El Teorema de Lax-Milgram, Generalizaciones y Aplicaciones

Autor: David Álvarez Melis

Director: Dr. Carlos Bosch Giral

Es imposible ser un matemático
sin ser un poeta en el alma.

Sofia Kovaleskaya (1850-1891)

La dernière démarche de la raison est de reconnaître
qu'il y a une infinité de choses qui la surprassent.

Blaise Pascal (1623-1662)

Agradecimientos

Quisiera agradecer...

Al Dr. Carlos Bosch por dedicar tiempo y esfuerzo a la dirección de esta tesis. El apoyo, consejos y amistad que me ha regalado durante estos años se quedarán conmigo como una de las cosas más valiosas que me llevo de la carrera. Le agradezco particularmente el que además de enseñarme matemáticas, me haya enseñado a disfrutar las matemáticas.

A la Dra. Nelia Charalambous, al Dr. César Luis García y a la Dra. Claudia Gómez por dedicar de su tiempo a leer esta tesis y por contribuir con sus valiosos comentarios, y además a esta última por toda la ayuda e interminables horas de plática que me ha brindado durante estos años.

A mi hermano y a mis padres, por todo el apoyo y el respaldo que me han brindado, y por todo lo que he aprendido de ellos.

A Elisa, que ha sido una compañera incondicional en todo este proceso, por toda la ayuda, amor y felicidad que me ha dado, y además por su contribución a este trabajo con su conocimiento sobre términos fisiológicos.

A mis amigos del Lancaster; Alex, Andrés, Andrea, Diego, Javier, Jesús, Jika, Jimena, Joan, Laura, Maria José, Mariana, Michelle, Paula, Santiago I., Santiago S. y Tatiana, por su amistad, apoyo e innumerables experiencias a lo largo de los años.

A los del dominó; Bernardo, Pablo, Raúl y Rodrigo, por tantos y tan divertidos sábados distrayéndome de esta tesis.

A Lorena y a Gerardo, sin los cuales mi años en la carrera no hubieran sido los mismos, por todo el tiempo que nos hemos divertido juntos, y a los demás del ITAM: Javier, Yuca, Juanjo, Luis, Teresa y Michelle.

Índice General

Introducción

El Teorema de Lax-Milgram ha sido, desde su formulación en 1954, una piedra angular en el campo del análisis funcional. Concebido por los matemáticos Peter D. Lax y Arthur N. Milgram [12] como una herramienta auxiliar a la teoría de existencia y unicidad de ecuaciones diferenciales parciales, este teorema tuvo una pronta repercusión tanto en este campo aplicado como en el campo teórico del análisis funcional.

Peter David Lax, nacido en una familia judía húngara que eventualmente emigraría a los Estados Unidos huyendo de la segunda guerra mundial, es un prolífico matemático que ha dedicado su vida a las ecuaciones diferenciales parciales desde el punto de vista teórico, aplicado y computacional. Entre sus múltiples contribuciones destacan los esquemas de Lax-Friedrichs y Lax-Wendorff, el Teorema de Equivalencia de Lax y los "Pares de Lax". Fuera de la academia, Lax ha colaborado en varias iniciativas de aplicación, llegando incluso a formar parte del proyecto Manhattan (1945-1946) en Los Álamos, en el grupo dirigido por John Von Neumann.

Arthur Norton Milgram, de nacionalidad estadounidense, hizo por su parte contribuciones en diversas áreas de las matemáticas, desde el análisis funcional hasta la teoría de Galois, pasando por la combinatoria, geometría diferencial y topología. Éste habría de morir tan sólo seis años después de la publicación del teorema que lleva su nombre, mientras que Lax ha continuado su trabajo en el Instituto Courant, eventualmente ganando el prestigioso premio Abel en 2005 por sus "innovadoras contribuciones a la teoría y la aplicación de ecuaciones diferenciales parciales y la obtención computacional de sus soluciones".

Desde el punto de vista del análisis funcional, el Teorema de Lax-Milgram tuvo una gran importancia no sólo netamente teórica, sino también histórica y temporal, en un momento en que las matemáticas se encontraban en un período de revolución y cambio profundo, con el nacimiento y desarrollo de esta rama del análisis. La idea de definir espacios de funciones, aunque trazaba sus antecedentes hasta finales del siglo XIX, no fue llevada acabo plenamente sino hasta entrado el siglo XX. Los espacios de Banach fueron introducidos hasta alrededor de 1920; los de Hilbert, introducidos cerca de 1910, pero axiomatizados y establecidos definitivamente hasta después de 1930; y los espacios de Sobolev en la década de los 50. En términos históricos de las matemáticas, estos espacios de funciones eran completamente nuevos, y los descubrimientos asociados a ellos se daban de manera vertiginosa.

Después de una primera etapa (primera mitad del siglo XX) en la que se definieron y axiomatizaron estos nuevos espacios, comenzaba una segunda etapa en la que se buscaba entenderlos profundamente y descubrir qué propiedades escondían. La torpeza con la que la intuición humana deambula en los espacios infinito-dimensionales hacía de esta búsqueda algo emocionante, incierto y lleno de sorpresas. Enclavado en este ámbito, el Teorema de Lax-Milgram vino a complementar otros teoremas de representación, formando uno de los bloques básicos sobre los que descansa el análisis funcional.

Desde el punto de vista de las ecuaciones diferenciales, la aparición del Teorema de Lax-Milgram se da en un contexto temporal en el que comenzaba el interés por formular problemas en espacios abstractos, que pudieran capturar sus propiedades y facilitaran su planteamiento y resolución. Las características intrínsecas que poseían espacios como los de Banach, Hilbert y Sobolev, todos ellos recientemente desarrollados, los hicieron candidatos idóneos para esto. Lax y Milgram formaron parte de este período de ebullición, contribuyendo con un resultado de existencia y unicidad que sustentaría una parte importante de esta teoría.

De esta manera, el Teorema de Lax-Milgram resultó ser de gran ayuda en la formulación de problemas con soluciones débiles, problemas variacionales y para verificar el buen planteamiento de problemas para aproximación con métodos de Galerkin. Sin embargo, el impacto mayor de dicho teorema no se dio por sus aplicaciones directas, sino por la teoría, métodos y generalizaciones desarrollados a partir de éste, que tuvieron -y siguen teniendo- una amplitud aún mucho mayor.

En este trabajo se busca explorar a detalle dicho teorema, analizando cuidadosamente tanto los antecedentes teóricos que lo sustentan, como las repercusiones que tiene. Se hará un recorrido por cada uno de los conceptos requeridos para la formulación del teorema, haciendo hincapié en la relación que tienen con éste.

A lo largo de este trabajo nos apoyaremos fundamentalmente en artículos que cubren las generalizaciones y las aplicaciones del teorema, aunque también, en menor medida, en libros y textos para sustentar la teoría básica. Los textos fundamentales para la realización de este trabajo fueron [2], [11],[12], [13] y [14]. La mayoría de las demostraciones aquí presentadas son adaptaciones realizadas a partir de las originales o de variantes de éstas. Para el Teorema de Lax-Milgram se presentan dos demostraciones: una adaptación de la variante de Evans y una original aquí propuesta, que sustituye la dependencia en el Teorema de Representación de Riesz, utilizando en su lugar una variante del Teorema de Hahn-Banach.

La distribución del trabajo se dará de la siguiente manera. El primer capítulo será dedicado a enunciar y demostrar el Teorema de Lax-Milgram, habiendo previamente presentado los conceptos teóricos requeridos para este propósito: principalmente los espacios de Banach, de Hilbert, formas bilineales y algunas propiedades de éstos. En el segundo capítulo presentaremos cuatro generalizaciones del Teorema que se desarrollaron en los años posteriores a la formulación de éste, siendo algunas de ellas usuales y ampliamente documentadas y otras menos conocidas. Demostraremos dichas generalizaciones con un enfoque similar a aquel del teorema original, con el fin de facilitar su comparación y dotar de coherencia al desarrollo teórico de este texto . Finalmente, en el tercer capítulo estudiaremos algunas aplicaciones prácticas del teorema original y de una de sus generalizaciones, habiendo expuesto previamente la teoría contextual de dichas aplicaciones. El siguiente diagrama ilustra un panorama general de las generalizaciones y aplicaciones presentadas, además de la sección en la que se encuentran.

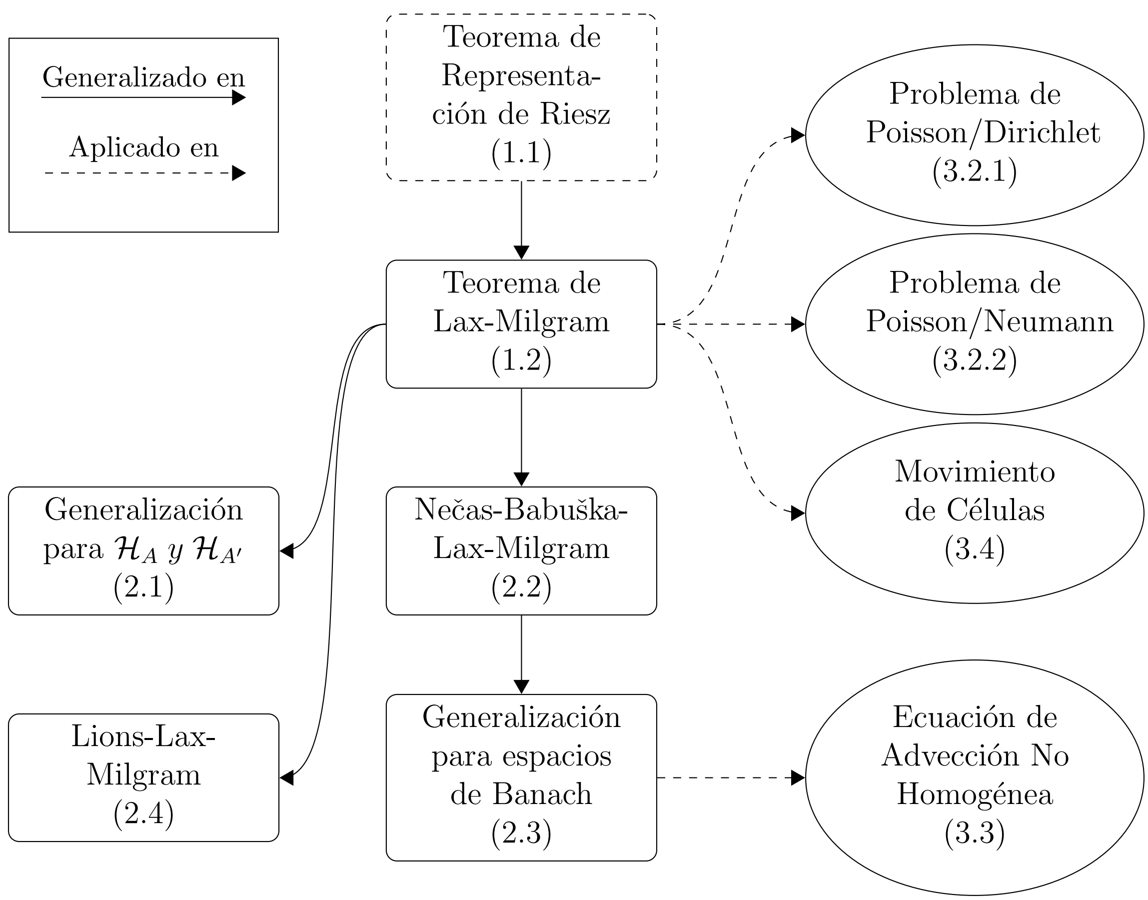

Figura 1: Dependencia Lógica entre Generalizaciones y Aplicaciones

Capítulo 1

El Teorema de Lax-Milgram

En este capítulo revisaremos los conceptos básicos necesarios para la formulación del teorema y posteriormente lo enunciaremos y demostraremos. Comenzaremos en la sección 1.1 con los conceptos mínimos necesarios para poder presentar a la brevedad posible los espacios de Banach y de Hilbert, y a partir de este punto recorreremos varias definiciones y proposiciones pensadas en el contexto de estos últimos espacios.

En la sección 1.2 expondremos y demostraremos el Teorema de Lax-Milgram, con una variante de la prueba usual. Veremos que en dicha demostración nos apoyaremos fuertemente en el Teorema de Representación de Riesz, además de que mostraremos que este último puede ser interpretado como un caso particular del primero.

En la sección 1.3 presentamos una manera distinta y original de demostrar el Teorema de Lax-Milgram, sin necesidad de utilizar el Teorema de Representación de Riesz, sino apoyándose directamente en el de Teorema de Hahn-Banach.

Finalizaremos este capítulo con una breve sección dedicada a analizar algunas implicaciones y consecuencias de este teorema, de carácter tanto teórico como práctico.

1.1. Definiciones y Conceptos

Comenzaremos por presentar algunas definiciones básicas del álgebra lineal y el análisis matemático, que sirven como sustento para la construcción de la teoría del análisis funcional. Supondremos cierta familiaridad del lector con los espacios vectoriales y las transformaciones lineales, aunque algunas propiedades de estos últimos son presentadas y demostradas en el Apéndice A. Los espacios vectoriales serán el punto de partida hacia los espacios que estudiaremos más adelante, obtenidos a partir de añadir progresivamente propiedades particulares y condiciones especiales a los primeros. Trataremos entonces con X un espacio vectorial dotado de un campo de escalares subyacente $\mathbb{K}$ (nos referiremos a X como un $\mathbb{K}$ - espacio vectorial). Supondremos que dicho campo será siempre ya sea $\mathbb{R}$ ó $\mathbb{C}$, a menos que especifiquemos el uso de alguno de los dos en particular.

Una de las relaciones más básicas que se busca establecer entre dos elementos de un espacio vectorial es la "distancia" entre ellos. Esta relación, aunque es ciertamente relativa, es una buena forma de entender la similitud entre los objetos en X. La noción de distancia nos será accesible a través de la definición de una métrica en el espacio.

Definición 1.1.1. Sea X un $\mathbb{K}$ - espacio vectorial[1], entonces $d : X \times X \to \mathbb{R}$ es una **métrica** si[2]:

(i) $d(x, y) = 0 \Leftrightarrow x = y$

(ii) $d(x, y) = d(y, x) \quad \forall x, y \in X$

(iii) $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$

Si d está definido para todo par $(x, y) \in X \times X$ entonces X es un **espacio métrico**. Se le denota en ese caso por la pareja (X, d).

Poder utilizar el concepto de distancia es fundamental, pero se muestra inútil cuando se busca conocer la naturaleza de un elemento $x \in X$ *per se*, y no en comparación a otro elemento. Para esto, se antoja necesario definir una propiedad que sea inherente a x y dependiente sólo de éste, a lo cual se piensa naturalmente en el "tamaño" o "magnitud" de un elemento. En términos formales, estamos interesados en definir la norma de un objeto.

Definición 1.1.2. Sea X un $\mathbb{K}$ - espacio vectorial, entonces $\|\cdot\| : X \to \mathbb{R}$ es una **norma** si:

(i) $\|x\| \geq 0 \quad y \quad \|x\| = 0 \Leftrightarrow x = 0$

[1]Formalmente, basta con que X sea un conjunto no vacío para poder definir una métrica en él. Dado que todos los espacios que utilizaremos tienen una estructura subyacente de espacio vectorial, definimos aquí la métrica para éstos.

[2]La propiedad $d \geq 0$ es consecuencia de (i), (ii) y (iii), pues $0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$ para cualquier pareja (x, y).

(ii) $\|\alpha x\| = |\alpha|\|x\| \quad \forall x \in X,\ \alpha \in \mathbb{K}$

(iii) $\|x+y\| \leq \|x\| + \|y\| \quad \forall x, y \in X$

Si $\|\cdot\|$ está definido para todo $x \in X$ entonces X es un **espacio normado**. Se le denota en ese caso por la pareja $(X, \|\cdot\|)$.

Definición 1.1.3. Se dice que $\|\cdot\|_1$ y $\|\cdot\|_2$ son **normas equivalentes** si existen constantes reales c_1 y c_2 tales que

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$$

Si sus normas son equivalentes, entonces los espacios $(X, \|\cdot\|_1)$ y $(X, \|\cdot\|_2)$ son topológicamente equivalentes.

Definición 1.1.4. Dos espacios normados $(X, \|\cdot\|_1)$ y $(Y, \|\cdot\|_2)$ son **isométricos e isomorfos** si existe un transformación $T : X \to Y$ biyectiva, es lineal y que además preserva normas, es decir:

$$\|T(v)\| = \|v\| \quad \forall v \in X$$

En ese caso, T es un **isomorfismo isométrico**, y decimos que $X \cong Y$.

Observación. Existe cierta ambigüedad en la literatura con respecto al término isomorfismo isométrico en el contexto de los espacios normados. Aunque gran parte de los autores ([18, 24], por ejemplo) definen este concepto como hicimos aquí y como se hace en otras áreas de las matemáticas, es común encontrar (en [20], por ejemplo) que en espacios normados se pide que T sea solamente inyectiva y no biyectiva. Esta ambigüedad cobrará particular importancia cuando más adelante definamos el concepto de reflexividad.

Es claro que a partir de una norma siempre se puede definir una métrica, pues basta con hacer $d(x,y) = \|x-y\|$ y se obtiene una métrica válida, que cumple con lo requerido por la definición (1.1.1):

(i) $d(x,y) = \|x-y\| = 0 \Leftrightarrow (x-y) = 0 \Leftrightarrow x = y$

(ii) $d(x,y) = \|x-y\| = \|(-1)(y-x)\| = |-1|\|y-x\| = \|y-x\| = d(y,x)$

(iii) $d(x,y) = \|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y)$

El converso de esta implicación no es cierto, es decir, a partir de una métrica no siempre es posible definir una norma. Por ejemplo, consideremos el espacio X de todas las sucesiones complejas $\{x_i\}$. La función

$$d(x,y)=\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|x_i-y_i|}{1+|x_i-y_i|}$$

es una métrica, pues las dos primeras propiedades de la definición (1.1.1) se cumplen trivialmente, y para la tercera tenemos que

$$\begin{aligned} d(x,y)=\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|x_i-y_i|}{1+|x_i-y_i|} &= \sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|x_i-z_i+z_i-y_i|}{1+|x_i-z_i+z_i-y_i|} \\ &\leq \sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|x_i-z_i|}{1+|x_i-z_i|}+\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|z_i-y_i|}{1+|z_i-y_i|}=d(x,z)+d(z,y) \end{aligned}$$

utilizando el hecho de que

$$\frac{|a+b|}{1+|a+b|}\leq\frac{|a|}{1+|a|}+\frac{|b|}{1+|b|}$$

para cualesquiera a y b números complejos. Sin embargo, X no puede ser un espacio normado, porque si existiese una norma $\|\cdot\|$ tal que $d(x,y)=\|x-y\|$ entonces

$$d(\alpha x,\alpha y)=\|\alpha x-\alpha y\|=|\alpha|\|x-y\|=|\alpha|d(x,y)$$

se debe satisfacer, pero para la métrica en cuestión esta relación no es valida pues

$$d(\alpha x,\alpha y)=\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|\alpha||x_i-y_i|}{1+|\alpha||x_i-y_i|}$$

y

$$|\alpha|d(x,y)=|\alpha|\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot\frac{|x_i-y_i|}{1+|x_i-y_i|}$$

por lo que dicha norma no puede existir.

Una tercera noción que buscaremos añadir a nuestro espacio vectorial X es alguna que relacione no sólo la distancia o magnitud de dos elementos del espacio, sino la conformación o posición relativa entre ellos. Estamos entrando ahora en el concepto de producto interno o interior, el cual nos permitirá definir nociones como la del "ángulo" y de la ortogonalidad entre dos elementos.

Definición 1.1.5. Sea X un $\mathbb{K}$ - espacio vectorial, entonces $(<,>) : X\times X\to\mathbb{K}$ es un **producto interno** si:

(i) $<x+x',y>=<x,y>+<x',y>$

(ii) $< \alpha x, y >= \alpha < x, y > \quad \forall x, y \in X \quad \forall \alpha \in K$

(iii) $< x, y >= \overline{< y, x >} \quad \forall x, y \in X$

(iv) $< x, x > \quad \geq 0 \quad \forall x \in X$ y $< x, x >= 0 \Leftrightarrow x = 0$

Si $(<,>)$ está definido para todo $(x, y) \in X \times X$ entonces X es un **espacio con producto interno**. Se le denota en ese caso por la pareja $\big(X, (<,>)\big)$.

Observación. Nótese que la tercera propiedad del producto interno es simplemente la simetría entre ambos argumentos del producto interno en el caso en que $\mathbb{K} = \mathbb{R}$.

Una de las propiedades más importantes en los espacios con producto interno es la muy conocida desigualdad de Cauchy-Bunyakovsky-Schwarz (CBS), que asegura que para x y y vectores arbitrarios se tiene

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

La relación entre el producto interno y la norma que otorga la desigualdad CBS es de crucial importancia en el análisis funcional, ya que, por ejemplo, nos ayuda a demostrar que todo espacio con producto interno es un espacio normado. Si a partir del producto interno se define:

$$\|x\| := \langle x, x\rangle^{\frac{1}{2}} \tag{1.1}$$

se tiene entonces que

(i) $\|x\| = \langle x, x\rangle^{\frac{1}{2}} \geq 0$

(i') $\|x\| = \langle x, x\rangle^{\frac{1}{2}} = 0 \Leftrightarrow \langle x, x\rangle = 0 \Leftrightarrow x = 0$

(ii) $\|\alpha x\| = \langle \alpha x, \alpha x\rangle^{\frac{1}{2}} = \left(|\alpha|^2\langle x, x\rangle\right)^{\frac{1}{2}} = |\alpha|\|x\|$

(iii) $\|x + y\| = \langle x + y, x + y\rangle^{\frac{1}{2}} \leq \|x\| + \|y\| \qquad$ (por la desigualdad CBS)

Es decir, $\langle x, x\rangle^{\frac{1}{2}}$ siempre es una norma en el espacio.

Las relaciones de inclusión entre los tipos de espacios definidos previamente se pueden visualizar fácilmente en el siguiente diagrama. En la parte superior se muestran la propiedades que se agregan a los espacios (cuando se pueda hacer de manera congruente) para dotarlos de una estructura intrínseca más robusta. Por el contrario, en la parte inferior se muestra cómo la estructura de los espacios más generales se encuentra presente implícitamente en los espacios mas complejos.

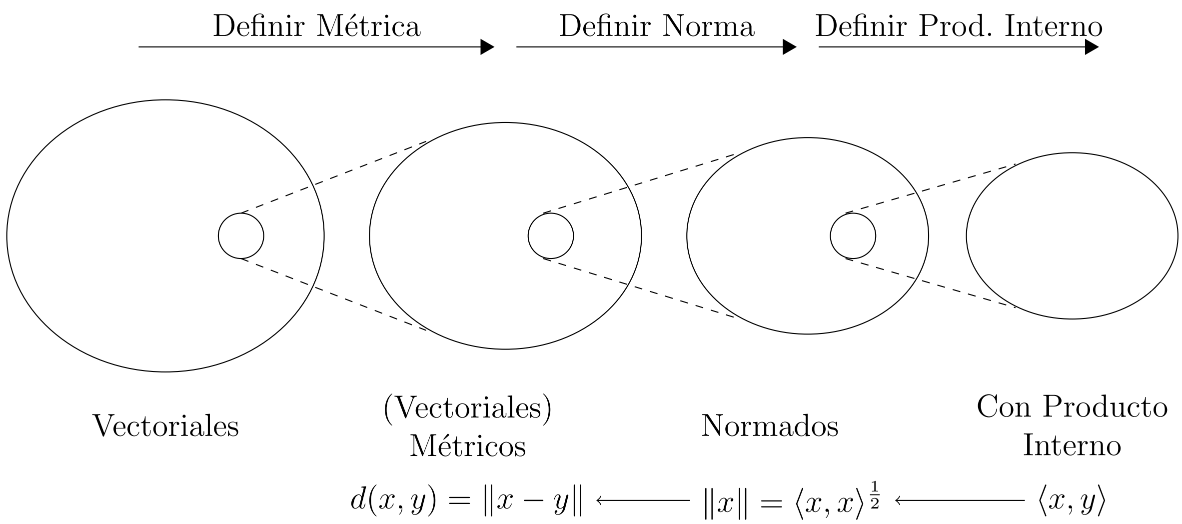

Figura 1.1: Inclusión entre familias de espacios

Como podemos ver, a partir de los espacios vectoriales hemos hecho progresivamente refinaciones a la estructura algebraica de los espacios, dotándolos de características que resultarán útiles (si no es que indispensables) para la teoría que se construirá más adelante. El producto interno, por ejemplo, nos permite definir la muy conocida noción de ortogonalidad. Esta propiedad, no obstante su aparente simplicidad, es de inmensa importancia, pues esencialmente determina la geometría de un espacio vectorial con producto interno.

Definición 1.1.6. Decimos que dos elementos x, y en un espacio vectorial con producto interior $\big(V, (<, >)\big)$ son **ortogonales** si se tiene que $\langle x, y\rangle = 0$. En ese caso escribimos $x \perp y$.

La ortogonalidad, que encierra una relación muy particular entre dos elementos de un espacio, se puede extender para incluir subconjuntos de vectores en el espacio.

Definición 1.1.7. Sea S un subconjunto no vacío y x un vector en un espacio X con producto interno. Decimos que x es ortogonal a S si es ortogonal a cada uno de los vectores en S, y lo denotamos por $x \perp S$. Además, denotamos por $S^{\perp}$ (llamado el **complemento ortogonal** de S) al conjunto de todos los vectores en X que son ortogonales a S.

Encaminaremos ahora nuestra atención hacia algunos conceptos básicos del análisis matemático, derivados de la existencia de la métrica, que nos permitirán finalmente completar la estructura de los espacios que serán motivo de nuestra atención a lo largo de este trabajo.

Definición 1.1.8. Sea (X, d) un espacio métrico, entonces para $\epsilon > 0$, la **ϵ-vecindad** de un punto x_0 en X es el conjunto $V_\epsilon(x_0) := \{x \in S \mid d(x_0, x) < \epsilon\}$.

Definición 1.1.9. Sea (X, d) un espacio métrico. Se dice que un subconjunto A de X es **abierto** en X si para cada punto $x \in A$ existe una vecindad V de x tal que $V \subseteq A$. Se dice que un subconjunto C de X es **cerrado** en X si el complemento $X \setminus C$ es un conjunto abierto en X.

Definición 1.1.10. Una sucesión (x_n) en un espacio métrico (X, d) es **convergente** si existe un elemento x en X para el cual $\lim\limits_{n\to\infty} d(x_n, x) = 0$. Se dice entonces que (x_n) converge a x y se escribe $x_n \to x$.

Observación. A la luz de esta definición, podemos dar una caracterización alternativa de los conjuntos cerrados. Un conjunto C es cerrado si y sólo si todas las sucesiones convergentes de elementos en C tienen su límite en C.

Definición 1.1.11. Una sucesión (x_n) en un espacio métrico (X, d) es **de Cauchy** si $\lim\limits_{n,m\to\infty} d(x_n, x_m) = 0$.

Es fácil demostrar que toda sucesión convergente es de Cauchy, pero el converso no siempre es cierto. De hecho, los espacios en los que esto sí se cumple tienen algunas propiedades interesantes y por ello reciben un nombre especial.

Definición 1.1.12. Un espacio métrico (X, d) es llamado **completo** si toda sucesión de Cauchy es convergente, y converge en el mismo espacio.

Intuitivamente, la noción de completez indica, como su nombre lo dice, que el espacio está "completo" en el sentido que no le "faltan puntos". En cierta forma, la completez es una propiedad similar a la cerradura, pues ambas aseguran que los puntos límite de sucesiones en un conjunto se encuentran dentro el conjunto mismo, esto es, que no tienen "agujeros". Sin embargo, difieren en el hecho de que la completez además asegura la convergencia de ciertas sucesiones. Es claro que la completez es una propiedad más fuerte que la cerradura; de hecho, es trivial demostrar que todo espacio métrico completo es en particular cerrado. Un teorema importante que relaciona estos dos conceptos es el siguiente.

Teorema 1.1.13. *Todo subconjunto cerrado de un espacio normado completo es también completo.*

Demostración. Sea X un espacio normado completo y $C \subset X$ cerrado. Sea además (x_n) una sucesión de Cauchy en C. Como X es completo, $(x_n) \to x \in X$. Entonces (x_n) es una sucesión en C que converge, por lo que $x \in C$ por ser éste cerrado. Concluimos así que C es completo. $\square$

Otra observación interesante es que la completez siempre se cumple para espacios vectoriales normados de dimensión finita (siempre y cuando estén definidos sobre un campo completo, sea $\mathbb{R}$ ó $\mathbb{C}$), cosa que no sucede con los de dimensión infinita. Es por esta razón que dicho concepto se vuelve particularmente interesante en el contexto de estos últimos espacios, y por consiguiente, en todo el análisis funcional.

La completez es una propiedad deseable por varias razones, siendo la principal de ellas el hecho que permite decidir sobre la convergencia de ciertas sucesiones sin conocer necesariamente su límite. Gran parte de la genialidad de Stephan Banach en este contexto consiste en ser el primero[3] en reconocer la importancia *per se* de la completez en los espacios que surgen usualmente en el análisis funcional. Así, Banach combinó la completez con las propiedades algebraicas de los espacios normados y definió, alrededor de 1922, una de las estructuras matemáticas más importantes desarrolladas en el siglo XX: los espacios que llevan su nombre.

Definición 1.1.14. Un espacio **de Banach** es un espacio vectorial normado y completo.

***Ejemplo* 1.** Uno de los espacios de Banach de dimensión infinita más comunes es el espacio de todas las funciones continuas en el intervalo unitario

$$C[0,1] = \big\{ f \mid f : [0,1] \to \mathbb{K},\ f \text{ continua} \big\}$$

dotado de la norma del supremo $\|f\| = \sup\big\{|f(x)| : x \in [0,1]\big\}$. Sin embargo, el mismo espacio dotado de la norma $\|f\| = \int_0^1 |f(x)|dx$ no es completo y por tanto no es de Banach. Este es un ejemplo claro donde vemos que la completez depende de la métrica utilizada (en este caso definida a partir de la norma) y no es inherente al espacio subyacente.

[3]Como veremos más adelante, Hilbert se apoyaba fuertemente en esta propiedad también, pero lo hacía porque el espacio con el que trabajaba, l_2, la poseía. Según Kirk y Khamsi [9, p.26], Banach fue "probablemente la primera persona en reconocer el papel verdaderamente fundamental que juega la completez". Como resultado de esto, la tomó como axioma al trabajar con lo que ahora conocemos como espacios de Banach.

La naturalidad con la que se definen funciones en los espacios de Banach, además de las propiedades que poseen, hicieron que se convirtieran pronto en los cimientos del análisis funcional, y parte central de su estudio. Se podría decir que el desarrollo de esta rama de las matemáticas no se podría entender de la misma manera sin los espacios de Banach.

Una primera propiedad interesante que caracteriza a algunos espacios de Banach y que tiende a ser bastante útil es aquella de la reflexividad. Sin embargo, para poder definir esta propiedad necesitaremos antes presentar el concepto de espacio dual. Recordemos que una funcional lineal es una transformación lineal de un espacio vectorial sobre su campo de escalares. Con esto en cuenta, damos a paso a la siguiente definición.

