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

**Tesis de Licenciatura**

# Orígenes y desarrollo del teorema de Dvoretzky

**Fernando Vische**

**Director:** Román Villafañe

Fecha de Presentación: 24/02/2025

*A mis abuelas, Elsa y Andresa, que eligieron la educación para sus hijas e hijos.*

# Agradecimientos

¿Qué es la matemática? ¿Es algo que existe con independencia de los seres humanos? ¿Es una construcción social? ¿Por qué están estas preguntas en los agradecimientos?

Quizás porque la experiencia vivida al realizar este trabajo, me ha resaltado que, independientemente de que usemos la palabra *descubrir* o *inventar*, la naturaleza en la que se lleva a cabo cualquiera de esos dos verbos *es social*. Es visceralmente social.

Y así, no tengo más que visceralmente agradecer...

Agradecerle a Roman, por aceptar dirigirme y acompañarme en un momento donde tal vez lo más lógico habría sido no hacerlo. Agradecer su constante apoyo y aliento, su paciencia y su bondad para conmigo, y mi personalidad un tanto errática.

Agradecerle a Pablo, por aceptar con completa solidaridad ayudarme en la finalización de este trabajo cuando se lo pedí, y por no soltarme la mano por más tentador que yo haya hecho eso.

Agradecerle a Esteban Andruchow y Martín Mansilla por aceptar ser jurados para esta tesis, y tomarse el trabajo de leerla.

Agradecerle a todos los docentes que tuve durante mis años de estudio en la universidad, en la secundaria, en la primaria, ¡y en el jardín! Atravesé todos estos espacios dentro de la educación pública, por lo que también agradezco a todos aquellos trabajadores y trabajadoras que en sus distintos roles la sostienen en el día a día.

Es sorprendente pensar como distintas personas, a lo largo de la vida de uno, lo van formando, con compromiso, con amor a su profesión. Quizás en muchas ocasiones sin sospechar que tan influyente puede llegar a ser su acción en la vida de un estudiante.

Particularmente recuerdo, y quisiera mencionar, a "la seño" María Inés de la salita de 5, a "la seño" Septimia de 3° grado, ambas de la escuela República de Nicaragua, por el amor con el que ejercían su tarea. A los docentes Francis La Greca, Susana Fiszman, Guillermo Rodriguez Pintos y Gabriela Retamal de la Escuela Técnica N°27 Hipólito Yrigoyen.

A Mercedes Perez Millan, Francisco Kordon y el Programa de Ingresantes CBC Exactas, que además de ayudarme a empezar el CBC con más seguridad (o menos inseguridad), me enseñaron que la formula resolvente de la cuadrática... ¡se deduce!

A Carlos Borches, que hizo que me pique el bichito de la matemática.

A los docentes que tuve en Exactas... Sebastián Freyre, Rafael Ferraro, Julian Haddad, Teresa Krick, Pablo Groisman, Dario Rodriguez, Matías Saucedo, Julián Fernández Bonder, Daniel Carando, Gabriel Minian, Estanislao Herscovich, Jonathan Barmak, Malena Becker y tantos otros... Por su compromiso e inspiración.

Agradecerle a todas aquellas personas con las que compartí años muy lindos en la

facultad, y que voy a llevar siempre conmigo. A Emi, Rago, Meli, Uli, Daiana, Maxi, Naza, Martin, Bili, Juampi, Maty Z, Eliana, Fiore, Pablo. Y tantas otras más con las que compartí, en tantos años...

Agradecerle a mi familia. A mi mamá, a mi papá, a mis hermanos, por su constante apoyo y su amor.

Agradecerle a Alejandra, por apoyarme y acompañarme en forma total.

No solo no hubiese podido realizar este trabajo (ni ningún otro) sin la formación y el apoyo que me brindaron estas personas, sino que le debo en parte algo de lo que soy a cada una de ellas.

Finalmente, agradezco también a las autoridades del CBC, por concederme una licencia que me fue de gran ayuda en la realización de este trabajo.

Por último, y casi en un mensaje para mi mismo, no quisiera olvidar que, como bien nos lo enseñan ciertas páginas de noticias en la internet de hoy en día, lo verdaderamente importante no está en ningún título, sino en el texto (¡y los sonidos e imágenes!) que tengamos dentro.

# Índice general

# Introducción

Ya en nuestro primer encuentro con las series, se nos menciona al pasar que una serie de números reales es absolutamente convergente si y solo si es incondicionalmente convergente, esto es, tras cualquier reordenamiento de los términos de la serie original, se obtiene una nueva serie que converge (y al mismo límite que la serie original).

Históricamente, se le atribuye a Dirichlet este *descubrimiento*. Notablemente, Dirichlet lo observa en el mismo trabajo [Dir37] en el que demuestra que toda sucesión aritmética $an+b$, con $a,b \in \mathbb{N}$ coprimos, contiene infinitos primos... ¡que gente increíble!

En verdad, parte de la crisis matemática que luego devino en una fundamentación más rigurosa del análisis (llevada a cabo durante el siglo XIX), estuvo causada por la manipulación de series (¡no siempre absolutamente!) convergentes como si fueran sumas, en las cuales se pueden reordenar los términos a gusto sin modificar el resultado.

Varios años después del resultado de Dirichlet, en 1854, Riemann demostró en su habilitación para un cargo de profesor auxiliar en la Universidad de Gotinga que si una serie numérica convergente no es absolutamente convergente, entonces para cada $\alpha \in \mathbb{R}$ existe algún reordenamiento en los términos de la serie, de manera que la serie así reordenada converja a $\alpha$ (notablemente, Riemann demuestra este hecho en el mismo trabajo en el que *explica* que significado tiene y cuando existe $\int_a^b f(x)\,dx$, definiendo en verdad lo que hoy conocemos como integral de Riemann... ¡sin palabras!). Este artículo sería publicado póstumamente, en 1867 [Rie67].

Así las cosas, desde el mismo momento en que se conciben los espacios de Banach, es natural preguntarse si se mantiene en dichos espacios la relación que encontramos en $\mathbb{R}$ entre la convergencia incondicional y la convergencia absoluta de series. Es decir, dada una sucesión $(x_n)_{n\in\mathbb{N}}$ en un espacio de Banach $X$ ¿Es equivalente la convergencia absoluta a la convergencia incondicional de la serie dada por la sucesión $(x_n)_{n\in\mathbb{N}}$?

Como es de esperarse, es sencillo dar respuesta a esta pregunta en el caso finito-dimensional: sí.

Esto se debe a que, como la convergencia en $\mathbb{R}^n$ con cualquier norma esta dada por la convergencia escalar coordenada a coordenada, la convergencia incondicional de una serie en $\mathbb{R}^n$ es equivalente a la convergencia incondicional en cada una de las coordenadas, lo cual es equivalente a la convergencia absoluta en cada coordenada. A su vez, por la desigualdad triangular (aplicada a cualquier norma sobre la descomposición canónica de un vector), y la continuidad de las funciones coordenadas, tenemos que para cada $1 \le j \le n$ existe una constante $C_j > 0$ tal que $\frac{1}{C_j}|x^j| \le \|x\| \le C(|x^1| + \ldots + |x^n|)$, donde $C = \text{máx}_{1\le i\le n}\|e_i\|$, siendo $e_n$ el $n$-ésimo vector de la base canónica, de donde la convergencia absoluta de la serie es equivalente a la convergencia absoluta (escalar) coordenada a coordenada.

Ahora bien, considerando el caso infinito-dimensional, si bien por un lado sigue siendo cierto que la convergencia absoluta de una serie implica la convergencia incondicional de la misma (y este hecho es, como en el caso escalar, equivalente a la completitud del espacio normado), es fácil ver que la equivalencia no es cierta en general. En efecto, por ejemplo podemos considerar en $l_2$ la sucesión $x_n = \frac{1}{n} e_n$, que da lugar a una serie incondicionalmente convergente, pero que no es absolutamente convergente. Ejemplos como este pueden producirse con relativa sencillez *en la mayoría* de los espacios clásicos, pero aún así, quedaba abierta la pregunta de si *existiría* algún espacio de Banach de dimensión infinita en el cual la convergencia absoluta y la convergencia incondicional de series sean equivalentes.

El propio Banach planteaba esta cuestión en los comentarios del 9° capítulo de su clásico y fundacional libro *Théorie des opérations linéaires (1932)* [Ban32]. Asimismo, esta pregunta constituye el problema número 122 del famoso *Scottish Book*, planteado allí por Orlicz y Mazur.

Notablemente, esta pregunta se mantuvo abierta durante aproximadamente 30 años (teniendo en cuenta que "el nacimiento definitivo de los espacios de Banach" fue a inicios de la segunda década del siglo XX). En 1950, Dvoretzky y Rogers [DR50] dan una respuesta para el caso infinito-dimensional: No.

Dedicaremos el Capítulo 1 principalmente a este trabajo, que si bien responde una pregunta muy importante e interesante, podría haber hecho simplemente eso, terminar o "zanjar" un problema abierto, y nada más. Y si bien no es un trabajo fundacional o revolucionario en sí mismo, es sorprendente como ejercerá una influencia histórica notable sobre el desarrollo de la teoría de espacios de Banach en la segunda mitad del siglo XX. Justamente, en esta tesis intentaremos recorrer tan solo una estrecha sección de la basta "catarata matemática" que encuentra a [DR50] como uno de sus afluentes principales (y lo haremos de forma necesariamente incompleta).

Más concretamente, en los Capítulos 2 y 3 nos concentraremos en el teorema de Dvoretzky, cuyas motivaciones originales podemos encontrar precisamente en [DR50].

El teorema de Dvoretzky dice que

**Teorema 1** (teorema de Dvoretzky)**.** *Dados $k \in \mathbb{N}$ y $\varepsilon > 0$ podemos encontrar en todo espacio normado* de dimensión suficientemente grande *(es decir, de dimensión mayor o igual a un cierto $n(k,\varepsilon)$) algún subespacio $(1+\varepsilon)$-isomorfo al espacio euclídeo $\left(\mathbb{R}^k, \| \ \|_2\right)$.*

En términos geométricos, dados $k \in \mathbb{N}$ y $\varepsilon > 0$, si $\| \ \|$ es una norma en $\mathbb{R}^n$ y $B$ es la bola unitaria de $(\mathbb{R}^n, \| \ \|)$, entonces, si la dimensión $n$ es suficientemente grande, existe alguna sección $k$-dimensional de $B$, digamos $B \cap S$ con $S$ subespacio de dimensión $k$ en $\mathbb{R}^n$, tal que para cierto elipsoide $\mathcal{E} \subseteq S$ (ver página xiv) se cumple que $\mathcal{E} \subseteq B \cap S \subseteq (1+\varepsilon)\mathcal{E}$ (en particular, $B \cap S$ *es muy parecido a* $\mathcal{E}$).

Como consecuencia de este teorema, se tiene el siguiente hecho acerca de los subespacios de dimensión finita de un espacio de Banach infinito-dimensional.

**Teorema 2.** *Si $X$ es un espacio de Banach de dimensión infinita, dado $\varepsilon > 0$, se cumple que para cada $n \in \mathbb{N}$ existe un subespacio $n$-dimensional $X_n \subseteq X$ tal que $X_n$ es $(1+\varepsilon)$-isomorfo al espacio euclídeo $(\mathbb{R}^n, \| \ \|_2)$.*

Notemos que, en términos isométricos, la única sucesión de espacios normados de dimensión finita $(Y_n)_{n\in\mathbb{N}}$, siendo $\dim(Y_n) = n$, que podría cumplir el Teorema 2 es la de los espacios euclídeos de dimensión finita $(\mathbb{R}^n, \| \|_2)_{n\in\mathbb{N}}$.

En efecto, si tomamos como espacio de Banach de dimensión infinita a un espacio de Hilbert, entonces *cada uno* de sus subespacios es isométrico a un Hilbert. En particular, todos sus subespacios de dimensión $n$ son isométricos a $\mathbb{R}^n$ con la norma euclídea. Por lo que, fijado $n \in \mathbb{N}$, $X$ tiene *esencialmente* un único subespacio de dimensión $n$, $(\mathbb{R}^n, \| \|_2)$, al cual $Y_n$ será $(1+\varepsilon)$-isomorfo para todo $\varepsilon > 0$. Así, $Y_n$ será necesariamente isométrico $(\mathbb{R}^n, \| \|_2)$.

Es algo notable el hecho de que, si bien quizás el enunciado del teorema de Dvoretzky puede sonar puramente intrínseco a la teoría de espacios normados de dimensión finita (quizás excepto por el "*de dimensión suficientemente grande*"), toda demostración del mismo (hasta donde sabemos) que dé alguna información cuantitativa sobre la dimensión $n(k,\varepsilon)$ más allá de su existencia, requiere de alguna u otra forma, y siempre en forma clave, de argumentos probabilísticos. Por otro lado, sí hay argumentos que se enmarcan enteramente dentro de la teoría de los espacios de Banach para demostrar el Teorema 2 (ver, por ejemplo, [AK16, Teorema 12.3.13]).

El teorema de Dvoretzky tiene una historia interesante, que podríamos iniciar con el propio teorema (y lema) de Dvoretzky-Rogers. En su tesis doctoral [Gro55], dedicada al desarrollo de la teoría de productos tensoriales entre espacios vectoriales topológicos, Grothendieck da una prueba del teorema de Dvoretzky-Rogers "*en su versión cualitativa*", es decir, muestra que si $X$ es un espacio de Banach tal que toda serie incondicionalmente convergente es absolutamente convergente, entonces $X$ es de dimensión finita. Obviamente, hace esto usando el lenguaje de los productos tensoriales, es decir, mediante técnicas totalmente distintas a las de [DR50].

Luego, un año más tarde, en [Gro56], Grothendieck vuelve a analizar en mayor profundidad estos hechos, y esta vez mediante el uso del lema de Dvoretzky-Rogers (ver Lema 1.1.7) y la teoría de normas en productos tensoriales de espacios de Banach, obtiene resultados "cuantitativos" que extienden al teorema de Dvoretzky-Rogers (ver Teorema 1.1.13). En este mismo artículo, Grothendieck comenta en la sección final, al pasar, que como consecuencia del lema de Dvoretzky-Rogers, se tiene el siguiente teorema (que demostraremos en la Sección 2.1).

**Teorema 3** (Factorización de Dvoretsky-Rogers)**.** *Dados $\varepsilon > 0$ y $k \in \mathbb{N}$, existe $n(k,\varepsilon) \in \mathbb{N}$ tal que en todo espacio normado $E$ de dimensión mayor o igual a $n(k,\varepsilon)$ se pueden encontrar vectores $x_1, \ldots, x_k$ de manera que, para cualesquiera $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ se cumple*

$$\frac{1}{1+\varepsilon} \|(\alpha_i)\|_\infty \leq \left\| \sum_{i=1}^{k} \alpha_i x_i \right\| \leq \|(\alpha_i)\|_2$$

Luego, conjetura si la misma afirmación será válida cambiando la norma infinito por la norma 2 en la cota inferior (esto es, exactamente, el teorema de Dvoretzky). Como el mundo es un pañuelo, el revisor de [Gro56] para su comentario en Mathematical Reviews no era otro más que Dvoretzky (esto es tan increíble, que más que un hecho digamos que es un rumor), que evidentemente, se sintió ampliamente seducido por dicha conjetura. Así, en 1959, Dvoretzky publica un artículo [Dvo59] donde afirma que esto es cierto, dando un esquema de la demostración. En [Dvo61] de 1961, publicaría su demostración completa.

Sin embargo, la demostración de Dvoretzky es considerada algo intrincada (de hecho, 11 años más tarde, en [Fig72], Figiel publicaría algunas aclaraciones y simplificaciones de la misma). A partir de allí, se dieron diversas demostraciones del teorema de Dvoretzky, buscando no solo exponerlo con mayor claridad, sino también con mayor precisión en la estimación de $n(k,\varepsilon)$.

Concretamente, si definimos

$$N(k,\varepsilon) = \text{mín}\left\{\begin{array}{l} n \in \mathbb{N}/ \text{ para todo espacio normado } (E, \|\ \|) \text{ con } \dim(E) = n \\ \qquad\qquad \exists\, S \subseteq E \text{ subespacio } k\text{-dimensional y } (1+\varepsilon)\text{-isomorfo a } (\mathbb{R}^k, \|\ \|_2) \end{array}\right\}, \tag{1}$$

entonces:

- En su demostración original [Dvo61], Dvoretzky obtuvo que $N(k,\varepsilon) \leq \exp\left(\frac{c}{\varepsilon^2} k^2 \ln^2(k)\right)$ con $c > 0$ constante absoluta ($c = 2^{15}$).

- En 1971, Milman [Mil71] dio una nueva (y extremadamente influyente) demostración, donde obtuvo que $N(k,\varepsilon) \leq \exp\left(c_1 \frac{|\ln(c_2\,\varepsilon)|}{\varepsilon^2} k\right)$, utilizando una interesante desigualdad descubierta por P. Lévy, a partir de la cual surgiría el estudio de lo que el mismo Milman dio en llamar "concentración de la medida".

  Más precisamente, esta desigualdad implica que existe una constante $c > 0$ tal que si $f : \mathbb{S}^{n-1} \to \mathbb{R}$ es una función Lipschitz de constante 1 sobre la esfera $(n-1)$-dimensional, y $M_f$ es una mediana de $f$ en $\mathbb{S}^{n-1}$ (i.e. $\lambda_n(f \geq M_f) \geq 1/2$ y $\lambda_n(f \leq M_f) \geq 1/2$, donde $\lambda_n$ es la medida de probabilidad invariante por rotaciones en $\mathbb{S}^{n-1}$) vale que

  $$\lambda_n\Big(|f - M_f| > \delta\Big) < 2\,e^{-c\,\delta^2 n}.$$

  Como veremos en la Proposición 2.5.1, la dependencia en $k$ de esta estimación es óptima, mostrando que en verdad $N(k,\varepsilon) = \exp(\eta(\varepsilon)k)$, donde $\eta$ es una función que solo depende de $\varepsilon$, definida en el intervalo $(0,1)$ (o algún intervalo $(0,\delta)$ con $\delta < 1$ de ser necesario) y cuyo comportamiento asintótico exacto es desconocido aún actualmente.

- En 1976, Figiel [Fig76] daría una nueva demostración (que él consideraría corta), en la que dada una norma $p$ en $\mathbb{R}^n$, se utiliza como medida de su distancia al espacio euclídeo $(\mathbb{R}^n, \|\ \|_2)$ la cantidad

  $$v_n(p) = \int_{\Sigma_n} \frac{|\langle \nabla p(x), y\rangle|^2}{p^2(x)}\, d\sigma_n \text{ donde } \Sigma_n = \Big\{(x,y) \in \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} / \langle x, y\rangle = 0\Big\}$$

y $\sigma_n$ es la medida de probabilidad invariante por la acción del grupo ortogonal (o unitario) en $\Sigma_n$.

Por este medio, solo se prueba la existencia de $n(k,\varepsilon)$, sin ninguna estimación cuantitativa.

- En 1977, Figiel, Lindenstrauss, y Milman publican un artículo [FLM77] donde dan una prueba "más limpia" que la de [Mil71], basándose en las mismas ideas. Lo notable de este artículo es que, dado un espacio normado $X$ de dimensión $n$, se estudia

$$k(X,\varepsilon) = \text{máx}\Big\{k \in \mathbb{N} \,/\, \exists\, S \subseteq X \text{ con } \dim(S) = k, S\ (1+\varepsilon)\text{-isomorfo a } (\mathbb{R}^k, \|\ \|_2)\Big\} \tag{2}$$

en relación a características específicas del espacio $X$ (como su tipo y cotipo, ó su distancia de Banach-Mazur a $(\mathbb{R}^n, \|\ \|_2)$, entre otras).

- En 1985, Gordon mostró en [Gor85], a través de ciertas desigualdades (establecidas en el mismo trabajo) entre variables aleatorias gaussianas, que $\eta(\varepsilon) \le c/\varepsilon^2$. Es decir, que se tiene

$$N(k,\varepsilon) \le \exp\left(\frac{c\,k}{\varepsilon^2}\right).$$

- En 1986, Pisier exhibió una demostración (que también puede encontrarse en el Capítulo 4 de su libro [Pis89]) análoga a la de [FLM77] (y que obtiene la misma cota que la obtenida allí para $N(k,\varepsilon)$), pero usando el espacio de medida $(\mathbb{R}^n, \gamma_n)$, donde $\gamma_n$ es la medida gaussiana estándar en $\mathbb{R}^n$. Reproduciremos esta demostración en el Capítulo 2.

- En 1988, Schechtman publicó una "observación" en [Sch89], donde nota que también puede deducirse que $\eta(\varepsilon) \le c/\varepsilon^2$ con algunos argumentos adicionales a la misma línea general que plantea la demostración dada por Pisier en el ítem anterior.

- En 2006 Schechtman probó en [Sch06] que $\eta(\varepsilon) \le c_1 \dfrac{|\ln(c_2\,\varepsilon)|^2}{\varepsilon}$. Para esto, utilizo un resultado dado en [AM83], que indica condiciones bajo las cuales se puede encontrar en un espacio normado $X$ un subespacio $(1+\varepsilon)$-isomorfo a $(\mathbb{R}^k, \|\ \|_\infty)$ (ver Teorema 3.3.7).

- En 2018, Paouris y Valettas demostraron en [PV18] que $\eta(\varepsilon) \le c_1 \dfrac{|\ln(c_2\,\varepsilon)|}{\varepsilon}$. Esta es la mejor estimación conocida para $\eta(\varepsilon)$ hasta el momento. Para ello utilizaron principalmente resultados de [AM83] (al igual Schechtman en [Sch06]) y una desigualdad conocida como desigualdad $L_1 - L_2$ de Talagrand para la medida gaussiana.

  El Capítulo 3 estará dedicado a contextualizar tanto el trabajo de Paouris y Valettas (más allá de lo ya expuesto en esta Introducción), como los resultados principales de los que estos autores hacen uso (mencionados en el párrafo anterior). También allí delinearemos las ideas principales seguidas en [PV18] para obtener su resultado principal.

Luego, en los apéndices finales de este trabajo, desarrollaremos con mayor detalle el camino delineado en el Capitulo 3 acerca del trabajo de Paouris y Valettas.

Finalmente, mencionemos que en la Proposición 2.5.2 veremos que $k((\mathbb{R}^n, \| \|_\infty), \varepsilon) \geq c_1 \dfrac{\ln(n)}{|\ln(c_2\,\varepsilon)|}$. A partir de esto, teniendo en cuenta que $(\mathbb{R}^n, \| \|_\infty)$ es un espacio de lo más no euclídeo (y otros hechos que no mencionamos aquí...) se ha conjeturado en [Mil88] que

$$k(n,\varepsilon) = \text{mín}\Big\{k(X,\varepsilon),\, X \text{ espacio normado de dim } n\Big\} \sim \frac{\ln(n)}{|\ln(c\,\varepsilon)|}[1]. \tag{3}$$

Esto es equivalente a

$$N(k,\varepsilon) = \exp\Big(c_1\,|\ln(c_2\,\varepsilon)|\,k\Big)$$

para ciertas constantes $c_1, c_2 > 0$. Es decir $\eta(\varepsilon) \sim |\ln(c_2\,\varepsilon)|$.

En 1989 Bourgain y Lindenstrauss demostraron en [BL89] que esto efectivamente se cumple si $X$ es un espacio con una base simétrica, es decir, con una base $v_1, \dots, v_n$ tal que para toda permutación $\sigma \in S_n$

$$\left\| \sum_{j=1}^{n} a_j\, v_j \right\| = \left\| \sum_{j=1}^{n} \pm a_{\sigma(j)}\, v_j \right\|.$$

Por último, comentemos que en 2023, Bo'az Klartag y Tomer Novikov publicaron en arxiv un artículo [KN23] en el que demuestran que $\eta(\varepsilon) \leq c|\ln(\varepsilon)|$, bajo la restricción más débil de que el subespacio de dimensión $k$ sea $(1+\varepsilon)$-isomorfo a un espacio normado con base 1-incondicional (en vez de a un espacio euclídeo). Esto es, a un espacio que posea una base $v_1, \dots, v_k$ tal que para cualesquiera $\alpha_1, \dots, \alpha_k \in \mathbb{R}$

$$\left\| \sum_{j=1}^{k} \alpha_j v_j \right\| = \left\| \sum_{j=1}^{k} \pm \alpha_j v_j \right\|.$$

[1] Dado un conjunto $S$ y dos funciones $f, g : S \to \mathbb{R}$ notamos $f \sim g$ si existen constantes $a, b > 0$ tales que $a\,f(s) \leq g(s) \leq b\,f(s)$ para todo $s \in S$.

# Notación y preliminares

Si bien la mayoría de los tópicos tratados en este trabajo se desarrollan de forma similar tanto si el cuerpo a considerar es $\mathbb{R}$ como si es $\mathbb{C}$, en general enunciaremos los resultados en el caso real. De ser necesario, especificaremos cuando surja una diferencia entre el caso real y el caso complejo. A su vez, utilizaremos $\mathbb{K}$ para referirnos a cualquiera de los dos cuerpos indistintamente. Siempre consideraremos la topología usual en $\mathbb{K}^n$ (que es la inducida por cualquier norma, por ejemplo la euclídea), y notaremos $(e_k)_{1\leq k\leq n}$ a la base canónica de $\mathbb{K}^n$.

Dado un espacio normado $(E, \| \|)$, notaremos:

- $B_{(E,\| \|)} = \{x \in E / \ \|x\| \leq 1\}$ a la bola unitaria de $(E, \| \|)$. Escribiremos $B_E = B_{(E,\| \|)}$ cuando la norma en $E$ quede clara por el contexto.
- Análogamente, si no hay confusión denotaremos $S_E = S_{(E,\| \|)} = \{x \in E \ / \ \|x\| = 1\}$ a la esfera unitaria de $E$.

Dados dos espacios normados $E$ y $F$ notaremos:

- $\mathcal{L}(E,F) = \{T : E \to F \text{ tales que } \ T \text{ es lineal y continuo}\}$.
- $E^* = \mathcal{L}(E,\mathbb{K})$ al conjunto de funcionales lineales y continuos en $E$.

### $C$-isomorfismos y distancia de Banach-Mazur

Dada una constante $C \geq 1$, diremos que dos espacios normados $E$ y $F$ son $C$-isomorfos si existe un isomorfismo $T \in \mathcal{L}(E,F)$ tal que $\|T\| \left\|T^{-1}\right\| \leq C$. En este caso, para todo $x \in E$ vale que

$$\frac{1}{\left\|T^{-1}\right\|} \|x\|_E \leq \|T(x)\|_F \leq \|T\| \|x\|_E .$$

Por lo que, definiendo en $E$ la norma $|x| = \left\|T^{-1}\right\| \|T(x)\|_F$ se cumple que $(E, | \ |)$ es isométricamente isomorfo a $(F, \| \|_F)$, siendo $u : (E, | \ |) \to (F, \| \|_F)$ definido por $u(x) = \left\|T^{-1}\right\| T(x)$ el isomorfismo isométrico. La relación entre las normas $\| \|_E$ y $| \ |_E$ en $E$ esta dada por

$$\|x\|_E \leq |x|_E \leq C \|x\|_E$$

de donde se ve claramente que en caso de ser $C$-isomorfos, las normas de $E$ y $F$ son *"$C$-equivalentes"*, y en este caso además $B_{(E,| \ |)} \subseteq B_{(E,\| \|_E)} \subseteq C\, B_{(E,| \ |)}$.

Dados dos espacios normados $E$ y $F$, isomorfos entre sí, definimos su distancia de Banach-Mazur como

$$\begin{aligned} d_{BM}(E,F) &= \text{ínf}\left\{\|T\|\left\|T^{-1}\right\|, T \in \mathcal{L}(E,F) \text{ isomorfismo}\right\} \\ &= \text{ínf}\{C \geq 1 \,/\, E \text{ y } F \text{ son } C\text{-isomorfos}\}. \end{aligned}$$

En caso de ser isométricamente isomorfos, vale que $d_{BM}(E,F) = 1$. La propiedad recíproca solo es cierta entre espacios de dimensión finita (por ejemplo, ver [Tom89, Capítulo 2]).

Notemos que dos espacios normados de dimensión $n$ son isométricos si y solo si existe $T \in \mathcal{L}(E,F)$ tal que $T(B_E) = B_F$.

### Espacios clásicos

Para cada $1 \leq p \leq \infty$ notaremos $l_p^n$ a $\mathbb{K}^n$ dotado con la norma

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \quad \text{si } p < \infty \quad \text{y} \quad \|x\|_\infty = \max_{1\leq i\leq n} |x_i|$$

y $B_p^n = B_{l_p^n}$ a la bola unitaria de estos espacios.

También denotaremos

$$\begin{aligned} l_p &= \left\{(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \,/\, \sum_{n=1}^{+\infty} |x_n|^p < +\infty\right\} \qquad \text{si } 1 \leq p < \infty \\ l_\infty &= \left\{(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \,/\, \sup_{n\in\mathbb{N}} |x_n| < +\infty\right\} \end{aligned}$$

con las normas

$$\|(x_n)\|_p = \left(\sum_{n=1}^{+\infty} |x_n|^p\right)^{1/p} \quad \text{si } 1 \leq p < \infty \quad \text{y} \quad \|(x_n)\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

### Convergencia débil y la propiedad de Schur

Recordemos que dado un espacio de Banach $(X, \| \ \|)$ podemos considerar en el mismo la topología débil, definida por ser la topología "más cruda" bajo la cual son continuos todos los funcionales en $X^*$. Esta topología está estrictamente incluida en la topología fuerte (aquella inducida por la norma en el espacio) siempre que $X$ sea infinito dimensional. Una base de abiertos para la misma está dada por los conjuntos

$$V_{x_0,\varphi_1,\ldots,\varphi_n,\varepsilon} = \{x \in X \,/\, |\varphi_i(x) - \varphi_i(x_0)| < \varepsilon \quad \forall\, 1 \leq i \leq n\}$$

donde $x_0 \in X$, $\varphi_1, \ldots, \varphi_n \in X^*$ y $\varepsilon > 0$. Diremos que una sucesión $(x_n)_{n\in\mathbb{N}} \subseteq X$ es débilmente convergente a un elemento $x \in X$ si converge hacia $x$ en la topología débil, y escribiremos $x = \omega\text{-}\lim_{n\to\infty} x_n$ o $x_n \rightharpoonup x$. Notemos que esto ocurre si y solo si $\varphi(x_n) \to \varphi(x)$ en $\mathbb{K}$ para todo $\varphi \in X^*$.

Decimos que un espacio de Banach $(X, \| \ \|)$ tiene *la propiedad de Schur* si toda sucesión débilmente convergente en $X$ es fuertemente convergente (es decir, converge con respecto a la topología inducida por la norma en $X$).

En 1921 (¡hace más de 100 años!) Schur demostró en [Sch21] que $l_1$ posee esta propiedad (de donde ésta lleva su nombre).

### Normas y cuerpos convexos balanceados en $\mathbb{K}^n$

Dado un conjunto $K \subseteq \mathbb{K}^n$ decimos que:

- es convexo si $\lambda x + (1-\lambda)y \in K$ siempre que $x, y \in K$, y $\lambda \in [0,1]$ .
- es un cuerpo convexo si es convexo, compacto, y con interior no vacío.
- es balanceado si $\lambda x \in K$ siempre que $x \in K$, y $|\lambda| \leq 1$.
- es simétrico si $-K = K$ (notemos que si $\mathbb{K} = \mathbb{R}$, entonces $K$ es convexo y simétrico si y solo si es convexo y balanceado).

