Lineal vs no lineal en espacios normados

Juan Guerrero Viu
Trabajo de fin de grado de Matemáticas
Universidad de Zaragoza

Director: Luis Carlos García Lirola
12 de Junio de 2024

Abstract

When we study vector spaces, the functions that arouse the most interest are those that respect the intrinsic structure of these spaces, i.e., linear functions. Additionally, it is easier to work with this type of functions since they have good properties. In the context of normed spaces, which are vector spaces where a norm has been defined, we must consider both the algebraic structure and the metric structure. For the latter, we are interested in functions that preserve the norm, that is, isometries. Therefore, to satisfy both requirements, it is natural to seek linear isometries.

In this sense, once we have found isometries, we would be interested in knowing under which conditions we can ensure that they are indeed linear. Or perhaps, if they are not, be able to construct certain linear isometries from them.

This work focuses on this task, specifically studying two very promising theorems in this regard: the Mazur-Ulam Theorem (1932) and, much more recently, the Godefroy-Kalton Theorem (2003). Thanks to these two results, we are able to obtain sufficient conditions to ensure the existence of linear isometries between normed spaces. In this way, we reflect how the vectorial structure and the metric structure are not independent but have a strong interaction with each other. Indeed, we see that imposing a function to preserve the metric is almost enough for it to respect both structures simultaneously.

The work is divided into 3 chapters.

In the first chapter, we recall a series of main concepts that have been studied in different subjects of the Bachelor Degree. On the one hand, we study notions of general topology such as density, separability, or the definition of homeomorphism. In particular, we focus on the topology of metric spaces, where we talk about convergence and completeness, and finally, we demonstrate the Baire Theorem (1899), stating that in a complete space any countable intersection of dense open sets is also dense.

On the other hand, we introduce concepts from Functional Analysis such as norms, Banach spaces, Lipschitz functions, or isometries, and we provide some properties about them. Additionally, we define the dual space and culminate the section by stating the Hahn-Banach Theorem, one of the fundamental principles of this discipline and one of the most useful tools throughout this work.

The second chapter focuses on the Mazur-Ulam Theorem, which guarantees that any surjective isometry between normed spaces is affine. It was proven in the 1930s, but we study the proof made by Bogdan Nica in 2012, as it is an ingenious and very concise proof for this classic theorem. Next, we give a second alternative proof, slightly longer but leading to a small modification of this result, in a theorem proven by Baker in 1971. In this theorem, the condition of surjectivity is replaced by another one on the target space, specifically, the condition of being strictly convex. We explain what this property means and give some examples of such spaces.

Finally, we also illustrate how neither of these two theorems are generally true if we remove the conditions of surjectivity or strict convexity, and that they also fail when extended to the complex case.

The third chapter is dedicated to formalize the proof of the Godefroy-Kalton Theorem, which states that if there exists a (non-linear) isometry between two Banach spaces, where the source space is also separable, then we can construct a linear isometry between them. To do this, we first give certain notions about convex functions. This type of functions have a very good relationship with respect to continuity and differentiability and are of vital importance to us since one of the examples of naturally convex functions are indeed norms.

Secondly, we introduce the concept of weak* topology, which is defined in the dual of a normed space and satisfies as one of its fundamental properties that the unit ball is compact with this topology, regardless of the dual space. This is known as the Alaoglu Theorem, and its proof is detailed in Appendix B. Thanks to the construction of this topology, we can prove a 1968 theorem attributed to Figiel, where the existence of a linear and continuous function serving as an «inverse» for each isometry defined in a separable Banach space is established. It is a theorem of great interest to us and indeed implies the truth of the Mazur-Ulam Theorem.

In the last section of this chapter, we address the Bochner integral, which is defined with the aim of formalizing the integral of a function that takes values in any Banach space. We study the necessary properties to make clear that the Bochner integral behaves just as expected. The proof of these properties is described in Appendix B. To conclude, we explain a constructive proof of the Godefroy-Kalton Theorem. That is, we explicitly calculate how the linear isometry should be, in terms of the original isometry, whose existence is guaranteed by this theorem. We conclude the chapter by showing how the theorem does not hold for the non-separable case.

As a final point, and to present a possible line of future research, we include an introduction to Lipschitz and Lipschitz-free spaces in Appendix A. The latter arises with the idea of studying Lipschitz functions between metric spaces through the analysis of linear functions defined in the Lipschitz-free space. This can be done because, as we see in this section, there is a correspondence between such functions. At the cost of losing some information about the source space, we gain the ease of working with linear functions. This is currently one of the most active topics of research in the field of Functional Analysis and has a close relationship with what we have seen in this work.

In conclusion, we note that the objective of this work has been to analyze, understand, and detail meticulously all those elements that are taken for granted in certain texts, trying to make the papers we based on more comprehensible and accesible for a wider audience. Specifically, a large part of these efforts has been dedicated to the paper [6] and the proof of Theorem 3.14 given in [12], where it has been necessary to adapt and break down most of the argumentation presented there, making it clearer for the reader. Similarly, we have studied the necessary tools for the presented proofs, such as weak* topology, the Bochner integral, or certain more advanced properties of convex functions than those appearing in the Bachelor Degree. However, all this knowledge does not end here, but serves to broaden the range of instruments we have to tackle future projects.

Resumen

Cuando estudiamos espacios vectoriales, las funciones que más interés despiertan son aquellas que respetan la estructura intrínseca de estos espacios, es decir las funciones lineales. Además, con este tipo de aplicaciones resulta más sencillo trabajar ya que tienen buenas propiedades. En el ámbito de los espacios normados, que son espacios vectoriales donde hemos definido una norma, tenemos que considerar tanto la estructura algebraica como la estructura métrica. Para la segunda, nos interesarán las aplicaciones que preservan la norma, es decir, las isometrías. Por lo tanto, para satisfacer ambas pretensiones, lo natural será ir en búsqueda de isometrías lineales.

En este sentido, una vez hemos encontrado isometrías, nos interesaría conocer bajo qué condiciones podemos asegurar que son de hecho lineales. O tal vez, si no lo fueran, ser capaces de construir ciertas isometrías que sí sean lineales, a partir de ellas.

Este trabajo centrará gran parte de sus esfuerzos en esta tarea, en concreto se estudiarán dos teoremas muy prometedores en ese aspecto. Estos son, el Teorema de Mazur-Ulam (1932) y también, mucho más reciente, el Teorema de Godefroy-Kalton (2003). Gracias a estos dos resultados, somos capaces de obtener condiciones suficientes bastante generales para asegurar la existencia de isometrías lineales entre espacios normados. De esta manera, reflejamos como la estructura vectorial y la estructura métrica no son independientes sino que tienen una fuerte interacción entre sí. En efecto, veremos que el hecho de imponer a una función que preserve la métrica, es casi suficiente para que respete las dos estructuras simultáneamente.

El trabajo estará dividido en 3 capítulos.

En el primer capítulo, recordaremos una serie de conceptos que han sido vistos en distintas asignaturas del Grado. Por un lado, estudiaremos nociones de topología general como densidad, separabilidad o la definición de homeomorfismo. En particular, nos centraremos en topología de espacios métricos, donde hablaremos sobre convergencia y completitud, y finalmente demostraremos el Teorema de Baire (1899) que afirma que en un espacio completo cualquier intersección numerable de abiertos densos, es también densa.

Por otro lado, introduciremos conceptos de Análisis Funcional como normas, espacios de Banach, funciones Lipschitz o isometrías y daremos alguna propiedad sobre ellas. Además, definiremos el espacio dual y culminaremos la sección enunciando el Teorema de Hahn-Banach, que es uno de los principios fundamentales de esta disciplina y una de las herramientas más útiles a lo largo de este trabajo.

Por su parte, el segundo capítulo se centrará en el Teorema de Mazur-Ulam, que garantiza que cualquier isometría sobreyectiva entre espacios normados es afín. Fue probado en los años 30 pero nosotros estudiaremos la demostración que realizó Bogdan Nica en el año 2012, ya que es una prueba ingeniosa y muy concisa para este clásico teorema. A continuación, daremos una segunda demostración alternativa, algo más extensa pero que conduce a una pequeña modificación de este resultado, en un teorema probado por Baker en 1971. En él, se reemplaza la condición de sobreyectividad por otra condición sobre el espacio de llegada, en concreto, la de ser estrictamente convexo. Veremos qué significa esta propiedad y daremos algunos ejemplos de espacios de este tipo.

Por último, también ilustraremos como ninguno de estos dos teoremas son ciertos en general si suprimimos las condiciones de sobreyectividad o de convexidad estricta, y que también fallan al extenderlos al caso complejo.

Finalmente, el tercer capítulo estará dedicado a formalizar la demostración del Teorema de Godefroy-Kalton, que afirma que si existe una isometría (no lineal) entre dos espacios de Banach, donde el de partida es además separable, entonces podemos construir una isometría lineal entre ellos. Para ello, daremos en primer lugar ciertas nociones sobre funciones convexas. Este tipo de funciones tienen muy buena relación con respecto a la continuidad y diferenciabilidad, y son de vital importancia para nosotros ya que uno de los ejemplos de funciones convexas por naturaleza son en efecto las normas.

En segundo lugar, introduciremos el concepto de topología débil*, que se define en el dual de un espacio normado, y que tiene como una de sus propiedades fundamentales que la bola unidad es compacta con dicha topología, sea cual sea el espacio dual. Esto se conoce como Teorema de Alaoglu, y su demostración la detallamos en el Apéndice B. Gracias a la construcción de esta topología, podemos demostrar un teorema de 1968 que se atribuye a Figiel, donde se establece la existencia de una función lineal y continua que sirve de «inversa» para cada isometría definida en un espacio de Banach separable. Es un teorema de gran interés para nosotros y que de hecho implica la veracidad del Teorema de Mazur-Ulam.

En la última sección de este capítulo, abordaremos la integral de Bochner, que se define con el objetivo de formalizar la integral de una función que toma valores en un espacio de Banach cualquiera. Estudiaremos algunas propiedades que necesitamos y que dejan claro que la integral de Bochner se comporta justamente como cabría esperar. La demostración de estas propiedades se describe en el Apéndice B. Para finalizar, probaremos el Teorema de Godefroy-Kalton y lo haremos de una manera constructiva. Es decir, calcularemos explícitamente como debe ser la isometría lineal, en términos de la isometría original, cuya existencia garantiza este teorema. Concluiremos el capítulo, viendo como el teorema no es cierto en el caso no separable.

Como último punto, y a modo de presentar una posible línea de investigación futura, incluimos en el Apéndice A una introducción a los espacios Lipschitz y Lipschitz-libres. Estos últimos surgen con la idea de tratar el estudio de las funciones Lipschitz entre espacios métricos, a través del análisis de funciones lineales definidas en el espacio Lipschitz-libre. Esto se puede realizar porque como vemos en esta sección existe una correspondencia entre dichas funciones. A costa de perder algo de información sobre el espacio de partida, ganamos la facilidad de trabajar con funciones lineales. Este tema es uno de los más activos en la investigación actual en el ámbito del Análisis Funcional, y guarda una estrecha relación con lo visto en este trabajo.

Para concluir, señalamos que el objetivo de este trabajo ha sido el de analizar, entender y detallar minuciosamente todos aquellos elementos que se dan por hechos en ciertos textos, tratando de hacer más comprensible y para un público más amplio algunos de los artículos en los que nos hemos basado. En concreto, una gran parte de estos esfuerzos han sido dedicados al artículo [6] y a la prueba del Teorema 3.14 que se daba en [12], donde ha sido necesario adaptar y desglosar la mayoría de la argumentación que allí se exponía, haciéndola más clara para el lector. De la misma manera, se han estudiado herramientas necesarias para las demostraciones de los resultados aquí vistos, como por ejemplo la topología débil*, la integral de Bochner, o ciertas propiedades de funciones convexas más avanzadas que las que aparecen en el Grado. Sin embargo, todo este conocimiento no termina aquí, sino que nos sirve para ampliar el repertorio de instrumentos con el que podemos afrontar futuros proyectos.

Índice general

Capítulo 1

Preliminares

En este capítulo, vamos a recordar conceptos y propiedades que han sido estudiados en distintas asignaturas del Grado de Matemáticas y de los que haremos uso durante este trabajo.

1.1. Topología

En primer lugar, repasaremos algunas nociones de Topología general y más en concreto sobre espacios métricos, ya que será el ámbito en el que trabajaremos en todo momento. Para más detalles ver [9].

Definición 1. Sea (X,τ) un espacio topológico. Decimos que un subconjunto $D\subseteq X$ es *denso* en X si la clausura de D es el total. Es decir si $\overline{D}=X$.

Además, si X contiene un subconjunto D denso y numerable, diremos que X es *separable*.

Definición 2. Sean $(X,\tau),(Y,\xi)$ dos espacios topológicos. Decimos que una función $f:X\to Y$ es continua, si para todo abierto $U\subseteq Y$ se tiene que su preimagen $f^{-1}(U)$ es abierto en X.

Definición 3. Sean $(X,\tau),(Y,\xi)$ dos espacios topológicos. Decimos que una aplicación $\Phi:X\to Y$ es un *homeomorfismo sobre su imagen* si cumple que

I. Φ es inyectiva.

II. Φ es continua.

III. $\Phi^{-1}:\Phi(X)\to Y$ es continua.

Si además $\Phi(X)=Y$ (Φ sobreyectiva) diremos simplemente que Φ es un *homeomorfismo*.

Definición 4. Sea (X,τ) un espacio topológico. Decimos que (X,τ) es un *espacio de Hausdorff* si verifica que para cualesquiera $x,y\in X$ con $x\neq y$, existen $U_x,U_y\subseteq X$ entornos de x e y respectivamente tales que $U_x\cap U_y=\emptyset$.

Definición 5. Sea (X,τ) un espacio topológico. Decimos que una sucesión $(x_n)_n\subseteq X$ es *convergente* a $x\in X$ si cumple que para todo entorno $U\subseteq X$ de x , existe un $n_0\in\mathbb{N}$ tal que $x_n\in U$ para todo $n\geq n_0$.

Si (M,d) es un espacio métrico, diremos que una sucesión $(x_n)_n\subseteq M$ es de *Cauchy* si para todo $\varepsilon>0$ existe $n_0\in\mathbb{N}$ tal que $d(x_n,x_m)<\varepsilon$ para todo $n,m\geq n_0$. En este caso la definición que hemos dado de sucesión convergente se puede reformular como que $(x_n)_n$ es convergente a $x\in M$ si y solo si para cualquier $\varepsilon>0$ existe un $n_0\in\mathbb{N}$ tal que $d(x_n,x)<\varepsilon$ para todo $n\geq n_0$.

Nota que gracias a la desigualdad triangular, toda sucesión convergente en un espacio métrico es automáticamente de Cauchy. Sin embargo, vamos a estudiar los espacios que satisfacen el recíproco.

Definición 6. Sea (M,d) un espacio métrico. Decimos que M es *completo* si toda sucesión de Cauchy es convergente en M.

Finalmente, enunciamos uno de los teoremas más relevantes sobre topología de espacios métricos, que data de 1899.

Teorema 1.1 (Baire). *Sea (M,d) un espacio métrico completo y sea $(U_n)_n \subseteq M$ una colección numerable de abiertos densos. Entonces*

$$\bigcap_{n\in\mathbb{N}} U_n$$

es denso en M.

Demostración. Vamos a demostrar que para cualquier $V \subseteq M$ abierto, se tiene que $V \cap (\cap_{n\in\mathbb{N}} U_n) \neq \emptyset$. Por la densidad de U_1, $V \cap U_1$ es un abierto no vacío, luego tomemos $x_1 \in M$ y $0 < r_1 < 1$ tal que la bola cerrada $\mathbb{B}[x_1, r_1] \subseteq V \cap U_1$. De nuevo, por la densidad de U_2, existen $x_2 \in M$ y $0 < r_2 < \frac{1}{2}$ tal que $\mathbb{B}[x_2, r_2] \subseteq \mathbb{B}[x_1, r_1] \cap U_2 \subseteq V \cap U_1 \cap U_2$. De esta manera, podemos construir por inducción sucesiones $(x_n)_n \subseteq M$ y $(r_n)_n$ con $0 < r_n < \frac{1}{n}$ para todo $n \in \mathbb{N}$, y de modo que $\mathbb{B}[x_n, r_n] \subseteq \mathbb{B}[x_{n-1}, r_{n-1}] \cap U_n \subseteq V \cap U_1 \cap \ldots \cap U_n$. La sucesión $(x_n)_n$ es de Cauchy ya que $\mathbb{B}[x_n, r_n] \subseteq \mathbb{B}[x_{n-1}, r_{n-1}]$ y $\lim_{n\to\infty} r_n = 0$. Como M es completo, sea $x \in M$ el límite de la sucesión $(x_n)_n$. Por tanto, como $\mathbb{B}[x_n, r_n]$ es cerrado, se tiene que $x \in \mathbb{B}[x_n, r_n] \subseteq V \cap U_n$ para cada $n \in \mathbb{N}$, y por tanto $x \in V \cap (\cap_{n\in\mathbb{N}} U_n)$. □

1.2. Análisis Funcional

Este trabajo está enmarcado dentro del ámbito del Análisis Funcional, y por ello debemos introducir algunos conceptos básicos al respecto. Para más información sobre esta sección consultar [2].

Definición 7. Sea X un espacio vectorial y $p : X \to [0,\infty)$ una función no negativa. Decimos que p es una *seminorma* si para todo $x, y \in X$ y todo $\lambda \in \mathbb{R}$ cumple que

I. $p(\lambda x) = |\lambda| p(x)$.

II. $p(x+y) \leq p(x) + p(y)$.

Si además $p(x) = 0$ si y solo si $x = 0$, entonces p es una *norma*.
Si existe una norma $p : X \to [0,\infty)$ diremos que X es un *espacio normado*.

Notación importante. De ahora en adelante (Capítulos 2 y 3 incluidos) X, Y, Z serán espacios normados reales a menos que se indique lo contrario. Denotaremos su norma por $\|\cdot\|$, y se entenderá por el contexto si $\|\cdot\|$ hace referencia a la norma definida en uno u otro espacio.

Definición 8. Si $(X, \|\cdot\|)$ es un espacio normado completo, decimos que es un *espacio de Banach*.

A continuación, definiremos dos tipos de funciones que se comportan bien entre espacios normados, y que por ello son un amplio campo de estudio.

