserva las distancias. En particular, para  $f \in L^2(\mathbb{R})$  y cualquier conjunto cerrado convexo  $\mathbb{K} \subset L^2(\mathbb{R})$ , la proyección ortogonal  $P_{\mathbb{K}}(f)$  es el único elemento de  $\mathbb{K}$  que minimiza la distancia a  $f$ . Por tanto, tenemos

$$\left\| \underset{\in L^2(\mathbb{R})}{f} - \underset{\in \hat{\mathbb{K}}}{P_{\hat{\mathbb{K}}}(f)} \right\| \stackrel{\mathcal{F}}{\underset{isom}{\equiv}} \left\| \underset{\in L^2(\mathbb{R})}{\mathcal{F}(f)} - \underset{\in \mathbb{K}}{\mathcal{F}(P_{\hat{\mathbb{K}}}(f))} \right\| \iff \mathcal{F}(P_{\hat{\mathbb{K}}}(f)) = P_{\mathbb{K}}(\mathcal{F}(f)) \implies P_{\hat{\mathbb{K}}}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}}(\mathcal{F}(f)))$$

□

**Observación 4.5.** Si  $\mathbb{K}$  es convexo y/o cerrado,  $\hat{\mathbb{K}}$  también es convexo y/o cerrado. Además, si añadimos a  $\hat{\mathbb{K}}$  la condición  $Im(\mathcal{F}(f)) = 0$  de forma que

$$\hat{\mathbb{K}}^+ = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K} \text{ con } Im(\mathcal{F}(f)) = 0\}$$

$\hat{\mathbb{K}}^+$  sigue siendo convexo y cerrado

*Demostración.* La transformada de Fourier  $\mathcal{F} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$  es lineal y biyectiva por lo que, si  $\mathbb{K}$  es convexo, entonces  $\hat{\mathbb{K}} = \mathcal{F}^{-1}(\mathbb{K})$  también lo es, ya que la preimagen de un conjunto convexo por una aplicación lineal es convexo. Además, sabemos que la transformada de Fourier es una isometría por el Teorema 4.3, luego en particular es continua, y la imagen de un conjunto cerrado por una función continua es cerrado, luego si  $\mathbb{K}$  es cerrado,  $\hat{\mathbb{K}}$  también lo es.

Veamos que la condición adicional  $Im(\mathcal{F}(f)) = 0$  no altera la convexidad ni el hecho de que  $\hat{\mathbb{K}}$  sea un conjunto cerrado. Sea  $V = \{g \in L^2(\mathbb{R}) : Im(g) = 0\}$  un subespacio cerrado de  $L^2(\mathbb{R})$ , ya que la condición de tener parte imaginaria nula es lineal y cerrada en  $L^2(\mathbb{R})$ . Entonces, tenemos que

$$\hat{\mathbb{K}}^+ = \mathcal{F}^{-1}(\mathbb{K}) \cap \mathcal{F}^{-1}(V)$$

como la intersección finita de conjuntos convexos y cerrados es convexo y cerrado,  $\hat{\mathbb{K}}$  es convexo y cerrado  $\square$

Con este argumento, podemos razonar sobre los siguientes conjuntos, relacionados con los conjuntos ya trabajados anteriormente, si bien hay que tener cuidado con algunos casos que comentaremos a continuación.

Consideramos el primer conjunto.

$$\hat{\mathbb{K}}_1 = \{f \in L^2(\mathbb{R}) : \hat{f}(\xi) = 0 \text{ c.t.p. } \xi \notin [-a, a]\}$$

es decir,  $\hat{\mathbb{K}}_1 = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K}_1\} = \mathcal{F}^{-1}(\mathbb{K}_1)$ .

En primer lugar, vemos que  $\hat{\mathbb{K}}_1$  es claramente un subespacio. Por la Observación 4.5, como  $\mathbb{K}_1$  es cerrado,  $\hat{\mathbb{K}}_1$  también es cerrado.

Para calcular su proyección, aplicamos el Teorema 4.4 y obtenemos

$$P_{\hat{\mathbb{K}}_1}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}_1}(\mathcal{F}(f))) = \mathcal{F}^{-1}\left(\hat{f}(\xi)\chi_{[-a,a]}(\xi)\right)$$

y realizando la transformada inversa de Fourier obtenemos

$$P_{\hat{\mathbb{K}}_1}(f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\xi)\chi_{[-a,a]}(\xi)e^{ix\xi} d\xi.$$

**Interpretación práctica:** Supongamos que tenemos un archivo de audio, como una canción, de la que queremos optimizar su tamaño de almacenamiento, minimizándolo al máximo, sin perder calidad de audio. La Transformada de Fourier toma esta señal en el dominio del tiempo y la convierte en una representación de frecuencias. El rango de frecuencia que puede percibir el oído humano oscila entre los 20 y los 20.000 Hz; luego, si tomamos  $\hat{\mathbb{K}}_1 = \{f \in L^2(\mathbb{R}) : \hat{f}(\xi) = 0 \text{ c.t.p. } \xi \notin [20, 20000]\}$  y hacemos la proyección del archivo de audio en  $\hat{\mathbb{K}}_1$ , eliminaremos todos los sonidos que el oído humano no puede percibir, pero que están grabados en la canción, ocupando espacio de almacenamiento. De esta manera, habremos cumplido el objetivo buscado: reducir el tamaño del archivo con una transformación imperceptible para el oído humano. Esta interpretación la hemos obtenido de [7], donde aparecen más casos que pueden resultar de interés.

A este proceso, aplicado en muchos algoritmos de compresión de audio como MP3, se le denomina **filtro de paso banda**, que elimina las frecuencias fuera del rango audible para reducir el tamaño del archivo sin una pérdida perceptible de calidad.

Esta es una de las interpretaciones más comunes de la transformada de Fourier. Todos los ejemplos que se ponen a continuación se pueden interpretar de la misma manera, como restricciones aplicadas a un archivo de audio sobre su frecuencia.

Consideramos ahora el conjunto

$$\hat{\mathbb{K}}_3 = \{f \in L^2(\mathbb{R}) : \hat{f}(\xi) = 0 \text{ c.t.p. } \xi \notin [-a, a]; \hat{f}(\xi) \geq 0, \text{Im}(\hat{f}(\xi)) = 0 \text{ c.t.p. } \xi \in [-a, a]\}$$

o lo que es lo mismo,  $\hat{\mathbb{K}}_3 = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K}_3 \text{ con } \text{Im}(\mathcal{F}(f)) = 0\}$ .

Como se observa, hemos introducido la condición adicional de que la parte imaginaria de la transformada sea cero. La condición de  $\text{Im}(\mathcal{F}(f)) = 0$  es necesaria porque la proyección solo afecta la parte real de la función en el dominio de Fourier. Aunque la función sobre la que se aplica Fourier sea una función real, esta puede tomar valores complejos.

Veamos un ejemplo. Sea

$$f(x) = \begin{cases} e^{-x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

aplicando la definición de transformada, obtenemos

$$\begin{aligned} \hat{f}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-x} e^{-i\xi x} dx = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-(1+i\xi)x} dx = -\frac{1}{\sqrt{2\pi}} \frac{e^{-(1+i\xi)x}}{1+i\xi} \Big|_{x=0}^{x=+\infty} = \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{1+i\xi} = \frac{1}{\sqrt{2\pi}} \frac{1-i\xi}{1+\xi^2} = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{1+\xi^2} - i \frac{\xi}{1+\xi^2} \right) \in \mathbb{C} \end{aligned}$$

que nos lleva a valores complejos.