Si $K$ es un cuerpo convexo balanceado en $\mathbb{K}^n$, entonces existe una única norma en $\mathbb{K}^n$, que notaremos $\| \ \|_K$, tal que $K$ es la bola unitaria con respecto a esa norma, es decir tal que $B_{(\mathbb{K}^n, \| \ \|_K)} = K$. Esta norma está dada por el funcional de Minkowski asociado a $K$, $\mu_K : \mathbb{K}^n \to \mathbb{R}$, definido por $\mu_K(x) = \inf\{t > 0 \text{ tales que } x \in tK\}$.

A su vez, cada norma $\| \ \|$ en $\mathbb{K}^n$ tiene un cuerpo convexo balanceado asociado, que es su bola unitaria, es decir podemos tomar como $K$ a $B_{(\mathbb{K}^n, \| \ \|)}$, de manera que la norma inducida por este cuerpo $K$ es la norma original $\| \ \|$ considerada en $\mathbb{K}^n$.

Por lo tanto, es equivalente hablar sobre normas o cuerpos convexos balanceados en $\mathbb{K}^n$. Notemos que si $\| \ \|$ es una norma en $\mathbb{R}^n$ cuya bola unitaria es el cuerpo $K$, entonces $\partial K = \{x \in \mathbb{R}^n / \ \|x\| = 1\}$ es la "esfera unitaria" asociada a dicha norma.

Por último, observemos que si $K$ es un cuerpo convexo y simétrico, entonces para cada $0 \leq \lambda \leq 1$ vale que

$$K = \lambda K - (1-\lambda)K. \tag{4}$$

En efecto, $K = \lambda K + (1-\lambda)K = \lambda K + (1-\lambda)(-K) = \lambda K - (1-\lambda)K$.

### Norma dual y cuerpo polar

Dado un cuerpo convexo balanceado $K \subseteq \mathbb{K}^n$, a cada $u \in \mathbb{K}^n$ podemos asociarle el funcional $\psi_u : (\mathbb{K}^n, \| \ \|_K) \to \mathbb{K}$ dado por $\psi_u(x) = \langle x, u \rangle = \sum_{j=1}^n x_j \overline{u_j}$. Asignándole a $u$ la norma de este funcional definimos $\| \ \|_*$, la *norma dual* a $\| \ \|_K$ en $\mathbb{K}^n$.

Más concretamente, definimos $\|u\|_* = \left\| \psi_u : (\mathbb{K}^n, \| \ \|_K) \to \mathbb{K} \right\|$. De esta manera nos queda que, para cualesquiera $u, v \in \mathbb{K}^n$

$$|\langle v, u \rangle| \leq \|v\|_K \|u\|_*.$$

Llamamos cuerpo polar de $K$ a la bola unitaria $K^* \subseteq \mathbb{K}^n$ asociada a la norma dual $\| \ \|_*$. Concretamente, tenemos

$$K^* = \{u \in \mathbb{K}^n / \ \|u\|_* \leq 1\} = \{u \in \mathbb{K}^n / \ |\langle x, u \rangle| \leq 1 \ \ \forall x \in K\}.$$

**Elipsoides en $\mathbb{K}^n$**

Un elipsoide en $\mathbb{K}^n$ es un conjunto dado por

$$\mathcal{E} = \left\{ x \in \mathbb{K}^n \,/\, \frac{|\langle x-p, v_1\rangle|^2}{a_1^2} + \cdots + \frac{|\langle x-p, v_n\rangle|^2}{a_n^2} \leq 1 \right\}$$

donde $\{v_1, \ldots, v_n\}$ es una base ortonormal de $\mathbb{K}^n$, $a_1, \ldots, a_n \in \mathbb{R}_{>0}$, y $p \in \mathbb{K}^n$ es el centro del elipsoide. En tal caso $\mathrm{Vol}(\mathcal{E}) = \mathrm{Vol}(B_2^n)\Pi_{i=1}^n a_i$.[2]

Si $u : \mathbb{K}^n \to \mathbb{K}^n$ es un isomorfismo lineal, entonces $u(B_2^n)$ es un elipsoide simétrico (cuyo centro está en el origen). Esto puede observarse al tomar la descomposición polar del isomorfismo lineal $u$. En efecto, escribiendo a $u$ como $u = PJ$, donde $P$ es un operador autoadjunto y positivo, y $J$ es una isometría de $l_2^n$ en $l_2^n$, vemos que

$$u(B_2^n) = PJ(B_2^n) = P(B_2^n) = \left\{ x \in \mathbb{K}^n \,/\, \frac{|\langle x, v_1\rangle|^2}{a_1^2} + \cdots + \frac{|\langle x, v_n\rangle|^2}{a_n^2} \leq 1 \right\}$$

donde los $v_1, \ldots, v_n$ forman una base ortonormal de autovectores de $P$, y los $a_1, \ldots, a_n$ son los correspondientes autovalores de $P$.

A su vez, se ve de la última igualdad que todo elipsoide simétrico puede obtenerse como la imagen de la bola euclídea por un isomorfismo (que puede tomarse positivo y autoadjunto) en $\mathbb{K}^n$.

Más en general, si $(E, \|\ \|)$ es un espacio normado de dimensión $n$, decimos que un elipsoide en $E$ es un conjunto de la forma $\mathcal{E} = p + u(B_2^n)$ donde $p \in E$, y $u : l_2^n \to (E, \|\ \|)$ es un isomorfismo lineal.

**Cápsula convexa**

Dado un conjunto $S \subseteq \mathbb{K}^n$, notamos $\mathrm{co}(S)$ a la cápsula convexa de $S$, definida como el menor conjunto convexo que contiene al conjunto $S$ (en el sentido de la inclusión). Más concretamente

$$\mathrm{co}(S) = \bigcap_{\substack{S \subseteq C \\ C \text{ convexo}}} C = \left\{ \alpha_1 x_1 + \cdots + \alpha_n x_n \text{ con } x_1, \ldots, x_n \in S,\ \alpha_i \geq 0,\ \sum_{i=1}^n \alpha_i = 1,\ n \in \mathbb{N} \right\}.$$

Por último, se puede ver que si $S$ y $T$ son conjuntos convexos en $\mathbb{K}^n$, entonces

$$\mathrm{co}(S \cup T) = \{\alpha s + \beta t \,/\, s \in S,\ t \in T,\ 0 \leq \alpha, \beta \leq 1,\ \alpha + \beta = 1\}. \tag{5}$$

[2]Dado $A \subseteq \mathbb{R}^n$ medible, denotamos con $\mathrm{Vol}(A)$ a la medida de Lebesgue $n$-dimensional de $A$.

# Capítulo 1

# Convergencia incondicional en espacios de Banach

## 1.1. Teorema de Dvoretzky-Rogers

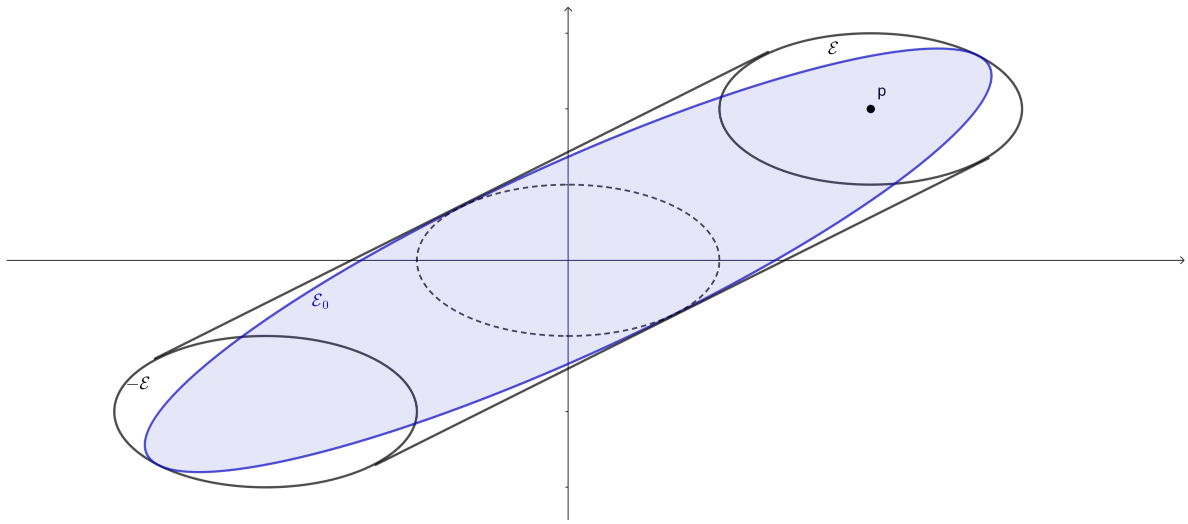
### 1.1.1. El elipsoide de John

En 1948 Fritz John publica, como regalo de los 60 para Courant, su artículo *Extremum problems with inequalities as subsidiary conditions* [Joh48] sobre el método de multiplicadores de Lagrange para una función de varias variables restringida a una región dada por inecuaciones. Este trabajo se enmarcaría en optimización, dentro del área de la programación no lineal (área cuyo nacimiento definitivo sería 3 años después, en 1951). Como una aplicación de este método, John demuestra ciertas propiedades del elipsoide de volumen mínimo que contiene un cierto cuerpo $S$ en $\mathbb{R}^n$. Así, pone sobre la escena matemática del momento los elipsoides extremales en relación a un cuerpo, objeto que usarán Dvoretzky y Rogers dos años más tarde (aunque en forma elemental, sin uso de las propiedades demostradas por John) para demostrar su celebrado lema.

Comenzamos entonces mostrando la existencia y unicidad del elipsoide de John[1] (aunque solo utilizaremos la existencia para demostrar la Proposición 1.1.7).

Para esto, veremos primero un lema, del cual quizás podríamos convencernos con el siguiente gráfico.

[1]Notemos que si bien John trabajo con el elipsoide de volumen mínimo que contiene un cierto cuerpo en [Joh48], suele llamarse (y llamaremos) elipsoide de John al elipsoide de volumen máximo contenido en un cuerpo dado.




**Lema 1.1.1.** *Sea $\mathcal{E} \subseteq \mathbb{R}^n$ un elipsoide con centro $p \neq 0$. Entonces existe un elipsoide simétrico $\mathcal{E}_0 \subseteq \mathrm{co}(\mathcal{E} \cup -\mathcal{E})$ con $\mathrm{Vol}(\mathcal{E}_0) > \mathrm{Vol}(\mathcal{E})$.*

*Demostración.* Podemos escribir $\mathcal{E} = p + u(B_2^n)$, con $u : \mathbb{R}^n \to \mathbb{R}^n$ isomorfismo. Por otro lado, al ser tanto $\mathcal{E}$ como $-\mathcal{E}$ convexos, tenemos que

$$\mathrm{co}(\mathcal{E} \cup -\mathcal{E}) = \{\lambda s + (1-\lambda)(-t) \,/\, s, t \in \mathcal{E},\ 0 \leq \lambda \leq 1\}$$

por lo observado en los preliminares (ver (5)). Es decir, dado $z \in co(\mathcal{E} \cup -\mathcal{E})$ existen $\lambda \in [0,1], x, y \in B_2^n$ tales que

$$z = \lambda(p + ux) + (1-\lambda)(-p - uy) = (2\lambda - 1)p + u(\lambda x - (1-\lambda)y).$$

Ahora, como $B_2^n$ es convexa y simétrica, $\lambda B_2^n - (1-\lambda)B_2^n = B_2^n$ para cada $\lambda \in [0,1]$ por la propiedad (4). De esta manera podemos escribir

$$\mathrm{co}(\mathcal{E} \cup -\mathcal{E}) = \bigcup_{\lambda \in [0,1]} (2\lambda - 1)p + u(B_2^n).$$

Si ahora tomamos $q \in \mathbb{R}^n$ tal que $u(q) = p$, entonces

$$\begin{aligned}
\mathrm{co}(\mathcal{E} \cup -\mathcal{E}) &= \bigcup_{\lambda \in [0,1]} (2\lambda - 1)\,u(q) + u(B_2^n) = \bigcup_{\lambda \in [0,1]} u((2\lambda - 1)q + B_2^n) \\
&= u\left(\bigcup_{\lambda \in [0,1]} (2\lambda - 1)q + B_2^n\right) = u\left(\bigcup_{\lambda \in [0,1]} (2\lambda - 1)q + \lambda B_2^n - (1-\lambda)B_2^n\right) \\
&= u\left(\bigcup_{\lambda \in [0,1]} \lambda(q + B_2^n) + (1-\lambda)(-q - B_2^n)\right) = u\Big(\mathrm{co}\Big((q + B_2^n) \cup (-q + B_2^n)\Big)\Big).
\end{aligned}$$

Así, basta con ver que si $q \in \mathbb{R}^n$, $q \neq 0$, entonces existe un elipsoide simétrico $\mathcal{E}_1 \subseteq \mathrm{co}\Big((q + B_2^n) \cup (-q + B_2^n)\Big)$ tal que $\mathrm{Vol}(\mathcal{E}_1) > \mathrm{Vol}(B_2^n)$, ya que en ese caso tomaríamos $\mathcal{E}_0 = u(\mathcal{E}_1)$, de manera que $\mathcal{E}_0$ es un elipsoide simétrico, y $\mathrm{Vol}(\mathcal{E}_0) = |\det(u)|\mathrm{Vol}(\mathcal{E}_1) > |\det(u)|\mathrm{Vol}(B_2^n) = \mathrm{Vol}(\mathcal{E})$[2].

Para esto, como los únicos objetos involucrados en este problema son $q + B_2^n, -q + B_2^n$ y $B_2^n$, podemos hacer una rotación de manera que sea $q = (q_1, 0, \ldots, 0)$, con $q_1 > 0$, a fin de facilitar la escritura.

Afirmamos entonces que el elipsoide

$$\mathcal{E}_1 : \frac{x_1^2}{(1+q_1)^2} + x_2^2 + \cdots + x_n^2 \leq 1$$

está incluido en la cápsula convexa de $-q_1 e_1 + B_2^n$ y $q_1 e_1 + B_2^n$. En efecto, si $(x_1, \ldots, x_n) \in \mathcal{E}_1$, entonces $|x_1| \leq (1+q_1)\Big(1 - (x_2^2 + \cdots + x_n^2)\Big)^{1/2} \leq \Big(1 - (x_2^2 + \cdots + x_n^2)\Big)^{1/2} + q_1$.

Por lo que existe $\lambda \in [0,1]$ tal que

$$x_1 = \lambda\Big(-(1 - (x_2^2 + \cdots + x_n^2))^{1/2} - q_1\Big) + (1-\lambda)\Big((1 - (x_2^2 + \cdots + x_n^2))^{1/2} + q_1\Big)$$

(ya que un intervalo $[a,b]$ es la cápsula convexa de sus puntos extremos). Así, tomando

$$\begin{aligned} y &= (y_1, x_2, \ldots, x_n) & &\text{con } y_1 = -(1 - (x_2^2 + \cdots + x_n^2))^{1/2} - q_1, \\ z &= (z_1, x_2, \ldots, x_n) & &\text{con } z_1 = (1 - (x_2^2 + \cdots + x_n^2))^{1/2} + q_1, \end{aligned}$$

tenemos que $y \in -q_1 e_1 + B_2^n$, $z \in q_1 e_1 + B_2^n$ y $x = \lambda y + (1-\lambda) z$.

Como $\mathrm{Vol}(\mathcal{E}_1) = (1+q_1)\mathrm{Vol}(B_2^n) > \mathrm{Vol}(B_2^n)$, tenemos lo buscado.

□

**Observación 1.1.2.**

(i) Dados $\mathcal{E} \subseteq \mathbb{R}^n$ elipsoide simétrico y $u : \mathbb{R}^n \to \mathbb{R}^n$ isomorfismo tal que $u(B_2^n) = \mathcal{E}$, notemos que cualquier otro isomorfismo $\widetilde{u} : \mathbb{R}^n \to \mathbb{R}^n$ tal que $\widetilde{u}(B_2^n) = \mathcal{E}$ cumple que $\widetilde{u} = u \circ G$, donde la matriz canónica de $G$ es ortogonal, es decir: $G : l_2^n \to l_2^n$ es una isometría. En efecto $G(B_2^n) = u^{-1} \circ \widetilde{u}(B_2^n) = u^{-1}(\mathcal{E}) = B_2^n$.

(ii) Dados $K$ un cuerpo convexo simétrico en $\mathbb{R}^n$, y $\mathcal{E}$ un elipsoide simétrico incluido en $K$, si tomamos $u : \mathbb{R}^n \to \mathbb{R}^n$ tal que $u(B_2^n) = \mathcal{E}$ , entonces $u$ considerado como un isomorfismo entre los espacios normados $l_2^n$ y $(\mathbb{R}^n, \| \|_K)$, $u : l_2^n \to (\mathbb{R}^n, \| \|_K)$, es un operador de norma menor o igual a 1. En efecto, tenemos que $u(B_2^n) = \mathcal{E} \subseteq K = B_{(\mathbb{R}^n, \| \|_K)}$, por lo que si $\|x\|_2 \leq 1$, entonces $\|u(x)\|_K \leq 1$.

Recordemos previo a la demostración del siguiente teorema que si $u : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal, se define la traza de $u$ como $\mathrm{tr}(u) = \mathrm{tr}([u]_\Gamma)$ con $\Gamma$ una base de $\mathbb{R}^n$. Además, si $g(t) = \det(Id + t\,u)$, entonces se tiene que $g'(0) = \mathrm{tr}(u)$.

[2] Recordemos que si $u : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal, entonces $\det(u) = \det([u]_\Gamma)$ donde $\Gamma$ es una base de $\mathbb{R}^n$. Además, si $A \subseteq \mathbb{R}^n$ es medible, entonces $u(A) \subseteq \mathbb{R}^n$ es medible, con $\mathrm{Vol}_n(u(A)) = |\det(u)|\mathrm{Vol}_n(A)$.

**Teorema 1.1.3.** *Sea $K$ un cuerpo convexo y simétrico en $\mathbb{R}^n$. Entonces existe un único elipsoide $\mathcal{E} \subseteq K$ (llamado el elipsoide de John de $K$) tal que*

$$\mathrm{Vol}(\mathcal{E}) = \sup_{\substack{\widetilde{\mathcal{E}} \subseteq K \\ \widetilde{\mathcal{E}} \text{ elipsoide}}} \mathrm{Vol}(\widetilde{\mathcal{E}}).$$

*Además, $\mathcal{E}$ resulta ser simétrico. Más aún, si $u_0 : l_2^n \to (\mathbb{R}^n, \| \|_K)$ es tal que $u_0(B_2^n) = \mathcal{E}$, entonces vale que*

$$\mathrm{tr}(u_0^{-1} v) \leq n \|v\| \quad \forall\, v : l_2^n \to (\mathbb{R}^n, \| \|_K).$$

*Demostración.* Notemos primero que, de existir un elipsoide $\mathcal{E}$ de volumen máximo contenido en $K$, este debe ser simétrico. En efecto, si $\mathcal{E}$ no fuera simétrico, tendríamos que como $K$ es simétrico $\mathcal{E} \cup -\mathcal{E} \subseteq K$, y al ser $K$ convexo valdría $\mathrm{co}(\mathcal{E} \cup -\mathcal{E}) \subseteq K$, pero entonces, por el Lema 1.1.1, tendríamos un elipsoide simétrico $\mathcal{E}_1 \subseteq \mathrm{co}(\mathcal{E} \cup -\mathcal{E}) \subseteq K$ de volumen mayor al de $\mathcal{E}$, lo cual sería un absurdo.

Sabiendo entonces que nuestro elipsoide debe ser simétrico, consideremos en $\mathcal{B} = B_{\mathcal{L}(l_2^n, (\mathbb{R}^n, \| \|_K))}$ la función determinante.

Al ser $\mathcal{B}$ compacto (pues $\mathcal{L}\left(l_2^n, (\mathbb{R}^n, \| \|_K)\right)$ es un espacio normado de dimensión finita) y la función $\det : \mathcal{B} \to \mathbb{R}$ continua, tenemos que det alcanza su máximo (y mínimo) absoluto en $\mathcal{B}$. Digamos que este máximo se alcanza en $u_0 \in \mathcal{B}$.

Notemos que $u_0$ debe ser inversible, ya que sino $\det(u_0) = 0$ (y existen $v \in \mathcal{B}$ tales que $\det(v) > 0$, ¡hay isomorfismos!). Más aún, por el mismo motivo debe ser $\det(u_0) > 0$. Además, claramente debe ser $\|u_0\| = 1$, pues sino $\det\left(\frac{1}{\|u_0\|} u_0\right) = \frac{1}{\|u_0\|^n} \det(u_0) > \det(u_0)$.

Entonces, por la Observación 1.1.2 (II), es claro que $\mathcal{E} = u_0(B_2^n)$ es un elipsoide contenido en $K$ de volumen máximo, ya que $\mathrm{Vol}(\mathcal{E}) = c_n |\det(u_0)| = c_n \det(u_0)$, donde $c_n = \mathrm{Vol}_n(B_2^n)$ (notemos que el máximo absoluto en $\mathcal{B}$ de la función $|\det|$ es el mismo que el de det, ya que la imagen de la función det en $\mathcal{B}$ es un intervalo simétrico, pues $v \in \mathcal{B} \iff v \circ U \in \mathcal{B}$, con $U$ ortogonal de determinante -1).

Así, tenemos la existencia. Veamos ahora que, si $v \in \mathcal{L}\left(l_2^n, (\mathbb{R}^n, \| \|_K)\right)$, entonces $\mathrm{tr}(u^{-1} v) \leq n \|v\|$.

Para esto claramente podemos suponer que $v \neq 0$. Entonces notemos que para $t$ pequeño (digamos $|t| < 1/\|v\|$), tenemos que $u_0 + tv \neq 0$, por lo que podemos escribir

$$\det\left(\frac{Id_n + t u_0^{-1} v}{\|u_0 + tv\|}\right) = \det(u_0^{-1}) \det\left(\frac{u_0 + tv}{\|u_0 + tv\|}\right) \leq \det(u_0^{-1}) \det(u_0) = 1.$$

De donde vemos que

$$\det(Id_n + t u_0^{-1} v) \leq \|u_0 + tv\|^n \leq (\|u_0\| + |t| \|v\|)^n = (1 + |t| \|v\|)^n.$$

Luego, $\det(Id_n + t u_0^{-1} v) \leq (1 + t\|v\|)^n$ para todo $t \in [0, \varepsilon)$ con $\varepsilon = 1/\|v\|$.

Definiendo $g : [0, \varepsilon) \to \mathbb{R}$ como $g(t) = \det(Id_n + t u_0^{-1} v)$ y $h : [0, \varepsilon) \to \mathbb{R}$ como $h(t) = (1 + t\|v\|)^n$, tenemos que $g(t) \leq h(t)$ para todo $t \in [0, \varepsilon)$ y que $g(0) = h(0) = 1$, por lo que debe ser (de la misma definición de derivada) $g'(0) \leq h'(0)$, es decir $\mathrm{tr}(u_0^{-1} v) \leq n \|v\|$ .

Veamos ahora la unicidad de $\mathcal{E}$. Supongamos que $\widetilde{\mathcal{E}}$ es un elipsoide simétrico contenido en $K$ de volumen máximo, y sea $v_0 \in \mathcal{B}$ tal que $v_0(B_2^n) = \widetilde{\mathcal{E}}$. Entonces sabemos que en $v_0$ se alcanza el máximo absoluto de $|\det|$, y podemos asumir, precomponiendo con una transformación ortogonal de determinante -1 de ser necesario, que en $v_0$ se alcanza el máximo de det. Queremos ver que entonces $u_0^{-1}v_0 : l_2^n \to l_2^n$ es una isometría.

Notemos que, de la misma manera que $u_0$, $v_0$ debe satisfacer que $\mathrm{tr}(v_0^{-1}w) \leq n\|w\|$ para todo $w : l_2^n \to (\mathbb{R}^n, \|\ \|_K)$.

Ahora, para ver que $u_0^{-1}v_0$ es una isometría en $l_2^n$, escribámoslo en forma polar: $u_0^{-1}v_0 = PU$, donde $U$ es una isometría y $P$ es un operador positivo (ya que $u_0^{-1}v_0$ es inversible) y autoadjunto. ¡Veremos que $P$ es la identidad! Para esto, como $P$ es diagonalizable, basta con ver que todos sus autovalores son 1. Sean entonces $\lambda_1, \ldots, \lambda_n > 0$ los autovalores de $P$ contados con multiplicidad. Tenemos que

$$\sum_{i=1}^{n} \lambda_i = \mathrm{tr}(P) = \mathrm{tr}(PU\,U^{-1}) = \mathrm{tr}(u_0^{-1}v_0U^{-1}) \leq n\left\|v_0U^{-1}\right\| = n\|v_0\| = n,$$
$$\sum_{i=1}^{n} \lambda_i^{-1} = \mathrm{tr}(P^{-1}) = \mathrm{tr}(UU^{-1}P^{-1}) = \mathrm{tr}(U\,v_0^{-1}u_0) = \mathrm{tr}(v_0^{-1}\,u_0U) \leq n\|u_0U\| = n\|u_0\| = n.$$

Sumando ambas desigualdades obtenemos

$$\sum_{i=1}^{n} \lambda_i + \lambda_i^{-1} \leq 2n.$$

Ahora bien, como cada término de esta suma es mayor o igual a 2, necesariamente debe ser $\lambda_i + \lambda_i^{-1} = 2$ para todo $1 \leq i \leq n$, por lo que $\lambda_i = 1$ para todo $1 \leq i \leq n$. En conclusión, $P = Id_n$ y, en consecuencia, $v_0 = u_0U$, con $U$ unitario en $l_2^n$. Luego,

$$\widetilde{\mathcal{E}} = v_0(B_2^n) = u_0(U(B_2^n)) = u_0(B_2^n) = \mathcal{E}. \quad \square$$

**Observación 1.1.4.**

(i) Si bien la demostración de la unicidad del elipsoide de John puede darse utilizando argumentos "más geométricos" (ver, por ejemplo, [AGM15]), seguimos los argumentos usados arriba por nuestra devoción a [DJT95].

(ii) Notemos que si $\mathcal{E}$ es el elipsoide de John de $K \subseteq \mathbb{R}^n$, entonces necesariamente debe existir algún punto de contacto entre $\partial K$ y $\partial\mathcal{E}$, ya que de lo contrario, tendríamos $\mathcal{E} \subsetneq K$, por lo que sería $d(\mathcal{E}, \partial K) > 0$, y existiría un $\delta > 0$ tal que $(1+\delta)\mathcal{E} \subseteq K$, lo cual es absurdo. Más aún, como veremos en instantes, deben existir al menos $2n$ puntos de contacto entre $\partial\mathcal{E}$ y $\partial K$.

**Definición 1.1.5.** Decimos que un cuerpo convexo y simétrico $K \subseteq \mathbb{R}^n$ *está en posición de John* si el elipsoide de John de $K$ coincide con la bola euclídea $B_2^n$.

**Observación 1.1.6.**

(i) Si $K$ está en posición de John, podemos tomar como $u_0 : l_2^n \to (\mathbb{R}^n, \| \, \|_K)$ cualquier isomorfismo que cumpla $u_0(B_2^n) = B_2^n \subseteq K$ (es decir, cualquier isometría de $l_2^n$), en particular podemos tomar a la identidad de $\mathbb{R}^n$, $Id_n : l_2^n \to (\mathbb{R}^n, \| \, \|_K)$. Así, tenemos que $tr(v) \leq n\|v\|$ para todo $v : l_2^n \to (\mathbb{R}^n, \| \, \|_K)$.

(ii) En particular, notemos que si $K$ está en posición de John, entonces $\|x\|_K \leq \|x\|_2$ para todo $x \in \mathbb{R}^n$ (ya que $\left\| Id_n : l_2^n \to (\mathbb{R}^n, \| \, \|_K) \right\| = 1$). Comentemos que además se cumple que $\|x\|_2 \leq \sqrt{n}\, \|x\|_K$ (para una demostración de insuperable elegancia sobre este hecho, ver [AGM15, Teorema 2.1.3]). Es decir, si $K$ esta en posición de John entonces vale que $B_2^n \subseteq K \subseteq \sqrt{n}\, B_2^n$.

En particular, dada una norma cualquiera $\| \, \|$ en $\mathbb{R}^n$, se tiene que $d_{BM}((\mathbb{R}^n, \| \, \|), l_2^n) \leq \sqrt{n}$.

(iii) Si $K \subseteq \mathbb{R}^n$ es un cuerpo convexo y simétrico, y $\mathcal{E}$ es su elipsoide de John, entonces podemos tomar un isomorfismo $v : \mathbb{R}^n \to \mathbb{R}^n$ tal que $v(\mathcal{E}) = B_2^n$. De esta manera, el cuerpo $v(K)$ está en posición John (ya que $\widetilde{\mathcal{E}} \subseteq K \iff v(\widetilde{\mathcal{E}}) \subseteq v(K)$, y $v$ preserva relaciones de volumen), y $v : (\mathbb{R}^n, \| \, \|_K) \to \left(\mathbb{R}^n, \| \, \|_{v(K)}\right)$ es una isometría.

Así, dada una norma $\| \, \|_{K_1}$ en $\mathbb{R}^n$ con bola unitaria $K_1$, podemos tomar una norma $\| \, \|_{K_2}$ con bola unitaria $K_2$, de manera que $K_2$ está en posición de John y los espacios $\left(\mathbb{R}^n, \| \, \|_{K_1}\right)$ y $\left(\mathbb{R}^n, \| \, \|_{K_2}\right)$ resulten isométricos.

(iv) En general, si $K$ es un cuerpo convexo y simétrico en $\mathbb{R}^n$ y $T : \mathbb{R}^n \to \mathbb{R}^n$ es un isomorfismo lineal, decimos que $T(K)$ es una *posición* del cuerpo $K$. Además de la posición de John, existen distintas posiciones clásicas de un cuerpo convexo $K$ (ver [AGM15]). En general, estas están definidas en términos de que se satisfaga alguna condición extremal, como ocurre por ejemplo con la posición de John.

### 1.1.2. Lema de Dvoretzky-Rogers

Para la prueba del lema de Dvoretzky-Rogers, seguiremos los pasos de la demostración original en [DR50], que solo hará uso de la existencia del elipsoide de John.

**Proposición 1.1.7** (lema de Dvoretzky-Rogers)**.** *Sea $K \subseteq \mathbb{R}^n$ un cuerpo convexo y simétrico en posición de John. Entonces existen una base ortonormal $\{z_1, \dots z_n\}$ de $l_2^n$ y una base $\{A_1, \dots, A_n\}$ de $\mathbb{R}^n$ tal que $\left\|A_j\right\|_K = \left\|A_j\right\|_2 = 1$ para todo $1 \leq j \leq n$, de manera que:*

(a) $A_i = \displaystyle\sum_{j=1}^{i} a_{ij} z_j$, *con* $a_{ii}^2 \geq 1 - \dfrac{i-1}{n}$ *para cada* $1 \leq i \leq n$.

(b) $\|A_i - z_i\|_K^2 \leq \|A_i - z_i\|_2^2 \leq \dfrac{2(i-1)}{n}$ *para cada* $1 \leq i \leq n$.

(c) *Si* $1 \leq r \leq n$*, para cualesquiera escalares* $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ *vale que*

$$\left\| \sum_{i=1}^{r} \alpha_i A_i \right\|_K \leq \left( 1 + \sqrt{\frac{r(r-1)}{n}} \right) \|(\alpha_1, \dots, \alpha_r)\|_2 .$$

*Demostración.* Veamos por inducción que existen $\{z_1,\dots,z_n\}$ y $\{A_1,\dots,A_n\}$ como en el enunciado, que cumplen (a). Luego deduciremos (b) y (c).

Para $i=1$, tomamos (con permiso de la Observación 1.1.4 (II) de la sección anterior) como $z_1 = A_1$ algún punto de contacto entre $\partial K$ y $\partial B_2^n = S^{n-1}$.