Definición 9. Sea una función $f : X \to Y$. Decimos que f es *Lipschitz* si existe una constante $K > 0$ tal que

$$\|f(x) - f(y)\| \leq K\|x - y\|, \quad \forall x, y \in X.$$

En tal caso también diremos que f es *K-Lipschitz*.

Definición 10. Sea una aplicación $\phi : X \to Y$. Decimos que ϕ es una *isometría* si para todo $x, y \in X$ se cumple que

$$\|\phi(x) - \phi(y)\| = \|x - y\|.$$

Observación 1. Nota que tanto las funciones Lipschitz como las isometrías son en particular funciones continuas.

Además, toda isometría es automáticamente inyectiva, puesto que si existen $x, y \in X$ con $\phi(x) = \phi(y)$ se tiene

$$0 = \|\phi(x) - \phi(y)\| = \|x - y\|$$

y por tanto por la condición de norma, se tiene $x = y$.

De hecho, veamos alguna propiedad más sobre isometrías.

Lema 1.2. *Sean $\alpha : X \to Y$, $\beta : Y \to Z$ isometrías, se tiene que*

I. *$\beta \circ \alpha$ es una isometría.*

II. *Si α es biyectiva (ó sobreyectiva), entonces α^{-1} es una isometría.*

Demostración. I. Si $x, y \in X$, entonces

$$\|\beta(\alpha(x)) - \beta(\alpha(y))\| = \|\alpha(x) - \alpha(y)\| = \|x - y\|.$$

II. Si $x, y \in Y$, entonces

$$\|\alpha^{-1}(x) - \alpha^{-1}(y)\| = \|\alpha^{-1}(\alpha(r)) - \alpha^{-1}(\alpha(s))\| = \|r - s\| = \|\alpha(r) - \alpha(s)\| = \|x - y\|,$$

donde gracias a la biyectividad de α, existen $r, s \in X$ tales que $\alpha(r) = x$ y $\alpha(s) = y$.

□

Por otro lado, dada una función lineal $f : X \to Y$, diremos que f es *acotada* (en norma) si cumple que

$$\sup_{x \in B_X} \|f(x)\| < \infty.$$

De hecho, una función lineal es acotada si y solo si es continua. Gracias a ello, podemos definir en el espacio de funciones lineales y continuas de X en Y la siguiente norma.

$$\|f\| := \sup_{x \in B_X} \|f(x)\| = \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|} = \text{ínf}\{M > 0 \ : \ \|f(x)\| \leq M\|x\|, \ \forall x \in X\}. \tag{1.1}$$

Es fácil ver la equivalencia de dichas expresiones y que de hecho, definen una norma. En particular, $\|f\| < \infty$ por ser lineal y continua como acabamos de ver. Por otra parte, atendiendo a la definición, concluimos que si f es lineal y continua se tiene que

$$\|f(x)\| \leq \|f\|\|x\|, \quad \forall x \in X.$$

Por todo esto, podemos definir un espacio normado en particular, que utilizaremos frecuentemente a lo largo de este escrito.

Definición 11. Denotamos por X^* el *espacio dual* de un espacio normado X, es decir el espacio vectorial

$$X^* := \{f : X \to \mathbb{R} \ : \ f \text{ lineal y continua}\}.$$

Se tiene que X^* es un espacio normado con la norma definida en (1.1).

Observación 2. En el caso en que $\dim(X) < \infty$, se puede probar como X^{**} es linealmente isométrico a X.

El teorema que presentamos a continuación, es uno de los principios fundamentales del Análisis Funcional y fue probado a finales de la década de 1920.

Teorema 1.3 (Hahn-Banach). *Sea Y un subespacio de X y $f : Y \to \mathbb{R}$ una aplicación lineal tal que $|f(x)| \leq \|x\|$ para todo $x \in Y$. Entonces existe una extensión lineal $\overline{f} : X \to \mathbb{R}$ de modo que $|\overline{f}(x)| \leq \|x\|$ para todo $x \in X$.*

En este trabajo, utilizaremos más frecuentemente la consecuencia que ahora enunciamos.

Corolario 1.4. *Si $x \neq 0$, entonces existe $f \in X^*$ tal que $\|f\| = 1$ y $f(x) = \|x\|$.*

Demostración. Dado $x \neq 0$, tomamos $Y = \text{span}(\{x\})$ y definimos la aplicación $f_0 : Y \to \mathbb{R}$ tal que $f_0(ax) = a\|x\|$ para todo $a \in \mathbb{R}$. Es claro que $f_0(x) = \|x\|$, y además f_0 es lineal, continua y cumple que

$$\|f_0\| = \sup_{y \neq 0} \frac{|f_0(y)|}{\|y\|} = \sup_{a \neq 0} \frac{|f_0(ax)|}{\|ax\|} = \sup_{a \neq 0} \frac{|a|}{|a|} = 1.$$

Luego por el Teorema 1.3 existe una extensión lineal f que conserva la norma y que por tanto satisface el enunciado. □

Capítulo 2

El Teorema de Mazur-Ulam

En este capítulo vamos a enunciar y demostrar el Teorema de Mazur-Ulam, que establece que toda isometría sobreyectiva entre espacios normados reales es afín.

2.1. Lemas previos

En esta sección se desarrollarán un par de lemas que nos van a servir de base para demostrar el Teorema de Mazur-Ulam.

Lema 2.1. *Sean X,Y espacios normados reales. Si $\alpha : X \to Y$ es continua y respeta los puntos medios, es decir, $\alpha(\frac{x+y}{2}) = \frac{\alpha(x)+\alpha(y)}{2}$ $\forall x,y \in X$, se tiene que*

I. $\alpha(x+y) = \alpha(x) + \alpha(y) - \alpha(0), \quad \forall x,y \in X.$

II. $\alpha(sx) = s\alpha(x) + (1-s)\alpha(0), \quad \forall x,y \in X, \forall s \in \mathbb{R}.$

Demostración. I. Sea $x \in X$. Como α conserva puntos medios, entonces

$$\alpha(x) = \alpha\left(\frac{1}{2}(2x+0)\right) = \frac{1}{2}(\alpha(2x) + \alpha(0))$$

y por tanto $\alpha(2x) = 2\alpha(x) - \alpha(0)$. Así pues,

$$\begin{aligned}\alpha(x+y) &= \alpha\left(\frac{1}{2}(2x+2y)\right) = \frac{1}{2}(\alpha(2x) + \alpha(2y)) \\ &= \frac{1}{2}(2\alpha(x) - \alpha(0) + 2\alpha(y) - \alpha(0)) = \alpha(x) + \alpha(y) - \alpha(0).\end{aligned}$$

II. Sea $x \in X$. Gracias al apartado anterior se prueba por inducción que se cumple para $s \in \mathbb{N}$ y también que

$$\alpha(0) = \alpha(x + (-x)) = \alpha(x) + \alpha(-x) - \alpha(0)$$

luego $\alpha(-x) = 2\alpha(0) - \alpha(x)$ y por tanto es cierto para $s = -1$. En definitiva, es cierto para $s \in \mathbb{Z}$. Sea $s = \frac{m}{n} \in \mathbb{Q}$ con $m,n \in \mathbb{Z}$. Tenemos que

$$\begin{aligned}m\alpha(x) &= \alpha(mx) + (m-1)\alpha(0) = \alpha\left(\frac{mn}{n}x\right) + (m-1)\alpha(0) \\ &= n\alpha\left(\frac{m}{n}x\right) + (m-1-n+1)\alpha(0) = n\left(\alpha\left(\frac{m}{n}x\right) + \left(\frac{m}{n}-1\right)\alpha(0)\right).\end{aligned}$$

Es decir,

$$\alpha(sx) = \alpha\left(\frac{m}{n}x\right) = \frac{m}{n}\alpha(x) + \left(1 - \frac{m}{n}\right)\alpha(0) = s\alpha(x) + (1-s)\alpha(0).$$

Visto que se cumple para los racionales, sea $s \in \mathbb{R}$ y tomamos una sucesión $(q_n)_n \subseteq \mathbb{Q}$ convergiendo a s. Así,

$$\begin{aligned} s\alpha(x) + (1-s)\alpha(0) &= \lim_{n\to\infty} q_n\alpha(x) + (1 - \lim_{n\to\infty} q_n)\alpha(0) \\ &= \lim_{n\to\infty} (q_n\alpha(x) + (1-q_n)\alpha(0)) = \lim_{n\to\infty} \alpha(q_n x) = \alpha(sx), \end{aligned}$$

donde la última igualdad se cumple puesto que α es continua.

□

Lema 2.2. *Sean X, Y espacios normados reales, $\alpha : X \to Y$ una aplicación. Son equivalentes:*

I. $\alpha - \alpha(0)$ *es lineal.*

II. α *es afín.*

Demostración. Para la primera implicación definimos $L(x) = \alpha(x) - \alpha(0)$, $\forall x \in X$, que es lineal por hipótesis. Entonces dados $x, y \in X$ y dado $t \in [0,1]$ se tiene que

$$\begin{aligned} \alpha(tx + (1-t)y) &= L(tx + (1-t)y) + \alpha(0) = tL(x) + (1-t)L(y) + \alpha(0) \\ &= t(\alpha(x) - \alpha(0)) + (1-t)(\alpha(y) - \alpha(0)) + \alpha(0) = t\alpha(x) + (1-t)\alpha(y) \end{aligned}$$

y por tanto α es afín.
Veamos ahora el recíproco. Definimos de nuevo $L(x) = \alpha(x) - \alpha(0)$, $\forall x \in X$. Fijamos $x \in X$ y vamos a ver que $L(tx) = tx, \forall t \in \mathbb{R}$. Primero sea $t \in [0,1]$, usando que α es afín, se tiene que

$$\begin{aligned} L(tx) &= \alpha(tx) - \alpha(0) = \alpha(tx + (1-t)0) - \alpha(0) \\ &= t\alpha(x) + (1-t)\alpha(0) - \alpha(0) = t(\alpha(x) - \alpha(0)) = tL(x). \end{aligned}$$

Ahora sea $t > 1$, es decir, $\frac{1}{t} \in (0,1)$ y por tanto, puesto que α es afín, se tiene

$$\alpha(x) = \frac{1}{t}\alpha(tx) + \left(1 - \frac{1}{t}\right)\alpha(0)$$

luego $\alpha(tx) = t\alpha(x) - (t-1)\alpha(0)$. Entonces

$$L(tx) = \alpha(tx) - \alpha(0) = t\alpha(x) - (t-1)\alpha(0) - \alpha(0) = t(\alpha(x) - \alpha(0)) = tL(x).$$

Por último sea $t = -1$. En este caso es claro que

$$\alpha(0) = \alpha\left(\frac{-x+x}{2}\right) = \frac{1}{2}(\alpha(-x) + \alpha(x))$$

y así $L(x) + L(-x) = \alpha(-x) + \alpha(x) - 2\alpha(0) = 0$. De aquí se sigue que $L(tx) = tL(x), \forall t \in \mathbb{R}$.
Ahora fijamos también $y \in X$ y veamos $L(x+y) = L(x) + L(y)$. Como α es afín,

$$\begin{aligned} L(x+y) &= 2L\left(\frac{x+y}{2}\right) = 2\alpha\left(\frac{x+y}{2}\right) - 2\alpha(0) = 2\frac{\alpha(x) + \alpha(y)}{2} - 2\alpha(0) \\ &= \alpha(x) - \alpha(0) + \alpha(y) - \alpha(0) = L(x) + L(y). \end{aligned}$$

De este modo ya está visto que L es lineal. □

2.2. Teorema de Mazur-Ulam

Una vez probados los tres lemas, podemos pasar a demostrar el Teorema de Mazur-Ulam basándonos en la demostración elaborada por Bogdan Nica en 2012 [10].

Teorema 2.3 (Mazur-Ulam, 1932)**.** *Sean X,Y espacios normados reales, $\alpha : X \to Y$ una isometría biyectiva. Entonces, α es afín.*

Demostración. Sean $x,y \in X$ y vamos a probar que $\alpha\left(\frac{x+y}{2}\right) = \frac{\alpha(x)+\alpha(y)}{2}$. Definimos el "defecto afín" como

$$\operatorname{def}(\alpha) = \left\| \alpha\left(\frac{x+y}{2}\right) - \frac{\alpha(x)+\alpha(y)}{2} \right\|.$$

Vamos a acotar esta cantidad de la siguiente forma:

$$\begin{aligned}
\operatorname{def}(\alpha) &= \left\| \frac{1}{2}\left(\alpha\left(\frac{x+y}{2}\right) - \alpha(x) + \alpha\left(\frac{x+y}{2}\right) - \alpha(y) \right) \right\| \\
&\leq \frac{1}{2}\left\| \alpha\left(\frac{x+y}{2}\right) - \alpha(x) \right\| + \frac{1}{2}\left\| \alpha\left(\frac{x+y}{2}\right) - \alpha(y) \right\| \\
&= \frac{1}{2}\left\| \frac{x+y}{2} - x \right\| + \frac{1}{2}\left\| \frac{x+y}{2} - y \right\| = \frac{1}{4}\left(\|x-y\| + \|y-x\|\right) = \frac{1}{2}\|x-y\|,
\end{aligned}$$

donde la desigualdad se debe a las propiedades de la norma y la desigualdad triangular, y la siguiente igualdad es gracias al hecho de que α es una isometría.
Ahora definimos otra isometría biyectiva α' en X tal que su defecto es el doble que el defecto de α. Para ello, tomamos $\alpha' = \alpha^{-1}\rho\alpha$ donde ρ es la simetría con respecto a $\frac{\alpha(x)+\alpha(y)}{2}$. Es decir, $\rho(z) = \alpha(x)+\alpha(y)-z, \forall z \in Y$. Nota que α^{-1} es una isometría (Lema 1.2) y ρ también lo es por ser una simetría. De nuevo por el Lema 1.2 tenemos que α' es isometría por ser composición de isometrías, y evidentemente biyectiva puesto que α y ρ lo son también. Además

$$\alpha'(x) = \alpha^{-1}(\rho(\alpha(x))) = \alpha^{-1}(\alpha(x)+\alpha(y)-\alpha(x)) = \alpha^{-1}(\alpha(y)) = y$$

y análogamente $\alpha'(y) = x$. Por tanto, se tiene que:

$$\begin{aligned}
\operatorname{def}(\alpha') &= \left\| \alpha'\left(\frac{x+y}{2}\right) - \frac{\alpha'(x)+\alpha'(y)}{2} \right\| = \left\| \alpha^{-1}\left(\alpha(x)+\alpha(y) - \alpha\left(\frac{x+y}{2}\right) \right) - \frac{x+y}{2} \right\| \\
&= \left\| \alpha^{-1}\left(\alpha(x)+\alpha(y) - \alpha\left(\frac{x+y}{2}\right) \right) - \alpha^{-1}\left(\alpha\left(\frac{x+y}{2}\right) \right) \right\|.
\end{aligned}$$

Y puesto que α^{-1} es una isometría,

$$\operatorname{def}(\alpha') = \left\| \alpha(x)+\alpha(y) - \alpha\left(\frac{x+y}{2}\right) - \alpha\left(\frac{x+y}{2}\right) \right\| = 2\left\| \frac{\alpha(x)+\alpha(y)}{2} - \alpha\left(\frac{x+y}{2}\right) \right\| = 2\operatorname{def}(\alpha).$$

Si $\operatorname{def}(\alpha) > 0$, iterando este proceso k veces, obtendríamos una isometría α_k cuyo defecto sería igual a $2^k \operatorname{def}(\alpha)$. Para un k suficientemente grande tendríamos que $\operatorname{def}(\alpha_k) = 2^k \operatorname{def}(\alpha) > \frac{\|x-y\|}{2}$ lo que contradice la acotación inicial que era válida para cualquier isometría definida en X.
Esto nos deja una única opción, que es $\operatorname{def}(\alpha) = 0$ y por tanto

$$\alpha\left(\frac{x+y}{2}\right) = \frac{\alpha(x)+\alpha(y)}{2}, \quad \forall x,y \in X,$$

luego α respeta los puntos medios, y además es continua por ser una isometría. Veamos ahora que $L(x) = \alpha(x) - \alpha(0)$ es lineal. Para ello le aplicamos el Lema 2.1 a α :

$$L(x+y) = \alpha(x+y) - \alpha(0) = \alpha(x)+\alpha(y)-\alpha(0)-\alpha(0) = L(x)+L(y), \quad \forall x,y \in X.$$

$$L(tx) = \alpha(tx) - \alpha(0) = t\alpha(x) + (1-t)\alpha(0) - \alpha(0) = t(\alpha(x)-\alpha(0)) = tL(x), \quad \forall x,y \in X, \forall t \in \mathbb{R}.$$

Por tanto $\alpha - \alpha(0)$ es lineal, y por el Lema 2.2, obtenemos que α es afín. $\square$

Nuestro objetivo ahora es generalizar el Teorema de Mazur-Ulam a isometrías no necesariamente biyectivas, en un teorema probado por Baker. Para ello, empezaremos por ver una prueba alternativa de que toda isometría biyectiva respeta los puntos medios, para después, apoyándonos sobre esa base, demostrar el Teorema de Baker. Para ello, nos basaremos en [5].

Antes de nada, definimos el *centro (métrico)* de los puntos $x,y \in X$ como el único punto perteneciente a $\bigcap_{n\in\mathbb{N}} H_n(x,y)$ (siempre que dicha intersección sea no vacía), donde los H_n vienen definidos de la siguiente forma:

$$H_1(x,y) = \left\{ u \in X : \|x-u\| = \|y-u\| = \frac{\|x-y\|}{2} \right\}$$
$$H_n(x,y) = \left\{ u \in H_{n-1}(x,y) : \|u-v\| \leq \frac{\delta(H_{n-1})}{2},\ \forall v \in H_{n-1} \right\}, \quad \forall n \geq 2,$$

donde $\delta(H_n)$ es el diámetro de H_n, es decir el supremo de las distancias entre cualesquiera dos elementos de dicho conjunto.
Usaremos en general la notación H_n en lugar de $H_n(x,y)$ si no hay dudas sobre x e y.