Por la Observación 4.5, como  $\mathbb{K}_3$  es convexo y cerrado,  $\hat{\mathbb{K}}_3$  es convexo y cerrado. Aplicando de nuevo el Teorema 4.4, obtenemos

$$P_{\hat{\mathbb{K}}_3}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}_3}(\mathcal{F}(f))) = \mathcal{F}^{-1} \left( \max(\text{Re}(\hat{f}(\xi)), 0) \chi_{[-a, a]}(\xi) \right)$$

y realizando la inversa de Fourier obtenemos

$$P_{\hat{\mathbb{K}}_3}(f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-a}^a \max(\text{Re}(\hat{f}(\xi)), 0) e^{ix\xi} d\xi.$$

Consideramos a continuación el siguiente subconjunto de  $\hat{\mathbb{K}}_3$ .

$$\hat{\mathbb{K}}_4 = \{f \in L^2(\mathbb{R}) : \hat{f}(\xi) = 0 \text{ c.t.p. } \xi \notin [-a, a], \hat{f}(\xi) \geq 0, \text{Im}(\hat{f}(\xi)) = 0 \text{ c.t.p. } \xi \in [-a, a], \frac{1}{\sqrt{2\pi}} \int_{-a}^a \hat{f}(\xi) d\xi = \rho, \rho > 0\}$$

o lo que es lo mismo  $\hat{\mathbb{K}}_4 = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K}_4 \text{ con } \text{Im}(\mathcal{F}(f)) = 0\}$ .

Como  $\mathbb{K}_4$  es convexo y cerrado, por la Observación 4.5,  $\hat{\mathbb{K}}_4$  también es convexo y cerrado. Aplicando de nuevo el Teorema 4.4, obtenemos

$$P_{\hat{\mathbb{K}}_4}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}_4}(\mathcal{F}(f))) = \mathcal{F}^{-1} \left( \max(\text{Re}(\hat{f}(\xi)) + c, 0) \chi_{[-a, a]}(\xi) \right)$$

y realizando la inversa se obtiene

$$P_{\mathbb{K}_4}(f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-a}^a \max(\operatorname{Re}(\hat{f}(\xi)) + c, 0) e^{ix\xi} d\xi.$$

donde  $c$  es una constante tal que

$$\frac{1}{\sqrt{2\pi}} \int_{-a}^a \max(\hat{f}(\xi) + c, 0) d\xi = \rho.$$

La existencia de tal constante  $c$  ya quedó probada en el conjunto  $\mathbb{K}_4$ .

Por último, consideramos otro subconjunto de  $\hat{\mathbb{K}}_3$ .

$$\hat{\mathbb{K}}_5 = \left\{ f \in L^2(\mathbb{R}) : \hat{f}(\xi) = 0 \text{ c.t.p. } \xi \notin [-a, a], \hat{f}(\xi) \geq 0 \text{ c.t.p. } \xi \in [-a, a], \operatorname{Im}(\hat{f}(\xi)) = 0 \text{ c.t.p. } \xi \in [-a, a], \frac{1}{\sqrt{2\pi}} \int_{-a}^a |\hat{f}(\xi)|^2 d\xi \leq \rho^2 \right\}$$

donde  $\rho > 0$ .  $\hat{\mathbb{K}}_5$  también se puede expresar como  $\hat{\mathbb{K}}_5 = \{f \in L^2(\mathbb{R}) : \mathcal{F}(f) \in \mathbb{K}_5 \text{ con } \operatorname{Im}(\mathcal{F}(f)) = 0\}$ .

Como  $\mathbb{K}_5$  es convexo y cerrado, por la Observación 4.5, sabemos que  $\hat{\mathbb{K}}_5$  también es convexo y cerrado. Aplicando el Teorema 4.4 tenemos

$$P_{\hat{\mathbb{K}}_5}(f) = \mathcal{F}^{-1}(P_{\mathbb{K}_5}(\mathcal{F}(f))) = \mathcal{F}^{-1}(c \max(\operatorname{Re}(\hat{f}(\xi)), 0) \chi_{[-a, a]}(\xi))$$

donde realizando la inversa obtenemos

$$P_{\hat{\mathbb{K}}_5}(f) = \frac{c}{\sqrt{2\pi}} \int_{-a}^a \max(\operatorname{Re}(\hat{f}(\xi)), 0) e^{ix\xi} d\xi$$

donde

$$c = \min \left( 1, \frac{\rho}{\sqrt{\frac{1}{2\pi} \int_{-a}^a \max(\operatorname{Re}(\hat{f}(\xi)), 0)^2 d\xi}} \right)$$

### 4.3. Ejemplos prácticos

En este último apartado se pretende ilustrar, mediante ejemplos concretos, la utilidad de las herramientas desarrolladas a lo largo del trabajo. Para ello, se abordarán dos situaciones reales en las que se aplica el marco teórico de las proyecciones sobre conjuntos convexos y cerrados.

#### 4.3.1. Lanzamiento de un cohete

Queremos lanzar un cohete de masa  $m$  (que suponemos constante) para que alcance una altura  $h$  (prefijada) en un instante de tiempo  $T$  (prefijado).

Por la Segunda Ley de Newton, tenemos

$$\begin{cases} mx''(t) = u(t) - mg \\ x(0) = 0 \\ x'(0) = 0 \end{cases}$$

donde  $x(t)$  es la altura del cohete en el instante  $t$ ,  $u(t)$  es la fuerza de empuje vertical en el instante  $t$  y  $g$  es la fuerza de la gravedad que suponemos  $g = 9,8 \text{ m/s}^2$ .

Definiendo nuestro problema de control óptimo, tendríamos

$$(PC) \begin{cases} \text{Min } J(u) = \frac{1}{2} \int_0^T u^2(t) dt \\ u(t) \geq 0 \\ x_u(T) = h \end{cases}$$

donde  $x_u$  denota la solución de la ecuación de estado asociada al control  $u$ . Para resolver este problema y calcular la fuerza de empuje necesaria, lo realizaremos como un problema de proyecciones. Para ello, definimos en primer lugar el conjunto sobre el que proyectar

$$\mathbb{K} = \{u \in L^2[0, T] : u(t) \geq 0 \text{ c.t.p. } t \in [0, T], x_u(T) = h\}$$

Usando la EDO, obtenemos

$$x''(t) = \frac{u(t)}{m} - g \xrightarrow{x'(0)=0} x'(t) = \frac{1}{m} \int_0^t u(s) ds - gt \xrightarrow{x(0)=0} x(t) = \frac{1}{m} \int_0^t \left[ \int_0^r u(s) ds \right] dr - g \frac{t^2}{2}$$

e integrando por partes obtenemos,

$$\int_0^t \int_0^r u(s) ds dr = r \int_0^r u(s) ds \Big|_{r=0}^{r=t} - \int_0^t ru(r) dr = t \int_0^t u(r) dr - \int_0^t ru(r) dr = \int_0^t (t-r)u(r) dr$$

luego tenemos

$$x(t) = \frac{1}{m} \int_0^t (t-r)u(r) dr - g \frac{t^2}{2}$$

y utilizando las condiciones del problema de control, podemos definir el siguiente conjunto

$$\mathbb{K} = \{u \in L^2[0, T] : u(t) \geq 0 \text{ en } [0, T], \int_0^T (T-r)u(r) dr = m(\frac{gT^2}{2} + h)\}$$

donde se cumple

$$\bar{u} \text{ es solución del (PC)} \iff P_{\mathbb{K}}(0) = \bar{u}$$

Sabemos que  $m, h, T$  y  $g$  son datos dados, luego  $\rho = m(\frac{gT^2}{2} + h)$  es una constante. De igual manera,  $p(r) = (T-r)$  es una función continua y decreciente en  $[0, T]$ . Luego podemos simplificar la notación de  $\mathbb{K}$  para facilitar el trabajo con él, quedando de la siguiente manera

$$\mathbb{K} = \{u \in L^2[0, T] : u \geq 0 \text{ en } [0, T], \int_0^T p(r)u(r) dr = \rho\}.$$

Veamos que el conjunto es convexo y cerrado. Para ver que es convexo, tomamos  $u_1, u_2 \in \mathbb{K}$  y  $\lambda \in [0, 1]$ . Sabemos que  $\int_0^T u_1(r)p(r)dr = \rho$  y  $\int_0^T u_2(r)p(r)dr = \rho$ . Por lo tanto

$$\int_0^1 (\lambda u_1(r) + (1-\lambda)u_2(r))p(r)dr = \lambda\rho + (1-\lambda)\rho = \rho$$

y es convexo. Veamos que es cerrado.