Supongamos ahora que para algún $1 \le k < n$ tenemos $z_1,\dots,z_k$ ortonormales y $A_1,\dots,A_k \in \partial K \cap \partial B_2^n$ l.i. que satisfacen (a). Entonces completamos $z_1,\dots,z_k$ a una base ortonormal de $l_2^n$ con $\tilde{z}_{k+1},\dots,\tilde{z}_n$, y notamos $\beta_1(x),\dots,\beta_n(x)$ a las coordenadas de un vector $x\in\mathbb{R}^n$ respecto de esta base (es decir, $\beta_i(x)$ es el producto interno entre $x$ y el $i$–ésimo elemento de la base ortonormal).

Consideremos ahora, para cada $\delta > 0$ el elipsoide

$$\mathcal{E}_\delta : \frac{\beta_1(x)^2}{(1+\delta)^{k-n}} + \cdots + \frac{\beta_k(x)^2}{(1+\delta)^{k-n}} + \frac{\beta_{k+1}(x)^2}{(1+\delta+\delta^2)^k} + \cdots + \frac{\beta_n(x)^2}{(1+\delta+\delta^2)^k} \le 1, \tag{1.1}$$

entonces

$$\mathrm{Vol}(\mathcal{E}_\delta) = \mathrm{Vol}(B_2^n)\Big((1+\delta)^{\frac{k-n}{2}}\Big)^k \Big((1+\delta+\delta^2)^{\frac{k}{2}}\Big)^{n-k} = \mathrm{Vol}(B_2^n)\frac{(1+\delta+\delta^2)^{\frac{k(n-k)}{2}}}{(1+\delta)^{\frac{k(n-k)}{2}}} > \mathrm{Vol}(B_2^n).$$

Por lo tanto, debe ser $\mathcal{E}_\delta \nsubseteq K$, por lo que existirá $p_\delta \in \partial K \cap \mathcal{E}_\delta$. Como $\partial K$ es compacto, tomando $\delta_m = 1/m$ y una subsucesión adecuada, obtenemos una sucesión convergente $(q_l)_{l\in\mathbb{N}}$ en $\partial K$, donde $q_l = p_{m_l}$ para todo $l\in\mathbb{N}$, de manera que $\delta_{m_l} \underset{l\to+\infty}{\longrightarrow} 0$. Sea $A_{k+1}\in\partial K$ el límite de $(q_l)_{l\in\mathbb{N}}$. Entonces $1 = \|A_{k+1}\|_K \le \|A_{k+1}\|_2$. Por otro lado, como para cada $l\in\mathbb{N}$ tenemos que $q_l \in \mathcal{E}_{\delta_{m_l}}$, tomando límite en (1.1) resulta

$$\beta_1(A_{k+1})^2 + \cdots + \beta_n(A_{k+1})^2 \le 1.$$

Es decir $\|A_{k+1}\|_2 \le 1$, de donde $A_{k+1}\in S^{n-1}$.

Ahora bien, como para cada $l\in\mathbb{N}$ es $1 = \|q_l\|_K \le \|q_l\|_2$ , a partir de (1.1) podemos escribir

$$\Big((1+\delta_{m_l})^{n-k}-1\Big)\Big(\beta_1(q_l)^2\cdots+\beta_k(q_l)^2\Big) + \Big((1+\delta_{m_l}+\delta_{m_l}^2)^{-k}-1\Big)\Big(\beta_{k+1}(q_l)^2+\cdots+\beta_n(q_l)^2\Big) \le 0.$$

Dividiendo por $\delta_{m_l}$ y tomando límite con $l\to+\infty$ tenemos

$$(n-k)\Big(\beta_1(A_{k+1})^2+\cdots+\beta_k(A_{k+1})^2\Big) - k\Big(\beta_{k+1}(A_{k+1})^2+\cdots+\beta_n(A_{k+1})^2\Big) \le 0. \tag{1.2}$$

Llamemos

$$a_{k+1\,k+1} = \Big(\beta_{k+1}(A_{k+1})^2+\cdots+\beta_n(A_{k+1})^2\Big)^{1/2} \quad \text{(de (1.2) se ve que debe ser } a_{k+1\,k+1}>0\text{), y}$$

$$z_{k+1} = \frac{1}{a_{k+1\,k+1}}\left(\beta_{k+1}(A_{k+1})\tilde{z}_{k+1}+\cdots+\beta_n(A_{k+1})\tilde{z}_n\right) \in \langle z_1,\dots,z_k\rangle^\perp.$$

Entonces, si para cada $1\le j\le k$ tomamos $a_{k+1\,j} = \beta_j(A_{k+1})$, tenemos que $A_{k+1} = \sum_{j=1}^{k+1} a_{k+1\,j}\, z_j$, con $\{z_1,\dots,z_{k+1}\}$ ortonormal, de manera que a partir de (1.2) escribimos

$$(n-k)\sum_{j=1}^k a_{k+1\,j}^2 \le k\, a_{k+1\,k+1}^2 \iff (n-k)(1-a_{k+1\,k+1}^2) \le k\, a_{k+1\,k+1}^2 \iff$$

$$n-k \le n\, a_{k+1\,k+1}^2 \iff 1-\frac{k}{n} \le a_{k+1\,k+1}^2.$$

Que es lo que deseábamos. Así, por inducción tenemos que existen $\{z_1,\dots,z_n\}$ y $\{A_1,\dots,A_n\}$ como en el enunciado, que satisfacen (a).

Veamos que se cumplen (b) y (c). Para (b) basta con ver que, si $1 \le i \le n$, como $0 < a_{ii} \le 1$

$$\|A_i - z_i\|_K^2 \le \|A_i - z_i\|_2^2 = \sum_{j=1}^{i-1} a_{ij}^2 + (1-a_{ii})^2 = (1-a_{ii}^2) + 1 - 2a_{ii} + a_{ii}^2 =$$

$$2(1-a_{ii}) \le 2(1-a_{ii})(1+a_{ii}) = 2(1-a_{ii}^2) \le \frac{2(i-1)}{n}.$$

Veamos (c). Sean $1 \le r \le n$ y $\alpha_1,\dots,\alpha_r \in \mathbb{R}$, entonces

$$\begin{aligned}
&\left\|\sum_{i=1}^r \alpha_i A_i\right\|_K \le \left\|\sum_{i=1}^r \alpha_i A_i\right\|_2 \le \left\|\sum_{i=1}^r \alpha_i z_i\right\|_2 + \left\|\sum_{i=1}^r \alpha_i (A_i - z_i)\right\|_2 \\
&\le \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2} + \sum_{i=1}^r |\alpha_i|\,\|A_i - z_i\|_2 \le \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2} + \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2}\left(\sum_{i=1}^r \|A_i - z_i\|_2^2\right)^{1/2} \\
&\le \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2}\left(1 + \left(\sum_{i=1}^r \frac{2(i-1)}{n}\right)^{1/2}\right) = \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2}\left(1 + \left(\frac{2}{n}\,\frac{r(r-1)}{2}\right)^{1/2}\right) \\
&= \left(\sum_{i=1}^r \alpha_i^2\right)^{1/2}\left(1 + \sqrt{\frac{r(r-1)}{n}}\right).
\end{aligned}$$

□

**Observación 1.1.8.**

(I) En [DR50] obtienen en el item (c) una constante ligeramente distinta. Más precisamente, obtienen que para cada $1 \le r \le n$

$$\left\|\sum_{i=1}^r \alpha_i A_i\right\|_K \le \sqrt{2 + \frac{r(r-1)}{n}}\,\|(\alpha_1,\dots,\alpha_r)\|_2,$$

llevando a cabo ciertos artilugios indexatorios que hemos evitado en esta ocasión.

(II) Notemos que $\left\|A_j\right\|_* = 1$ para cada uno de los $A_j$ del lema anterior. Es decir que los $A_j$ *también* viven en el borde del cuerpo polar a $K$. Esto se debe a que si $K$ está en posición de John, entonces $\partial K \cap \partial B_2^n \subseteq \partial K^*$.

En efecto, si $A \in \partial K \cap \partial B_2^n$, como $K$ es convexo, por la versión geométrica del teorema de Hahn-Banach, debe existir algún hiperplano $H_{v,\alpha} = \{x \in \mathbb{R}^n / \langle x, v\rangle = \alpha\}$ que pase por $A$ y que separe a $\mathring{K}$ (el interior de $K$) de $A$, es decir tal que $\langle x, v\rangle < \alpha = \langle A, v\rangle$ para todo $x \in \mathring{K}$ . Ahora bien, como $B_2^n \subseteq K$, este hiperplano realizará el mismo trabajo para $\mathring{B}_2^n$ y $A \in \partial B_2^n$. Pero como existe un único hiperplano que separa a $\mathring{B}_2^n$ de $A$, a saber, $H = \{x \in \mathbb{R}^n / \langle x, A\rangle = 1\}$, tenemos que debe ser $H_{v,\alpha} = H$, de donde $\langle x, A\rangle \le 1$ para todo $x \in K$. Luego, $\|A\|_* = 1$ (ya que $\langle A, A\rangle = 1$) .

(iii) Teniendo en cuenta la Observación 1.1.6 (iii), dado un cuerpo convexo simétrico $K \subseteq \mathbb{R}^n$ y $\mathcal{E}$ su elipsoide de John, tomando la isometría correspondiente entre $(\mathbb{R}^n, \| \, \|_K)$ y $(\mathbb{R}^n, \| \, \|_{\widetilde{K}})$, donde $\widetilde{K}$ está en posición de John, aplicando lo visto en la Proposición 1.1.7 sobre $\widetilde{K}$, y "volviendo" a $K$, arribamos a que existen $A_1, \ldots, A_n \in \partial K \cap \partial \mathcal{E}$ tales que

$$\left\| \sum_{i=1}^{r} \alpha_i A_i \right\|_K \leq \left( 1 + \sqrt{\frac{r(r-1)}{n}} \right) \|(\alpha_1, ..., \alpha_r)\|_2$$

para $1 \leq r \leq n$, y $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$.

(iv) En particular, obtenemos que, como $\{A_1, \ldots, A_n\}$ es l.i., si $K$ es un cuerpo convexo simétrico y $\mathcal{E}$ es su elipsoide de John, entonces debe haber como mínimo $2n$ puntos de contacto entre $\partial K$ y $\partial \mathcal{E}$. A saber: $\{\pm A_1, \ldots, \pm A_n\}$.

Pasando de un espacio normado $(E, \| \, \|_E)$ de dimensión $n$ a $\mathbb{R}^n$ por medio de un isomorfismo $T : E \to \mathbb{R}^n$, y considerando en $\mathbb{R}^n$ la norma (isométrica a $\| \, \|_E$) inducida por $T$, llegamos entonces a la siguiente proposición.

**Proposición 1.1.9.** *Si $(E, \| \, \|_E)$ es un espacio normado de dimensión $n$, entonces existe una base de vectores unitarios $x_1, \ldots, x_n$ de $E$ tal que para cada $1 \leq r \leq n$ y $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ vale que*

$$\left\| \sum_{i=1}^{r} \alpha_i x_i \right\|_E \leq \left( 1 + \sqrt{\frac{r(r-1)}{n}} \right) \|(\alpha_1, ..., \alpha_r)\|_2 .$$

A partir de lo obtenido en la Proposición 1.1.7, conseguimos el siguiente corolario.

**Corolario 1.1.10.** *Sea $(E, \| \, \|_E)$ un espacio normado de dimensión $n$, entonces existe una base de vectores $x_1, \ldots, x_n \in B_E$ tales que:*

*(a)* $\left\| \sum_{i=1}^{n} \alpha_i \, x_i \right\|_E \leq \|(\alpha_1, ..., \alpha_n)\|_2$ *para cualesquiera* $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$.

*(b)* $\sqrt{1 - \frac{r-1}{n}} \leq \|x_r\|_E \leq 1$ *para cada* $1 \leq r \leq n$.

*Demostración.* Considerando una norma isométrica a la de $(E, \| \, \|_E)$ en $\mathbb{R}^n$, cuya bola unitaria $K$ este en posición de John, basta con tomar en la Proposición 1.1.7 a $x_i = z_i$. Entonces, dados $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$

$$\left\| \sum_{i=1}^{n} \alpha_i x_i \right\|_K \leq \left\| \sum_{i=1}^{n} \alpha_i \, z_i \right\|_2 = \left( \sum_{i=1}^{n} \alpha_i^2 \right)^{1/2} .$$

De aquí, en particular, tenemos que cada $x_i \in B_E$. Por otro lado, tenemos que

$$a_{rr} = |\langle z_r, A_r \rangle| = |\langle x_r, A_r \rangle| \leq \|x_r\|_K \, \|A_r\|_* = \|x_r\|_K \, . \, 1$$

y como $a_{rr} \geq \sqrt{1 - \frac{r-1}{n}}$, se obtiene (b). □

En relación a la Proposición 1.1.9 el corolario anterior nos dice que podemos obtener una base de $E$ cuyos vectores no necesariamente son unitarios, sino que viven dentro de $B_E$, satisfaciendo la cota (b) sobre sus normas, y como contrapartida de esta "relajación", tenemos que se satisface (a).

### 1.1.3. Teorema de Dvoretzky-Rogers

Recordemos la noción de sumabilidad incondicional de una sucesión.

**Definición 1.1.11.** Decimos que una sucesión $(x_n)$ en un espacio de Banach $X$ es incondicionalmente sumable si para toda biyección $\sigma : \mathbb{N} \to \mathbb{N}$ la serie $\sum_{j=1}^{+\infty} x_{\sigma(j)}$ es convergente.

Para demostrar el Teorema 1.1.13 utilizaremos el lema de Dvoretzky-Rogers "en su versión del Corolario 1.1.10", aunque podríamos igualmente usarlo "en su versión original" (Proposición 1.1.7).

Más precisamente, usaremos el siguiente lema, que se obtiene al tomar los primeros $n$ vectores que nos entrega el Corolario 1.1.10 en un espacio de dimensión $2n$.

**Lema 1.1.12.** *Si $(E, \| \ \|_E)$ es un espacio normado de dimensión $2n$, entonces existen $x_1, \ldots, x_n \in B_E$ l.i. tales que $\|x_i\|_E \geq 1/\sqrt{2}$ para todo $1 \leq i \leq n$, y $\left\|\sum_{i=1}^{n} \alpha_i x_i\right\|_E \leq \left(\sum_{i=1}^{n} |\alpha_i|^2\right)^{1/2}$ para cualesquiera escalares $\alpha_1, \ldots, \alpha_n$.*

**Teorema 1.1.13** (teorema de Dvoretzky-Rogers)**.** *Sea $X$ un espacio de Banach de dimensión infinita. Dada una sucesión $(\lambda_n)_{n\in\mathbb{N}} \in l_2$, existe una sucesión incondicionalmente sumable $(x_n)_{n\in\mathbb{N}}$ en $X$ tal que $\|x_n\| = |\lambda_n|$ para todo $n \in \mathbb{N}$.*

*Demostración.* Como $(\lambda_n)_{n\in\mathbb{N}} \in l_2$, tenemos que $\sum_{n=1}^{+\infty} |\lambda_n|^2 < +\infty$, por lo que la sucesión de sumas parciales de dicha serie es de Cauchy. Así, podemos armar una sucesión de naturales estrictamente creciente $(n_k)_{k\in\mathbb{N}}$ de manera que para cada $k \in \mathbb{N}$ tenemos que $\sum_{n=n_k}^{N} |\lambda_n|^2 < 1/2^{2k}$ si $N > n_k$.

Ahora bien, como $X$ es de dimensión infinita, usando el Lema 1.1.12, podemos tomar una sucesión de vectores l.i. $(y_n)_{n\in\mathbb{N}}$ en $X$ de manera que $\left\|y_n\right\| \geq 1/\sqrt{2}$ para todo $n \in \mathbb{N}$, y $\left\|\sum_{j=n_k}^{N} \alpha_j y_j\right\| \leq \left(\sum_{j=n_k}^{N} |\alpha_j|^2\right)^{1/2}$ para todo $n_k \leq N < n_{k+1}$.

En efecto, inicialmente escogemos $y_1, \ldots, y_{n_1-1} \in X$ unitarios y l.i., y notamos $E_0 = \langle y_1, \ldots, y_{n_1-1}\rangle$. Luego, para $k = 1$, tomamos en $X$ un subespacio $E_1$ en suma directa con $E_0$, y de dimensión finita $2(n_2 - n_1)$ , donde podemos encontrar vectores l.i $y_{n_1}, \ldots, y_{n_2-1} \in B_X$ de norma mayor o igual a $1/\sqrt{2}$ tales que $\left\|\sum_{j=n_1}^{n_2-1} \alpha_j y_j\right\| \leq \left(\sum_{j=n_1}^{n_2-1} |\alpha_j|^2\right)^{1/2}$. Prosiguiendo de igual manera, inductivamente obtenemos la sucesión $(y_n)_{n\in\mathbb{N}}$ con las propiedades enunciadas.

Ahora, para cada $n \in \mathbb{N}$, definimos $x_n = \lambda_n \dfrac{y_n}{\left\|y_n\right\|}$. Entonces es claro que $\|x_n\| = |\lambda_n|$ para cada $n \in \mathbb{N}$.

Veamos que la sucesión $(x_n)_{n\in\mathbb{N}}$ es incondicionalmente sumable. Dados una biyección $\sigma : \mathbb{N} \to \mathbb{N}$ y $\varepsilon > 0$, sea $k_0 \in \mathbb{N}$ tal que $1/2^{k-1} < \varepsilon$ para todo $k \geq k_0$. Tomemos $a_0 \in \mathbb{N}$ de manera que $\{1,2,\ldots,n_{k_0+1}-1\} \subseteq \sigma(\{1,2,\ldots,a_0\})$. Entonces, si $b \geq a > a_0$, definimos para cada $n \in \mathbb{N}$

$$\alpha_n = \begin{cases} 1 & \text{si } n \in \sigma(\{a, a+1, \ldots, b\}) \\ 0 & \text{si } n \notin \sigma(\{a, a+1, \ldots, b\}). \end{cases}$$

De esta manera, tenemos que

$$\left\| \sum_{r=a}^{b} x_{\sigma(r)} \right\| = \left\| \sum_{n=n_{k_0+1}}^{+\infty} \alpha_n x_n \right\| \leq \sum_{i=k_0+1}^{+\infty} \left\| \sum_{n=n_i}^{n_{i+1}-1} \alpha_n x_n \right\| = \sum_{i=k_0+1}^{+\infty} \left\| \sum_{n=n_i}^{n_{i+1}-1} \alpha_n \lambda_n \frac{y_n}{\|y_n\|} \right\|$$
$$\leq \sum_{i=k_0+1}^{+\infty} \left( \sum_{n=n_i}^{n_{i+1}-1} \frac{|\alpha_n|^2 |\lambda_n|^2}{\|y_n\|^2} \right)^{1/2} \leq \sqrt{2} \sum_{i=k_0+1}^{+\infty} \left( \sum_{n=n_i}^{n_{i+1}-1} |\lambda_n|^2 \right)^{1/2} < 2 \sum_{i=k_0+1}^{+\infty} \frac{1}{2^i} = \frac{1}{2^{k_0-1}} < \varepsilon.$$

Por lo que, al ser $X$ completo, resulta ser $(x_n)_{n\in\mathbb{N}}$ incondicionalmente sumable. □

En particular podemos extraer, dada su importancia, el siguiente resultado "cualitativo".

**Teorema 1.1.14.** *Sea $X$ un espacio de Banach de dimensión infinita. Entonces existe una sucesión incondicionalmente sumable $(x_n)_{n\in\mathbb{N}} \subseteq X$ , que no es absolutamente sumable.*

*Demostración.* Tomando cualquier sucesión $(\lambda_n)_{n\in\mathbb{N}} \in l_2 \setminus l_1$ (por ejemplo $\lambda_n = 1/n$), mediante el Teorema 1.1.13 obtenemos la existencia de una sucesión incondicionalmente sumable en $X$, que no es absolutamente sumable. □

## 1.2. Convergencia incondicional de series

Notemos que, con respecto al hecho asentado en el teorema de Dvoretzky-Rogers, surgen varias cuestiones:

(a) Nos insinuá que todo espacio de Banach $X$ de dimensión infinita contiene "cierto aire euclídeo". Algo similar podríamos decir acerca de los lemas de Dvoretzky-Rogers, con respecto a los espacios normados de dimensión finita. En 1961 (once años después del teorema anterior) Dvoretzky mostraría que, como veremos en el Capítulo 2, esto efectivamente es así.

(b) Hace que nos preguntemos si existirá algún $r > 2$ de manera que, dada $(\lambda_n)_{n\in\mathbb{N}} \in l_r$, exista en todo espacio de Banach $X$ de dimensión infinita alguna sucesión incondicionalmente sumable $(x_n)_{n\in\mathbb{N}}$ tal que $\|x_n\| = |\lambda_n|$ para todo $n \in \mathbb{N}$.

Para justificar la "insinuación" mencionada en (a) y responder a la pregunta realizada en (b), enunciamos la siguiente proposición.

**Proposición 1.2.1.** *Sean $H$ un espacio de Hilbert y $(x_n)_{n\in\mathbb{N}}$ una sucesión incondicionalmente sumable en $H$. Entonces $\sum_{n\in\mathbb{N}} \|x_n\|^2 < +\infty$.*

**Observación.** En verdad, esta propiedad no es únicamente característica de los espacios (isomorfos a espacios) de Hilbert. Más adelante, comentaremos en qué tipos de espacios podemos encontrarla.

Para poder demostrar la anterior proposición, sin embargo, requeriremos de una noción equivalente a la convergencia incondicional de una serie. Aprovechando la ocasión, enunciamos y demostramos diversas formulaciones que admite la convergencia incondicional de series.

**Teorema 1.2.2** (Teorema de la convergencia incondicional 1). *Sean $X$ un espacio de Banach y $(x_n)_{n\in\mathbb{N}}$ una sucesión en $X$. Son equivalentes:*

*(I)* $(x_n)_{n\in\mathbb{N}}$ *es incondicionalmente sumable.*

*(II)* *Dado* $\varepsilon > 0$, *existe* $n_\varepsilon > 0$ *tal que para todo* $F \subseteq \mathbb{N}$ *finito, con* $min(F) > n_\varepsilon$, *vale que* $\left\|\sum_{n\in F} x_n\right\| < \varepsilon$.

*(III)* *Si consideramos al conjunto* $\Lambda = \{F \subseteq \mathbb{N}/F \text{ es finito}\}$ *como un conjunto dirigido con la relación de orden dada por la inclusión. Entonces la red* $\left(\sum_{n\in F} x_n\right)_{F\in\Lambda}$ *es convergente en $X$.*

*(IV)* $\lim_{n\to+\infty} \sup_{\varphi\in B_{X^*}} \sum_{j\geq n} |\varphi(x_j)| = 0$.

*(V)* *Dada* $(b_n)_{n\in\mathbb{N}} \in l_\infty$, *la serie* $\sum_{n=1}^{+\infty} b_n x_n$ *es convergente.*

*(VI)* *El operador* $T_2 : l_\infty \to X$ *dado por* $T_2(b_n) = \sum_{n\in\mathbb{N}} b_n x_n$ *está bien definido y es acotado. Llamamos constante de incondicionalidad de la sucesión* $(x_n)_{n\in\mathbb{N}}$ *a la norma de* $T_2$, *y la notamos* $C_{inc}$.

*(VII)* *Dada una sucesión* $(\varepsilon_n)_{n\in\mathbb{N}}$ *con* $\varepsilon_n \in \{-1, 1\}$ *para todo* $n \in \mathbb{N}$, *la serie* $\sum_{n=1}^{+\infty} \varepsilon_n x_n$ *es convergente (*$(x_n)_{n\in\mathbb{N}}$ *es* signo sumable*).*

*(VIII)* *Dada una sucesión estrictamente creciente de números naturales* $(n_k)_{k\in\mathbb{N}}$, *la serie* $\sum_{k=1}^{+\infty} x_{n_k}$ *es convergente (*$(x_n)_{n\in\mathbb{N}}$ *es* subserie sumable*).*

*Demostración.* Veremos por un lado que (I), (II), (III) son equivalentes, y otro lado que (II), (IV), (V), (VI), (VII), (VIII) son equivalentes.

$(I) \Rightarrow (II)$ **y** $(VIII) \Rightarrow (II)$. Supongamos que no vale lo afirmado en (II). Entonces, existiría un $\widetilde{\varepsilon} > 0$ ("*el epsilon maldito*") tal que, para cada $n \in \mathbb{N}$ tendríamos un $\widetilde{F}_n \subseteq \mathbb{N}$ finito con $\text{mín}(\widetilde{F}_n) > n$ y $\left\| \sum_{k \in \widetilde{F}_n} x_k \right\| \geq \widetilde{\varepsilon}$.

En particular, tenemos un $F_1 \subseteq \mathbb{N}$ finito con $\text{mín}(F_1) > 1$ y $\left\| \sum_{k \in F_1} x_k \right\| \geq \widetilde{\varepsilon}$.

De igual manera, tenemos un $F_2 \subseteq \mathbb{N}$ finito con $\text{mín}(F_2) > \text{máx}(F_1)$ y $\left\| \sum_{k \in F_2} x_k \right\| \geq \widetilde{\varepsilon}$.

Así, inductivamente, nos formamos una sucesión de conjuntos finitos $F_k \subseteq \mathbb{N}$ tales que $\text{mín}(F_n) > \text{máx}(F_{n-1})$ y $\left\| \sum_{k \in F_n} x_k \right\| \geq \widetilde{\varepsilon}$.

Notemos $a_n = \text{mín}(F_n), b_n = \text{máx}(F_n)$ y $r_n = |F_n|$.[3]

Para $(VIII) \Rightarrow (II)$, tomamos la sucesión estrictamente creciente de naturales $(n_k)_{k \in \mathbb{N}}$ de manera que $\{n_k, k \in \mathbb{N}\} = \bigcup_{n \in \mathbb{N}} F_n$, siendo ésta unión disjunta. Notemos $s_n = \sum_{i=1}^{n} r_i$ para $n > 0$, y $s_0 = 0$, entonces $(s_n)_{n \in \mathbb{N}_0}$ es estrictamente creciente y para cada $n \in \mathbb{N}$ tenemos que $\{n_k \,/\, s_{n-1} + 1 \leq k \leq s_n\} = F_n$, por lo que

$$\left\| \sum_{k=s_{n-1}+1}^{s_n} x_{n_k} \right\| = \left\| \sum_{k \in F_n} x_k \right\| \geq \widetilde{\varepsilon},$$

lo cual es absurdo.

Para $(I) \Rightarrow (II)$, observemos que $(a_n)_{n \in \mathbb{N}}$ es estrictamente creciente. Más precisamente

$$a_n \leq a_n + r_n - 1 \leq b_n < a_{n+1}.$$

Dados dos naturales $a < b$ notaremos $I(a,b) = \{a, \dots, b\}$. Definiremos entonces una biyección $\sigma : \mathbb{N} \to \mathbb{N}$ de manera que para cada $n \in \mathbb{N}$, $\sigma_n = \sigma|_{I(a_n, a_{n+1}-1)} : I(a_n, a_{n+1} - 1) \to I(a_n, a_{n+1} - 1)$ sea una biyección.

Comenzamos tomando a $\sigma$ como la identidad en $I(1, a_1 - 1)$. Luego, para cada $n \in \mathbb{N}$, tomamos $\sigma_n$ de manera que

$$\sigma_n(I(a_n, a_n + r_n - 1)) = F_n \quad \text{y} \quad \sigma_n(I(a_n + r_n, a_{n+1} - 1)) = I(a_n, a_{n+1} - 1) \setminus F_n.$$

Entonces tenemos que la serie $\sum_{n \in \mathbb{N}} x_{\sigma(n)}$ no es de Cauchy, ya que para cada $n \in \mathbb{N}$ vale que

$$\left\| \sum_{j=a_n}^{a_n+r_n-1} x_{\sigma(j)} \right\| = \left\| \sum_{j \in F_n} x_j \right\| \geq \widetilde{\varepsilon}.$$

[3]Dado un conjunto finito $A$, denotamos $|A|$ al cardinal de $A$.

$(II) \Rightarrow (III)$. Dado $\varepsilon > 0$, basta con tomar el $n_\varepsilon \in \mathbb{N}$ del cual nos provee (II), y $F_0 = \{1, \ldots, n_\varepsilon\}$, de manera que, por (II), si $F_1, F_2 \subseteq \mathbb{N}$ son finitos y $F_0 \subseteq F_1 \cap F_2$, entonces

$$\left\| \sum_{n \in F_1} x_n - \sum_{n \in F_2} x_n \right\| = \left\| \left( \sum_{n \in F_1} x_n - \sum_{n=1}^{n_\varepsilon} x_n \right) - \left( \sum_{n \in F_2} x_n - \sum_{n=1}^{n_\varepsilon} x_n \right) \right\|$$
$$\leq \left\| \sum_{n \in F_1 \setminus F_0} x_n \right\| + \left\| \sum_{n \in F_2 \setminus F_0} x_n \right\| \leq 2\varepsilon.$$

Así, la red $(\sum_{n \in F} x_n)_{F \in \Lambda}$ es de Cauchy en $X$, por lo que es convergente.

$(III) \Rightarrow (I)$. Sean $x = \lim_F \sum_{n \in F} x_n$ y $\sigma : \mathbb{N} \to \mathbb{N}$ una biyección. Dado $\varepsilon > 0$, tenemos que existe $F_0 \subseteq \mathbb{N}$ finito tal que, si $F \subseteq \mathbb{N}$ es finito con $F_0 \subseteq F$, entonces $\left\| x - \sum_{n \in F} x_n \right\| < \varepsilon$.

Sea $n_0 \in \mathbb{N}$ tal que $F_0 \subseteq \sigma(\{1, \ldots, n_0\})$. Entonces, es claro que si $n \geq n_0$ tenemos que $\left\| x - \sum_{j=1}^{n} x_{\sigma(j)} \right\| < \varepsilon$. Por lo que $\sum_{j=1}^{+\infty} x_{\sigma(j)} = x$.

Tenemos así que (I), (II), (III) son equivalentes. Veamos ahora la equivalencia entre (II) y los ítems restantes, teniendo en cuenta que ya probamos que $(VIII) \Rightarrow (II)$.

$(II) \Rightarrow (IV)$. Dado $\varepsilon > 0$ sea $n_\varepsilon > 0$ tal que si $F \subseteq \mathbb{N}$ es finito con $\min(F) > n_\varepsilon$, entonces $\left\| \sum_{n \in F} x_n \right\| < \varepsilon/2$. Fijemos $\varphi \in B_{X^*}$ y $N_2 \geq N_1 > n_\varepsilon$. Notemos

$$F^+ = \{n \in \mathbb{N} \text{ tales que } N_1 \leq n \leq N_2 \text{ y } \varphi(x_n) \geq 0\}$$
$$F^- = \{n \in \mathbb{N} \text{ tales que } N_1 \leq n \leq N_2 \text{ y } \varphi(x_n) < 0\}.$$

Entonces

$$\sum_{n=N_1}^{N_2} |\varphi(x_n)| = \sum_{n \in F^+} \varphi(x_n) - \sum_{n \in F^-} \varphi(x_n) = \varphi\left( \sum_{n \in F^+} x_n \right) - \varphi\left( \sum_{n \in F^-} x_n \right)$$
$$\leq \left\| \sum_{n \in F^+} x_n \right\| + \left\| \sum_{n \in F^-} x_n \right\| < \varepsilon.$$

Así, $\sum_{n=N_1}^{N_2} |\varphi(x_n)| < \varepsilon$ para cualesquiera $\varphi \in B_{X^*}$ y $N_2 \geq N_1 > n_\varepsilon$, por lo que $\sum_{j \geq N_1} |\varphi(x_j)| \leq \varepsilon$ para toda $\varphi \in B_{X^*}$ y todo $N_1 > n_\varepsilon$. Luego, tenemos que

$$\lim_{n \to +\infty} \sup_{\varphi \in B_{X^*}} \sum_{j \geq n} |\varphi(x_j)| = 0.$$

$(IV) \Rightarrow (V)$. Sea $(b_n)_{n\in\mathbb{N}} \in l_\infty$. Veamos que $\sum_{n=1}^{+\infty} b_n x_n$ es de Cauchy. Dados $N_1 \leq N_2 \in \mathbb{N}$, aplicando $(IV)$ se tiene que

$$\left\|\sum_{n=N_1}^{N_2} b_n x_n\right\| = \sup_{\varphi\in B_{X^*}} \left|\varphi\left(\sum_{n=N_1}^{N_2} b_n x_n\right)\right| \leq \sup_{\varphi\in B_{X^*}} \sum_{n=N_1}^{N_2} |b_n|.|\varphi(x_n)|$$
$$\leq \|(b_n)\|_\infty \sup_{\varphi\in B_{X^*}} \sum_{n=N_1}^{N_2} |\varphi(x_n)| \leq \|(b_n)\|_\infty \sup_{\varphi\in B_{X^*}} \sum_{n\geq N_1} |\varphi(x_n)| \xrightarrow[N_1\to+\infty]{} 0.$$

Luego, la serie $\sum_{n=1}^{+\infty} b_n x_n$ es de Cauchy, y por lo tanto es convergente.