Para entenderlo mejor, vamos a visualizar los conjuntos H_1 (en color negro) y H_2 (en color azul) para diferentes puntos (x,y) (en color gris) en $\mathbb{R}^2$ con las normas $\|\cdot\|_1$ (en color amarillo) y $\|\cdot\|_\infty$ (en color rojo). Observamos que H_1 es la intersección de las bolas de radio $\frac{\|x-y\|}{2}$ y centros x e y, y que H_2 ya es solamente el punto $\frac{x+y}{2}$, es decir:

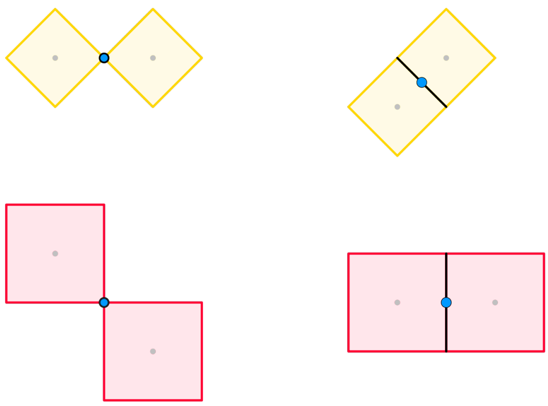

Figura 2.1: Conjuntos H_1 y H_2

Nota que si tomásemos la norma $\|\cdot\|_2$, tendríamos que, independientemente de los (x,y) que escogiéramos, H_1 estaría siempre formado por un único punto $\frac{x+y}{2}$ que sería el centro métrico de x e y. Esto se debe a una característica de $(\mathbb{R}^2, \|\cdot\|_2)$ que más adelante comentaremos.

Volviendo a la definición de centro métrico, veamos que tiene sentido. Probamos por inducción que $\delta(H_n) \leq \frac{\|x-y\|}{2^{n-1}}, \forall n \in \mathbb{N}$. Para $u,v \in H_1$ se tiene que

$$\|u-v\| \leq \|u-x\| + \|x-v\| = \frac{\|x-y\|}{2} + \frac{\|x-y\|}{2} = \|x-y\|,$$

luego $\delta(H_1) \leq \|x-y\|$. Supongamos cierto para n y veamos qué ocurre para $n+1$. Para $u,v \in H_{n+1}$ tenemos que

$$\|u-v\| \leq \frac{\delta(H_n)}{2} \leq \frac{1}{2}\frac{\|x-y\|}{2^{n-1}} = \frac{\|x-y\|}{2^n}.$$

Por tanto $\delta(H_n) \leq \frac{\|x-y\|}{2^{n-1}}, \forall n \in \mathbb{N}$, y en particular tiende a 0, luego la intersección $\bigcap_{n\in\mathbb{N}} H_n$ está formada por un único punto o es igual al vacío.

Lema 2.4 (Banach, 1932)**.** *Si $x,y \in X$ entonces el punto $\frac{1}{2}(x+y)$ es el centro métrico del par x,y.*

Demostración. Para cada $u \in X$ tomamos $u' = x + y - u$. Vamos a probar por inducción que

$$u \in H_n(x,y) \Rightarrow u' \in H_n(x,y), \quad \forall n \in \mathbb{N}.$$

Si $u \in H_1(x,y)$, entonces $\|u' - x\| = \|y - u\|$ y $\|u' - y\| = \|x - u\|$ luego

$$\|u' - x\| = \|u' - y\| = \frac{\|x - y\|}{2}$$

puesto que $u \in H_1$. Asumimos que es cierto para $n-1$ y veamos que se cumple también para n. Supongamos $u \in H_n$. Si $v \in H_{n-1}$, entonces

$$\|u' - v\| = \|(x + y - v) - u\| = \|v' - u\| \leq \frac{\delta(H_{n-1})}{2}$$

ya que $v' \in H_{n-1}$ y $u \in H_n$. Por tanto, $u' \in H_n$ y queda demostrado para todo $n \in \mathbb{N}$.
Ahora vamos a demostrar, de nuevo por inducción, que $z = \frac{1}{2}(x+y) \in H_n$, $\forall n \in \mathbb{N}$. Es claro que $z \in H_1$ porque $\|z - x\| = \|z - y\| = \frac{\|x-y\|}{2}$. Supongamos ahora que $z \in H_{n-1}$ y sea $u \in H_{n-1}$. Como hemos visto antes $u' \in H_{n-1}$ y así

$$2\|z - u\| = \|x + y - 2u\| = \|u' - u\| \leq \delta(H_{n-1}).$$

Por tanto, $\|z - u\| \leq \frac{\delta(H_{n-1})}{2}$ y $z \in H_n$. En definitiva $z \in \bigcap_{n\in\mathbb{N}} H_n(x,y)$ y por tanto es el centro de x, y. □

Proposición 2.5. *Si $\alpha : X \to Y$ es una isometría sobreyectiva, entonces $\alpha\left(\frac{1}{2}(x+y)\right)$ es el centro de $\alpha(x), \alpha(y)$ en Y, para cualesquiera $x, y \in X$.*

Demostración. Si $u \in H_1(x,y)$, entonces

$$\|\alpha(u) - \alpha(x)\| = \|u - x\| = \frac{\|x - y\|}{2} = \frac{\|\alpha(x) - \alpha(y)\|}{2}$$

y análogamente $\|\alpha(u) - \alpha(y)\| = \frac{\|\alpha(x)-\alpha(y)\|}{2}$. Luego $\alpha(u) \in H_1(\alpha(x), \alpha(y))$ y por tanto
$\alpha(H_1(x,y)) \subseteq H_1(\alpha(x), \alpha(y))$. Recíprocamente, sea $v \in H_1(\alpha(x), \alpha(y))$ y gracias a la sobreyectividad de α sea $u \in X$ tal que $v = \alpha(u)$. Entonces

$$\|u - x\| = \|\alpha(u) - \alpha(x)\| = \|v - \alpha(x)\| = \frac{\|\alpha(x) - \alpha(y)\|}{2} = \frac{\|x - y\|}{2}.$$

Así pues $u \in H_1(x,y)$ y se tiene $H_1(\alpha(x), \alpha(y)) \subseteq \alpha(H_1(x,y))$. Es decir,

$$H_1(\alpha(x), \alpha(y)) = \alpha(H_1(x,y)). \tag{2.1}$$

Además como α es isometría, la distancia entre cualesquiera dos puntos de $H_1(\alpha(x), \alpha(y))$ es igual a la distancia entre sus preimágenes por α, que por lo probado anteriormente están en $H_1(x,y)$, luego el diámetro de dichos conjuntos es idéntico. Es decir

$$\delta(H_1(x,y)) = \delta(H_1(\alpha(x), \alpha(y))). \tag{2.2}$$

Asumimos que (2.1) y (2.2) son ciertas para $n-1$ y tomamos $u \in H_n(x,y)$ y $v \in H_{n-1}(\alpha(x), \alpha(y))$. Entonces por hipótesis de inducción existe $v \in H_{n-1}(x,y)$ tal que $\alpha(v) = w$, luego

$$\|\alpha(u) - w\| = \|\alpha(u) - \alpha(v)\| = \|u - v\| \leq \frac{\delta(H_{n-1}(x,y))}{2} = \frac{\delta(H_{n-1}(\alpha(x), \alpha(y)))}{2}.$$

Por tanto, $\alpha(u) \in H_n(\alpha(x), \alpha(y))$. Recíprocamente, si $u \in H_{n-1}(x,y)$, $w \in H_n(\alpha(x), \alpha(y))$ y $v \in X$ tal que $\alpha(v) = w$, obtenemos

$$\|u - v\| = \|\alpha(u) - w\| \leq \frac{\delta(H_{n-1}(\alpha(x), \alpha(y)))}{2} = \frac{\delta(H_{n-1}(x,y))}{2}.$$

Luego $v \in H_n(x,y)$. En definitiva (2.1) es cierta para todo $n \in \mathbb{N}$ y como consecuencia (2.2) también. Ahora bien, como por el Lema 2.4, se cumple que $\frac{1}{2}(x+y) \in H_n(x,y)$ para todo $n \in \mathbb{N}$, se tiene que $\alpha\left(\frac{1}{2}(x+y)\right) \in H_n(\alpha(x), \alpha(y)), \forall n \in \mathbb{N}$. De esta forma $\alpha\left(\frac{1}{2}(x+y)\right)$ es el único elemento de $\bigcap_{n\in\mathbb{N}} H_n(x,y)$, y por tanto es el centro de $\alpha(x), \alpha(y)$ y es igual a $\frac{1}{2}(\alpha(x) + \alpha(y))$ de nuevo por el Lema 2.4. □

Demostración alternativa del Teorema de Mazur-Ulam. Puesto que α es una isometría sobreyectiva, se tiene que es continua y conserva los puntos medios gracias a la Proposición 2.5 y al Lema 2.4. Aplicando el Lema 2.1 deducimos que $\alpha - \alpha(0)$ es lineal y por el Lema 2.2, tenemos que α es afín. □

Ejemplo 1 (Figiel)**.** Veamos como la hipótesis de sobreyectividad es necesaria para el Teorema de Mazur-Ulam. Para ello, consideramos $X = \mathbb{R}$ con la norma euclídea usual, $Y = \mathbb{R}^2$ con la norma $\|(x,y)\|_\infty = \text{máx}\{|x|,|y|\}$.
Tomamos la aplicación $U(x) = (x, \text{sen}(x))$ definida en X y que claramente tiene su imagen en Y pero no es sobreyectiva. La función U no es lineal, luego tampoco es afín ya que $U(0) = (0,0)$. Sin embargo U es una isometría, puesto que

$$\|U(x) - U(y)\|_\infty = \text{máx}\{|x-y|, |\text{sen}(x) - \text{sen}(y)|\} = |x-y|, \quad \forall x,y \in \mathbb{R},$$

donde la última igualdad se debe a que por el teorema del valor medio, aplicado a la función seno, existe $z \in \mathbb{R}$ tal que

$$|\text{sen}(x) - \text{sen}(y)| = |\text{sen}'(z)||x-y| = |\cos(z)||x-y| \leq |x-y|.$$

En definitiva, hemos encontrado una isometría no sobreyectiva que no es afín.

2.3. Teorema de Baker

Vamos a ver que podemos suprimir la condición de sobreyectividad si exigimos que nuestro espacio de llegada Y sea estrictamente convexo. De nuevo, nos basaremos principalmente en [5].

Definición 12. Decimos que Y es *estrictamente convexo* si dados $x,y \in Y$ tales que $\|x+y\| = \|x\| + \|y\|$, necesariamente x,y son linealmente dependientes. En ese caso, si $y \neq 0$, existe $t \geq 0$ tal que $x = ty$.

Lema 2.6. *Sean $u,v \in Y$ tales que $\|u+v\| = \|u\| + \|v\|$. Entonces, para todo $t \geq 0$, se tiene que*

$$\|tu+v\| = t\|u\| + \|v\|.$$

Demostración. Definimos $\varphi_1(t) = \|tu+v\|$ y $\varphi_2(t) = t\|u\| + \|v\|$ para $t \geq 0$. Para $\lambda \in [0,1]$, $r,s \geq 0$, tenemos que

$$\begin{aligned}\varphi_1(\lambda r + (1-\lambda)s) &= \|(\lambda r + (1-\lambda)s)u + v\| = \|\lambda(ru+v) + (1-\lambda)(su+v)\| \\ &\leq \lambda\|ru+v\| + (1-\lambda)\|su+v\| = \lambda\varphi_1(r) + (1-\lambda)\varphi_1(s).\end{aligned}$$

Por tanto φ_1 es convexa. Además $\varphi_1(0) = \varphi_2(0)$ y $\varphi_1(1) = \varphi_2(1)$, y finalmente se tiene por la desigualdad triangular que $\varphi_1(t) \leq \varphi_2(t)$, para todo $t \geq 0$. Por todo ello debe ser $\varphi_1(t) = \varphi_2(t)$, para todo $t \geq 0$, o equivalentemente $\|tu+v\| = t\|u\| + \|v\|$ para todo $t \geq 0$. □

Lema 2.7. *Si Y es estrictamente convexo, entonces $H_1(x,y) = \left\{\frac{1}{2}(x+y)\right\}$ para cualesquiera $x,y \in Y$.*

Demostración. La prueba es trivial si $x = y$, luego supongamos $x \neq y$. Ya vimos en el Lema 2.4 que $\frac{1}{2}(x+y) \in H_1(x,y)$. Supongamos que existen $u,v \in H_1(x,y)$. Entonces

$$\begin{aligned}\left\|x - \frac{1}{2}(u+v)\right\| &= \left\|\frac{1}{2}(x-u) + \frac{1}{2}(x-v)\right\| \leq \frac{1}{2}(\|x-u\| + \|x-v\|) \\ &= \frac{1}{2}\left(\frac{1}{2}\|x-y\| + \frac{1}{2}\|x-y\|\right) = \frac{1}{2}\|x-y\|.\end{aligned}$$

Análogamente se obtiene $\left\|y - \frac{1}{2}(u+v)\right\| \leq \frac{1}{2}\|x-y\|$. Además dichas desigualdades son en realidad igualdades puesto que en caso contrario $\|x-y\| \leq \left\|x - \frac{1}{2}(u+v)\right\| + \left\|y - \frac{1}{2}(u+v)\right\| < \|x-y\|$, que es claramente imposible. Así pues, obtenemos en realidad que

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

Pero como Y es estrictamente convexo, $x - u = t(x-v)$ para algún $t \geq 0$. Como además $\|x-u\| = \frac{1}{2}\|x-y\| = \|x-v\|$ necesariamente $t = 1$, luego $u = v = \frac{1}{2}(x+y)$ como queríamos demostrar. □

Observación 3. $(\mathbb{R}^n, \|\cdot\|_2)$ es estrictamente convexo, mientras que no lo es con las normas $\|\cdot\|_1$ ó $\|\cdot\|_\infty$. Esto justifica lo comentado en la Figura 2.1.

Teorema 2.8 (Baker, 1971)**.** *Si $\alpha : X \to Y$ es una isometría y además Y es estrictamente convexo, entonces α es afín.*

Demostración. Sean $x, y \in X$. Como vimos en la Proposición 2.5, $\alpha(H_1(x,y)) \subseteq H_1(\alpha(x), \alpha(y))$ (para probar dicho contenido no se usó la sobreyectividad). Así, puesto que $\frac{1}{2}(x+y) \in H_1(x,y)$, se tiene que $\alpha\left(\frac{1}{2}(x+y)\right) \in H_1(\alpha(x), \alpha(y))$, que gracias al Lema 2.7 es un conjunto formado por un único punto $\frac{1}{2}(\alpha(x)+\alpha(y))$. En definitiva

$$\alpha\left(\frac{1}{2}(x+y)\right) = \frac{1}{2}(\alpha(x)+\alpha(y)).$$

Es decir, α respeta puntos medios. Ahora aplicamos los Lemas 2.2 y 2.1 como en la demostración del Teorema de Mazur-Ulam 2.3 y obtenemos que α es afín. $\square$

De nuevo veamos con un ejemplo que la condición de que Y sea estrictamente convexo es necesaria.

Ejemplo 2. Supongamos que Y no es estrictamente convexo y tomemos $X = \mathbb{R}$. Como Y no es estrictamente convexo, tomamos $u, v \in Y$ linealmente independientes con $\|u\| = \|v\| = 1$ y tal que $\|u+v\| = \|u\| + \|v\|$. Definimos

$$\alpha(t) = \begin{cases} tu & \text{si} \quad t \leq 1 \\ u + (t-1)v & \text{si} \quad t > 1. \end{cases}$$

Veamos que α es una isometría:
Si $t, r \leq 1$,

$$\|\alpha(t) - \alpha(r)\| = \|tu - ru\| = |t-r|\|u\| = |t-r|.$$

Si $t, r > 1$,

$$\|\alpha(t) - \alpha(r)\| = \|u + (t-1)v - u - (r-1)v\| = \|(t-r)v\| = |t-r|\|v\| = |t-r|.$$

Si $t \leq 1, r > 1$ (análogamente si $t > 1, r \leq 1$),

$$\begin{aligned} \|\alpha(t) - \alpha(r)\| &= \|tu - u - (r-1)v\| = \|(t-1)u + (1-r)v\| = (r-1)\left\|\frac{t-1}{1-r}u + v\right\| \\ &= (r-1)\left(\frac{t-1}{1-r}\|u\| + \|v\|\right) = (r-1)\left(\frac{t-1}{1-r} + 1\right) = r - t = |t-r|, \end{aligned}$$

donde hemos aplicado el Lema 2.6 ya que $\frac{t-1}{1-r} \geq 0$ puesto que $t \leq 1$ y $r > 1$.
Luego α es una isometría. Además

$$0 = \alpha(0) = \alpha(-2+2) \neq \alpha(-2) + \alpha(2) = -2u + u + (2-1)v = v - u$$

ya que $u \neq v$ por ser linealmente independientes. Por tanto α no es lineal, luego no es afín ($\alpha(0) = 0$).

Ejemplo 3. Veamos además que ni el Teorema de Mazur-Ulam ni el de Baker pueden extenderse al caso complejo. Para ello, notemos que la aplicación $U : \mathbb{C} \to \mathbb{C}$ tal que $U(z) = \overline{z}$ es una isometría sobreyectiva no lineal, con $U(0) = 0$ y $\mathbb{C}$ es estrictamente convexo.
Para todo $z, w \in \mathbb{C}$,

$$|U(z) - U(w)| = |\overline{z} - \overline{w}| = |\overline{z-w}| = |z-w|.$$

luego U es una isometría. Por otro lado,

$$-i = U(i) \neq iU(1) = i$$

luego U no es $\mathbb{C}$-lineal.
Y obviamente $U(0) = 0$ y U es sobreyectiva (de hecho biyectiva con inversa ella misma).
Además $\mathbb{C}$ se identifica con $(\mathbb{R}^2, \|\cdot\|_2)$ luego es estrictamente convexo ya que $\mathbb{R}^2$ lo es.
En definitiva ambos teoremas fallan cuando pasamos al cuerpo complejo.

Capítulo 3

El Teorema de Godefroy-Kalton

Este capítulo lo dedicaremos a estudiar si la existencia de una isometría (no lineal) implica que podemos encontrar otra isometría que sí sea lineal entre los mismos espacios normados. En concreto probaremos el Teorema de Godefroy-Kalton, que demuestra que esto ocurre siempre que el espacio de partida sea separable. Este capítulo se basará en [6] y [12] aunque también serán necesarias otras herramientas que iremos introduciendo a lo largo del mismo.

3.1. Convexidad

En esta sección, estudiaremos el comportamiento de las funciones convexas. Nos basaremos en [1].