Para ello, consideramos una sucesión  $\{u_n\}$  en  $\mathbb{K}$  tal que  $u_n \rightarrow u$  en  $L^2[0, T]$ . Para cada  $n$  sabemos que

$$\int_0^T p(r)u_n(r) dr = \rho$$

y pasando al límite

$$\lim_{n \rightarrow \infty} \int_0^T p(r)u_n(r) dr = \int_0^T p(r)u(r) dr = \rho$$

Por otro lado, tenemos que ver que si  $u_n \geq 0$  para todo  $n$ , entonces  $u \geq 0$ . Aplicando el Teorema A.15, sabemos que existe una subsucesión  $\{u_{n_k}\}$  que converge a  $u$ . Dado que  $u_{n_k}(x) \geq 0$  en c.t.p. en  $[0, T]$ , el límite puntual de  $u_{n_k}(x)$ , que es  $u(x)$ , también cumple  $u(x) \geq 0$  en  $[0, 1]$ . Por lo tanto, concluimos que es cerrado.

Como hemos probado que es un conjunto cerrado y acotado, podemos aplicar el Teorema 2.3 y se tiene que cumplir que, sea  $\bar{u} = P_{\mathbb{K}}(0)$

$$(0 - \bar{u}(r), v(r) - \bar{u}(r)) \leq 0 \quad \forall v \in \mathbb{K}.$$

Veamos cuál es la proyección. De forma análoga a lo realizado en los ejemplos de la Sección 2.3 consideramos una corrección lineal en la dirección de  $(T - r)$ , es decir,  $\bar{u}(r) = C(T - r)$ , donde  $C$  es una constante que cumple  $\int_0^T (T - r)\bar{u}(r) dr = \rho$ . Calculemos esta constante. Sustituyendo

$$\int_0^T (T - r)\bar{u}(r) dr = \int_0^T (T - r)C(T - r) dr = C \int_0^T (T - r)^2 dr = \rho.$$

La integral

$$\int_0^T (T - r)^2 dr = \frac{T^3}{3}$$

y despejando  $C$  obtenemos

$$C = \frac{3\rho}{T^3}$$

es decir,  $\bar{u}(r) = \frac{3\rho}{T^3}(T - r)$ .  $C$  está claramente bien definida ya que  $T > 0$  y al ser  $\rho > 0$ ,  $C$  también es mayor a cero, luego  $\bar{u}(r) \geq 0$ . Veamos que la proyección es correcta y satisface el Teorema 2.3

$$\int_0^T (0 - \bar{u}(r))(v(r) - \bar{u}(r)) dr = \int_0^T -\frac{3\rho}{T^3}(T - r) \left( v(r) - \frac{3\rho}{T^3}(T - r) \right) dr = -\frac{3\rho}{T^3} \int_0^T (T - r)v(r) dr + \frac{9\rho^2}{T^6} \int_0^T (T - r)^2 dr$$

y como por la restricción  $\int_0^T (T - r)v(r) dr = \rho$

$$-\frac{3\rho^2}{T^3} + \frac{9\rho^2}{T^6} \frac{T^3}{3} = -\frac{3\rho^2}{T^3} + \frac{3\rho^2}{T^3} = 0 \quad \forall v \in \mathbb{K}$$

por lo que se cumple la desigualdad.

Este sería un ejemplo muy sencillo de una aplicación de las proyecciones. Sin embargo, no se ajusta a la realidad, ya que estamos suponiendo una masa constante, cuando el 85 – 90 % de la masa de un cohete es combustible, que se va consumiendo a medida que se eleva.

A continuación, estudiaremos el caso donde la masa del cohete varía en función del tiempo,  $m(t)$ , a medida que el combustible se va consumiendo. Supondremos que el combustible se agota en el momento  $t = T$  y la masa final del cohete  $m_f$  será igual a la masa original  $m_0$  menos la masa del combustible expulsado. Además, llamamos  $v_e$  a la velocidad de empuje de los gases en su expulsión, que consideramos constante.

Sabiendo todo esto, y modificando el problema anterior tenemos, por la segunda ley de Newton

$$\begin{cases} m(t)x''(t) = -m(t)g - v_e m'(t) \\ \quad \quad \quad \begin{matrix} >0 & & <0 & & >0 \end{matrix} \\ x(0) = 0 \\ x'(0) = 0 \end{cases}$$

donde despejando obtenemos

$$x''(t) = -g - \frac{v_e m'(t)}{m(t)}.$$

En este caso  $-\frac{v_e m'(t)}{m(t)}$  será nuestro  $u(t)$ , es decir, nuestra fuerza de empuje vertical. Por paralelismo con el primer ejemplo tenemos

$$\frac{-v_e m'(t)}{m(t)} = \bar{u}(t) = C(T - t) \implies \frac{m'(t)}{m(t)} = \frac{-C}{v_e}(T - t)$$

y resolviendo la EDO obtendríamos que  $m(t) = m_0 \cdot \exp\left(-\frac{C}{v_e}\left(Tt - \frac{t^2}{2}\right)\right)$ .

Sabiendo esto, calculamos  $x(t)$ . Sabemos que

$$x''(t) = -g + C(T - t)$$

que integrando una vez nos da

$$x'(t) = -gt + C \int (T - t) dt = -gt + C\left(Tt - \frac{t^2}{2}\right) + C_1 \stackrel{x'(0)=0}{=} -gt + C\left(Tt - \frac{t^2}{2}\right).$$

Integrando nuevamente

$$x(t) = -\frac{gt^2}{2} + C\left(\frac{Tt^2}{2} - \frac{t^3}{6}\right) + C_2 \stackrel{x(0)=0}{=} -\frac{gt^2}{2} + C\left(\frac{Tt^2}{2} - \frac{t^3}{6}\right).$$

Sabiendo que  $x(T) = h$  obtenemos que la constante  $C$  es

$$h = -\frac{gT^2}{2} + C\left(\frac{T^3}{2} - \frac{T^3}{6}\right) = -\frac{gT^2}{2} + C\left(\frac{T^3}{3}\right) \implies C = \frac{3}{T^3}\left(h + \frac{gT^2}{2}\right)$$

por lo tanto

$$\bar{u}(t) = \frac{3}{T^3}\left(h + \frac{gT^2}{2}\right)(T - t).$$

Si dividimos ambos lados de la ecuación entre la gravedad, obtenemos lo que en física se denomina **relación empuje-peso**, que es un parámetro característico de los motores de los cohetes y aviones que sirve para hacer comparaciones cuantitativas entre ellos. Existen muchos factores que afectan a la relación empuje-peso. Para poder compararlas, el empuje debería ser medido bajo condiciones controladas. Veamos nuestro caso

$$\frac{\bar{u}(t)}{g} = C(T - t) = \frac{3}{gT^3}\left(h + \frac{gT^2}{2}\right)(T - t)$$

y como nos interesa la fuerza de empuje en  $t = 0$  obtenemos

$$\frac{\bar{u}(0)}{g} = \frac{3}{gT^2}\left(h + \frac{gT^2}{2}\right) = \frac{3}{2} + \frac{3h}{gT^2} \geq \frac{3}{2}.$$

Es decir, concluimos que el cohete deberá tener una relación empuje-peso de al menos 1,5 para que despegue. Teóricamente, si la relación empuje-peso es mayor a 1, el cohete despegue; sin embargo, en la realidad no solo queremos que despegue, sino que lo haga con ciertas garantías y una cierta estabilidad, por lo que tendrá que ser algo mayor a 1, dato que coincide con lo que hemos obtenido. A continuación, se presenta una tabla con datos de cohetes reales.