$(V) \Rightarrow (VI)$. Por $(V)$, tenemos que el operador $T_2 : l_\infty \to X$ dado por $T_2(b) = \sum\limits_{j=1}^{+\infty} b_j x_j$ está bien definido para cada $b \in l_\infty$. Por otro lado, es claro que $T_2$ es lineal, veamos que es acotado. Para esto, basta con ver que $T_2$ es límite puntual de una sucesión de operadores lineales y acotados $u_n : l_\infty \to X$. En efecto, si consideramos $u_n(b) = \sum\limits_{j=1}^{n} b_j x_j$, entonces claramente $u_n$ es lineal y acotado (y de rango finito), y el hecho de que la serie $\sum\limits_{j=1}^{+\infty} b_j x_j$ sea convergente nos dice que $\lim\limits_{n\to\infty} u_n(b) = T_2(b)$ para cada $b \in l_\infty$. Luego $T_2$ es acotado.

$(VI) \Rightarrow (VII)$**.** Esto es claro ya que, dada una sucesión $(\varepsilon_n)_{n\in\mathbb{N}}$ con $\varepsilon_n \in \{-1, 1\}$ para todo $n \in \mathbb{N}$, se tiene que $(\varepsilon_n)_{n\in\mathbb{N}} \in l_\infty$.

$(VII) \Rightarrow (VIII)$**.** Sea $(n_k)_{k\in\mathbb{N}}$ una sucesión estrictamente creciente de números naturales. Sea $M = \{n_k, k \in \mathbb{N}\}$, y definamos

$$\varepsilon_n = \begin{cases} 1 & \text{si } n \in M \\ -1 & \text{si } n \notin M. \end{cases}$$

Entonces, por $(VII)$, $S_N = \sum_{n=1}^{N} x_n + \sum_{n=1}^{N} \varepsilon_n x_n$ es convergente, pero $S_{n_k} = 2\sum_{j=1}^{k} x_{n_j}$. Por lo que $\sum_{j=1}^{+\infty} x_{n_j}$ es convergente. □

Ahora sí, demostremos la Proposición 1.2.1.

*Dem (Proposición 1.2.1):* Sea $(x_n)_{n\in\mathbb{N}} \subseteq H$ una sucesión incondicionalmente sumable. Para ver que la serie $\sum_{n=1}^{+\infty} \|x_n\|^2$ es convergente, basta con ver que la sucesión de sumas parciales es acotada.

Partamos pues de la ley del paralelogramo: Dados $x, y \in H$, tenemos que

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

Como cambiando $x$ por $-x$ las cantidades de ambos lados de la igualdad no se modifican, podemos escribir:

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

o más concisamente:

$$\frac{1}{2^2}\sum_{(\varepsilon_1,\varepsilon_2)\in\{-1,1\}^2}\left\|\varepsilon_1 x+\varepsilon_2 y\right\|^2=\|x\|^2+\left\|y\right\|^2.$$

Veamos por inducción que, en general, el promedio de las cantidades $\|\varepsilon_1 x_1+\cdots+\varepsilon_n x_n\|^2$, con $\varepsilon_i=\pm 1$, es $\|x_1\|^2+\cdots+\|x_n\|^2$. Para $n=1$ esto es claro. Para $n=2$, es lo observado anteriormente. Supongamos que es cierto para $n-1$, donde $n>1$. Entonces, dados $x_1,\dots,x_n\in H$, y usando nuevamente la ley del paralelogramo, tenemos que

$$\begin{aligned}
&\frac{1}{2^n}\sum_{(\varepsilon_1,\dots,\varepsilon_n)\in\{-1,1\}^n}\|\varepsilon_1 x_1+\cdots+\varepsilon_n x_n\|^2\\
&=\frac{1}{2^{n-1}}\sum_{(\varepsilon_1,\dots,\varepsilon_{n-1})\in\{-1,1\}^{n-1}}\frac{1}{2}\left(\|\varepsilon_1 x_1+\cdots+\varepsilon_{n-1}x_{n-1}+x_n\|^2+\|\varepsilon_1 x_1+\cdots+\varepsilon_{n-1}x_{n-1}-x_n\|^2\right)\\
&=\frac{1}{2^{n-1}}\sum_{(\varepsilon_1,\dots,\varepsilon_{n-1})\in\{-1,1\}^{n-1}}\|\varepsilon_1 x_1+\cdots+\varepsilon_{n-1}x_{n-1}\|^2+\|x_n\|^2\\
&=\left(\frac{1}{2^{n-1}}\sum_{(\varepsilon_1,\dots,\varepsilon_{n-1})\in\{-1,1\}^{n-1}}\|\varepsilon_1 x_1+\cdots+\varepsilon_{n-1}x_{n-1}\|^2\right)+\|x_n\|^2=\|x_1\|^2+\cdots+\|x_n\|^2.
\end{aligned}$$

Ahora bien, usando (VI) del Teorema de la convergencia incondicional 1, dado $n\in\mathbb{N}$, tenemos que

$$\begin{aligned}
\sum_{j=1}^{n}\left\|x_j\right\|^2&=\frac{1}{2^n}\sum_{(\varepsilon_1,\dots,\varepsilon_n)\in\{-1,1\}^n}\|\varepsilon_1 x_1+\cdots+\varepsilon_n x_n\|^2\\
\leq&\frac{1}{2^n}\sum_{(\varepsilon_1,\dots,\varepsilon_n)\in\{-1,1\}^n}C_{inc}^2\|(\varepsilon_1,\dots,\varepsilon_n,0,\dots)\|_\infty^2=C_{inc}^2.
\end{aligned}$$

Por lo que la serie es convergente, y más aún, nos queda

$$\sum_{j=1}^{+\infty}\left\|x_j\right\|^2\leq C_{inc}^2. \quad \square$$

**Observación 1.2.3.**

(I) Vimos en la anterior demostración que $H$ es Hilbert si y solo si para cada $n \in \mathbb{N}$ y $x_1, \dots, x_n \in H$ vale que

$$\frac{1}{2^n} \sum_{(\varepsilon_j) \in \{-1,1\}^n} \|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2 .$$

(II) Para realizar la misma cuenta de la demostración en un espacio de Banach $X$, nos bastaría con tener que exista alguna constante $C(X) > 0$ tal que para todo $n \in \mathbb{N}$ y $x_1, \dots, x_n \in X$ valga que

$$\|x_1\|^2 + \cdots + \|x_n\|^2 \leq C(X) \frac{1}{2^n} \sum_{(\varepsilon_j) \in \{-1,1\}^n} \|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\|^2 .$$

Los espacios $X$ que satisfacen esta propiedad se llaman *espacios de cotipo 2* (con constante $C(X)$). Ejemplo de tales espacios son los espacios $L_p(\Omega, \Sigma, \mu)$, donde $1 \leq p \leq 2$ y $(\Omega, \Sigma, \mu)$ es un espacio de medida (ver, por ejemplo, [DJT95, Capítulo 11]).

(III) Dado un espacio de Banach $X$, *en general*, la convergencia incondicional de una serie de términos $(x_n)_{n \in \mathbb{N}}$ no implica nada (más allá de la necesaria condición de que $\|x_n\| \to 0$) sobre el decaimiento de las normas $\|x_n\|$. Por ejemplo, si $X = c_0$, dada $a = (a_n)_{n \in \mathbb{N}} \in c_0$ cualquiera, tenemos que la serie de término general $x_n = a_n e_n$ converge incondicionalmente (¡hacia la misma $a$!), mientras que $\|x_n\| = |a_n|$ para todo $n \in \mathbb{N}$.

Para finalizar, asentemos otras formas de establecer la convergencia incondicional de series.

**Teorema 1.2.4** (Teorema de la convergencia incondicional 2)**.** *Sean $X$ un espacio de Banach y $(x_n)_{n \in \mathbb{N}}$ una sucesión en $X$. Son equivalentes:*

*(I) $(x_n)_{n \in \mathbb{N}}$ es incondicionalmente sumable.*

*(II) $(b_n x_n)_{n \in \mathbb{N}}$ es débilmente sumable para cada $(b_n)_{n \in \mathbb{N}} \in l_\infty$. Es decir, para cada $(b_n)_{n \in \mathbb{N}} \in l_\infty$, la serie $\displaystyle\sum_{k=1}^{+\infty} b_n x_n$ es convergente para la topología débil.*

*(III) $(x_n)_{n \in \mathbb{N}}$ es débilmente signo sumable. Es decir, dada una sucesión $(\varepsilon_n)_{n \in \mathbb{N}}$ con $\varepsilon_n \in \{-1, 1\}$ para todo $n \in \mathbb{N}$, la serie $\displaystyle\sum_{n=1}^{+\infty} \varepsilon_n x_n$ es convergente para la topología débil.*

*(IV) $(x_n)_{n \in \mathbb{N}}$ es débilmente subserie sumable. Es decir, dada sucesión estrictamente creciente de naturales $(n_k)_{k \in \mathbb{N}}$, la serie $\displaystyle\sum_{k=1}^{+\infty} x_{n_k}$ es convergente para la topología débil.*

*(V) El operador $T_1 : X^* \to l_1$ dado por $T_1(\varphi) = (\varphi(x_n))_{n \in \mathbb{N}}$ está bien definido y es compacto.*

*(VI)* *El operador* $T_2 : l_\infty \to X$ *dado por* $T_2(b_n) = \sum_{n\in\mathbb{N}} b_n x_n$ *está bien definido y es compacto.*

*(VII)* *El operador* $T_3 : c_0 \to X$ *dado por* $T_3(a_n) = \sum_{n\in\mathbb{N}} a_n x_n$ *está bien definido y es compacto.*

*Además, en caso de valer alguna (y por lo tanto todas) de las afirmaciones anteriores, se cumple que* $\|T_1\|_{\mathcal{L}(X^*,l_1)} = \|T_2\|_{\mathcal{L}(l_\infty,X)} = \|T_3\|_{\mathcal{L}(c_0,X)}$.

*Demostración.* Veremos primero que $(I)$, $(II)$, $(III)$, $(IV)$, $(V)$ son equivalentes. Luego veremos la equivalencia de $(VI)$ y $(VII)$ con los ítems anteriores. Notemos que, como $(I)$, podemos tomar como equivalente cualquiera de los ítems vistos en el Teorema 1.2.2.

$(I) \Rightarrow (II)$**.** Como para cada $(b_n)_{n\in\mathbb{N}} \in l_\infty$ la serie $\sum_{j=1}^{+\infty} b_j x_j$ converge en norma, entonces es convergente para la topología débil.

$(II) \Rightarrow (III)$**.** Basta con notar que $(\varepsilon_n)_{n\in\mathbb{N}} \in l_\infty$ si $\varepsilon_n \in \{-1, 1\}$ para todo $n \in \mathbb{N}$.

$(III) \Rightarrow (IV)$**.** Es similar a la demostración de $(VII) \Rightarrow (VIII)$ en el Teorema 1.2.2, solo que en este caso la convergencia es con respecto a la topología débil en X.

$(IV) \Rightarrow (V)$**.** Sea $S = \overline{\langle x_n \rangle}_{n\in\mathbb{N}}$. Veremos que $T : S^* \to l_1$ dado por $T(\eta) = (\eta(x_n))_{n\in\mathbb{N}}$ está bien definido y es compacto. Luego, notando $r : X^* \to S^*$ a la restricción $r(\varphi) = \varphi|_S$, tendremos que $T_1 = T \circ r : X^* \to l_1$ dado por $T_1(\varphi) = (\varphi|_S(x_n))_{n\in\mathbb{N}} = (\varphi(x_n))_{n\in\mathbb{N}}$ está bien definido y es compacto.

Dada $\eta \in S^*$, teniendo en cuenta que $\eta$ es la restricción de algún funcional definido en $X^*$, tenemos que para cada sucesión de naturales estrictamente creciente $(n_k)_{k\in\mathbb{N}}$ la serie $\sum_{k=1}^{+\infty} \eta(x_{n_k})$ es convergente, por lo que la serie $\sum_{n=1}^{+\infty} \eta(x_n)$ es incondicionalmente convergente, y por lo tanto absolutamente convergente, ya que es una serie de escalares. Así, con tranquilidad podemos definir al operador $T : S^* \to l_1$ dado por $T(\eta) = (\eta(x_n))_{n\in\mathbb{N}}$. Como este operador es límite puntual de los operadores lineales y acotados $u_k : S^* \to l_1$ dados por $u_k(\eta) = ((\eta(x_n))_{1\le n\le k}$ (esto no es más que la convergencia de la serie $\sum_{n\in\mathbb{N}} |\eta(x_n)|$), entonces $T$ es acotado. Veamos que es compacto.

Para eso, tomemos una sucesión $(\eta_n)_{n\in\mathbb{N}}$ en $B_{S^*}$, y veamos que $(T(\eta_n))_{n\in\mathbb{N}}$ tiene alguna subsucesión convergente en $l_1$. Ahora bien, como $S$ es separable, tenemos que $(B_{S^*}, \omega^*)$ es un compacto metrizable, por lo que existe una subsucesión $(\eta_{n_k})_{k\in\mathbb{N}}$ $\omega^*$–convergente a un $\eta \in B_{S^*}$. Veremos que $T(\eta_{n_k}) \to T(\eta)$ en $l_1$. Para esto, como $l_1$ tiene la propiedad de Schur, basta con ver que $T(\eta_{n_k}) \rightharpoonup T(\eta)$ en $l_1$.

Dados $b \in l_\infty$ y $a \in l_1$, notemos $b \cdot a$ a la acción de $b$ como un funcional en $l_1$. Queremos entonces ver que $b \cdot T(\eta_{n_k}) \to b \cdot T(\eta)$ para cada $b \in l_\infty$. Como $\left(T(\eta_{n_k})\right)_{k\in\mathbb{N}}$ es acotada (pues $T$ es acotado y $\eta_{n_k} \in B_{S^*}\ \forall\, k \in \mathbb{N}$) alcanza con ver que $b \cdot T(\eta_{n_k}) \to b \cdot T(\eta)$ para todo $b \in D$ donde $D$ es algún subespacio denso (en la topología fuerte) de $l_\infty$.

En efecto, en tal caso, dados $\varepsilon > 0$ y $b \in l_\infty$ tendríamos un $\widetilde{b} \in D$ con $\|\widetilde{b} - b\|_\infty < \varepsilon$, por lo que, para todo $k$ mayor o igual a un cierto $k_0 \in \mathbb{N}$ nos quedaría

$$|b \cdot T(\eta_{n_k}) - b \cdot T(\eta)| \leq |(b - \widetilde{b})(T(\eta_{n_k}) - T(\eta))| + |\widetilde{b}\Big(T(\eta_{n_k}) - T(\eta)\Big)| \leq 2\,\|T\|\,\varepsilon + \varepsilon.$$

Afirmamos que $D = \langle b_F \rangle_{F \subseteq \mathbb{N}}$, donde para cada $F \subseteq \mathbb{N}$

$$b_F(k) = \begin{cases} 1 & \text{si } k \in F \\ 0 & \text{si } k \notin F \end{cases}$$

es denso en $l_\infty$. Veremos esto tras finalizar el argumento de la demostración.

Ahora, dado $F \subseteq \mathbb{N}$ escribamos $F = \{m_k, \text{ con } k \in \mathbb{N} \text{ y } 1 \leq k \leq |F|\}$ donde $(m_k)$ es estrictamente creciente, y notemos

$$z_F = \omega\text{-}\lim_{l \to \infty} \sum_{j=1}^{l} x_{m_j}.$$

Observemos que si $F$ es finito, nos queda simplemente $z_F = \sum_{m \in F} x_m$.

Entonces tenemos que

$$\begin{aligned} b_F \cdot T(\eta_{n_k}) &= \sum_{m \in F} \eta_{n_k}(x_m) = \lim_{l \to \infty} \sum_{j=1}^{l} \eta_{n_k}(x_{m_j}) = \lim_{l \to \infty} \eta_{n_k}\left(\sum_{j=1}^{l} x_{m_j}\right) \\ &= \eta_{n_k}(z_F) \underset{k \to \infty}{\longrightarrow} \eta(z_F) = \lim_{l \to \infty} \eta\left(\sum_{j=1}^{l} x_{m_j}\right) = \sum_{m \in F} \eta(x_m) = b_F \cdot T(\eta). \end{aligned}$$

Así, tenemos que $T$ es compacto, lo que implica lo enunciado en $(V)$ por lo comentado al inicio de la demostración.

Veamos ahora que el subespacio $D = \langle b_F \rangle_{F \subseteq \mathbb{N}}$ definido anteriormente es denso en $l_\infty$. En efecto, dados $a \in l_\infty$ y $\varepsilon > 0$, existe $L \in \mathbb{N}$ tal que $a_k \in [-L\varepsilon, L\varepsilon)\ \forall\, k \in \mathbb{N}$. Si para cada $l \in \mathbb{Z}$ definimos $M_l = \{k \in \mathbb{N} /\ l\varepsilon \leq a_k < (l+1)\varepsilon\}$ y $b = \sum_{-L \leq l < L} l\,\varepsilon\, b_{M_l} \in D$, entonces es claro que $\|a - b\|_\infty < \varepsilon$.

$(V) \Rightarrow (I)$. Como $T_1 : X^* \to l_1$ es compacto, tenemos que $\overline{T_1(B_{X^*})} \subseteq l_1$ es compacto. Ahora, dado $\varepsilon > 0$ consideramos en $l_1$ los abiertos

$$V_n(\varepsilon) = \left\{ a \in l_1 /\ \sum_{j \geq n} |a_j| < \varepsilon \right\}.$$

Notemos que $V_n(\varepsilon) \subseteq V_{n+1}(\varepsilon)$ y $l_1 = \bigcup_{n \in \mathbb{N}} V_n(\varepsilon)$, por lo que debe ser $T_1(B_{X^*}) \subseteq V_{n_0}(\varepsilon)$ para algún $n_0 \in \mathbb{N}$.

Es decir que $\sum_{j\geq n} |\varphi(x_j)| < \varepsilon$ para toda $\varphi \in B_{X^*}$ y todo $n \geq n_0$. Luego

$$\lim_{n\to+\infty} \sup_{\varphi\in B_{X^*}} \sum_{j\geq n} |\varphi(x_j)| = 0.$$

Por lo que $(x_n)_{n\in\mathbb{N}}$ es incondicionalmente sumable.

$(I\text{-}V) \Rightarrow (VI)$. Teniendo que $T_1 : X^* \to l_1$, $T_1(\varphi) = (\varphi(x_n))_{n\in\mathbb{N}}$, es compacto, sabemos que $T_1^* : l_\infty \to X^{**}$ es compacto. Ahora bien, dadas $b \in l_\infty$ y $\varphi \in X^*$, como sabemos que $(b_n x_n)_{n\in\mathbb{N}}$ es sumable (pues asumimos que $(x_n)_{n\in\mathbb{N}}$ es incondicionalmente sumable) podemos escribir

$$T_1^*(b)(\varphi) = b \cdot T_1(\varphi) = \sum_{n\in\mathbb{N}} b_n \varphi(x_n) = \varphi\left(\sum_{n\in\mathbb{N}} b_n x_n\right) = J(x)(\varphi)^{4},$$

donde $x = \sum_{n\in\mathbb{N}} b_n x_n \in X$. Luego podemos considerar $T_2 = T_1^*|^X : l_\infty \to X$, la correstricción de $T_1^*$ a $X$, dada por $T_2(b) = \sum_{n\in\mathbb{N}} b_n x_n$ y tenemos que $T_2$ es un operador lineal compacto.

$(VI) \Rightarrow (VII)$. Simplemente notamos que $T_3 = T_2|_{c_0}$.

$(VII) \Rightarrow (I\text{-}V)$. Chequeemos que tomando adjunto nos queda $T_3^* = T_1 : X^* \to l_1$.
En efecto, dadas $\varphi \in X^*$ y $a \in c_0$

$$T_3^*(\varphi)(a) = \varphi(T_3(a)) = \varphi\left(\sum_{n\in\mathbb{N}} a_n x_n\right) = \sum_{n\in\mathbb{N}} a_n \varphi(x_n) = (\varphi(x_n)) \cdot a.$$

Luego $T_1 : X^* \to l_1$ dado por $T_1(\varphi) = (\varphi(x_n))_{n\in\mathbb{N}}$ está bien definido y es compacto.

Por último, observemos que en caso de cumplirse las afirmaciones anteriores, vale que $\|T_1\| = \|T_2\| = \|T_3\|$. Esto es claro ya que $T_1 = T_3^*$ y $T_2 = T_1^*|^X$. □

Notemos que, como consecuencia del teorema anterior, tenemos la siguiente proposición:

**Proposición 1.2.5.** *Sean $X$ un espacio de Banach y $T : c_0 \to X$ un operador lineal compacto. Entonces $T$ se extiende a un operador compacto $\widetilde{T}$ en $l_\infty$, con $\|\widetilde{T}\| = \|T\|$.*

*Demostración.* En efecto, si $T : c_0 \to X$ es un operador compacto, entonces, definiendo $x_n = T(e_n)$ tenemos, por lo visto en el teorema anterior, que $(x_n)_{n\in\mathbb{N}}$ es una sucesión incondicionalmente sumable en $X$, por lo que el operador $\widetilde{T} : l_\infty \to X$ dado por $\widetilde{T}(a_n) = \sum_{n\in\mathbb{N}} a_n x_n$ resulta estar bien definido y ser compacto. Además, $\widetilde{T}|_{c_0} = T$ y $\|\widetilde{T}\| = \|T\|$. □

Este hecho puede verse a la luz del Teorema 6.1 en [Lin64] (observando en particular la equivalencia entre los ítems (2) y (8) allí expresados).

[4] Aquí $J : X \to X^{**}$ es la inclusión canónica de $X$ en su bidual.

**Observación 1.2.6.**

(I) Vimos que si $(x_n)_{n\in\mathbb{N}} \subseteq X$ es incondicionalmente sumable, entonces el operador $T : c_0 \to X$ dado por $T(e_n) = x_n$ es compacto. Surge entonces la siguiente pregunta: En el caso en el que la sucesión $(x_n)_{n\in\mathbb{N}}$ es *absolutamente* sumable, ¿que más podemos decir acerca del operador $T$?

(II) Mirando nuevamente el ítem $(VII)$ del Teorema 1.2.4, notamos que al estar cada sucesión incondicionalmente sumable $(x_n)_{n\in\mathbb{N}}$ en $X$ asociada a un único operador compacto $T_{(x_n)} \in \mathcal{L}(c_0, X)$, y viceversa, podemos dotar al espacio de sucesiones incondicionalmente sumables en $X$ de una norma, de manera que resulte un espacio de Banach, digamos $(\mathcal{S}(X), \|\ \|_{\text{inc}})$, donde $\mathcal{S}(X) = \{(x_n)_{n\in\mathbb{N}} \subseteq X \,/\, (x_n)_{n\in\mathbb{N}}$ es incondicionalmente sumable$\}$ y $\|(x_n)_{n\in\mathbb{N}}\|_{\text{inc}} = \left\|T_{(x_n)}\right\| = C_{inc}$. Este espacio resulta claramente isométrico a $\mathcal{K}(c_0, X)$, el espacio de los operadores compactos de $c_0$ en $X$.

## Comentario

Vimos como, si $X$ es un espacio de dimensión infinita, entonces la convergencia absoluta y la convergencia incondicional de series no son equivalentes en $X$. Por otro lado, si estamos en $\mathbb{R}$, dada una serie de términos $(a_n)_{n\in\mathbb{N}}$ absolutamente (ó incondicionalmente) convergente, podemos pensar

$$\sum_{n=1}^{+\infty} a_n = \int_{\mathbb{N}} a_n \, d\mu_c(n),$$

donde $\mu_c$ es la medida de contar en $(\mathbb{N}, \mathcal{P}(N))$. Notemos que la expresión de la izquierda está entendida como un límite de sumas parciales, mientras que la expresión de la derecha no, sino que más bien está dada como... bueno, como una integral de Lebesgue sobre $\mathbb{N}$. Es decir, podemos considerar la expresión de la derecha sin dar ningún orden en particular en $\mathbb{N}$. Notemos que esa es la sutil diferencia entre escribir

$$\sum_{n=1}^{+\infty} a_n \quad \text{y} \quad \sum_{n\in\mathbb{N}} a_n.$$

Ahora bien, en $\mathbb{R}$, para que una función medible $f$ sea Lebesgue integrable, requerimos (*necesitamos*) que se cumpla $\int_{\mathbb{R}} |f(x)| \, dx < +\infty$ de manera que $\int_{\mathbb{R}} f(x) \, dx$ sea un número real de forma totalmente irreprochable.

Por otro lado, en un espacio de Banach de dimensión infinita $X$ tenemos la integral de Bochner (ver [DU77]), que nos pide que $f : (\Omega, \Sigma, \mu) \to X$ sea fuertemente medible y "absolutamente integrable", es decir que se cumpla $\int_{\Omega} \|f(\omega)\|_X \, d\mu(\omega) < +\infty$ para poder tener $\int_{\Omega} f(\omega) \, d\mu(\omega)$ como un elemento de $X$. Ahora, poniendo por ejemplo $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_c)$ como espacio de medida, dada una sucesión $(x_n)_{n\in\mathbb{N}}$ en $X$, queda claro que este requerimiento es excesivo para la impecable existencia del elemento $\int_{\mathbb{N}} x_n \, d\mu_c(n)$ en $X$ (en forma fuerte). Esto es, para que la serie converja incondicionalmente (que es lo que entendemos que nos da la indudable existencia del elemento $\sum_{n\in\mathbb{N}} x_n$ en $X$), es excesivo requerir que la serie de término general $(x_n)_{n\in\mathbb{N}}$ converja absolutamente. Vale recordar aquí la Observación 1.2.3 (III).

Así, podemos preguntarnos si habrá alguna noción de "*integral incondicional*", distinta a la de integral de Bochner, de manera que para una $f : (\Omega, \Sigma, \mu) \to X$ integrable en este sentido, se tenga la existencia (o convergencia) en norma, o en un sentido fuerte, de la integral $\int_\Omega f(\omega)\, d\mu(\omega) \in X$, más allá de que la función $f$ sea o no Bochner integrable.

Pues bien, aunque no es tan conocida como la integral de Bochner, existe una noción (entre varias otras en verdad...) de integral en espacios de Banach que refleja los aspectos mencionados anteriormente, la llamada *integral de Birkhoff*, introducida por Birkhoff en [Bir35]. Para un estudio completo de la misma, y de su relación con otras nociones de integrales vectoriales, puede verse [Rui06].

# Capítulo 2

# El teorema de Dvoretzky

En la Sección 2.4 de este capítulo daremos una demostración del teorema de Dvoretzky siguiendo a [Pis89].

La misma hará uso del teorema de Factorización de Dvoretsky-Rogers, que veremos en la Sección 2.1. También tendrá como uno de sus ingredientes principales a la medida gaussiana en $\mathbb{R}^n$, que introduciremos en la Sección 2.2. En dicha sección veremos el Teorema 2.2.11, de particular importancia para la demostración del teorema de Dvoretzky. Este teorema consiste en una desigualdad que acota la medida gaussiana del conjunto de puntos en los que una función Lipschitz en $\mathbb{R}^n$ se desvía de su media.

Por último, necesitaremos ciertos lemas geométricos (Lema 2.3.1 y Lema 2.3.2), que veremos en la Sección 2.3. Estos lemas nos dicen a grandes rasgos que, dado $\varepsilon > 0$, si dos normas $\| \ \|_1$ y $\| \ \|_2$ en $\mathbb{R}^n$ "están lo suficientemente cerca" sobre un conjunto finito de puntos (cuyo cardinal estará controlado), entonces los espacios $(\mathbb{R}^n, \| \ \|_1)$ y $(\mathbb{R}^n, \| \ \|_2)$ serán $(1+\varepsilon)$-isomorfos.

Luego, mostraremos en la Sección 2.5 dos hechos comentados en la Introducción: El primero es que $N(k,\varepsilon)$, definido en la Introducción, página VIII, tiene un comportamiento asintótico exponencial en la dimensión $k$ (Proposición 2.5.1). El segundo es que si $k(X,\varepsilon)$ se define como en la Introducción, página IX, entonces $k(l_\infty^n, \varepsilon) \geq c_1 \dfrac{\ln(n)}{|\ln(c_2\,\varepsilon)|}$ (Proposición 2.5.2).

## 2.1. Factorización de Dvoretzky-Rogers

Como mencionamos en la Introducción, Grothendieck comenta en [Gro56] que a partir del lema de Dvoretzky-Rogers puede obtenerse el teorema de Factorización de Dvoretsky-Rogers. Este teorema suele llamarse *factorización de Dvoretzky-Rogers*, ya que podríamos enunciarlo de la siguiente manera

*Dados $\varepsilon > 0$ y $k \in \mathbb{N}$, existe $n(k,\varepsilon) \in \mathbb{N}$ tal que si $E$ es un espacio normado de dimensión mayor o igual a $n(k,\varepsilon)$, entonces existen un subespacio $k$-dimensional $E_k \subseteq E$ y operadores $u : l_2^k \to E_k$, $v : E_k \to l_\infty^k$ de manera que $v \circ u$ es la identidad (formal) $I : l_2^k \to l_\infty^k$, con $\|u\| \leq 1$,*

$\|v\| \leq 1+\varepsilon$.

Por lo que estamos factorizando la identidad $I : l_2^k \to l_\infty^k$ pasando a través del subespacio $E_k \subseteq E$, con apenas una pequeña distorsión en la norma de $I$.

Veamos entonces como, a partir del lema de Dvoretzky-Rogers, en efecto puede obtenerse el teorema de Factorización de Dvoretsky-Rogers. Más precisamente, obtendremos el siguiente resultado.

**Teorema 2.1.1.** *Sean $\varepsilon > 0$ y $n \in \mathbb{N}$. Entonces, existe $\lambda(\varepsilon) > 0$ tal que si $E$ es un espacio normado de dimensión $n$ y $k \leq 1 + \lambda(\varepsilon)\sqrt{n}$, se pueden encontrar vectores $x_1, \dots, x_k \in E$ de manera que, para cualesquiera $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ se cumple*

$$\frac{1}{1+\varepsilon}\|(\alpha_i)\|_\infty \leq \left\|\sum_{j=1}^k \alpha_i x_i\right\|_E \leq \|(\alpha_i)\|_2.$$

*Más aún, alcanza con tomar $\lambda(\varepsilon) = \dfrac{\varepsilon}{2(1+\varepsilon)}$.*

*Demostración.* Tomando un isomorfismo lineal $T : E \to \mathbb{R}^n$ de manera que $T(B_E) = K$ esté en posición de John, podemos trabajar con el espacio $(\mathbb{R}^n, \|\ \|_K)$ (ya que es isométrico a $E$).