Definición 13. Sea $C \subseteq \mathbb{R}^n$ convexo y sea $f : C \to \mathbb{R}$ una función. Decimos que f es *convexa* si

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \quad \forall x, y \in C, \forall \lambda \in [0,1].$$

Una vez recordada su definición, veamos algunas propiedades interesantes sobre estas funciones.

Lema 3.1. *Sea $C \subseteq \mathbb{R}^n$ convexo. Si $f : C \to \mathbb{R}$ es convexa, entonces para todo $k \in \mathbb{N}$, cualesquiera $x_1, \dots, x_k \in C$ y $\lambda_i \geq 0$ tal que $\lambda_1 + \cdots + \lambda_k = 1$, se tiene*

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k).$$

Demostración. Se sigue de aplicar la definición de convexidad primero para $\lambda = \frac{\lambda_1}{\lambda_1+\lambda_2}$ luego para $\lambda = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3}$ y así sucesivamente hasta llegar a $\lambda = \frac{\lambda_1+\cdots+\lambda_{k-1}}{\lambda_1+\cdots+\lambda_{k-1}+\lambda_k} = \lambda_1 + \cdots + \lambda_{k-1}$, obteniendo el resultado que queremos gracias a que se van cancelando denominador con numerador consecutivos. □

Lema 3.2. *Sea $f : \mathbb{R} \to \mathbb{R}$ convexa, entonces para cualesquiera $x < z < y$ se cumple que*

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(z)}{y - z}.$$

Demostración. Tenemos que $z = \lambda y + (1-\lambda)x$, con $\lambda = \frac{z-x}{y-x} \in (0,1)$. Por tanto

$$f(z) \leq \lambda f(y) + (1-\lambda)f(x)$$

Luego $f(z) - f(x) \leq \lambda(f(y) - f(x))$, y entonces

$$\frac{f(z) - f(x)}{z - x} \leq \frac{\lambda(f(y) - f(x))}{z - x} = \frac{f(y) - f(x)}{y - x}.$$

Análogamente se demuestra la segunda desigualdad. □

Vamos a ilustrar el significado del Lema 3.2 con la siguiente figura. Dibujamos la función $f(x) = x^2$ y tomamos tres puntos $x < z < y$. Es fácil apreciar como la pendiente de la recta entre x y z (color amarillo) es menor que la de x e y (color marrón) que a su vez es menor que la de z e y (color rojo).

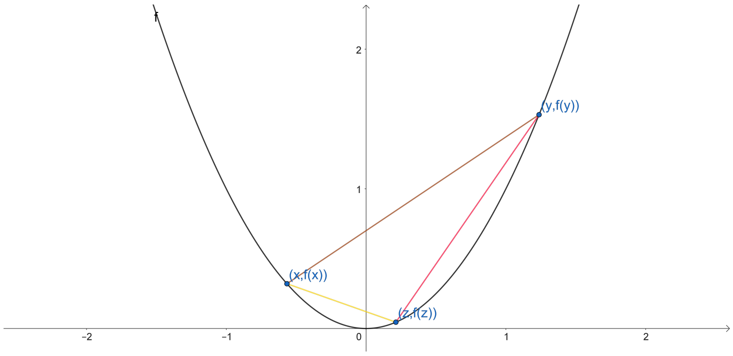

Figura 3.1: Visualización del Lema 3.2

Lema 3.3. *Sea $f:(a,b)\to\mathbb{R}$ convexa. Entonces f es continua.*

Demostración. Sea $x\in(a,b)$ y tomemos $a<x<v<b$. Así, gracias al Lema 3.2 se tiene

$$\frac{f(x)-f(a)}{x-a}\leq\frac{f(v)-f(x)}{v-x}\leq\frac{f(b)-f(x)}{b-x},$$

o lo que es lo mismo,

$$f(x)+\frac{f(x)-f(a)}{x-a}(v-x)\leq f(v)\leq f(x)+\frac{f(b)-f(x)}{b-x}(v-x).$$

Por tanto, $\lim_{v\to x^+} f(v)=f(x)$. Análogamente tomando $a<v<x<b$ se obtiene $\lim_{v\to x^-} f(v)=f(x)$. En definitiva, f es continua en x. □

Lema 3.4. *Si $f:(a,b)\to\mathbb{R}$ es convexa, entonces para cada $z\in(a,b)$ se tiene que la función*

$$\varphi_z(x)=\frac{f(x)-f(z)}{x-z},\quad \forall x\in I_z=(a,z)\cup(z,b)$$

es monótona creciente.

Demostración. Sean $x,y\in I_z$ con $x<y$. Tenemos tres casos:

I. $x<y<z,\quad \varphi_z(x)=\frac{f(z)-f(x)}{z-x}\leq\frac{f(z)-f(y)}{z-y}=\varphi_z(y).$

II. $z<x<y,\quad \varphi_z(x)=\frac{f(x)-f(z)}{x-z}\leq\frac{f(y)-f(z)}{y-z}=\varphi_z(y).$

III. $x<z<y,\quad \varphi_z(x)=\frac{f(z)-f(x)}{z-x}\leq\frac{f(y)-f(z)}{y-z}=\varphi_z(y).$

Nota que las desigualdades anteriores provienen del Lema 3.2. En todos los casos obtenemos $\varphi_z(x)\leq\varphi_z(y)$ como queríamos. □

Así pues la función $f_x(h)=\frac{f(x+h)-f(x)}{h}$ es creciente en $h\in\mathbb{R}\setminus\{0\}$ para cada x. Por tanto para h cercano a cero, digamos $h\in(-\theta,0)\cup(0,\theta)$ se tiene que

$$-\infty<f_x(-\theta)\leq f_x(h)\leq f_x(\theta)<+\infty.$$

Luego $f_x(h)$ es una función monótona y acotada cuando h tiende a cero, así que existen y son finitos los siguientes límites:

$$\lim_{h\to 0^+} f_x(h)=f'_+(x)$$
$$\lim_{h\to 0^-} f_x(h)=f'_-(x).$$

Proposición 3.5. *Sea $f:(a,b)\to\mathbb{R}$ convexa. Se tiene que*

I. *$f'_+(x), f'_-(x)$ existen y son finitas para todo $x\in(a,b)$.*

II. *$f'_-(x)\le f'_+(x)\le f'_-(y)\le f'_+(y)$ para todo $x,y\in(a,b)$ con $x<y$.*

III. *f'_-, f'_+ son crecientes.*

IV. *f es derivable en (a,b) excepto a lo sumo en un conjunto contable de puntos.*

Demostración. I. Acabamos de probarlo.

II. Por un lado, gracias a la monotonía de $f_x(h)$, tenemos que

$$f'_-(x)=\lim_{h\to 0^-} f_x(h)\le \lim_{h\to 0^+} f_x(h)=f'_+(x),\quad \forall x\in(a,b).$$

Siendo $x<y$, tomamos $0<h_1<\frac{y-x}{2}$ y $0>h_2>\frac{x-y}{2}$. Por tanto $x<x+h_1<y+h_2<y$, luego aplicando el Lema 3.2 se obtiene

$$f_x(h_1)=\frac{f(x+h_1)-f(x)}{h_1}\le\frac{f(y+h_2)-f(x+h_1)}{y+h_2-x-h_1}\le\frac{f(y)-f(y+h_2)}{h_2}=f_y(h_2)$$

y así,

$$f'_+(x)=\lim_{h\to 0^+} f_x(h)\le\lim_{h\to 0^-} f_y(h)=f'_-(y).$$

III. Obvio gracias a II.

IV. Como f'_+, f'_- son crecientes, sus únicas discontinuidades (si es que las hay) son de tipo salto. Supongamos que f'_+ es discontinua en $x\in(a,b)$. Entonces, tomamos $q_x\in\mathbb{Q}$ tal que $\lim_{y\to x^-} f'_+(y)<q_x<\lim_{y\to x^+} f'_+(y)$. Notando que $q_x<q_y$ si $x<y$, construimos una aplicación inyectiva del conjunto de discontinuidades de f'_+ en $\mathbb{Q}$, luego dicho conjunto es contable. Veamos ahora que si f'_+ es continua en x, entonces $f'_+(x)=f'_-(x)$ y por tanto f es derivable en x. Si f'_+ es continua en x,

$$\begin{aligned}f'_+(x)&=\lim_{\delta\to 0^-} f'_+(x+\delta)=\lim_{\delta\to 0^-}\lim_{h\to 0^+}\frac{f(x+\delta+h)-f(x+\delta)}{h}\\&\le\lim_{\delta\to 0^-}\lim_{h\to 0^+}\frac{f(x+h)-f(x+\delta+h)}{-\delta}=\lim_{\delta\to 0^-}\frac{f(x+\delta)-f(x)}{\delta}=f'_-(x)\le f'_+(x),\end{aligned}$$

donde la primera desigualdad se deduce aplicando el Lema 3.2, y la igualdad posterior gracias a la continuidad de f (Lema 3.3). Por tanto, $f'_+(x)=f'_-(x)$ como queríamos.

□

Lema 3.6. *Sea $f:\mathbb{R}\to\mathbb{R}$ convexa y sea $x\in\mathbb{R}$. Entonces f es derivable en x si y solo si*

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h}=0.$$

Demostración. Tenemos la siguiente cadena de afirmaciones equivalentes.

$$\begin{aligned}&f\text{ es derivable en }x\iff f'_+(x)=f'_-(x)\iff\lim_{h\to 0^+}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0^-}\frac{f(x+h)-f(x)}{h}\\
&\iff\begin{cases}\lim_{h\to 0^+}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0^+}\frac{f(x-h)-f(x)}{-h}\\ \lim_{h\to 0^-}\frac{f(x-h)-f(x)}{-h}=\lim_{h\to 0^-}\frac{f(x+h)-f(x)}{h}\end{cases}\\
&\iff\begin{cases}\lim_{h\to 0^+}\left(\frac{f(x+h)-f(x)}{h}-\frac{f(x-h)-f(x)}{-h}\right)=0\\ \lim_{h\to 0^-}\left(\frac{f(x-h)-f(x)}{-h}-\frac{f(x+h)-f(x)}{h}\right)=0\end{cases}\\
&\iff\begin{cases}\lim_{h\to 0^+}\frac{f(x+h)+f(x-h)-2f(x)}{h}=0\\ \lim_{h\to 0^-}\frac{f(x+h)+f(x-h)-2f(x)}{h}=0\end{cases}\\
&\iff\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h}=0.\end{aligned}$$

Nota que todos los límites que aparecen existen y son finitos gracias a la convexidad de f.

□

Lema 3.7 ([6]). *Sea C un subconjunto de $\mathbb{R}^n$ convexo no vacío y tal que $(-x) \in C$, para todo $x \in C$. Sea también $F : C \to \mathbb{R}$ una función convexa acotada superiormente, tal que $F(0) = 0$. Entonces se tiene que*

$$\sup_{x\in C} F(x) = \sup_{x\in C} |F(x)|.$$

Demostración. Dado $x \in C$, tenemos que

$$0 = F(0) = F\left(\frac{x+(-x)}{2}\right) \leq \frac{F(x)+F(-x)}{2}.$$

Luego $F(-x) \geq -F(x)$, y en consecuencia probamos que:

$$\sup_{x\in C} |F(x)| = \sup_{x\in C}\{\text{máx}\{F(x), -F(x)\}\} \leq \sup_{x\in C}\{\text{máx}\{F(x), F(-x)\}\} = \sup_{x\in C\cup(-C)} F(x) = \sup_{x\in C} F(x).$$

□

Enunciamos a continuación un resultado muy interesante que refleja que las funciones convexas pueden llegar a ser muy similares a funciones diferenciables.

Proposición 3.8 ([6]). *Sea $g : \mathbb{R}^n \to \mathbb{R}$ convexa. Entonces g es continua en $\mathbb{R}^n$ y diferenciable en un conjunto denso de $\mathbb{R}^n$.*

Demostración. Sea $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y sea $\alpha > 0$. Denotamos por $e_1, \ldots, e_n$ la base canónica de $\mathbb{R}^n$. Sea $0 \neq h \in \mathbb{R}^n$ con $\|h\|_1 = a \leq \alpha$. Es claro que $x + h = \sum_{i=1}^n \lambda_i E_i$, con $\lambda_i = \frac{|h_i|}{a}$ y $E_i = x + \text{sgn}(h_i) a e_i$. Así pues usando el Lema 3.1 y notando que $\sum_{i=1}^n \lambda_i = 1$ obtenemos que

$$g(x+h) \leq \sum_{i=1}^n \lambda_i g(E_i) \leq \text{máx}_{1\leq j\leq n} g(E_j) \sum_{i=1}^n \lambda_i \leq \text{máx}_{1\leq j\leq n, |\varepsilon|=1} g(x+\varepsilon a e_j).$$

Además, para $0 \leq a \leq \alpha$ se tiene que $x + \varepsilon a e_i = \lambda(x - \alpha e_i) + (1-\lambda)(x + \alpha e_i)$ para algún $\lambda \in [0,1]$ luego $g(x + \varepsilon a e_i) \leq \text{máx}\{g(x - \alpha e_i), g(x + \alpha e_i)\}$, para cada $i = 1, \ldots, n$. En definitiva podemos concluir que

$$\sup_{\|h\|_1 \leq \alpha} g(x+h) - g(x) = \text{máx}_{1\leq i\leq n, |\varepsilon|=1} g(x+\varepsilon\alpha e_i) - g(x).$$

Por otro lado, es evidente que la función $F(h) = g(x+h) - g(x)$ y el conjunto $C_\alpha = \{h \in \mathbb{R}^n : \|h\|_1 \leq \alpha\}$ cumplen las hipótesis del Lema 3.7, así que

$$\sup_{\|h\|_1 \leq \alpha} |g(x+h) - g(x)| = \text{máx}_{1\leq i\leq n, |\varepsilon|=1} g(x+\varepsilon\alpha e_i) - g(x).$$

Notemos que fijado $x \in \mathbb{R}^n$ las funciones $f_i^x : \mathbb{R} \to \mathbb{R}$ con $f_i^x(a) = g(x + a e_i)$ son convexas porque g lo es y por tanto continuas por el Lema 3.3 (en particular continuas en 0). Así, dado $\rho > 0$ existe $\delta > 0$ tal que si $|a| \leq \delta$, entonces $\text{máx}_{1\leq i\leq n} g(x + a e_i) - g(x) < \rho$. Luego

$$\sup_{\|h\|_1 \leq \delta} |g(x+h) - g(x)| = \text{máx}_{1\leq i\leq n, |\varepsilon|=1} g(x+\varepsilon\delta e_i) - g(x) < \rho$$

y así g es continua en x, para todo $x \in \mathbb{R}^n$.
Pasemos ahora a estudiar su diferenciabilidad. Para $k \in \mathbb{N}, 1 \leq i \leq n$ y $t > 0$, construimos

$$O_{k,i}(t) = \left\{x \in \mathbb{R}^n : \frac{g(x+te_i) + g(x-te_i) - 2g(x)}{t} < \frac{1}{k}\right\}$$

$$V_{k,i} = \bigcup_{t>0} O_{k,i}(t).$$

Como g es continua, los conjuntos $O_{k,i}(t)$ son abiertos por ser preimágenes continuas de $(-\infty, \frac{1}{k})$, que es abierto. Así, los conjuntos $V_{k,i}$ son también abiertos por ser unión de conjuntos abiertos. Gracias al Lema 3.6 y a la monotonía que nos proporciona el Lema 3.4, se tiene que

$$\Delta_i := \left\{ x \in \mathbb{R}^n : \frac{\partial g}{\partial x_i}(x) \text{ existe} \right\} = \bigcap_{k \geq 1} V_{k,i}.$$

Si existiese un abierto $U \subseteq \mathbb{R}^n$ tal que $U \cap \Delta_i = \emptyset$, en particular, dado $x \in U$ habría un conjunto infinito no numerable de puntos de la forma $(x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n)$ contenidos en U y por tanto cuya derivada parcial $\frac{\partial g}{\partial x_i}$ no existiría. Considerando ahora la función $f_i(t) = g(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$, que sería no derivable en un conjunto infinito no numerable, se contradiría la Proposición 3.5 (4). Así pues, concluimos que Δ_i es denso en $\mathbb{R}^n$, para todo $i = 1, \ldots, n$ y en consecuencia los $V_{k,i}$ también lo son (ya que $\Delta_i \subseteq V_{k,i}$). Sea

$$\Omega_g = \bigcap_{1 \leq i \leq n} \Delta_i.$$

Como Δ_i es intersección numerable de conjuntos abiertos y densos ($V_{k,i}$), se tiene que Ω_g también lo es, y por el Teorema de Baire 1.1, concluimos que Ω_g es denso en $\mathbb{R}^n$.
Tomemos $x \in \Omega_g$ y definamos la función $G(y) = g(y) - g(x) - \sum_{i=1}^{n} (y_i - x_i) \frac{\partial g}{\partial x_i}(x)$. Por construcción de Ω_g resulta que para cada $i = 1, \ldots, n$, dado $\delta > 0$, existe θ_i tal que

$$|G(x + te_i)| = \left| g(x + te_i) - g(x) - t \frac{\partial g}{\partial x_i}(x) \right| \leq \delta |t|, \quad \forall |t| < \theta_i.$$

Tomando $\theta = \min\{\theta_1, \ldots, \theta_n\}$ y notando que de nuevo la función G y el conjunto $C_\theta = \{h \in \mathbb{R}^n : \|h\|_1 \leq \theta\}$ cumplen las hipótesis del Lema 3.7, tenemos que

$$|G(x+h)| \leq \max_{1 \leq i \leq n, |\varepsilon| = 1} G(x + \varepsilon \|h\|_1 e_i) \leq \delta \|h\|_1, \quad \forall h \in \mathbb{C}_\theta.$$

Así, $\lim_{h \to 0} \frac{|G(x+h)|}{\|h\|_1} = 0$, luego g es diferenciable en x. Y obviamente si g es diferenciable en x, entonces $x \in \Omega_g$. Por tanto, Ω_g es el conjunto de puntos de $\mathbb{R}^n$ donde g es diferenciable, y dicho conjunto es denso, como queríamos probar.

□

Vamos ahora a calcular explícitamente un conjunto de diferenciabilidad Ω_g como el de la Proposición 3.8, en concreto para la función $g = \| \cdot \|_1$.