Definición 1.1.15. Sea X un espacio normado[4]. El **espacio dual** de X, denotado por X^*, es el espacio de las funcionales lineales continuas de X en $\mathbb{K}$, es decir $X^* = \{f | f : X \to \mathbb{K}, \text{f lineal y continua}\}$. Análogamente, X^{**} denota el espacio de funcionales lineales continuas en X^* y es llamado el espacio **doble dual** de X.

***Ejemplo* 2.** Consideremos c_0 el espacio de las sucesiones $x = (x_1, \dots, x_n, \dots)$ convergentes a cero. Con la norma $\|x\| = \sup_n |x_n|$, c_0 es un espacio de Banach. El dual de c_0 es l_1, el espacio de todas las sucesiones absolutamente sumables $f = (f_1, \dots, f_n, \dots)$ con la norma $\|f\| = \sum_{n=1}^{\infty} |f_n|$. Para ver esto, veamos que cualquier sucesión $f \in l_1$ define en el espacio c_0 una funcional lineal acotada f mediante la fórmula

$$\tilde{f}(x) = \sum_{n=1}^{\infty} f_n x_n \tag{1.2}$$

Podemos ver que $|\tilde{f}(x)| \leq \|x\| \sum_{n=1}^{\infty} |f_n x_n|$, por lo que $\|\tilde{f}\| \leq \|f\|$. Consideremos los elementos e_i en c_0, donde e_i es una sucesión de ceros, excepto en el i-ésimo término, cuyo valor es 1. Si definimos $x^{(N)} = \sum_{n=1}^{N} \frac{f_n}{|f_n|}$ (suponemos que $\frac{f_n}{|f_n|} = 0$ si $f_n = 0$), entonces $x^{(N)} \in c_0$, $\|x^{(N)}\| \leq 1$ y

$$\tilde{f}(x^{(N)}) = \sum_{n=1}^{N} \frac{f_n}{|f_n|} \tilde{f}(e_n) = \sum_{n=1}^{N} |f_n|$$

de manera que $\lim_{N\to\infty} \tilde{f}(x^{(N)}) = \sum_{n=1}^{\infty} |f_n| = \|f\|$. Por lo tanto $\|\tilde{f}\| \geq \|f\|$, y concluimos así que la transformación $f \mapsto \tilde{f}$ es isométrica entre l_1 y c_0. Falta simplemente ver que toda funcional en c_0 se puede representar en la forma (1.2). Para todo $x = (x_n) \in c_0$ tenemos que $x = \sum_{n=1}^{\infty} x_n e_n$. Esta serie converge al elemento x pues

[4]Aunque el dual se puede definir para cualquier espacio vectorial (en cuyo caso es llamado *dual algebraico*), normalmente nos interesa el *dual continuo* de un espacio (i.e. de funcionales continuas), por lo que se define para espacios normados (o métricos).

$\|x - \sum_{n=1}^{N} x_n e_n\| = \sup_{n>N} |x_n| \to 0$ si $N \to \infty$. Como $\tilde{f} \in c_0^*$ es continua, tenemos que $\tilde{f}(x) = \sum_{n=1}^{\infty} x_n \tilde{f}(e_n)$, por lo que basta comprobar que $|\sum_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$. Tomando $x^{(N)} = \sum_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} e_n$ y observando que $x^{(N)} \in c_0$, $\|x^{(N)}\| \leq 1$ tenemos

$$\sum_{n=1}^{N} \|\tilde{f}(e_n)\| = \sum_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} \tilde{f}(e_n) = \tilde{f}(x^{(N)}) \leq \|\tilde{f}\|$$

Finalmente, debido a la arbitrariedad de N, concluimos que $\sum_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$, es decir, la transformación entre c_0 y l_1 que hemos encontrado es un isomorfismo isométrico. En otras palabras, $(c_o)^* = l_1$.

El dual de un espacio es de gran importancia para entender la naturaleza del éste, pues revela características algebraicas que pueden no ser evidentes directamente en el espacio. Además, es utilizado para definir conceptos como medidas, distribuciones y tensores. Aquí nos será útil para definir la noción de reflexividad.

Definición 1.1.16. Sea X un espacio vectorial normado. La **inmersión canónica** asociada a X es la aplicación

$$J : X \to X^{**}$$

definida por la relación

$$(J(x))(f) := f(x) \quad x \in X,\ f \in X^*$$

Definición 1.1.17. Un espacio de Banach $\mathcal{B}$ es **reflexivo** si la inmersión canónica entre $\mathcal{B}$ y $\mathcal{B}^{**}$ es un isomorfismo isométrico (en el sentido de (1.1.4)).

Observación. Con los Teoremas de Hahn-Banach (Apéndice A), se puede ver que la inmersión canónica en un espacio de Banach es siempre lineal e isométrica. Además, la segunda de estas propiedades implica directamente la inyectividad, pues si $\|J(x)\| = \|x\|$ y además $J(x_1) = J(x_2)$ entonces $0 = \|J(x_1) - J(x_2)\| = \|x_1 - x_2\|$, por lo que $x_1 = x_2$. Entonces, el único requisito para que un espacio de Banach sea reflexivo es que J sea sobre en el doble dual, es decir $J(\mathcal{B}) = \mathcal{B}^{**}$.

La reflexividad es otra propiedad que los espacios de dimensión finita poseen invariablemente, pues el espacio, su dual y su doble dual todos tienen la misma dimensión lineal, por lo que el conocido Teorema de la Dimensión en espacios de dimensión finita asegura que la inmersión canónica es trivialmente biyectiva.

A continuación presentamos otro tipo de espacio fundamental en el análisis funcional, que surge a partir de sustituir el esqueleto del espacio de Banach (el espacio normado) por otro tipo de espacio menos general (espacio con producto interno), pero manteniendo la - crucial - propiedad de la completez.

Definición 1.1.18. Se dice que un espacio X dotado de un producto interno $(<,>)$ es de **Hilbert** si con la norma inducida por éste es un espacio de Banach.

***Ejemplo* 3.** Posiblemente el espacio de Hilbert de dimensión infinita más conocido es el espacio de sucesiones reales (o complejas) sumables al cuadrado, es decir

$$l_2 = \left\{ (x_n) \ : \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}$$

equipado con el producto interno $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$. Este espacio fue descubierto por David Hilbert alrededor de 1910 y es por lo tanto el precursor histórico del resto de los espacios de Hilbert. Dentro de la familia de espacios l_p (definidos de manera similar al anterior, cambiando 2 por $1 \leq p < \infty$) l_2 es el único cuya norma es inducida por un producto interno, y por lo tanto es el único que puede ser un espacio de Hilbert.

Junto con los de Banach, los espacios de Hilbert son probablemente la estructura espacial más importante desarrollada en el siglo XX. La posesión de un producto interno, la noción de ortogonalidad y la completez podrían llevarnos a considerar éstos últimos análogos de dimensión infinita a los espacios Euclidianos. De hecho, su importancia es tal que los espacios con producto interno son comúnmente llamados también espacios "pre-Hilbert", manifestando el hecho que se encuentran en una etapa anterior a aquella ideal, esto es, a la posesión de la completez. A pesar de que los espacios de Banach y de Hilbert comparten la propiedad de la completez, difieren en que los de Hilbert ostentan dicha propiedad indirectamente, pues al no estar dotados explícitamente de una norma, se utiliza aquella inducida por la operación con la que sí cuentan inherentemente; el producto interno.

La definición de espacio de Hilbert (1.1.18) requiere que a partir del producto interno se defina una norma, lo cual sabemos que es posible dado que todo espacio con producto interno es en particular normado, por la identidad (1.1) de la página 8. Por lo tanto, directamente a partir de su definición es claro que todos los espacios de Hilbert son de Banach. Sin embargo, el converso de esta propiedad no es cierto. Es decir, no todo espacio de Banach es de Hilbert, de hecho, lo son sólo aquellos en los que se cumple una conocida identidad algebraica.

Proposición 1.1.19. *Sea $\mathcal{B}$ un espacio de Banach real. Entonces $\mathcal{B}$ es un espacio de Hilbert si y sólo si para todo par de elementos $x, y \in \mathcal{B}$ se cumple*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad \textit{(identidad del paralelogramo)} \tag{1.3}$$

Demostración. Supongamos que $\mathcal{B}$ es un espacio de Hilbert. Entonces es un espacio con producto interno, y su norma es la inducida por éste, por lo cual:

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \tag{1.4}$$

similarmente, vemos que

$$\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2 - 2\langle x,y\rangle + \|y\|^2 \tag{1.5}$$

por lo tanto, sumando las ecuaciones (1.4) y (1.5) obtenemos

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\langle x,y\rangle - 2\langle x,y\rangle + 2\|y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Para demostrar el converso, es necesario demostrar que $\mathcal{B}$ es un espacio con producto interno. Para esto, consideramos la Identidad de Polarización (real):

$$\zeta(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \tag{1.6}$$

Afirmamos que ζ es un producto interno en $\mathcal{B}$. Para esto, mostraremos que se cumplen las cuatro propiedades enunciadas en su definición (1.1.5). Para empezar, vemos que

$$\zeta(x,x) = \frac{1}{4}\left(\|2x\|^2 - \|0\|^2\right) = \|x\|^2$$

por lo que claramente se cumple la propiedad (iv). Además, es fácil ver que

$$\zeta(x,y) = \zeta(y,x)$$

Por otra parte, veamos que

$$\zeta(x,z) + \zeta(z,y) = \frac{1}{4}\left(\|x+z\|^2 - \|x-z\|^2 + \|z+y\|^2 - \|z-y\|^2\right) \tag{1.7}$$

pero

$$\|x+z\|^2 = \left\|\left(\frac{x+y}{2} + z\right) + \left(\frac{x}{2} - \frac{y}{2}\right)\right\|^2$$
$$\|y+z\|^2 = \left\|\left(\frac{x+y}{2} + z\right) - \left(\frac{x}{2} - \frac{y}{2}\right)\right\|^2$$

utilizando (1.3) vemos que

$$\|x+z\|^2 + \|y+z\|^2 = 2\left(\left\|\frac{x+y}{2} + z\right\|^2 + \left\|\frac{x}{2} - \frac{y}{2}\right\|^2\right)$$

Análogamente

$$\|x-z\|^2 + \|y-z\|^2 = 2\left(\left\|\frac{x+y}{2} - z\right\|^2 + \left\|\frac{x}{2} - \frac{y}{2}\right\|^2\right)$$

Por lo tanto, retomando (1.7)

$$\begin{aligned}\zeta(x,z)+\zeta(z,y) &= \frac{1}{4}\left[2\left(\left\|\frac{x+y}{2}+z\right\|^2+\left\|\frac{x}{2}-\frac{y}{2}\right\|^2\right)-2\left(\left\|\frac{x+y}{2}-z\right\|^2+\left\|\frac{x}{2}-\frac{y}{2}\right\|^2\right)\right] \\ &= \frac{1}{2}\left(\left\|\frac{x+y}{2}+z\right\|^2-\left\|\frac{x+y}{2}-z\right\|^2\right)=2\,\zeta\left(\frac{x+y}{2},z\right) \qquad (1.8)\end{aligned}$$

Si tomamos $y=0$, obtenemos $\zeta(x,z)=2\zeta\left(\frac{x}{2},z\right)$, ya que $\zeta(0,z)=0$ por (1.6). Utilizando esta identidad, sustituyendo x por $x+y$, obtenemos $\zeta(x+y,z)=\zeta(x,z)+\zeta(y,z)$. Con un argumento inductivo, podemos ver entonces que

$$\zeta(\alpha x,y)=\alpha\zeta(x,y) \qquad (1.9)$$

se cumple para números racionales diádicos de la forma $\alpha=m/2^n$. Dado que $\|\alpha x+y\|$ y $\|\alpha x-y\|$ son continuos con respecto a α, entonces, por (1.6) vemos que $\zeta(\alpha x,y)$ también lo es. Finalmente, por la densidad de los racionales diádicos en los reales, (1.9) se cumple para toda $\alpha\in\mathbb{R}$. □

Obsérvese que la proposición anterior se demostró para el caso de espacios de Banach y Hilbert reales. La demostración para espacios de complejos es análoga, definiendo la identidad de polarización compleja

$$\zeta(x,y)=\frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\|x+iy\|^2-i\|x-iy\|^2\right)$$

Los cálculos requeridos en ese caso son bastante más abundantes, y fuera de los objetivos de este trabajo, pero cabe mencionar que la Proposición (1.1.19), cuando es formulada para espacios complejos, es conocida como el Teorema de Jordan-Von Neumann, demostrado en un artículo de 1935 [8]. Ya sea en el caso real o complejo, es interesante cómo una sola identidad tan simple como la del paralelogramo resume completamente la compatibilidad entre el producto interno y la norma en un espacio.

Es curioso el hecho de que David Hilbert mismo no definió los espacios que llevan su nombre de la manera que se hizo aquí (1.1.18). De hecho, la estructura particular de espacio de Hilbert la investigó únicamente en el espacio l_2 (mostrado en el ejemplo (3)), que abarcó la mayoría de su tiempo y trabajo, razón por la cual es conocido como el "espacio de Hilbert" original. El primero en definir dichos espacios de manera axiomática fue J. Von Neumann[5]. De hecho, fue él mismo quien utilizó por primera vez el término "espacio de Hilbert complejo" en un artículo de 1929 [23] para referirse a dichos espacios, práctica que en los años subsecuentes se volvió común.

[5]Estos y algunos otros detalles del desarrollo de los espacios de Hilbert se puede consultar en [24, pp. 91-92].

Otro dato interesante es que hasta antes de 1930 (incluyendo la axiomatización de Von Neumann) la definición de espacio de Hilbert incluía una condición adicional: la separabilidad del espacio.

Definición 1.1.20. Se dice que un espacio normado X es **separable** si contiene un subconjunto numerable que es denso en X.

En un principio, se pensó necesario suponer separabilidad para definir los espacios de Hilbert pues l_2 poseía esta propiedad. De esta manera, hasta antes de la década de los 30 un espacio de Hilbert (aún si no era llamado con este nombre) era por definición un espacio métrico completo y separable. Como consecuencia de esto, todos esos espacios (de dimensión infinita) eran isométricamente isomorfos a l_2, y por tanto bastaba con investigar éste, el espacio de Hilbert primordial. Riesz fue uno de los primeros en ir en contra de esta tendencia, pues probó la versión de su Teorema de Representación para espacios de Hilbert [16] sin utilizar dicha hipótesis de separabilidad.

Dejaremos de lado las propiedades de los espacios de Hilbert y de Banach para dirigir nuestra atención hacia las transformaciones lineales y formas bilineales en espacios normados, con las cuales completaremos la teoría necesaria para presentar el Teorema de Lax-Milgram. Para esto, revisaremos a continuación algunas definiciones a las cuales recurriremos habitualmente a lo largo de este trabajo. Seguiremos la convención de llamar *transformaciones* a las aplicaciones $T : X \to Y$ entre espacios X y Y distintos, y *operadores* a aquellas sobre el mismo espacio, es decir, de la forma $T : X \to X$. Además, supondremos siempre que las transformaciones y operadores son lineales, aunque no lo expresemos explícitamente.

Definición 1.1.21. Una transformación lineal T en un espacio normado está **acotada** si existe una constante positiva k tal que $\|Tx\| \leq k\|x\| \quad \forall x \in X$.

Definición 1.1.22. La transformación lineal T es **continua** en $D(T)$ si para todo $x_0 \in D(T)$ dado un $\epsilon > 0$, existe $\delta > 0$ tal que $\|Tx - Tx_0\| < \epsilon$ siempre que $\|x - x_0\| < \delta$.

Definición 1.1.23. La **norma** de una transformación lineal T está dada por

$$\|T\| = \sup \left\{ \frac{\|Tx\|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

Observaciones.
1. En el Teorema (A.5) del Apéndice A mostramos la equivalencia entre esta y otras normas posibles para transformaciones.

2. En el caso particular de una funcional, la norma se define análogamente, a saber

$$\|f\| = \sup \left\{ \frac{|f(x)|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

Definición 1.1.24. Para la transformación lineal $T : X \to Y$, definimos los conjuntos $R(T) = \{T(x) \in Y \,|\, x \in X\}$ y $N(T) = \{x \in X \,|\, T(x) = 0\}$, llamados el **rango** y **núcleo** de T, respectivamente.

El concepto del operador adjunto, de gran utilidad en el álgebra lineal y la teoría de operadores, requiere en su definición que el espacio X posea además un producto interno. Varios tipos de operadores que se basan en esta noción, junto con sus propiedades, se presentan en el Apéndice A.

Definición 1.1.25. Sea T un operador lineal sobre un espacio de con producto interno X. El **adjunto** del operador T, denotado T^*, se define por

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad \forall x, y \in X$$

Observación. Aunque el concepto de dualidad se suele definir para operadores, se puede generalizar fácilmente para transformaciones de un espacio con producto interno X a otro espacio con producto interno Y distintos, de la siguiente manera

$$\langle Tx, y\rangle_Y = \langle x, T^*y\rangle_X \quad \forall x \in X,\, \forall y \in Y$$

con $T : X \to Y$ y $T^* : Y \to X$.

En el contexto de las transformaciones lineales, los conceptos de continuidad y acotamiento son equivalentes, como lo muestra la siguiente proposición.

Proposición 1.1.26. *Sean X, Y espacios normados y T una transformación lineal de X en Y, entonces T es continua si y sólo si es acotada.*

Demostración. Supongamos que T es continua y que no está acotada. Esto implica que para cualquier $n \in \mathbb{N}$ existe $x_n \neq 0$ tal que $\|T(x_n)\| \geq n\|x_n\|$. Es decir

$$\frac{1}{n\|x_n\|}\|T(x_n)\| \geq 1$$

que por la linealidad de T es equivalente a

$$\left\| T\left(\frac{x_n}{n\|x_n\|}\right) \right\| \geq 1$$

Definimos entonces

$$y_n := \frac{x_n}{n\|x_n\|}$$

Es claro que $y_n \to 0$, pero $\|Ty_n\| \geq 1$ por lo que $Ty_n \nrightarrow 0$. Esto implica que T no es continua en 0, lo cual resulta en una contradicción.

Para probar ahora el converso, supongamos que T es acotada. Esto implica que existe k positiva tal que $\|Tx\| \leq k\|x\|$. Ahora, si $x_n \to x$, para x arbitraria en X, entonces

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq k\|x_n - x\| \to 0$$

por lo que $Tx_n \to Tx$ y concluimos que T es continua para todo $x \in X$. □

El siguiente concepto que definiremos será el de las formas sesquilineales y bilineales, que serán el objeto principal del Teorema de Lax-Milgram y sobre las cuales se aplicará la conclusión de éste. Enunciaremos primero el más general de estos casos: el de las formas sesquilineales. Estas formas son funciones de dos argumentos que son lineales en uno de ellos y antilineales en el otro, esto es, los escalares salen de ella conjugados.

Definición 1.1.27. Sea X un espacio normado sobre $\mathbb{C}$. Definimos $B(\cdot,\cdot) : X \times X \to \mathbb{K}$. Decimos que B es una forma **sesquilineal**[6] si $\forall\, x, y, u, v \in X$, $\forall\, \alpha, \beta \in \mathbb{K}$ se cumplen las dos siguientes condiciones:

1. $B(\alpha u + v, y) = \alpha B(u, y) + B(v, y)$
2. $B(x, \beta u + v) = \overline{\beta} B(x, u) + B(x, v)$

Las formas bilineales, en cambio, son lineales en ambos argumentos, esto es, el escalar β en la definición anterior aparece si conjugar. Cabe mencionar aquí que existe ambigüedad al respecto de estos dos términos, pues muchos autores se refieren a las formas sesquilineales como bilineales, posiblemente porque cuando el campo subyacente es un campo real, ambas son equivalentes. En este trabajo, sin embargo, nos acataremos a las definiciones formales y trataremos por lo general con formas bilineales, a menos que se especifique lo contrario.

La primera propiedad fundamental de las formas bilineales que presentaremos es la del acotamiento. Esta propiedad, que en términos generales tiende a ser siempre deseable, será de particular importancia en el Teorema de Lax-Milgram, pues determinará el tipo de formas bilineales sobre las cuales se puede utilizar dicho teorema.

[6] Del latín, *sēsqui-*, que significa “uno y medio”.

Definición 1.1.28. Sea $B(\cdot,\cdot) : X \times X \to \mathbb{K}$ forma bilineal. Entonces se dice que B es **acotada** si existe una constante positiva γ tal que:

$$|B(x, y)| \leq \gamma \|x\| \cdot \|y\|$$

para cualesquiera $x, y \in X$

A continuación presentaremos el concepto de coercitividad (fuerte) de una forma bilineal, que será crucial en el enunciado del teorema de Lax-Milgram. De hecho, veremos más adelante que la condición de coercitividad será la hipótesis más importante para obtener la conclusión de dicho teorema. Más aún, variaciones sutiles de este concepto darán paso a generalizaciones que analizaremos en el segundo capítulo. Por esta razón hay que prestar especial atención a su definición formal.

Definición 1.1.29. Sea X un espacio normado sobre $\mathbb{K}$ y $B(\cdot,\cdot) : X \times X \to \mathbb{R}$ forma bilineal. Se dice que B es **fuertemente**[7] **coercitiva** si existe una constante positiva δ tal que:

$$B(x, x) \geq \delta \|x\|^2$$

para todo $x \in X$.

Observación. Una propiedad de la forma bilineal más débil que la coercitividad fuerte es la positividad, que simplemente significa que $B(x, x) \geq 0 \quad \forall x \in X$.

***Ejemplo* 4.** Para una forma bilineal simétrica B sobre $\mathbb{R}^n$, la coercitividad fuerte es análoga a la noción de positividad definida para matrices. Si expresamos a B en términos de su forma cuadrática asociada, tenemos que

$$B(x, y) = x^T A y$$

donde A es una matriz cuadrada de tamaño n. Entonces B es fuertemente coercitiva si existe $\delta > 0$ tal que $x^T A x \geq \delta \|x\|^2$ para toda $x \in \mathbb{R}^n$. Se puede demostrar que esto si cumple si y sólo si la matriz A es positiva definida (i.e. todos sus valores propios son positivos). De hecho, la constante δ de coercitividad tiene un valor particular en este caso. Dado que la forma bilineal es simétrica, la matriz A también lo es, entonces, utilizando el cociente de Rayleigh vemos que

$$\min_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_1$$

[7] La propiedad que se define aquí es a veces denominada simplemente "coercitividad", cosa que evitaremos para distinguirla de la coercitividad débil, que se presentará más adelante. Otro nombre común con el que se le conoce es el de *X-elipticidad*, por su relación con las ecuaciones diferenciales elípticas.

donde λ_1 es el valor propio más chico de A. Entonces

$$B(x,x) = x^T A x \geq \lambda_1 \|x\|^2 \quad \forall x \in \mathbb{R}^n$$

Es decir, la coercitividad fuerte se cumple con constante $\delta = \lambda_1$.

Los dos últimos resultados que presentaremos en esta sección son probablemente las caracterizaciones más importantes que hay de los espacios de Hilbert. Ambos son resultados particulares de estos espacios, pues dependen de la existencia del producto interno y de la ortogonalidad. Además, serán las herramientas fundamentales que utilizaremos para demostrar el Teorema de Lax-Milgram en la próxima sección.

El primero de estos resultados, el Teorema de la Proyección, es importante porque además de garantizar la existencia de la proyección ortogonal en los espacios de Hilbert, muestra una caracterización del espacio a través de subespacios y sus complementos ortogonales. Dicha propiedad es común a todos los espacios con producto interno de dimensión finita, y la demostración en ese caso es bastante directa. Del caso para los espacios de Hilbert, un tanto más complicado, existen varias versiones, entre las cuales está la que presentamos a continuación.

Teorema 1.1.30. (de Proyección) *Si $\mathcal{H}$ es un espacio de Hilbert y M es un subespacio cerrado de $\mathcal{H}$ entonces*

$$\mathcal{H} = M \oplus M^{\perp}$$

Demostración. 1. Dado que $\langle \cdot, \cdot \rangle$ es lineal, vemos que $M^{\perp}$ es un subespacio de $\mathcal{H}$, y la desigualdad CBS implica que además es cerrado. Si $x \in M$ y $x \in M^{\perp}$ entonces $\langle x, x \rangle = 0$, lo que implica que $x = 0$. Por lo tanto, $M \cap M^{\perp} = \{0\}$.

Basta mostrar ahora que $M + M^{\perp} = \mathcal{H}$.

2. Sea $x \in \mathcal{H}$ arbitrario. Mostraremos que el conjunto $x - M = \{y \in \mathcal{H} \mid y = x - m, m \in M\}$ tiene un elemento de norma mínima. Sea $d = \text{ínf}\{\|u\| : u \in x - M\}$. Escojamos $u_n \in x - M$ tal que $\|u\| \to d$. Dado que $\frac{1}{2}(u_n + u_m) \in x - M$, $\|u_n + u_m\| \geq 4d^2$. Esto, combinado con la ley del paralelogramo, muestra que

$$2\|u_n\|^2 + 2\|u_m\|^2 - \|u_n - u_m\|^2 = \|u_n + u_m\|^2 \geq 4d^2$$

Pero $2\|u_n\|^2 + 2\|u_m\|^2 \to 4d^2$, por lo que $\|u_n - u_m\|^2 \to 0$, luego (u_n) es de Cauchy, y por la completez del espacio y la cerradura de $x - M$, converge a algún $u \in x - M$, con $\|u\| = d$. Por lo tanto, existe $x_1 \in M$ que minimiza $\|x - x_1\|$.

3. Sea $x_2 := x - x_1$, entonces $\|x_2\| \leq \|x_2 + y\|$ para todo $y \in M$. Podemos decir equivalentemente que

$$\|x_2\| \leq \|x_2 - \lambda y\| \quad \forall y \in M$$

donde $\lambda = \frac{\overline{\langle x_2, y\rangle}}{\|y\|^2}$. Elevando al cuadrado y desarrollando vemos que

$$\begin{aligned} \|x_2\|^2 \leq \|x_2 - \lambda y\|^2 &= \|x_2\|^2 + 2\Re(\lambda\langle x_2, y\rangle) + \lambda^2\|y\|^2 \\ &= \|x_2\|^2 - 2\frac{\langle x_2, y\rangle^2}{\|y\|^2} + \langle x_2, y\rangle^2\frac{\|y\|^2}{\|y\|^4} = \|x_2\|^2 - \frac{\langle x_2, y\rangle^2}{\|y\|^2} \end{aligned} \tag{1.10}$$

Es claro entonces que $\langle x_2, y\rangle$ debe ser 0, por lo que $x_2 \in M^\perp$. Hemos demostrado entonces que $x = x_1 + x_2$ con $x_1 \in M$ y $x_2 \in M^\perp$, es decir, $\mathcal{H} = M + M^\perp$. □

Como consecuencia directa del teorema anterior, si M es un subespacio cerrado de $\mathcal{H}$, y es distinto de $\mathcal{H}$, entonces existe un elemento $x \in \mathcal{H}$ tal que $x \perp M$. Este sencillo razonamiento será de vital importancia tanto para la demostración del Teorema de Lax-Milgram, como para casi todas sus generalizaciones.

El siguiente teorema, además de ser esencial en los espacios de Hilbert, es probablemente uno de los más importantes en todo el análisis funcional. Conocido como el Teorema de Representación de Riesz, este resultado en realidad pertenece a un grupo de varios teoremas fundamentales conocidos con este nombre. La primera versión de este teorema, en la que se prueba que todo funcional lineal en $L^2[0,1]$ es representable por integración, fue descubierto en 1907 paralelamente por los matemáticos Frigyes Riesz [15] y Maurice Fréchet [7], húngaro y francés, respectivamente. La versión general para espacios de Hilbert, que será la que mostraremos a continuación, se encuentra en un artículo posterior de Riesz [16].

El Teorema de Representación de Riesz jugó un rol fundamental en los albores del análisis funcional, siendo uno de los primeros resultados básicos y fuertes de éste. En ese momento, ayudó a formalizar los cimientos de esta rama de las matemáticas y a entender mejor los espacios de funciones y la relación con sus duales. Aún hoy en día es uno de los resultados más utilizados en el análisis funcional.

Sin más preámbulo, mostramos a continuación dicho teorema.

Teorema 1.1.31. (Representación de Riesz) *Sea $\mathcal{H}$ un espacio de Hilbert y $f : \mathcal{H} \to \mathbb{K}$ una funcional lineal continua, entonces existe un único $y \in \mathcal{H}$ tal que $f(x) = \langle x, y\rangle \quad \forall x \in \mathcal{H}$ y además $\|f\| = \|y\|$.*

Demostración. 1. Si $f = 0$ entonces tomando $y = 0$ tenemos que

$$f(x) = \langle x, 0 \rangle = 0 \quad \forall x \in \mathcal{H}$$

Si por el contrario $f \neq 0$, entonces $N(f) \subset \mathcal{H}$. Por la continuidad de f, $N(f)$ es un subespacio cerrado de $\mathcal{H}$, por lo que el Teorema de la Proyección (1.1.30) implica que existe $u \neq 0$ con $u \in \mathcal{H} \setminus N(f)$ tal que $u \perp N(f)$. Es claro que se puede tomar en particular u con $\|u\| = 1$. Notemos que para toda $x \in \mathcal{H}$ se tiene

$$f\left(f(x)u - f(u)x\right) = f(x)f(u) - f(u)f(x) = 0$$

por lo que $f(x)u - f(u)x \in N(f)$. Además, como $u \in N(f)^{\perp}$ entonces

$$0 = \langle f(x)u - f(u)x, u \rangle = f(x)\|u\|^2 - \langle x, \overline{f(u)}u \rangle$$

es decir

$$f(x) = \langle x, \overline{f(u)}u \rangle = \langle x, y \rangle \quad \forall x \in \mathcal{H}$$

tomando $y = \overline{f(u)}u$.

2. Mostraremos a continuación que dicho y es único. Si existiesen $y_1, y_2 \in \mathcal{H}$ tales que $f(x) = \langle x, y_1 \rangle = \langle x, y_2 \rangle$ entonces

$$\langle x, y_1 - y_2 \rangle = 0 \quad \forall x \in \mathcal{H} \quad \Longrightarrow \quad y_1 - y_2 = 0 \quad \Longrightarrow \quad y_1 = y_2$$

Es decir, y es único.

3. Para probar la segunda afirmación del teorema vemos que

$$f(x) = \langle x, y \rangle \quad \Rightarrow |f(x)| = |\langle x, y \rangle| \leq \|x\|\|y\|$$

por lo que

$$\|f\| = \sup_{\|x\|=1} |f(x)| \leq \|y\|$$

Si $\|y\| = 0$, entonces $y = 0$ y $f = 0$, por lo que trivialmente $\|f\| = \|y\|$. Si $\|y\| \neq 0$,

$$\|f\| = \sup_{\|x\|=1} |f(x)| \geq \left| f\left(\frac{y}{\|y\|}\right) \right| = \left| \left\langle \frac{y}{\|y\|}, y \right\rangle \right| = \frac{1}{\|y\|} |\langle y, y \rangle| = \frac{\|y\|^2}{\|y\|} = \|y\|$$

siendo la desigualdad cierta pues $\|\frac{y}{\|y\|}\| = 1$. Es decir, hemos visto que en cualquier caso $\|f\| = \|y\|$ □

Observación. Si $\mathbb{K} = \mathbb{R}$, es claro que la conclusión del teorema anterior se puede obtener invirtiendo el papel que juegan los elementos de la primera y segunda entrada del producto interno. Sin embargo, cuando $\mathbb{K} = \mathbb{C}$ la otra versión posible del teorema es

$$\exists\, x \in \mathcal{H} \text{ única, tal que} \quad f(y) = \overline{\langle x, y \rangle} \quad \forall y \in \mathcal{H}$$

resultado que se obtiene directamente al conjugar ambos lados de la igualdad en la conclusión original del teorema.