Sean $z_1, \dots, z_n$ los vectores ortonormales que nos son provistos por el lema de Dvoretzky-Rogers. Tomando los primeros $k$ elementos de esta base, obtendremos los vectores con las propiedades deseadas.

Desde ya, tenemos que si $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ son escalares cualesquiera, entonces

$$\left\|\sum_{j=1}^n \alpha_j z_j\right\|_K \leq \left\|\sum_{j=1}^n \alpha_j z_j\right\|_2 = \left(\sum_{j=1}^n \alpha_j^2\right)^{1/2}.$$

Lo que nos requerirá cierto trabajo es ver que, restringiéndonos a combinaciones lineales de $z_1, \dots, z_k$, vale la cota inferior del enunciado.

Comenzaremos mostrando por inducción que $\|z_j\|_* \leq 1+\varepsilon$ para todo $1 \leq j \leq k$ [1].

Para $j = 1$, según lo visto en el lema de Dvoretzky-Rogers, tenemos que $z_1 = A_1$. Por lo que, según lo mencionado en la observación 1.1.8 (II), tenemos que $\|z_1\|_* = 1$.

Supongamos entonces que $\|z_j\|_* \leq 1+\varepsilon$ para todo $1 \leq j \leq r$, con $r \leq k-1 \leq \lambda(\varepsilon)\sqrt{n}$.

Entonces, siguiendo las notaciones de los elementos en el lema de Dvoretzky-Rogers

$$\begin{aligned}\left\|\sum_{j=1}^r a_{r+1\,j}\, z_j\right\|_* &\leq \sum_{j=1}^r |a_{r+1\,j}|\, \|z_j\|_* \leq (1+\varepsilon)\sum_{j=1}^r |a_{r+1\,j}| \leq (1+\varepsilon)\sqrt{r}\left(\sum_{j=1}^r |a_{r+1\,j}|^2\right)^{1/2} \\ &= (1+\varepsilon)\sqrt{r}\left(1 - a_{r+1\,r+1}^2\right)^{1/2} \leq (1+\varepsilon)\sqrt{r}\sqrt{\frac{r}{n}} = \frac{(1+\varepsilon)r}{\sqrt{n}}.\end{aligned}$$

[1] Recordemos que cada elemento $z_j$, con $1 \leq j \leq n$, induce un funcional en $\mathbb{R}^n$ dado por $\varphi_j(x) = \langle x, z_j\rangle$ y notamos $\|z_j\|_* = \|\varphi_j\|_{(\mathbb{R}^n, \|\ \|_K)^*}$ a la norma de este funcional en $(\mathbb{R}^n, \|\ \|_K)^*$.

Usando que $\|A_i\|_* = 1$ para todo $1 \le i \le n$ y la desigualdad triangular, nos queda que

$$\|z_{r+1}\|_* = \frac{\left\|A_{r+1} - \sum_{j=1}^r a_{r+1\,j}\, z_j\right\|_*}{a_{r+1\,r+1}} \le \sqrt{\frac{n}{n-r}} \left(1 + \frac{(1+\varepsilon)\,r}{\sqrt{n}}\right)$$

$$\le \sqrt{\frac{n}{n-r}}\,(1+(1+\varepsilon)\,\lambda(\varepsilon)) = \frac{1}{\sqrt{1-\frac{r}{n}}}\left(1+(1+\varepsilon)\frac{\varepsilon}{2(1+\varepsilon)}\right) \le \frac{1}{\sqrt{1-\frac{\lambda(\varepsilon)}{\sqrt{n}}}}\left(1+\frac{\varepsilon}{2}\right)$$

$$\le \frac{1}{1-\frac{\lambda(\varepsilon)}{\sqrt{n}}}\left(1+\frac{\varepsilon}{2}\right) \le \frac{1}{1-\lambda(\varepsilon)}\left(1+\frac{\varepsilon}{2}\right) = \frac{2(1+\varepsilon)}{2+2\varepsilon-\varepsilon}\left(\frac{2+\varepsilon}{2}\right) = 1+\varepsilon.$$

De esta manera, llegamos a que si $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, para cada $1 \le i \le k$ vale que

$$\left\|\sum_{j=1}^r \alpha_j\, z_j\right\|_K \ge \frac{1}{1+\varepsilon}\left|\langle \sum_{j=1}^r \alpha_j\, z_j, z_i \rangle\right| = \frac{1}{1+\varepsilon}|\alpha_i|.$$

Por lo que $\max\limits_{1\le i\le k} |\alpha_i| \le (1+\varepsilon)\left\|\sum_{j=1}^r \alpha_j\, z_j\right\|_K$. □

**Observación 2.1.2.**

(I) Este resultado nos dice que, dados $\varepsilon > 0$ y $k \in \mathbb{N}$, el $n(k,\varepsilon)$ del teorema de Factorización de Dvoretsky-Rogers puede tomarse como

$$n(k,\varepsilon) = \left[\frac{4(1+\varepsilon)^2}{\varepsilon^2}(k-1)^2\right] + 1.$$

En particular, la dimensión necesaria para el espacio $E$ de manera de poder encontrar $k$ vectores que satisfagan lo enunciado en el teorema de Factorización de Dvoretsky-Rogers es (a lo sumo) cuadrática en $k$.

(II) Notablemente, en 1988 Bourgain y Szarek demostraron en [BS88] que es posible tomar la dimensión $k$ proporcional a $n = \dim(E)$ para lograr una factorización de la identidad formal $I : l_2^k \to l_\infty^k$ pasando a través de un subespacio de dimensión $k$ en $E$. Eso si, esta factorización no produce *distorsiones tan pequeñas como sean deseadas* en la norma de $I : l_2^k \to l_\infty^k$.

Más precisamente, demostraron que dado $0 < \delta < 1$ y $n \in \mathbb{N}$ existe $M(\delta) \ge 1$ tal que si $k \le (1-\delta)n$ entonces existen $x_1, \ldots, x_k$ que satisfacen

$$\frac{1}{M(\delta)}\|(\alpha_i)\|_\infty \le \left\|\sum_{j=1}^k \alpha_i x_i\right\|_E \le \|(\alpha_i)\|_2.$$

La dependencia de $M(\delta)$ en $\delta$ fue estudiada posteriormente, y aún siendo desconocido su comportamiento asintótico preciso, se ha demostrado que puede tomarse $M(\delta) \sim 1/\delta$. Por otro lado, se sabe que debe satisfacerse $M(\delta) \ge c\,\delta^{-1/10}$.

(III) Mencionemos por último una lectura geométrica del resultado obtenido: Si $K \subseteq \mathbb{R}^n$ es un cuerpo convexo simétrico en posición de John y $k \leq \lambda(\varepsilon)\sqrt{n}+1$, entonces existen una sección $k$-dimensional de $K$ (dada por $K \cap E_k$, con $E_k \subseteq \mathbb{R}^n$ subespacio de dimensión $k$), y un cubo (de lado 2) $Q_k$ $k$-dimensional en $E_k$, de manera que

$$B_2^n \cap E_k \subseteq K \subseteq (1+\varepsilon)Q_k.$$

## 2.2. Preliminares gaussianos

Denotaremos con $\mathcal{M}_n$ a los conjuntos medibles Lebesgue de $\mathbb{R}^n$. Dado $A \in \mathcal{M}_n$ notaremos con $m(A)$ o $m_n(A)$ a la medida de Lebesgue de $A$. Trabajaremos con la medida gaussiana estándar en $(\mathbb{R}^n, \mathcal{M}_n)$, que hace del mismo un espacio de probabilidad. Esta medida es absolutamente continua con respecto a la medida de Lebesgue, y está definida por

$$\gamma_n(A) = \frac{1}{(2\pi)^{n/2}} \int_A e^{\frac{-\|x\|_2^2}{2}}\, dm(x) = \frac{1}{(2\pi)^{n/2}} \int_A e^{\frac{-\|x\|_2^2}{2}}\, dx$$

para $A \in \mathcal{M}_n$.

Notemos que si $n = 1$, en particular nos queda $\gamma_1([a,b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx$, y que

$$\gamma_n\Big([a_1,b_1] \times \cdots \times [a_n,b_n]\Big) = \prod_{j=1}^{n} \gamma_1([a_j,b_j]).$$

Es decir $\gamma_n = \gamma_1 \times \cdots \times \gamma_1 = \gamma_k \times \gamma_{n-k}$ para cada $1 \leq k \leq n$.

Por su definición, es claro que $\gamma_n$ es invariante por isometrías lineales en $\mathbb{R}^n$. Más concretamente, si $u : \mathbb{R}^n \to \mathbb{R}^n$ es ortogonal, y $A \subseteq \mathbb{R}^n$ es medible, entonces, por el teorema de cambio de variables

$$\gamma_n(u(A)) = \int_{u(A)} e^{\frac{-\|x\|_2^2}{2}}\, dx = |\det(u)| \int_A e^{\frac{-\|u(y)\|_2^2}{2}}\, dy = \int_A e^{\frac{-\|y\|_2^2}{2}}\, dy = \gamma_n(A).$$

Luego, dadas $g : \mathbb{R}^n \to [0,+\infty]$ medible y absolutamente integrable respecto de $\gamma_n$ y $u : \mathbb{R}^n \to \mathbb{R}^n$ una transformación ortogonal, vale que

$$\int_{\mathbb{R}^n} g \circ u \; d\gamma_n = \int_{\mathbb{R}^n} g \; d\gamma_n.$$

Recordemos que si $(\Omega, \Sigma, \mu)$ es un espacio de probabilidad, y $g : \Omega \to \mathbb{R}$ es una variable aleatoria, decimos que $g$ es gaussiana estándar (o que tiene distribución gaussiana estándar) si dado $A \subseteq \mathbb{R}$ abierto vale que $\mu(g \in A) = \gamma_1(A)$.

En general, si $G : \Omega \to \mathbb{R}^n$ es un vector de variables aleatorias, decimos que $G$ es gaussiano estándar (o tiene distribución gaussiana estándar) si $\mu(G \in A) = \gamma_n(A)$ para todo $A \subseteq \mathbb{R}^n$ boreliano. Si $G = (g_1, \ldots, g_n)$, esto es equivalente a que las variables aleatorias $g_1, \ldots, g_n$ sean independientes y tengan distribución gaussiana estándar en $\mathbb{R}$.

En particular observemos que las funciones coordenadas $p_j : (\mathbb{R}^n, \gamma_n) \to \mathbb{R}$, $p_j(x) = x_j$ con $1 \leq j \leq n$, satisfacen trivialmente lo anterior, por lo que conforman un vector gaussiano estándar. Más aún, para nunca caer en falta de variables aleatorias gaussianas estándar independientes, podemos tomar el producto de espacios de probabilidad de numerables copias de $(\mathbb{R}, \mathcal{B}, \gamma_1)$ (donde $\mathcal{B}$ son los borelianos en $\mathbb{R}$) para obtener $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}_\infty, \gamma_\infty)$, donde $\mathcal{B}_\infty$ son los borelianos de $\mathbb{R}^{\mathbb{N}}$ con la topología producto. De esta manera, las proyecciones (o funciones coordenadas) $p_j : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, $p_j(x) = x_j$ con $j \in \mathbb{N}$, constituyen una sucesión de variables aleatorias gaussianas estándar independientes.

Como consecuencia de la invariancia por rotaciones de la medida gaussiana en $\mathbb{R}^n$, tenemos el siguiente lema.

**Lema 2.2.1.** *Sean $(\Omega, \Sigma, \mu)$ un espacio de probabilidad, $G : \Omega \to \mathbb{R}^n$ un vector aleatorio gaussiano estándar y $U \in \mathbb{R}^{n \times n}$ una matriz ortogonal. Entonces $U.G : \Omega \to \mathbb{R}^n$ es un vector aleatorio con distribución gaussiana estándar.*

*Es decir, si $g_1, \ldots, g_n : \Omega \to \mathbb{R}$ son variables aleatorias gaussianas estándar independientes, entonces, dada $(a_{ij}) \in \mathbb{R}^{n \times n}$ ortogonal, las variables aleatorias $\widetilde{g_i} : \Omega \to \mathbb{R}$ definidas por*

$$\widetilde{g_i} = \sum_{j=1}^{n} a_{ij}\, g_j$$

*son gaussianas estándar independientes entre sí.*

*En particular, si $\sum_{j=1}^{n} a_j^2 = 1$ entonces $\widetilde{g} = \sum_{j=1}^{n} a_j\, g_j$ tiene distribución gaussiana estándar.*

*Demostración.* En efecto, dado $A \subseteq \mathbb{R}^n$ boreliano, basta con notar que

$$\mu\Big(U.G \in A\Big) = \mu\Big(G \in U^{-1}.A\Big) = \gamma_n(U^{-1}.A) = \gamma_n(A). \quad \square$$

Más que con vectores gaussianos estándar en $\mathbb{R}^n$, trabajaremos con la siguiente ligera generalización:

**Definición 2.2.2.** Sean $(\Omega, \Sigma, \mu)$ un espacio de probabilidad y $E$ un espacio normado de dimensión $n$. Decimos que una variable aleatoria $E$-valuada $X : \Omega \to E$ es *$n$-gaussiana* (o *gaussiana* a secas) si existen una base $\{x_1, \ldots, x_n\}$ de $E$, y $g_1, \ldots, g_n : \Omega \to \mathbb{R}$ variables aleatorias gaussianas estándar independientes tales que para cada $\omega \in \Omega$,

$$X(\omega) = \sum_{j=1}^{n} g_j(\omega)\, x_j.$$

Por otro lado, si $Y : \Omega \to E$ es una variable aleatoria $E$-valuada, notaremos $dist(Y)$ a la medida que $Y$ induce en los borelianos de $E$, definida por $dist(Y)(A) = \mu\Big(Y^{-1}(A)\Big)$ para cada $A \subseteq E$ boreliano. A esta medida la llamaremos *la distribución de $Y$*.

**Observación 2.2.3.** Supongamos que $E$ es un espacio normado de dimensión $n$ y $\{x_1, \ldots, x_n\} \subseteq E$ es una base de $E$. Supongamos además que $(\Omega_1, \Sigma_1, \mu_1)$ y $(\Omega_2, \Sigma_2, \mu_2)$ son dos espacios de probabilidad, en los cuales tenemos definidas a las funciones $g_i : \Omega_1 \to \mathbb{R}$ gaussianas estándar independientes y $h_i : \Omega_2 \to \mathbb{R}$ gaussianas estándar independientes (donde

$1 \leq i \leq n$). Notemos entonces que las variables aleatorias gaussianas $E$-valuadas definidas por

$$X_1 : \Omega_1 \to E \quad X_1(\omega) = \sum_{i=1}^{n} g_i(\omega) x_i \qquad \text{y} \qquad X_2 : \Omega_2 \to E \quad X_2(\omega) = \sum_{i=1}^{n} h_i(\omega) x_i$$

tienen la misma distribución.

En efecto, dado $A \subseteq E$ abierto, si definimos $T : E \to \mathbb{R}^n$ por $T(x_i) = e_i$ (donde $e_1, \ldots, e_n$ son los vectores de la base canónica de $\mathbb{R}^n$), entonces

$$\begin{aligned} &dist(X_1)(A) = \mu_1(X_1 \in A) = \mu_1(T(X_1) \in T(A)) = \mu_1((g_1, \ldots, g_n) \in T(A)) = \gamma_n(T(A)) \\ &= \mu_2((h_1, \ldots, h_n) \in T(A)) = \mu_2(T(X_2) \in T(A)) = \mu_2(X_2 \in A) = dist(X_2)(A). \end{aligned}$$

Es decir, la distribución de una variable aleatoria gaussiana $E$-valuada depende únicamente de los vectores $x_1, \ldots, x_n$.

De forma análoga a las variables aleatorias escalares, decimos que las variables aleatorias $E$-valuadas $X_1, \ldots, X_k : \Omega \to E$ son independientes entre sí, si dados $A_1, \ldots, A_k \subseteq E$ borelianos, vale que

$$\mu(X_1 \in A_1, \ldots, X_k \in A_k) = \prod_{i=1}^{k} \mu(X_i \in A_i).$$

**Observación 2.2.4.** En particular, si $E = \mathbb{R}^n$, tomando cada $A_i$ como un producto cartesiano de intervalos en $\mathbb{R}$

$$A_i = \prod_{j=1}^{n} I_i^j \quad I_i^j \subseteq \mathbb{R} \text{ intervalo}$$

puede verse que $G_1, \ldots, G_k : \Omega \to \mathbb{R}^n$ son vectores aleatorios independientes y con distribución gaussiana estándar si y solo si para cada $1 \leq i \leq k$ vale que $G_i = (g_{i1}, \ldots, g_{in})$ donde las variables aleatorias (escalares) $g_{ij} : \Omega \to \mathbb{R}$ con $1 \leq i \leq k$, $1 \leq j \leq n$ son gaussianas estándar independientes entre sí.

Al igual que con las funciones borelianas escalares [2] veamos que vale la siguiente proposición.

**Proposición 2.2.5.** *Dados un espacio de probabilidad $(\Omega, \Sigma, \mu)$ , $E$ y $F$ espacios normados de dimensión finita, $\varphi : F \to E$ boreliana y $X_1, \ldots, X_k : \Omega \to F$ variables aleatorias $F$-valuadas, vale que*

*(a) Si $X_1, \ldots, X_k$ tienen la misma distribución, entonces $\varphi(X_1), \ldots, \varphi(X_k)$ tienen la misma distribución.*

*(b) Si $X_1, \ldots, X_k$ son independientes, entonces $\varphi(X_1), \ldots, \varphi(X_k)$ son independientes entre sí.*

[2] Recordemos que si $X, Y$ son dos espacios topológicos, $\varphi : X \to Y$ es boreliana (o medible Borel), si $\varphi^{-1}(A) \in \mathcal{B}_X$ para todo $A \subseteq Y$ abierto, siendo $\mathcal{B}_X$ la $\sigma$–álgebra generada por los abiertos en $X$.

*Demostración.* Por simplicidad en la escritura, veamos ambos ítems para $k = 2$. Los mismos argumentos aplican para cualquier $k \in \mathbb{N}$.

(a) Dado $A \subseteq E$ abierto, tenemos que

$$\mu(\varphi(X_1) \in A) = \mu\left(X_1 \in \varphi^{-1}(A)\right) = \mu\left(X_2 \in \varphi^{-1}(A)\right) = \mu(\varphi(X_2) \in A).$$

(b) De manera similar, si $A, B \subseteq E$ son abiertos, entonces

$$\begin{aligned}\mu(\varphi(X_1) \in A, \varphi(X_2) \in B) &= \mu\left(X_1 \in \varphi^{-1}(A), X_2 \in \varphi^{-1}(B)\right) \\ &= \mu\left(X_1 \in \varphi^{-1}(A)\right) \mu\left(X_2 \in \varphi^{-1}(B)\right) = \mu(\varphi(X_1) \in A)\, \mu(\varphi(X_2) \in B).\end{aligned}$$

□

En la demostración del teorema de Dvoretzky utilizaremos el siguiente lema.

**Lema 2.2.6.** *Sean $F$ un espacio normado de dimensión $r$ y $\{z_1, \ldots, z_r\} \subseteq F$ una base de $F$. Dados $(\Omega, \Sigma, \mu)$ un espacio de probabilidad y $g_{ij} : \Omega \to \mathbb{R}$ variables aleatorias gaussianas estándar independientes entre sí, con $1 \leq i \leq k$ (donde $k \in \mathbb{N}$) y $1 \leq j \leq r$. Si para cada $1 \leq i \leq k$ $X_i : \Omega \to F$ es la variable aleatoria $F$-valuada gaussiana definida por $X_i(\omega) = \sum_{j=1}^{r} g_{ij}(\omega) z_j$ entonces*

*(a) Las variables aleatorias $X_i$ son independientes e igualmente distribuidas.*

*(b) Si $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ son tales que $\sum_{i=1}^{k} \alpha_i^2 = 1$, entonces la variable $F$-valuada $X : \Omega \to F$ definida por $X = \sum_{i=1}^{k} \alpha_i X_i$ es gaussiana, y $dist(X) = dist(X_i)$ para cada $1 \leq i \leq k$.*

*Demostración.* (a) Que las variables aleatorias $X_i$ tienen la misma distribución se deduce de la Observación 2.2.3, y que son independientes se deduce de la Proposición 2.2.5 aplicada a la transformación lineal $\varphi : \mathbb{R}^r \to F$ definida por $\varphi(e_i) = z_i$, y a los vectores gaussianos estándar $G_i : \Omega \to \mathbb{R}^r$, $G_i = (g_{i1}, \ldots, g_{ir})$, que por la Observación 2.2.4 son independientes entre sí.

(b) Basta con escribir

$$X = \sum_{i=1}^{k} \alpha_i X_i = \sum_{j=1}^{r} \left( \sum_{i=1}^{k} \alpha_i g_{ij} \right) z_j = \sum_{j=1}^{r} \widetilde{g_j} z_j,$$

donde para cada $1 \leq j \leq r$ la variable aleatoria $\widetilde{g_j} : \Omega \to \mathbb{R}$ viene dada por $\widetilde{g_j} = \sum_{i=1}^{k} \alpha_i g_{ij}$. En efecto, tenemos que $\widetilde{g_1}, \ldots, \widetilde{g_r}$ son gaussianas estándar por el Lema 2.2.1 y son independientes por la Proposición 2.2.5, ya que $\widetilde{g_j} = \varphi(\widetilde{G_j})$ donde $\varphi : \mathbb{R}^k \to \mathbb{R}$ esta dada por $\varphi(x_1, \ldots, x_k) = \sum_{i=1}^{k} \alpha_i x_i$ y los vectores aleatorios $\widetilde{G_j} = (g_{1j}, \ldots, g_{kj})$, con $1 \leq j \leq r$, son vectores gaussianos estándar independientes entre sí.

□

Si $(\Omega, \Sigma, \mu)$ es un espacio de probabilidad y $f \in L_1(\Omega, \mu)$, notaremos a la esperanza de $f$ como $\mathbb{E}(f) = \int_\Omega f(\omega)\, d\mu(\omega)$. A veces, para remarcar cuál es la medida (o el espacio) con respecto a la cual tomamos la esperanza, notaremos $\mathbb{E}(f) = \mathbb{E}_\mu(f)$.

Utilizaremos más adelante el siguiente resultado

**Lema 2.2.7.** *Sean $v \in \mathbb{R}^n$, y $g : \mathbb{R}^n \to \mathbb{R}$ dada por $g(x) = e^{\langle v, x\rangle}$. Entonces*

$$\mathbb{E}_{\gamma_n}(g) = \int_{\mathbb{R}^n} e^{\langle v, x\rangle}\, d\gamma_n(x) = e^{\frac{\|v\|_2^2}{2}}.$$

*Demostración.* Sea $u : \mathbb{R}^n \to \mathbb{R}^n$ ortogonal tal que $u^*(v) = \|v\|_2\, e_1$. Entonces

$$\begin{aligned}
&\int_{\mathbb{R}^n} e^{\langle v, x\rangle}\, d\gamma_n(x) = \int_{\mathbb{R}^n} e^{\langle v, u(x)\rangle}\, d\gamma_n(x) = \int_{\mathbb{R}^n} e^{\langle u^*(v), x\rangle}\, d\gamma_n(x) = \int_{\mathbb{R}^n} e^{\|v\|_2\, x_1}\, d\gamma_n(x)\\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{\|v\|_2\, x_1 - \frac{\|x\|_2^2}{2}}\, dx\\
&= \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\|v\|_2\, x_1 - \frac{x_1^2}{2}}\, dx_1\right)\left(\frac{1}{(2\pi)^{(n-1)/2}} \int_{\mathbb{R}^{n-1}} e^{-\frac{\|\widetilde{x}\|_2^2}{2}}\, d\widetilde{x}\right)\\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x_1 - \|v\|_2)^2}{2}}\, e^{\frac{\|v\|_2^2}{2}}\, dx_1 = e^{\frac{\|v\|_2^2}{2}}.
\end{aligned}$$

□

Veamos ahora como se comporta asintóticamente $\gamma_1$.

**Lema 2.2.8.** *Para cada $x > 0$*

$$\frac{1}{\sqrt{2\pi}}\, \frac{x}{1+x^2}\, e^{-x^2/2} < \gamma_1((x, +\infty)) < \frac{1}{\sqrt{2\pi}}\, \frac{1}{x}\, e^{-x^2/2}.$$

*Demostración.* Veamos primero la cota superior, para eso, basta integrar por partes

$$\int_x^{+\infty} e^{-t^2/2}\, dt = \int_x^{+\infty} \frac{1}{t}\,(t e^{-t^2/2})\, dt = \left(\frac{1}{t}\left(-e^{-t^2/2}\right)\right)\Big|_x^{+\infty} - \int_x^{+\infty} \frac{1}{t^2}\, e^{-t^2/2}\, dt < \frac{1}{x}\, e^{-x^2/2}.$$

De donde obtenemos la cota superior. Por otro lado, para la cota inferior basta con notar que la función

$$h(x) = \int_x^{+\infty} e^{-t^2/2}\, dt - \frac{x}{1+x^2}\, e^{-x^2/2}$$

tiene derivada negativa en $[0, +\infty)$, ya que al derivarla encontramos $h'(x) = \dfrac{-2e^{-x^2/2}}{(1+x^2)^2}$, y que su límite cuando $x$ tiende a $+\infty$ es cero, por lo que debe ser $h(x) > 0$ para todo $x > 0$. □

Utilizaremos también el siguiente lema (para una justificación acerca del número 22 millones en el enunciado del lema, ver la observación posterior a su demostración).

**Lema 2.2.9.** *Sean $(\Omega, \Sigma, \mu)$ un espacio de probabilidad y $g_1, \ldots, g_n : \Omega \to \mathbb{R}$ variables aleatorias gaussianas estándar independientes, con $n \geq 2{,}2.10^7$. Entonces*

$$\mathbb{E}\left(\max_{1 \leq i \leq n} |g_i|\right) \geq \frac{5}{4}\sqrt{\ln(n)}.$$

*Demostración.* Dado $a > 0$

$$\mu\left(\max_{1\le i\le n}|g_i| > a\right) = 1-\mu\left(\max_{1\le i\le n}|g_i| \le a\right) = 1-\prod_{i=1}^{n}\mu\left(|g_i|\le a\right) = 1-\mu\left(|g_1|\le a\right)^n$$
$$= 1-\left(1-\mu(|g_1|>a)\right)^n = 1-\left(1-\gamma_1(|x|>a)\right)^n = 1-\left(1-2\,\gamma_1(x>a)\right)^n.$$

Luego, usando la desigualdad de Markov y el lema anterior tenemos que

$$\mathbb{E}\left(\max_{1\le i\le n}|g_i|\right) \ge a\,\mu\left(\max_{1\le i\le n}|g_i| > a\right) = a\left(1-\left(1-2\,\gamma_1(x>a)\right)^n\right)$$
$$\ge a\left(1-\exp\left(-2n\,\gamma_1(x>a)\right)\right) \ge a.\left(1-\exp\left(-\sqrt{\frac{2}{\pi}}\,n\,\frac{a}{1+a^2}e^{-a^2/2}\right)\right).$$

Tomando $a = c\sqrt{\ln(n)}$, con $c > 0$, nos queda

$$\mathbb{E}\left(\max_{1\le i\le n}|g_i|\right) \ge c\sqrt{\ln(n)}.\left(1-\exp\left(-\sqrt{\frac{2}{\pi}}\,\frac{c\sqrt{\ln(n)}}{1+c^2\ln(n)}\,n^{1-c^2/2}\right)\right). \tag{2.1}$$

Si bien no usaremos esto para el hecho que queremos demostrar, puede observarse numéricamente que para cada $n\in\mathbb{N}$ existe un $c_n$ "óptimo" que maximiza la expresión que acompaña a $\sqrt{\ln(n)}$,

$$c\left(1-\exp\left(-\sqrt{\frac{2}{\pi}}\,\frac{c\sqrt{\ln(n)}}{1+c^2\ln(n)}\,n^{1-c^2/2}\right)\right).$$

Este $c_n$ tiene la forma

$$c_n = \sqrt{2-(1+\alpha_n)\frac{\ln(\ln(n))}{\ln(n)}}$$

donde $\alpha_n$ es decreciente y tiende (muy lentamente) a cero.

Reemplazando la expresión de $c_n$ en (2.1), llagamos a la desigualdad

$$\mathbb{E}\left(\max_{1\le i\le n}|g_i|\right) \ge c_n\sqrt{\ln(n)}\left(1-\exp\left(-\sqrt{\frac{2}{\pi}}\,\frac{c_n\ln(n)}{1+c_n^2\ln(n)}\,\ln^{\frac{\alpha_n}{2}}(n)\right)\right).$$

Es interesante notar que más allá de la validez de esta observación numérica, tomando un valor de $\alpha$ fijo y positivo (por ejemplo $\alpha = 1$) y la sucesión $c_n$ que le corresponde, podemos ver que

$$\mathbb{E}\left(\max_{1\le i\le n}|g_i|\right) \ge d_n\sqrt{\ln(n)} \quad \text{con} \quad d_n \xrightarrow[n\to+\infty]{} \sqrt{2}.$$

Ahora bien, volviendo al hecho que queremos demostrar, observemos que para $\alpha$ y $c > 0$ fijos, la expresión $\dfrac{c\ln(n)}{1+c^2\ln(n)}\ln^{\frac{\alpha}{2}}(n)$ es monótona creciente en $n$. En efecto, esto puede verse mediante el cálculo de la derivada de $\varphi(x) = \dfrac{c\ln(x)}{1+c^2\ln(x)}\ln^{\frac{\alpha}{2}}(x)$, por ejemplo.

Así, para obtener la cota concreta del enunciado, tomamos $\alpha = 1{,}47$. Entonces vemos que para $n = 2{,}2.10^7$ la expresión que acompaña a $\sqrt{\ln(n)}$ es mayor a $5/4$ (usando $c = c_{2,2.10^7}$), de donde tenemos que

$$\mathbb{E}\left(\max_{1\leq i\leq n} |g_i|\right) \geq \frac{5}{4}\sqrt{\ln(n)}$$

para todo $n \geq 2{,}2.10^7$.

$\square$

**Observación 2.2.10.**

(I) Con una cuenta similar (y un tanto menos rebuscada), podríamos haber visto que existe una constante $C > 0$ tal que para todo $n \geq 2$ se tiene

$$\mathbb{E}\left(\max_{1\leq i\leq n} |g_i|\right) \geq C\sqrt{\ln(n)}.$$

Esto vale, por ejemplo, para $C = 2/5$. Nos permitimos la "incomodidad" de obtener una desigualdad válida para $n$ mayor a 22 millones, ya que esto nos permitirá obtener mejores cotas para la dimensión $N(k,\varepsilon)$ en el teorema de Dvoretzky, siendo que las dimensiones involucradas en dicho teorema son intrínsecamente mucho mayores a la restricción de que $n \geq 2, 2.10^7$ (por lo que esta hipótesis se cumpliría más adelante aunque no la tomáramos).