Ejemplo 4. Sea la función $g(x) = \|x\|_1$, para $x \in \mathbb{R}^n$. Por ser una norma, tenemos que dicha función es convexa. Así pues, estamos en las condiciones de la Proposición 3.8 y como hemos visto en la demostración, x es diferenciable si y solo si existe $\frac{\partial g}{\partial x_i}(x)$ para $i = 1, \ldots, n$. Para $i \in \{1, \ldots, n\}$, tenemos que

$$\lim_{t \to 0} \frac{g(x + te_i) - g(x)}{t} = \lim_{t \to 0} \frac{|x_i + t| + \sum_{j \neq i} |x_j| - \sum_{k=1}^{n} |x_k|}{t} = \lim_{t \to 0} \frac{|x_i + t| - |x_i|}{t}$$

y sabemos que el último límite existe si y solo si $x_i \neq 0$. Por tanto, $\frac{\partial g}{\partial x_i}(x)$ existe si y solo si $x_i \neq 0$, para $i = 1, \ldots, n$. De esta manera concluimos que

$$\Omega_g = \{x = (x_1, \ldots, x_n) \mid x_i \neq 0, \ i = 1, \ldots, n\},$$

que en efecto es denso en $\mathbb{R}^n$.

Uno de los ejemplos de funciones convexas más sencillos de encontrar en espacios normados, son en efecto las normas. Veamos alguna propiedad sobre normas en dimensión finita, que se enuncian en [6].

Sea $\|\cdot\|$ una norma en $\mathbb{R}^n$. Denotamos por $\Omega_{\|\cdot\|}$ el conjunto de puntos donde dicha norma es diferenciable, que es un abierto denso gracias a la Proposición 3.8. Si $x \in \Omega_{\|\cdot\|}$ denotamos por $\partial_{\|x\|} : \mathbb{R}^n \to \mathbb{R}$ la diferencial de la norma en x. Así,

$$\partial_{\|x\|}(x) = \lim_{t\to 0} \frac{\|x+tx\| - \|x\|}{t} = \lim_{t\to 0} \frac{(|1+t|-1)}{t}\|x\| = \|x\|.$$

Por otro lado, para cada $y \in X$ se tiene

$$|\partial_{\|x\|}(y)| = \lim_{t\to 0} \frac{|\|x+ty\| - \|x\||}{|t|} \leq \lim_{t\to 0} \frac{\|x+ty-x\|}{|t|} = \|y\|$$

luego es claro que

$$\|\partial_{\|x\|}\| = \sup_{y\neq 0} \frac{|\partial_{\|x\|}(y)|}{\|y\|} = \partial_{\|x\|}\left(\frac{x}{\|x\|}\right) = 1. \tag{3.1}$$

Por otra parte, sea $(x_n)_n \subseteq \Omega_{\|\cdot\|}$ una sucesión convergiendo a $z \in \mathbb{R}^n$. Si $t > 0$ entonces

$$\begin{aligned}\left|\frac{\|x_n+tz\| - \|x_n\|}{t} - \|z\|\right| &= \frac{\|x_n\| + \|tz\| - \|x_n+tz\|}{t} \leq \frac{\|x_n\| + \|tz-tx_n\| + \|tx_n\| - \|x_n+tz\|}{t}\\ &= \frac{\|(1+t)x_n\| - \|x_n+tz\| + t\|z-x_n\|}{t} \leq 2\|z-x_n\|,\end{aligned}$$

donde la última desigualdad se debe a la desigualdad triangular entre los dos primeros términos del numerador. Así, es claro que la convergencia es uniforme en t y justificamos (como se ve en [11], pág. 159-160) el intercambio del orden de los siguientes límites:

$$\lim_{n\to\infty} \partial_{\|x_n\|}(z) = \lim_{n\to\infty}\lim_{t\to 0^+} \frac{\|x_n+tz\| - \|x_n\|}{t} = \lim_{t\to 0^+}\lim_{n\to\infty} \frac{\|x_n+tz\| - \|x_n\|}{t} = \lim_{t\to 0^+} \frac{\|z+tz\| - \|z\|}{t} = \|z\|.$$

En resumen, obtenemos que

$$\lim_{n\to\infty} \partial_{\|x_n\|}(z) = \|z\|. \tag{3.2}$$

Observación 4. Como vimos en la prueba de la Proposición 3.8, $x \in \Omega_{\|\cdot\|} \neq \emptyset$ si y solo si existe $\frac{\partial\|\cdot\|}{\partial x_i}(x)$ para $i = 1,\dots,n$. Además, si $x \in \Omega_{\|\cdot\|}$, $x \neq 0$ y tomamos $z = \frac{1}{\|x\|}x$, se tiene que para cualquier $i \in \{1,\dots,n\}$

$$\lim_{t\to 0} \frac{\|z+te_i\| - \|z\|}{t} = \lim_{t\to 0} \frac{\|(x+t\|x\|e_i)\| - \|x\|}{t\|x\|} = \lim_{u\to 0} \frac{\|x+ue_i\| - \|x\|}{u} = \frac{\partial\|\cdot\|}{\partial x_i}(x),$$

aplicando el cambio $u = t\|x\|$. Por tanto existen $\frac{\partial\|\cdot\|}{\partial x_i}(z)$, $i = 1,\dots,n$. En definitiva, $z \in \Omega_{\|\cdot\|}$ luego podemos tomar, siempre que nos sea conveniente, $x \in \Omega_{\|\cdot\|}$ con $\|x\| = 1$.

Lema 3.9. *Sea E un espacio normado de dimensión finita, con norma $\|\cdot\|$. Tomamos $x \in \Omega_{\|\cdot\|}$ con $\|x\| = 1$. Entonces, $\partial_{\|x\|}$ es la única función 1-Lipschitz $\varphi : E \to \mathbb{R}$ tal que $\varphi(tx) = t$, para todo $t \in \mathbb{R}$.*

Demostración. Notemos primero que gracias a (3.1) ,

$$|\partial_{\|x\|}(y) - \partial_{\|x\|}(z)| = |\partial_{\|x\|}(y-z)| \leq \|\partial_{\|x\|}\|\|y-z\| = \|y-z\|, \quad \forall y,z \in E$$

y además

$$\partial_{\|x\|}(tx) = t\partial_{\|x\|}(x) = t\|x\| = t, \quad \forall t \in \mathbb{R}.$$

Luego $\partial_{\|x\|}$ cumple lo que queremos. Ahora probemos la unicidad. Para ello sea φ con las condiciones del enunciado, y sea $y \in E$. Tenemos que

$$1 = \left|t\varphi(y) - t\varphi\left(\left(\varphi(y) + \frac{1}{t}\right)x\right)\right| \leq |t|\left\|y - \varphi(y)x - \frac{1}{t}x\right\| = \|x - t(y-\varphi(y)x)\|, \quad \forall t \neq 0.$$

Luego la función $h(t) = \|x - t(y - \varphi(y)x)\|$ alcanza su mínimo en $t = 0$ ya que $\|x\| = 1$. Gracias a que $x \in \Omega_{\|\cdot\|}$, la regla de la cadena nos garantiza que h es derivable en 0 y por tanto

$$0 = h'(0) = \partial_{\|x\|}(y - \varphi(y)x) = \partial_{\|x\|}(y) - \varphi(y)\partial_{\|x\|}(x) = \partial_{\|x\|}(y) - \varphi(y),$$

obteniendo la unicidad que buscábamos. □

3.2. Topología débil* y el Teorema de Figiel

Para poder continuar con el estudio de los teoremas de este capítulo necesitamos introducir el concepto de *topología débil** que definiremos a continuación. Esta primera parte se ha basado en [13].

Definición 14. Sea X^* el espacio dual de las aplicaciones lineales y continuas de X en $\mathbb{R}$. Construimos la *topología débil** como aquella topología en X^* que tiene por base de entornos de $x_0^* \in X^*$ la formada por los conjuntos

$$U(x_0^*; \varepsilon, x_1, \ldots, x_n) := \left\{x^* \in X^* : |x^*(x_j) - x_0^*(x_j)| < \varepsilon \text{ para } j = 1, \ldots, n\right\},$$

donde $x_1, \ldots, x_n$ son un número finito de elementos de X y $\varepsilon > 0$.

Veamos que en efecto esta base define una topología. Se cumplen las siguientes propiedades.

I. Es obvio que $x_0^* \in U(x_0^*; \varepsilon, x_1, \ldots, x_n)$ para cualesquiera $\varepsilon > 0$, $x_1, \ldots, x_n \in X$.

II. Sean $U(x_0^*; \varepsilon_1, x_1, \ldots, x_n)$, $U(x_0^*; \varepsilon_2, y_1, \ldots, y_m)$ dos entornos básicos de $x_0^* \in X$. Entonces es claro que
$$U(x_0^*; \varepsilon, x_1, \ldots, x_n, y_1, \ldots, y_m) \subseteq U(x_0^*; \varepsilon_1, x_1, \ldots, x_n) \cap U(x_0^*; \varepsilon_2, y_1, \ldots, y_m)$$
donde $\varepsilon = \text{mín}\{\varepsilon_1, \varepsilon_2\} > 0$.

III. Sea $U(x_0^*; \varepsilon, x_1, \ldots, x_n)$ entorno de x_0^*. Entonces para cada $y^* \in U(x_0^*; \varepsilon/2, x_1, \ldots, x_n)$ se tiene que $U(y^*; \varepsilon/2, x_1, \ldots, x_n) \subseteq U(x_0^*; \varepsilon, x_1, \ldots, x_n)$ gracias a la desigualdad triangular.

Por tanto, se tiene que dicha base de entornos define una única topología en X^*, la topología débil*. Además, gracias al Teorema 1.3 es fácil ver que la topología débil* es de hecho Hausdorff.

Observación 5. Vamos ahora a caracterizar la convergencia de sucesiones en esta topología.
Si tenemos $(x_n^*)_n$ una sucesión en X^* entonces es convergente (en la topología débil*) a $x^* \in X^*$ si y solo si para cualquier entorno U de la base existe un $n_0 \in \mathbb{N}$ tal que $x_n^* \in U$ para $n \geq n_0$. Esto es equivalente a que para cada $x \in X$ y cada $\varepsilon > 0$ exista $n_0 \in \mathbb{N}$ tal que $|x_n^*(x) - x^*(x)| \leq \varepsilon$ para $n \geq n_0$. Y a su vez, esto último equivale a que para todo $x \in X$, la sucesión $(x_n^*(x))_n$ converja a $x^*(x)$.
En definitiva, $(x_n^*)_n$ converge a $x^* \in X^*$ en la topología débil* si y solo si $(x_n^*(x))_n$ converge a $x^*(x)$ para todo $x \in X$.

Considerando entonces la topología débil* podemos obtener un resultado que nos será de mucha utilidad en lo que sigue y que no es cierto, si X es infinito dimensional, con la topología de la norma.

Teorema 3.10 (Alaoglu)**.** *Sea el dual X^* de un espacio normado X. Entonces la bola (cerrada) unidad $B_{X^*} \subseteq X^*$ es débil*-compacto.*

Demostración. Incluimos la prueba en el Apéndice B (Teorema B.1). □

Ahora sí podemos demostrar el resultado más relevante de esta sección y que se atribuye a Figiel. En primer lugar, lo probaremos para dimensión finita en el siguiente lema, para después generalizarlo a dimensión infinita en un teorema posterior.

Lema 3.11 (Figiel, 1968 [6]). *Supongamos que* $\dim(X) < \infty$ *y sea* $\phi : X \to Y$ *una isometría con* $\phi(0) = 0$. *Supongamos también que* $\overline{\text{span}}(\phi(X)) = Y$. *Entonces existe una única función lineal y continua* $T : Y \to X$ *tal que* $T \circ \phi = Id_X$ *y* $\|T\| = 1$.

Demostración. Consideramos primero el caso $\dim(X) = 1$. Como X es isométrico a $\mathbb{R}$ podemos asumir que ϕ está definida en $\mathbb{R}$. Para cada $k \in \mathbb{N}$ se tiene que

$$\|\phi(k) - \phi(-k)\| = |k - (-k)| = 2k \neq 0.$$

Luego gracias al Corolario 1.4 podemos asegurar la existencia de $y_k^* \in Y^*$ con $\|y_k^*\| = 1$ y tal que $y_k^*(\phi(k) - \phi(-k)) = 2k$. Ahora bien, para $0 \leq t \leq k$,

$$\begin{aligned} 2k = y_k^*(\phi(k) - \phi(-k)) &= y_k^*(\phi(k) - \phi(t)) + y_k^*(\phi(t) - \phi(0)) + y_k^*(\phi(0) - \phi(-k)) \\ &\leq \|\phi(k) - \phi(t)\| + \|\phi(t) - \phi(0)\| + \|\phi(0) - \phi(-k)\| = k - t + t - 0 + k - 0 = 2k. \end{aligned}$$

Luego las desigualdades deben ser igualdades y por tanto $y_k^*(\phi(t)) = t$. Análogamente se cumple para $-k \leq t \leq 0$, luego

$$y_k^*(\phi(t)) = t, \quad \forall t \in [-k, k].$$

Gracias al Teorema 3.10 y al hecho de que cualquier sucesión contenida en un compacto tiene un punto de acumulación, tenemos que existe $y^* \in B_{Y^*}$ punto de acumulación de la sucesión $(y_k^*)_k \subseteq B_{Y^*}$. Teniendo en cuenta la definición de entorno débil*, esto supone que para cada $t \in \mathbb{R}$ y para cualquier $\varepsilon > 0$ podemos encontrar $k \geq |t|$ tal que

$$|t - y^*(\phi(t))| = |y_k^*(\phi(t)) - y^*(\phi(t))| < \varepsilon.$$

Y como ε es arbitrario, $y^*(\phi(t)) = t$, $\forall t \in \mathbb{R}$ y tenemos lo que buscamos tomando $T = y^*$.
Pasamos ahora al caso general considerando la función ϕ del enunciado. Tomamos $x \in \Omega_{\|\cdot\|}$ con $\|x\| = 1$ (gracias a la observación 4) y consideramos $j(t) = \phi(tx)$, $\forall t \in \mathbb{R}$. Es claro que j es una isometría con $j(0) = 0$ luego por el caso unidimensional existe $f_x^* \in B_{Y^*}$ tal que $f_x^*(\phi(tx)) = f_x^*(j(t)) = t$, $\forall t \in \mathbb{R}$. Además

$$\|f_x^*(\phi(z)) - f_x^*(\phi(y))\| = \|f_x^*(\phi(z) - \phi(y))\| \leq \|\phi(z) - \phi(y)\| = \|z - y\|, \quad \forall z, y \in X.$$

Por tanto la función $f_x^* \circ \phi$ cumple las hipótesis del Lema 3.9 y gracias a la unicidad se tiene $\partial_{\|x\|} = f_x^* \circ \phi$. Por otro lado, se sigue de la Proposición 3.8 y de (3.2) que para cualquier $z \in X \setminus \{0\}$ existe $x' \in \Omega_{\|\cdot\|}$ tal que $\partial_{\|x'\|}(z) \neq 0$. Entonces, haciendo $z_1 = z$, $x_1 = x'$ podemos tomar inductivamente $z_r \in \ker(\partial_{\|x_1\|}) \cap \cdots \cap \ker(\partial_{\|x_{r-1}\|})$ para $2 \leq r \leq n$ ya que $\dim(\ker(\partial_{\|x_1\|}) \cap \cdots \cap \ker(\partial_{\|x_{r-1}\|})) = n - r$ por ser intersección de subespacios de X de dimensión $n - 1$. En consecuencia, eligiendo también de manera inductiva x_r tal que $\partial_{\|x_r\|}(z_r) \neq 0$ para $1 \leq r \leq n$, obtenemos por construcción que $(\partial_{\|x_r\|})_{1 \leq r \leq n}$ es base de X^*. Por las propiedades de espacio dual, sabemos que existe $(\omega_i)_{1 \leq i \leq n}$ base de $X^{**} = X$ tal que

$$\partial_{\|x_r\|}(\omega_i) = \delta_{i,r} := \begin{cases} 1 & \text{si} \quad i = r \\ 0 & \text{si} \quad i \neq r. \end{cases}$$

Y como hemos visto que para cada $1 \leq r \leq n$ existe $f_{x_r}^* \in Y^*$ cumpliendo $\partial_{\|x_r\|} = f_{x_r}^* \circ \phi$, definimos la función $T : Y \to X$ de la siguiente manera:

$$T(y) = \sum_{i=1}^{n} f_{x_i}^*(y)\omega_i.$$

La función T es obviamente lineal y continua y cumple que

$$T(\phi(y)) = \sum_{i=1}^{n} \partial_{\|x_i\|}(y)\omega_i = \sum_{i=1}^{n} (a_1 \partial_{\|x_i\|}(\omega_1) + \cdots + a_n \partial_{\|x_i\|}(\omega_n))\omega_i = \sum_{i=1}^{n} a_i \omega_i = y,$$

donde hemos usado la descomposición de y en términos de la base $(\omega_i)_{1 \leq i \leq n}$ con coeficientes $a_1, \ldots, a_n$ reales.

Además, si existiese otra función T' cumpliendo las mismas condiciones, se tendría que T y T' coincidirían por construcción en $\phi(X)$. Como ambas son lineales coincidirían también en $\text{span}(\phi(X))$ y puesto que son continuas serían iguales en $\overline{\text{span}}(\phi(X)) = Y$. Luego queda probada la unicidad.
Por último, para todo $x' \in \Omega_{\|\cdot\|}$ tenemos que

$$f_{x'}^* = \partial_{\|x'\|} \circ T$$

ya que son ambas funciones continuas que coinciden en el conjunto denso $\text{span}(\phi(X))$. Así, por un lado,

$$1 = \|f_{x'}^*\| \leq \|\partial_{\|x'\|}\| \|T\| = \|T\|.$$

Y por otro lado, si le aplicamos (3.2) a $z = T(y)$ para algún $y \in Y$, obtenemos lo siguiente:

$$\|T(y)\| = \lim_{n\to\infty} \partial_{\|x_n\|}(T(y)) = \lim_{n\to\infty} f_{x_n}^*(y) \leq \lim_{n\to\infty} \|f_{x_n}^*\| \|y\| = \|y\|.$$

Es decir, $\|T\| \leq 1$. Por tanto, necesariamente $\|T\| = 1$. □

Antes de continuar, veamos un ejemplo que corresponde al Lema 3.11.