Cohete	Empuje (kN)	Peso en despegue (kN)	Relación Empuje-Peso
Saturno V [9]	34.020	25.420	1,34
Transbordador Espacial [10]	30.160	20.000	1,5
Falcon 9 [11]	7.607	5.497	1,38
Starship (Space X) [12]	74.000	45.000	1,64

Cuadro 4.1: Relación empuje-peso de diferentes cohetes en el despegue.

Como podemos comprobar, en la mayoría de los lanzamientos, esta relación oscila entre 1,3 y 1,7, con lo que podemos concluir que nuestro cálculo, a pesar de no contemplar distintas variables, como los cambios de gravedad, la resistencia del viento y la velocidad de escape, es una buena aproximación que concuerda con la realidad. Además, observamos que cuanto mayor sea  $h$  o menor sea  $T$ , mayor deberá ser la relación empuje-peso, lo cual parece muy lógico.

### 4.3.2. Problema control óptimo

En este apartado, introducimos los problemas de control óptimo, que permiten encontrar la mejor forma de controlar un sistema dinámico para optimizar cierto criterio de rendimiento. Este tipo de problemas son ampliamente utilizados en diversos campos como la economía, la física, o incluso, la medicina, particularmente en el tratamiento del cáncer, para EDOs más complicadas que la EDO lineal que presentamos. En este contexto, se busca determinar una política de administración de dosis que permita alcanzar un estado deseado final del paciente, representado por  $x_T$ , al mismo tiempo que se minimiza el coste económico y se procura, en la mayor medida posible, preservar el bienestar y la calidad de vida del paciente durante el transcurso del tratamiento,  $x_d$ . Estos aspectos se ponderan mediante los pesos, que reflejan la importancia relativa que se otorga a cada uno. De esta forma, se busca encontrar la estrategia óptima que proporcione el mejor tratamiento posible, equilibrando eficacia clínica y viabilidad económica.

Consideramos la Ecuación del Estado

$$\begin{cases} x'(t) = a(t)x(t) + b(t)u(t) \\ x(0) = x_0 \end{cases}$$

donde  $a, b \in C[0, T]$  son funciones conocidas y  $x_0$  y  $T$  son datos del problema. Definiendo nuestro problema de control óptimo, tenemos

$$(PCO) \underset{u \in \mathbb{K}}{\text{Minimizar}} J(u) = \frac{\alpha}{2} \int_0^T (x_u(t) - x_d(t))^2 dt + \frac{\beta}{2} \int_0^T u^2(t) dt + \frac{\gamma}{2} (x_u(T) - x_T)^2$$

donde la primera integral representa la cercanía a un estado deseado  $x_d$  a lo largo del proceso, la segunda el coste del proceso y la última el resultado al final del proceso. La importancia que se le da a cada una de ellas se puede controlar mediante los “pesos”  $\alpha, \beta, \gamma \in [0, +\infty)$ , que serán conocidos. Además,  $x_d \in L^2[0, T]$  y  $x_T \in \mathbb{R}$  también serán conocidos.

El conjunto de posibles soluciones será

$$\mathbb{K} = \{u \in L^2[0, T] : \delta \leq u(t) \leq \rho \text{ c.t.p. } t \in [0, T]\}$$

con  $\delta$  y  $\rho$  conocidos.  $x_u$  denotará la única solución continua en  $[0, T]$  de la Ecuación del Estado para cada  $u \in L^2[0, T]$ .

**Teorema 4.6.** *Si  $\beta > 0$  o  $\delta, \rho \in \mathbb{R}$  con  $\delta < \rho$ , el problema de control óptimo (PCO) posee una única solución  $\bar{u}$  (que se denomina control óptimo).*

*Demostración.* En primer lugar, sabemos que  $\mathbb{K}$  es convexo y cerrado, como se ha demostrado para conjuntos análogos en la Sección 2.3, al estar definido por cotas en  $L^2[0, T]$ .

Sabemos que  $C = \{J(u) : u \in \mathbb{K}\} \subset [0, +\infty)$ . Por el axioma del ínfimo, existe  $\inf C = \bar{c} \in [0, +\infty)$ . Por tanto, para cada  $n \in \mathbb{N}$ , existe  $u_n \in \mathbb{K}$  tal que

$$\bar{c} \leq J(u_n) \leq \bar{c} + \frac{1}{n}$$

luego  $\{u_n\}_{n \in \mathbb{N}}$  es una sucesión minimizante y  $\lim_{n \rightarrow \infty} J(u_n) = \bar{c}$ .

- Si  $\beta > 0$ ,  $\frac{\beta}{2} \|u_n\|_{L^2[0, T]}^2 \leq J(u_n) \leq \bar{c} + 1 \implies \{u_n\}_{n \in \mathbb{N}} \subset L^2[0, T]$  es una sucesión acotada.
- Si  $\beta = 0$ , pero  $\delta, \rho \in \mathbb{R}$  con  $\delta \leq u_n(t) \leq \rho \forall n$  c.t.p.  $t \in [0, T] \implies \{u_n\}_{n \in \mathbb{N}} \subset L^2[0, T]$  es una sucesión acotada.

Como  $L^2[0, T]$  es un espacio de Hilbert, existe una subsucesión  $\{u_{n_k}\}_{k \in \mathbb{N}}$  débilmente convergente hacia  $\bar{u} \in L^2[0, T]$ . Además, como  $\mathbb{K}$  es convexo y cerrado, entonces también es débilmente cerrado, luego  $\bar{u} \in \mathbb{K}$  (Teorema 3.7, [1]).

Queremos ver que  $J(\bar{u}) = \bar{c}$ . Denotamos  $x_{u_{n_k}} = x_{n_k}$ . Tenemos

$$J(u_{n_k}) = \frac{\alpha}{2} \int_0^T (x_{n_k}(t) - x_d(t))^2 dt + \frac{\beta}{2} \int_0^T u_{n_k}^2(t) dt + \frac{\gamma}{2} (x_{n_k}(T) - x_T)^2$$

Por la Proposición 3.5, [1], sabemos que  $\liminf \|u_{n_k}\|_{L^2}^2 \geq \|\bar{u}\|_{L^2}^2$ . Por otro lado, por el método de variación de constantes calculamos que

$$x_{n_k}(t) = x_0 e^{\int_0^t a(s) ds} + e^{\int_0^t a(s) ds} \int_0^t b(r) u_{n_k}(r) e^{-\int_0^r a(s) ds} dr$$

y como  $u_{n_k} \rightarrow \bar{u}$  débilmente en  $L^2[0, T]$ , tenemos que

$$\lim_{n \rightarrow \infty} \int_0^T u_{n_k}(t) f(t) dt = \int_0^T \bar{u}(t) f(t) dt \quad \forall f \in L^2[0, T]$$

por lo tanto, aplicando la regla del sándwich, tenemos

$$\bar{c} = \liminf J(u_{n_k}) \geq J(\bar{u}) \geq \bar{u} \implies J(\bar{u}) = \bar{c}$$

□

Veamos ahora que cumple las condiciones necesarias de optimalidad de primer orden. Para simplificar la notación, denotamos  $\bar{x} = x_{\bar{u}}$  y  $x_\lambda = x_{\lambda u + (1-\lambda)\bar{u}}$  con  $u \in \mathbb{K}$  y  $\lambda \in (0, 1)$ . A  $\bar{x}$  se le denomina estado óptimo.