(II) Por otro lado, usando que si $g$ es una variable aleatoria gaussiana estándar $\mathbb{E}\left(e^{t|g|}\right) = e^{t^2/2}$, puede verse aplicando logaritmo a la cadena de desigualdades

$$e^{t\mathbb{E}\left(\max\limits_{1\leq i\leq n}|g_i|\right)} \leq \mathbb{E}\left(e^{t\max\limits_{1\leq i\leq n}|g_i|}\right) = \mathbb{E}\left(\max_{1\leq i\leq n} e^{t|g_i|}\right) \leq \mathbb{E}\left(\sum_{i=1}^{n} e^{t|g_i|}\right) = \sum_{i=1}^{n}\mathbb{E}\left(e^{t|g_i|}\right) = ne^{t^2/2}$$

y tomando $t = \sqrt{2\ln(n)}$, que si las variables aleatorias $(g_i)_{i\in\mathbb{N}}$ son gaussianas estándar independientes entonces $\mathbb{E}\left(\max\limits_{1\leq i\leq n}|g_i|\right) \leq \sqrt{2}\sqrt{\ln(n)}$ para todo $n \geq 2$.

Por lo que

$$\frac{\mathbb{E}\left(\max\limits_{1\leq i\leq n}|g_i|\right)}{\sqrt{\ln(n)}} \xrightarrow[n\to\infty]{} \sqrt{2}.$$

(III) Notemos que, bajo las condiciones enunciadas en el lema, $\mathbb{E}\left(\max\limits_{1\leq i\leq n}|g_i|\right) = \mathbb{E}_{\gamma_n}(\|x\|_\infty)$.

### 2.2.1. Concentración de la medida gaussiana

Como mencionamos anteriormente, seguiremos el camino dado en [Pis89]. Así como la prueba dada por Milman, esta demostración usa la concentración de la medida (o más precisamente, cierta inecuación que limita la desviación de una función Lipschitz de su media), pero en $\mathbb{R}^n$ con la medida gaussiana estándar, en vez de utilizar la esfera unitaria con la medida de probabilidad invariante por rotaciones.

En la demostración del siguiente teorema, utilizaremos el teorema de Rademacher, que nos dice que dados $A \subseteq \mathbb{R}^n$ abierto y $f : A \to \mathbb{R}$ Lipschitz se tiene que $f$ es diferenciable en casi todo punto de $A$ (para una demostración ver [Eva15, Capítulo 3]).

**Teorema 2.2.11.** *Si $f : \mathbb{R}^n \to \mathbb{R}$ es Lipschitz*[3] *con constante $L > 0$ (con respecto a la norma euclídea), entonces para cada $t > 0$ vale que*

$$\gamma_n\Big( \big|f - \mathbb{E}_{\gamma_n}(f)\big| > t \Big) \leq 2\exp\left( -\frac{2}{\pi^2}\frac{t^2}{L^2} \right).$$

*Demostración.* Fijemos dos puntos $x, y \in \mathbb{R}^n$ y definamos la curva $c_{x,y} = c : [0, \pi/2] \to \mathbb{R}^n$ dada por $c(\theta) = x\sin(\theta) + y\cos(\theta)$. Observemos que $c$ es $\mathcal{C}^1$, con $c'(\theta) = x\cos(\theta) - y\sin(\theta)$.

A su vez, tenemos que como $f$ es Lipschitz es diferenciable en casi todo punto de $\mathbb{R}^n$ (tanto para la medida de Lebesgue como para $\gamma_n$) y $f \circ c : [0, \pi/2] \to \mathbb{R}$ es Lipschitz, por lo que es absolutamente continua.

Asumamos momentáneamente que $f$ es diferenciable en $c(\theta)$ para casi todo $\theta \in [0, \pi/2]$. Entonces tenemos que

$$f(x) - f(y) = f\Big(c(\pi/2)\Big) - f\Big(c(0)\Big) = \int_0^{\pi/2} (f \circ c)'(\theta)\, d\theta = \int_0^{\pi/2} \langle \nabla f(c(\theta)), c'(\theta)\rangle\, d\theta.$$

Ahora bien, considerando el espacio de probabilidad $I = \left([0, \pi/2], \mathcal{M}([0, \pi/2]), \frac{2}{\pi} m\right)$, podemos escribir la anterior igualdad como

$$f(x) - f(y) = \mathbb{E}_I\left( \frac{\pi}{2} \langle \nabla f(c(\theta)), c'(\theta)\rangle \right).$$

Ahora, si dado $\lambda \in \mathbb{R}$ notamos $\varphi_\lambda : \mathbb{R} \to \mathbb{R}_{>0}$ a $\varphi_\lambda(t) = e^{\lambda t}$, entonces $\varphi_\lambda$ es convexa, y por la desigualdad de Jensen tenemos que

$$\varphi_\lambda(f(x) - f(y)) \leq \mathbb{E}_I\left( \varphi_\lambda\left( \frac{\pi}{2}\langle \nabla f(c(\theta)), c'(\theta)\rangle \right)\right) = \frac{2}{\pi}\int_0^{\pi/2} \varphi_\lambda\left( \frac{\pi}{2}\langle \nabla f(c(\theta)), c'(\theta)\rangle \right) d\theta. \quad (2.2)$$

Veamos que la expresión de la derecha se puede integrar en $(\mathbb{R}^n \times \mathbb{R}^n, \gamma_n \times \gamma_n)$. Para eso, nos basta con ver que para casi todo $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $f$ es diferenciable en $c_{x,y}(\theta)$ para casi todo $\theta \in [0, \pi/2]$.

A tal fin, notemos $Z = \{z \in \mathbb{R}^n / f \text{ no es dif. en } z\}$ (sabemos que $m_n(Z) = 0$) y

$$N = \Big\{ (x, y, \theta) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, \pi/2] \text{ tales que } c_{x,y}(\theta) \in Z \Big\}.$$

Basta con ver entonces que $N$ es un conjunto de medida cero.

[3] Recordemos que una función $f : (\mathbb{R}^n, \|\,\|_2) \to (\mathbb{R}^m, \|\,\|_2)$ es Lipschitz con constante $L > 0$ si para todo $x, y \in \mathbb{R}^n$ vale que $\|f(x) - f(y)\|_2 \leq L\|x - y\|_2$.

Para eso, definamos la función $T : \mathbb{R}^n \times \mathbb{R}^n \times [0,\pi/2] \to \mathbb{R}^n \times \mathbb{R}^n \times [0,\pi/2]$ dada por $T(x,y,\theta) = (c_{x,y}(\theta), c'_{x,y}(\theta), \theta)$. Puede verse que $T$ diferenciable, con $|\det(DT(x,y,\theta))| = 1$ para todo $(x,y,\theta) \in \mathbb{R}^n \times \mathbb{R}^n \times [0,\pi/2]$, y es inversible, con $T^{-1} = T$. Entonces, de nuestro curso de análisis real tenemos que si $A \subseteq \mathbb{R}^n \times \mathbb{R}^n \times [0,\pi/2]$ es medible, entonces $T(A)$ lo es y $m(T(A)) = m(A)$. Luego

$$m_{2n+1}(N) = m_{2n+1}(T^{-1}(Z, \mathbb{R}^n, [0,\pi/2])) = m_{2n+1}(Z, \mathbb{R}^n, [0,\pi/2]) = m_n(Z) = 0.$$

Ahora sí, integrando (2.2) como habíamos anunciado, y usando Fubini, tenemos

$$\begin{aligned} &\iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda(f(x)-f(y))\, d\gamma_n \times \gamma_n(x,y) \\ &\leq \frac{2}{\pi}\int_0^{\pi/2} \left( \iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda\left(\frac{\pi}{2}\left\langle \nabla f(c_{x,y}(\theta)), c'_{x,y}(\theta)\right\rangle\right) d\gamma_n \times \gamma_n(x,y)\right) d\theta. \end{aligned}$$

Notemos que, como para cada $\theta \in [0,\pi/2]$ fijo la transformación $u : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ dada por $u(x,y) = (c_{x,y}(\theta), c'_{x,y}(\theta))$ es ortogonal, y a su vez la medida gaussiana $\gamma_{2n} = \gamma_n \times \gamma_n$ es invariante por transformaciones ortogonales, tenemos que el último término de la inecuación anterior es igual a

$$\begin{aligned} &\frac{2}{\pi}\int_0^{\pi/2} \left( \iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda\left(\frac{\pi}{2}\langle \nabla f(x), y\rangle\right) d\gamma_n \times \gamma_n(x,y)\right) d\theta = \iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda\left(\frac{\pi}{2}\langle \nabla f(x), y\rangle\right) d\gamma_n \times \gamma_n(x,y) \\ &= \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \varphi_\lambda\left(\frac{\pi}{2}\langle \nabla f(x), y\rangle\right) d\gamma_n(y)\right) d\gamma_n(x) \qquad \text{(usando el Lema 2.2.7)} \\ &= \int_{\mathbb{R}^n} \exp\left( \frac{\lambda^2\pi^2}{2^2} \frac{\|\nabla f(x)\|_2^2}{2}\right) d\gamma_n(x) \leq \exp\left(\frac{\lambda^2\pi^2}{2^3}L^2\right). \end{aligned}$$

Ya que si $f$ es diferenciable en $x$, $\|\nabla f(x)\|_2 \leq L$, puesto que para todo $y \in B_2^n$

$$|\langle \nabla f(x), y\rangle| = \lim_{t\to 0^+} \left| \frac{f(x+ty)-f(x)}{t} \right| \leq L.$$

Hasta aquí, obtuvimos entonces que

$$\iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda(f(x)-f(y))\, d\gamma_n \times \gamma_n(x,y) \leq \exp\left(\frac{\lambda^2\pi^2}{2^3}L^2\right).$$

Por otro lado, usando la desigualdad de Jensen nuevamente (notando $\mathbb{E}_{\gamma_n, y}$ a la esperanza con respecto a la medida gaussiana en la variable $y$)

$$\begin{aligned} &\varphi_\lambda\left(f(x) - \int_{\mathbb{R}^n} f(y)\, d\gamma_n(y)\right) = \varphi_\lambda\Big(\mathbb{E}_{\gamma_n,y}(f(x)-f(y))\Big) \leq \mathbb{E}_{\gamma_n,y}\Big(\varphi_\lambda(f(x)-f(y))\Big) \\ &= \int_{\mathbb{R}^n} \varphi_\lambda(f(x)-f(y))\, d\gamma_n(y). \end{aligned}$$

E integrando esta última desigualdad respecto a $x$

$$\int_{\mathbb{R}^n} \varphi_\lambda\left(f(x) - \int_{\mathbb{R}^n} f(y)\,d\gamma_n(y)\right) d\gamma_n(x)$$
$$\leq \iint_{\mathbb{R}^n\times\mathbb{R}^n} \varphi_\lambda(f(x)-f(y))\,d\gamma_n\times\gamma_n(x,y) \leq \exp\left(\frac{\lambda^2\pi^2}{2^3}L^2\right).$$

Luego, como para $\lambda > 0$ vale que $\varphi_\lambda$ es creciente, por la desigualdad de Markov tenemos que, si $\lambda, t > 0$

$$\gamma_n\left(x\in\mathbb{R}^n/\, f(x) - \mathbb{E}_{\gamma_n}(f) > t\right) \leq \frac{\mathbb{E}_{\gamma_n}\left(\varphi_\lambda(f - \mathbb{E}_{\gamma_n}(f))\right)}{\varphi_\lambda(t)} \leq \exp\left(\frac{\lambda^2\pi^2}{2^3}L^2 - \lambda t\right).$$

Para cada $t > 0$ fijo, podemos buscar el $\lambda > 0$ que minimice esta expresión (que es, en esencia, una cuadrática en $\lambda$), encontrando $\lambda = \frac{4t}{\pi^2L^2}$, de donde obtenemos que

$$\gamma_n\left(x\in\mathbb{R}^n/\, f(x) - \mathbb{E}_{\gamma_n}(f) > t\right) \leq \exp\left(\frac{-2}{\pi^2}\frac{t^2}{L^2}\right).$$

Repitiendo el procedimiento, pero integrando primero respecto de $x$ y después respecto a $y$ (o aplicando lo visto para $-f$), también tenemos que

$$\gamma_n\left(x\in\mathbb{R}^n/\, -f(x) + \mathbb{E}_{\gamma_n}(f) > t\right) \leq \exp\left(\frac{-2}{\pi^2}\frac{t^2}{L^2}\right).$$

De donde finalmente podemos deducir

$$\gamma_n\left(\left|f - \mathbb{E}_{\gamma_n}(f)\right| > t\right) \leq 2\exp\left(-\frac{2}{\pi^2}\frac{t^2}{L^2}\right). \quad \square$$

Observemos que la cota para la medida de $\gamma_n\left(\left|f - \mathbb{E}_{\gamma_n}(f)\right| > t\right)$ es independiente de la dimensión.

## 2.3. Lemas geométricos

Durante esta sección, si $\|\ \|$ es una norma en $\mathbb{R}^n$, notaremos $B = B_{(\mathbb{R}^n,\|\ \|)}$ a la bola unitaria, y $S = S_{(\mathbb{R}^n,\|\ \|)}$ a la esfera unitaria de la norma correspondiente.

**Lema 2.3.1.** *Sea $\|\ \|$ una norma en $\mathbb{R}^n$. Dado $\delta > 0$, existe un conjunto finito $A \subseteq S$ con a lo sumo $\left(1+\frac{2}{\delta}\right)^n$ puntos, tal que $d(x,A) \leq \delta\ \forall\, x \in S$.*

*Demostración.* Sea $A = \{y_1,\ldots,y_r\} \subseteq S$ un conjunto maximal con respecto a la propiedad de que $d(y,\widetilde{y}) > \delta\ \forall\, y \neq \widetilde{y}$ en $A$.

Para construir un tal conjunto, tomamos $y_1 \in S$. Si existe $y_2 \in S$ con $d(y_2, y_1) > \delta$, agregamos $y_2$ a nuestro conjunto. Si existe $y_3 \in S$ con $d(y_3, y_j) > \delta$ para todo $1 \leq j \leq 2$, agregamos $y_3$ al conjunto. Continuamos de esta manera hasta no poder más (por la compacidad de $S$, *debemos* no poder más en algún momento).

Sea $r \in \mathbb{N}$ el cardinal de $A$. Por la maximalidad de $A$, se cumple que $d(x, A) \leq \delta \ \forall x \in S$. Por otro lado, como $d(y, \widetilde{y}) > \delta \ \forall y \neq \widetilde{y}$ en $A$, las bolas $(y_j + \delta/2\, B)_{1 \leq j \leq r}$ son disjuntas, y están todas incluidas en $(1 + \delta/2)B$, por lo que

$$r\left(\frac{\delta}{2}\right)^n \mathrm{Vol}_n(B) = \mathrm{Vol}_n\left(\bigcup_{j=1}^r y_j + \frac{\delta}{2}B\right) \leq \mathrm{Vol}_n\left(\left(1 + \frac{\delta}{2}\right)B\right) = \left(1 + \frac{\delta}{2}\right)^n \mathrm{Vol}_n(B),$$

de donde

$$r \leq \left(\frac{2}{\delta} + 1\right)^n. \quad \square$$

Por brevedad (y también por convención) dado un espacio métrico $M$, diremos que un conjunto finito $A \subseteq M$ es una $\delta$-red en $M$ si $d(x, A) \leq \delta$ para todo $x \in M$. Así, el lema anterior nos dice que existe una $\delta$-red en $S$ cuyo cardinal "está controlado".

**Lema 2.3.2.** *Para cada $\varepsilon > 0$, existe $0 < \delta(\varepsilon) < 1/3$ tal que:*

*Dados $n \in \mathbb{N}$, $\| \ \|$ una norma en $\mathbb{R}^n$, y una $\delta$-red $A$ en la esfera unitaria $S = S_{(\mathbb{R}^n, \| \ \|)}$, si $E$ es un espacio de Banach y $x_1, \ldots, x_n \in B_E$ son tales que para todo $\alpha = (\alpha_1, \ldots, \alpha_n) \in A$ vale que*

$$1 - \delta(\varepsilon) \leq \left\| \sum_{i=1}^n \alpha_i x_i \right\|_E \leq 1 + \delta(\varepsilon), \tag{2.3}$$

*entonces*

$$\frac{1}{(1+\varepsilon)^{1/2}} \|\alpha\| \leq \left\| \sum_{i=1}^n \alpha_i x_i \right\|_E \leq (1+\varepsilon)^{1/2} \|\alpha\| \qquad \forall \alpha \in \mathbb{R}^n.$$

*En particular, si $F = \langle x_1, \ldots, x_n \rangle \subseteq E$, entonces $F$ es $(1+\varepsilon)$-isomorfo a $(\mathbb{R}^n, \| \ \|)$.*

*Demostración.* Sea $u : (\mathbb{R}^n, \| \ \|) \to E$ dada por $u(e_i) = x_i$. Veamos que $\|u\| \leq (1+\varepsilon)^{1/2}$ si escogemos $\delta$ de manera que se satisfaga (2.3) de forma adecuada.

Tomemos $\alpha \in S$, entonces existe $\widetilde{\alpha} \in A$ tal que $\|\alpha - \widetilde{\alpha}\| \leq \delta$. De donde podemos escribir

$$\|u(\alpha)\|_E = \|u(\widetilde{\alpha}) - u(\widetilde{\alpha} - \alpha)\|_E \leq \|u(\widetilde{\alpha})\|_E + \|u(\widetilde{\alpha} - \alpha)\|_E \leq 1 + \delta + \|u\|\,\delta.$$

Luego tenemos que $\|u\| \leq 1 + \delta + \|u\|\,\delta$, y por lo tanto $\|u\| \leq \dfrac{1+\delta}{1-\delta}$. Más adelante diremos como tomar $\delta$ de manera que $\dfrac{1+\delta}{1-\delta} \leq (1+\varepsilon)^{1/2}$.

Por otro lado, dado $\alpha \in S$, y escogiendo $\widetilde{\alpha} \in A$ como antes

$$\begin{aligned}\|u(\alpha)\|_E &= \|u(\widetilde{\alpha}) - u(\widetilde{\alpha} - \alpha)\|_E \geq \|u(\widetilde{\alpha})\|_E - \|u(\widetilde{\alpha} - \alpha)\|_E \\ &\geq 1 - \delta - \|u\|\,\delta \geq 1 - \delta - \frac{1+\delta}{1-\delta}\delta = \frac{1 - 2\delta + \delta^2 - \delta - \delta^2}{1-\delta} = \frac{1-3\delta}{1-\delta}.\end{aligned}$$

Por lo que, exigiendo $\delta < 1/3$, tenemos que $u : \mathbb{R}^n \to u(\mathbb{R}^n)$ es inversible y $\|u(\alpha)\|_E \geq \frac{1-3\delta}{1-\delta}\|\alpha\|$ para todo $\alpha \in \mathbb{R}^n$.

Ahora, queremos escoger $\delta$ de manera que

$$\frac{1+\delta}{1-\delta} \leq (1+\varepsilon)^{1/2} \quad \text{y} \quad \frac{1-3\delta}{1-\delta} \geq \frac{1}{(1+\varepsilon)^{1/2}}.$$

Si bien la demostración no nos entra en el margen de esta hoja, afirmamos que esto equivale a tomar

$$\delta \leq \frac{\varepsilon}{\left(1+\sqrt{1+\varepsilon}\right)^2} \quad \text{y} \quad \delta \leq \frac{\varepsilon}{\left(3\sqrt{1+\varepsilon}-1\right)\left(\sqrt{1+\varepsilon}+1\right)}.$$

Por lo que en general, bastará con tomar $\delta(\varepsilon)$ como el mínimo de estas dos cotas. Ahora, para rescatar una expresión más amigable, si consideramos $0 < \varepsilon \leq 1$, entonces ambas desigualdades se cumplen tomando $\delta(\varepsilon) = \varepsilon/8$. □

**Observación.** Los lemas que enunciamos en esta sección pueden utilizarse para demostrar el teorema de Dvoretzky a partir de la concentración de la medida en la esfera de $l_2^n$ (o más precisamente, a partir de una inecuación que, al igual que la del Teorema 2.2.11, limite la medida en la que una función Lipschitz en la esfera se desvía de su media, o su mediana).

## 2.4. Demostración del teorema de Dvoretzky

Demostraremos la siguiente versión, un poco más precisa, y mucho menos elegante, del teorema de Dvoretzky enunciado en la Introducción.

**Teorema 2.4.1.** *Sea $\|\ \|$ una norma en $\mathbb{R}^n$. Entonces, dado $\varepsilon > 0$, para cada $k \in \mathbb{N}$ que satisfaga*

$$k < \frac{\frac{9}{400\pi^2}\ln(n)\,\varepsilon^2 - \frac{9\ln(50)}{200\pi^2}\,\varepsilon^2 - \ln(2)}{\ln(1+16/\varepsilon)}$$

*existe un subespacio $k$-dimensional $S \subseteq \mathbb{R}^n$ tal que $(S, \|\ \|)$ es $(1+\varepsilon)$-isomorfo a $l_2^k$.*

*Demostración.* Podemos asumir que $0 < \varepsilon \leq 1$, y tomar entonces $\delta(\varepsilon) = \varepsilon/8 < 1/3$, de manera que se satisface lo enunciado en el Lema 2.3.2.

En este caso, la cota para $k$ es negativa si $n \leq 2^{400\pi^2/9} \sim 1{,}1\,.\,10^{132}$, ya que resulta $\frac{9}{400\pi^2}\ln(n)\,\varepsilon^2 - \ln(2) \leq 0$. Así, podemos asumir $n \geq 10^{132}$.

Sea $r = \left[\sqrt{n}/50 + 1\right] \gg 2{,}2\,.\,10^7$. Entonces, por el teorema de factorización de Dvoretzky-Rogers 2.1.1 con $\widetilde{\varepsilon} = 1/24$, podemos tomar $r$ vectores $z_1, \ldots, z_r \in \mathbb{R}^n$ de manera que se cumpla

$$\frac{24}{25}\|(\alpha_i)\|_\infty \leq \left\|\sum_{j=1}^{r}\alpha_i z_i\right\| \leq \|(\alpha_i)\|_2 \tag{2.4}$$

para cualesquiera escalares $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$. Llamemos $F = \langle z_1, \ldots, z_r\rangle$.

Proveámonos ahora de un espacio de probabilidad $(\Omega, \Sigma, \mu)$ junto con variables aleatorias gaussianas estándar independientes $g_{ij} : \Omega \to \mathbb{R}$, con $1 \leq i \leq k$ (donde $k$ es como en el enunciado) y $1 \leq j \leq r$. Consideremos entonces las variables aleatorias $F$-valuadas $X_i : \Omega \to F$, con $1 \leq i \leq k$, dadas por

$$X_i = \sum_{j=1}^{r} g_{ij}\, z_j.$$

Entonces, por el Lema 2.2.6 las variables aleatorias $F-$valuadas $X_i$ son independientes e igualmente distribuidas.

Ahora, sea $A \subseteq \partial B_2^k$ una $\delta(\varepsilon)$-red, de manera que se satisfaga lo enunciado en el Lema 2.3.1, es decir tal que $|A| \leq (1+2/\delta)^k$, donde $|A|$ es el cardinal de $A$.

Para cada $\alpha = (\alpha_1, \ldots, \alpha_k) \in A$, consideremos la variable aleatoria $F$-valuada $G_\alpha = \sum_{i=1}^k \alpha_i\, X_i$.

Entonces, nuevamente por el Lema 2.2.6, para cada $\alpha \in A$, $G_\alpha$ tiene la misma distribución que las $X_i$ (pues $\alpha \in \partial B_2^k$).

Queremos ver que para algún $\omega \in \Omega$, y algún $M > 0$ vale que

$$\left| \frac{1}{M} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| \leq \delta(\varepsilon) \quad \forall\, \alpha \in A,$$

ya que entonces, tomando $y_i = X_i(\omega)/M$ para $1 \leq i \leq k$, tendríamos por el Lema 2.3.2 que $S = \langle y_1, \ldots, y_k \rangle \subseteq \mathbb{R}^n$ es $(1+\varepsilon)$-isomorfo a $l_2^k$.

Veremos que, en efecto, si $M = \mathbb{E}(\|X_1\|)$

$$\mu\left( \omega \in \Omega / \; \forall\, \alpha \in A \;\; \left| \frac{1}{M} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| \leq \delta(\varepsilon) \right) > 0.$$

De donde obtendremos lo buscado.

Para esto, notemos que

$$\begin{aligned} &\mu\left( \omega \in \Omega / \; \forall\, \alpha \in A \;\; \left| \frac{1}{M} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| \leq \delta(\varepsilon) \right) \\ &= 1 - \underbrace{\mu\left( \omega \in \Omega / \; \exists\, \alpha \in A \text{ con } \left| \frac{1}{M} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| > \delta(\varepsilon) \right)}_{(*)}. \end{aligned}$$

Basta con ver entonces que $(*) < 1$.

Comencemos notando que si $f : (l_2^r, \gamma_r) \to \mathbb{R}$ esta dada por

$$f(x_1, \ldots, x_r) = \left\| \sum_{j=1}^{r} x_j\, z_j \right\|.$$

Entonces, por (2.4), $f$ es Lipschitz con constante $L = 1$, y como variable aleatoria, por la Observación 2.2.3, $f$ tiene la misma distribución que $\|X_1\|$. Por lo que, usando el Teorema 2.2.11 con $t = \mathbb{E}_{\gamma_r}(f)\delta(\varepsilon)$, nos queda

$$\mu\Big(\omega \in \Omega/ \mid \|X_1(\omega)\| - M\mid > M\delta(\varepsilon)\Big) = \gamma_r\Big(x \in \mathbb{R}^r / \left|f(x) - \mathbb{E}_{\gamma_r}(f)\right| > \mathbb{E}_{\gamma_r}(f)\delta(\varepsilon)\Big)$$
$$\leq 2\exp\left(\frac{-2}{\pi^2}\left(\mathbb{E}_{\gamma_r}(f)\delta(\varepsilon)\right)^2\right).$$

Ahora bien, de vuelta por (2.4), tenemos que $\frac{24}{25}\|x\|_\infty \leq f(x)$ para cada $x \in \mathbb{R}^r$, por lo que, Lema 2.2.9 mediante

$$\mathbb{E}_{\gamma_r}(f) \geq \frac{24}{25}\mathbb{E}_{\gamma_r}(\|x\|_\infty) \geq \frac{24}{25}\frac{5}{4}\sqrt{\ln(r)} \geq \frac{6}{5}\sqrt{\ln\left(\frac{\sqrt{n}}{50}\right)}. \tag{2.5}$$

Así, llegamos a que

$$\mu\Big(\omega \in \Omega/ \frac{1}{M}\mid \|X_1(\omega)\| - 1\mid > \delta(\varepsilon)\Big) \leq 2\exp\left(\frac{-72}{25\pi^2}\ln\left(\frac{\sqrt{n}}{50}\right)\delta^2(\varepsilon)\right).$$

Luego

$$\mu\left(\omega \in \Omega/ \exists\, \alpha \in A \text{ con } \left|\frac{1}{M}\left\|\sum_{i=1}^{k}\alpha_i X_i(\omega)\right\| - 1\right| > \delta(\varepsilon)\right)$$
$$\leq \sum_{\alpha\in A}\mu\left(\omega \in \Omega/ \left|\frac{1}{M}\left\|\sum_{i=1}^{k}\alpha_i X_i(\omega)\right\| - 1\right| > \delta(\varepsilon)\right)$$
$$= |A|\,\mu\left(\omega \in \Omega/ \left|\frac{1}{M}\|X_1(\omega)\| - 1\right| > \delta(\varepsilon)\right)$$
$$\leq 2\left(1 + \frac{2}{\delta(\varepsilon)}\right)^k \exp\left(\frac{-72}{25\pi^2}\ln\left(\frac{\sqrt{n}}{50}\right)\delta^2(\varepsilon)\right)$$
$$= \exp\left(\ln(2) + k\ln\left(1 + \frac{16}{\varepsilon}\right) + \frac{9}{200\pi^2}\ln(50)\,\varepsilon^2 - \frac{9}{400\pi^2}\ln(n)\,\varepsilon^2\right) < 1$$

si $k$ es como en el enunciado. $\square$

**Observación 2.4.2.** Esto nos dice que, dados $k \in \mathbb{N}$ y $\varepsilon > 0$ alcanza con tomar

$$n(k,\varepsilon) \geq 2500\,.\,2^{\frac{400\pi^2}{9\varepsilon^2}}\,.\,e^{\frac{400\pi^2}{9}k\frac{\ln(1+16/\varepsilon)}{\varepsilon^2}}. \tag{2.6}$$

Por ejemplo, si buscamos un subespacio de dimensión 4 a distancia 1 de $l_2^4$, en general debemos buscar en un espacio de dimensión $n \geq 2{,}4.\,10^{2294}$, y si buscamos un subespacio de dimensión 100 a distancia 0,1 de $l_2^{100}$, debemos buscar en general en un espacio de dimensión $n \geq 10^{9693430}$.

Notemos que, para llegar a una expresión más sencilla, de (2.6) podemos decir que, según la definición de $N(k,\varepsilon)$ dada en la página viii de la Introducción,

$$N(k,\varepsilon) \leq \exp\left(549\frac{k\ln(17/\varepsilon)}{\varepsilon^2}\right)$$

si $0 < \varepsilon \leq 1$ y $k \geq 1$ (este es el mismo tipo de comportamiento asintótico obtenido por Milman en [Mil71]). Escrito de otra forma, tenemos que, si $k(n,\varepsilon)$ es definido como en la página x de la Introducción,

$$k(n,\varepsilon) \geq \frac{\varepsilon^2}{549\ln(17/\varepsilon)}\ln(n).$$

Por último, como también mencionamos en la Introducción, el comportamiento asintótico con respecto a $\varepsilon$ de $N(k,\varepsilon)$ ó $k(n,\varepsilon)$ obtenido en esta demostración no es óptimo (así como tampoco las constantes obtenidas).

## 2.5. Algunas observaciones acerca de $k(X,\varepsilon)$

Vimos que para todo espacio de Banach $X$ de dimensión $n$, $k(X,\varepsilon) \geq \dfrac{\varepsilon^2}{549\ln(17/\varepsilon)}\ln(n)$.

Veremos en esta sección dos resultados, ambos mencionados en la Introducción.

El primer resultado nos dirá que el comportamiento logarítmico en $n$ de $k(n,\varepsilon)$ obtenido en la demostración anterior del teorema de Dvoretzky es óptimo.

**Proposición 2.5.1.** *Si $0 < \varepsilon \leq 1$, entonces se cumple $k(n,\varepsilon) \leq k(l_\infty^n,\varepsilon) \leq 8\pi e\ln(n)$.*

El segundo resultado nos indica con mayor precisión el comportamiento de $k(l_\infty^n,\varepsilon)$.

**Proposición 2.5.2.** *Para todo $0 < \varepsilon \leq 1$ vale que*

$$k(l_\infty^n,\varepsilon) \geq \frac{1}{2\ln(1+4/\varepsilon)}\ln(n) \geq \frac{1}{2\ln(5/\varepsilon)}\ln(n).$$

### 2.5.1. Comportamiento óptimo con respecto a $n$

Para lo que sigue, utilizaremos la función gamma, $\Gamma : (0,+\infty) \to (0,+\infty)$, dada por

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\,dt.$$

De las tantas cosas que podrían decirse respecto a esta función, mencionemos que satisface la ecuación funcional $\Gamma(x+1) = x\Gamma(x)$ para todo $x > 0$ (como debería toda buena extensión holomorfa de $(n-1)!$), lo cual puede verse fácilmente integrando por partes.

Por otro lado, gracias a la Proposición 2.5.3 tendremos que $\Gamma$ es dos veces derivable, y que

$$\Gamma''(x) = \int_0^{+\infty} t^{x-1}\ln^2(t)e^{-t}\,dt > 0,$$

por lo que $\Gamma$ es convexa.