Como podemos ver, detrás del enunciado -aparentemente simple- del Teorema de Riesz se esconde una demostración un tanto más delicada, que se apoya principalmente en el Teorema de Proyección (1.1.30) en espacios de Hilbert.

Para hacer una lectura un poco más profunda del teorema, podemos pensar en el producto interno como una función lineal de una variable con el espacio dual de $\mathcal{H}$ como contradominio, es decir $\langle \cdot, \rangle : \mathcal{H} \to \mathcal{H}^*$, que es de la forma

$$M_{\langle,\rangle}(y) = \langle \cdot, y \rangle \in \mathcal{H}^* \tag{1.11}$$

Bajo este planteamiento se hace evidente que el teorema anterior asegura la existencia de la inversa de dicha función. Es decir, si tenemos una funcional lineal continua $f \in \mathcal{H}^*$, el teorema asegura que existe un elemento $y \in \mathcal{H}$ cuya imagen bajo $M_{\langle,\rangle}$ es precisamente esa funcional. Más aún, el teorema asegura que dicho elemento es único, lo cual indica que la inversa está bien definida.

Además, el teorema en su enunciado nos otorga otra propiedad de la función $M_{\langle,\rangle}$ que puede pasar desapercibida, pero es de una importancia crucial. Nos referimos al hecho de que $\|f\| = \|y\|$. Esta equivalencia implica que la solución y que otorga el teorema depende continuamente de la función f que se elige, o equivalentemente, que la inversa de la función $M_{\langle,\rangle}$ es una continua.

Proposición 1.1.32. *La función $M_{\langle,\rangle}^{-1}$ es continua.*

Demostración. Sea $M_{\langle,\rangle}^{-1} : \mathcal{H}^* \to \mathcal{H}$ la inversa de la función definida en (1.11) y $f \in \mathcal{H}^*$ funcional lineal y arbitraria. Sea $y = M_{\langle,\rangle}^{-1}(f)$ según asegura el teorema (1.1.31). Por ese mismo teorema se tiene que

$$\|M_{\langle,\rangle}^{-1}(f)\| = \|y\| = \|f\|$$

por lo que $M_{\langle,\rangle}^{-1}$ es acotada, o equivalentemente, continua. □

Tomando en cuenta además esta propiedad, podemos ver que la inversa encontrada, siendo bien definida y continua, es prácticamente ideal. En pocas palabras, el Teorema de Riesz asegura que el producto interno, como función de una variable, es bicontinuo (continuo, biyectivo y con inversa continua). En otros términos, la función $M_{\langle,\rangle}$ es un homeomorfismo. Más aún, la proposición anterior muestra además que dicha función es una isomorfismo isométrico.

El Teorema de Representación de Riesz oculta aún un resultado más: permite demostrar que todos los espacios de Hilbert son reflexivos, pues otorga una identificación entre $\mathcal{H}$, $\mathcal{H}^*$ y $\mathcal{H}^{**}$. Presentamos a continuación el teorema que prueba este resultado, omitiendo algunos pasos sencillos en la demostración.

Teorema 1.1.33. *Sea $\mathcal{H}$ un espacio de Hilbert, entonces $\mathcal{H}$ es reflexivo.*

Demostración. Sea $J : \mathcal{H} \to \mathcal{H}^{**}$ la inmersión canónica. A la luz de la observación de la definición de reflexividad (1.1.17), sabemos que basta con demostrar que $R(J) = \mathcal{H}^{**}$. Esto equivale a mostrar que para todo elemento $F \in \mathcal{H}^{**}$ existe $x \in \mathcal{H}$ tal que $J(x) = F$. Adoptaremos la notación $J(x) = J_x$ por simplicidad.

Sea entonces $F \in \mathcal{H}^{**}$ arbitraria. Si y es un elemento fijo de $\mathcal{H}$, sea $\psi : \mathcal{H} \to \mathcal{H}^*$ dada por $\psi(y) = f_y$, donde f_y es el funcional asociado a y por el Teorema de Representación de Riesz. Es fácil ver que ψ es biyectiva e isométrica. Si el campo subyacente es $\mathbb{R}$, es además lineal, pero si el campo es $\mathbb{C}$, es anti-lineal (es decir, saca escalares conjugados). Por otra parte, definamos $g \in \mathcal{H}^*$ la funcional dada por

$$g(x) = \overline{F(\psi(x))} \quad x \in \mathcal{H},\ F \in \mathcal{H}^{**}$$

Utilizando las propiedades de ψ y el hecho que $F \in \mathcal{H}^{**}$ vemos que

$$\begin{aligned} g(\alpha x_1 + x_2) = \overline{F(\psi(\alpha x_1 + x_2))} &= \overline{F(\psi(\alpha x_1))} + \overline{F(\psi(x_2))} \\ &= \overline{\bar{\alpha} F(\psi(x_1))} + \overline{F(\psi(x_2))} = \alpha F(\psi(x_1)) + F(\psi(x_2)) \end{aligned}$$

y por otra parte

$$|g(x)| = |\overline{F(\psi(x))}| = |F(\psi(x))| \leq \|F\|\|\psi(x)\| = \|F\|\|x\|$$

Entonces g es una funcional lineal acotada en $\mathcal{H}$. Una vez más, por el Teorema de Representación de Riesz, existe una única $z \in \mathcal{H}$ tal que

$$g(x) = \langle x, z \rangle \quad \forall x \in \mathcal{H}$$

es decir

$$F(\psi(x)) = \overline{g(x)} = \overline{\langle x, z \rangle} = \langle z, x \rangle \tag{1.12}$$

Finalmente, de la definición de la inmersión canónica vemos que

$$J_z(\psi(x)) = \psi(x)(z) = f_x(z) = \langle z, x \rangle \tag{1.13}$$

De las ecuaciones (1.12) y (1.13), tenemos que $J_z = F$, que es lo que queríamos demostrar. □

Habiendo alcanzado el Teorema de representación de Riesz para tenerlo a nuestra disposición, tenemos finalmente las herramientas necesarias para dar paso al resultado fundamental de este trabajo: el Teorema de Lax-Milgram.

1.2. El Teorema

El Teorema de Lax-Milgram fue presentado por primera vez en 1954, en una recopilación de artículos publicada por la Universidad de Princeton con un objetivo en común: contribuir con elementos teóricos a la teoría de ecuaciones diferenciales parciales, que se encontraba en pleno apogeo en ese momento. Denominado originalmente “lema”, y presentado casi como uno técnico por sus autores, la importancia *per se* de este resultado haría que con el tiempo se le denominara cada vez menos de esa manera y cada vez más como “teorema”.

Como hemos mencionado anteriormente, el Teorema de Lax-Milgram es uno de representación. Este tipo de teoremas, a grandes rasgos, aseguran que existe una identificación entre una estructura matemática y otra estructura posiblemente más conocida o fácil de manejar, y que ciertas propiedades esenciales se mantienen en la transición entre ellos. El nombre proviene de la idea de que un elemento en la segunda estructura “representa” a aquel con el cuál está identificado en la primera. En el contexto del análisis funcional, posiblemente el teorema de representación más conocido es el de Riesz (1.1.31) (y algunas de sus variantes). Por su parte, el Teorema de Lax-Milgram asegura que las formas bilineales sobre un espacio de Hilbert son representables por elementos del dual de dicho espacio.

La demostración del teorema que se presentará aquí es una adaptación de la existente en [6], que posee un enfoque mucho más intuitivo que las demostraciones usuales que utilizan un operador contractivo y el Teorema del Punto Fijo de Banach (véase por ejemplo [20, 24, 18]). Se eligió basarse en esa demostración pues muestra claramente la influencia del Teorema de Representación de Riesz, además de que al requerir menor manipulación algebraica, permite percibir más directamente las ideas fundamentales detrás de la prueba. La demostración fue modificada para homogeneizarla a aquellas de las generalizaciones que presentaremos más adelante, y así hacer los paralelismos entre ellas evidentes.

Antes de dar paso al teorema, comenzaremos por probar un par de lemas que serán de gran ayuda en la demostración siguiente y en varias de las generalizaciones. Éstos se basan en el concepto de acotamiento por debajo de una transformación, estrechamente ligado a la coercitividad de una forma bilineal.

Definición 1.2.1. Una transformación T en un espacio normado X es **acotada por debajo** si existe una constante $\beta > 0$ tal que $\beta\|x\| \leq \|Tx\| \quad \forall x \in X$.

Aunque el acotamiento por debajo no asegura la continuidad de la transformación, sí nos permite hacer algunas otras conclusiones interesantes sobre ésta, como muestran los dos lemas siguientes. No obstante utilizaremos estos dos resultados para espacios de Hilbert en el Teorema de Lax-Milgram, los formulamos para espacios más generales para poder utilizarlos también en las generalizaciones subsecuentes.

Lema 1.2.2. *Sean X, Y espacios normados. Entonces una transformación lineal T de $D(T) \subseteq X$ a Y admite una inversa continua T^{-1} si y sólo si es acotada por debajo.*

Demostración. Supongamos que T es acotada por debajo, entonces

$$\beta\|x\| \leq \|Tx\| \quad \forall x \in X \tag{1.14}$$

A partir de esto vemos que

$$x \in N(T) \Leftrightarrow Tx = 0 \Leftrightarrow \|Tx\| = 0 \Leftrightarrow \|x\| = 0 \Leftrightarrow x = 0$$

por lo que $N(T) = \{0\}$ y T es inyectiva. Por lo tanto, $T^{-1} : R(T) \to X$ existe. Además, (1.14) es equivalente a

$$\|T^{-1}y\| \leq \frac{1}{\beta}\|y\| \quad \forall y \in Y$$

es decir, T^{-1} es acotada, y por lo tanto continua. $\square$

Directamente relacionado a este resultado se puede también obtener el siguiente.

Lema 1.2.3. *Sea $\mathcal{B}$ un espacio de Banach y $T : \mathcal{B} \to \mathcal{B}$ un operador lineal y continuo. Si T es acotado por debajo, entonces el rango de T es cerrado en $\mathcal{B}$.*

Demostración. Supongamos que $\{z_n\}$ es una sucesión en $R(T)$, el rango de T, y que $z_n \to z \in \mathcal{B}$. Al ser T acotado por debajo, el lema anterior implica que debe ser inyectivo. Así, sabemos que

$$\exists\,!\{v_n\} \subset X \quad \text{tal que} \quad z_n = Tv_n \quad \text{para cada} \quad n \in \mathbb{N}$$

Entonces vemos que

$$\|z_n - z_m\| = \|Tv_n - Tv_m\| = \|T(v_n - v_m)\| \geq \beta\|v_n - v_m\|$$

Pero $\{z_n\}$ en particular es de Cauchy, por lo que

$$\lim_{n,m\to\infty} \|z_n - z_m\| = 0$$

luego, por la desigualdad anterior, $\{v_n\}$ también es de Cauchy. Más aún, como $\mathcal{B}$ es un espacio de Hilbert, $\{v_n\}$ es convergente con la norma usual a un elemento, digamos v, en $\mathcal{B}$. Para terminar utilizamos la continuidad de T para ver que

$$z = \lim_{n\to\infty} z_n = \lim_{n\to\infty} Tv_n = T(\lim_{n\to\infty} v_n) = Tv$$

Es decir $z = Tv \in R(T)$. Hemos probado entonces que $R(T)$ es cerrado. $\square$

Con los conceptos y resultados de la sección anterior más estos dos lemas adicionales, nos encontramos finalmente listos para definir y demostrar el teorema principal de este trabajo. Como veremos a continuación, esta demostración se basa en definir un operador T asociado a la forma bilineal B a través del producto interno. Demostrando que dicho operador es invertible, se le combina con el Teorema de Representación de Riesz para llegar a la conclusión.

Teorema 1.2.4. (Lax-Milgram)(TLM) *Sea $\mathcal{H}$ un espacio de Hilbert y $B : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ forma bilineal, acotada y fuertemente coercitiva. Si $F : \mathcal{H} \to \mathbb{R}$ es una funcional lineal acotada, entonces existe un único $y_0 \in \mathcal{H}$ tal que*[8]

$$B(x, y_0) = F(x) \qquad \forall x \in \mathcal{H} \tag{1.15}$$

Demostración. 1. Fijando el elemento $y \in \mathcal{H}$, la aplicación $B_y = B(x, y) : \mathcal{H} \to \mathbb{R}$ es una funcional lineal sobre $\mathcal{H}$ para x. Como B es trivialmente continua para cada uno de sus argumentos, entonces B_y lo es también, y por tanto por el Teorema de Representación de Riesz (1.1.31) existe un único $z \in \mathcal{H}$ que satisface

$$B(x, y) = B_y(x) = \langle x, z \rangle \quad \forall\, x \in \mathcal{H} \tag{1.16}$$

Como se puede repetir este proceso para cada $y \in \mathcal{H}$, podemos definir el operador T de $\mathcal{H}$ en $\mathcal{H}$ dado por $Ty = z$ tal que $\langle x, z \rangle = B(x, y)$ para todo $x \in \mathcal{H}$.

2. Es claro que T es un operador lineal, pues B y el producto interno lo son. Además, es acotado pues como

$$\|Ty\|^2 = \langle Ty, Ty \rangle = B(Ty, y) \leq \gamma \|Ty\| \|y\|$$

por ser B acotada, entonces

$$\|Ty\| \leq \gamma \|y\| \quad \forall y \in \mathcal{H}$$

Utilizando el hecho de que B es coercitiva y la definición de T, existe $\delta \in \mathbb{R}^+$ tal que

$$\delta \|y\|^2 \leq B(y, y) = \langle y, Ty \rangle \leq \|y\| \|Ty\|$$

por lo que

$$\delta \|y\| \leq \|Ty\| \tag{1.17}$$

es decir, T es acotado por debajo. Esto implica que T^{-1} existe y es continuo, por el Lema (1.2.2), y además $R(T)$ es cerrado por el Lema (1.2.3).

[8]En el enunciado original del teorema, la consecuencia presentada es $F(x) = B(x, y_0) = B(y_1, x)$ para dos elementos únicos y_0,y_1. Con esto Lax y Milgram muestran explícitamente que el teorema puede ser aplicado simétricamente sobre las dos entradas de la forma bilineal.

Demostraremos ahora que $R(T) = \mathcal{H}$. Supongamos lo contrario, esto es, $R(T) \subset \mathcal{H}$. Entonces en virtud del Teorema de Proyección (1.1.30) (que podemos utilizar dado que $R(T)$ es cerrado) existe $w \in \mathcal{H}$ con $w \neq 0$ tal que $w \in R(T)^{\perp}$. Pero esto implica que

$$\delta \|w\|^2 \leq B(w, w) = \langle w, Tw \rangle = 0$$

luego $w = 0$, que implica una contradicción. Vemos entonces que $R(T) = \mathcal{H}$. Combinando esto con el hecho que T es invertible, podemos concluir que la ecuación $Ty = z$ tiene una solución única para cada $z \in \mathcal{H}$.

3. Por otra parte, utilizamos una vez más el Teorema (1.1.31) para ver que

$$F(x) = \langle x, z_F \rangle \quad \forall x \in \mathcal{H}$$

para un único $z_F \in \mathcal{H}$. Por lo visto en el punto anterior, podemos encontrar un único $y_0 \in \mathcal{H}$ que satisface $Ty_0 = z_F$. Entonces retomando (1.16) vemos que

$$B(x, y_0) = \langle x, Ty_0 \rangle = \langle x, z_F \rangle = F(x) \quad \forall x \in \mathcal{H}$$

que es precisamente lo que se buscaba demostrar. $\square$

A la luz de la fuerte dependencia de la demostración anterior en el Teorema de Representación de Riesz, se hace evidente que el Teorema de Lax-Milgram puede ser interpretado como una generalización (en los reales) de este último, para incluir funciones bilineales más generales que el producto interno. En otras palabras, basta con tomar la forma bilineal requerida por el teorema anterior como el producto interno en el espacio $\mathcal{H}$, es decir $B(x, y) = \langle x, y \rangle$, para obtener el Teorema de Riesz como caso particular. Es evidente que el producto interno cumple las hipótesis de acotamiento y coercitividad, pues

$$|\langle x, x \rangle| = \|x\|^2 \qquad \text{y} \qquad |\langle x, y \rangle| \leq \|x\| \|y\|$$

por lo que ambas condiciones se cumplen con constantes respectivas iguales a 1.

En este punto nos detenemos, al igual que lo hicimos para el de Riesz, para hacer un análisis más profundo del significado del teorema recién demostrado. Recordando la función $M_{\langle , \rangle}$ definida en (1.11), podemos ahora pensar en cualquier forma bilineal como una función lineal de una variable $B(, \cdot) : \mathcal{H} \to \mathcal{H}^*$. Entonces, podemos ahora definir para cada $y \in \mathcal{H}$ la función

$$M_B(y) = B(\cdot, y) \in \mathcal{H}^* \tag{1.18}$$

De esta forma, el Teorema de Lax-Milgram asegura la existencia de la inversa de dicha función, siempre y cuando cumpla condiciones de acotamiento y coercitividad. Es decir, si tenemos una funcional lineal continua $F \in \mathcal{H}^*$, el teorema asegura que existe un elemento $y_0 \in \mathcal{H}$ cuya imagen bajo M_B es precisamente esa funcional. Más aún, el teorema asegura que dicho elemento es único, lo cual indica que la inversa está bien definida. Nos gustaría que, como sucedió para el Teorema de Representación de Riesz, la función M_B^{-1} fuese también continua. Para investigar dicha propiedad, necesitaremos un corolario más.

El siguiente corolario es un resultado directo que en ocasiones se suele incluir como un resultado adicional en el enunciado del Teorema de Lax-Milgram. Sin embargo, éste tiene una importancia particular por sí solo -sin mencionar que en el artículo original de Lax y Milgram no está incluido dentro de la formulación del teorema- y por esta razón lo hemos tratado separadamente como un corolario en este trabajo. Esto permite explicar su importancia de manera independiente, además de no confundirlo con el resultado principal del teorema.

Corolario 1.2.5. *Con la notación del Teorema* (1.2.4)*, sea y_0 el elemento asociado a la funcional F. Entonces se tiene que $\|y_0\|_{\mathcal{H}} \leq \frac{1}{\delta}\|F\|_{\mathcal{H}^*}$, donde δ es la constante de coercitividad de la forma bilineal B.*

Demostración. Sea $z = Ty$. Según vimos en la demostración del teorema, el Lema (1.2.2) asegura que T^{-1} es acotada, es decir, continua. Más aún, ese lema muestra que la cota del operador inverso es el inverso de la cota inferior del operador, que en este caso es δ por (1.17). Entonces

$$\|y\| = \|T^{-1}z\| \leq \frac{1}{\delta}\|z\|$$

Sabemos que z proviene de aplicar el Teorema de Representación de Riesz al funcional F, por lo que, también según ese teorema, se tiene que $\|z\| = \|F\|$. Es decir

$$\|y\| \leq \frac{1}{\delta}\|z\| = \frac{1}{\delta}\|F\|$$

que es lo que se buscaba demostrar. □

Esta propiedad es de una importancia enorme, pues asegura que el elemento y_0 que se obtiene como consecuencia del Teorema de Lax-Milgram está acotado de antemano a través de la funcional lineal F que se escoge. El conocimiento de esta cota *a priori* para el elemento y_0 será fundamental en las aplicaciones del teorema, como veremos más adelante. En otros términos, el corolario anterior indica que y_0 depende continuamente de F, ó -retomando la interpretación dada anteriormente de M_B - que la función M_B^{-1} es además continua. Agrupando todas estas propiedades, concluimos que la función M_B es un homeomorfismo entre $\mathcal{H}$ y $\mathcal{H}^*$ (en este caso no es una isometría).

1.3. Una demostración alternativa

Como mencionamos en el apartado anterior, las demostraciones del Teorema de Lax-Milgram usualmente son variantes de aquella mostrada en la sección anterior, utilizando como apoyo el Teorema de Representación de Riesz, o pertenecientes a la familia de demostraciones que utiliza este teorema en combinación con el Teorema del Punto Fijo de Banach. Sin embargo, es posible realizar dicha demostración sin utilizar el Teorema de Riesz, utilizando en su lugar una variante del Teorema de Hahn-Banach, atribuida a S. Mazur, que presentamos a continuación.

Teorema 1.3.1. (Hahn-Banach) *Sea M un subespacio cerrado de un espacio de Banach $\mathcal{B}$. Entonces para todo $x_0 \in \mathcal{B} \setminus M$ existe una funcional lineal continua $f_0 \in \mathcal{B}^*$ tal que $f_0(x_0) > 1$ y $f_0 = 0$ en M.*

Demostración. La prueba de este teorema, así como la de otros teoremas de Hahn-Banach, se realiza a detalle en el Apéndice B. $\square$

Utilizando este resultado, proponemos la siguiente demostración del Teorema de Lax-Milgram.

Demostración Alternativa de (1.2.4).

1. Fijando el elemento $y \in \mathcal{H}$, la aplicación $B(\cdot, y) : \mathcal{H} \to \mathbb{R}$ es una funcional lineal sobre $\mathcal{H}$ para la primera entrada. Definimos entonces la transformación $T : \mathcal{H} \to \mathcal{H}^*$ dada por $Ty = B(\cdot, y)$.

2. T es automáticamente lineal y continua ya que B lo es. T es además acotada por debajo pues

$$\delta\|y\|^2 \leq |B(y,y)| = |Ty(y)| = \|y\| \left| Ty\left(\frac{y}{\|y\|}\right)\right| \leq \|y\| \sup_{\|y\|=1} \{y \in \mathcal{H}\} = \|y\|\|Ty\|$$

es decir

$$\delta\|y\| \leq \|Ty\|$$

Entonces los Lemas (1.2.2) y (1.2.3) aseguran que T es inyectiva y además $R(T)$ es cerrado en $\mathcal{H}^*$.

3. $R(T) = \mathcal{H}^*$. Supongamos lo contrario, es decir, que existe $\hat{F} \in \mathcal{H}^*$ tal que $\hat{F} \notin R(T)$ (obviamente $\hat{F} \neq 0$). Por el Teorema (1.3.1), existe $z \in \mathcal{H}^{**}$ tal que $z(\hat{F}) > 1$ y

$z(F) = 0$ para todo $F \in R(T)$. Como $\mathcal{H}$ es reflexivo, podemos identificar z con $x \in \mathcal{H}$, entonces las propiedades anteriores se convierten en $\hat{F}(x) > 1$ y $B(x, y) = Ty(x) = 0$ para toda y en $\mathcal{H}$. La coercitividad fuerte de B y la segunda de estas dos propiedades implican que $x = 0$, y por la linealidad de $\hat{F}$ se tiene que $\hat{F}(x) = 0$, una contradicción. Entonces vemos que $R(T) = \mathcal{H}^*$.

4. Finalmente, el punto anterior implica que para cada $F \in \mathcal{H}^*$, existe un elemento $y_0 \in \mathcal{H}$ tal que $F = Ty_0$, es decir, $F(x) = Ty_0(x) = B(x, y_0)$ para todo x en $\mathcal{H}$. Dicho elemento y_0 es único por la inyectividad de T. Esto es precisamente lo que buscábamos demostrar. □

Esta demostración Teorema de Lax-Milgram es similar a la original, presentada en la sección anterior, en el hecho que define una transformación T asociada a la forma bilineal B. Sin embargo, difiere en el hecho que en este caso la transformación T tiene como contradominio el espacio dual de $\mathcal{H}$, y no $\mathcal{H}$ mismo como el caso anterior (en el cual T era un operador), además de que en este caso la transformación no está definida a través del producto interno sino directamente como una de las entradas de B. Este hecho es de vital importancia, pues sugiere que una prueba de este estilo puede ser utilizada para espacios que no poseen un producto interno, como los de Banach.

Otra idea que surge a partir de esta demostración alternativa es que si el Teorema de Hahn-Banach puede fungir como sustituto del de Representación de Riesz en este caso, debe existir algún tipo de relación entre ellos. En efecto, de la familia de teoremas de extensión de Hahn-Banach, algunos de ellos situados en espacios de Hilbert implican resultados muy similares al de Riesz. De hecho, el Teorema (1.3.1) aquí utilizado, cuando se toma a $\mathcal{B}$ como un espacio de Hilbert, es esencialmente el Teorema de la Proyección (1.1.30), que hemos utilizado en la sección (1.1) para demostrar el Teorema de Representación de Riesz. Otro momento en la demostración en que el Teorema de Riesz aparece implícitamente es cuando utilizamos la hipótesis de reflexividad del espacio de Hilbert. Sin embargo, no estamos formalmente utilizando el Teorema de Riesz pues la propiedad de reflexividad de los espacios de Hilbert puede ser también demostrada - aunque más tortuosamente - con el Teorema de Hahn-Banach.

En conclusión, la demostración alternativa del Teorema de Lax-Milgram que presentamos aquí es simplemente un *atajo*, que se apoya en un resultado más general que el de Riesz, pero que al hacer esto (y evitando el uso del producto interno) otorga la ventaja de volverse potencialmente útil para variantes más generales del Teorema de Lax-Milgram. De cualquier manera, el hecho de que esta demostración se pueda hacer apoyada ya sea en el Teorema de Riesz o en el de Hahn-Banach, ejemplifica una característica interesante del análisis funcional: gran parte de la teoría se puede sustentar simplemente sobre un grupo pequeño de resultados trascendentales, todos ellos íntimamente relacionados.

1.4. Implicaciones

En esta sección analizaremos algunas consecuencias, observaciones y sutiles resultados paralelos al Teorema de Lax-Milgram. Comenzaremos investigando las nociones de suficiencia y necesidad en éste.

Según lo que nos muestra el Teorema (1.2.4), la coercitividad fuerte (1.1.29) de B es una condición suficiente para que exista una solución única al problema:

$$\begin{cases} \textit{Encontrar} \quad y \in \mathcal{H} \quad \text{tal que} \\ B(x, y) = F(x), \quad \forall x \in \mathcal{H} \end{cases} \tag{1.19}$$

Sin embargo, dicha condición no es una condición necesaria. En caso de que lo fuese, el Teorema de Lax-Milgram sería una implicación en ambos sentidos (como en efecto sucederá para algunas de las generalizaciones). En el contexto de las aplicaciones este es un hecho crucial: muestra que la coercitividad fuerte no es en general una condición óptima. No obstante esto, en el caso en que la forma bilineal B es simétrica y positiva (además de acotada), la coercitividad fuerte de B sí es una condición necesaria para la existencia de la solución al problema (1.19). Esto se resume en el siguiente lema.

Lema 1.4.1. *Sea B una forma bilineal simétrica, positiva y acotada. Si para cada $F \in \mathcal{H}^*$ existe una única $y \in \mathcal{H}$ que cumple $B(x, y) = F(x)$ para toda $x \in \mathcal{H}$, entonces la forma B es fuertemente coercitiva.*

Demostración. Para empezar, demostraremos que la forma B cumple una desigualdad análoga a CBS. Por la positividad de B sabemos que $B(u - \lambda v, u - \lambda v) \geq 0$ para cualesquiera $u, v \in \mathcal{H}$ y $\lambda \in \mathbb{R}$. Utilizando la bilinealidad y simetría de de B vemos que:

$$0 \leq B(u - \lambda v, u - \lambda v) = B(u, u) - 2\lambda B(u, v) + \lambda^2 B(v, v)$$

Si suponemos que $B(v, v) \neq 0$ (ese caso es trivial), tomemos $\lambda = \frac{B(u,v)}{B(v,v)}$. Entonces a partir de lo anterior obtenemos

$$0 \leq B(u, u) - 2\frac{B(u, v)^2}{B(v, v)} + \frac{B(u, v)^2}{B(v, v)} = B(u, u) - \frac{B(u, v)^2}{B(v, v)}$$

Esto equivale a

$$B(u, v) \leq B(u, u)^{1/2} B(v, v)^{1/2} \tag{1.20}$$

Sea $T : \mathcal{H} \to \mathcal{H}^*$ la transformación dada por $Ty = B(\cdot, y)$. Según las hipótesis, para cada $F \in \mathcal{H}^*$, la preimagen bajo T de F existe y es única. Esto equivale a decir que T es biyectiva. Si una transformación en un espacio de Hilbert es biyectiva, entonces debe ser acotada por debajo (probaremos esto para el caso más general de los espacios de Banach en el Capítulo 2, Lema (2.3.5)), por lo que existe δ positiva tal que $\delta\|y\| \leq \|Ty\|$ para toda $y \in \mathcal{H}$. Desarrollando esto vemos que

$$\begin{aligned}\delta\|y\| \leq \|Ty\| = \sup\left\{\frac{Ty(x)}{\|x\|}, x \neq 0\right\} \\ = \sup\left\{\frac{B(x,y)}{\|x\|}, x \neq 0\right\} \leq \frac{B(x,x)^{1/2}B(y,y)^{1/2}}{\|x\|} \qquad \forall x \in \mathcal{H}\end{aligned}$$

Dado que lo anterior se cumple para toda y, podemos en particular tomar $y = x$, dejando x libre, y obtener

$$\delta\|x\|^2 \leq B(x,x) \quad \forall x \in \mathcal{H}$$

es decir, la coercitividad fuerte de B. □

La razón por la cual la coercitividad fuerte no es una condición necesaria radica en que es demasiado impositiva, por lo que es suficiente para asegurar el resultado, pero demasiado para que el converso del teorema se cumpla. Intuitivamente, esto se podría resolver imponiendo a la forma bilineal alguna otra condición menos fuerte. Dicha condición, la *coercitividad débil* que definiremos más adelante, será fundamental para muchas de las generalizaciones.

Existe otro hecho que a pesar de no ser evidente, es crucial en las hipótesis del Teorema de Lax-Milgram: la forma bilineal B puede no ser simétrica. Esta sutileza es precisamente la que distingue este teorema del de Representación de Riesz, como veremos a continuación.

Si B cumple las hipótesis del Teorema (1.2.4) y además es simétrica, entonces es un producto interno en $\mathcal{H}$. Esto es fácil de ver, pues las primeras dos condiciones de la definición (1.1.5) son evidentes por la linealidad de B, la tercera es precisamente la simetría y

$$(iv) \qquad B(x,x) \geq 0 \quad \forall x \in \mathcal{H} \quad \text{y} \quad B(x,x) = 0 \Leftrightarrow x = 0$$

es consecuencia directa de la coercitividad fuerte, pues

$$B(x,x) \geq \delta\|x\|^2 \geq 0 \quad \forall x \in \mathcal{H}, \quad 0 = B(x,x) \geq \delta\|x\| \Leftrightarrow x = 0$$

Además, la norma inducida por dicho producto interno $\|v\|_B^2 = B(v,v)$ es equivalente a la norma original en $\mathcal{H}$ pues

$$\delta\|v\|_{\mathcal{H}}^2 \leq B(v,v) = \|v\|_B^2 \leq M\|v\|_{\mathcal{H}}\|v\|_{\mathcal{H}} = M\|v\|_{\mathcal{H}}^2$$

por las condiciones de acotamiento y coercitividad. De esta manera, vemos claramente que cuando la forma B es simétrica, el Teorema de Lax-Milgram es simplemente una reformulación del Teorema de Riesz, asegurando que la solución $y \in \mathcal{H}$ es el representante de F para el producto interno $B(\cdot,\cdot)$. Este hecho muestra que la verdadera importancia del Teorema de Lax-Milgram radica en ser válido aún para formas bilineales no simétricas. Sin embargo, si la forma B es simétrica entonces además existe una forma explícita de construir el elemento y cuya existencia está garantizada por este teorema.