Por ser  $\bar{u}$  la solución del (PC) se verifica que

$$\begin{aligned} J(\bar{u}) &\leq J(\lambda u + (1-\lambda)\bar{u}) \quad \forall \lambda \in (0, 1) \quad \forall u \in \mathbb{K} \\ \iff &\frac{\alpha}{2} \int_0^T (\bar{x}(t) - x_d(t))^2 dt + \frac{\beta}{2} \int_0^T \bar{u}^2(t) dt + \frac{\gamma}{2} (\bar{x}(T) - x_T)^2 \leq \\ &\leq \frac{\alpha}{2} \int_0^T (x_\lambda(t) - x_d(t))^2 dt + \frac{\beta}{2} \int_0^T (\lambda u(t) + (1-\lambda)\bar{u}(t))^2 dt + \frac{\gamma}{2} (x_\lambda(T) - x_T)^2 \end{aligned}$$

y desarrollando los cuadrados

$$0 \leq \frac{\alpha}{2} \int_0^T x_\lambda^2(t) - 2x_\lambda(t)x_d(t) - \bar{x}^2(t) + 2\bar{x}(t)x_d(t) dt +$$

$$+ \frac{\beta}{2} \int_0^T \lambda^2 (u(t) - \bar{u}(t))^2 + 2\lambda(u(t) - \bar{u}(t))\bar{u}(t) dt + \frac{\gamma}{2} (x_\lambda^2(T) - 2x_\lambda(T)x_T - \bar{x}^2(T) + 2\bar{x}(T)x_T). \quad (4.1)$$

Llamamos  $z(t) = \frac{x_\lambda(t) - \bar{x}(t)}{\lambda}$  y vemos que **no depende** de  $\lambda$  porque es la única solución del problema,

$$\begin{cases} z'(t) = a(t)z(t) + (u(t) - \bar{u}(t))b(t) \\ z(0) = 0 \end{cases}$$

Dividiendo la ecuación (4.1) por  $\lambda \in (0, 1)$  obtenemos

$$\begin{aligned} & \frac{\alpha}{2} \int_0^T z(t)(x_\lambda(t) + \bar{x}(t)) - 2x_d(t)z(t) dt + \\ & + \frac{\beta}{2} \int_0^T \lambda (u(t) - \bar{u}(t))^2 + 2(u(t) - \bar{u}(t))\bar{u}(t) dt + \frac{\gamma}{2} (z(T)(x_\lambda(T) + \bar{x}(T)) - 2x_T z(T)) \geq 0. \end{aligned}$$

Tomando el límite  $\lambda \rightarrow 0$  en esta expresión,  $x_\lambda \rightarrow \bar{x}$  y resulta

$$\alpha \int_0^T z(t)(\bar{x}(t) - x_d(t)) dt + \beta \int_0^T (u(t) - \bar{u}(t))\bar{u}(t) dt + \gamma z(T)(\bar{x}(T) - x_T) \geq 0 \quad \forall u \in \mathbb{K}. \quad (4.2)$$

Para facilitar la expresión y solo tener que calcularla una vez independientemente de cuál sea  $u$ , introducimos el **estado adjunto**  $p(t)$  como la única solución de

$$\begin{cases} -p'(t) = a(t)p(t) + \alpha(\bar{x}(t) - x_d(t)) & t \in [0, T] \\ p(T) = \gamma(\bar{x}(T) - x_T) \end{cases}$$

Integrando por partes tenemos

$$\begin{aligned} \int_0^T z(t)p'(t) dt &= z(t)p(t) \Big|_{t=0}^{t=T} - \int_0^T p(t)z'(t) dt \stackrel{z(0)=0}{\iff} \\ &= \int_0^T z(t)p'(t) dt + \int_0^T p(t)z'(t) dt = z(T)p(T) = z(T)\gamma(\bar{x}(T) - x_T) \end{aligned}$$

donde sustituyendo y simplificando obtenemos

$$\begin{aligned} \int_0^T z(t)[-a(t)p(t) - \alpha(\bar{x}(t) - x_d)] dt + \int_0^T p(t)[a(t)z(t) + (u(t) - \bar{u}(t))b(t)] dt &= \gamma z(T)(\bar{x}(T) - x_T) \iff \\ \alpha \int_0^T z(t)(\bar{x}(t) - x_d(t)) dt + \gamma z(T)(\bar{x}(T) - x_T) &= \int_0^T p(t)(u(t) - \bar{u}(t))b(t) dt. \end{aligned}$$

Por lo tanto (4.2) es equivalente a

$$\int_0^T (p(t)b(t) + \beta\bar{u}(t))(u(t) - \bar{u}(t)) dt \geq 0 \quad \forall u \in \mathbb{K} \iff \bar{u}(t) = P_{\mathbb{K}} \left( \frac{-p(t)b(t)}{\beta} \right) \quad (4.3)$$

**Ejemplo 4.7.** Consideramos la Ecuación del Estado

$$(EE) = \begin{cases} \bar{x}'(t) = -\bar{x}(t) + \bar{u}(t) \\ \bar{x}(0) = \bar{x}_0 \end{cases}$$

queremos minimizar

$$J(u) = \frac{1}{2} \int_0^T u^2(t) dt + \frac{1}{2} x(T)^2$$

con restricción de cota  $-1 \leq u(t) \leq 0$ . Vemos que  $a(t) \equiv -1, b(t) \equiv 1, \alpha = 0, \beta = 1, \gamma = 1$ . La ecuación del estado adjunto, al ser  $\alpha = 0$ , se simplifica a

$$\begin{cases} -p'(t) = -p(t) \\ p(T) = \bar{x}(T) \end{cases}$$

y por lo tanto  $p(t) = \bar{x}(T)e^{t-T}$ . Aplicando ahora el resultado (4.3) obtenemos

$$\bar{u}(t) = P_{[-1,0]} \left( \frac{-p(t)b(t)}{\beta} \right) = P_{[-1,0]}(-p(t)) = \begin{cases} 0 & \text{si } p(t) < 0 \\ -p(t) & \text{si } p(t) \in [0, 1] = \text{máx}(-1, \text{mín}(0, -p(t))) \\ -1 & \text{si } p(t) > 1 \end{cases}$$

Ahora resolvemos la ecuación del estado estado

$$p(t) \in [0, 1] \iff 0 \leq \bar{x}(T)e^{t-T} \leq 1$$

$$\bar{x}'(t) = -\bar{x}(t) - \bar{x}(T)e^{t-T} \implies \bar{x}(t) = x_0e^{-t} + e^{-t} \int_0^t e^s \bar{u}(s) ds \implies$$

$$\implies \bar{x}(T) = x_0e^{-T} + e^{-T} \int_0^T e^s \bar{u}(s) ds \stackrel{\bar{u} \leq 0}{\leq} x_0e^{-T}$$

$$\begin{cases} \text{si } x_0 \leq e^T & \implies \bar{x}(T) \leq 1 & \implies \bar{p}(t) \leq 1 \\ \text{si } e^T - 1 \leq x_0 & \implies \bar{x}(T) \geq 0 & \implies \bar{p}(t) \geq 0 \end{cases} \implies \bar{u}(t) = -p(t)$$

es decir, si  $x_0 \in [e^T - 1, e^T] \implies \bar{x}(T) \in [0, 1] \implies p(t) \in [0, 1] \implies \bar{u}(t) = -p(t)$  que implica que el control óptimo no está saturado. Por tanto, bajo esta condición, integrando obtenemos

$$\implies \bar{x}(t) = x_0e^{-t} - \bar{x}(T)e^{-t} \int_0^t e^{2s-T} ds = x_0e^{-t} - \bar{x}(T)e^{-t-T} \frac{e^{2t} - 1}{2} \stackrel{t=T}{\implies} \bar{x}(T) = \frac{2x_0e^{-T}}{3 - e^{-2T}}$$

Por tanto

$$\bar{u}(t) = -p(t) = -\bar{x}(T)e^{t-T} = \frac{2x_0e^{t-2T}}{e^{-2T} - 3} = \frac{2x_0e^{-2T}}{e^{-2T} - 3}e^t$$

y podemos obtener  $\bar{x}(t)$  sustituyendo

$$\bar{x}(t) = x_0e^{-t} - \frac{x_0e^{-2T-t}}{3 - e^{-2T}}(e^{2t} - 1)$$

**Interpretación del resultado:** Supongamos de nuevo que nos encontramos en un contexto médico, en el que  $-\bar{u}(t)$  representa la tasa óptima de administración de un medicamento en el tiempo  $t$  dentro del intervalo  $[0, T]$ . Vemos que la solución es una función exponencial creciente en el tiempo. Esto significa que el medicamento debe administrarse cada vez más intensamente a medida que avanza el tratamiento. La dosis óptima empieza pequeña y aumenta de forma acelerada hacia el final. Algunas interpretaciones de esto pueden ser que el efecto del medicamento decae con el tiempo (con lo que es necesario compensarlo con un aumento progresivo de la dosis), o que el objetivo clínico está fuertemente concentrado al final del tratamiento, por ejemplo, eliminar por completo el tumor en un tiempo dado, o que se trata de una preparación para una operación, lo que obliga a reforzar el tratamiento al final.