Ejemplo 5. Retomemos el Ejemplo 1. Ya vimos que la función $U : \mathbb{R} \to \mathbb{R}^2$ tal que $U(x) = (x, \sin(x))$ era una isometría con $U(0) = 0$. Además, $U(\pi) = (\pi, 0)$ y $U(\frac{\pi}{2}) = (\frac{\pi}{2}, 1)$ son linealmente independientes en $\mathbb{R}^2$ y por tanto $\overline{\text{span}}(U(\mathbb{R})) = \mathbb{R}^2 = Y$. Por tanto, sabemos por el Lema 3.11 que existe una única función lineal y continua T tal que $T \circ U = Id_{\mathbb{R}}$ y $\|T\| = 1$. De hecho, es fácil ver que en este caso dicha función es la proyección sobre la primera coordenada, es decir,

$$T(x,y) = x, \quad \forall (x,y) \in \mathbb{R}^2.$$

Para verlo, nota que para $x, y \in \mathbb{R}$ se cumple que

$$\frac{|T(x,y)|}{\|(x,y)\|_\infty} = \frac{|x|}{\|(x,y)\|_\infty} \leq 1,$$

y además $\frac{|T(1,0)|}{\|(1,0)\|_\infty} = 1$ luego $\|T\| = 1$. Por último, es obvio que $T \circ U = Id_{\mathbb{R}}$.

Teorema 3.12 (Figiel, 1968 [6]). *Supongamos que X es Banach, separable y de dimensión infinita. Sea $\phi : X \to Y$ con $\phi(0) = 0$ y asumimos que $\overline{\text{span}}(\phi(X)) = Y$. Entonces existe una única función lineal y continua $T : Y \to X$ tal que $T \circ \phi = Id_X$ y $\|T\| = 1$.*

Demostración. Si $D = \{v_n \mid n \in \mathbb{N}\}$ es un denso de X, que existe por la hipótesis de separabilidad, tomemos $E_n = \text{span}(v_1, \ldots, v_n)$ para cada $n \in \mathbb{N}$. De esta forma $(E_n)_n$ es una sucesión creciente de subespacios de dimensión finita y $X = \overline{\bigcup_{n\in\mathbb{N}} E_n}$.
Tomando $F_n = \text{span}(\phi(E_n))$, por el Lema 3.11 existe $T_n : F_n \to E_n$ tal que $T_n \circ \phi = Id_{E_n}$ y $\|T_n\| = 1$. Además, gracias a la unicidad, podemos definir $T : \bigcup_{n\in\mathbb{N}} F_n \to X$ tal que $T(y) = T_n(y)$, si $y \in F_n$. Como $\|T_n\| = 1$ para todo $n \in \mathbb{N}$ se tiene que $\|T\| = 1$.
Finalmente como $\overline{\text{span}}(\phi(X)) = Y$ tenemos que

$$Y = \overline{\text{span}}(\phi(\overline{\cup_{n\in\mathbb{N}} E_n})) \subseteq \overline{\text{span}}\overline{(\phi(\cup_{n\in\mathbb{N}} E_n))} = \overline{\text{span}}\overline{(\cup_{n\in\mathbb{N}} \phi(E_n))},$$

usando que ϕ es continua y la propiedad $\phi(\overline{A}) \subseteq \overline{\phi(A)}$ para un conjunto A cualquiera. Por otro lado, $\overline{\text{span}}\overline{(\cup_{n\in\mathbb{N}} \phi(E_n))} \subseteq \overline{\text{span}}(\cup_{n\in\mathbb{N}} \phi(E_n))$, ya que las sucesiones se comportan bien con las combinaciones lineales finitas. El contenido es de hecho una igualdad y por tanto

$$Y = \overline{\text{span}}(\cup_{n\in\mathbb{N}} \phi(E_n)) = \overline{\{\sum_{i=1}^{r} \lambda_i x_i \mid \text{para algún } r \in \mathbb{N}, \text{ unos } \lambda_i \in \mathbb{R}, \text{ y con } x_i \in \phi(E_n) \text{ para algún } n \in \mathbb{N}\}}$$

y gracias a que $(E_n)_n$ es creciente,

$$Y = \overline{\bigcup_{n\in\mathbb{N}} \operatorname{span}(\phi(E_n))} = \overline{\bigcup_{n\in\mathbb{N}} F_n} \tag{3.3}$$

Es claro que la sucesión de espacios vectoriales $(F_n)_n$ es creciente, por tanto $F = \cup_{n\in\mathbb{N}} F_n$ es también un espacio vectorial. Entonces como T es lineal, se tiene que para $x, y \in F$,

$$\|T(x) - T(y)\| = \|T(x-y)\| \leq \|T\|\|x-y\| = \|x-y\|$$

luego T es 1-Lipschitz. Así pues, si $(y_n)_n \subseteq F$ es convergente a $y \in Y$, se tiene que $(T(y_n))_n$ es convergente por ser de Cauchy y X completo. Además, si tenemos otra sucesión $(z_n)_n \subseteq F$ que converge al mismo punto y, tenemos que

$$\lim_{n\to\infty} T(y_n) - T(z_n) = \lim_{n\to\infty} T(y_n - z_n) = T\left(\lim_{n\to\infty} y_n - z_n\right) = T(0) = 0$$

y por tanto $\lim_{n\to\infty} T(y_n) = \lim_{n\to\infty} T(z_n)$.
De esta forma y gracias a (3.3) podemos extender la función T a todo Y, demostrando así el teorema. □

Observación 6. El Teorema 3.12 implica directamente el Teorema de Mazur-Ulam 2.3 por lo siguiente. Si ϕ es sobreyectiva, necesariamente T debe ser inyectiva. Además, para cualesquiera $x, y \in X$, $\lambda \in \mathbb{R}$,

$$T(\phi(x) + \phi(y)) = T(\phi(x)) + T(\phi(y)) = x + y = T(\phi(x+y))$$
$$T(\lambda\phi(x)) = \lambda T(\phi(x)) = \lambda x = T(\phi(\lambda x))$$

y la inyectividad de T nos proporciona la linealidad de ϕ.

3.3. Integral de Bochner y el Teorema de Godefroy-Kalton

Para poder demostrar los dos últimos teoremas de este trabajo necesitamos saber integrar funciones que toman valores en un espacio de Banach cualquiera. Para ello, definimos la *integral de Bochner*.

Esta sección se ha basado principalmente en [12].

Definición 15. Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida e Y un espacio de Banach. Decimos que $s : \Omega \to Y$ es una función *simple* si existen $n \in \mathbb{N}$, $x_1, \ldots, x_n \in Y$ y $E_1, \ldots, E_n \subseteq \Omega$ disjuntos y de medida μ finita, tal que

$$s(x) = \sum_{i=1}^{n} x_i 1_{E_i}(x), \quad \text{para todo } x \in \Omega.$$

Además dada una función simple s, definimos su integral como

$$\int_X s \, \mathrm{d}\mu := \sum_{i=1}^{n} x_i \mu(E_i).$$

Definición 16. Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida e Y un espacio de Banach. Dada una función medible $f : \Omega \to Y$, decimos que es *Bochner integrable* si existe una sucesión $(s_n)_n$ de funciones simples de Ω en Y tales que

$$\lim_{n\to\infty} \left(\int_\Omega \|f - s_n\| \, \mathrm{d}\mu \right) = 0.$$

En ese caso, definimos la integral de f como

$$\int_\Omega f \, \mathrm{d}\mu := \lim_{n\to\infty} \left(\int_\Omega s_n \, \mathrm{d}\mu \right).$$

Además usaremos una serie de propiedades sobre la integral de Bochner que se enuncian a continuación y cuya prueba aparece en el Apéndice B (Lema B.2).

Lema 3.13 (Propiedades de la integral de Bochner)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, Y un espacio de Banach y sea $f : \Omega \to Y$ medible.*

I. *Si f es Bochner integrable, entonces para cualquier sucesión $(s_n)_n$ de funciones simples tal que $\lim_{n\to\infty} \int_\Omega \|s_n - f\|\, d\mu = 0$, existe el límite $\lim_{n\to\infty} (\int_\Omega s_n\, d\mu)$ y su valor es independiente de la sucesión $(s_n)_n$ escogida.*

II. *Si f es Bochner integrable, se cumple que $\| \int_\Omega f\, d\mu\| \leq \int_\Omega \|f\|\, d\mu$.*

III. *Si Ω es compacto y Hausdorff, μ es una medida finita Borel, y f es continua, entonces f es Bochner integrable.*

IV. *Si f es Bochner integrable, Z es un espacio de Banach y $T : Y \to Z$ una función lineal y acotada, entonces $T \circ f$ es Bochner integrable y $\int_\Omega T \circ f\, d\mu = T(\int_\Omega f\, d\mu)$.*

Demostración. Incluimos la prueba en el Apéndice B (Lema B.2). ☐

Observación 7. Gracias a la propiedad (IV) del Lema 3.13, y al hecho de que las proyecciones en una coordenada son funciones lineales y continuas, se tiene que si $f : \Omega \to \mathbb{R}^n$ es Bochner integrable con $f = (f_1, \ldots, f_n)$ entonces su integral es

$$\int_\Omega f\, \mathrm{d}\mu = \left(\int_\Omega f_1 \mathrm{d}\mu, \ldots, \int_\Omega f_n \mathrm{d}\mu \right).$$

Ahora sí, podemos enunciar uno de los resultados más importantes de este trabajo, y cuya demostración es de hecho constructiva, potenciando más aún su utilidad.

Teorema 3.14. *Sean X, Y espacios de Banach tal que X es separable y sea $f : X \to Y$ una función K-Lipschitz para algún $K > 0$. Entonces existe una función lineal y acotada $F : X \to Y$ satisfaciendo que*

I. *$\|F\| \leq K$,*

II. *para todo espacio de Banach Z y toda función lineal y acotada $T : Y \to Z$, si $T \circ f$ es lineal entonces $T \circ f = T \circ F$.*

Demostración. Sea $\{x_i : i \in I\}$ un conjunto contable de vectores linealmente independientes en X tal que $\overline{\operatorname{span}}(\{x_i : i \in I\}) = X$ y tales que $\sum_{i \in I} \|x_i\| < \infty$. Es suficiente considerar el caso $\operatorname{span}(\{x_i : i \in I\}) = X$ ya que podemos extender F por continuidad (F es lineal y acotada, luego continua).
Sea $H = \prod_{i \in I} \left[-\frac{1}{2}, \frac{1}{2}\right]$, y para cada $i \in I$ denotamos $H_i = \prod_{j \in I, j \neq i} \left[-\frac{1}{2}, \frac{1}{2}\right]$. Sean μ y μ_i las medidas definidas en H y H_i respectivamente, construidas como el producto de las medidas de Lebesgue en cada factor. Esta construcción se realiza de la siguiente forma. Partimos de los espacios de medida $(\left[-\frac{1}{2}, \frac{1}{2}\right], \beta, \lambda)$, donde λ y β son la medida Lebesgue y la σ-álgebra de los borelianos respectivamente, restringidas a $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Entonces, consideramos la σ-álgebra $\mathscr{A}$ engendrada por los conjuntos $A = \prod_{i \in I} A_i$, donde $A_i \in \beta$ para todo $i \in I$ y además todos los conjuntos A_i son $\left[-\frac{1}{2}, \frac{1}{2}\right]$ excepto para un número finito de ellos. Entonces definimos $\mu(A) = \prod_{i \in I} \lambda(A_i)$ y como $\left[-\frac{1}{2}, \frac{1}{2}\right]$ tiene medida 1 ese producto está perfectamente definido ya que de hecho es un producto finito. Además, se puede probar que μ se extiende a todo $\mathscr{A}$ de manera natural (para más detalles consultar [8], pág. 154-160). De la misma manera se razona con μ_i. Nota entonces que ambos conjuntos H y H_i tienen medida igual a 1. Además, a los elementos de H los denotaremos como $\vec{t} = (t_i)_{i \in I} \in H$. Indistintamente denotaremos también de la misma forma a los elementos de H_i.

Es claro que H_i con la topología producto es Hausdorff y compacto por el Teorema de Tychonoff, y que f es continua por ser Lipschitz. Así, gracias al Lema 3.13 (III) podemos, para cada $i \in I$, definir

$$F(x_i) = \int_{H_i} f\left(\sum_{j \neq i} t_j x_j + \frac{1}{2} x_i \right) \mathrm{d}\mu_i(\vec{t}) - \int_{H_i} f\left(\sum_{j \neq i} t_j x_j - \frac{1}{2} x_i \right) \mathrm{d}\mu_i(\vec{t})$$

y extender F linealmente a X. Así, para cualquier Z y T como en la condición (II) con $T \circ f$ lineal se tiene que

$$\begin{aligned} T(F(x_i)) &= \int_{H_i} \left(T\left(f\left(\sum_{j\neq i} t_j x_j + \frac{1}{2} x_i \right) \right) - T\left(f\left(\sum_{j\neq i} t_j x_j - \frac{1}{2} x_i \right) \right) \right) \mathrm{d}\mu_i(\vec{t}) \\ &= \int_{H_i} T\left(f\left(\sum_{j\neq i} t_j x_j + \frac{1}{2} x_i - \sum_{j\neq i} t_j x_j + \frac{1}{2} x_i \right) \right) \mathrm{d}\mu_i(\vec{t}) \\ &= \int_{H_i} T\big(f(x_i)\big) \mathrm{d}\mu_i(\vec{t}) = T\big(f(x_i)\big), \quad \text{para cada } i \in I, \end{aligned}$$

donde la primera igualdad es gracias al Lema 3.13 (IV), la segunda se debe a la linealidad de $T \circ f$, y la última se justifica con que la medida de H_i es igual a 1. Así, y gracias a que $T \circ f$ y $T \circ F$ son ambas lineales, es claro que $T \circ f = T \circ F$ en todo X.

Vamos ahora a demostrar el apartado (I). Para ello fijamos un elemento de X que por tanto es de la forma $a = \sum_{i \in J} a_i x_i$, donde J es un subconjunto finito y no vacío de I con digamos N elementos, y $a_i \in \mathbb{R}$ para todo $i \in J$. Vamos a probar que $\|F(a)\| \leq K\|a\|$, así que consideramos para $h > 0$, la integral

$$S_a(h) := \int_H \frac{f\left(\sum_{i\in I} t_i x_i + ha\right) - f\left(\sum_{i\in I} t_i x_i\right)}{h} \mathrm{d}\mu(\vec{t}).$$

Es claro que como f es K-Lipschitz, la norma del integrando está acotada por $K\|a\|$ y por tanto la norma de la integral también lo está para cualquier $h > 0$. Por otro lado, vamos a ver que dicha integral converge a $F(a)$ cuando h tiende a 0, y así tendremos probado el enunciado. En primer lugar, fijamos un $y^* \in B_{Y^*}$ y definimos $g = y^* \circ f$. Es fácil ver que g es también K-Lipschitz y ahora la imagen de g está en $\mathbb{R}$. De esta forma, podemos trabajar con g en lugar de f y así todas las técnicas que usemos para operar con integrales vendrán justificadas por lo que sabemos sobre la integral de Lebesgue. En consecuencia, consideraremos a partir de ahora la integral

$$y^*(S_a(h)) = \int_H \frac{g\left(\sum_{i\in I} t_i x_i + ha\right) - g\left(\sum_{i\in I} t_i x_i\right)}{h} \mathrm{d}\mu(\vec{t}). \tag{3.4}$$

Para demostrar la convergencia que necesitamos, realizaremos dos aproximaciones. Respecto a la primera de ellas, observemos que en la integral de (3.4) estamos integrando $g\left(\sum_{i\in I} t_i x_i\right)$ con $\vec{t}$ sobre el conjunto $H' = \prod_{i\in I}[-\frac{1}{2} + ha_i, \frac{1}{2} + ha_i]$ y restando la misma integral pero sobre el conjunto H. En definitiva, la integral que aparece en (3.4) se puede reescribir como

$$\int_{H'\cap \mathscr{C}H} \frac{g\left(\sum_{i\in I} t_i x_i\right)}{h} \mathrm{d}\mu(\vec{t}) - \int_{H\cap \mathscr{C}H'} \frac{g\left(\sum_{i\in I} t_i x_i\right)}{h} \mathrm{d}\mu(\vec{t}), \tag{3.5}$$

donde $\mathscr{C}H$ y $\mathscr{C}H'$ son los complementarios de H y H' respectivamente.

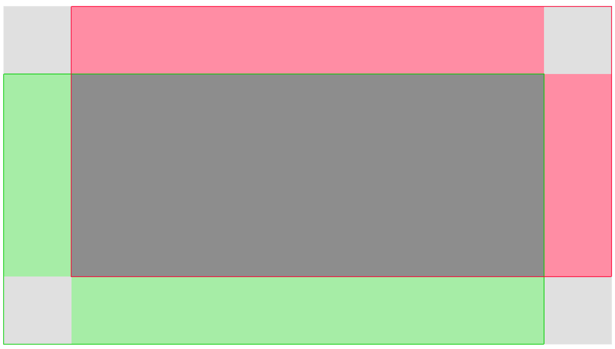

Figura 3.2: Conjuntos de Integración H y H' para el caso $\dim(X) = 2$.