Teorema 1.4.2. *Sea X un espacio con producto interno, B una forma bilineal acotada, positiva y simétrica. Entonces $y \in X$ cumple*

$$B(x,y) = F(x) \quad \forall x \in X \tag{1.21}$$

si y sólo si minimiza la siguiente funcional:

$$J(v) = \frac{1}{2}B(v,v) - F(v)$$

Demostración. Sean $x, y \in X$ y $t \in \mathbb{R}$. Veamos que

$$\begin{aligned} J(tx+y) &= \frac{1}{2}B(tx+y, tx+y) - F(tx+y) \\ &= \frac{1}{2}B(y,y) + B(tx,y) + \frac{t^2}{2}B(x,x) - tF(x) - F(y) \\ &= J(y) + t[B(x,y) - F(x)] + \frac{t^2}{2}B(x,x) \end{aligned} \tag{1.22}$$

Ahora, supongamos que $y \in X$ cumple (1.21). Entonces utilizando la ecuación anterior con $t = 1$ obtenemos

$$J(x+y) = J(y) + \frac{1}{2}B(x,x) \geq J(y) \quad \forall x \in X \tag{1.23}$$

es decir, y minimiza J. Conversamente, si J tiene un mínimo en $y \in X$ entonces la derivada de la función $\phi(t) = J(tx+y)$ debe ser nula en $t = 0$ para toda $x \in X$. Vemos así a partir de (1.22) que

$$\phi'(0) = B(x,y) - F(x) = 0 \quad \forall x \in X$$

es decir, se cumple (1.21). $\square$

Si en el teorema anterior se sustituye la condición de positividad por la coercitividad fuerte se obtiene un resultado adicional.

Corolario 1.4.3. *Supongamos que en el Teorema* (1.4.2) *la forma bilineal B es además fuertemente coercitiva con constante $\delta > 0$. Entonces la funcional J tiene un único mínimo.*

Demostración. Retomando (1.23) de la demostración del Teorema (1.4.2), vemos que para cualquier x diferente de cero se tiene

$$J(x+y) = J(y) + \frac{1}{2}B(x,x) \geq J(y) + \delta\|x\|^2 > J(y)$$

es decir, y es el único mínimo de la funcional J. □

Podemos concluir a partir de los dos resultados anteriores que encontrar una solución al problema (1.19) puede ser interpretado como un problema de optimización, siempre y cuando la forma bilineal B sea simétrica. Más aún, cuando B es además fuertemente coercitiva, este problema de optimización tiene una solución única. Esto sugiere que podemos interpretar la coercitividad fuerte de B como una propiedad de *convexidad estricta*[9] de la funcional J, pues garantiza la existencia de a lo más un punto mínimo.

La funcional J tiene aún otra aplicación en la teoría circundante al Teorema de Lax-Milgram: permite demostrar una variante de dicho teorema para conjuntos convexos. Esta versión, extraída de la teoría de elementos finitos[10], consiste en utilizar como dominio de la funcional F (del las hipótesis del Teorema de Lax-Milgram) cualquier subconjunto convexo $\mathcal{W}$ del espacio de Hilbert principal. El interés de esta generalización radica no sólo en su utilidad para varias aplicaciones prácticas dentro y fuera del método de elementos finitos, sino también en el hecho que ejemplifica la estrecha relación que existe en entre los teoremas de existencia y unicidad (como lo es el de Lax-Milgram) con los problemas de optimización de funcionales. Además, es importante porque, a diferencia del Teorema (1.4.2), no requiere que la forma bilineal sea simétrica.

Teorema 1.4.4. *Sea $\mathcal{W}$ un conjunto convexo y cerrado en un espacio de Hilbert $\mathcal{H}$ y $B : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ un operador bilineal, acotado y fuertemente coercitivo. Entonces para cada $F \in \mathcal{H}^*$ existe un único $y_0 \in \mathcal{W}$ tal que $J(y_0) = \inf_{w\in\mathcal{W}} J(w)$ donde*

$$J(w) = \frac{1}{2}B(w,w) - F(w)$$

[9]En optimización convexa, una función $f : C \to \mathbb{R}$, donde C es un subconjunto convexo de un espacio vectorial, es llamada **estrictamente convexa** si $f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2)$ para toda $t \in (0,1)$. Entre otras cosas, esta propiedad asegura que si la función tiene un mínimo, este será único.

[10]Ver por ejemplo [21, pp. 24-25].

Demostración. Veamos que la funcional J es acotada por debajo pues:

$$J(w) \geq \frac{1}{2}\delta\|w\|^2 - \|F\|\|w\| = \frac{1}{2\delta}(\delta\|w\| - \|F\|)^2 - \frac{\|F\|^2}{2\delta} \geq -\frac{\|F\|^2}{2\delta}$$

Sea $a = \text{ínf}_{w\in\mathcal{W}} J(w)$ y $(w_n)_{n=1}^{\infty}$ una sucesión minimizadora, es decir

$$\lim_{n\to\infty} J(w_n) = a,\ w_n \in \mathcal{W}$$

Entonces

$$\begin{aligned}
\delta\|w_n - w_m\|^2 &\leq B(w_n - w_m, w_n - w_m) \\
&= 2B(w_n, w_n) + 2B(w_m, w_m) - B(w_n + w_m, w_n + w_m) \\
&= 4J(w_n) + 4J(w_m) - 8J\left(\frac{w_n + w_m}{2}\right) \\
&\leq 4J(w_n) + 4J(w_m) - 8a
\end{aligned} \tag{1.24}$$

donde $\frac{1}{2}(w_n + w_m) \in \mathcal{W}$ por la convexidad de $\mathcal{W}$. Ahora, como ambas $J(w_n) \to a$ y $J(w_m) \to a$, entonces se tiene que $\|w_n - w_m\| \to 0$ cuando $n, m \to \infty$. Por lo tanto, $(w_n)_{n=1}^{\infty}$ es una sucesión de Cauchy en $\mathcal{H}$ y por tanto existe $y_0 \in \mathcal{H}$ tal que $w_n \to y_0$. Además, como $\mathcal{W}$ es cerrado, podemos concluir que $y_0 \in \mathcal{W}$. La continuidad de J implica que

$$J(y_0) = \lim_{n\to\infty} J(w_n) = \inf_{w\in\mathcal{W}} J(w)$$

Ahora, para probar que la solución $y_0 \in \mathcal{W}$ es única, supongamos que y_0 y y_1 son dos soluciones. Claramente la sucesión $y_0, y_1, y_0, y_1, ...$ es una sucesión minimizadora, pero vimos antes que toda sucesión minimizadora debe ser de Cauchy. Esto implica que $y_0 = y_1$. □

Como podemos ver, la funcional $J(v) = \frac{1}{2}B(v, v) - F(v)$ tiene gran importancia teórica. Dicha importancia se manifiesta además en diversas aplicaciones, pues la función J comúnmente es utilizada para modelar ciertos efectos físicos que intervienen en la formulación de un problema. Por ejemplo, en varias aplicaciones a ecuaciones diferenciales parciales de tipo elíptico (normalmente derivadas de la física) que describen estados de equilibrio o de mínima energía de un sistema, dicha funcional representa una función de energía abstracta. Existe además un esquema numérico basado en la minimización de funciones de energía de este tipo, llamado el Método de Ritz. Retomaremos la funcional J cuando analicemos aplicaciones del Teorema de Lax-Milgram a la ecuación de Poisson en el tercer capítulo.

Capítulo 2

Generalizaciones

Casi desde el momento en que el Teorema de Lax-Milgram irrumpió en el campo del análisis funcional, surgió el interés por encontrar posibles generalizaciones a tan poderoso resultado. Habría de transcurrir menos de una década para que las primeras de estas generalizaciones hicieran su aparición a manos de Jacques-Louis Lions (1961) y Jindřich Nečas (1962), y en las décadas subsecuentes surgieron cada vez más.

Estas generalizaciones abarcan una gran variedad de posibles modificaciones a las hipótesis del teorema original, casi todas ellas involucrando un dominio de la forma bilineal B más general, ya sea permitiendo que sea el producto cartesiano de espacios distintos o que estos espacios sean de clases más generales que los espacios de Hilbert. Muchas otros autores se enfocaron en descubrir condiciones adicionales que permitirían obtener la conclusión de dicho teorema para espacios más particulares, con ciertas propiedades específicas.

Dentro de las posibles generalizaciones del teorema, la que probablemente parece más natural e intuitiva es aquella de permitir que el dominio de la forma bilineal B sea el producto cartesiano de dos espacios de Hilbert no necesariamente iguales. Por esta razón, esta idea será la primera que analizaremos en las dos primeras secciones. Posteriormente, lo haremos con otras generalizaciones que parten de ideas distintas a ésta, variando de alguna manera u otra el dominio de dicha forma bilineal. De cualquier manera, es digno de notar la gran fecundidad de resultados a los que dio pie el Teorema de Lax Milgram a raíz del interés inmediato que suscitó, siendo algunos de estos resultados posteriores muy importantes para el desarrollo de diferentes áreas de las matemáticas.

2.1. Una generalización para $\mathcal{H}_A$ y $\mathcal{H}_{A'}$

Edward Landesman realizó un primer acercamiento hacia la posibilidad de utilizar el Teorema de Lax-Milgram para una forma bilineal B cuyo dominio constara de espacios de Hilbert distintos. Dicho acercamiento fue presentado en [11], desarrollando una generalización para este teorema en la cual la forma bilineal tiene como dominio el producto cartesiano $\mathcal{H}_A \times \mathcal{H}_{A'}$, espacios de Hilbert asociados a una transformación A y su $*$-recíproco $A' = A^{*-1}$. Estos espacios, introducidos por él mismo en otro artículo anterior [10], son de gran importancia para el estudio de ecuaciones diferenciales parciales, en las cuales A es un operador diferencial. Mostraremos a continuación cómo se obtienen dichos espacios.

Supongamos que $\mathcal{H}$ y $\mathcal{H}_1$ son dos espacios de Hilbert cualesquiera, y A una transformación lineal de $\mathcal{H}$ en $\mathcal{H}_1$, con rango cerrado en $\mathcal{H}_1$ y con dominio denso en $\mathcal{H}$. A continuación definimos la siguiente norma en el rango

$$\|x\|_A = \|Ax\| \qquad \forall x \in D_A \tag{2.1}$$

con la cual el dominio $D(A) = D_A$ es un espacio completo. Para que $\|\cdot\|_A$ sea efectivamente una norma, supondremos en este punto que $N(A)$, el núcleo de A, es únicamente $\{0\}$. El espacio $(D_A, \|\cdot\|)$ que se obtiene como resultado de esto es un espacio de Hilbert y lo llamaremos $\mathcal{H}_A$.

Sea A' el pseudo-inverso del adjunto de A (también llamado $*$-recíproco), es decir $A' = A^{*-1}$. Como se hizo antes, se completa el dominio $D(A') = D_{A'}$ con respecto a la norma

$$\|x\|_{A'} = \|A'x\| \qquad \forall x \in D_{A'} \tag{2.2}$$

y se bautiza al espacio de Hilbert así obtenido como $\mathcal{H}_{A'}$.

Por otro lado, la cerradura $\overline{R}_A$ del rango de A es un espacio de Hilbert con la norma natural en $\mathcal{H}_1$. Además, A es claramente una transformación de D_A en R_A que preserva la norma, tomando la norma en D_A como (2.1). Entonces extendemos A linealmente a una transformación unitaria[1] $\tilde{A} : \mathcal{H}_A \to \overline{R}_A$. Completar $D_{A'}$ con respecto a la norma (2.2) es equivalente a hacerlo con respecto a

$$\|x\|_{A'} = \sup_{y \in D_A} \frac{\langle x, y\rangle}{\|Ay\|} \tag{2.3}$$

pues si escribimos $y = A^{-1}z$ para $z \in D_{A^{-1}}$ entonces

$$\sup_{y \in D_A} \frac{\langle x, y\rangle}{\|Ay\|} = \sup_{z \in D_{A^{-1}}} \frac{\langle x, A^{-1}z\rangle}{\|z\|} = \sup_{z \in D_{A^{-1}}} \frac{\langle A'x, z\rangle}{\|z\|} = \|A'x\|$$

[1]Una transformación es unitaria si su adjunto coincide con su inverso. La definición formal de esto, además de algunas propiedades de los operadores unitarios se presentan en el Apéndice A.

Llamamos $\tilde{A}'$ a la transformación unitaria con dominio $\mathcal{H}_{A'}$ e imagen $\overline{R}_{A'}$. El primer resultado que se puede obtener a partir de esto es

Lema 2.1.1. $\overline{R}_A = \overline{R}_{A'}$

Demostración. Para empezar veamos que

$$(\overline{R}_A)^{\perp} = R_A^{\perp} = N_{A^{-1}} = N_{A^*} = R_{A'}^{\perp} = (\overline{R}_{A'})^{\perp}$$

y como $\overline{R}_A$ y $\overline{R}_{A'}$ son cerrados se tiene que

$$\overline{R}_A = [(\overline{R}_A)^{\perp}]^{\perp} = [(\overline{R}_{A'})^{\perp}]^{\perp} = \overline{R}_{A'}$$

Observación. A partir de lo anterior, se vuelve evidente que $\tilde{A}'$ transforma $\mathcal{H}_{A'}$ en $\overline{R}_A$, y viceversa para $\tilde{A}$. Esto nos lleva a notar que, por la forma en la que se construyeron los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$, dado un elemento $y \in \mathcal{H}_{A'}$ cualquiera, siempre se puede encontrar $x \in \mathcal{H}_A$ tal que $\tilde{A}x = \tilde{A}'y$. Además, se tiene en ese caso que

$$\|x\|_{\mathcal{H}_A} = \|\tilde{A}x\| = \|\tilde{A}'y\| = \|y\|_{\mathcal{H}_{A'}}$$

Esta idea será crucial tanto para el resto de la teoría referente a dichos espacios, como para la demostración del teorema principal.

Por la manera en que hemos definido los conceptos anteriores, podemos definir un producto interno para $x \in \mathcal{H}_A, y \in \mathcal{H}_{A'}$ de la siguiente manera

$$\langle x, y\rangle_* := \langle \tilde{A}x, \tilde{A}'y\rangle_{\mathcal{H}_1} \tag{2.4}$$

Notemos que dicha definición es congruente con el producto interno original en $\mathcal{H}$, pues si x está en D_A y además y está en $D_{A'}$ entonces

$$\langle x, y\rangle_* = \langle \tilde{A}x, \tilde{A}'y\rangle = \langle Ax, A'y\rangle = \langle A^{-1}Ax, y\rangle = \langle x, y\rangle_{\mathcal{H}}$$

En el contexto de los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$, definiremos la coercitividad de una forma bilineal de la siguiente manera, para que sea congruente con la definición usual.

Definición 2.1.2. Sea $B(\cdot,\cdot) : \mathcal{H}_A \times \mathcal{H}_{A'} \to \mathbb{R}$. Decimos que B es **fuertemente coercitiva** si existe una constante δ positiva tal que

$$B(x, y) \geq \delta\|y\|^2_{\mathcal{H}_{A'}}$$

para todas las x, y tales que $\tilde{A}x = \tilde{A}'y$.

Observación. Es claro que la definición anterior se puede formular también en términos de la norma de x, pues como $\tilde{A}x = \tilde{A}'y$, se tiene que $\|x\|_{\mathcal{H}_A} = \|y\|_{\mathcal{H}_{A'}}$, entonces

$$B(x,y) \geq \delta\|y\|^2_{\mathcal{H}_{A'}} = \delta\|x\|^2_{\mathcal{H}_{A'}}$$

Una vez que hemos construido la estructura básica de los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$, podemos pasar a presentar algunos resultados que serán necesarios para la demostración del teorema que nos concierne. El primero de ellos es una ligera generalización de la desigualdad de Cauchy-Bunyakovsky-Schwarz.

Teorema 2.1.3. (Desigualdad CBS Generalizada) *Si $x \in \mathcal{H}_A$, $y \in \mathcal{H}_{A'}$ entonces*

$$|\langle x, y\rangle| \leq \|x\|_{\mathcal{H}_A}\|y\|_{\mathcal{H}_{A'}}$$

Demostración. El resultado se obtiene directamente a partir de la definición del producto interno (2.4) y la desigualdad CBS original

$$|\langle x, y\rangle| = |\langle \tilde{A}x, \tilde{A}'y\rangle| \leq \|\tilde{A}x\|\|\tilde{A}'y\| = \|x\|_{\mathcal{H}_A}\|y\|_{\mathcal{H}_{A'}}$$

□

Otro resultado que se puede generalizar fácilmente de esta manera y que además resultará, como anteriormente, crucial para demostrar la generalización del Teorema de Lax-Milgram en cuestión, es el Teorema de Representación de Riesz, cuya versión original presentamos en (1.1.31).

Teorema 2.1.4. (Representación de Riesz Generalizado) *Si f es una funcional lineal continua sobre $\mathcal{H}_A$ entonces existe un único $y_0 \in \mathcal{H}_{A'}$ tal que $f(x) = \langle x, y_0\rangle_*$ para todo $x \in \mathcal{H}_A$.*

Demostración. Dado que $\mathcal{H}_A$ es un espacio de Hilbert, por el Teorema de Riesz original se tiene que

$$f(x) = \langle x, z\rangle_{\mathcal{H}_A} \qquad \text{para algún } z \in \mathcal{H}_A$$

Por la observación al Lema (2.1.1), podemos escoger $y_0 \in \mathcal{H}_{A'}$ tal que $\tilde{A}z = \tilde{A}'y_0$. Así,

$$\langle x, z\rangle_{\mathcal{H}_A} = \langle \tilde{A}x, \tilde{A}z\rangle_{\mathcal{H}_1} = \langle \tilde{A}x, \tilde{A}'y_0\rangle_{\mathcal{H}_1} = \langle x, y_0\rangle_*$$

Entonces $f(x) = \langle x, y_0\rangle_*$ para algún $y_0 \in \mathcal{H}_{A'}$. □

Finalmente, daremos paso a la ya mencionada generalización propuesta por Landesman. La demostración del teorema que utilizaremos a continuación es una adaptación de la presentada originalmente por él mismo, procediendo en una manera similar a como lo hicimos para el teorema original (1.2.4).

Teorema 2.1.5. (Lax-Milgram Generalizado) *Sean $\mathcal{H}$ y $\mathcal{H}_1$ dos espacios de Hilbert reales. Sean $\mathcal{H}_A$ y $\mathcal{H}_{A'}$ los espacios de Hilbert definidos como en lo anterior a partir de una transformación A de $\mathcal{H}$ en $\mathcal{H}_1$ y sea $B(x,y) : \mathcal{H}_A \times \mathcal{H}_{A'} \to \mathbb{R}$ forma bilineal, acotada y fuertemente coercitiva en el sentido de* (2.1.2). *Entonces dada una funcional lineal acotada F en $\mathcal{H}_A$, existe una única y_0 en $\mathcal{H}_{A'}$ tal que*

$$B(x, y_0) = F(x) \quad \forall x \in \mathcal{H}_A$$

Demostración. 1. Fijando $y \in \mathcal{H}_{A'}$ cada $B_y(x) = B(x,y)$ es una funcional lineal en $\mathcal{H}_A$. Claramente B_y es continuo pues

$$\|B_y\| = \sup_{\substack{x\in\mathcal{H}_A \\ \|x\|=1}} |B(x,y)| \leq \gamma\|y\|_{\mathcal{H}_{A'}}$$

Entonces por el teorema (2.1.4) existe un único $z \in \mathcal{H}_{A'}$ tal que

$$B(x,y) = B_y(x) = \langle x, z\rangle \quad \forall\, x \in \mathcal{H}_A \tag{2.5}$$

Por lo tanto para cada $y \in \mathcal{H}_{A'}$ podemos definir el operador T de $\mathcal{H}_{A'}$ en $\mathcal{H}_{A'}$ tal que $Ty = z$ si $\langle x, z\rangle = B(x,y)$.

2. El operador T es trivialmente lineal. Además es continuo, pues si tomamos $y \in \mathcal{H}_{A'}$ arbitrario, por la definición de T, tenemos que $B(x,y) = \langle x, Ty\rangle \quad \forall x \in \mathcal{H}_A$, por lo que según una observación anterior podemos escoger x tal que $\tilde{A}x = \tilde{A}'Ty$. Entonces

$$\begin{aligned}\|Ty\|^2_{\mathcal{H}_{A'}} &= \langle Ty, Ty\rangle_{\mathcal{H}_{A'}} = \langle \tilde{A}'(Ty), \tilde{A}'(Ty)\rangle = \langle \tilde{A}x, \tilde{A}'Ty\rangle \\ &= \langle x, Ty\rangle = B(x,y) \leq \gamma\|x\|_{\mathcal{H}_A}\|y\|_{\mathcal{H}_{A'}} = \gamma\|Ty\|_{\mathcal{H}_{A'}}\|Ty\|_{\mathcal{H}_{A'}}\end{aligned} \tag{2.6}$$

por lo que

$$\|Ty\|_{\mathcal{H}_{A'}} \leq \gamma\|y\|_{\mathcal{H}_{A'}}$$

T además es acotado por debajo pues por un argumento similar al anterior, escogiendo x tal que $\tilde{A}x = \tilde{A}'y$, utilizando la propiedad de coercitividad (2.1.2) y la desigualdad de Cauchy-Bunyakovsky-Schwarz generalizada (2.1.3), tenemos que

$$\delta\|y\|^2_{\mathcal{H}_{A'}} \leq B(x,y) = \langle x, Ty\rangle \leq \|x\|_{\mathcal{H}_A}\|Ty\|_{\mathcal{H}_{A'}}$$

pero $\|x\|_{\mathcal{H}_A} = \|y\|_{\mathcal{H}_{A'}}$, por lo que

$$\delta\|y\|_{\mathcal{H}_{A'}} \leq \|Ty\|_{\mathcal{H}_{A'}} \tag{2.7}$$

Entonces, por los Lemas (1.2.2) y (1.2.3) sabemos que T^{-1} existe y es continuo (acotada por la constante $\frac{1}{\delta}$) y además $R(T)$ es cerrado en $\mathcal{H}_A$. Entonces, supongamos que $R(T) \neq \mathcal{H}_{A'}$. Así, existe $w \in \mathcal{H}_{A'}$ con $w \neq 0$ tal que $w \in R(T)^{\perp}$. Escojamos $x \in \mathcal{H}_A$ tal que $\tilde{A}x = \tilde{A}'w$. Pero esto implica que

$$\delta\|w\|^2 \leq B(x,w) = \langle x, Tw\rangle = \langle \tilde{A}x, \tilde{A}'(Tw)\rangle = \langle \tilde{A}'w, \tilde{A}'Tw\rangle = \langle w, Tw\rangle_{\mathcal{H}_{A'}} = 0$$

luego, $w = 0$ y concluimos entonces que $R(T) = \mathcal{H}_{A'}$.

3. Sea $F \in (\mathcal{H}_A)^*$, usando una vez más (2.1.4) tenemos que $F(x) = \langle x, y^*\rangle$ para un único $y^* \in \mathcal{H}_{A'}$, con $\|y^*\|_{\mathcal{H}_{A'}} = \|F\|_{(\mathcal{H}_A)^*}$. Tomando $y_0 := T^{-1}y^*$ obtenemos el resultado deseado. Además, vemos que $\|y_0\| = \|T^{-1}y^*\| \leq \frac{1}{\delta}\|y_0\| = \frac{1}{\delta}\|F\|_{(\mathcal{H}_A)^*}$

4. Ahora para demostrar la unicidad de y_0, supongamos que existen y_0 y y_1 en $\mathcal{H}_{A'}$ tales que

$$B(x,y_0) = F(x) \qquad \text{y} \qquad B(x,y_1) = F(x) \qquad \forall x \in \mathcal{H}_A$$

entonces

$$B(x, y_0 - y_1) = B(x,y_0) - B(x,y_1) = F(x) - F(x) = 0 \quad \forall x \in \mathcal{H}_A$$

como lo anterior se cumple para todo x en $\mathcal{H}_A$, por la forma en que fueron construidos $\mathcal{H}_A$ y $\mathcal{H}_{A'}$, podemos escoger x^* en $\mathcal{H}_A$ tal que

$$\tilde{A}x^* = \tilde{A}'(y_0 - y_1)$$

por lo que utilizando la coercitividad una vez más vemos que

$$\delta\|y_0 - y_1\|^2_{\mathcal{H}_A} \leq B(x^*, y_0 - y_1) = 0$$

es decir, $y_0 = y_1$. Con esto concluimos la demostración. □

En este punto, se muestra evidente que para obtener el Teorema de Lax-Milgram original como caso particular del teorema anterior, basta con tomar $\mathcal{H} = \mathcal{H}_1$ y la transformación A como la identidad I. De esta manera, $A = A' = I$ y $\mathcal{H}_A = \mathcal{H}_{A'} = \mathcal{H}$ y por consiguiente las hipótesis y conclusiones del Teorema (2.1.5) original coinciden con las de (1.2.4).

La generalización que tratamos en esta sección es sin duda interesante por el hecho que relaja la restricción - bastante limitativa - que impone el Teorema de Lax-Milgram sobre el dominio de la forma bilineal (esto es, que sea el producto cartesiano del mismo espacio de Hilbert). Es digno de notar también que no fue necesario exigir hipótesis adicionales o más fuertes que aquellas heredadas del teorema original, sino simplemente modificarlas (la coercitividad por ejemplo) para que fueran congruentes con las otras nociones definidas en los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$. Además, volver el teorema válido en dichos espacios que resultan ser, como dice Landesman, "de importancia fundamental cuando se estudian ecuaciones diferenciales parciales, particularmente las lineales elípticas"[2], es sin duda de gran utilidad práctica.

No obstante lo anterior, es claro que un Teorema de Lax-Milgram para los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$ no es un resultado significativamente más global que el original, y fuera del contexto de ciertas aplicaciones en particular, parece no ofrecer una gran amplitud teórica. Por esta razón, la importancia de esta generalización no radica tanto en volver el Teorema de Lax-Milgram un resultado más general, si no precisamente en hacerlo válido en un contexto específico, útil para las aplicaciones de ecuaciones diferenciales parciales.

2.2. Nečas-Babuška-Lax-Milgram

Una generalización subsecuente a la anterior, aunque más sencilla en su formulación, finalmente nos permitirá abarcar las formas bilineales sobre el producto cartesiano de cualesquiera dos espacios de Hilbert. Estos espacios podrán ser distintos y no tendrán que estar necesariamente definidos a través de transformaciones, como en el caso de la sección anterior.

El teorema que trata con esta generalización fue planteado originalmente por Jindřich Nečas en 1962 [14], tan solo 8 años después de la aparición del Teorema de Lax-Milgram, y contempla, en efecto, una forma bilineal B sobre el espacio $\mathcal{H}_1 \times \mathcal{H}_2$, donde $\mathcal{H}_1$ y $\mathcal{H}_2$ son espacios de Hilbert cualesquiera. Nečas, matemático checo especializado en ecuaciones diferenciales parciales y análisis funcional no lineal, mostró en dicha publicación que basta con sustituir la condición de la coercitividad fuerte por un tipo de coercitividad más débil para obtener el resultado del teorema original para dicho dominio de la forma bilineal.

Sin embargo, este resultado fue en gran parte popularizado por otro matemático checo, Ivo Babuška, en el contexto de la teoría de elementos finitos, inicialmente en 1971 [2] y posteriormente en 1972 [3], razón por la cual tiende a atribuírsele erróneamente a éste la introducción de esta generalización. En algunas fuentes ([17], por ejemplo) se hace referencia a este resultado con el nombre de "Teorema de Babuška-Lax-Milgram", cosa que evitaremos aquí para no restar importancia a su creador original. Es importante reconocer, de cualquier manera, el mérito de Babuška, un prolífico matemático del campo de las ecuaciones diferenciales parciales, en ser uno de los principales responsables de introducir teoremas de existencia (como el de Lax-Milgram y sus variantes) al método

[2] [11, p. 339]

de elementos finitos, y por consiguiente, a un conocimiento general mucho mayor.

Habiendo aclarado estos detalles importantes, pasaremos, a continuación, a definir el concepto de coercitividad débil.

Definición 2.2.1. Sean X, Y espacios normados sobre $\mathbb{K}$ y $B(\cdot,\cdot) : X \times Y \to \mathbb{K}$ una forma bilineal. Se dice que B es **débilmente coercitiva**[3] si existe una constante positiva δ tal que:

$$\sup_{\|y\|=1} \quad |B(x,y)| \geq \delta\|x\| \quad \forall x \in X,$$

y para toda $y \in Y \setminus \{0\}$

$$\sup_{x\in X} |B(x,y)| > 0$$

Observación. Como su nombre lo sugiere, la coercitividad fuerte en efecto implica la coercitividad débil, en el caso en que $X = Y$. Para ver esto, supongamos que B es fuertemente coercitiva, es decir, cumple

$$|B(x,x)| \geq \delta\|x\|^2 \quad \forall x \in X$$

entonces para $x \neq 0$ se tiene

$$\delta\|x\| \leq \frac{|B(x,x)|}{\|x\|} = \left|B\left(x, \frac{x}{\|x\|}\right)\right| \leq \sup_{\|y\|=1} |B(x,y)|$$

por lo tanto se cumple la primera condición de (2.2.1), pues el caso $x = 0$ se cumple también trivialmente. Por otra parte, si $y \neq 0$

$$0 < \delta\|y\| \leq |B(y,y)| \leq \sup_{x\in X} |B(x,y)|$$

Entonces B cumple también la segunda condición de (2.2.1), es decir, es débilmente coercitiva.

Resulta claro de antemano que la coercitividad fuerte no sería útil para el caso cuando los espacios subyacentes de la forma bilineal no son iguales, pues la definición misma de ésta implica que la forma pueda ser evaluada en un mismo elemento en ambas entradas. Vemos entonces que la coercitividad débil es, en cierta manera, a lo más que se puede aspirar en términos de coercitividad en este caso. Sin embargo, como veremos más adelante, resulta ser justamente suficiente para lo que a esta generalización concierne.

[3]En la formulación original del teorema, Nečas utiliza $\sup_{\|x\|=1} |B(x,y)| \geq \delta_2\|y\| \quad \forall y \in \mathcal{H}_2$ en lugar de la segunda de estas condiciones. Posteriormente, se hizo evidente que bastaba con la condición aquí expuesta para que el resultado fuera válido, y por consiguiente, la aquí presentada se volvió la definición estándar de coercitividad débil. Es por esto que el teorema tiende a ser enunciado con esta hipótesis en vez de en los términos originales de su creador.