# Apéndice A

## Preliminares

En este apéndice, se presentan una serie de **definiciones** y **resultados teóricos fundamentales**, obtenidos en su totalidad de [1], muchos de los cuales han sido estudiados a lo largo del grado, y que serán empleados en el desarrollo del presente trabajo. Asimismo, se establecerá la **notación** que se utiliza a lo largo del documento, con el objetivo de fijar un marco teórico claro y situar el contexto en el que se enmarca nuestro estudio.

### A.1. Espacios Hilbert

**Definición A.1.** Sea  $H$  un espacio vectorial sobre el cuerpo  $\mathbb{C}$ , y  $(\cdot, \cdot) : H \times H \rightarrow \mathbb{C}$  un producto interno que satisface las siguientes propiedades para todo  $x, y, z \in H$  y  $\alpha \in \mathbb{C}$ :

1. **Sesquilineal:**

$$(\alpha x + y, z) = \alpha(x, z) + (y, z)$$

2. **Hermítica:**

$$(x, y) = \overline{(y, x)}$$

donde  $\overline{(y, x)}$  denota el conjugado complejo. En  $\mathbb{C}$ , el producto interno es conjugado simétrico. En el caso real, esta propiedad se reduce a la simetría usual,  $(x, y) = (y, x)$ . Todos los resultados válidos en  $\mathbb{C}$  siguen siendo correctos en  $\mathbb{R}$  con esta simplificación.

3. **Definida positiva:**

$$(x, x) \geq 0 \quad \text{y} \quad (x, x) = 0 \iff x = 0$$

Se dice que  $(H, (\cdot, \cdot))$  es un **espacio prehilbert**.

Consideramos  $H$  un espacio vectorial sobre el cuerpo  $\mathbb{C}$ , dotado de un producto interno  $(\cdot, \cdot) : H \times H \rightarrow \mathbb{C}$ , el cual induce una norma dada por

$$\|x\| := \sqrt{(x, x)}$$

**Definición A.2.** Se dice que una sucesión  $\{x_n\}_{n \in \mathbb{N}} \subset H$  es una **sucesión de Cauchy** si

$$\forall \epsilon > 0, \text{ existe } n_\epsilon \in \mathbb{N} \text{ tal que } \|x_n - x_m\| < \epsilon \quad \forall n, m > n_\epsilon.$$

**Definición A.3.** Decimos que  $H$  es **completo** si toda sucesión de Cauchy en  $H$  con respecto a esta norma converge en  $H$ .

**Definición A.4.** Sea  $H$  un espacio prehilbert. Se dice que  $H$  es un **espacio de Hilbert** si es completo con respecto a la norma inducida por su producto interno.

**Corolario A.5. Identidad del paralelogramo.**

Sea  $H$  un espacio de Hilbert. Sean  $u, v \in H$  se cumple,

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

*Demostración.* Desarrollamos ambos cuadrados:

$$\begin{aligned}\|u + v\|^2 &= (u + v, u + v) = (u, u) + (u, v) + (v, u) + (v, v) \\ \|u - v\|^2 &= (u - v, u - v) = (u, u) - (u, v) - (v, u) + (v, v)\end{aligned}$$

Sumando ambas expresiones

$$\begin{aligned}\|u + v\|^2 + \|u - v\|^2 &= [(u, u) + (u, v) + (v, u) + (v, v)] + [(u, u) - (u, v) - (v, u) + (v, v)] = \\ &= 2(u, u) + 2(v, v) = 2\|u\|^2 + 2\|v\|^2.\end{aligned}$$

□

**Lema A.6. Desigualdad de Cauchy-Schwarz.**

Sea  $H$  un espacio de Hilbert. Entonces,  $\forall u, v \in H$  se tiene que

$$|(u, v)| \leq \|u\| \|v\| \tag{A.1}$$

*Demostración.* Si  $v = 0$ , entonces se cumple que  $(u, v) = 0 \quad \forall u$ , por lo que  $0 = |(u, v)| \leq \|u\| \cdot 0 = 0$ . Si  $v \neq 0$ , consideramos la siguiente función. Sea  $\lambda \in \mathbb{C}$

$$f(\lambda) = \|u - \lambda v\|^2 = (u - \lambda v, u - \lambda v) = \|u\|^2 - \lambda(u, v) - \bar{\lambda}(v, u) + |\lambda|^2 \|v\|^2 = \|u\|^2 - 2\Re(\overline{\lambda}(u, v)) + |\lambda|^2 \|v\|^2$$

Y sabemos que  $f(\lambda) \geq 0$ . Tomamos

$$\lambda = \frac{(v, u)}{\|v\|^2} \in \mathbb{C}$$

que es válido, ya que  $v \neq 0$ . Sustituyendo en  $f(\lambda)$  tenemos

$$f(\lambda) = \|u\|^2 - 2\frac{|(u, v)|^2}{\|v\|^2} + \frac{|(u, v)|^2}{\|v\|^2} = \|u\|^2 - \frac{|(u, v)|^2}{\|v\|^2}$$

y dado que  $f(\lambda) \geq 0$  concluimos

$$\|u\|^2 - \frac{|(u, v)|^2}{\|v\|^2} \geq 0 \implies |(u, v)|^2 \leq \|u\|^2 \|v\|^2 \implies |(u, v)| \leq \|u\| \|v\|$$

□

**Teorema A.7.** Sea  $H$  un espacio de Hilbert y  $\mathbb{K}$  un subespacio de  $H$  de dimensión finita, entonces  $\mathbb{K}$  es un subespacio cerrado en  $H$ .

*Demostración.* Supongamos que  $\mathbb{K}$  tiene dimensión finita  $n$ . Entonces existe una base ortonormal tal que

$$\mathbb{K} = \left\{ f : f = \sum_{i=1}^n c_i \phi_i; \quad c_i \in \mathbb{C} \right\}$$

Por lo tanto, cualquier elemento de  $\mathbb{K}$  puede escribirse como combinación lineal finita de estas funciones. Como  $\mathbb{K}$  es de dimensión finita, la norma inducida por  $H$  en  $\mathbb{K}$  es equivalente a cualquier otra norma en  $\mathbb{K}$ . Esto implica que la norma de toda sucesión de Cauchy respecto de la norma de  $H$  también es de Cauchy respecto a la topología finito-dimensional de  $\mathbb{K}$ .