Además, observamos que la integral sobre $H'\cap\mathscr{C}H$ la podemos expresar como la integral en la unión, con i en J, de los conjuntos

$$H_i\times[\frac{1}{2},\frac{1}{2}+ha_i],\quad \text{si } a_i\geq 0 \quad \text{ó} \quad H_i\times[-\frac{1}{2}+ha_i,-\frac{1}{2}],\quad \text{si } a_i<0$$

además de integrar sobre una cantidad de conjuntos finita y dependiente de N que llamaremos $C(N)$ y donde todos ellos tienen, en el peor de los casos, medida del orden de una $O(h^2)$. Esto se debe a que todos estos conjuntos tienen dos o más lados de longitud múltiplo de h. Análogamente ocurre lo mismo con $H\cap\mathscr{C}H'$, cuya integral la podemos escribir como integrar sobre la unión con i en J de

$$H_i\times[-\frac{1}{2},-\frac{1}{2}+ha_i],\quad \text{si } a_i\geq 0 \quad \text{ó} \quad H_i\times[\frac{1}{2}+ha_i,\frac{1}{2}],\quad \text{si } a_i<0$$

más integrar sobre una cantidad que llamamos $C'(N)$ de conjuntos de medida, a lo sumo, $O(h^2)$. Si observamos que g es K-Lipschitz y que todos los conjuntos que aparecen en la descomposición están acotados, se tiene que las integrales respecto a los conjuntos de medida $O(h^2)$ tienen un valor del orden de $O(h^2)$. Y puesto que hay una cantidad finita $(C(N)+C'(N))$ de ellas, el valor de la suma de esas integrales es una $O(h^2)$. Nota que en el caso 2-dimensional, estos conjuntos de medida $O(h^2)$ están representados en gris claro en la Figura 3.2. En rojo y verde están representados H' y H respectivamente, y en gris oscuro la intersección de estos últimos.
En definitiva, tenemos que la expresión en (3.5), para h cercano a 0, es equivalente a

$$\frac{1}{h}\sum_{i\in J}\left(\int_{H_i}\int_{\frac{1}{2}}^{\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)\mathrm{d}s\mathrm{d}\mu_i(\vec{t})-\int_{H_i}\int_{-\frac{1}{2}}^{-\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)\mathrm{d}s\mathrm{d}\mu_i(\vec{t})\right)+\frac{O(h^2)}{h}. \quad (3.6)$$

Observa que esta expresión se obtiene gracias a que invertir el orden de los límites de integración es equivalente a multiplicar por (-1). Para la segunda aproximación, vemos que para cada $i\in J$ se tiene

$$\begin{aligned}
&\left|\int_{H_i}\int_{\frac{1}{2}}^{\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)\mathrm{d}s\mathrm{d}\mu_i(\vec{t})-ha_i\int_{H_i} g\left(\sum_{j\neq i}t_jx_j+\frac{1}{2}x_i\right)\mathrm{d}\mu_i(\vec{t})\right|\\
&\leq\int_{H_i}\left|\int_{\frac{1}{2}}^{\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)\mathrm{d}s-ha_ig\left(\sum_{j\neq i}t_jx_j+\frac{1}{2}x_i\right)\right|\mathrm{d}\mu_i(\vec{t})\\
&=\int_{H_i}\left|\int_{\frac{1}{2}}^{\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)-g\left(\sum_{j\neq i}t_jx_j+\frac{1}{2}x_i\right)\mathrm{d}s\right|\mathrm{d}\mu_i(\vec{t})\\
&\leq\int_{H_i}K\|x_i\|\left|\int_{\frac{1}{2}}^{\frac{1}{2}+ha_i}\left(s-\frac{1}{2}\right)\mathrm{d}s\right|\mathrm{d}\mu_i(\vec{t})=K\|x_i\|\frac{h^2a_i^2}{2},
\end{aligned}$$

donde la primera desigualdad se sigue del Teorema de Fubini, y la última desigualdad es gracias a que g es K-Lipschitz. Calculando la integral en una variable real y observando que la medida de H_i es 1 se tiene el último paso. Con el mismo procedimiento, de nuevo para cada $i\in J$, se obtiene que

$$\left|\int_{H_i}\int_{-\frac{1}{2}}^{-\frac{1}{2}+ha_i} g\left(\sum_{j\neq i}t_jx_j+sx_i\right)\mathrm{d}s\mathrm{d}\mu_i(\vec{t})-ha_i\int_{H_i} g\left(\sum_{j\neq i}t_jx_j-\frac{1}{2}x_i\right)\mathrm{d}\mu_i(\vec{t})\right|\leq K\|x_i\|\frac{h^2a_i^2}{2}.$$

Notando que el término $K\|x_i\|\frac{h^2a_i^2}{2}$ es una $O(h^2)$ concluimos que para h cercano a 0, la integral de (3.6) es equivalente a

$$\frac{1}{h}\sum_{i\in J}ha_i\left(\int_{H_i} g\left(\sum_{j\neq i}t_jx_j+\frac{1}{2}x_i\right)\mathrm{d}\mu_i(\vec{t})-\int_{H_i} g\left(\sum_{j\neq i}t_jx_j-\frac{1}{2}x_i\right)\mathrm{d}\mu_i(\vec{t})\right)+\frac{O(h^2)}{h}.$$

Así, haciendo h tender a 0 y observando que y^* es continua y conmuta con las integrales, tenemos que

$$y^*\left(S_a(h)\right) = \int_H \frac{g\left(\sum_{i\in I} t_i x_i + ha\right) - g\left(\sum_{i\in I} t_i x_i\right)}{h} \mathrm{d}\mu(\vec{t}) \quad \xrightarrow{h\to 0} \quad \sum_{i\in J} y^*(a_i F(x_i)) = y^*(F(a)).$$

Además, es claro que para todo $h > 0$ se tiene

$$|y^*(S_a(h))| \leq \|y^*\| \|S_a(h)\| \leq \|S_a(h)\| \leq K\|a\|,$$

luego $|y^*(F(a))| \leq K\|a\|$. Finalmente como esto ocurre para toda $y^* \in B_{Y^*}$ obtenemos que

$$\|F(a)\| = \sup_{y^*\in B_{Y^*}} |y^*(F(a))| \leq K\|a\|.$$

Luego $\|F\| \leq K$ y tenemos demostrado el teorema.

□

Finalmente, enunciaremos el Teorema de Godefroy-Kalton que concluye esta sección con un resultado extraordinario.

Teorema 3.15 (Godefroy-Kalton, 2003 [6]). *Sean X,Y espacios de Banach, con X separable. Si existe una isometría $\phi : X \to Y$, entonces Y contiene un subespacio vectorial cerrado que es linealmente isométrico a X.*

Demostración. En primer lugar, podemos asumir que $\phi(0) = 0$ ya que en otro caso, la función $\varphi := \phi - \phi(0)$ es una isometría con $\varphi(0) = 0$. Además, también supondremos que $\overline{\text{span}}(\phi(X)) = Y$. Nota que si el contenido fuese estricto, puesto que $\overline{\text{span}}(\phi(X))$ es Banach podríamos aplicar este teorema, y como $\overline{\text{span}}(\phi(X))$ también es cerrado encontraríamos el subespacio de Y que buscamos cumpliendo el enunciado. Con estas consideraciones, gracias al Teorema 3.12, existe una función lineal $T : Y \to X$ de norma 1 y tal que $T \circ \phi = Id_X$. Ahora, podemos aplicar el Teorema 3.14 tomando $f = \phi$ y obtener una función lineal $F : X \to Y$ con $\|F\| \leq 1$ y $T \circ F = T \circ \phi = Id_X$. Luego necesariamente $\|F\| = 1$ y de hecho, F es una isometría ya que

$$\|F(x) - F(y)\| = \|F(x-y)\| \leq \|F\|\|x-y\| = \|x-y\|, \quad \forall x,y \in X$$
$$\|x-y\| = \|T(F(x-y))\| \leq \|T\|\|F(x)-F(y)\| = \|F(x)-F(y)\|, \quad \forall x,y \in X.$$

En definitiva, F es una isometría lineal entre X e Y, luego $F(X)$ es el subespacio vectorial de Y que satisface el enunciado. Nota que al ser X completo y F isometría, tenemos que $F(X)$ es completo, y además está contenido en Y que también es completo, luego necesariamente $F(X)$ es cerrado. □

Vamos a exponer un ejemplo sencillo en el que se ilustran los Teoremas 3.14 y 3.15.

Ejemplo 6. De nuevo, volvemos al Ejemplo 1, con la función $U : \mathbb{R} \to \mathbb{R}^2$ tal que $U(x) = (x, \sin(x))$. Por el Teorema 3.15 sabemos que existe una isometría lineal F de $\mathbb{R}$ en $\mathbb{R}^2$. De hecho no sólo eso, sino que además, el Teorema 3.14 nos proporciona el modo de calcularla. En este caso, $\{1\}$ es base de $\mathbb{R}$, así que F viene definida por su valor en 1, que es

$$F(1) = U\left(\frac{1}{2}\right) - U\left(-\frac{1}{2}\right) = \left(\frac{1}{2}, \sin\frac{1}{2}\right) - \left(-\frac{1}{2}, \sin\left(-\frac{1}{2}\right)\right) = \left(1, 2\sin\frac{1}{2}\right).$$

En definitiva, como F es lineal se tiene que

$$F(t) = tF(1) = \left(t, 2t\sin\frac{1}{2}\right), \quad \forall t \in \mathbb{R}.$$

Y en consecuencia, es claro que el conjunto $\left\{\left(t, 2t\sin\frac{1}{2}\right) \;:\; t \in \mathbb{R}\right\}$ es un subespacio vectorial cerrado de $\mathbb{R}^2$ linealmente isométrico a $\mathbb{R}$.

Observación 8. Cabe mencionar que el Teorema 3.15 no es cierto para el caso no separable. De hecho, un contraejemplo es el dado por la isometría $\delta : \ell_\infty \to \mathscr{F}(\ell_\infty)$, donde δ y $\mathscr{F}(\ell_\infty)$ están definidos en el Apéndice A. La prueba de este contraejemplo se puede consultar en [7].

Apéndice A

Espacios de funciones Lipschitz y espacios Lipschitz-libres

Los espacios Lipschitz-libres fueron introducidos por Godefroy y Kalton, y son uno de los temas de investigación más activos en la actualidad en el contexto del Análisis Funcional, con aplicaciones relevantes en la Teoría de la Computación y el Transporte Óptimo. Vamos a ver entonces en qué consisten y daremos algunas propiedades acerca de ellos. Para más información sobre este tema, consultar [4].

En primer lugar, nota que la definición de función Lipschitz que dimos en el Capítulo 1, puede extenderse también cuando X, Y sean espacios métricos (no necesariamente normados), cambiando en la desigualdad las normas por las respectivas distancias. Así pues, sea (M, d) un espacio métrico cualquiera, y sea $f : M \to \mathbb{R}$ una función Lipschitz. Definimos entonces

$$\|f\|_L = \sup\left\{ \frac{|f(x) - f(y)|}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\} < \infty. \tag{A.1}$$

Es fácil ver que $\|\cdot\|_L$ respeta tanto la multiplicación por escalares como la desigualdad triangular, y por ello define una seminorma en el espacio de funciones Lipschitz de M en $\mathbb{R}$. Además, observando la expresión de (A.1) podemos interpretar $\|f\|_L$ como la mejor constante Lipschitz de la función f.

Por otro lado, fijamos un punto de M y lo llamamos 0. A continuación consideramos el espacio

$$\mathrm{Lip}_0(M) = \{f : M \to \mathbb{R} \ : \ f \text{ es Lipschitz y } f(0) = 0\}.$$

Observamos que al restringir $\|\cdot\|_L$ al espacio $\mathrm{Lip}_0(M)$, como la única función constante que hay en dicho espacio es la función nula, se tiene que $\|\cdot\|_L$ define una norma en $\mathrm{Lip}_0(M)$. Y aún más, se tiene el siguiente resultado.

Proposición A.1. *El espacio* $(\mathrm{Lip}_0(M), \|\cdot\|_L)$ *es un espacio de Banach.*

Demostración. Consultar en [12], página 37. □

Veamos ahora una consideración sobre la elección del punto 0 de M.

Observación 9. Si $0_1, 0_2$ son dos puntos elegidos de M, la aplicación $\Phi : \mathrm{Lip}_{0_1}(M) \to \mathrm{Lip}_{0_2}(M)$ tal que

$$\Phi(f) = f - f(0_2), \quad \forall f \in \mathrm{Lip}_{0_1}(M)$$

está bien definida puesto que $\Phi(f) = f - f(0_2)$ sigue siendo Lipschitz y además se anula en el punto 0_2. Además, Φ es una isometría por ser una simple traslación. Por ello, podemos concluir que la elección del punto distinguido $0 \in M$ es irrelevante.

Con estas definiciones, dado $x \in M$, consideramos la aplicación

$$\delta(x) : \mathrm{Lip}_0(M) \to \mathbb{R} \quad \text{tal que} \quad \delta(x)(f) = f(x), \quad \forall f \in \mathrm{Lip}_0(M). \tag{A.2}$$

La función $\delta(x)$ es evidentemente lineal y además verifica que

$$|\delta(x)(f)| = |f(x)| = |f(x) - f(0)| \leq \|f\|_L d(x,0), \quad \forall f \in \mathrm{Lip}_0(M).$$

Luego $\delta(x)$ es continua y por tanto, para cualquier $x \in M$, se tiene que $\delta(x) \in \mathrm{Lip}_0(M)^*$ con $\|\delta(x)\| \leq d(x,0)$. De esta manera, podemos introducir ya el concepto de espacio Lipschitz-libre.

Definición 17. Dado (M,d) un espacio métrico, llamamos espacio *Lipschitz-libre* sobre M al espacio de Banach

$$\mathscr{F}(M) := \overline{\mathrm{span}}\{\delta(x) \ : \ x \in M\} \subseteq \mathrm{Lip}_0(M)^*.$$

Observación 10. Veamos que el conjunto $\{\delta(x) \ : \ x \in M\}$ es linealmente independiente, es decir, que dados $N \in \mathbb{N}$, $x_1, \ldots, x_N \in M$, y $a_1, \ldots, a_N \in \mathbb{R}$ siempre se tiene $\sum_{i=1}^N a_i \delta(x_i) \neq 0$. Esto es cierto, porque si no fuese así, y existiera una combinación lineal que se anulase, significaría que $\sum_{i=1}^N a_i f(x_i) = 0$ para cualquier $f \in \mathrm{Lip}_0(M)$. Y esto es imposible, ya que podríamos tomar las funciones

$$f_i(t) = d(t, \{x_1, \ldots, x_i\}) - d(0, \{x_1, \ldots, x_i\}), \quad i = 1, \ldots, N-1$$

que claramente están en $\mathrm{Lip}_0(M)$ y así concluir que $a_i = 0$ para $i = 1, \ldots, N$.
En resumen, $\{\delta(x) \ : \ x \in M\}$ es un conjunto linealmente independiente.

A continuación, vamos a dar una serie de propiedades en relación con los espacios Lispchitz-libres.

Proposición A.2. *Dado (M,d) un espacio métrico, sea la aplicación $\delta : M \to \mathscr{F}(M)$ tal que $\delta(x)$ es la definida en (A.2) para todo $x \in M$. Entonces δ es una isometría.*

Demostración. Sean $x, y \in M$ con $x \neq y$. Por un lado, tenemos que

$$\|\delta(x) - \delta(y)\| = \sup\{|(\delta(x) - \delta(y))(f)| : f \in B_{\mathrm{Lip}_0(M)}\} = \sup\{|f(x) - f(y)| : f \in B_{\mathrm{Lip}_0(M)}\} \leq d(x,y).$$

Por otra parte, considera la aplicación $f(t) = d(x,t) - d(x,0)$ para $t \in M$. Es claro que $f(0) = 0$ y además por la desigualdad triangular $|f(t) - f(s)| \leq d(t,s)$, para cualesquiera $t, s \in M$. Por tanto, $f \in B_{\mathrm{Lip}_0(M)}$, luego

$$\|\delta(x) - \delta(y)\| \geq |(\delta(x) - \delta(y))(f)| = |f(x) - f(y)| = |d(x,x) - d(x,0) - d(x,y) + d(x,0)| = d(x,y),$$

ya que la distancia es no negativa y $d(x,x) = 0$. En conclusión, $\|\delta(x) - \delta(y)\| = d(x,y)$ y así δ es una isometría. $\square$

Proposición A.3. *Dado (M,d) un espacio métrico, Y un espacio de Banach y $f : M \to Y$ una función Lipschitz con $f(0) = 0$, existe un único operador lineal $T : \mathscr{F}(M) \to Y$ cumpliendo $T \circ \delta = f$, y $\|T\| = \|f\|_L$.*

Demostración. Vamos a construir la función T en cuestión. Definimos $T : \mathrm{span}\{\delta(x) \ : \ x \in M\} \to Y$, por linealidad de la siguiente forma

$$T\left(\sum_{i=1}^N a_i \delta(x_i)\right) = \sum_{i=1}^N a_i f(x_i), \quad \text{para cualesquiera} \quad N \in \mathbb{N},\ a_i \in \mathbb{R},\ x_i \in M,\ i = 1, \ldots, N.$$

Vamos a ver que T es acotado y $\|T\| = \|f\|_L$. Por un lado, dado $\mu = \sum_{i=1}^N a_i \delta(x_i) \in \mathrm{span}\{\delta(x) \ : \ x \in M\}$ con $\mu \neq 0$, tenemos que $T(\mu) \in Y$, luego por el Corolario 1.4 existe $y^* \in B_{Y^*}$ con $y^*(T(\mu)) = \|T(\mu)\|$.

Además, $y^* \circ f \in \mathrm{Lip}_0(M)$ con $\|y^* \circ f\|_L \leq \|f\|_L$. De esta manera, puesto que $\mu \in \mathscr{F}(M) \subseteq \mathrm{Lip}_0(M)^*$ obtenemos

$$|\mu(y^* \circ f)| \leq \|\mu\| \|y^* \circ f\| \leq \|\mu\| \|f\|_L.$$

Y también se tiene

$$|\mu(y^* \circ f)| = \left| \left(\sum_{i=1}^{N} a_i \delta(x_i) \right) (y^* \circ f) \right| = \left| \sum_{i=1}^{N} a_i y^*(f(x_i)) \right| = |y^*(T(\mu))| = \|T(\mu)\|.$$

En definitiva, se obtiene que $\|T(\mu)\| \leq \|\mu\| \|f\|_L$. Por tanto T es acotado, y lo podemos extender por continuidad a todo $\mathscr{F}(M)$. Ahora, sí tiene sentido considerar $\|T\|$ y concluir que $\|T\| \leq \|f\|_L$.
Para la desigualdad contraria, observamos que para cualesquiera $x, y \in M$, con $x \neq y$, se tiene

$$T\left(\frac{\delta(x) - \delta(y)}{d(x,y)} \right) = \frac{f(x) - f(y)}{d(x,y)},$$

y por la Proposición A.2, $\left\| \frac{\delta(x)-\delta(y)}{d(x,y)} \right\| = 1$. De esta forma, obtenemos

$$\|T\| = \sup\{|T(\mu)| \ : \ \mu \in B_{\mathscr{F}(M)}\} \leq \sup\left\{ \frac{|f(x) - f(y)|}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\} = \|f\|_L.$$

Así pues, $\|T\| = \|f\|_L$. La unicidad se sigue inmediatamente de la definición de T.