Otra formulación de la definición anterior son las conocidas condiciones *inf-sup*.[4]

Definición 2.2.2. Sea $B(\cdot,\cdot) : X \times Y \to C$ forma bilineal. Se dice que B cumple las condiciones *inf-sup* si existe $\delta > 0$ tal que:

$$\inf_{\substack{x\in X\\ x\neq 0}} \sup_{\substack{y\in Y\\ y\neq 0}} \frac{B(x,y)}{\|x\|_X\,\|y\|_Y} \geq \delta$$

y

$$\inf_{\substack{y\in Y\\ y\neq 0}} \sup_{\substack{x\in X\\ x\neq 0}} \frac{B(x,y)}{\|x\|_X\,\|y\|_Y} > 0$$

Esta otra formulación de la coercitividad, un poco más confusa a primera vista, varía en el hecho que muestra dos condiciones simétricas (en una se toma el ínfimo sobre el espacio X y el supremo sobre el espacio Y, y en la otra es al revés), una de las cuales contiene una cota $\delta > 0$ y otra cumple simplemente la positividad. Sin embargo, ambas formulaciones son similares en que la primera condición es una de acotamiento inferior para los elementos en X y la segunda para elementos de Y. Mostramos a continuación que, en efecto, ambas formulaciones son intercambiables.

Proposición 2.2.3. *Las condiciones de la definición* (2.2.1) *y* (2.2.2) *son equivalentes.*

Demostración. Partiendo de la primera condición en (2.2.2) tenemos las siguientes equivalencias:

$$\begin{aligned}
\inf_{\substack{x\in X\\ x\neq 0}} \sup_{\substack{y\in Y\\ y\neq 0}} \frac{\|B(x,y)|}{\|x\|\|y\|} \geq \delta \quad &\Longleftrightarrow \quad \sup_{\substack{y\in Y\\ y\neq 0}} \frac{|B(x,y)|}{\|x\|\|y\|} \geq \delta \quad \forall x \in X \\
&\Longleftrightarrow \quad \sup_{\substack{y\in Y\\ y\neq 0}} \left| B\left(x, \frac{y}{\|y\|}\right)\right| \geq \delta\|x\| \quad \forall x \in X \\
&\Longleftrightarrow \quad \sup_{\|x\|=1} |B(x,y)| \geq \delta\|x\| \quad \forall x \in X
\end{aligned} \tag{2.8}$$

Para la segunda condición se tiene

$$\begin{aligned}
\inf_{\substack{y\in Y\\ y\neq 0}} \sup_{\substack{x\in X\\ x\neq 0}} \frac{|B(x,y)|}{\|x\|\|y\|} > 0 \quad &\Longleftrightarrow \quad \sup_{\substack{x\in X\\ x\neq 0}} \frac{|B(x,y)|}{\|x\|\|y\|} > 0 \quad \forall y \in Y \setminus \{0\} \\
&\Longleftrightarrow \quad \sup_{x\in X} |B(x,y)| > 0 \quad \forall y \in Y \setminus \{0\}
\end{aligned} \tag{2.9}$$

[4]La primera de estas condiciones, también llamada *condición de Babuška-Brezzi*, *condición LBB* (Ladyshenskaja-Babuška-Brezzi) o simplemente la *condición inf-sup*, juega un rol fundamental en la teoría de métodos de elementos finitos y ha sido fundamental en aplicaciones de flujos incompresibles, problemas de Navier-Stokes y flujos de Stokes. Ejemplos interesantes se pueden ver en [4].

Por lo tanto ambas condiciones son equivalentes. □

A grandes rasgos, la imposición de estas condiciones asegura que el valor de la forma bilineal B crece suficientemente rápido conforme el primero de sus argumentos crece. Además, asegura que B es no degenerado con respecto a su segundo elemento (es decir, que la forma B no es la forma nula cuando es evaluada en un elemento sobre Y no nulo). Habiendo ya expuesto e investigado la coercitividad débil, no queda más que dar paso al teorema que fundamenta esta generalización.

La demostración del teorema presentada a continuación sigue el esquema de la utilizada para la versión original del teorema (1.2.4), y es prácticamente idéntica a ésta, a excepción de algunos cambios necesarios en los lemas adicionales para tomar en cuenta el hecho de que la transformación T tendrá como contradominio al espacio $\mathcal{H}_2$, (posiblemente) distinto a $\mathcal{H}_1$.

Teorema 2.2.4. *Sean $\mathcal{H}_1$ y $\mathcal{H}_2$ espacios de Hilbert y $B : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{R}$ una función bilineal continua. Supongamos adicionalmente que B es débilmente coercitiva. Entonces, para toda $F \in \mathcal{H}_2^*$ existe un único $x_0 \in \mathcal{H}_1$ tal que*

$$B(x_0, y) = F(y) \quad \textit{para todo} \quad y \in \mathcal{H}_2$$

y además:

$$\|x_0\|_{\mathcal{H}_2} \leq \frac{1}{\delta}\|F\|_{\mathcal{H}_1^*}$$

Demostración. 1. Fijando $x \in \mathcal{H}_1$ cada $B_x(y) = B(x, y)$ es una funcional lineal en $\mathcal{H}_2$. Claramente B_x es continua pues

$$\|B_x\| = \sup_{\substack{y \in \mathcal{H}_2 \\ \|y\|=1}} |B(x, y)| \leq \gamma\|x\|_{\mathcal{H}_1}$$

Entonces por el Teorema (1.1.31) existe un único $z \in \mathcal{H}_2$ tal que

$$B(x, y) = B_x(y) = \langle y, z\rangle \quad \forall\, y \in \mathcal{H}_2 \tag{2.10}$$

Por lo tanto podemos definir para cada $x \in \mathcal{H}_1$ la transformación Tx de $\mathcal{H}_1$ en $\mathcal{H}_2$ tal que $Tx = z$ si $\langle y, z\rangle = B(x, y)$.

2. T es lineal y además es continua pues

$$\|Tx\|_{\mathcal{H}_2} = \sup_{\substack{y \in \mathcal{H}_2 \\ \|y\|\leq 1}} |\langle y, Tx\rangle| = \sup_{\substack{y \in \mathcal{H}_2 \\ \|y\|\leq 1}} |B(x, y)| \leq \sup_{\substack{y \in \mathcal{H}_2 \\ \|y\|\leq 1}} \gamma\|x\|_{\mathcal{H}_1}\|y\|_{\mathcal{H}_2} = \gamma\|x\|_{\mathcal{H}_1}$$

T es también acotada por debajo pues

$$\beta\|x\|_{\mathcal{H}_1} \leq \sup_{\substack{y\in\mathcal{H}_2\\ \|y\|\leq 1}} |B(x,y)| = \sup_{\substack{y\in\mathcal{H}_2\\ \|y\|\leq 1}} |\langle y, Tx\rangle| = \|Tx\|$$

Entonces, por los lemas (1.2.2) y (1.2.3), T^{-1} existe y es continua, y además $R(T)$ es cerrado en $\mathcal{H}_2$. Supongamos entonces que $R(T) \neq \mathcal{H}_2$. Entonces $\exists\, w \in \mathcal{H}_2$ con $w \neq 0$ tal que $w \in R(T)^{\perp}$. Pero esto implica que

$$|B(u,w)| = |\langle w, Tu\rangle| = 0 \quad \forall u \in \mathcal{H}_1$$

pero entonces

$$\sup_{u\in\mathcal{H}_1} |B(u,w)| = 0$$

que contradice la segunda condición de la coercitividad débil. Concluimos entonces que $R(T) = \mathcal{H}_2$.

3. Sea $F \in \mathcal{H}_2^*$, usando una vez más (1.1.31) tenemos que $F(y) = \langle y, y_0\rangle$ para una única $y_0 \in \mathcal{H}_2$, con $\|y_0\|_{\mathcal{H}_2} = \|F\|_{\mathcal{H}_2^*}$. Tomando $x_0 := T^{-1}y_0$ obtenemos el resultado deseado. Además, vemos que $\|x_0\| = \|T^{-1}y_0\| \leq \frac{1}{\delta}\|y_0\| = \frac{1}{\delta}\|F\|_{\mathcal{H}_2^*}$

4. Para mostrar ahora la unicidad de x_0, supongamos que existen x_0 y x_1 en $\mathcal{H}_1$ tales que

$$B(x_0, y) = F(y) \quad \text{y} \quad B(x_1, y) = F(y) \quad \forall\, y \in \mathcal{H}_2$$

entonces

$$B(x_0 - x_1, y) = B(x_0, v) - B(x_1, v) = \overline{F(v)} - \overline{F(v)} = 0 \quad \forall\, v \in \mathcal{H}_2$$

esto implica que

$$0 = \sup_{\substack{v\in\mathcal{H}_2\\ \|v\|=1}} |B(x_0 - x_1, v)| \geq \delta\|x_0 - x_1\|$$

por lo tanto $\|x_0 - x_1\| = 0$, ó $x_0 = x_1$. Es decir, el elemento x_0 es único. □

Observación. Hemos mostrado la versión del teorema en la cual $F \in \mathcal{H}_2^*$. En el caso en que $F \in \mathcal{H}_1^*$, ésta sería, como antes, $B(x,y) = F(x)$ para un único $y \in \mathcal{H}_2$, y en las condiciones de coercitividad débil se deberían invertir los roles de las variables de $\mathcal{H}_1$ y $\mathcal{H}_2$. Se escogió mostrar el primero de estos casos porque es el más común en la literatura, cosa que ha estandarizado la definición de la coercitividad débil de la manera que la presentamos aquí (2.2.1).

Como hemos podido notar, la demostración de esta generalización no resulta ser muy diferente de aquella utilizada en el Teorema de Lax-Milgram original, a excepción de ciertos detalles sutiles. La principal diferencia radica en incorporar las condiciones de coercitividad débil para poder demostrar el acotamiento inferior y continuidad de la transformación T que hemos definido en todas las demostraciones, pero que en este caso posee un dominio y contradominio distintos.

Llama la atención el hecho que no fue necesario definir un producto interno especial entre los espacios $\mathcal{H}_1$ y $\mathcal{H}_2$, como ocurrió en la generalización se la sección 2.1, para los espacios $\mathcal{H}_A$ y $\mathcal{H}_{A'}$. Esto se debe a que en este caso, la demostración se hizo de manera que todos los productos internos utilizados se encontraban en sólo uno de los espacios ($\mathcal{H}_2$), por lo que nos bastó con utilizar el producto interno ahí definido. En la generalización de Landesman, la teoría se construyó de manera que se permitiera utilizar esta operación entre espacios diferentes, cosa que después simplemente se aprovecha en la demostración del Teorema de Lax-Milgram generalizado. Sin embargo, se muestra evidente en este punto que una demostración de dicho teorema se podría haber hecho sin necesidad de definir un producto interno entre espacios.

2.3. Una generalización para espacios de Banach

La siguiente generalización que estudiaremos consta realmente de una leve modificación a la generalización de la sección anterior, que permitirá plantear el Teorema de Lax-Milgram en espacios de Banach, a condición de que uno de ellos sea un espacio reflexivo. Precisamente por esta similitud, estas dos generalizaciones tienden a asimilarse (e incluso a confundirse), pues en varios textos donde se presenta el Teorema (2.2.4), como en [3] y [14], se menciona en una nota aparte la posibilidad de generalizar dicho resultado para dos espacios de Banach distintos. Por esta razón, el desarrollo de la generalización que estudiaremos en esta sección es frecuentemente asociado a Jindřich Nečas y a Ivo Babuška. Dicha asociación es tan fuerte que en [5] el teorema que presentaremos a continuación es llamado el *Teorema BNB* (Babuška-Nečas-Banach). La razón del último de estos nombres es que, según el autor, este resultado consiste, desde el punto de vista del análisis funcional, de una reformulación de dos teoremas fundamentales debidos a Banach: el Teorema de la Gráfica Cerrada y el Teorema de la Aplicación Abierta.

Al plantearse un equivalente del Teorema de Lax-Milgram en espacios de Banach, lo primero que llama la atención es que en los casos anteriores el producto interno que poseían los espacios de Hilbert aparecía en repetidas ocasiones en la demostración, cosa que definitivamente tendrá que cambiar para este caso, dada la falta de dicho producto en los espacios de Banach. Como explicamos en el primer capítulo, el Teorema de Representación de Riesz juega un rol central en el de Lax-Milgram en espacios de Hilbert, pero dado que este teorema utiliza el producto interno en su misma formulación, no estará a nuestra disposición. Sin embargo, la demostración alternativa que presentamos en la sección (1.3) nos da una pista de hacia dónde nos tenemos que dirigir: sustituir el producto interno y el Teorema de Riesz por conceptos más generales. Es evidente, entonces, que el caso para espacios de Banach requerirá modificaciones fuertes en la teoría utilizada para demostrarlo.

Presentaremos a continuación una notación utilizada comúnmente en la literatura, que no obstante su similitud con el producto interno (veremos más adelante que en varios momentos desempeñarán funciones similares), no debe ser confundido con éste. Debemos recordar en todo momento que en un espacio de Banach no siempre se puede definir un producto interno compatible con la norma en éste, por lo tanto no podemos recurrir a dicho producto. Entonces, utilizaremos la notación $\langle \cdot, \cdot \rangle$ para simbolizar la acción de una funcional sobre un elemento, es decir, si $x \in \mathcal{B}$ y $x' \in \mathcal{B}^*$ entonces

$$\langle x, x' \rangle_{\mathcal{B},\mathcal{B}^*} := x'(x) \tag{2.11}$$

Como primer uso de esta notación definimos a continuación el concepto del dual de una transformación lineal en espacios de Banach, que nos será de gran ayuda más adelante.

Definición 2.3.1. Sea $T : \mathcal{B}_1 \to \mathcal{B}_2$ una transformación lineal entre dos espacios de Banach. El **dual** $T' : \mathcal{B}_2^* \to \mathcal{B}_1^*$ de T es la transformación que satisface:

$$\langle Tx, y' \rangle_{\mathcal{B}_2,\mathcal{B}_2^*} = \langle x, T'y' \rangle_{\mathcal{B}_1,\mathcal{B}_1^*} \quad \forall\, x \in \mathcal{B}_1,\ y' \in \mathcal{B}_2^*$$

Observación. Nótese cómo la Definición (2.3.1) imita aquella de los operadores adjuntos (1.1.25) en espacios con producto interno. De hecho, no es raro encontrar en la literatura la utilización del término *adjunto* en vez de *dual* para transformaciones en espacios de Banach y la notación T^* en vez de T'. Evitaremos esto aquí para no confundir aquellas transformaciones que son definidas a través del producto interno con aquellas que no lo son.

El siguiente es uno de los principales resultados sobre propiedades de transformaciones lineales en espacios de Banach.

Teorema 2.3.2. (Aplicación Abierta de Banach) *Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ espacios de Banach, $T : \mathcal{B}_1 \to \mathcal{B}_2$ lineal y suprayectiva, y U un conjunto abierto en $\mathcal{B}_1$, entonces $T(U)$ es abierto en $\mathcal{B}_2$.*

Demostración. *Ver [24, pp. 75-77].* □

El siguiente lema es una reformulación del teorema anterior, pero será de gran utilidad en la demostración de la generalización.

Lema 2.3.3. *Sea $T : \mathcal{B}_1 \to \mathcal{B}_2$ una transformación lineal entre dos espacios de Banach. Son equivalentes:*

(i) R(T) es cerrado.

(ii) Existe $\alpha > 0$ tal que

$$\forall w \in R(T), \quad \exists v_w \in \mathcal{B}_1 \qquad \textit{tal que} \qquad Tv_w = w \quad \textit{y} \quad \alpha \|v_w\|_{\mathcal{B}_1} \leq \|w\|_{\mathcal{B}_2}$$

Demostración. Supongamos que se cumple (i). Como $R(T)$ es cerrado en $\mathcal{B}_2$, entonces $R(T)$ es un espacio de Banach. Como consecuencia del Teorema de la Aplicación Abierta (2.3.2) para $T : \mathcal{B}_1 \to R(T)$ y $U = B_{\mathcal{B}_1}(0,1)$ (la bola unitaria en $\mathcal{B}_1$), se tiene que $T(B_{\mathcal{B}_1}(0,1))$ es abierto en $R(T)$. Dado que $0 \in T(B_{\mathcal{B}_1}(0,1))$, entonces existe $\gamma > 0$ tal que $B_{\mathcal{B}_2}(0,\gamma) \subset T(B_{\mathcal{B}_1}(0,1))$. Sea $w \in R(T)$. Dado que $\frac{\gamma}{2}\frac{w}{\|w\|} \in B_{\mathcal{B}_2}(0,\alpha)$ para alguna $\alpha \in \mathbb{R}$, existe $z \in B_{\mathcal{B}_1}(0,1)$ tal que $Tz = \frac{\gamma}{2}\frac{w}{\|w\|}$. En otras palabras, definiendo $v = \frac{2\|w\|}{\gamma} z$, se tiene que $Tv = w$ y $\frac{\gamma}{2}\|v\|_{\mathcal{B}_1} \leq \|w\|_{\mathcal{B}_2}$.

Para demostrar el converso, sea (w_n) una sucesión en $R(T)$ que converge a algún w en $\mathcal{B}_2$. Usando la hipótesis (ii), deducimos que existe una sucesión (v_n) en $\mathcal{B}_1$ tal que $Tv_n = w_n$ y $\alpha\|v_n\|_{\mathcal{B}_1} \leq \|w_n\|_{\mathcal{B}_2}$. Dado que (v_n) es de Cauchy en $\mathcal{B}_1$ y $\mathcal{B}_1$ es completo, v_n converge a un $v \in \mathcal{B}_1$. En virtud de la continuidad de T, Tv_n converge a Tv. Por lo tanto, $w = Tv \in R(T)$, demostrando que $R(T)$ es cerrado. □

Otro de los grandes teoremas debidos a Banach sobre transformaciones lineales es el siguiente.

Teorema 2.3.4. (Gráfica Cerrada de Banach) *Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ espacios de Banach, $T : \mathcal{B}_1 \to \mathcal{B}_2$ tal que $D(T)$ es denso en $\mathcal{B}_1$. Entonces las siguientes proposiciones son equivalentes:*

- *$R(T)$ es cerrado en $\mathcal{B}_2$*
- *$R(T')$ es cerrado en $\mathcal{B}_2^*$*
- $R(T) = N(T')^{\perp} = \{y \in \mathcal{B}_2 \mid \langle y, y' \rangle_{\mathcal{B},\mathcal{B}^*} = 0 \quad \forall y' \in N(T')\}$
- $R(T') = N(T)^{\perp} = \{x' \in \mathcal{B}_1^* \mid \langle x, x' \rangle_{\mathcal{B},\mathcal{B}^*} = 0 \quad \forall x \in N(T)\}$

Demostración. *Ver [24, pp. 205-208].* □

Combinando los teoremas de la Gráfica Cerrada y Aplicación Abierta de Banach (este último a través del Lema (2.3.3)) y los lemas técnicos que utilizamos para el Teorema de Lax-Milgram original, probamos el siguiente lema, que será la pieza central de la demostración de la generalización en cuestión.

Lema 2.3.5. *Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ espacios de Banach, $T : \mathcal{B}_1 \to \mathcal{B}_2^*$ transformación lineal. Entonces T' es inyectiva y T es acotada por debajo si y sólo si T es biyectiva.*

Demostración. Supongamos que T es acotada por debajo, entonces por el Lema (1.2.2) T es inyectiva. Por otra parte, el Lema (1.2.3) implica que $R(T)$ es cerrado en $\mathcal{B}_2^*$. Ahora utilizamos el Teorema de la Gráfica Cerrada (2.3.4) para ver que

$$R(T) = N(T')^{\perp} = \{y' \in \mathcal{B}_2^* \mid y'(y) = 0 \quad \forall y \in N(T') = \{0\}\} = \mathcal{B}_2^*$$

Entonces $R(T) = \mathcal{B}_2^*$, es decir, T es suprayectiva. Esto es, T es biyectiva.

Conversamente, si T es biyectiva entonces la suprayectividad implica que $N(T') = R(T)^{\perp} = \{0\}$, es decir, T' es inyectiva. Por otra parte, como $R(T) = \mathcal{B}_2^*$ es cerrado, utilizamos el lema (2.3.3) para ver que existe $\delta > 0$ tal que:

$$\forall y \in R(T) = \mathcal{B}_2^*, \quad \exists x_y \in \mathcal{B}_1 \qquad \text{tal que} \qquad Tx_y = y \quad \text{y} \quad \delta\|x_y\|_{\mathcal{B}_1} \leq \|y\|_{\mathcal{B}_2^*}$$

pero por la biyectividad de T, podemos expresar esta afirmación en términos de las x_y:

$$\forall x_y \in \mathcal{B}_1 \quad \alpha\|x_y\|_{\mathcal{B}_1} \leq \|Tx_y\|_{\mathcal{B}_2^*}$$

es decir, T es acotada por debajo. □

A partir de las definiciones y resultados anteriores, podemos finalmente dar paso a la generalización del Teorema de Lax-Milgram de esta sección. Supongamos entonces que se tiene una forma bilineal B sobre $\mathcal{B}_1 \times \mathcal{B}_2$, siendo $\mathcal{B}_1$ un espacio de Banach cualquiera y $\mathcal{B}_2$ un espacio de Banach reflexivo. Se tendrá además que F es una funcional sobre $\mathcal{B}_2^*$, el dual del espacio reflexivo, y las siguientes condiciones, que son una reformulación de la coercitividad débil (2.2.1).

$$\inf_{x \in \mathcal{B}_1} \sup_{y \in \mathcal{B}_2} \frac{B(x,y)}{\|x\|_{\mathcal{B}_1} \|y\|_{\mathcal{B}_2}} \geq \delta \tag{2.12}$$

$$\forall y \in \mathcal{B}_2 \quad (\forall x \in \mathcal{B}_1, B(x,y) = 0) \Longrightarrow (y = 0) \tag{2.13}$$

Vemos una vez más que la primera de las condiciones impone una cota inferior sobre el crecimiento del valor de la forma bilineal B, y la segunda de ellas es una de no-degeneración sobre el espacio reflexivo $\mathcal{B}_2$. Estas dos condiciones bastarán para obtener la generalización del Teorema de Lax-Milgram para espacios de Banach.

Teorema 2.3.6. *Sea $\mathcal{B}_1$ un espacio de Banach y $\mathcal{B}_2$ un espacio de Banach reflexivo. Sea además $B : \mathcal{B}_1 \times \mathcal{B}_2 \to \mathbb{R}$ una función bilineal continua. Entonces las condiciones* (2.12) *y* (2.13) *se cumplen si y sólo si para toda $F \in \mathcal{B}_2^*$ existe un único $x_0 \in \mathcal{B}_1$ tal que*

$$B(x_0, y) = F(y) \quad \textit{para todo} \quad y \in \mathcal{B}_2 \tag{2.14}$$

y además:

$$\|x_0\|_{\mathcal{B}_1} \leq \frac{1}{\delta}\|F\|_{\mathcal{B}_2^*}$$

Demostración. 1. Fijando el elemento $x \in \mathcal{B}_1$, $B(x, \cdot)$ es una funcional lineal en $\mathcal{B}_2^*$. Sea $T_x : \mathcal{B}_2' \to \mathcal{B}_2$ dicho funcional, es decir $T_x(y) = B(x, y)$ para todo $y \in \mathcal{B}_2$. Hemos definido entonces una función $T : \mathcal{B}_1 \to \mathcal{B}_2^*$ dada por $x \mapsto T_x$. Usando la notación de (2.11) podemos decir equivalentemente $B(x, y) = \langle y, Tx\rangle_{\mathcal{B}_2,\mathcal{B}_2^*}$.

Ahora, veamos que el dual de T es $T' : \mathcal{B}_2^{**} \to \mathcal{B}_1^*$, pero como $\mathcal{B}_2$ es reflexivo, $\mathcal{B}_2 \cong \mathcal{B}_2^{**}$ y podemos decir entonces que $T' : \mathcal{B}_2 \to \mathcal{B}_1^*$. Tenemos entonces que $\langle y, Tx\rangle_{\mathcal{B}_2,\mathcal{B}_2^*} = \langle T'y, x\rangle_{\mathcal{B}_2,\mathcal{B}_2^*}$, y denotamos análogamente $T'(y) = T_y'$.

2. Demostraremos ahora que las condiciones (2.12) y (2.13) se cumplen si y sólo si T' es inyectiva y T es acotada por debajo. Partiendo de la primera de estas condiciones, observamos las siguientes equivalencias:

$$\begin{aligned}
\inf_{x\in\mathcal{B}_1}\sup_{y\in\mathcal{B}_2}\frac{B(x,y)}{\|x\|_{\mathcal{B}_1}\|y\|_{\mathcal{B}_2}} \geq \delta \quad &\Longleftrightarrow \quad \sup_{y\in\mathcal{B}_2}\frac{B(x,y)}{\|y\|_{\mathcal{B}_2}} \geq \delta\|x\|_{\mathcal{B}_1} \qquad \forall x \in \mathcal{B}_1\\
&\Longleftrightarrow \quad \sup_{y\in\mathcal{B}_2}\left\langle \frac{y}{\|y\|_{\mathcal{B}_2}}, Tx\right\rangle_{\mathcal{B},\mathcal{B}^*} \geq \delta\|x\|_{\mathcal{B}_1} \quad \forall x \in \mathcal{B}_1\\
&\Longleftrightarrow \quad \|Tx\| \geq \delta\|x\|_{\mathcal{B}_1} \qquad \forall x \in \mathcal{B}_1
\end{aligned}$$

que es precisamente el acotamiento por debajo. Ahora, retomando la segunda condición vemos que para todo $y \in \mathcal{B}_2$ se cumplen las siguientes implicaciones lógicas:

$$\begin{aligned}
(\forall x \in \mathcal{B}_1,\ B(x,y) = 0) \Rightarrow (y = 0) \quad &\Longleftrightarrow \quad (\forall x \in \mathcal{B}_1,\ \langle y, Tx\rangle_{\mathcal{B}_2,\mathcal{B}_2'} = 0) \Rightarrow (y = 0)\\
&\Longleftrightarrow \quad (\forall x \in \mathcal{B}_1,\ \langle T'y, x\rangle_{\mathcal{B}_2,\mathcal{B}_2'} = 0) \Rightarrow (y = 0)\\
&\Longleftrightarrow \quad (\forall x \in \mathcal{B}_1,\ T_y'(x) = 0) \Rightarrow (y = 0)
\end{aligned}$$

entonces tenemos que

$$T'(y) = 0 \quad \forall y \in \mathcal{B}_2 \quad \Longrightarrow \quad y = 0$$

es decir $N(T') = \{0\}$, que es simplemente la inyectividad de T'.

3. Por el Lema (2.3.5) la condición de que T' sea inyectiva y T acotada por debajo es equivalente a que T sea biyectiva.

4. Finalmente demostraremos que T es biyectiva si y sólo si se cumple (2.14). Para probar la implicación, supongamos que T es biyectiva y sea $F \in \mathcal{B}_2^*$ arbitrario. Como T es suprayectiva, $F = Tx_0$ para alguna $x_0 \in \mathcal{B}_1$. Más aún, como T es además inyectiva ese x_0 es único. Entonces para un único $x_0 \in \mathcal{B}_1$ se cumple

$$F(y) = \langle y, Tx_0 \rangle_{\mathcal{B},\mathcal{B}^*} = B(x_0, y) \qquad \forall y \in \mathcal{B}_2$$

Conversamente, si se cumple (2.14) se tiene entonces que para todo $F \in \mathcal{B}_2^*$ existe un único $x_0 \in \mathcal{B}_1$ tal que

$$\langle y, Tx_0 \rangle_{\mathcal{B},\mathcal{B}^*} = B(x_0, y) = F \quad \text{para todo} \quad y \in \mathcal{B}_2$$

Es decir, para cada elemento de $\mathcal{B}_2^*$ existe un único elemento en $\mathcal{B}_1$ que es su imagen inversa bajo T. En otras palabras, T es biyectiva. Para probar la segunda parte de la conclusión, vemos que

$$\delta \|x_0\|_{\mathcal{B}_1} \leq \sup_{y \in \mathcal{B}_2} \frac{B(x_0, y)}{\|y\|} = \sup_{y \in \mathcal{B}_2} \frac{F(y)}{\|y\|} = \|F\|_{\mathcal{B}_2^*}$$

Con esto concluimos la demostración. $\square$

Después de observar esta demostración y los resultados paralelos a ella que fueron utilizados, se vuelve evidente que al no tener disponible el producto interno para esta generalización, la teoría necesaria para demostrar el Teorema de Lax-Milgram se reduce a las propiedades básicas de las transformaciones lineales en espacios de Banach. Esto sin duda ocasiona que la demostración se torne mucho más elaborada y extendida que la del teorema original, pero nos permite ver que en esencia dicho teorema es simplemente una reformulación de las propiedades de acotamiento, biyectividad y dualidad de las transformaciones.

Debe hacerse hincapié en el hecho de que este teorema es una implicación en ambos sentidos, cosa que no ocurría para el Teorema de Lax-Milgram original. Esto muestra que, a diferencia de la coercitividad fuerte (ver sección 1.4), las condiciones (2.12) y (2.13) son condiciones necesarias y suficientes para la existencia del elemento x_0 cuya existencia asegura el teorema. Para las aplicaciones esto es de gran importancia, pues muestra que dichas condiciones son óptimas.

Un último aspecto que llama la atención, como mencionamos al definir la notación $\langle x, f \rangle_{\mathcal{B},\mathcal{B}^*}$, es el rol tan parecido que desempeña dicho operador comparado con el producto interno. En varios fragmentos de la demostración, pareciera que este operador de dualidad simplemente hereda ciertas propiedades de este producto. Sin duda, es una de las más sencillas y fieles aproximaciones a un producto interno que se puede definir en un espacio de Banach, al punto que es frecuente encontrar en la literatura abundante teoría que busca imitar las propiedades de uno en el otro.

Para finalizar esta sección, mostramos el siguiente lema, que asegura que el Teorema (2.3.6) en efecto implica el de Lax-Milgram (1.2.4).

Lema 2.3.7. *Supongamos que $\mathcal{B}_1 = \mathcal{B}_2$ es un espacio de Hilbert. Entonces la coercitividad fuerte implica las condiciones* (2.12) *y* (2.13).

Demostración. Supongamos que la forma bilineal B es fuertemente coercitiva. Entonces la condición (2.12) se deduce directamente de

$$\delta\|x\|_{\mathcal{B}_1} \leq \frac{B(x,x)}{\|x\|_{\mathcal{B}_1}} \leq \sup_{y\in\mathcal{B}_2} \frac{B(x,y)}{\|y\|}$$

Por otra parte, sea $y \in \mathcal{B}_2$. Tomando $x = y$ obtenemos

$$\sup_{x\in\mathcal{B}_1} B(x,y) \geq B(x,x) \geq \delta\|y\|^2_{\mathcal{B}_2}$$

por lo tanto, $\sup_{x\in\mathcal{B}_1} = 0$ implica que $y = 0$, probando lo condición (2.13). $\square$

2.4. Lions-Lax-Milgram

Pocos años después de que Lax y Milgram publicaran su teorema, el renombrado matemático francés Jacques-Louis Lions utilizó el concepto de proyección en un subespacio para obtener una ligera generalización de dicho resultado. Publicada originalmente en un libro sobre ecuaciones diferenciales operacionales [13] en 1961, esta generalización es particularmente interesante por la combinación que hace entre los conceptos de coercitividad y proyección ortogonal en un espacio de Hilbert. Debemos notar que la versión formulada originalmente por Lions comprende un espacio de Hilbert $\mathcal{H}$ y un subespacio $\mathcal{V}$ de éste, dotado de una norma $\|\cdot\|_{\mathcal{V}}$ para la cual $\mathcal{V}$ puede ser completo o no. La forma bilineal B se define entonces sobre el producto cartesiano de estos dos espacios.