Dada  $\{f_k\}_{k \in \mathbb{N}} \subset \mathbb{K}$  una sucesión de Cauchy en  $H$ , existe un conjunto finito de coordenadas  $(c_i^{(k)})$  tales que

$$f_k = \sum_{i=1}^n c_i^{(k)} \phi_i$$

Como  $\{f_k\}$  es de Cauchy, entonces para cualquier  $\epsilon > 0$  existe un  $s \in \mathbb{N}$  tal que para  $k, m \geq s$ :

$$\|f_k - f_m\| \leq \epsilon$$

utilizando la representación de  $f_k$  y  $f_m$  tenemos

$$f_k - f_m = \sum_{i=1}^n (c_i^{(k)} - c_i^{(m)}) \phi_i$$

entonces la norma de  $H$  satisface

$$\|f_k - f_m\| = \left\| \sum_{i=1}^n (c_i^{(k)} - c_i^{(m)}) \phi_i \right\|$$

Como  $\{\phi_1, \phi_2, \dots, \phi_n\}$  es una base ortonormal, la norma se puede expresar en términos de las coordenadas

$$\|f_k - f_m\| = \sqrt{\sum_{i=1}^n |c_i^{(k)} - c_i^{(m)}|^2}$$

Como  $\|f_k - f_m\| \leq \epsilon$  para  $k, m \geq s$  se deduce que cada término de la suma  $|c_i^{(k)} - c_i^{(m)}|$  debe satisfacer

$$|c_i^{(k)} - c_i^{(m)}| \leq \|f_k - f_m\|$$

por lo tanto, para cada  $i = 1, 2, \dots, n$ , la sucesión  $\{c_i^{(k)}\}$  también satisface la condición de Cauchy.

Como  $\mathbb{C}$  es completo, las sucesiones  $\{c_i^{(k)}\}$  convergen a ciertos valores  $c_i \in \mathbb{C}$  y si definimos

$$f = \sum_{i=1}^n c_i \phi_i \in \mathbb{K}$$

se tiene que  $\{f_k\}$  converge a  $f$ . Por lo tanto, hemos demostrado que toda sucesión de Cauchy en  $\mathbb{K}$  con respecto a la norma de  $H$  converge a un punto en  $\mathbb{K}$ , y por tanto, es cerrado.  $\square$

**Definición A.8.** Dado un conjunto medible  $\Omega \subseteq \mathbb{R}^n$ , el espacio  $L^2(\Omega)$  se define como el conjunto de funciones medibles  $f: \Omega \rightarrow \mathbb{C}$  tales que

$$\int_{\Omega} |f(x)|^2 dx < +\infty$$

Es decir,  $L^2(\Omega)$  está formado por las clases de equivalencia de funciones medibles cuadrado-integrables. Este es un espacio vectorial sobre  $\mathbb{C}$ , y se puede dotar de un producto interno definido por:

$$(f, g) = \int_{\Omega} f(x) \cdot \overline{g(x)} dx$$

para todas las funciones  $f, g$  pertenecientes a  $L^2(\Omega)$ . La norma inducida por este producto interno es:

$$\|f\|_{L^2} = \sqrt{\int_{\Omega} |f(x)|^2 dx}$$

Con este producto interno, el espacio  $L^2(\Omega)$  es un **espacio de Hilbert**, es decir, es completo con respecto a la norma definida por ese producto interno.

A continuación, presentamos los casos que hemos empleado a lo largo del trabajo, y daremos un ejemplo de cada uno de ellos.

**1. Caso real en  $\Omega = [a, b]$**

El espacio  $L^2[a, b]$  se define como el conjunto de funciones reales  $f : [a, b] \rightarrow \mathbb{R}$  tales que

$$\int_a^b |f(x)|^2 dx < \infty$$

**Ejemplo A.9.** Sea  $f(x) = x$  definida en el intervalo  $[0, 1]$ .

$$\int_0^1 x^2 dx = \frac{1}{3} < \infty$$

Por lo tanto,  $f \in L^2[0, 1]$ .

**2. Caso complejo en  $\Omega = [a, b]$**

El espacio  $L^2([a, b], \mathbb{C})$  se compone de funciones complejas  $f : [a, b] \rightarrow \mathbb{C}$  tales que

$$\int_a^b |f(x)|^2 dx < \infty$$

**Ejemplo A.10.** Sea  $f(x) = e^{i\pi x}$ , función compleja definida en el intervalo  $[0, 1]$ . Esta función tiene módulo constante

$$|f(x)| = |e^{i\pi x}| = 1$$

Entonces

$$\int_0^1 |f(x)|^2 dx = 1 < \infty$$

Por lo tanto,  $f \in L^2([0, 1], \mathbb{C})$ .

**3. Caso real en  $\Omega = \mathbb{R}$**

El espacio  $L^2(\mathbb{R})$  contiene todas las funciones medibles reales  $f : \mathbb{R} \rightarrow \mathbb{R}$  tales que:

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty$$

Es decir, la función debe tener cuadrado integrable en toda la recta real.

**Ejemplo A.11.** Sea  $f(x) = e^{-\frac{x^2}{2}}$  una función real definida en toda la recta real. Para comprobar que  $f \in L^2(\mathbb{R})$ , calculamos la integral de su módulo al cuadrado,

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} e^{-x^2} dx$$

que es la conocida integral de Gauss

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} < +\infty.$$

Este resultado muestra que  $f \in L^2(\mathbb{R})$ , ya que su norma cuadrada es finita.

**4. Caso complejo en  $\Omega = \mathbb{R}$**

El espacio  $L^2(\mathbb{R}, \mathbb{C})$  contiene todas las funciones medibles complejas  $f : \mathbb{R} \rightarrow \mathbb{C}$  tales que

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$$

**Ejemplo A.12.** Sea  $f(x) = \frac{e^{ix}}{1+x^2}$ , función compleja en  $\mathbb{R}$ .  
Aquí

$$|f(x)|^2 = \left| \frac{e^{ix}}{1+x^2} \right|^2 = \frac{1}{(1+x^2)^2}$$

entonces

$$\int_{-\infty}^{\infty} \frac{1}{(1+x^2)^2} dx = \frac{\pi}{2} < +\infty$$

Esta integral también es conocida, así que  $f \in L^2(\mathbb{R}, \mathbb{C})$ .

A continuación, recordamos el teorema de la convergencia dominada. Este teorema ha sido obtenido de [1]. Su demostración ha sido vista durante el grado, por lo que no la incluimos.

**Teorema A.13. Teorema de la Convergencia Dominada.**

Sea  $\{f_n\}$  una sucesión de funciones  $L^1(\Omega)$ . Supongamos que

1.  $f_n(x) \rightarrow f(x)$  c.t.p. en  $\Omega$ .
2. existe una función  $g \in L^1$  tal que para cada  $n$ ,  $|f_n(x)| \leq g(x)$  c.t.p. en  $\Omega$ .

Entonces  $f \in L^1(\Omega)$  y  $\|f_n - f\|_{L^1} \rightarrow 0$ .

**Teorema A.14.**  $L^2(\Omega)$  es un espacio completo.

*Demostración.* Sea  $\{f_n\}$  una sucesión de Cauchy en  $L^2(\Omega)$ . Veamos que posee una subsucesión convergente en  $L^2(\Omega)$ . Tomamos  $\{f_{n_k}\}$  tal que

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^2(\Omega)} \leq \frac{1}{2^k} \quad \forall k \geq 1 \quad (\text{A.2})$$

(Llegamos a  $\frac{1}{2^k}$  de la siguiente manera: existe  $n_1$  tal que  $\|f_m - f_n\|_{L^2(\Omega)} \leq \frac{1}{2}$  para  $n, m > n_1$ ; después se toma  $n_2 > n_1$  tal que  $\|f_m - f_n\|_{L^2(\Omega)} \leq \frac{1}{2^2}$ , y continuando la recurrencia llegamos a  $n_k$  donde  $\|f_m - f_n\|_{L^2(\Omega)} \leq \frac{1}{2^k}$ )