**Proposición 2.5.3.** *Sean $(\Omega,\Sigma,\mu)$ un espacio de medida, $A \subseteq \mathbb{R}$ abierto, y $f:\Omega\times A \to \mathbb{R}$ una función tal que:*

- *Para cada $x \in A$ la función $f(\cdot,x) \in L_1(\Omega,\mu)$.*
- *Para casi todo $\omega \in \Omega$ $f(\omega,\cdot)$ es derivable en $A$.*
- *Para cada $x_0 \in A$ existen un entorno compacto $K \subseteq A$ de $x_0$ y una función $g_K \in L_1(\Omega,\mu)$ tal que*
$$\left|\frac{\partial f}{\partial x}(\omega,x)\right| \leq g_K(\omega)$$
*para casi todo $\omega \in \Omega$, y para todo $x \in K$.*

*Entonces la función $h: A \to \mathbb{R}$ dada por*

$$h(x) = \int_\Omega f(\omega,x)\,d\mu(\omega)$$

*es derivable en $A$, y vale que*
$$h'(x) = \int_\Omega \frac{\partial f}{\partial x}(\omega,x)\,d\mu(\omega).$$

*Demostración.* Sea $D = \{\omega \in \Omega \,/\, f(\omega,\cdot) \text{ es derivable en } A\}$. Dado $x_0 \in A$, sean $r > 0$ tal que $K = [x_0 - r, x_0 + r] \subseteq A$ y $g_K \in L_1(\Omega,\mu)$ como en el enunciado. Entonces, si $0 < |s| < r$

$$\frac{h(x_0+s)-h(x_0)}{s} = \int_\Omega \frac{f(\omega,x_0+s)-f(\omega,x_0)}{s}\,d\mu(\omega).$$

Si $\omega \in D$, entonces existe $t_\omega$ entre $x_0+s$ y $x_0$ tal que

$$\left|\frac{f(\omega,x_0+s)-f(\omega,x_0)}{s}\right| = \left|\frac{\partial f}{\partial x}(\omega,t_\omega)\right| \leq g_K(\omega),$$

donde la última desigualdad vale para casi todo $\omega \in D$ (y por lo tanto para casi todo $\omega \in \Omega$). Luego, por el teorema de convergencia dominada, existe

$$h'(x_0) = \lim_{s\to 0} \frac{h(x_0+s)-h(x_0)}{s} = \int_\Omega \frac{\partial f}{\partial x}(\omega,x_0)\,d\mu(\omega). \quad \square$$

**Observación 2.5.4.** En verdad vale que $\Gamma$ es logarítmicamente convexa, es decir

$$\Gamma(\alpha x + (1-\alpha)y) \leq \Gamma^\alpha(x)\Gamma^{1-\alpha}(y)$$

para todo $0 \leq \alpha \leq 1$ y $x,y > 0$.

**Lema 2.5.5.** *Para todo $x > 0$ vale que*

$$\Gamma(x+1)^{1/x} \leq \frac{x}{2} + 1.$$

*Demostración.* Comencemos notando que $\Gamma(x) \leq 1$ para todo $x \in [1,2]$ ya que $\Gamma(1) = \Gamma(2) = 1$ y $\Gamma$ es convexa. Por estos mismos hechos (teorema de Rolle mediante) tenemos que $\Gamma$ es creciente y mayor o igual a 1 en $[2,+\infty)$.

Ahora, si $0 < x < 1$, por lo observado arriba tenemos que $\Gamma(x+1) \leq 1 < 1 + x/2$.

Por otro lado, si $x \geq 1$, notemos con $n = [x]$ a la parte entera de $x$, y $\alpha = \{x\}$ a la parte fraccionaria de $x$. Así, $x = n + \alpha$, con $n \in \mathbb{N}$ y $0 \leq \alpha < 1$. Entonces

$$\Gamma(x+1) = \Gamma(n+\alpha+1) = \left(\prod_{j=0}^{n-1} n - j + \alpha\right)\Gamma(1+\alpha) \leq \left(\prod_{j=0}^{n-1} n - j + \alpha\right).$$

Luego, si $x \geq 1$

$$\Gamma(x+1)^{1/x} \leq \Gamma(x+1)^{1/n} \leq \left(\prod_{j=0}^{n-1} n - j + \alpha\right)^{1/n} \leq \frac{1}{n}\sum_{j=0}^{n-1} n - j + \alpha = \frac{1}{n}\left(\frac{n(n+1)}{2} + n\alpha\right)$$
$$= \frac{n+1}{2} + \alpha = \frac{n+1+2\alpha}{2} = \frac{x+1+\alpha}{2} \leq \frac{x}{2} + 1. \qquad \square$$

**Observación.** Sin bien no utilizaremos esto, se puede ver que $\lim_{x\to+\infty} \frac{\Gamma(x+1)^{1/x}}{x} = \frac{1}{e}$, por el teorema de Stirling.

En lo que sigue, dados un espacio de probabilidad $(\Omega, \Sigma, \mu)$, una variable aleatoria gaussiana estándar $g : \Omega \to \mathbb{R}$ y $p \geq 1$, notaremos

$$\gamma(p) = \|g\|_{L_p(\Omega)} = \left(\int_\Omega |g(\omega)|^p \, d\mu(\omega)\right)^{1/p}$$

al $p$-ésimo momento de $g$.

**Lema 2.5.6.** *Sean $(\Omega, \Sigma, \mu)$ un espacio de probabilidad, $g : \Omega \to \mathbb{R}$ una variable aleatoria gaussiana estándar y $p \geq 1$. Entonces $\gamma(p) \leq \sqrt{2p}$.*

*Demostración.* Usando la medida gaussiana estándar en $\mathbb{R}$, podemos escribir mediante la sustitución $s = t^2/2$

$$\gamma(p)^p = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |t|^p \, e^{-t^2/2} \, dt = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} t^p \, e^{-t^2/2} \, dt = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} (2s)^{(p-1)/2} e^{-s} \, ds$$
$$= \frac{2^{p/2}}{\sqrt{\pi}} \int_0^{+\infty} s^{(p+1)/2-1} \, e^{-s} \, ds = \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right).$$

De donde vemos que

$$\gamma(p) = \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}\right)^{1/p}. \tag{2.7}$$

Usando el lema anterior, nos queda que

$$\gamma(p) = \frac{\sqrt{2}}{\pi^{1/(2p)}}\,\Gamma\left(\frac{p-1}{2}+1\right)^{1/p} \leq \sqrt{2}\left(1+\frac{p-1}{4}\right)^{(p-1)/2p} \leq \sqrt{2}\left(\frac{p+3}{4}\right)^{1/2} \leq \sqrt{2p}. \qquad \square$$

Recordemos que una variable aleatoria $g : (\Omega,\Sigma,\mu) \to \mathbb{R}$ tiene distribución gaussiana con media $m$ y varianza $\sigma^2$ (lo cual notamos $g \sim \mathcal{N}(m,\sigma^2)$) si

$$\mu(g \in (a,b)) = \frac{1}{\sqrt{2\pi}\sigma}\int_a^b e^{-(t-m)^2/2\sigma^2}\,dt.$$

En tal caso tenemos que $\mathbb{E}(g) = m$ y $\mathbb{E}((g-m)^2) = \sigma^2$. Además vale que $g_s = \frac{g-m}{\sigma}$ tiene distribución gaussiana estándar. Por lo que podemos escribir $g = \sigma\, g_s + m$, con $g_s$ gaussiana estándar.

**Proposición 2.5.7.** *Sean $n \geq 2$, $(\Omega,\Sigma,\mu)$ un espacio de probabilidad y $g_1,\ldots,g_n : \Omega \to \mathbb{R}$ variables aleatorias gaussianas (no necesariamente independientes) con media cero y varianzas $\sigma_1^2,\ldots,\sigma_n^2$ respectivamente. Entonces*

$$\mathbb{E}\left(\max_{1\leq i\leq n}|g_i|\right) \leq 2e^{1/2}\sqrt{\ln(n)}\max_{1\leq i\leq n}\sigma_i.$$

*Demostración.* Podemos escribir a cada $g_i$ como $g_i = \sigma_i\,\widetilde{g_i}$, donde $\widetilde{g_i}$ es una variable aleatoria gaussiana estándar.

En particular, $\|g_i\|_{L_p(\Omega)} = \sigma_i\gamma(p)$ para cada $p \geq 1$.

Ahora bien, si notamos

$$X = \left(\sum_{i=1}^n |g_i|^p\right)^{1/p},$$

entonces por la desigualdad de Hölder resulta $\|X\|_{L_1(\Omega)} \leq \|X\|_{L_p(\Omega)}$. Es decir

$$\mathbb{E}\left(\left(\sum_{i=1}^n |g_i|^p\right)^{1/p}\right) \leq \left(\sum_{i=1}^n \mathbb{E}\left(|g_i|^p\right)\right)^{1/p} = \left(\sum_{i=1}^n \sigma_i^p\,\gamma(p)^p\right)^{1/p} \leq n^{1/p}\,\gamma(p)\max_{1\leq i\leq n}\sigma_i.$$

En particular tenemos que

$$\mathbb{E}\left(\max_{1\leq i\leq n}|g_i|\right) \leq \mathbb{E}\left(\left(\sum_{i=1}^n |g_i|^p\right)^{1/p}\right) \leq n^{1/p}\,\gamma(p)\max_{1\leq i\leq n}\sigma_i \leq \sqrt{2}\,n^{1/p}\sqrt{p}\max_{1\leq i\leq n}\sigma_i$$

para todo $p \geq 1$. Tomando $p = 2\ln(n) \geq 1$ para todo $n \geq 2$, nos queda

$$\mathbb{E}\left(\max_{1\leq i\leq n}|g_i|\right) \leq 2e^{1/2}\sqrt{\ln(n)}\max_{1\leq i\leq n}\sigma_i. \qquad \square$$

**Definición 2.5.8.** Sean $(\Omega,\Sigma,\mu)$ un espacio de probabilidad y $g_j : \Omega \to \mathbb{R}$, con $j \in \mathbb{N}$, una sucesión de variables aleatorias gaussianas estándar e independientes. Dado un espacio normado $E$ de dimensión $n$, definimos

$$\delta(E) = \sup\left\{\left(\mathbb{E}\left(\left\|\sum_{i=1}^k g_i\,u(e_i)\right\|_E\right)\right)^2, \text{ con } k \leq n,\ u : l_2^k \to E \text{ inyectiva y } \|u\| = 1\right\}.$$

**Observación 2.5.9.**

(I) Por lo mencionado en la Observación 2.2.3, $\delta(E)$ no depende del espacio de probabilidad o las variables aleatorias escogidas, siempre que satisfagan lo enunciado en la definición.

(II) Durante la demostración del teorema de Dvoretzky, vimos en (2.5) que si $E$ es un espacio normado de dimensión $n$, con $n$ tal que $\sqrt{n}/50 > 2,2.10^7$, entonces

$$\delta(E) \geq \frac{36}{25}\ln(\sqrt{n}/50).$$

(III) Más aún, en la demostración del teorema de Dvoretzky se ve (*mutatis mutandis*) que si

$$k < \frac{\frac{\varepsilon^2}{32\pi^2}\delta(E) - \ln(2)}{\ln(1+16/\varepsilon)},$$

entonces existe un subespacio $k$-dimensional $(1+\varepsilon)$-isomorfo a $l_2^k$ en $E$.

La cantidad $\delta(E)$ nos es de interés por la siguiente proposición.

**Proposición 2.5.10.** *Sea E un espacio normado, entonces para todo $0 < \varepsilon \leq 1$ vale que*

$$k(E,\varepsilon) \leq 2\pi\,\delta(E).$$

*Demostración.* Sean $k = k(E,\varepsilon)$, $u : l_2^k \to E$ inyectiva con $\|u\| = 1$, $\left\|u^{-1}\right\| \leq 1+\varepsilon$. Notemos $S = u(l_2^k) = \langle u(e_1),\dots,u(e_k)\rangle \subseteq E$.

Tomemos un espacio de probabilidad $(\Omega,\Sigma,\mu)$ junto con variables aleatorias gaussianas estándar independientes $g_1,\dots,g_k : \Omega \to \mathbb{R}$. Definamos la variable aleatoria gaussiana $S$-valuada $X : \Omega \to S$ dada por

$$X(\omega) = \sum_{i=1}^{k} g_i(\omega)\,u(e_i).$$

Entonces, usando la desigualdad de Jensen (que en este caso es la desigualdad triangular integral) y que $\gamma(1) = \sqrt{2/\pi}$,

$$\left\|u^{-1}\right\|\mathbb{E}\left(\|X\|_E\right) = \mathbb{E}\left(\left\|u^{-1}\right\|\|X\|_E\right) \geq \mathbb{E}\left(\left\|u^{-1}X\right\|_2\right) = \mathbb{E}\left(\|(|g_1|,\dots,|g_k|)\|_2\right) \tag{2.8}$$

$$\geq \|\mathbb{E}(|g_1|,\dots,|g_k|)\|_2 = \left\|\sqrt{\frac{2}{\pi}}(1,\dots,1)\right\|_2 = \sqrt{\frac{2}{\pi}}\sqrt{k}. \tag{2.9}$$

Luego

$$k \leq \frac{\pi}{2}\mathbb{E}(\|X\|_E)^2(1+\varepsilon)^2 \leq 2\pi\,\delta(E). \quad \square$$

Para terminar con la demostración de la Proposición 2.5.1, veamos que $\delta(l_\infty^n) \leq 4e\ln(n)$. Para esto, utilizaremos el siguiente lema acerca de la norma de un operador $u : l_2^k \to l_\infty^n$.

**Lema 2.5.11.** *Sea $u : l_2^k \to l_\infty^n$ un operador lineal. Entonces $\|u\| = \max_{1\le i\le n} \|u^*(e_i)\|_2$.*

*Es decir, si la matriz de $u$ en las bases canónicas de $l_2^k$ y $l_\infty^n$ es*

$$[u] = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix},$$

*entonces*

$$\|u\| = \max_{1\le i\le n} \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2}.$$

*Demostración.* Podemos asumir $u \neq 0$. Sea $(t_1,\dots,t_k) \in S_2^k = \partial B_2^k$, entonces, usando la desigualdad de Hölder

$$\|u(t_1,\dots,t_k)\|_\infty = \max_{1\le i\le n} \left| \sum_{j=1}^{k} a_{ij}\, t_j \right| \le \max_{1\le i\le n} \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2}.$$

Por lo que

$$\|u\| \le \max_{1\le i\le n} \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2}.$$

Por otro lado, si $1 \le i_0 \le n$ es tal que

$$\left( \sum_{j=1}^{k} a_{i_0 j}^2 \right)^{1/2} = \max_{1\le i\le n} \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2},$$

tomando $(t_1,\dots,t_k) = \dfrac{1}{\left\|(a_{i_0 1},\dots,a_{i_0 k})\right\|_2}(a_{i_0 1},\dots,a_{i_0 k})$, tenemos que

$$\|u\| \ge \|u(t_1,\dots,t_k)\|_\infty = \max_{1\le i\le n} \left| \sum_{j=1}^{k} a_{ij}\, t_j \right| \ge \left( \sum_{j=1}^{k} a_{i_0 j}^2 \right)^{1/2} = \max_{1\le i\le n} \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2}.$$

De donde obtenemos la igualdad deseada. □

**Proposición 2.5.12.** *Para cada $n \in \mathbb{N}$ vale que $\delta(l_\infty^n) \le 4e\ln(n)$.*

*Demostración.* Sean $k \le n$ y $u : l_2^k \to l_\infty^n$ lineal e inyectiva con $\|u\| = 1$, cuya matriz canónica es $[u] = (a_{ij}) \in \mathbb{R}^{n\times k}$. Dados un espacio de probabilidad $(\Omega, \Sigma, \mu)$ junto con variables aleatorias $g_1,\dots,g_k : \Omega \to \mathbb{R}$ gaussianas independientes estándar, observemos que si

$$\sigma_i = \left( \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2},$$

entonces para cada $1 \leq i \leq n$ la variable aleatoria

$$X_i = \sum_{j=1}^{k} a_{ij}\, g_j = \sigma_i \sum_{j=1}^{k} \frac{a_{ij}}{\sigma_i}\, g_j$$

es gaussiana (por el Lema 2.2.1) con media cero y varianza $\sigma_i$ (notemos que las $X_i$ pueden no ser independientes entre sí). Así, usando la Proposición 2.5.7 y el lema anterior, tenemos que

$$\mathbb{E}\left(\left\|\sum_{i=1}^{k} g_i\, u(e_i)\right\|_\infty\right) = \mathbb{E}\left(\max_{1\leq i\leq n}\left|\sum_{j=1}^{k} a_{ij}\, g_j\right|\right) \leq 2e^{1/2}\sqrt{\ln(n)}\, \max_{1\leq i\leq n} \sigma_i$$
$$= 2e^{1/2}\sqrt{\ln(n)}\,\|u\| = 2e^{1/2}\sqrt{\ln(n)}.$$

De donde obtenemos lo buscado. □

La demostración de la Proposición 2.5.1 se sigue de combinar la anterior proposición con la Proposición 2.5.10.

### 2.5.2. Comportamiento de $k(l_\infty^n, \varepsilon)$

La demostración de la Proposición 2.5.2 estará basada en el siguiente resultado.

**Proposición 2.5.13.** *Sean $F$ un espacio normado de dimensión $k$ y $0 < \delta < 1$. Entonces existen $n \in \mathbb{N}$, con $n \leq (1 + 2/\delta)^k$, y $\widetilde{F} \subseteq l_\infty^n$ con $\dim(\widetilde{F}) = k$ tal que $d_{BM}(F, \widetilde{F}) \leq (1-\delta)^{-1}$.*

*Demostración.* Tomemos, con permiso del Lema 2.3.1, una $\delta$-red $(\varphi_i)_{1\leq i\leq n}$ en $S_{F^*}$, con $n \leq (1+2/\delta)^k$.

Ahora, dado $x \in F$, sabemos que existe $\varphi \in S_{F^*}$ tal que $\varphi(x) = \|x\|$. Sea $1 \leq i_0 \leq n$ tal que $\left\|\varphi - \varphi_{i_0}\right\| \leq \delta$. Entonces

$$\|x\| = \varphi_{i_0}(x) + (\varphi - \varphi_{i_0})(x) \leq \max_{1\leq i\leq n} |\varphi_i(x)| + \delta\,\|x\|,$$

de donde

$$(1-\delta)\,\|x\| \leq \max_{1\leq i\leq n} |\varphi_i(x)| \leq \|x\|.$$

Por lo que si consideramos $T : F \to l_\infty^n$ dado por $T(x) = (\varphi_i(x))_{1\leq i\leq n}$, entonces $T$ es inyectiva, y si $\widetilde{F} = T(F)$, tenemos que $d_{BM}(\widetilde{F}, F) \leq 1/(1-\delta)$. □

**Observación 2.5.14.** La anterior proposición tiene varias consecuencias:

(I) Si $F$ es un espacio de dimensión finita $k$, entonces dado $\varepsilon > 0$ existe un subespacio $k$-dimensional $\widetilde{F} \subseteq c_0$, tal que $d_{BM}(F, \widetilde{F}) < 1 + \varepsilon$.

(II) Si $F = l_2^k$ y $0 < \varepsilon \leq 1$, tomando $n = [(1+2/\delta)^k]$ y $0 < \delta < 1$ tal que $(1-\delta)^{-1} \leq 1+\varepsilon$ (más precisamente, podemos tomar $\delta = \varepsilon/2 \leq \varepsilon/(1+\varepsilon)$, siendo esta última cantidad el valor exacto para el cual $(1-\delta)^{-1} = 1+\varepsilon$), tenemos que existe $\widetilde{F} \subseteq l_\infty^n$ con $d_{BM}(\widetilde{F}, l_2^k) \leq 1+\varepsilon$.

A partir de esta última observación podemos demostrar la Proposición 2.5.2.

*Demostración de la Proposición 2.5.2.* Fijemos $n \in \mathbb{N}$ y $0 < \varepsilon \leq 1$. Sea $k \in \mathbb{N}$ el único natural tal que

$$[(1+4/\varepsilon)^k] \leq n < [(1+4/\varepsilon)^{k+1}].$$

Entonces, por el punto (II) observado anteriormente, podemos encontrar un subespacio $(1+\varepsilon)$-isomorfo a $l_2^k$ en $l_\infty^n$. Luego

$$\ln(n) < (k+1)\ln(1+4/\varepsilon) \leq 2k\ln(1+4/\varepsilon).$$

De donde

$$k(l_\infty^n, \varepsilon) \geq k > \frac{1}{2\ln(1+4/\varepsilon)}\ln(n). \quad \square$$

# Comentario

Mencionemos por último una forma interesante, aunque quizás poco útil (al menos en el presente trabajo), de "leer" el teorema de Dvoretzky, a través del compacto de Banach-Mazur, que introducimos a continuación.

En los preliminares definimos la distancia de Banach-Mazur entre dos espacios normados isomorfos entre sí, $E$ y $F$, como

$$d_{BM}(E,F) = \text{ínf}\left\{\|T\|\,\left\|T^{-1}\right\|,\ T : E \to F \text{ isomorfismo}\right\}.$$

No es difícil ver que la distancia de Banach-Mazur satisface lo siguiente: Dados $E, F$ y $G$ espacios normados de dimensión $n$, vale que

- $d_{BM}(E,F) = d_{BM}(F,E)$.
- $d_{BM}(E,G) \leq d_{BM}(E,F)\, d_{BM}(F,G)$.
- $d_{BM}(E,F) = 1$ si y solo si $E$ y $F$ son isométricos.

Como cada espacio normado $E$ de dimensión $n$ es isométrico a $\mathbb{R}^n$ con alguna norma adecuada, podemos considerar únicamente distintas normas en $\mathbb{R}^n$, fijando así el espacio vectorial subyacente. Entonces, los tres ítems mencionados anteriormente nos dicen que el logaritmo natural (o el logaritmo en cualquier base mayor a 1) de la distancia de Banach-Mazur define una pseudo-métrica en el conjunto de normas en $\mathbb{R}^n$ (o, mejor dicho, en el conjunto de espacios normados que tienen como espacio vectorial subyacente a $\mathbb{R}^n$). Si en este conjunto tomamos la relación de equivalencia dada por

$$X \sim Y \iff X \text{ es isométrico a } Y,$$

entonces puede definirse (con cierto abuso de notación) $d_{BM}([X],[Y]) = d_{BM}(X,Y)$. Así obtenemos un espacio métrico, $(\mathcal{Q}(n), \ln \circ d_{BM})$, llamado el compacto de Banach-Mazur. Como lo indica su nombre, este espacio métrico resulta compacto. Los puntos de $\mathcal{Q}(n)$ son clases de isometría de espacios normados de dimensión $n$, y la distancia entre dos puntos es el logaritmo natural de la distancia de Banach-Mazur entre cualesquiera de sus respectivos representantes (este numero nos indica *que tan no isométricos* son los espacios entre sí). Este espacio a sido objeto de estudio por distintos autores, y en los últimos 25 años se han realizado importantes avances en el entendimiento de su topología (ver [SBR02]).

Así, una primera lectura del teorema de Dvoretzky en términos del compacto de Banach-Mazur es la siguiente.

*Dados $\varepsilon > 0$ y $k \in \mathbb{N}$, existe $n(k,\varepsilon) \in \mathbb{N}$ tal que para todo $n \geq n(k,\varepsilon)$ se tiene que si $E$ es un espacio de dimensión $n$, entonces existe algún subespacio $k$-dimensional $L \subseteq E$ tal que la clase de isometría de $L$ esta en la bola centrada en $[l_2^k]$ de radio $\varepsilon$. Es decir, tal que $[L] \in B([l_2^k],\varepsilon) \subseteq \mathcal{Q}(k)$.*

Ahora, si notamos

$$Gr(k,E) = \{L \subseteq E,\ L \text{ subespacio } k\text{-dimensional}\}$$

a la $k$-grassmaniana de $E$ (que resulta ser una variedad diferenciable compacta y de dimensión $k(n-k)$, siendo $n = \dim(E)$), podemos pensar en la *traza*, o en la imagen, de $Gr(k,E)$ en $\mathcal{Q}(k)$ mediante la aplicación

$$\begin{array}{ccc} Gr(k,E) & \rightarrow & \mathcal{Q}(k) \\ L & \rightarrow & [L]. \end{array}$$

Si notamos $\mathcal{Q}(k,E)$ a la imagen de esta aplicación (al conjunto de los puntos en $\mathcal{Q}(k)$ asociados a todos los subespacios $k$–dimensionales de $E$), entonces tenemos que $\mathcal{Q}(k,E)$ interseca a la bola $B([l_2^k],\varepsilon)$ si $\dim(E) \geq n(k,\varepsilon)$.

Aún más, dado un espacio de Banach de dimensión infinita $X$, podemos considerar igualmente al conjunto $Gr(k,X)$ (que en tal caso ya no será una variedad diferenciable compacta y de dimensión finita) y su *traza* $\mathcal{Q}(k,X)$ sobre $\mathcal{Q}(k)$. Así, lo que nos dice el Teorema 2 enunciado en la Introducción, es que $[l_2^k]$ está en la clausura de $\mathcal{Q}(k,X) \subseteq \mathcal{Q}(k)$, para cada $k \in \mathbb{N}$.

# Capítulo 3

# Bajando un $\varepsilon$

## 3.1. Menú y entrada

En este capítulo veremos que $k(n,\varepsilon) \geq \dfrac{c_1\varepsilon}{|\ln(c_2\varepsilon)|}\ln(n)$, siguiendo el trabajo de Grigoris Paouris y Petros Valettas [PV18]. Esto mejora en un $\varepsilon$ la estimación asintótica para $k(n,\varepsilon)$ obtenida en el capítulo anterior. Aún así, podríamos decir que ese no es el resultado más importante del trabajo que seguiremos.

Para explicar un poco esto, recordemos que para la demostración del teorema de Dvoretzky dada en el capítulo anterior, fue clave el uso de la desigualdad del Teorema 2.2.11. Esto es, dada $f : \mathbb{R}^n \to \mathbb{R}$ Lipschitz con $|f(x)-f(y)| \leq b\left\|x-y\right\|_2$ se tiene que

$$\gamma_n\Big(\left|f-\mathbb{E}_{\gamma_n}(f)\right| > t\Big) \leq 2\exp\left(-\frac{2}{\pi^2}\frac{t^2}{b^2}\right).$$

Dada una norma $\|\ \|$ en $\mathbb{R}^n$, usando esta desigualdad para la función $f(x) = \|x\|$, se tiene la siguiente proposición. [1]

**Proposición 3.1.1.** *Sea $E = (\mathbb{R}^n, \|\ \|)$ un espacio normado con $b > 0$ tal que $\|x\| \leq b\|x\|_2$ para todo $x \in \mathbb{R}^n$. Entonces para cada $t > 0$ vale que*

$$\gamma_n\Big(\left|\|x\| - \mathbb{E}_{\gamma_n}(\|x\|)\right| > t\,\mathbb{E}_{\gamma_n}(\|x\|)\Big) \leq 2\exp\left(-\frac{2}{\pi^2}t^2 k(E)\right),$$

*donde* $k(E) = \dfrac{\left(\mathbb{E}_{\gamma_n}(\|x\|)\right)^2}{b^2}$.

Notemos que el $k(E)$ definido en la proposición anterior satisface $k(E) \leq \delta(E)$, con $\delta(E)$ como en la Definición 2.5.8.

Podríamos decir que en la demostración del teorema de Dvoretzky del capítulo anterior, dado un espacio normado $E = (\mathbb{R}^n, \|\ \|)$, buscamos en $E$ un subespacio $F$ con "$k(F)$ grande", o más precisamente, buscamos vectores $z_1, \ldots, z_r \in E$ tales que si $b$ es la constante

[1] Notemos que si $b$ es la constante Lipschitz con respecto a la norma euclídea de $f(x) = \|x\|$, entonces $b$ coincide con la norma de la identidad formal $\left\|Id : l_2^n \to (\mathbb{R}^n, \|\ \|)\right\|$. Es decir $\|x\| \leq b\,\|x\|_2$.

Lipschitz de $f : l_2^r \to \mathbb{R}$, $f(x) = \left\| \sum_{i=1}^{r} x_i z_i \right\|$, entonces $\frac{\left(\mathbb{E}_{\gamma_r}(f)\right)^2}{b^2}$ "sea grande". Nuevamente, notemos que $\frac{\left(\mathbb{E}_{\gamma_r}(f)\right)^2}{b^2} \leq \delta(E)$. Esta búsqueda es porque, como fue mencionado en la Observación 2.5.9, un $\delta(E)$ "grande" nos permite encontrar subespacios casi euclídeos en $E$ de dimensión "grande", mediante el uso de la inecuación de concentración de la Proposición 3.1.1.

Y en efecto, asumiendo que $n = \dim(E)$ es suficientemente grande, en dicha demostración encontramos un subespacio (muchos en verdad) $F =< z_1, \ldots, z_r >$, con $r = \dim(F) \geq C\sqrt{n}$, en el que se tiene $\frac{(\mathbb{E}_{\gamma_r}(f))^2}{b^2} \geq C\ln(r) \geq \widetilde{C}\ln(n)$, siendo este el contenido de la Observación 2.5.9 (II)[2].

Ahora bien, en verdad puede verse (...y veremos en la Proposición 3.1.6) que dado un espacio normado $E = (\mathbb{R}^n, \|\ \|_E)$, existe un isomorfismo $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que si notamos $\widetilde{E}$ al espacio isométrico a $E$ cuya norma esta dada por $\|x\|_{\widetilde{E}} = \|T(x)\|_E$, entonces $k\left(\widetilde{E}\right) \geq C\ln(n)$, donde $C > 0$ es una constante absoluta. A partir de esto y la Proposición 3.1.1, se obtiene la siguiente proposición.

**Proposición 3.1.2.** *Sea $E = (\mathbb{R}^n, \|\ \|)$ un espacio normado. Entonces existe un isomorfismo $T : \mathbb{R}^n \to \mathbb{R}^n$ de manera que se cumple*

$$\gamma_n\left( \left| \|T(x)\| - \mathbb{E}_{\gamma_n}(\|T(x)\|) \right| > t\,\mathbb{E}_{\gamma_n}(\|T(x)\|) \right) \leq 2\exp\left(-ct^2\ln(n)\right)$$

*para todo $t > 0$, donde $c > 0$ es una constante absoluta.*

El isomorfismo $T$ de la proposición anterior no es más que alguno que ponga a $B_E$ en posición de John, como veremos más adelante.

Si bien no fue el enfoque dado en el capítulo anterior, podríamos pensar que la demostración del teorema de Dvoretzky allí dada está, sino basada, al menos ligada a esta desigualdad. Para explicitar esto, y a riesgo de ser densos o demasiado repetitivos, demostremos el teorema de Dvoretzky explícitamente basándonos en la misma.

En efecto veremos que una gran parte de los argumentos se repiten tal cual según lo realizado en la Sección 2.4, en particular el uso de los lemas geométricos de la Sección 2.3.

*Demostración del Teorema de Dvoretzky usando la Proposición 3.1.2.* Tenemos dados un espacio normado $E = (\mathbb{R}^n, \|\ \|)$ y un $\varepsilon > 0$. En este caso tomaremos $1 \leq k \leq n$ a determinar más adelante. Buscamos la existencia de un subespacio $S \subseteq E$, con $k = \dim(S)$, y $S$ $(1+\varepsilon)$-isomorfo a $l_2^k$.