□

Visualmente, podemos representar el Teorema A.3 mediante el siguiente diagrama.

$$\begin{tikzcd}
M \arrow[r, "f"] \arrow[d, "\delta"'] & Y \\
\mathscr{F}(M) \arrow[ur, dashed, "T"'] &
\end{tikzcd}$$

Observación 11. En particular, aplicando el Teorema A.3 al caso $Y = \mathbb{R}$, se obtiene que toda función $f \in \mathrm{Lip}_0(M)$ se corresponde a una única función $T \in \mathscr{F}(M)^*$, con la misma norma y tal que $T \circ \delta = f$. Recíprocamente, dada $T \in \mathscr{F}(M)^*$, la función $f = T \circ \delta$ está en $\mathrm{Lip}_0(M)$ y tiene la misma norma que T. Es decir, se tiene que

$$\mathrm{Lip}_0(M) = \mathscr{F}(M)^*,$$

entendiendo dicha igualdad como que son espacios isométricos.

Finalmente, daremos una descripción de la bola (cerrada) unidad de $\mathscr{F}(M)$.

Proposición A.4. *Dado (M,d) un espacio métrico, se tiene que*

$$B_{\mathscr{F}(M)} = \overline{\mathrm{conv}}\left\{ \frac{\delta(x) - \delta(y)}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\}.$$

Demostración. Gracias a la proposición A.2, el contenido de derecha a izquierda es obvio. Veamos el otro contenido entonces. Supongamos que existe $\mu \in B_{\mathscr{F}(M)} \setminus \overline{\mathrm{conv}}\left\{ \frac{\delta(x)-\delta(y)}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\}$. Por un argumento de separación como consecuencia del Teorema de Hahn-Banach 1.3 (consultar en [2], pág. 215), existe $f \in \mathscr{F}(M)^* = \mathrm{Lip}_0(M)$ tal que

$$f(\mu) > \sup\left\{ f(\nu) \ : \ \nu \in \overline{\mathrm{conv}}\left\{ \frac{\delta(x) - \delta(y)}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\} \right\}.$$

Y observando que el término de la izquierda es menor o igual que $\|f\|_L$ porque $\mu \in B_{\mathscr{F}(M)}$, y el término de la derecha es mayor o igual que $\|f\|_L$ por definición, se obtiene la contradicción. Así pues, concluimos que

$$B_{\mathscr{F}(M)} \subseteq \overline{\text{conv}} \left\{ \frac{\delta(x) - \delta(y)}{d(x,y)} \ : \ x, y \in M \text{ con } x \neq y \right\}$$

y por tanto se obtiene el enunciado. □

Observación 12. Para terminar esta sección, veamos que si X es un espacio de Banach separable, podemos encajarlo en $\mathscr{F}(X)$ de manera lineal e isométrica. Esto ocurre, como consecuencia del Teorema 3.15 aplicado a la isometría (no lineal) $\delta : X \to \mathscr{F}(X)$. Así, existe un subespacio de $\mathscr{F}(X)$ que es linealmente isométrico a X.

Apéndice B

Resultados complementarios

Este apéndice lo dedicaremos a dar las demostraciones de algunos resultados que hemos utilizado a lo largo del trabajo. En concreto, el Teorema de Alaoglu y ciertas propiedades sobre la integral de Bochner.

En primer lugar, vamos a enunciar y demostrar el Teorema 3.10, en el que nos hemos apoyado para probar algunas cuestiones de este trabajo. Esta demostración está apoyada sobre [13], pág. 29.

Teorema B.1. *Sea el dual X^* de un espacio normado X. Entonces la bola (cerrada) unidad $B_{X^*} \subseteq X^*$ es débil*-compacto.*

Demostración. En primer lugar, consideramos el conjunto $\prod_{x\in X}[-\|x\|,\|x\|]$, que lo podemos identificar con el conjunto de aplicaciones $A = \{f : X \to \mathbb{R} : |f(x)| \leq \|x\|, \forall x \in X\}$. Así, consideramos la función

$$\Phi : B_{X^*} \to \prod_{x\in X}[-\|x\|,\|x\|] \quad \text{tal que} \quad \Phi(x^*) = (x^*(x))_{x\in X},\ \forall x^* \in B_{X^*}$$

donde en el dominio tomamos la topología débil* y en el conjunto de llegada se toma la topología producto. Veamos que Φ es un homeomorfismo sobre su imagen.

- Φ es inyectiva. Sean $x^*, y^* \in B_{X^*}$ tal que $\Phi(x^*) = \Phi(y^*)$. Entonces para todo $x \in X$, $x^*(x) = y^*(x)$. Es decir, $x^* = y^*$.

- Φ es continua. Sea V un abierto básico de $\prod_{x\in X}[-\|x\|,\|x\|]$, que será de la forma

  $$V = \prod_{x\in X} V_x$$

  donde $V_x = [-\|x\|,\|x\|]$ para todo $x \in X$ excepto un conjunto finito $\{x_1,\dots,x_n\} \subseteq X$ donde $V_{x_i} = [-\|x_i\|,\|x_i\|] \cap (a_i - \varepsilon_i, a_i + \varepsilon_i)$ con $a_i \in \mathbb{R}$, $\varepsilon_i > 0$, $i = 1,\dots,n$. Ahora, sea $x_0^* \in \Phi^{-1}(V)$, es decir, $x_0^*(x) \in V_x$, $\forall x \in X$. Por tanto, $x_0^*(x_i) \in (a_i - \varepsilon_i, a_i + \varepsilon_i)$, $i = 1,\dots,n$. Entonces podemos tomar $\varepsilon > 0$ tal que

  $$(x_0^*(x_i) - \varepsilon, x_0^*(x_i) + \varepsilon) \subseteq (a_i - \varepsilon_i, a_i + \varepsilon_i), \quad i = 1,\dots,n.$$

  En consecuencia, se tiene que $W = U(x_0^*; \varepsilon, x_1, \dots, x_n) \cap B_{X^*}$ es entorno de x_0^* en B_{X^*} (con la topología débil*) y cumple que $W \subseteq \Phi^{-1}(V)$. Luego $\Phi^{-1}(V)$ es abierto y por tanto Φ es continua.

- Φ^{-1} es continua. Sea U un abierto básico en B_{X^*}. Por tanto

  $$U = U(x_0^*; \varepsilon, x_1 \dots, x_n) \cap B_{X^*}, \quad \text{para ciertos } x_1,\dots,x_n \in X,\ \varepsilon > 0,\ x_0^* \in X^*.$$

  Recordando la identificación del inicio, es claro que

  $$\Phi(U) = \{f \in A : |f(x_i) - x_0^*(x_i)| < \varepsilon,\ i = 1,\dots,n\} \cap \Phi(B_{X^*}) = \left(\prod_{x\in X} U_x\right) \cap \Phi(B_{X^*})$$

donde

$$U_x = [-\|x\|, \|x\|], \quad \text{si } x \notin \{x_1, \ldots, x_n\}$$

$$U_{x_i} = [-\|x_i\|, \|x_i\|] \cap (x_0^*(x_i) - \varepsilon, x_0^*(x_i) + \varepsilon), \quad i = 1, \ldots, n.$$

Por tanto, $\Phi(U)$ es abierto en la topología producto relativa a $\Phi(B_{X^*})$. Así, Φ^{-1} es continua.

En definitiva Φ es un homeomorfismo sobre su imagen.
Veamos ahora que $\Phi(B_{X^*})$ es cerrado, o equivalentemente, que $(\prod_{x\in X}[-\|x\|, \|x\|]) \setminus \Phi(B_{X^*})$ es abierto. Para ello, (de nuevo gracias a la identificación inicial) sea $f \in A$ con $f \notin \Phi(B_{X^*})$. Necesariamente f no es lineal (si lo fuera $f \in \Phi(B_{X^*})$), luego se cumple una de las dos opciones siguientes.

I. Existen $x_0 \in B_X$ y $\lambda \in \mathbb{R}$ con $f(\lambda x_0) \neq \lambda f(x_0)$.
Sea $\delta = \frac{1}{2}|f(\lambda x_0) - \lambda f(x_0)| > 0$ y $V = \{g \in A : |g(x_0) - f(x_0)| < \delta,\ |g(\lambda x_0) - f(\lambda x_0)| < \delta\}$. Nota que si $g \in V$ entonces $g(\lambda x_0) \neq \lambda g(x_0)$, luego

$$f \in V \subseteq \left(\prod_{x\in X}[-\|x\|, \|x\|]\right) \setminus \Phi(B_{X^*}).$$

Además, V es abierto en $\prod_{x\in X}[-\|x\|, \|x\|]$ con la topología producto.

II. Existen $x_0, y_0 \in B_X$ con $f(x_0 + y_0) \neq f(x_0) + f(y_0)$. Sea $\delta = \frac{1}{3}|f(x_0) + f(y_0) - f(x_0 + y_0)| > 0$ y $V = \{g \in A : |g(x_0) - f(x_0)| < \delta,\ |g(y_0) - f(y_0)| < \delta,\ |g(x_0 + y_0) - f(x_0 + y_0)| < \delta\}$. De nuevo se tiene que V es abierto en $\prod_{x\in X}[-\|x\|, \|x\|]$ y además

$$f \in V \subseteq \left(\prod_{x\in X}[-\|x\|, \|x\|]\right) \setminus \Phi(B_{X^*}).$$

En definitiva, en cualquiera de las dos opciones tenemos que $\Phi(B_{X^*})$ es cerrado y puesto que el conjunto $\prod_{x\in X}[-\|x\|, \|x\|]$ es compacto por el Teorema de Tychonoff, se deduce que $\Phi(B_{X^*})$ es compacto. Como Φ^{-1} es continua, es claro que $B_{X^*} = \Phi^{-1}(\Phi(B_{X^*}))$ es compacto, como queríamos probar. □

En segundo lugar, vamos a demostrar las propiedades de la integral de Bochner que se han utilizado en este trabajo y que reflejan como esta integral tiene un comportamiento bastante similar a la integral de Lebesgue, a excepción de algunos resultados más avanzados.

Lema B.2 (Propiedades de la integral de Bochner)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, Y un espacio de Banach y sea $f : \Omega \to Y$ medible.*

I. *Si f es Bochner integrable, entonces para cualquier sucesión $(s_n)_n$ de funciones simples tal que $\lim\limits_{n\to\infty} \int_\Omega \|s_n - f\|\, d\mu = 0$, existe el límite $\lim\limits_{n\to\infty} (\int_\Omega s_n\, d\mu)$ y su valor es independiente de la sucesión $(s_n)_n$ escogida.*

II. *Si f es Bochner integrable, se cumple que $\| \int_\Omega f\, d\mu\| \leq \int_\Omega \|f\|\, d\mu$.*

III. *Si Ω es compacto y Hausdorff, μ es una medida finita Borel, y f es continua, entonces f es Bochner integrable.*

IV. *Si f es Bochner integrable, Z es un espacio de Banach y $T : Y \to Z$ una función lineal y acotada, entonces $T \circ f$ es Bochner integrable y $\int_\Omega T \circ f\, d\mu = T(\int_\Omega f\, d\mu)$.*

Demostración. I. En primer lugar, como $\lim\limits_{n\to\infty} \|f - s_n\|\, \mathrm{d}\mu = 0$, se tiene que la sucesión $\left(\int_\Omega s_n\, \mathrm{d}\mu\right)_n$ es de Cauchy, y puesto que Y es de Banach, existe el límite deseado. Por otro lado, sea t una función simple de la forma $t(x) = \sum_{i=1}^n x_i 1_{F_i}(x)$, donde los $F_i \subseteq \Omega$ son disjuntos. Entonces se tiene que

$$\left\|\int_\Omega t\, \mathrm{d}\mu\right\| = \left\|\sum_{i=1}^n x_i \mu(F_i)\right\| \leq \sum_{i=1}^n \|x_i\| \mu(F_i) = \sum_{i=1}^n \int_{F_i} \|x_i\|\, \mathrm{d}\mu = \sum_{i=1}^n \int_{F_i} \|t\|\, \mathrm{d}\mu = \int_\Omega \|t\|\, \mathrm{d}\mu. \quad \text{(B.1)}$$

Por tanto, si $(t_n)_n$ es otra sucesión de funciones simples cumpliendo $\lim_{n\to\infty} \|f-t_n\|\ \mathrm{d}\mu = 0$, tenemos que gracias a (B.1) se cumple

$$\left\| \lim_{n\to\infty} \int_\Omega (s_n - t_n)\ \mathrm{d}\mu \right\| \leq \lim_{n\to\infty} \int_\Omega \|s_n - t_n\|\ \mathrm{d}\mu \leq \lim_{n\to\infty} \left(\int_\Omega \|f-s_n\|\ \mathrm{d}\mu + \int_\Omega \|f-t_n\|\ \mathrm{d}\mu \right) = 0.$$

Así pues, se tiene que $\lim_{n\to\infty} \int_\Omega s_n\ \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega t_n\ \mathrm{d}\mu$ y por tanto es independiente de la sucesión de funciones simples escogida.

II. Ya hemos visto en (B.1) que esta propiedad era cierta para funciones simples. Ahora, para f se tiene que

$$\left\| \int_\Omega f\ \mathrm{d}\mu \right\| = \left\| \lim_{n\to\infty} \int_\Omega s_n\ \mathrm{d}\mu \right\| \leq \lim_{n\to\infty} \int_\Omega \|s_n\|\ \mathrm{d}\mu \leq \lim_{n\to\infty} \int_\Omega \|s_n - f\|\ \mathrm{d}\mu + \int_\Omega \|f\|\ \mathrm{d}\mu = \int_\Omega \|f\|\ \mathrm{d}\mu$$

donde $(s_n)_n$ es una sucesión de funciones simples cumpliendo $\lim_{n\to\infty} \int_\Omega \|s_n - f\|\ \mathrm{d}\mu = 0$.

III. Como f es continua y Ω compacto, $\|f\|$ está acotada y como μ es finita, $\|f\|$ es Lebesgue integrable. Además, $\|f\|$ es medible y por tanto existe una sucesión de funciones $(t_n)_n$ que toman una cantidad numerable de valores y que para cada $n \in \mathbb{N}$ cumplen $\|f - t_n\| < \frac{1}{n}$, para casi todo punto de Ω (ver en [3], pág. 179). Por tanto $\|t_n\|$ está acotada, luego es integrable Lebesgue para cada $n \in \mathbb{N}$. Además, dichas funciones serán de la forma $t_n = \sum_{m=1}^{\infty} x_{n,m} 1_{E_{n,m}}$ donde $x_{n,m} \in Y$ y $E_{n,m}$ los podemos suponer disjuntos entre sí para cada n fijo.
Ahora bien, como $\|t_n\|$ es integrable para cada $n \in \mathbb{N}$, existe $p_n \in \mathbb{N}$ tal que

$$\int_{\bigcup_{m=p_n+1}^{\infty} E_{n,m}} \|t_n\| \leq \frac{\mu(\Omega)}{n}, \quad \forall n \in \mathbb{N}.$$

Por tanto, definimos la sucesión de funciones simples $(s_n)_n$ tales que $s_n = \sum_{m=1}^{p_n} x_{n,m} 1_{E_{n,m}}$ para cada $n \in \mathbb{N}$. Entonces

$$\int_\Omega \|f - s_n\|\ \mathrm{d}\mu \leq \int_\Omega \|f - t_n\|\ \mathrm{d}\mu + \int_\Omega \|t_n - s_n\|\ \mathrm{d}\mu \leq 2\frac{\mu(\Omega)}{n}, \quad \forall n \in \mathbb{N}.$$

En definitiva se tiene que $\lim_{n\to\infty} \int_\Omega \|f - s_n\|\ \mathrm{d}\mu = 0$, luego f es Bochner integrable.

IV. Tomamos $(s_n)_n$ una sucesión de funciones simples tales que $\lim_{n\to\infty} \int_\Omega \|s_n - f\|\ \mathrm{d}\mu = 0$. Como T es lineal, las funciones $t_n = T \circ s_n = \sum_{i=1}^{n} T(x_i) 1_{E_i}$ son simples. Por tanto

$$\lim_{n\to\infty} \int_\Omega \|T(f) - t_n\|\ \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega \|T(f - s_n)\|\ \mathrm{d}\mu \leq \lim_{n\to\infty} \|T\| \int_\Omega \|f - s_n\|\ \mathrm{d}\mu = 0,$$

puesto que T es acotada y lineal. Luego $T \circ f$ es Bochner integrable. Además,

$$\int_\Omega T \circ f\ \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega T \circ s_n\ \mathrm{d}\mu = \lim_{n\to\infty} T\left(\int_\Omega s_n\ \mathrm{d}\mu \right) = T\left(\lim_{n\to\infty} \int_\Omega s_n\ \mathrm{d}\mu \right) = T\left(\int_\Omega f\ \mathrm{d}\mu \right).$$

Nota que para la segunda igualdad se ha usado que T permuta con la integral de s_n puesto que es lineal y s_n es una función simple. La tercera igualdad se debe a que T es continua, por el hecho de ser lineal y acotada.

□

Bibliografía

[1] J.M. BORWEIN, J.D. VANDERWERFF, Convex Functions: Constructions, Characterizations and Counterexamples. Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK (2010).

[2] B. CASCALES, J.M. MIRA, Análisis Funcional. Electrolibris, versión 1.0 (2012).

[3] B. CASCALES Y S. TROYANSKI, Fundamentos de Análisis Matemático. Universidad de Murcia (2007).

[4] M. CÚTH, M. DOUCHA, P. WOJTASZCZYK, On the structure of Lipschitz-free spaces. Proc. Amer. Math. Soc. 144 (2016), no. 9, 3833-3846.

[5] R.J. FLEMING, J.E. JAMISON, Isometries on Banach spaces: function spaces. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, vol. 129, Chapman & Hall/CRC, Boca Raton, FL (2003).

[6] G. GODEFROY, Linearization of isometric embeddings between Banach spaces: an elementary approach. Oper. Matrices, 6, 2 (2012), 339-345.

[7] G. GODEFROY, N.J. KALTON, Lipschitz-free Banach spaces, Studia Math. 159 (1) (2003), 121-141.

[8] P.R. HALMOS, Graduate texts in Mathematics, Measure Theory. Springer-Verlag, NewYork-Heidelberg-Berlin (1974).

[9] J.R. MUNKRES, Topología. Prentice Hall, Pearson Eduación S.A., segunda edición (2002).

[10] B. NICA, The Mazur-Ulam theorem. Expo. Math., 30 (2012), 397-398.

[11] W. RUDIN, Principios de Análisis Matemático. Libros McGrawhill de Mexico, 3ª edición (1980).

[12] N. WEAVER, Lipschitz Algebras. World Scientific, second edition (2018).

[13] P. WOJTASZCYK, Banach Spaces For Analysts. Cambridge University Press (1991).