Lions formuló su teorema de esta manera (teniendo a $\mathcal{V}$ como subespacio de $\mathcal{H}$) retomando la idea del Teorema de Lax-Milgram de tener un producto cartesiano del mismo espacio, previo a la posibilidad de hacerlo en espacios distintos gracias a las generalizaciones de las secciones anteriores (el Teorema de Lions fue publicado un año antes que la generalización de Nečas). La única condición que Lions impone al espacio $\mathcal{V}$ es que la aplicación $v \to v$ sea continua de $\mathcal{H}$ en $\mathcal{V}$, es decir

$$\|v\|_{\mathcal{H}} \leq c_1 \|v\|_{\mathcal{V}} \quad \forall v \in \mathcal{V} \quad \text{para alguna constante} \quad c_1 > 0 \tag{2.15}$$

Esto equivale a decir que $\mathcal{V}$ es continuamente encajado en $\mathcal{H}$. Notamos aquí, como lo hace Lions, que el primero de estos espacios puede o no ser denso en el segundo. Las condiciones anteriores, dice Lions, se cumplen por ejemplo si $\mathcal{V}$ es un subespacio vectorial no cerrado y no denso de $\mathcal{H}$, dotado del producto interno inducido.

Posteriormente, se recurre a las siguientes dos hipótesis sobre la forma bilineal B:

- Para toda $v \in \mathcal{V}$, la transformación $T : x \mapsto B(x, v)$ es continua sobre $\mathcal{H}$ (decimos que B es continua sobre la primera entrada).
- B es fuertemente coercitiva sobre $\mathcal{V}$.

Vemos que, a diferencia del Teorema de Lax-Milgram, en este caso no se pide que la forma bilineal B sea globalmente continua, es decir, $v \to B(x, v)$ para un $x \in \mathcal{H}$ puede no serlo. Si F es una funcional lineal, basta con estas hipótesis para concluir la existencia (aunque no la unicidad) del elemento $x \in \mathcal{H}$ para el cual

$$B(x, v) = F(v) \quad \forall x \in \mathcal{H}$$

Exploraremos más adelante los casos en los cuales sí se puede asegurar la unicidad.

Este teorema se encuentra normalmente en la literatura bajo el nombre de "Teorema de Lions" ó "Teorema Lions-Lax-Milgram", aunque estos nombres son utilizados para designar un abanico de variantes de éste. La versión que presentamos a continuación, tomada de [19], es a su vez una ligera generalización del resultado original de Lions, que permite omitir la hipótesis que $\mathcal{V}$ sea un subespacio de $\mathcal{H}$. Al estar conformado el domino de B de dos espacios que no son iguales ($\mathcal{H}$ y $\mathcal{V}$), se puede deducir que la condición impuesta a la forma bilineal será una variante de la coercitividad débil. Además, utilizaremos el concepto de acotamiento débil, que presentamos a continuación.

Definición 2.4.1. Un conjunto S en un espacio X normado es **débilmente acotado** en X si para cada $f \in X^*$ existe $r > 0$ tal que $\|f(x)\| < r$ para todo $x \in S$.

El siguiente teorema es una caracterización del acotamiento débil en espacios completos.

Teorema 2.4.2. *Un conjunto S en un espacio de Banach $\mathcal{B}$ es débilmente acotado si y sólo si es acotado.*

Demostración. El resultado es una consecuencia directa de los Teoremas de Banach-Alaoglu y Banach-Steinhaus. Para una prueba detallada ver [18, pp. 68-69].

Con esto, estamos listos para dar paso al teorema principal de esta sección. Para la demostración nos basamos en aquella original de Lions [13] y en la variante de Showalter [19].

Teorema 2.4.3. (Lions-Lax-Milgram) *Sea $\mathcal{H}$ un espacio de Hilbert y $\mathcal{V}$ un espacio normado. Supongamos que $B : \mathcal{H} \times \mathcal{V} \to \mathbb{R}$ es bilineal y que $B(\cdot, v)$ está en $\mathcal{H}^*$ para cada $v \in \mathcal{V}$. Entonces los siguientes son equivalentes:*

(i) Para alguna constante $\delta > 0$

$$\inf_{\|v\|_{\mathcal{V}}=1} \sup_{\|x\|_{\mathcal{H}} \leq 1} |B(x, v)| \geq \delta \tag{2.16}$$

(ii) Para cada $F \in \mathcal{V}^*$, existe un elemento $x \in \mathcal{H}$ tal que

$$B(x, v) = F(v) \quad \forall v \in \mathcal{V} \tag{2.17}$$

Demostración. Para demostrar la implicación, definamos primero $T : \mathcal{V} \to \mathcal{H}^*$ a través de $B(x, v) = \langle x, Tv \rangle_{\mathcal{H},\mathcal{H}'} \quad \forall x \in \mathcal{H}$. T es lineal mas no necesariamente continua. La propiedad (2.16) implica que

$$\inf_{\|v\|_{\mathcal{V}}=1} \sup_{\|h\|_{\mathcal{H}} \leq 1} |T_v(h)| \geq \delta$$

equivalentemente

$$\|T(v)\| \geq \delta \|v\| \quad \forall v \in \mathcal{V}$$

es decir, T es inyectiva. Por lo tanto, T es además invertible. Sea entonces $T^{-1} : R(T) \to \mathcal{V}$, transformación que se puede extender por continuidad a $\overline{T^{-1}} : \overline{R(T)} \to \hat{\mathcal{V}}$ lineal, siendo $\hat{\mathcal{V}}$ el "completado" de $\mathcal{V}$. Sea $F \in \mathcal{V}^*$, entonces $x \in \mathcal{H}$ satisface $B(x, v) = F(v)$ para todo $v \in \mathcal{V}$ si $\langle x, g \rangle_{\mathcal{H},\mathcal{H}^*} = F(T^{-1}g)$ para todo $g \in R(T)$. Entonces, si P es la proyección ortogonal de $\mathcal{H}^*$ en el subespacio $\overline{R(T)}$, lo anterior sería equivalente a que

$$\langle x, g \rangle_{\mathcal{H},\mathcal{H}^*} = F(\overline{T^{-1}}Pg) \quad \forall g \in \mathcal{H}^*$$

La composición $\overline{T^{-1}}P : \mathcal{H}^* \to \hat{\mathcal{V}}$ es continua y su dual es $(\overline{T^{-1}}P)^* : \mathcal{V}^* = \hat{\mathcal{V}}^* \to \mathcal{H}^{**} = \mathcal{H}$. Entonces podemos obtener la solución x directamente poniendo $x = (\overline{T^{-1}}P)^* F$.

Para demostrar el converso, supongamos que se cumple (ii) y la condición (i) es falsa. Entonces existe una sucesión $\{v_n\} \subset \mathcal{V}$ tal que

$$\|v_n\| = 1 \quad \text{y además} \quad \sup_{\|x\| \leq 1} |B(x, v_n)| < \frac{1}{n} \quad \text{para} \quad n \geq 1$$

Como se cumple (ii), se tiene que para toda $f \in \mathcal{V}^*$ existe un elemento $x_f \in \mathcal{H}$ tal que

$$|f(nv_n)| = |B(x_f, nv_n)| \leq \|x_f\| \qquad \forall n \geq 1$$

esto significa que $\{nv_n\}$ está débilmente acotada en $\hat{\mathcal{V}}$. Entonces, el teorema (2.4.2) implica que $\{nv_n\}$ está acotada en $\mathcal{V}$. Esto es claramente una contradicción. $\square$

Como en los casos anteriores, el elemento x cuya existencia asegura el teorema está acotado de antemano por la funcional F elegida.

Corolario 2.4.4. *La solución x en el Teorema* (2.4.3) *satisface el estimado a priori* $\|x\|_{\mathcal{H}} \leq \frac{1}{\delta}\|F\|_{\mathcal{V}'}$.

Demostración. La transformación $\overline{T^{-1}}P : \mathcal{H}^* \to \hat{\mathcal{V}}$ tiene norma igual a $\frac{1}{\delta}$, y su dual tiene la misma norma.

Debe notarse en este punto que al no ser B continua en ambas entradas, $R(T)$ no necesariamente es denso en $\mathcal{H}^*$, por lo que no es posible asegurar la unicidad de la solución. La condición necesaria y suficiente para que exista dicha unicidad es en efecto que $R(T)$ sea denso en $\mathcal{H}^*$. Esta falta de unicidad nos lleva a notar que la cota mostrada por el corolario anterior se cumple para *alguna* solución y no para *toda* solución.

Como mencionamos al principio de esta sección, la versión original de la generalización que propuso Lions habla de un subespacio $\mathcal{V}$ de $\mathcal{H}$. De esta manera, el Teorema de Lions original se puede obtener del anterior como un caso particular, tomando $\mathcal{V}$ como un espacio encajado en $\mathcal{H}$ y sustituyendo la condición (2.16) por la coercitividad fuerte. El siguiente corolario muestra esto, además de la constante de acotamiento a priori que se obtiene.

Corolario 2.4.5. *Sean las hipótesis del Teorema* (2.4.3). *Si $\mathcal{V}$ es continuamente encajado en $\mathcal{H}$, es decir, si $\|v\|_{\mathcal{H}} \leq c_1\|v\|_{\mathcal{V}}$ para todo $v \in \mathcal{V}$, y además B es fuertemente coercitiva sobre $\mathcal{V}$ con constante δ_1, entonces el Teorema* (2.4.3) *se cumple con la constante $\delta = \frac{\delta_1}{c_1}$.*

Demostración. Dado que $\mathcal{V}$ es un subconjunto de $\mathcal{H}$, es en particular normado. Además, la coercitividad fuerte cumple en particular la condición (2.16) de coercitividad débil. Entonces se tiene la conclusión del Teorema (2.4.3). Por otra parte, a partir de la coercitividad fuerte vemos que

$$\delta_1\|v\|_{\mathcal{V}}^2 = B(v,v) \leq \langle v, Tv\rangle \leq \|v\|_{\mathcal{H}}\|Tv\|_{\mathcal{H}'} \leq c_1\|v\|_{\mathcal{V}}\|Tv\|_{\mathcal{H}'}$$

por lo que $\frac{\delta_1}{c_1}\|v\|_{\mathcal{V}} \leq \|Tv\|_{\mathcal{H}'}$ $\square$

Si ahora suponemos que se toma el espacio $\mathcal{V}$ como el espacio $\mathcal{H}$ mismo (la función identidad de $\mathcal{H}$ en $\mathcal{H}$ es un encaje continuo), es claro que se obtiene el Teorema de Lax-Milgram como caso particular. En este caso, como la función B será automáticamente continua en $\mathcal{H} \times \mathcal{H}$, $R(T)$ será denso en $\mathcal{H}^*$ y, según lo visto anteriormente, se tendrá efectivamente la unicidad de la solución, como vimos en la versión original del Teorema (1.2.4).

Capítulo 3

Aplicaciones

Desde finales del siglo XVIII, cuando las ecuaciones diferenciales parciales comenzaron a hacer su aparición a manos de grandes matemáticos como Jean le Rond d'Alembert, Daniel Bernoulli y Pierre Simon Laplace, y hasta principios del siglo XX, éstas se formulaban en los espacios de funciones continuas usuales (C^1, C^2). Eventualmente, se observó que estos espacios no eran los ideales para estudiar las soluciones de ecuaciones diferenciales, entre otras razones porque las derivadas de una función en alguno de esos espacios podía ni siquiera estar en algún otro espacio de la misma familia. Para este propósito, se comenzaron a utilizar diversos espacios con propiedades adecuadas para el estudio de dichas soluciones. Una de las familias de espacios más importantes en este sentido son los espacios de Sobolev.

Sergei Sobolev, matemático ruso, revolucionó el campo de las ecuaciones diferenciales introduciendo[1] en la década de los 30 una familia de espacios basados en el concepto de la derivada débil. Éstos constaban de funciones extraídas de algún espacio $L_p(\Omega)$ de Lebesgue, cuyas derivadas parciales también pertenecieran a dicho espacio. Aquellos espacios mostraron ser el ambiente natural para el estudio de las soluciones y en pocos años se habían vuelto el lenguaje universal de las ecuaciones diferenciales parciales.

En el contexto de los espacios de Sobolev, en conjunción con los espacios de Banach y Hilbert, el teorema de Lax-Milgram probó tener su mayor repercusión. Por esto, dedicaremos este capítulo a mostrar algunos ejemplos de aplicaciones de éste en el ámbito de las ecuaciones diferenciales parciales y espacios de Sobolev, repasando antes brevemente la teoría que sustenta dichas aplicaciones. Gran parte de las nociones que definiremos y utilizaremos en este capítulo requieren del conocimiento de distribuciones

[1] Existe cierta controversia sobre el crédito de este descubrimiento: en la literatura suelen mencionarse también los nombres de otros matemáticos, como Fichera, Friedrichs y Leray, que trabajaron con espacios similares. Para más detalles sobre la historia de los espacios de Sobolev ver [22].

y sus propiedades principales, por lo que se presupone cierta familiaridad con estos conceptos.

3.1. Teoría Previa

3.1.1. Formulaciones Débiles

La idea de las formulaciones débiles, proveniente del cálculo de variaciones, es crucial para entender el planteamiento de ecuaciones diferenciales en espacios de funciones. En una formulación débil no se requiere que una ecuación se cumpla de manera absoluta, sino que se cumpla solamente *de cierta forma*, o indirectamente. Usualmente, esta *forma* se da a través de la interacción con “funciones de prueba”; por ejemplo a través del producto interno o de formas bilineales, que relacionen las potenciales soluciones y los elementos de prueba.

Supongamos que x, y son elementos del espacio X, y que T es un operador en X. Entonces la ecuación

$$Tx = y$$

puede ser planteada en forma débil de la siguiente manera

$$\langle Tx, z\rangle = \langle y, z\rangle \qquad \forall z \in X$$

En el contexto de los espacios de funciones, si tenemos un elemento $f \in \mathcal{B}'$ y una transformación $T : \mathcal{B} \to \mathcal{B}'$ y nos interesa encontrar una solución $u \in \mathcal{B}$ a

$$Tu = f$$

el cálculo de variaciones nos dice que esto equivale a encontrar $u \in \mathcal{B}$ tal que para toda $v \in \mathcal{B}$ se cumpla la formulación débil del problema

$$Tu(v) = f(v)$$

Si la solución a la formulación débil existe, se dice que ésta es una *solución débil* del problema. En contraste, a la solución de la formulación original se le llama *solución clásica o fuerte*. Debemos recalcar aquí que la existencia de una solución débil no garantiza la existencia de una clásica, de hecho, se puede dar el caso que la solución encontrada con la formulación débil no esté siquiera definida para la formulación fuerte. Entonces, la existencia de una solución débil tiene dos ventajas: puede ser útil tener *al menos* una forma de satisfacer el problema en *cierta manera* y además la existencia de ésta puede sugerir que *tal vez* exista una solución clásica. Así, vemos que toda solución clásica es también una solución débil, pero el converso no es cierto.

Las formulaciones débiles son de gran importancia en el estudio de ecuaciones diferenciales parciales y ordinarias, pues muchos problemas (aún aquellos provenientes de modelar fenómenos del mundo real) no admiten soluciones suficientemente suaves, por lo que resolverlos de la manera clásica no es posible. Por la facilidad que presentan los métodos para soluciones débiles, a veces es conveniente primero encontrar éstas y posteriormente probar que son suficientemente suaves (y por lo tanto coinciden con las soluciones clásicas). La teoría de las formulaciones y soluciones débiles de ecuaciones diferenciales parciales se sustenta en los conceptos de derivadas distribucionales y derivadas débiles, nociones fundamentales de la teoría de distribuciones.

3.1.2. Distribuciones y Espacios de Sobolev

La teoría de las formulaciones débiles encontró su hogar predilecto en los espacios de Sobolev principalmente porque éstos incorporaban intrínsecamente las nociones de derivadas débiles. Además, tenían la propiedad de que las derivadas de una función en un espacio de Sobolev siempre se encuentran en otro espacio de Sobolev, por lo que resultaban ideales para plantear ecuaciones diferenciales de manera natural. Las primeras nociones de funciones generalizadas aparecieron también en el contexto de los espacios de Sobolev, conceptos que después retomaría Laurent Schwartz algunos años después para construir su renombrada teoría de las distribuciones. Curiosamente, a manera de un sutil *bucle autorreferencial*, esta teoría posteriormente ayudaría a concretar la de los espacios de Sobolev, permitiendo definir derivadas parciales para cualquier función localmente integrable, cuando en un principio se podía hacer sólo para funciones en $L_p(\Omega)$.

Recordemos que se denota por $\mathcal{D}(\Omega)$ al espacio de funciones de prueba en la teoría de distribuciones, que se obtiene de la siguiente manera. Sea un conjunto $\Omega \subset \mathbb{R}^n$ abierto y no vacío. Para cada subconjunto compacto K de Ω, denotamos por $\mathcal{D}_K$ al espacio de todas las funciones infinitamente diferenciables con soporte en K. Así, tomando la unión de estos espacios conforme K abarca todos los subconjuntos compactos de Ω obtenemos el espacio $\mathcal{D}(\Omega)$. Claramente $\mathcal{D}(\Omega)$ es un espacio vectorial, y dotándolo de una topología particular, se puede ver que es además un espacio localmente convexo. Así, una distribución T es una funcional lineal en $\mathcal{D}(\Omega)$, y al espacio de todas las distribuciones se le denota por $\mathcal{D}(\Omega)^*$. Es común también utilizar la notación $\langle T, \varphi \rangle$ en lugar de $T(\phi)$, la aplicación de T a la función de prueba φ.

Como vimos anteriormente, para poder definir los espacios de Sobolev es necesario contar con las nociones de derivadas débiles y distribucionales. Comenzaremos por la más directa de ellas, la de derivada distribucional, para la cual simplemente nos hace falta una definición preliminar.

Definición 3.1.1. Un **multi-índice** α es una n-tupla ordenada $\alpha = (\alpha_1, \dots, \alpha_n)$ de enteros no negativos α_i. Con cada multi-índice α está asociado el operador diferencial

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$$

cuyo orden es

$$|\alpha| = \alpha_1 + \cdots + \alpha_n$$

Definición 3.1.2. Sea T una distribución en $\mathcal{D}(\Omega)$. Se define la **α-derivada distribucional** de T como $\langle D^\alpha T, \varphi\rangle := (-1)^{|\alpha|}\langle T, D^\alpha \varphi\rangle$, para cada $\varphi \in D(\Omega)$.

Antes de dar paso al otro tipo de derivadas, presentamos a continuación una de las familias de espacios más importantes del análisis funcional, que serán de vital importancia en lo sucesivo.

Definición 3.1.3. Denotamos por $L_p(\Omega)$ (y llamamos **el Espacio p de Lebesgue**) al espacio de todas las funciones medibles[2] f definidas en Ω tales que $|f|^p$ es también medible en Ω. $L_p(\Omega)$ es un espacio normado si se define

$$\|f\|_p = \left(\int_\Omega |f|^p dx\right)^{1/p}$$

Observación. Como sucedió para l_2, L_2 es el único espacio de la familia L_p que es un espacio de Hilbert, cuando se define el producto interno

$$\langle f, g\rangle = \left(\int_\Omega f(x)g(x)dx\right)$$

A partir de los espacios de Lebesgue, se define también $L^1_{loc}(\Omega)$, el espacio de funciones localmente integrables sobre Ω, es decir, la funciones cuya integral sobre cualquier subconjunto compacto K de Ω es finita. Habiendo definido este espacio, podemos finalmente definir las derivadas débiles.

Definición 3.1.4. Sea $f \in L^1_{loc}(\Omega)$. Si existe $u \in L^1_{loc}(\Omega)$ tal que

$$\int_\Omega f(x) D^\alpha \varphi(x) dx = (-1)^{|\alpha|} \int_\Omega u(x)\varphi(x)dx$$

[2]Seguimos la convención usual de identificar funciones f y g si $f = g$ *casi en todas partes.*

entonces decimos que u es la **α-derivada débil** de f.

Observación. A partir de las dos definiciones de derivadas generalizadas anteriores es fácil ver que todas las derivadas clásicas son derivadas débiles y todas las derivadas débiles son derivadas distribucionales.

Definición 3.1.5. Sea Ω un conjunto abierto de $\mathbb{R}^n$, y k un entero positivo. Para $1 \leq p < \infty$, denotamos por $\mathcal{W}^{k,p}(\Omega)$ al conjunto de todas las funciones complejas $f(x) = f(x_1, x_2, ..., x_n)$ definidas en Ω tales que f y sus derivadas distribucionales $D^s f$ de orden $|s| = \sum_{j=1}^{n} |s_j| \leq k$ pertenecen todas a $L^p(\Omega)$. $\mathcal{W}^{k,p}(\Omega)$ es el **Espacio de Sobolev** de orden k, p.

Es fácil ver que $\mathcal{W}^{k,p}(\Omega)$ es un espacio vectorial normado con la norma

$$\|f\|_{k,p} = \left(\sum_{|s| \leq k} \int_{\Omega} |D^s f(x)|^p dx \right)^{\frac{1}{p}} \qquad dx = dx_1 dx_2 ... dx_n$$

Además, dado que las derivadas son tomadas de la manera distribucional, el espacio $\mathcal{W}^{k,p}(\Omega)$ es completo, y por lo tanto, de Banach.

En el caso en que $p = 2$, si definimos

$$\langle f, g \rangle_{W^{k,2}(\Omega)} = \sum_{|s| \leq k} \langle D^s f, D^s g \rangle_{L_2(\Omega)}$$

entonces $W^{k,2}(\Omega)$ es un espacio de Hilbert.

Observación. El hecho de ser espacios de Hilbert, además de su estrecha relación con las series de Fourier[3], vuelve a los espacios de Sobolev con $p = 2$ particularmente importantes. Por este motivo, se creó una notación específica para designarlos:

$$H^k := \mathcal{W}^{k,2}$$

Vemos entonces que la norma inducida por el producto interno que hemos definido para H^k cumple

$$\|f\|_{H^k} = \left(\sum_{i=0}^{k} \|D^i f\|_{L_2}^2 \right)^{1/2} = \left(\|f\|_{L_2}^2 + \cdots + \|D^k f\|_{L_2}^2 \right)^{1/2}$$

[3]La teoría que relaciona estos espacios con las series de Fourier es extensa y bien documentada. Véase por ejemplo [22].

Si Γ es la frontera del dominio Ω, entonces la restricción de $f \in H^k(\Omega)$ a Γ es llamada la traza de f y es denotada por Γ_f. Definimos entonces el espacio de funciones en $H^k(\Omega)$ que son límites de sucesiones de funciones infinitamente diferenciables con soporte compacto en Ω:

$$H_0^k(\Gamma) = \left\{ f \in H^k(\Omega) \quad | \quad \Gamma_f = 0 \right\}$$

Es decir, el espacio $H_0^k(\Omega)$ es la cerradura de $D(\Omega)$ en $H^k(\Omega)$. Cuando Ω tiene una frontera regular, $H_0^k(\Omega)$ consta de las funciones en $H^k(\Omega)$ que se anulan en la frontera. Notemos las siguientes propiedades:

- $D(\Omega) \subset H_0^k(\Omega)$
- $H_0^k(\mathbb{R}^n) = H^k(\mathbb{R}^n)$

La siguiente propiedad, conocida como la desigualdad de Poincaré[4], es útil para establecer contenciones entre diferentes espacios de Sobolev.

Teorema 3.1.6. *Sea Ω un subconjunto abierto de $\mathbb{R}^n$. Entonces existe una constante positiva $C = C(\Omega, p)$ tal que*

$$\|v\|_{L^p(\Omega)} \leq C \|\nabla v\|_{L^p(\Omega)}$$

Demostración. Ver [6, p. 265]. □

Como corolario de esta desigualdad tenemos la siguiente propiedad.

Corolario 3.1.7. *Sea Ω un subconjunto abierto de $\mathbb{R}^n$. Entonces existe una constante positiva C_1 tal que para todo $v \in H^1(\Omega)$ se tiene*

$$\|v\|_{H^1(\Omega)} \leq C_1 \|\nabla v\|_{L^2(\Omega)}$$

Demostración. Para cualquier $v \in H^1(\Omega)$ utilizando la desigualdad de Poincaré vemos que

$$\|v\|^2_{H^1(\Omega)} = \|v\|^2_{L^2(\Omega)} + \|\nabla v\|^2_{L^2(\Omega)} \leq (1 + C^2) \|\nabla v\|^2_{L^2(\Omega)} = C_1^2 \|\nabla v\|^2_{L^2(\Omega)}$$

□

[4]Esta desigualdad, que es presentada en diversas variantes, es conocida en ocasiones como la de Friedrichs o de Poincaré-Friedrichs, dado que existen resultados muy similares de ambos matemáticos.

A partir del corolario anterior y la definición de $\|\cdot\|_{H^1(\Omega)}$ vemos que

$$\|\nabla v\|^2_{L^2(\Omega)} \leq \|v\|^2_{L^2(\Omega)} + \|\nabla v\|^2_{L^2(\Omega)} = \|v\|^2_{H^1(\Omega)} \leq C\|\nabla v\|^2_{L^2(\Omega)} \tag{3.1}$$

por lo que las normas $\|v\|_{H^1(\Omega)}$ y $\|v\|_{H^1_0(\Omega)} := \|\nabla v\|_{L^2(\Omega)}$ son equivalentes.

Aunque la teoría de los espacios de Sobolev es de una gran profundidad y amplitud, los conceptos anteriores nos bastarán para poder formular y resolver los problemas presentes en las aplicaciones posteriores. Falta ahora simplemente mencionar algunos conceptos básicos de las ecuaciones diferenciales parciales.

3.1.3. Ecuaciones Diferenciales Parciales

Naturalmente, una de las características más deseables de un problema que involucre ecuaciones diferenciales parciales es que tenga solución. Sin embargo, comúnmente esto no es suficiente y se espera que dicha solución sea única y dependa continuamente de los datos del problema. Si esto sucede, se dice que el problema está *bien condicionado*. En este sentido, el Teorema de Lax-Milgram es una herramienta útil para asegurar el buen condicionamiento de problemas diferenciales.

Recordemos que la forma general de una ecuación diferencial parcial lineal de segundo orden en dos variables independientes es

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} = f(x, y, u, u_x, u_y) \tag{3.2}$$

donde los coeficientes A, B, C pueden depender de x y de y. Esta forma recuerda a aquella forma general de las secciones cónicas en el plano, y la similitud no termina ahí, pues las ecuaciones diferenciales parciales de segundo orden también se suelen catalogar como *parabólicas*, *elípticas* o *hiperbólicas* según su discriminante. Sin embargo, en este caso el discriminante está dado por $B^2 - AC$, por la convención de escribir $2B$ en vez de B en (3.2) . Entonces las ecuaciones son de tipo

- Elípticas si $B^2 - AC < 0$
- Parabólicas si $B^2 - AC = 0$
- Hiperbólicas si $B^2 - AC > 0$

En el caso en que la ecuación diferencial parcial lineal de segundo orden esté en $\mathbb{R}^n$ (i.e. tenga n variables independientes), la clasificación se define en términos de los valores propios de la matriz de coeficientes de las derivadas. Por ejemplo, se dice que la ecuación es elíptica si los valores propios son todos negativos o todos positivos, parabólica si hay sólo uno nulo e hiperbólica si hay uno negativo (positivo) y el resto positivos (negativos).

La clasificación de ecuaciones diferenciales parciales de mayor orden no es tan obvia, y suelen existir diferentes definiciones dependiendo del contexto. Si L es un operador diferencial de orden k de la forma

$$Lu = \sum_{|\alpha|\leq k} a_\alpha(x)\delta^\alpha u$$

donde α es un multi-índice y δ^α es la derivada parcial con respecto a las variables en α, la ecuación $Lu = f$ es usualmente catalogada como elíptica si el operador L lo es, es decir, si para todo $x \in \Omega$ y todo $\xi \in \mathbb{R}^d$ diferente de cero se tiene

$$\sum_{|\alpha|\leq k} a_\alpha(x)\xi^\alpha \neq 0$$

Recordemos que, como vimos en el primer capítulo, la coercitividad fuerte suele ser llamada también X-elipticidad. Este nombre alternativo proviene directamente de este contexto, el de las ecuaciones diferenciales, pues los operadores diferenciales elípticos cumplen una propiedad análoga a la coercitividad fuerte de formas bilineales. De hecho, el operador diferencial L es llamado fuertemente elíptico si existe una constante $c > 0$ tal que

$$\langle [\sigma_L(\xi)](v), v\rangle \geq c\|v\|^2 \qquad \forall\, \|\xi\| = 1,\ \forall v \in \Omega$$

donde $\sigma_L(\xi)$ es el *símbolo principal* de L, es decir,

$$\sigma_L(\xi) = \sum_{|\alpha|\ =k} a_\alpha \xi^\alpha$$

Como veremos en las próximas secciones, el Teorema de Lax-Milgram es utilizado principalmente para asegurar la existencia y unicidad de soluciones para ecuaciones diferenciales de tipo elípticas, y en menor medida, parabólicas. Con esto, nos encontramos finalmente listos para dar paso a la primera aplicación de dicho teorema.

3.2. Ecuación de Poisson

La ecuación de Poisson es probablemente el ejemplo más común de la aplicación del Teorema de Lax-Milgram a las ecuaciones diferenciales parciales, puesto que es, junto con la de Laplace (caso particular de la de Poisson), la ecuación de tipo elíptico por excelencia. Además, esta ecuación fue la razón por la cual Sergei Sobolev definió el espacio $H^1(\Omega)$ en un principio, pues era el espacio natural para resolverla.

La forma general de la ecuación es la siguiente

$$\Delta\varphi = f(\bar{x}) \tag{3.3}$$

donde Δ es el operador de Laplace, es decir

$$\Delta\varphi = \nabla^2\varphi = \sum_{i=1}^{n} \frac{\partial^2\varphi}{\partial x_i^2}$$

Esta ecuación, planteada originalmente por el reconocido matemático francés Siméon-Denis Poisson, es de gran utilidad en la física e ingeniería, particularmente en las áreas de electroestática y mecánica. Por ejemplo, la ecuación de Poisson describe la relación entre el potencial eléctrico φ y la función de densidad de la carga ρ_f. Según la ley de Gauss tenemos

$$\nabla \cdot \mathbf{B} = \rho_f$$

donde $\mathbf{B}$ es la campo de desplazamiento eléctrico. Haciendo algunos supuestos técnicos sobre el medio (que sea *isotrópico*[5] y *homogéneo*[6]), se puede concluir que $\mathbf{B} = \epsilon\mathbf{F}$, donde ϵ es la permitividad del medio y $\mathbf{F}$ el campo eléctrico. Así, la ecuación toma la forma

$$\nabla \cdot \mathbf{F} = \frac{\rho_f}{\epsilon}$$

Además, la Ley de Inducción de Faraday implica que el campo y potencial eléctricos están relacionados a través de $\mathbf{F} = -\nabla\varphi$. Por lo tanto, tenemos que

$$\nabla^2\varphi = \nabla \cdot \nabla\varphi = -\frac{\rho_f}{\epsilon}$$

Tomando $f = -\frac{\rho_f}{\epsilon}$, la relación entre potencial y este múltiplo escalar de la carga corresponde a la ecuación de Poisson en su forma estándar.