Demostramos ahora que  $\{f_{n_k}\}$  converge en  $L^2(\Omega)$ . Para simplificar la notación escribiremos  $f_k$  en lugar de  $f_{n_k}$ . Denotamos

$$g_n(x) = \sum_{k=1}^n |f_{k+1}(x) - f_k(x)|$$

de donde

$$\|g_n\|_{L^2(\Omega)} \leq \sum_{k=1}^n \frac{1}{2^k} \leq 1 \quad (\text{A.3})$$

Aplicamos ahora el teorema de la convergencia monótona para concluir que  $g_n(x)$  converge a un límite, que denotamos  $g(x)$ , con  $g \in L^2[\Omega]$ . Veamos que se cumplen las condiciones de dicho teorema:

1) No negatividad: Por construcción  $g_n(x) = \sum_{k=1}^n |f_{k+1}(x) - f_k(x)| \geq 0$

2) Monótona creciente: Cada término nuevo es el anterior más la suma de un valor no negativo

$$g_{n+1}(x) = g_n(x) + |f_{n+2}(x) - f_{n+1}(x)| \geq g_n(x)$$

3) Acotación en norma  $L^2(\Omega)$ : Como ya hemos visto  $\|g_n\|_{L^2(\Omega)} \leq 1$ .  
Por lo tanto, del teorema se deduce que

$$\int_{\Omega} g(x)^2 dx = \lim_{n \rightarrow \infty} \int_{\Omega} g_n(x)^2 dx$$

y utilizando (A.3) llegamos a

$$\int_{\Omega} g(x)^2 dx = \lim_{n \rightarrow \infty} \int_{\Omega} g_n(x)^2 dx \leq 1$$

luego  $g \in L^2(\Omega)$ . Por otra parte, para cada  $m \geq n \geq 2$  se verifica, por la desigualdad triangular,

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f_{m-1}(x)| + \dots + |f_{n+1}(x) - f_n(x)| \leq \sum_{k=n}^{\infty} |f_{k+1}(x) - f_k(x)| = g(x) - g_{n-1}(x)$$

Por tanto,  $\forall \varepsilon > 0$  existe  $n_0$  tal que  $g(x) - g_{n-1}(x) < \varepsilon$  para  $n \geq n_0$  y por lo tanto llegamos a que  $|f_m(x) - f_n(x)| \leq g(x) - g_{n-1}(x) < \varepsilon$ . Y por ello obtenemos que en c.t.p. en  $\Omega$ ,  $(f_n(x))$  es de Cauchy y converge a un límite denotado por  $f(x)$ .

Tenemos que en c.t.p. en  $\Omega$

$$|f(x) - f_n(x)| \leq g(x) - g_{n-1}(x) \leq g(x) \quad \text{para } n \geq 2 \quad (\text{A.4})$$

de donde resulta que  $f \in L^2(\Omega)$ , ya que,  $|f(x) - f_n(x)| \leq g(x)$ , luego  $|f(x) - f_n(x)|^2 \leq g(x)^2$ , y  $g(x)^2$  actúa como una función mayorante integrable, lo que asegura que  $|f(x)|^2$  es integrable, y por ende,  $f \in L^2(\Omega)$ . Por último, veamos que  $\|f_n - f\|_{L^2(\Omega)} \rightarrow 0$ ,

1) Como  $f_n(x) \rightarrow f(x)$  en c.t.p., entonces  $|f_n(x) - f(x)|^2 \rightarrow 0$  en c.t.p. en  $\Omega$ .

2)  $g^2$  actúa como una función mayorante integrable, como ya hemos visto.

Por lo tanto, cumple las hipótesis del teorema de la convergencia dominada y concluimos que

$$\|f_n - f\|_{L^2(\Omega)}^2 = \int_{\Omega} |f_n(x) - f(x)|^2 dx \rightarrow 0$$

□

**Teorema A.15.** Sea  $\{f_n\}_{n \in \mathbb{N}}$  una sucesión en  $L^2(\Omega)$  y  $f \in L^2(\Omega)$ , tales que  $\|f_n - f\|_{L^2(\Omega)} \rightarrow 0$ . Entonces existe una subsucesión  $\{f_{n_k}\}$  tal que

(a)  $f_{n_k}(x) \rightarrow f(x)$  c.t.p. en  $\Omega$ .

(b)  $|f_{n_k}(x)| \leq h(x) \quad \forall k$  y c.t.p. en  $\Omega$ , con  $h \in L^2(\Omega)$ .

*Demostración.* Como la sucesión  $\{f_n\}$  es de Cauchy, la demostración es análoga a la del teorema precedente, tomando una subsucesión  $\{f_{n_k}\}$  cumpliendo (A.2). De esta forma vemos que  $f_{n_k}(x)$  converge en c.t.p. a un límite, que designamos por  $f^*(x)$ . Además, obtenemos por (A.4) que

$$|f^*(x) - f_{n_k}(x)| \leq g(x) \quad \forall k, \quad \text{c.t.p. en } \Omega, \quad \text{con } g \in L^2(\Omega)$$

de donde se obtiene que  $f^* \in L^2(\Omega)$  y que  $f_{n_k} \rightarrow f^*$  en c.t.p. en  $L^2(\Omega)$ . Por consiguiente,  $f = f^*$  c.t.p. y se deduce (a). Para (b) basta elegir  $h = |f^*| + g$  ya que, de nuevo por el teorema anterior y la desigualdad triangular

$$|f_{n_k}(x)| \leq |f^*(x)| + |f^*(x) - f_{n_k}(x)| \leq |f^*(x)| + g(x)$$

y  $h$  pertenece claramente a  $L^2(\Omega)$  por ser suma de dos elementos que pertenecen a  $L^2(\Omega)$ . □

# Bibliografía

- [1] H. Brézis; *Análisis Funcional*, Alianza Universidad Textos, 1984.
- [2] F. Deutsch; *Best Approximation in Inner Product Spaces*, Canadian Mathematical Society, 2000.
- [3] Encyclopedia of Mathematics, *Plancherel theorem*. [https://encyclopediaofmath.org/index.php?title=Plancherel\\_theorem](https://encyclopediaofmath.org/index.php?title=Plancherel_theorem). Accedido: 2025-6-16.
- [4] L.A. Fernández; *Optimal control of a Fisher-KPP type model arising in chemotherapy for brain tumors*, Pure and Applied Functional Analysis, Volume 7, Number 5, 1648-1650, 2022.
- [5] L.A. Fernández; *Apuntes asignatura Introducción a las Ecuaciones en Derivadas Parciales*, Universidad de Cantabria, 2022.
- [6] E. Kopecká y S. Reich; *A Note on the Von Neumann Alterating Projections Algorithm*, Journal of Nonlinear and Convex Analysis, Volume 5, Number 3, 379-386, 2004.
- [7] D.C. Youla y H. Webb; *Image Restoration by the Method of Convex Projections*, IEEE Transactions on Medical Imaging, Volume MI-1, Number 2, 1982.
- [8] Wikipedia, *Projections onto convex sets*. [https://en.wikipedia.org/wiki/Projections\\_onto\\_convex\\_sets](https://en.wikipedia.org/wiki/Projections_onto_convex_sets). Accedido: 2024-10-15.
- [9] Wikipedia, *Saturno V*. [https://es.wikipedia.org/wiki/Saturno\\_Vs](https://es.wikipedia.org/wiki/Saturno_Vs). Accedido: 2025-3-14.
- [10] Wikipedia, *Transbordador STS*. [https://es.wikipedia.org/wiki/Transbordador\\_STS](https://es.wikipedia.org/wiki/Transbordador_STS). Accedido: 2025-3-14.
- [11] Wikipedia, *Falcon 9* [https://es.wikipedia.org/wiki/Falcon\\_9](https://es.wikipedia.org/wiki/Falcon_9). Accedido: 2025-3-14.
- [12] Wikipedia, *Starship (Space X)*. [https://es.wikipedia.org/wiki/Starship\\_\(SpaceX\)](https://es.wikipedia.org/wiki/Starship_(SpaceX)). Accedido: 2025-3-14.