Analizaremos la ecuación de Poisson en dos de sus variantes: el problema de Dirichlet y el de Neumann. La diferencia entre estos dos radica en el tipo de condiciones de frontera que se agregan al ecuación original de Poisson; en el primero de ellos se especifican condiciones sobre las soluciones, mientras que en el de Neumann se hace para las derivadas de las soluciones.

[5]Un medio es **isotrópico** si la permitividad y la permeabilidad son uniformes en todas las direcciones, como sucede, por ejemplo, en el espacio libre.

[6]Un medio es **homogéneo** si sus propiedades físicas no varían en diferentes localizaciones.

3.2.1. Condiciones de Frontera de Dirichlet

El Problema de Dirichlet consiste en agregar a una ecuación diferencial parcial condiciones de frontera en las cuales la función u buscada se anula en la frontera del dominio Ω. Buscamos resolver dicha ecuación en un dominio $\Omega \subset \mathbb{R}^d$ con la condición de frontera

$$u(x) = 0 \quad \forall x \in \partial\Omega = \Gamma$$

Este tipo de restricción requiere que escojamos el espacio de soluciones V como $H_0^1(\Omega)$, espacio que incluye intrínsecamente sólo funciones que cumplen la condición de frontera.

En el caso unidimensional con $\Omega = (0,1)$ el problema de Dirichlet toma la forma

$$\begin{cases} u'' = -f & 0 < x < 1 \\ u(0) = u(1) = 0. \end{cases} \tag{3.4}$$

Multiplicando ambos lados por la función de prueba $v(x) \in \mathcal{D}(\Omega)$ e integrando por partes obtenemos

$$\int_0^1 fv dx = -\int_0^1 u''v dx = -vu'|_0^1 + \int_0^1 u'v' dx = \int_0^1 u'v' dx$$

dado que el término $vu'|_0^1$ se anula pues buscamos funciones de prueba en $H_0^1(0,1)$. Esta es la forma variacional del problema (3.4).

Definamos ahora

$$B(u,v) = \int_0^1 u'(x)v'(x)dx \quad , \quad F(v) = \int_0^1 f(x)v(x)dx$$

entonces el problema toma la forma

$$B(u,v) = F(v) \quad \forall v \in H_0^1(0,1)$$

Como podemos ver, esto tiene la forma usual del problema que resuelve el Teorema de Lax-Milgram (1.2.4), puesto que B es claramente bilineal. Sin embargo, antes de poder utilizar este teorema, debemos verificar las condiciones de acotamiento y coercitividad fuerte.

Utilizando la desigualdad CBS para el producto interno $\langle f, g\rangle = \int_0^1 f \cdot g$ vemos que:

$$\begin{aligned} |B(u,v)| \leq \int_0^1 |u'(x)v'(x)| dx &\leq \left(\int_0^1 |u'(x)|^2 dx\right)^{1/2} \left(\int_0^1 |v'(x)|^2 dx\right)^{1/2} \\ &= \|u'\|_{L_2(0,1)} \|v'\|_{L_2(0,1)} \\ &\leq \|u\|_{H^1(0,1)} \|v\|_{H^1(0,1)} \quad \text{(por la definición de } \|\cdot\|_{H^1(0,1)}) \end{aligned}$$

Es decir, B es acotada. Por otra parte, por el corolario de la desigualdad de Poincaré-Friedrichs (3.1.7), existe $C_1 > 0$ tal que:

$$B(v,v) = \int_0^1 |v'(x)|^2 dx = \|v'\|^2_{L_2(0,1)} \geq C_1^{-\frac{1}{2}} \|v\|^2_{H^1(0,1)}$$

Por lo tanto B es también fuertemente coercitiva. Finalmente, vemos que

$$|F(v)| = \left| \int_0^1 f(x)v(x)dx \right| \leq \|f\|_{L_2(0,1)} \|v\|_{L_2(0,1)} \leq \|f\|_{L_2(0,1)} \|v\|_{H^1(0,1)}$$

por lo que podemos acotar F con $\|f\|_{L_2(0,1)}$, suponiendo obviamente que $f \in L_2(0,1)$. Entonces el Teorema de Lax-Milgram asegura que el problema de Dirichlet unidimensional tiene solución única. Más aún, como B es simétrica, la solución corresponde (según lo visto en la sección 1.4) a la minimización de la siguiente funcional

$$J(v) = \frac{1}{2}B(v,v) - F(v) = \frac{1}{2}\int_0^1 u'(x)v'(x)dx - \int_0^1 f(x)v(x)dx$$

Analizaremos ahora el caso multidimensional. En este caso, el problema toma la forma

$$\begin{cases} \Delta\phi = -f & \text{en} \quad \Omega \\ \phi = 0 & \text{en} \quad \Gamma = \partial\Omega \end{cases} \tag{3.5}$$

Transformaremos nuestro problema, como en el caso anterior, en uno variacional. Probando con funciones diferenciales ψ que se anulan en Γ, podemos obtener la siguiente formulación débil del problema original.

$$-\int_\Omega \psi\Delta\phi dx = \int_\Omega f\psi d\Omega$$

La Primera Identidad de Green asegura que

$$\int_\Omega (\psi\Delta\phi + \nabla\phi\cdot\nabla\psi)d\Omega = \int_\Gamma \psi\frac{\partial\phi}{\partial n}d\Gamma \tag{3.6}$$

Sin embargo, dado que $\psi = 0$ sobre Γ, el lado derecho de esta igualdad se anula, por lo que obtenemos

$$\int_\Omega \nabla\phi\cdot\nabla\psi d\Omega = -\int_\Omega \psi\Delta\phi d\Omega = \int_\Omega f\psi d\Omega$$

Entonces si definimos

$$B(\phi,\psi) = \int_\Omega \nabla\phi\cdot\nabla\psi d\Omega \quad , \quad F(\psi) = \int_\Omega f\psi d\Omega$$

nuestro problema se reduce a encontrar $\phi \in H_0^1(\Omega)$ tal que $B(\phi, \psi) = F(\psi) \quad \forall \psi \in H_0^1(\Omega)$. En este caso, para funciones $\psi \in H_0^1(\Omega)$ que se anulan en la frontera tenemos $\|\psi\|_{L_2(\Gamma)} = 0$, y de nuevo por la desigualdad de Poincaré-Friedrichs (3.1.7) tenemos

$$\|\psi\|^2_{H^1(\Omega)} \leq C_1^2 \|\nabla\psi\|^2_{L_2(\Omega)} = C_1^2 B(\psi, \psi)$$

Entonces concluimos que B es fuertemente coercitiva, con constante $\delta = C^{-1/2}$.

Para ver que F es continuo, vemos que una vez más podemos acotarlo con $\|f\|_{L^2(\Omega)}$:

$$|F(\psi)| = \int_\Omega f\psi d\Omega \leq \|f\|_{L_2(\Omega)}\|\psi\|_{L_2(\Omega)} \leq \|f\|_{L_2(\Omega)}\|\psi\|_{H^1(\Omega)}$$

Entonces podemos finalmente concluir que el problema de Dirichlet multidimensional tiene solución única.

3.2.2. Condiciones de Frontera de Neumann

Consideraremos ahora la ecuación de Poisson con un término lineal de absorción y condiciones de frontera de Neumann[7], es decir, sobre la derivada normal de ϕ en la frontera. Utilizaremos la notación

$$\partial_n \phi = \frac{\partial \phi}{\partial n}(x) = \nabla\phi \cdot \mathbf{n}$$

donde n es el vector normal externo a la frontera Γ del dominio acotado $\Omega \subset \mathbb{R}$. Entonces el problema es de la forma

$$\begin{cases} -\Delta\phi + \phi = f & \text{en} \quad \Omega \\ \partial_n u = 0 & \text{en} \quad \Gamma \end{cases} \tag{3.7}$$

En este caso, la forma variacional del problema es

$$\int_\Omega (-\psi\Delta\phi + \psi\phi) d\Omega = \int_\Omega f\psi d\Omega$$

donde tomamos a ϕ y ψ en el espacio usual $H^1(\Omega)$. Utilizando la Identidad de Green (3.6) obtenemos

$$\int_\Omega \nabla\phi \cdot \nabla\psi d\Omega + \int_\Gamma \psi\frac{\partial\phi}{\partial n} d\Gamma + \int_\Omega \phi\psi d\Omega = \int_\Omega f\psi d\Omega$$

[7]Carl Neumann (1832-1925), matemático alemán, es considerado uno de los fundadores de la teoría de ecuaciones integrales. Junto con Alfred Clebsch, fundó la aclamada revista Mathematische Annalen, de la cual sería editor David Hilbert algunos años después.

pero la condición de frontera de Neumann implica que el segundo término del lado izquierdo se anula, por lo que el problema finalmente se convierte en

$$B(\phi, \psi) = \int_\Omega (\nabla\phi \cdot \nabla\psi + \phi\psi) d\Omega = \int_\Omega f\psi d\Omega = F(\psi)$$

Notemos que en este caso la manera en la que hemos definido B coincide con el producto interno en $H^1(\Omega)$, por lo que B es trivialmente acotada y fuertemente coercitiva con constantes $\delta = \gamma = 1$.

Por otra parte, según vimos en el apartado anterior, $F(\psi)$ es continua. Entonces el Teorema de Lax-Milgram garantiza la existencia y unicidad de la solución también en este caso.

3.3. Ecuación de Advección No Homogénea

La advección es un término común en la química y en la ingeniería que se refiere a un mecanismo de transporte de una sustancia por un fluido. Ejemplos comunes de este efecto es el transporte de contaminantes en un río, el calor debido al viento o los cambios en salinidad debidos a las corrientes marinas. En términos concretos, la advección se refiere al efecto de un campo vectorial sobre una función escalar en un punto.

La ecuación diferencial parcial que encarna esta propiedad es una consecuencia directa de conocidas identidades físicas, como la Ley de Gauss y la Ley de Conservación, y tiene la forma

$$\frac{\partial u}{\partial t} + \nabla \cdot (u\beta) = h$$

donde β es el campo vectorial. Si en particular el flujo es incompresible y constante, entonces la ecuación de advección toma la forma

$$\beta \cdot \nabla u = h$$

en donde además se supone que $-\frac{1}{2}\nabla \cdot \beta > \mu_0 > 0$ para alguna constante μ_0.

Para poder definir las condiciones de frontera que utilizaremos, supongamos que Ω es un dominio acotado. La frontera Γ de Ω se puede dividir en los dos siguientes conjuntos

$$\begin{aligned} \Gamma^- &= \{x \in \partial\Omega \quad | \quad (\beta \cdot n)(x) < 0\} \\ \Gamma^+ &= \{x \in \partial\Omega \quad | \quad (\beta \cdot n)(x) \geq 0\} \end{aligned}$$

siendo n el vector normal exterior a $\partial\Omega$.

Con esto, podemos plantear la ecuación de manera formal. Sea $\beta \in [C^1(\overline{\Omega})]^d$ un campo vectorial y consideraremos el problema con condición de frontera no homogénea siguiente

$$\begin{cases} \text{Encontrar } u \text{ tal que} \\ \beta \cdot \nabla u = h \qquad \text{en} \quad \Omega \\ u = g \qquad \text{en} \quad \Gamma^- \end{cases} \tag{3.8}$$

Si definimos

$$B(u,v) := \int_\Omega v(\beta \cdot \nabla u)\,, \qquad \gamma_0(u) := u|_{\Gamma^-}, \qquad f(v) = \int_\Omega hv$$

entonces el problema (3.8) pertenece a una familia general de problemas no homogéneos de la forma

$$\begin{cases} \text{Encontrar } u \in W \text{ tal que} \\ B(u,v) = f(v) \quad \forall v \in V \\ \gamma_0(u) = g \quad \text{en} \quad Z \end{cases} \tag{3.9}$$

en la cual, en este caso, tenemos que

$$W = \{u \in L^2(\Omega) \mid \beta \cdot \nabla u \in L^2(\Omega)\}\,, \qquad V = L^2(\Omega)$$

Para caracterizar el espacio Z, haremos uso de la siguiente definición.

Definición 3.3.1. Sea $P_0(\Omega)$ el conjunto de todos los subconjuntos compactos de Ω. Entonces definimos el espacio de funciones localmente p-integrables sobre Ω para la medida μ como

$$L^p_{loc}(\Omega) = \left\{ f : \Omega \to \mathbb{C} \text{ medible} \;\middle|\; \int_K |f|^p \mu(\Omega) < \infty \quad \forall K \in P_0(\Omega) \right\}$$

Para que la función γ_0 esté bien definida, basta con tomar $Z = L^2_{loc}(\Gamma^-)$ para la medida $|\beta \cdot n|$.

Sabemos que $V = L^2(\Omega)$ es un espacio de Hilbert (y por lo tanto uno de Banach reflexivo). Además W (un subespacio cerrado de V) y Z son también espacios de Banach. Por otra parte, la función $\gamma_0 : W \to Z$ es claramente lineal y continua, pues

$$\|\gamma_0(u)\| = \left\| u|_{\Gamma^-} \right\| \leq \|u\|$$

y además es suprayectiva. Entonces, por el Teorema de la Aplicación Abierta de Banach (2.3.2), existe una constate $c > 0$ tal que

$$\forall g \in Z \quad \exists u_g \in W \text{ tal que } \gamma_0(u_g) = g \quad , \quad \|u_g\|_W \leq c\|g\|_Z$$

Si definimos $W_0 = N(\gamma_0)$, podemos reformular el problema (3.9) de la siguiente manera

$$\begin{cases} \text{Encontrar } \phi = u - u_g \in W_0 \text{ tal que} \\ B(\phi, v) = f(v) - B(u_g, v) \quad \forall v \in V \end{cases} \tag{3.10}$$

Basta con tomar ahora $F(v) = f(v) - B(u_g, v)$ para que el problema (3.10) finalmente tenga la forma usual de las formulaciones variacionales. De esta manera, la funcional F es continua pues

$$\begin{aligned} \|F(v)\| = \|f(v) - B(u_g, v)\| &\leq (\|f\|_{V'} + \|B\|\|u_g\|_W)\,\|v\|_V \\ &\leq (\|f\|_{V'} + c\|B\|\|g\|_Z)\,\|v\|_V \end{aligned} \tag{3.11}$$

Además, la restricción de B a $W_0 \times V$ con $W_0 = N(\gamma_0)$ cumple las condiciones de coercitividad débil (2.12) y (2.13) de la sección 2.3. Para demostrar esto, veamos primero que

$$\begin{aligned} 2\int_\Omega u(\beta\cdot\nabla u)\,d\Omega &+ \int_\Omega (\nabla\cdot\beta)u^2 \\ &= \int_\Omega \nabla u^2\cdot\beta + (\nabla\cdot\beta)u^2 d\Omega = \int_\Omega \nabla\cdot u^2\beta\,d\Omega = \int_\Gamma u^2\beta\cdot\mathbf{n}d\Gamma \end{aligned}$$

Por lo que

$$\int_\Omega u(\beta\cdot\nabla u)\,d\Omega = -\frac{1}{2}\int_\Omega (\nabla\cdot\beta)u^2\,d\Omega + \frac{1}{2}\int_\Gamma u^2\beta\cdot\mathbf{n}d\Gamma$$

De esta manera, se tiene que

$$\begin{aligned} B(u,u) &= -\frac{1}{2}\int_\Omega (\nabla\cdot\beta)u^2 d\Omega + \frac{1}{2}\int_\Gamma u^2\beta\cdot\mathbf{n}d\Gamma \\ &= -\frac{1}{2}\int_\Omega (\nabla\cdot\beta)u^2\,d\Omega + \frac{1}{2}\int_{\Gamma^+} u^2\beta\cdot\mathbf{n}d\Gamma \geq -\frac{1}{2}\int_\Omega (\nabla\cdot\beta)u^2\,d\Omega \geq \mu_0\int_\Omega u^2 d\Omega \end{aligned}$$

Finalmente, vemos que para todo $u \in W_0$ se tiene

$$\sup_{\|v\|=1} |B(u,v)| \geq \left|B\big(u, \frac{u}{\|u\|}\big)\right| \geq \|u\|^{-1}\left|\mu_0\int_\Omega u^2 d\Omega\right| = \mu_0\|u\|_{L^2(\Omega)}$$

es decir, B cumple la primera condición de la coercitividad débil. Por otra parte, para un $v \in V$ arbitrario, si suponemos que $B(u,v) = 0$ para todo $u \in W_0$ y tomamos

$$u = \begin{cases} v & \text{en} \quad \Omega\setminus\Gamma^- \\ 0 & \text{en} \quad \Gamma^- \end{cases}$$

entonces por un argumento análogo al anterior vemos que

$$B(u,v) = 0 \Rightarrow v = 0$$

Concluimos así que el problema (3.10) (y por lo tanto (3.9) también) tiene una solución única, que según vimos antes, está acotada a priori de la siguiente manera

$$\|u\|_W \leq c(\|f\|_{V^*} + \|g\|_Z)$$

En resumen, hemos visto que para resolver el problema no-homogéneo de Dirichlet basta con hacer un cambio de variable para *levantar* la condición de frontera, y posteriormente plantearlo de la manera usual con la generalización (2.3.6) del Teorema de Lax-Milgram.

3.4. Unicidad del Movimiento de Células Endoteliales

La última aplicación del Teorema de Lax-Milgram que analizaremos es una que aunque resulta ser bastante directa en términos matemáticos, es interesante por el contexto en el que se presenta: la fisiología celular. En un artículo reciente [1], Pamuk y Altunaç utilizaron este teorema para demostrar la existencia y unicidad del movimiento de células endoteliales[8] hacia regiones con altos niveles de enzima proteolítica o bajos niveles de fibronectinas.

El movimiento de las células del endotelio se estudia en el contexto de la angiogénesis, proceso fisiológico en el cual se forman nuevos vasos sanguíneos a partir de los preexistentes. Aunque esta es una función natural y necesaria para el crecimiento del organismo y la cicatrización, también juega un rol fundamental en el desarrollo de tumores cancerígenos. Las células endoteliales, actores principales de la angiogénesis, han despertado interés en el mundo de la medicina por el potencial que existe para tratamientos que controlen su multiplicación, y por lo tanto el crecimiento de un tumor.

Para que la angiogénesis tenga lugar, las células endoteliales deben ser estimuladas por un factor externo, proveniente por ejemplo de un tumor. Después de esta estimulación, las células muestran una tendencia de largo plazo hacia la función de densidad de probabilidad de transición asociada a ciertas enzimas y a la fibronectina, un componente principal de la matriz extracelular.

Para modelar este comportamiento, se trabaja en un segmento de l micrones de longitud colocado sobre el eje horizontal y se reescala de manera que esté descrito por el intervalo $0 \leq x \leq 1$.

[8] Este y otros términos fisiológicos utilizados en esta sección son explicados brevemente en el Glosario, al final de este trabajo.

Utilizaremos la siguiente notación para las concentraciones de los diferentes factores sobre la pared capilar, todas ellas en términos de micro moles por litro cúbico:

$$\begin{aligned}
v(x,t) &= \text{factor angiogénico},\\
c(x,t) &= \text{enzima proteolítica},\\
c_a(x,t) &= \text{enzima proteolítica activa},\\
c_i(x,t) &= \text{enzima proteolítica inhibida},\\
i_a(y,t) &= \text{inhibidor de proteasa},\\
f(x,t) &= \text{fibronectina},\\
a(x,t) &= \text{angiostatina},\\
u(x,t) &= \text{densidad de células endoteliales}
\end{aligned}$$

A partir de los procesos químicos que se llevan a cabo en el proceso de la angiogénesis, los autores llegan al siguiente sistema de ecuaciones que describen las relaciones entre los diferentes factores involucrados.

$$\begin{aligned}
\frac{\partial v}{\partial t} &= -\frac{\lambda_1 v}{1+\nu_1 v}\frac{u}{u_0} + v_r(x,t),\\
\frac{\partial c}{\partial t} &= \frac{\lambda_1 v}{1+\nu_1 v}\frac{u}{u_0} - \mu c,\\
\frac{\partial f}{\partial t} &= \frac{4}{T_f} f\left(1-\frac{f}{f_0}\right)\frac{u}{u_0} - \frac{\lambda_3 c_a f}{1+\nu_3 f},\\
\frac{\partial a}{\partial t} &= -\frac{\lambda_2 a}{1+\nu_2 a v}\frac{u}{u_0} + a_r(x,t),\\
\frac{\partial \iota_a}{\partial t} &= \frac{\lambda_2 a u}{1+\nu_2 a} - \frac{\iota_a(x,t)}{T_{rel}},\\
c &= c_i + c_a,\\
c_i &= \nu_e \iota_e c_e
\end{aligned}$$

donde λ_i, ν_i son parámetros cinéticos, u_0, f_0 son constantes de referencia, T_f es el tiempo de generación de la fibronectina, T_{rel} es el tiempo de relajación de la angiostatina, v_r es la tasa a la que el factor de crecimiento está siendo proveído por el tumor y a_r es el término fuente para la angiostatina. Utilizando la teoría de las caminatas aleatorias reforzadas, a partir de lo anterior podemos deducir la siguiente ecuación que describe el movimiento de las células endoteliales:

$$\frac{\partial u}{\partial t} = D\frac{\partial}{\partial x}\left(u\frac{\partial}{\partial x}\left(ln\frac{u}{\hat{\tau}}\right)\right)$$

donde D es el coeficiente de difusión de las células endoteliales y $\hat{\tau}$ es la ya mencionada densidad de probabilidad de transición, que sesga la caminata aleatoria de dichas células. Se sabe que este sesgo se presenta tanto por la enzima proteolítica como por la fibronectina en la lámina basal, por lo que $\hat{\tau} = \hat{\tau}(c_a, f)$. Según los autores, una probabilidad de transición que refleja adecuadamente este efecto es $\hat{\tau}(c_a, f) = \tau_1(c_a)\tau_2(f)$, donde

$$\tau_1(c_a) = \left(\frac{a_1 + c_a}{a_2 + c_a}\right)^{\xi_1}, \quad \tau_2(f) = \left(\frac{b_1 + f}{b_2 + f}\right)^{\xi_2}$$

Las constantes a_i, b_i son valores positivos tales que a_1 y b_2 son cercanas a cero, a_2 y b_1 son mayores que 1, por lo que reflejan una relación positiva de $\hat{\tau}$ con respecto a c_a y negativa con respecto a f. Esto se debe al hecho de que las células endoteliales "prefieren" moverse hacia regiones con alto nivel de enzimas activas o bajo nivel de fibronectina.

Impondremos condiciones de frontera de cero flujo para las células capilares, que en términos matemáticos se expresa como

$$Du\frac{\partial}{\partial x}ln\left(\frac{u}{\hat{\tau}(c_a, f)}\right) = 0 \quad x \in \{0, 1\}$$

y se supone además que $u(x, 0) = 1$, ya que el capilar se encuentra inicialmente en estado de reposo. Una hipótesis adicional es el hecho de que la cantidad de enzima activa y de fibronectina no dependen del tiempo, lo que nos lleva a definir la función $\kappa(x) = \hat{\tau}(c_a(x), f(x))$, independiente de t. Al suponer que $c_a'(0) = c_a'(1) = f'(0) = f'(1)$, tenemos entonces que $\kappa'(0) = \kappa'(1) = 0$. Manipulando la ecuación diferencial e incorporando las condiciones de frontera anteriores llegamos al siguiente problema

$$\begin{gathered} u_t = D\big(u_{xx} - (uL(x))_x\big) \quad (x, t) \in \Omega_T \\ u(x, 0) = 1 \quad x \in (0, 1) \\ u_x(x, t) = 0 \quad \forall t \in (0, T],\ x \in \{0, 1\} \end{gathered}$$

donde $L(x) = \frac{\kappa'(x)}{\kappa(x)}$. En este punto, transformamos el problema en uno independiente del tiempo, pues el objetivo es investigar la solución de estado estacionario. Entonces tenemos

$$\begin{gathered} -\frac{d^2u}{dx^2} + \frac{d}{dx}(uL(x)) = 0 \quad x \in (0, 1), \\ u'(0) = u'(1) = 0 \end{gathered} \tag{3.12}$$

Aunque el problema anterior se puede fácilmente resolver en el caso en que $L(x) \in C^1[0, 1]$, nos interesa investigar una solución generalizada, para el caso en que L no sea necesariamente continua. Entonces, asumiremos simplemente que L' es una función medible en $[0, 1]$ acotada de la siguiente manera

$$0 < M_1 \leq L'(x) \leq M_2$$

Así, transformamos el problema en uno variacional, multiplicando por v e integrando por partes

$$\int_0^1 -u''v + \frac{d}{dx}(uL(x))vdx = -u'v|_0^1 + \int_0^1 u'v'dx + \int_0^1 vu'L + vuL'dx$$

Las condiciones de frontera sobre u' implican que el primero de estos términos es cero. Por otra parte, de la ecuación diferencial original (3.12) se concluye que cualquier solución u debe cumplir

$$-u' + uL(x) = 0$$

Si imponemos esta misma condición a las funciones de prueba v, tenemos que

$$\int_0^1 vu'L + vuL' = \int_0^1 u'v' + vuL'dx$$

por lo que el problema variacional toma finalmente la forma

$$\int_0^1 2u'v' + L'uv\,dx = 0$$

Entonces, si tomamos $\mathcal{H} = H^1(0,1)$ y definimos

$$B(u,v) = \int_0^1 (2u'v' + F'(x)uv)dx$$

y tomamos F como el funcional lineal nulo, entonces el problema es encontrar $u \in H^1(0,1)$ tal que

$$B(u,v) = F(v) \quad \forall v \in H^1(0,1)$$

que es la forma usual del Teorema de Lax-Milgram. Vemos entonces que

$$\begin{aligned} |B(u,v)| = \int_0^1 (2u'v' + L'(x)uv)dx &\leq 2\int_0^1 |u'||v'|dx + M_2 \int_0^1 |u||v|dx \\ &\leq 2\|u'\|_{L_2(0,1)}\|v'\|_{L_2(0,1)} + M_2\|u\|_{L_2(0,1)}\|v\|_{L_2(0,1)} \end{aligned}$$

Utilizando el hecho que $\|u'\|_{L^2(0,1)} \leq \|u\|_{H^1(0,1)}$ y $\|u\|_{L^2(0,1)} \leq \|u\|_{H^1(0,1)}$, la desigualdad anterior nos lleva a

$$|B(u,v)| \leq 2\|u\|_{H^1(0,1)}\|v\|_{H^1(0,1)} + M_2\|u\|_{H^1(0,1)}\|v\|_{H^1(0,1)} \leq \gamma\|u\|_{H^1(0,1)}\|v\|_{H^1(0,1)}$$

con $\gamma = 2\,\text{máx}(2, M_2)$, por lo tanto B es acotada. Por otra parte, utilizamos la desigualdad de Poincaré-Friedrichs para ver que existe $C_1 > 0$ tal que

$$B(u,u) = \int_0^1 \left((u')^2 + L'(x)u^2\right)dx \geq 2\|u'\|^2_{L_2(0,1)} + M_1\|u\|^2_{L_2(0,1)} \geq (2C_1^{-\frac{1}{2}} + M_1)\|u\|^2_{H^1(0,1)}$$

es decir, B es también fuertemente coercitiva. Podemos concluir entonces, por el Teorema de Lax-Milgram (1.2.4) que el problema (3.12) tiene una solución única.

En conclusión, el resultado anterior muestra que la tendencia de las células endoteliales hacia la función de transición de probabilidad es única. En otras palabras, el movimiento de estas células hacia las regiones donde hay una concentración alta de enzimas activas o baja de fibronectina es único.

Conclusiones

La idea original de este trabajo comenzó siendo simplemente la de explicar a fondo el Teorema de Lax-Milgram y presentar algunas de sus aplicaciones. Pronto, dos cosas se volvieron evidentes. Para empezar, la teoría sobre la que se apoya el teorema resultó ser mucho más rica y abundante que sólo un grupo reducido de definiciones y teoremas. Por otra parte, sobre la marcha fue claro que para mostrar realmente la fuerza de este resultado iba a ser necesario explorar sus diversas generalizaciones.

Así, a lo largo de este trabajo se presentó la teoría circundante al Teorema de Lax-Milgram, se le demostró y se estudiaron cuatro de sus generalizaciones. Posteriormente se mostraron algunas aplicaciones del teorema a las ecuaciones diferenciales parciales. Fue particularmente interesante encontrar una aplicación tan concreta y específica como la que se presentó para el contexto de la fisiología celular. Se presentó también una prueba original del Teorema de Lax-Milgram, que a diferencia de la prueba usual basada en el Teorema de Representación de Riesz, se apoya sobre variantes del Teorema de Hahn-Banach. En esta prueba quedó en evidencia la fuerte relación que existe entre estos dos teoremas fundamentales del análisis funcional.

Sin embargo, como en todo trabajo de estas características, fue inevitable cometer omisiones y dejar algunas ideas inexploradas. Fuera de las que se presentaron aquí, existen otras generalizaciones interesantes del Teorema de Lax-Milgram que se podrían investigar. Por otra parte, el paso natural después de las aplicaciones del teorema que se presentaron aquí sería investigar el método de elementos finitos, el cual debe su existencia en gran parte al Teorema de Lax-Milgram. El lector interesado puede consultar [5], [4] ó [21] para una introducción a esta teoría. Finalmente, sería interesante también investigar posibles adaptaciones de la prueba alternativa del Teorema de Lax-Milgram para casos más generales, como aquellos que fueron presentados aquí en las generalizaciones.

A pesar de las omisiones, el objetivo principal de este trabajo no fue mutilado. El autor confía en que la belleza del Teorema de Lax-Milgram, uno de los puentes más apasionantes entre las matemáticas puras y las aplicadas, haya impregnado cada una de estas páginas.

Apéndice A

Transformaciones Lineales

Dedicaremos este apartado a definir conceptos y demostrar propiedades relacionados a las transformaciones lineales y operadores. Como dijimos anteriormente, utilizaremos el término “transformación” para referirnos a funciones de un espacio normado X a otro espacio Y, y utilizaremos el término “operador” solo cuando la función tenga como dominio y contradominio el mismo espacio. Para gran parte de este apéndice nos hemos basado en [20].

Definición A.1. $T : X \to Y$ es una **transformación lineal** si cumple

1. $T(x + y) = Tx + Ty \qquad \forall x, y \in X$
2. $T(\alpha x) = \alpha T x \qquad \forall \alpha \in \mathbb{R}$

Definición A.2. Decimos que I es la **transformación identidad** de X en Y si $Ix = x$ para todo x en X.

Definición A.3. Se denota por $\mathscr{L}(X, Y)$ el **espacio de todas la transformaciones lineales** del espacio X al espacio Y.

Definición A.4. Se denota por $\mathscr{B}(X, Y)$ el **espacio de todas la transformaciones lineales acotadas** del espacio normado X al espacio normado Y.

Teorema A.5. *Para $T \in \mathscr{L}(X, Y)$, donde X, Y son espacios normados, las siguientes condiciones son equivalentes:*

1. $\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} \mid x \neq 0\right\}$

2. $\|T\| = \text{ínf}\{k \mid \|Tx\| \leq k\|x\|\}$

3. $\|T\| = \sup\{\|Tx\| \mid \|x\| \leq 1\}$

4. $\|T\| = \sup\{\|Tx\| \mid \|x\| = 1\}$

Demostración. (*1* $\Leftrightarrow$ *4*):

$$\sup_{x\neq 0}\frac{\|Tx\|}{\|x\|} = \sup_{x\neq 0}\left\|T(\frac{x}{\|x\|})\right\| = \sup_{\|x\|=1}\|Tx\|$$

(*1* $\Leftrightarrow$ *2*): A partir de 1., vemos que

$$\sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\} \geq \frac{\|Tx\|}{\|x\|} \quad \forall x \neq 0$$

Por lo tanto,

$$\|x\| \cdot \sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\} \geq \|Tx\|$$

entonces

$$\text{ínf}\{k \mid \|Tx\| \leq k\|x\|\} \geq \sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\} \tag{A.I}$$

Por otra parte, si $x \neq 0$ y k cumple la relación $\|Tx\| \leq k\|x\|$, entonces $\frac{\|Tx\|}{\|x\|} \leq k$. Esto implica que

$$\sup : x \neq 0 \left\{\frac{\|Tx\|}{\|x\|}\right\} \leq k$$

Como esta relación se cumple para toda k obtenemos que

$$\text{ínf}\{k \mid \|Tx\| \geq k\|x\|\} \geq \sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\} \tag{A.II}$$

Por (A.I) y (A.II), tenemos entonces que

$$\sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\} = \text{ínf}\{k \mid \|Tx\| \leq k\|x\|\}$$

(*1* $\Leftrightarrow$ *3*): Si suponemos que $\|x\| \leq 1$ entonces

$$\|Tx\| \leq \sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\}\|x\| \leq \sup_{x\neq 0}\left\{\frac{\|Tx\|}{\|x\|}\right\}$$

por lo tanto

$$\sup_{\|x\|\leq 1} \|Tx\| \leq \sup_{x\neq 0} \left\{ \frac{\|Tx\|}{\|x\|} \right\} \tag{A.iii}$$

Por otra parte, por las propiedades del supremo tenemos que para cualquier $\epsilon > 0$ existe $x_1 \neq 0$ tal que

$$\sup_{x\neq 0} \left\{ \frac{\|Tx_\epsilon\|}{\|x_\epsilon\|} \right\} - \epsilon < \frac{\|Tx_\epsilon\|}{\|x\|} = \left\| T\left(\frac{x_\epsilon}{\|x_\epsilon\|} \right) \right\| \leq \sup_{\|x\|\leq 1} \{\|Tx\|\}$$

Por lo tanto

$$\sup_{x\neq 0} \left\{ \frac{\|Tx\|}{\|x\|} \right\} \leq \sup_{\|x\|\leq 1} \{\|Tx\|\}\} \tag{A.iv}$$

Combinando (A.iii) y (A.iv) tenemos que

$$\sup_{x\neq 0} \left\{ \frac{\|Tx\|}{\|x\|} \right\} = \sup_{\|x\|\leq 1} \{\|Tx\|\}$$

Con esto concluimos la demostración. □

Teorema A.6. *El espacio $\mathscr{B}(X,Y)$, como definido en* (A.4)*, es un espacio normado. Más aún, es un espacio de Banach si Y lo es.*

Demostración. 1. Para empezar, definiremos las operaciones de suma y multiplicación en $\mathscr{B}(X,Y)$ de la siguiente manera. Para $T, S \in \mathscr{B}(X,Y)$,

- $(T+S)(x) = Tx + Sx$
- $(\alpha T)(x) = \alpha T(x) \quad \alpha \in \mathbb{R}$

Es fácil demostrar que con estas operaciones $\mathscr{B}(X,Y)$ es un espacio vectorial.

2. Demostraremos ahora que

$$\|T\| = \sup \left\{ \frac{\|Tx\|}{\|x\|} \mid x \neq 0 \right\}$$

es una norma en $\mathscr{B}(X,Y)$, por lo que este espacio será normado con cualquiera de las normas equivalentes definidas en el Teorema (A.5). Es claro que $\|T\|$ existe. Como $\sup \left\{ \frac{\|Tx\|}{\|x\|} \mid x \neq 0 \right\}$ es un subconjunto acotado de los reales, su supremo debe existir.

(a) $\|T\| \geq 0$ pues $\|T\|$ es el supremo de números no negativos. Ahora, si $\|T\| = 0$ entonces $Tx = 0$ para todo $x \in X$, por lo que T es la transformación 0. Conversamente, si $T = 0$ entonces $\|Tx\| = \|Ox\| = \|0\|$ y así $\sup \left\{ \frac{\|Tx\|}{\|x\|} \mid x \neq 0 \right\} = 0$. Por lo tanto

$\|T\| = 0$ si y sólo si $T = 0$.

(b) Para demostrar la segunda propiedad de la norma vemos que

$$\begin{aligned}\|\alpha T\| = \sup\left\{\frac{\|\alpha Tx\|}{\|x\|} \mid x \neq 0\right\} &= \sup\left\{\frac{|\alpha|\|Tx\|}{\|x\|} \mid x \neq 0\right\} \\ &= |\alpha| \sup\left\{\frac{\|Tx\|}{\|x\|} \mid x \neq 0\right\} = |\alpha|\|T\|\end{aligned}$$

(c) Por otra parte

$$\begin{aligned}\|T + S\| &= \sup\left\{\frac{\|(T+S)x\|}{\|x\|} \mid x \neq 0\right\} \\ &= \sup\left\{\frac{\|Tx + Sx\|}{\|x\|} \mid x \neq 0\right\} \\ &\leq \sup\left\{\frac{\|Tx\| + \|Sx\|}{\|x\|} \mid x \neq 0\right\} \\ &\leq \sup\left\{\frac{\|Tx\|}{\|x\|} \mid x \neq 0\right\} + \sup\left\{\frac{\|Sx\|}{\|x\|} \mid x \neq 0\right\} = \|T\| + \|S\|\end{aligned}$$

Por lo tanto $\mathscr{B}(X, Y)$ es un espacio normado.

3. Finalmente, demostraremos que $\mathscr{B}(X, Y)$ es un espacio de Banach si Y lo es. Sea (T_n) una sucesión de Cauchy en $\mathscr{B}(X, Y)$. Esto significa que para todo $\epsilon > 0$ existe un número natural N tal que $\|T_n - T_m\| < \epsilon$ para cualesquiera $n, m > N$. Esto implica que para cada elemento $x \in X$ se tiene

$$\|T_n x - T_m x\| \leq \|(T_n - T_m)\|\|x\| < \epsilon\|x\| \quad \forall n, m > N \tag{A.v}$$

esto significa que $(T_n x)$ es un sucesión de Cauchy en Y. Dado que Y es de Banach, el límite de esta sucesión existe, que denotaremos por Tx. Para demostrar que en efecto $T \in \mathscr{B}(X, Y)$ tenemos que probar que es (a) lineal y (b) acotada, y posteriormente hay que verificar que (c) T_n converge a T.

(a) Como T está definida para un $x \in X$ arbitrario, es en efecto una transformación de X en Y. Además:

$$\begin{aligned}T(x+y) &= \lim_{n\to\infty} T_n(x+y) = \lim_{n\to\infty} T_n(x) + T_n(y) = \lim_{n\to\infty} T_n x + \lim_{n\to\infty} T_n y = Tx + Ty \\ T(\alpha x) &= \lim_{n\to\infty} T_n(\alpha x) = \alpha \lim_{n\to\infty} T_n x = \alpha T x\end{aligned}$$

(b) Como las T_n's son acotadas, existe $M > 0$ tal que $\|T_n x\| \leq M$ para toda n. Esto implica que para toda n y toda $x \in X$ se tiene $\|T_n x\| \leq \|T_n\|\|x\| \leq M\|x\|$. Tomando el límite tenemos que

$$\lim_{n\to\infty} \|T_n x\| \leq M\|x\|$$

o equivalentemente, dado que la norma es una función continua, tenemos

$$\|Tx\| \leq M\|x\|$$

(c) Retomando (A.v), si m suficientemente grande se tiene que

$$\|Tx - T_n x\| \leq \epsilon\|x\|$$

Por lo tanto

$$\|T - T_n\| = \sup\left\{\frac{\|Tx - T_n x\|}{\|x\|}\right\} \leq \frac{\epsilon\|x\|}{\|x\|} = \epsilon$$

tomando el límite vemos que $T_n \to T$, por lo que $\mathscr{B}(X, Y)$ es completo, y por lo tanto, de Banach. Con esto concluimos la demostración. □

Teorema A.7. *Sea T un operador lineal acotado sobre un espacio de Hilbert $\mathcal{H}$. Entonces el operador adjunto T^* (como definido en* (1.1.25) *en el Capítulo 1) tiene las siguientes propiedades.*

1. *T^* siempre existe.*
2. *T^* es acotado.*
3. *T^* es único.*

Demostración. 1. Sea $\langle Tx, y\rangle = f_y(x)$. Entonces

$$f_Y(x_1 + x_2) = \langle T(x_1 + x_2), y\rangle = \langle Tx_1, y\rangle + \langle Tx_2, y\rangle = f_y(x_1) + f_y(x_2)$$

Es decir, f_y es lineal. Además, utilizando la desigualdad CBS vemos que

$$|f_y(x)| = |\langle Tx, y\rangle| \leq \|Tx\|\|y\| \leq \|T\|\|x\|\|y\| \leq k\|x\|$$

para toda $y \in \mathcal{H}$ fija. Entonces $f_y \in \mathcal{H}^*$ y por el Teorema de Representación de Riesz, existe $y^* \in \mathcal{H}$ tal que

$$\langle Tx, y\rangle = f_y(x) = \langle x, y^*\rangle \quad \forall x \in \mathcal{H}$$

Por lo tanto, T induce una transformación lineal $y \to y^*$, que escribimos $y^* = T^*y$, donde T^* es definido de $\mathcal{H}$ en sí mismo.

2. Una vez más por la desigualdad CBS tenemos

$$\|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle T(T^*x), x\rangle \leq \|T(T^*x)\|\|x\|$$

Como T es acotado, existe $k > 0$ tal que $\|T(T^*x)\| \leq k\|T^*x\|$. Por lo tanto

$$\|(T^*x)\|^2 \leq k\|x\|\|T^*x\|$$

o equivalentemente $\|(T^*x)\| \leq k\|x\|$, es decir, T^* es acotado.

3. Supongamos que existen T_1^* y T_2^* tales que

$$\langle Tx, y\rangle = \langle x, T_1^* y\rangle \qquad y \qquad \langle Tx, y\rangle = \langle x, T_2^* y\rangle$$

Entonces restando ambas igualdades tenemos que

$$\langle x, T_1^* y\rangle - \langle x, T_2^* y\rangle = 0 \forall x \in \mathcal{H}$$

es decir

$$\langle x, (T_1^* - T_2^*)y\rangle \quad \forall x \in \mathcal{H}$$

Esto implica que $T_1^* y = T_2^* y$ para toda $y \in \mathcal{H}$, o $T_1^* = T_2^*$. Por lo tanto el adjunto de T es único. □

Para el resto de este apartado, supondremos que T es un operador lineal de un espacio de Hilbert $\mathcal{H}$ sobre sí mismo.

Definición A.8. T es llamado **auto-adjunto** si $T = T^*$.

Definición A.9. Un operador auto-adjunto T es llamado **positivo** si $\langle Tx, x\rangle \geq 0$ para toda $x \in X$. Es llamado **estrictamente positivo** si $\langle Tx, x\rangle = 0$ sólo si $x = 0$.

La definición anterior permite establecer un orden parcial entre los operadores auto-adjuntos, diciendo que $T \geq S$ si $T - S$ es un operador positivo.

Definición A.10. T es llamado **normal** si $TT^* = T^*T$.

Definición A.11. T es llamado **unitario** si $TT^* = T^*T = I$, donde I es el operador identidad.

Observación. Es claro a partir de las definiciones que todo operador auto-adjunto es normal y todo operador unitario es normal. Los conversos no son necesariamente verdaderos.

Lema A.12. *Sea T un operador lineal acotado en un espacio de Hilbert $\mathcal{H}$, entonces las siguientes condiciones son equivalentes:*

1. $T^*T = I$
2. $\langle Tx, Ty\rangle = \langle x, y\rangle \quad \forall x, y \in \mathcal{H}$

3. $\|Tx\| = \|x\| \quad \forall x \in \mathcal{H}$

Demostración. (*1* $\Rightarrow$ *2*): Supongamos que $T^*T = I$. Entonces

$$\langle Tx, Ty\rangle = \langle T^*Tx, y\rangle = \langle x, y\rangle \forall x, y \in \mathcal{H}$$

(*2* $\Rightarrow$ *3*): Supongamos ahora que $\langle Tx, Ty\rangle = \langle x, y\rangle \quad \forall x, y \in \mathcal{H}$. Si tomamos $x = y$ vemos que $\|Tx\|^2 = \|x\|^2$, o equivalentemente, $\|Tx\| = \|x\|$.

(*3* $\Rightarrow$ *3*): Si $\|Tx\| = \|x\|$ para toda $x \in \mathcal{H}$, entonces $\|Tx\|^2 = \|x\|^2$ implica que $\langle Tx, Tx\rangle = \langle x, x\rangle$. Entonces:

$$\langle T^*Tx, x\rangle = \langle x, x\rangle \quad \Leftrightarrow \quad \langle (T^*T - I)x, x\rangle = 0 \quad \forall x \in \mathcal{H}$$

Por lo tanto $T^*T - I = 0$, es decir, $T^*T = I$. $\square$

Teorema A.13. *Un operador T acotado sobre un espacio de Hilbert $\mathcal{H}$ es unitario si y sólo si es un isomorfismo isométrico de $\mathcal{H}$ en sí mismo.*

Demostración. Sea T unitario. Por el Lema (A.12) T es un isomorfismo isométrico de $\mathcal{H}$ en sí mismo. Conversamente, si T es un isomorfismo isométrico de $\mathcal{H}$ en sí mismo, entonces existe el inverso T^{-1} y por el mismo lema tenemos que

$$T^*T = I \Rightarrow (T^*T)T^{-1} = IT^{-1} \Rightarrow T^*(TT^{-1}) = IT^{-1}$$

es decir $T^* = T^{-1}$, o equivalentemente $TT^* = TT^{-1} = I$, es decir T es unitario. $\square$

Apéndice B

Teoremas de Hahn-Banach

Dedicaremos este apartado a enunciar y demostrar algunos resultados de la *familia* de Teoremas de Hahn-Banach. Dicha familia consta de varios teoremas para los cuales se utiliza el término "Teorema de Hahn-Banach", todos ellos íntimamente relacionados y provenientes de los resultados obtenidos paralelamente por Hans Hahn y Stephan Banach[9] a finales de la década de los 20. Comenzaremos con algunas definiciones necesarias para este propósito.

Definición B.1. Sea X un espacio vectorial y $\mathbb{K}$ su campo subyacente. Se dice que un conjunto S en X es **equilibrado** si

$$\lambda S \subset S \qquad \forall\, \lambda \in \mathbb{K} \quad \text{tal que} \quad |\lambda| \leq 1$$

donde

$$\lambda S := \{\lambda x \mid x \in S\}$$

Definición B.2. Sea X un espacio vectorial. Se dice que un conjunto S en V es **absorbente** si para todo $x \in X$ existe un número real r tal que:

$$\forall \lambda \in \mathbb{K} \quad |\lambda| \geq r \quad \Longrightarrow \quad x \in \lambda S$$

Definición B.3. Un espacio vectorial topológico X es llamado **localmente convexo** si existe una base local de vecindades U en el origen que consiste de conjuntos convexos, equilibrados y absorbentes.

[9] Aunque Hahn y Banach dieron al teorema que lleva su nombre la forma que conocemos hoy en día, sus orígenes se remontan casi veinte años atrás a trabajos de Eduard Helly y Frigyes Riesz.

Definición B.4. Sea X un espacio vectorial. La función real $p(x)$ es llamada **sub-lineal** en X si para cualesquiera x, y en X y $\alpha \geq 0$ cumple

(i) $p(x+y) \leq p(x) + p(y)$ (es sub-aditiva)

(ii) $p(\alpha x) = \alpha p(x)$

Si cambiamos (ii) por la condición más fuerte $p(\alpha x) = |\alpha| p(x)$ para todo escalar α, la función p es llamada una **semi-norma** en X.

Definición B.5. La funcional

$$p_M(x) = \text{ínf}\{\alpha > 0 \,:\, \alpha^{-1}x \in M, \alpha \in \mathbb{R}\} \tag{B.I}$$

es llamada la **Funcional de Minkowski** para el conjunto convexo, equilibrado y absorbente M.

Teorema B.6. (Hahn-Banach) *Sea X un espacio vectorial real y M un subespacio vectorial de X. Sean además $p(x)$ una función sub-lineal y f_0 una funcional real definida en M que satisface $f_0(x) \leq p(x)$ en M. Entonces existe una funcional lineal F definida en todo X tal que*

(i) F es una extensión de f_0, es decir $F(x) = f_0(x) \quad \forall x \in M$

(ii) $F(x) \leq p(x)$ en X.

Demostración. Supongamos primero que X es generado por M y un elemento $x_0 \notin M$, es decir

$$X = \{x = m + \alpha x_0 \mid m \in M, \alpha \text{ real}\} \tag{B.II}$$

Como $x_0 \notin M$, la representación anterior de $x \in X$ en la forma $x = m + ax_0$ es única. Se sigue que si para cualquier número real c definimos

$$F(x) := F(m + \alpha x_0) = f_0(m) + \alpha c$$

entonces F es una funcional lineal en X que es una extensión de f_0. Debemos escoger c de manera que $F(x) \leq p(x)$, esto es, $f_0(m) + \alpha c \leq p(m + \alpha x_0)$. Esta condición es equivalente a las dos siguientes:

$$f_0\left(\frac{m}{\alpha}\right) + c \leq p\left(\frac{m}{\alpha} + x_0\right) \quad \text{para} \quad \alpha > 0$$
$$f_0\left(\frac{m}{(-\alpha)}\right) - c \leq p\left(\frac{m}{(-\alpha)} - x_0\right) \quad \text{para} \quad \alpha < 0$$

Para satisfacer estas condiciones, debemos escoger c tal que

$$f_0(m') - p(m' - x_0) \leq c \leq p(m'' + x_0) - f_0(m'') \quad \forall\, m', m'' \in M$$

Dicha elección es posible pues como

$$\begin{aligned} f_0(m') + f_0(m'') = f_0(m' + m'') = &\leq p(m' + m'') = p(m' - x_0 + m'' + x_0) \\ &\leq p(m' - x_0) + p(m'' + x_0) \end{aligned}$$

debemos simplemente escoger c entre los números

$$\sup_{m' \in M} [f_0(m') - p(m' - x_0)] \qquad \text{e} \qquad \inf_{m'' \in M} [p(m'' + x_0) - f_0(m'')]$$

Consideremos ahora la familia de todas las extensiones lineales reales g de f_0 para las cuales la desigualdad $g(x) \leq p(x)$ se cumple para todo x en $D(g)$, el dominio de g. Daremos un orden parcial a esta familia definiendo la relación $\succ$, donde $h \succ g$ si h es una extensión de g. Así, el Lema de Zorn (ver [24, p.3]) asegura la existencia de una extensión g lineal maximal de f_0 para la cual la desigualdad $g(x) \leq p(x)$ se cumple para todo x en $D(g)$.

Basta con probar ahora que $D(g)$ coincide con X mismo. Si suponemos lo contrario, y dejamos que $D(g)$ juegue el rol de M y g el de f_0 en el razonamiento anterior, obtenemos una extensión propia F de g que satisface $F(x) \leq p(x)$ para todo $x \in D(F)$. Esto contradice la maximalidad de la extensión lineal g. $\square$

La siguiente generalización del Teorema de Hahn-Banach para espacios vectoriales complejos se debe a H. F. Bohnenblust y A. Sobczyk.

Teorema B.7. *Sea X un espacio vectorial complejo y p una semi-norma definida en X. Sea M un subespacio de X y f una funcional lineal compleja definida en M tal que $|f(x)| \leq p(x)$ en M. Entonces existe una funcional compleja F definida en X tal que*

(i) F es una extensión de f

(ii) $|F(x)| \leq p(x)$ en X.

Demostración. Observemos primero que un espacio vectorial complejo es también un espacio real si la multiplicación escalar se restringe a los números reales. Si $f(x) = g(x) + ih(x)$, donde $g(x)$ y $h(x)$ son respectivamente las partes real e imaginaria de $f(x)$, entonces g y h son funcionales lineales reales definidas sobre M. Así

$$|g(x)| \leq |f(x)| \leq p(x) \qquad \text{y} \qquad |h(x)| \leq |f(x)| \leq p(x) \qquad \text{en M}$$

Por otra parte, como

$$g(ix) + ih(ix) = f(ix) = if(x) = i(g(x) + ih(x)) = -h(x) + ig(x) \quad \forall x \in M$$

equiparado partes reales e imaginarias de cada lado podemos concluir que

$$h(x) = -ig(x) \quad \forall x \in M$$

Ahora, por el teorema anterior podemos extender g a una funcional lineal G definida sobre X con la propiedad de $G(x) \leq p(x)$ en X. Por lo tanto, $-G(x) = G(-x) \leq p(-x) = p(x)$, y así, $|G(x)| \leq p(x)$. Definimos

$$F(x) := G(x) - iG(ix)$$

F es claramente una funcional lineal compleja en X pues

$$F(ix) = G(ix) - iG(-x) = iG(x) + G(ix) = i[G(x) - iG(x)] = iF(x)$$

Además F es una extensión de f, pues si $x \in M$ entonces

$$F(x) = G(x) - iG(ix) = g(x) - ig(ix) = g(x) + ih(x) = f(x)$$

Finalmente, para probar que $|F(x)| \leq p(x)$, escribimos $F(x) = re^{-i\theta}$ de manera que $|F(x)| = e^{i\theta}F(x) = F(e^{i\theta}x)$ es real y positivo, por lo que $|F(x)| = |G(e^{i\theta}x)| \leq p(e^{i\theta}x) = |e^{i\theta}|p(x) = p(x)$. Concluimos así la demostración. □

Teorema B.8. *Sea X un espacio vectorial topológico real o complejo, x_0 un punto en X y $p(x)$ una semi-norma continua en X. Entonces existe una funcional continua F en X tal que $F(x_0) = p(x_0)$ y además $|F(x)| \leq p(x)$ en X.*

Demostración. Sea M el conjunto de todos los elementos αx_0, y definamos f en M a través de $f(\alpha x_0) = \alpha p(x_0)$. Entonces f es lineal en M y $|f(\alpha x_0)| = |\alpha p(x_0)| = |p(\alpha x_0)|$ en M. Por lo tanto existe, en virtud del Teorema (B.7), una extensión F de f tal que $|F(x)| \leq p(x)$ en X. Así, F es continua en $x = 0$ con respecto a $p(x)$, que es continua, y por la linealidad de F, $F(x)$ es continua en todo punto de X. □

Corolario B.9. *Sea X un espacio localmente convexo y sea $x_0 \neq 0$ un elemento de X. Entonces existe una funcional lineal f_0 en X tal que*

$$f_0(x_0) \neq 0 \quad y \quad |f_0(x)| \leq p(x) \quad \forall x \in X$$

Demostración. Como $x_0 \neq 0$, debe existir[10] alguna semi-norma continua p en X para la cual $p(x_0) \neq 0$. Entonces, por el teorema anterior, existe una funcional lineal f_0 tal que $f_0(x_0) = p(x_0) \neq 0$ y $|f_0(x)| \leq p(x)$. □

[10]Esto se debe a que en todo espacio localmente convexo existe una familia $\mathcal{P}$ de seminormas continuas que *separa puntos*, es decir, tales que si $x_1 \neq x_2$, existe $p \in \mathcal{P}$ tal que $p(x_1) \neq p(x_2)$.

El siguiente resultado es una ligera variación del Teorema de Hahn-Banach obtenida por el matemático polaco Stanisław Mazur. Aunque usualmente este teorema se enuncia para espacios vectoriales topológicos localmente convexos, bastará para nuestros propósitos presentarlo para el caso particular de los espacios de Banach.

Teorema B.10. (Mazur) *Sea X un espacio de Banach y M un subconjunto equilibrado, cerrado y convexo de X. Entonces para cualquier $x_0 \notin M$ existe una funcional lineal continua $f_0 \in X^*$ tal que $f_0(x_0) > 1$ y $|f_0(x)| \leq 1$ en M.*

Demostración. Como M es cerrado, existe una vecindad convexa y equilibrada V de 0 tal que

$$M \cap (x_0 + V) = \emptyset$$

Al ser V equilibrada y convexa, tenemos

$$\left(M + \frac{V}{2}\right) \cap \left(x_0 + \frac{V}{2}\right) = \emptyset$$

Siendo el conjunto $\left(x_0 + \frac{V}{2}\right)$ una vecindad de x_0, la cerradura U de $\left(M + \frac{V}{2}\right)$ no contiene a x_0. Como $0 \in M$, el conjunto convexo equilibrado U es una vecindad de 0, pues U contiene a $\frac{V}{2}$ como subconjunto.

Sea p la Funcional de Minkowski de U. Como U es cerrado, tenemos que para cualquier $x_0 \notin U$, $p(x_0) > 1$ y por el contrario $p(x) \leq 1$, si $x \in U$. Entonces existe, por el Corolario (B.9), una funcional lineal continua f_0 en X tal que $f_0(x_0) = p(x_0) > 1$ y $|f_0(x)| \leq p(x)$ en X. □

Observación. El Teorema (1.3.1), utilizado para la demostración alternativa del Teorema de Lax-Milgram presentada en la sección 1.3, es realmente un corolario del resultado anterior. La demostración es bastante directa; si M es un subespacio vectorial cerrado de X (en particular convexo y equilibrado), la conclusión del teorema implica que $f_0(M)$ debe ser un subespacio propio del campo escalar, es decir, debe ser $\{0\}$. Por lo tanto, en ese caso se concluye que $f_0(x) = 0$ para todo x en M.

Glosario de Términos Fisiológicos

angiogénesis Proceso fisiológico que consiste en la formación de nuevos vasos sanguíneos, formados, en una fenómeno estrictamente regulado, a partir de los vasos preexistentes. La angiogénesis da inicio cuando los factores angiogénicos activan receptores presentes en las células endoteliales. Como consecuencia, estas células producen enzimas proteolíticas que degradan la membrana basal, permitiendo la liberación de las células hacia otras regiones.

angiostatina Proteína inhibidora de la angiogénesis; bloquea el crecimiento de nuevos vasos sanguíneos. Es actualmente utilizada en ensayos clínicos para su uso en terapias contra el cáncer.

endotelio Delgada capa de células que recubre el interior de los vasos sanguíneos, formando una interfase entre la sangre fluyendo en el lumen y el resto del vaso sanguíneo. Estas células, llamadas células endoteliales, tienen varias funciones, entre ellas la de regular el paso de materiales desde/hacia el flujo sanguíneo.

enzima proteolítica También conocida como *proteasa*, interviene en la digestión molecular y la descomposición de proteínas en productos más sencillos, debido a que rompe los enlaces peptídicos que unen a los aminoácidos.

factor angiogénico Cualquiera de las substancias (en su mayoría polipéptidos) que fomentan la angiogénesis, al estimular la producción de proteasas en las células endoteliales.

fibronectina Glicoproteína adhesiva presente en la matriz extracelular de la mayoría de los tejidos celulares animales. Incrementa la coagulación de la sangre, la cicatrización y la fagocitosis. La fibronectina no solo juega un papel importante en la adhesión de las células a la matriz sino que también actúa como guía de las migraciones celulares que tiene lugar en los embriones de los vertebrados.

matriz extracelular Conjunto de moléculas, proteínas y carbohidratos sintetizados y secretados por la misma célula al medio intercelular. Aporta propiedades mecánicas a los tejidos, mantiene la forma celular, provee adhesión a las células para formar tejidos y permite la comunicación intercelular.

Bibliografía

[1] E. Altunaç and S. Pamuk, *On the qualitative analysis of the uniqueness of the movement of endothelial cells*, Turk J. Math., 34 (2010), pp. 367–375.

[2] I. Babuška, *Error-bounds for the finite element method*, Numer. Math., 16 (1971), pp. 322–333.

[3] I. Babuška and A. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. Aziz, ed., Academic Press, New York,, 1972, pp. 3–359.

[4] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[5] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Springer, 2004.

[6] L. C. Evans, *Partial differential equations*, Americal Mathematical Society, second ed., 1998.

[7] M. Fréchet, *Sur les ensembles de fonctions et les opérations linéaires*, C. R. Acad. Sci. Paris, 144 (1907), pp. 1414–1416.

[8] P. Jordan and J. von Neumann, *On inner products in linear metric spaces*, Annals of Mathematics, (1935), pp. 719–723.

[9] M. A. Khamsi and W. A. Kirk, *An Introduction to Metric Spaces and Fixed Point Theory*, Pure and Applied Mathematics, John Wiley and Sons, 2001.

[10] E. M. Landesman, *Hilbert-space methods in elliptic partial differential equations*, Pacific Journal of Mathematics, 21 (1967), pp. 113–131.

[11] ——, *A generalized Lax-Milgram theorem*, in Proceedings of the American Mathematical Society, vol. 19, American Mathematical Society, April 1968, pp. 339–344.

[12] P. Lax and A. Milgram, *Parabolic equations*, in Contributions to the theory of partial differential equations, no. 33, Princeton University Press, 1954, pp. 167–190.

[13] J. L. Lions, *Equations différentielles opérationnelles et problèmes aux limites*, vol. 111 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1961.

[14] J. Nečas, *Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la varationelle*, Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 3era serie, 16 (1962), pp. 305–326.

[15] F. Riesz, *Sur une espèce de géométrie analytique des systèmes de fonctions sommables*, C. R. Acad. Sci. Paris, 144 (1907), pp. 1409–1411.

[16] F. Riesz, *Zur theorie des hilbertschen raumes*, Acta Sci. Math., (1934), pp. 34–38.

[17] I. Roşca, *Babuška–Lax–Milgram theorem*, in Encyclopaedia of Mathematics, M. Hazewinkel, ed., Kluwer Academic Publishers, 2001.

[18] W. Rudin, *Functional Analysis*, McGraw-Hill, 1973.

[19] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, vol. 49 of Mathematical Surveys and Monographs, American Mathematical Society, 1997.

[20] A. H. Siddiqi, *Applied Functional Analysis: Numerical Methods, Wavelet Methods and Image Processing*, Marcel Dekker, Inc., 2004.

[21] P. Šolín, *Partial differential equations and the finite element method*, John Wiley and Sons, 2006.

[22] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Springer, 2007.

[23] J. von Neumann, *Allgemeine eigenwerttheorie hermitescher funktionaloperatoren*, Mathematische Annalen, (1929), pp. 49–131.

[24] K. Yosida, *Functional Analysis*, Springer-Verlag, fourth ed., 1974.