Sean $T : \mathbb{R}^n \to \mathbb{R}^n$ el isomorfismo lineal de la Proposición 3.1.2, y para cada $1 \leq i \leq n$, $u_i = T(e_i)$, donde $(e_i)_{1 \leq i \leq n}$ es la base canónica de $\mathbb{R}^n$. Nuevamente, nos munimos de un espacio de probabilidad $(\Omega, \Sigma, \mu)$ junto con variables aleatorias gaussianas estándar independientes $g_{ij} : \Omega \to \mathbb{R}$, donde $1 \leq i \leq k$ y $1 \leq j \leq n$. Para cada $1 \leq i \leq k$ consideramos

[2] Durante esta discusión más cualitativa, notamos indistintamente como $C$ (o similares) a constantes absolutas, que en cada inecuación pueden representar números distintos.

ahora las variables aleatorias $\mathbb{R}^n$-valuadas (o simplemente, los vectores aleatorios) $X_i : \Omega \to \mathbb{R}^n$, dados por

$$X_i = \sum_{j=1}^{n} g_{ij}\, u_j.$$

De idéntica manera a lo realizado en la Sección 2.4, tomemos ahora $\delta(\varepsilon) = \varepsilon/8$ como en el Lema 2.3.2, y una $\delta(\varepsilon)$-red $A \subseteq \partial B_2^k$ (con respecto a la norma euclídea en $\mathbb{R}^k$) como en el Lema 2.3.1, de manera que $|A| \le (1 + 2/\delta)^k$. Para cada $\alpha = (\alpha_1, \dots, \alpha_k) \in A$ consideramos el vector aleatorio $G_\alpha : \Omega \to \mathbb{R}^n$

$$G_\alpha = \sum_{i=1}^{k} \alpha_i\, X_i.$$

Por el Lema 2.2.6, los vectores aleatorios $X_i$, para $1 \le i \le k$ son igualmente distribuidos e independientes, mientras que los vectores aleatorios $G_\alpha$, con $\alpha \in A$, son igualmente distribuidos a cualquiera de los $X_i$, $1 \le i \le k$. A su vez, notemos que los $X_i$ tienen la misma distribución que $T : (\mathbb{R}^n, \gamma_n) \to \mathbb{R}^n$, ya que $T(x) = \sum_{j=1}^n x_j u_j$.

Nuevamente, nos basta con ver que para algún $M > 0$

$$\mu\Big(\omega \in \Omega / \ \forall\, \alpha \in A \ \left| \frac{1}{M} \|G_\alpha(\omega)\| - 1 \right| \le \delta(\varepsilon)\Big) > 0,$$

para entonces poder asegurar, según el Lema 2.3.2, que existe $\omega \in \Omega$ para el que $S = \langle X_1(\omega), \dots, X_k(\omega)\rangle$ es $(1+\varepsilon)$-isomorfo a $l_2^k$.

Para eso, vemos nuevamente que

$$\mu\left(\omega \in \Omega / \ \exists\, \alpha \in A \text{ con } \left| \frac{1}{\mathbb{E}(\|X_1\|)} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| > \delta(\varepsilon)\right) < 1.$$

Y ahora sí, por la Proposición 3.1.2, podemos escribir

$$\begin{aligned}
&\mu\left(\omega \in \Omega / \ \exists\, \alpha \in A \text{ con } \left| \frac{1}{\mathbb{E}(\|X_1\|)} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| > \delta(\varepsilon)\right) \\
\le& \sum_{\alpha \in A} \mu\left(\omega \in \Omega / \ \left| \frac{1}{\mathbb{E}(\|X_1\|)} \left\| \sum_{i=1}^{k} \alpha_i\, X_i(\omega) \right\| - 1 \right| > \delta(\varepsilon)\right) \\
=& |A|\ \gamma_n\Big( \big| \|T(x)\| - \mathbb{E}_{\gamma_n}(\|T(x)\|) \big| > \delta(\varepsilon)\, \mathbb{E}_{\gamma_n}(\|T(x)\|) \Big) \\
\le& (1 + 2/\delta(\varepsilon))^k\, 2 \exp\big(-c\delta^2(\varepsilon) \ln(n)\big) \\
=& \exp\big(k \ln(1 + 16/\varepsilon) + \ln(2) - c_1 \varepsilon^2 \ln(n)\big) < 1
\end{aligned}$$

tomando $k < \dfrac{c_1 \varepsilon^2 \ln(n) - \ln(2)}{\ln(1 + 16/\varepsilon)}$. □

De aquí se ve que el orden para la dimensión $k$ del subespacio $(1+\varepsilon)$-isomorfo obtenido es $k \sim \dfrac{\varepsilon^2}{|\ln(\varepsilon)|}\ln(n)$, siendo que el $\varepsilon^2$ proviene del orden cuadrático en $t$ en la Proposición 3.1.2, mientras que el $|\ln(\varepsilon)|$ proviene de la técnica utilizada en la demostración, que involucra una red cuyo cardinal esta acotado por $(1+16/\varepsilon))^k$.

Finalmente, podemos enunciar el resultado principal de [PV18], y por lo tanto de este mismo capítulo.

**Teorema 3.1.3.** *Sea $E = (\mathbb{R}^n, \|\ \|)$ un espacio normado. Entonces existe un isomorfismo $T : \mathbb{R}^n \to \mathbb{R}^n$ de manera que se cumple*

$$\gamma_n\Big(\big|\|T(x)\| - \mathbb{E}_{\gamma_n}(\|T(x)\|)\big| > t\,\mathbb{E}_{\gamma_n}(\|T(x)\|)\Big) \leq c_1 \exp\left(-c_2\, t \ln(n)\right)$$

*para todo $0 < t \leq 1$, donde $c_1, c_2 > 0$ son constantes absolutas.*

De lo realizado anteriormente, se ve que, usando el Teorema 3.1.3, obtendríamos en la demostración del teorema de Dvoretzky recién realizada un subespacio $S \subseteq (\mathbb{R}^n, \|\ \|)$ de dimensión $k$, $(1+\varepsilon)$-isomorfo a $l_2^k$, cuya dimensión se comportaría asintóticamente como $k \sim \frac{\varepsilon}{|\ln(\varepsilon)|}\ln(n)$, que es lo que enunciamos a continuación.

**Teorema 3.1.4.** *Existen $c_1, c_2 > 0$ tales que si $E = (\mathbb{R}^n, \|\ \|)$ es un espacio normado, dados $\varepsilon > 0$ y $k \leq c_1 \dfrac{\varepsilon}{|\ln(c_2\,\varepsilon)|}\ln(n)$, existe en $E$ un subespacio $S$ $k$-dimensional que es $(1+\varepsilon)$-isomorfo a $l_2^k$.*

Veamos ahora, como prometimos, el ingrediente necesario para la demostración de la Proposición 3.1.2. Para ello requeriremos de una ligera pero notable variación del lema de Dvoretzky-Rogers.

**Lema 3.1.5.** *Sea $E = (\mathbb{R}^n, \|\ \|)$ un espacio normado cuya bola $B_E$ esta en posición de John. Entonces existe una base ortonormal $\{u_1, \ldots, u_n\}$ de $l_2^n$ de manera que $\|u_j\| \geq 1/(2+\sqrt{2}) > 1/4$ para todo $1 \leq j \leq n$.*

*Demostración.* Comenzamos tomando la base ortonormal $z_1, \ldots, z_n$ de $l_2^n$ de la que nos provee el lema de Dvoretzky-Rogers. Sabemos entonces que $\|z_j\| \geq \sqrt{1 - \frac{j-1}{n}}$ para cada $1 \leq j \leq n$, según lo visto en la Demostración del Corolario 1.1.10.

Esto nos asegura que los vectores $z_1, \ldots, z_{[(n+1)/2]}$ cumplen con la condición $\|z_j\| \geq 1/\sqrt{2} > 1/(2+\sqrt{2})$.

Ahora bien, para los $i$ entre 1 y $[(n-1)/2]$, si $z_{n-i}$ tiene norma mayor a $1/(\sqrt{2}+2)$, tomamos $u_i = z_i$ y $u_{n-i} = z_{n-i}$. Por otro lado, si $z_{n-i}$ tiene norma menor a $1/(2+\sqrt{2})$, entonces *reemplazamos* el par $(z_i, z_{n-i})$ por el par $u_i = \dfrac{z_i + z_{n-i}}{\sqrt{2}}$ y $u_{n-i} = \dfrac{z_i - z_{n-i}}{\sqrt{2}}$. Notemos que $\{u_i, u_{n-i}\}$ es base ortonormal del espacio generado por $\{z_i, z_{n-i}\}$. Por otro lado, usando la desigualdad triangular, tenemos que

$$\min\left(\|u_i\|, \|u_{n-i}\|\right) \geq \frac{1}{\sqrt{2}}\left(\|z_i\| - \|z_{n-i}\|\right) \geq \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - \frac{1}{2+\sqrt{2}}\right) = \frac{1}{2+\sqrt{2}}.$$

Así, por construcción, tenemos que los vectores $u_1, \ldots, u_n$ cumplen lo enunciado. $\square$

Con el Lema 3.1.5 y recordando, según lo comentamos en la Observación 2.2.10, que existe $C > 0$ tal que $\mathbb{E}_{\gamma_n}(\|x\|_\infty) \geq C\sqrt{\ln(n)}$ para cada $n \geq 2$ (donde puede tomarse $C = 2/5$), demostremos ahora lo afirmado previamente a la Proposición 3.1.2.

**Proposición 3.1.6.** *Sea $E = (\mathbb{R}^n, \|\ \|_E)$ un espacio normado. Entonces existe un isomorfismo lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que si $\widetilde{E} = \left(\mathbb{R}^n, \|\ \|_{\widetilde{E}}\right)$ es el espacio isométrico a $E$ con norma $\|x\|_{\widetilde{E}} = \|T(x)\|_E$, entonces $k\left(\widetilde{E}\right) \geq \frac{1}{100}\ln(n)$.*

*Demostración.* Tomamos como isomorfismo $T : \mathbb{R}^n \to \mathbb{R}^n$ aquel que coloque al cuerpo convexo $B_E$ en su posición de John, es decir, de manera que $T(B_E) = B_{\widetilde{E}}$ esté en posición de John.

Así, según el lema anterior tenemos una base ortonormal $\{u_1, \ldots, u_n\}$ de $l_2^n$ que cumple $\left\|u_j\right\|_{\widetilde{E}} \geq 1/4$ para todo $1 \leq j \leq n$. Por otro lado, al estar $\widetilde{E}$ en posición de John tenemos que $\|x\|_{\widetilde{E}} \leq \|x\|_2$ para todo $x \in \mathbb{R}^n$, por lo que la constante Lipschitz de la función $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|_{\widetilde{E}}$ es $b = 1$.

Ahora, notando $e_1, \ldots, e_n$ a los vectores de la base canónica de $\mathbb{R}^n$, por la invariancia de la medida gaussiana por transformaciones ortogonales tenemos que

$$\begin{aligned}
&\mathbb{E}_{\gamma_n}\left(\|x\|_{\widetilde{E}}\right) = \int_{\mathbb{R}^n} \left\| \sum_{i=1}^n x_i e_i \right\|_{\widetilde{E}} d\gamma_n(x) = \int_{\mathbb{R}^n} \left\| \sum_{i=1}^n x_i u_i \right\|_{\widetilde{E}} d\gamma_n(x) \\
&= \int_{\mathbb{R}^n} \frac{1}{2} \left( \left\| \sum_{i=1}^{n-1} x_i u_i + x_n u_n \right\|_{\widetilde{E}} + \left\| \sum_{i=1}^{n-1} x_i u_i - x_n u_n \right\|_{\widetilde{E}} \right) d\gamma_n(x) \\
&\geq \int_{\mathbb{R}^n} \max\left( \left\| \sum_{i=1}^{n-1} x_i u_i \right\|_{\widetilde{E}}, \|x_n u_n\|_{\widetilde{E}} \right) d\gamma_n(x) \\
&= \int_{\mathbb{R}^n} \frac{1}{2} \left( \max\left( \left\| \sum_{i=1}^{n-2} x_i u_i + x_{n-1} u_{n-1} \right\|_{\widetilde{E}}, \|x_n u_n\|_{\widetilde{E}} \right) + \max\left( \left\| \sum_{i=1}^{n-2} x_i u_i - x_{n-1} u_{n-1} \right\|_{\widetilde{E}}, \|x_n u_n\|_{\widetilde{E}} \right) \right) d\gamma_n(x) \\
&\geq \int_{\mathbb{R}^n} \max\left( \frac{1}{2} \left( \left\| \sum_{i=1}^{n-2} x_i u_i + x_{n-1} u_{n-1} \right\|_{\widetilde{E}} + \left\| \sum_{i=1}^{n-2} x_i u_i - x_{n-1} u_{n-1} \right\|_{\widetilde{E}} \right), \|x_n u_n\|_{\widetilde{E}} \right) d\gamma_n(x) \\
&\geq \int_{\mathbb{R}^n} \max\left( \left\| \sum_{i=1}^{n-2} x_i u_i \right\|_{\widetilde{E}}, \|x_{n-1} u_{n-1}\|_{\widetilde{E}}, \|x_n u_n\|_{\widetilde{E}} \right) d\gamma_n(x) \geq \ldots \\
&\geq \int_{\mathbb{R}^n} \max_{1 \leq i \leq n} \left( |x_i|\, \|u_i\|_{\widetilde{E}} \right) d\gamma_n(x) \geq \frac{1}{4} \int_{\mathbb{R}^n} \|x\|_\infty\, d\gamma_n(x) \geq \frac{1}{10}\sqrt{\ln(n)}.
\end{aligned}$$

Luego,

$$k\left(\widetilde{E}\right) = \frac{\left(\mathbb{E}_{\gamma_n}\left(\|x\|_{\widetilde{E}}\right)\right)^2}{b^2} \geq \frac{1}{100}\ln(n). \quad \square$$

Teniendo presente entonces que el resultado principal de [PV18] es el Teorema 3.1.3, mencionemos que allí se utilizan para demostrar este teorema dos resultados “fuertes”.

Uno es la desigualdad $L_1 - L_2$ de Talagrand para la medida gaussiana, y otro es un resultado de Alon y Milman en [AM83], aunque en verdad utilizaremos una versión mejorada de este último resultado, demostrada por Talagrand en [Tal95].

La segunda sección de este capítulo estará dedicada a hablar con más detalle de la desigualdad $L_1 - L_2$ de Talagrand. En la tercera sección de este capítulo hablaremos con más detalle de los trabajos de Alon-Milman [AM83] y Talagrand [Tal95]. Finalmente, en la cuarta sección de este capítulo discutiremos el camino llevado a cabo en [PV18] para demostrar el Teorema 3.1.3, dejando las demostraciones y detalles pendientes para los tres apéndices finales de este trabajo.

## 3.2. La desigualdad $L_1 - L_2$ de Talagrand

Esta desigualdad *nace* en el trabajo [Tal94] de Talagrand, en el área de la teoría de probabilidades, cuyo interés residía en dar ciertas aplicaciones a la teoría de percolación y grafos aleatorios.

Veamos un poco mejor el contexto en el que surge. Consideremos el $n$-cubo $\{0,1\}^n = \{(x_1,\dots,x_n) \,/\, x_i = \pm 1\}$. Se dice que un subconjunto $A \subseteq \{0,1\}^n$ es monótono si vale que para todo $x \in A$ se cumple que, si $y \in \{0,1\}^n$ satisface $x \leq y$ (es decir, $x_i \leq y_i$ para cada $1 \leq i \leq n$), entonces $y \in A$.

Ahora, para cada $0 \leq p \leq 1$, consideramos en $\{0,1\}^n$ la medida de probabilidad producto de las medidas $p$-Bernoulli en $\{0,1\}$. Es decir, para cada $x \in \{0,1\}^n$ es $\mu_p(\{x\}) = p^k(1-p)^{n-k}$, donde $k = |\{1 \leq i \leq n \,/\, x_i = 1\}|$. En tal caso se tiene que si $\emptyset \subsetneq A \subsetneq \{0,1\}^n$ es un subconjunto monótono, entonces la función $\mu_p(A)$ es creciente en $p$, con $\mu_0(A) = 0$ y $\mu_1(A) = 1$. Más aún, para ciertos conjuntos monótonos, se observa un comportamiento de "cambio de fase" o "efecto umbral", en el que la función $\mu_p(A)$ es casi constantemente 0 y casi constantemente 1 en ciertos intervalos, siendo que en un intervalo muy pequeño, centrado alrededor de algún punto o valor critico $p_*$ en el que $\mu_{p_*}(A) = 1/2$, la función $\mu_p(A)$ crece muy abruptamente.

En 1982 Russo [Rus82] muestra que esto ocurre para un conjunto monótono $A$, si este "no depende mucho de ninguna coordenada en particular", con esto queremos decir algo del siguiente estilo: Para cada $1 \leq i \leq n$ tenemos la función $U_i : \{0,1\}^n \to \{0,1\}^n$ dada por $(U_i(x))_j = x_j$ si $j \neq i$, y $(U_i(x))_i = 1 - x_i$. Es decir, $U_i(x)$ deja todas las coordenadas de $x$ intactas, excepto la $i$-ésima. Ahora, dado un conjunto monótono $A \subseteq \{0,1\}^n$ podemos definir para cada $1 \leq i \leq n$ las cantidades

$$I_i(p) = \mu_p\Big(x \in \{0,1\}^n \,/\, \mathbb{1}_A(x) \neq \mathbb{1}_A(U_i(x))\Big)^3,$$

entonces podemos pensar que $I_i(p)$ mide que tan "no invariante" por cambios en la $i$-coordenada es el conjunto $A$, o que tanto depende que un punto esté o no en $A$ de la $i$-ésima coordenada, respecto de $\mu_p$. Por lo que podríamos pensar que $A$ "no depende mucho de ninguna coordenada en particular" si las cantidades $I_i(p)$ son pequeñas.

En [Tal94], Talagrand simplifica y mejora los resultados obtenidos por Russo, utilizando en forma clave para esto una desigualdad, que demuestra en el mismo trabajo utilizando desarrollos de Fourier en el $n$-cubo.

[3] Notamos aquí con $\mathbb{1}_A$ a la función indicadora de $A$.

**Teorema 3.2.1** (Desigualdad $L_1 - L_2$ de Talagrand)**.** *Existe $C > 0$ tal que para toda $f : \{0,1\}^n \to \mathbb{R}$ se tiene*

$$\mathrm{Var}_{\mu_p}(f) \leq C \ln\left(\frac{2}{p(1-p)}\right) \sum_{i=1}^{n} \frac{\|\Delta_i f\|^2_{L_2(\mu_p)}}{1 + \ln\left(\frac{\|\Delta_i f\|_{L_2(\mu_p)}}{\|\Delta_i f\|_{L_1(\mu_p)}}\right)},$$

*donde $\Delta_i f = (1-p)(f - f \circ U_i)$ si $x_i = 1$ y $\Delta_i f = p(f - f \circ U_i)$ si $x_i = 0$.*

En particular, si consideramos la medida de probabilidad uniforme en $\{0,1\}^n$, que es la medida que nos queda al tomar $p = 1/2$, tenemos la siguiente proposición.

**Proposición 3.2.2.** *Existe $C > 0$ tal que para toda $f : \{0,1\}^n \to \mathbb{R}$ se tiene que*

$$\mathrm{Var}(f) \leq C \sum_{i=1}^{n} \frac{\|\Delta_i f\|_2^2}{1 + \ln\left(\frac{\|\Delta_i f\|_2}{\|\Delta_i f\|_1}\right)},$$

*donde $\Delta_i f = f - f \circ U_i$.*

Dado que en un espacio de probabilidad $(\Omega, \Sigma, \mu)$ las normas $p$ de una función $h : \Omega \to \mathbb{R}$ son crecientes en $p$ (es decir, $\|\Delta_i f\|_2 \geq \|\Delta_i f\|_1$), la desigualdad de la Proposición 3.2.2 puede verse como una mejora, salvo constantes, de la siguiente desigualdad (ver [BLM13, página 119]), válida para toda $f : \{0,1\}^n \to \mathbb{R}$, tomando nuevamente en $\{0,1\}^n$ la probabilidad uniforme

$$\mathrm{Var}(f) \leq \frac{1}{4} \sum_{i=1}^{n} \|\Delta_i f\|_2^2.$$

Ahora bien, la desigualdad anterior tiene una bien conocida y clásica versión continua en $\mathbb{R}^n$ con la medida de probabilidad gaussiana, llamada la desigualdad gaussiana de Poincaré.

**Teorema 3.2.3** (Desigualdad gaussiana de Poincaré)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función $\mathcal{C}^1$, tal que $f \in L^2(\gamma_n)$. Entonces*

$$\mathrm{Var}_{\gamma_n}(f) \leq \sum_{i=1}^{n} \|\partial_i f\|^2_{L^2(\gamma_n)} = \mathbb{E}_{\gamma_n}\left(\|\nabla f\|_2^2\right).$$

Uno podría preguntarse si en el caso de la medida gaussiana en $\mathbb{R}^n$, vale una mejora análoga a la de la desigualdad $L_1 - L_2$ de Talagrand en el hipercubo... y efectivamente, la hay, y es la desigualdad en la que estamos interesados.

Si bien, según indican Cordero-Erausquin y Ledoux en [CL12] , una versión continua de la desigualdad $L_1 - L_2$ de Talagrand para la medida gaussiana en $\mathbb{R}^n$ ha sido conocida como parte del "folklore probabilístico" (al menos desde 1999), ellos mismos dan una demostración del siguiente teorema en dicho trabajo (en verdad, en [CL12] se demuestra la validez de desigualdades tipo $L_1 - L_2$ de Talagrand para una clase mucho más amplia de espacios de probabilidad).

**Teorema 3.2.4.** . *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función $\mathcal{C}^1$ tal que $f \in L_2(\mathbb{R}^n, \gamma_n)$ y $\partial_i f \in L_2(\mathbb{R}^n, \gamma_n)$ para todo $1 \leq i \leq n$. Entonces*

$$\mathrm{Var}_{\gamma_n}(f) \leq C_T \sum_{i=1}^{n} \frac{\|\partial_i f\|_2^2}{1 + \ln\left(\frac{\|\partial_i f\|_2}{\|\partial_i f\|_1}\right)},$$

*donde $C_T > 0$ es una constante absoluta.*

En [Cha14, página 48] se da una demostración de esta última desigualdad a partir de la cual puede verse (si bien eso no se observa explícitamente en [Cha14]) que es posible tomar $C_T = \sqrt[3]{3}$.

Por último, la aplicación de la desigualdad $L_1 - L_2$ de Talagrand para la medida gaussiana en [PV18] no se da en la forma directa que tiene esta desigualdad en el Teorema 3.2.4, sino que se utiliza la siguiente consecuencia de la misma.

**Proposición 3.2.5.** *Sea $f \in \mathcal{C}^1$ tal que $f \in L_2(\mathbb{R}^n, \gamma_n)$ y $\partial_i f \in L_2(\mathbb{R}^n, \gamma_n)$ para todo $1 \leq i \leq n$. Entonces*

$$\mathrm{Var}_{\gamma_n}(f) \leq C_T \sum_{i=1}^{n} \frac{\|\partial_i f\|_{L_2(\gamma_n)}^2}{1 + \ln\left(\sqrt{R(f)}\right)}, \quad \text{donde } R(f) = \frac{\sum_{j=1}^{n} \left\|\partial_j f\right\|_{L_2(\gamma_n)}^2}{\sum_{j=1}^{n} \left\|\partial_j f\right\|_{L_1(\gamma_n)}^2},$$

*y $C_T > 0$ es la constante del Teorema 3.2.4 (en particular, podemos tomar $C_T = \sqrt[3]{3}$).*

Para una demostración de esta última desigualdad se puede ver [Cha14, Teorema 5.4]. Aclaremos que allí se demuestran dos desigualdades, y para poder ver una de esas dos desigualdades se asume que $f$ es monótona en cada coordenada. Sin embargo, la desigualdad de la Proposición 3.2.5 (que es en la que estamos interesados) no requiere de tal hipótesis.

## 3.3. Alon, Milman y subespacios no tan lejanos a $l_\infty^k$

En 1976, Maurey y Pisier demuestran en [MP76] que si $X$ es un espacio de Banach de dimensión infinita entonces para todo $k \in \mathbb{N}$ se pueden encontrar en $X$ subespacios de dimensión finita $k$ tan cercanos como se quiera a $l_\infty^k$ (con respecto a la distancia de Banach-Mazur), si y solo si $X$ no tiene cotipo finito[4].

Ahora bien... aclaremos que significa la noción de cotipo.

**Definición 3.3.1.** Se dice que $X$ tiene cotipo $q \in [2, +\infty)$ si existe $C > 0$ tal que para cualesquiera $n \in \mathbb{N}$ y $x_1, \ldots, x_n \in X$ se cumple que

$$\left(\sum_{i=1}^{n} \|x_i\|^q\right)^{1/q} \leq C \frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\|\sum_{i=1}^{n} \varepsilon_i x_i\right\| \quad \text{donde } \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n).$$ [5]

[4] En verdad, el trabajo de Maurey y Pisier contiene resultados aún más profundos sobre la teoría de espacios de Banach de dimensión infinita. Además de [MP76], ver [MS86].

[5] Esta definición se realiza solo para $q \in [2, +\infty)$ ya que puede verse que si $1 \leq \widetilde{q} < 2$ entonces ningún espacio normado no trivial puede cumplir esta propiedad para un tal $\widetilde{q}$.

El espacio $X$ tiene cotipo finito, si tiene cotipo $q$ para algun $q \in [2, +\infty)$.

Con respecto a esta definición, si consideramos en $\{-1,1\}^n$ la probabilidad uniforme y definimos $h\colon \{-1,1\}^n \to \mathbb{R}$ por

$$h(\varepsilon) = \left\| \sum_{i=1}^{n} \varepsilon_i x_i \right\| \quad \text{con } \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n),$$

entonces vemos que la expresión tras la constante $C$ en la Definición 3.3.1 es la esperanza de $h$, que notaremos $\mathbb{E}_\varepsilon(h)$.

**Observación 3.3.2.** Notemos que, teniendo en cuenta la Proposición 2.5.13, el resultado obtenido por Maurey y Pisier nos dice que si $X$ es un espacio de Banach de dimensión infinita que no tiene cotipo finito, entonces dado *cualquier* espacio normado $E$ de dimensión finita, existen en $X$ subespacios tan cercanos a $E$ como se quiera. Por ejemplo, como lo mencionamos en la Observación 2.5.14 (I), $c_0$ es un espacio con tales condiciones.

El resultado de Maurey y Pisier sin embargo deja pendiente, de alguna manera, la cuestión de bajo qué condiciones podemos encontrar *en un* espacio normado de dimensión $n$ subespacios cercanos a $l_\infty^k$ (o, equivalentemente, para qué valores de $k$ es posible encontrar estos subespacios). Esta pregunta fue la que motivó el trabajo de Alon y Milman [AM83] de 1983. Obviamente, no podemos esperar encontrar *en cualquier* espacio normado de dimensión $n$ subespacios de una dimensión “apreciable” $k$ (digamos, ¡mayor a 1!) cercanos a $l_\infty^k$, como nos lo muestra la inigualable familia de espacios euclídeos $l_2^n$. En efecto, todo subespacio de dimensión $k$ de $l_2^n$ es isométrico a $l_2^k$, y es sabido que $d_{BM}(l_2^k, l_\infty^k) = \sqrt{k}$.

Asentemos entonces los resultados que nos son de interés para el presente trabajo, obtenidos por Alon y Milman en [AM83]. Antes, aclaremos que dados $n$ vectores $x_1, \ldots, x_n$ en un espacio normado $X$, notaremos

$$M_n = M_n((x_i)_{1\le i\le n}) = \mathbb{E}_\varepsilon\left( \left\| \sum_{i=1}^{n} \varepsilon_i x_i \right\| \right). \tag{3.1}$$

**Teorema 3.3.3** (Teorema 4.1 de [AM83]). *Sea $X$ un espacio normado y sean $x_1, \ldots, x_n$ vectores unitarios en $X$. Entonces existe $A \subseteq \{1, \ldots, n\}$, con $k = |A| \ge \left[ \frac{\sqrt{n}}{2^7 M_n} \right]$, tal que para cualesquiera $(\alpha_i)_{i\in A}$ se tiene*

$$\frac{1}{2} \max_{i\in A} |\alpha_i| \le \left\| \sum_{i\in A} \alpha_i x_i \right\| \le 8 M_n \max_{i\in A} |\alpha_i|.$$

*En particular, el subespacio generado por $\{x_i,\, i \in A\}$ es $16M_n$-isomorfo a $l_\infty^k$.*

El Teorema 3.3.3 se obtiene en [AM83] a partir de una serie de proposiciones de naturaleza combinatoria, que omitimos por completo. De hecho, en un trabajo de 1995 ([Tal95]), Talagrand mejora el resultado enunciado en el Teorema 3.3.3, mediante una demostración relativamente más sencilla, que omite completamente las cuestiones combinatorias mencionadas anteriormente, y que hace uso de ciertas propiedades que satisfacen un conjunto de variables aleatorias independientes con distribución de Bernoulli, demostradas por él mismo en otro trabajo ([Tal93]). Esta versión del Teorema 3.3.3, que enunciamos a continuación, será la que utilizaremos más adelante.

**Teorema 3.3.4.** *Sean $X$ un espacio de Banach y $x_1,\dots,x_n \in X$ tales que $\|x_i\| \geq 1$ para cada $1 \leq i \leq n$. Notemos*

$$\tau = \tau\left((x_i)_{1\leq i\leq n}\right) = \|T : l_\infty^n \to X\|, \quad \text{con } T \text{ definido por } T(e_i) = x_i, \tag{3.2}$$

*donde $(e_i)_{1\leq i\leq n}$ es la base canónica de $\mathbb{R}^n$.*

*Entonces existe $A \subseteq \{1,\dots,n\}$ con $k = |A| \geq \dfrac{n}{32\tau}$ tal que para cualesquiera $(\alpha_i)_{i\in A}$ se tiene*

$$\frac{1}{2}\max_{i\in A}|\alpha_i| \leq \left\|\sum_{i\in A}\alpha_i x_i\right\| \leq 4M_n \max_{i\in A}|\alpha_i|.$$

*En particular, el subespacio generado por $\{x_i,\, i\in A\}$ es $8M_n$-isomorfo a $l_\infty^k$.*

Más allá de las mejoras en las constantes, el Teorema 3.3.4 es capaz de dar en ciertas situaciones una mejora en el orden sobre el cardinal del conjunto $A$. En efecto, esto se basa en la desigualdad $\dfrac{1}{\tau} \geq \dfrac{1}{\sqrt{2n}M_n}$. Para ver este hecho, utilizaremos la desigualdad de Khinchin, que enunciamos a continuación.

**Teorema 3.3.5** (Desigualdad de Khinchin, caso p=1)**.** *Para cualesquiera $n\in\mathbb{N}$ y $a_1,\dots,a_n\in\mathbb{R}$ vale que*

$$\left(\sum_{i=1}^n a_i^2\right)^{1/2} \leq \sqrt{2}\,\mathbb{E}_\varepsilon\left(\left|\sum_{i=1}^n \varepsilon_i a_i\right|\right)$$

Para (el enunciado completo y) la demostración de *las desigualdades* de Khinchin ver, por ejemplo, [DJT95, Capítulo 1].

Con esto, veamos la relación entre $\tau$ y $M_n$ afirmada más arriba.

**Proposición 3.3.6.** *Sean $X$ espacio normado y $x_1,\dots,x_n\in X$ tales que $\|x_i\|\geq 1$ para cada $1\leq i\leq n$. Entonces, si $M_n$ y $\tau$ son como en (3.1) y (3.2) respectivamente, se tiene que $\tau \leq \sqrt{2n}M_n$.*

*Demostración.* Comencemos notando que como $\tau = \|T : l_\infty^n \to X\|$, con $T$ definido como en (3.2), entonces $\tau = \left\|T^* : X^* \to l_1^n\right\|$ con $T^*$ dado por $T^*(\varphi) = (\varphi(x_i))_{1\leq i\leq n}$. Así,

$$\begin{aligned}\tau &= \sup_{\varphi\in B_{X^*}}\sum_{i=1}^n|\varphi(x_i)| \leq \sup_{\varphi\in B_{X^*}}\sqrt{n}\left(\sum_{i=1}^n\varphi(x_i)^2\right)^{1/2} \leq \sqrt{2n}\sup_{\varphi\in B_{X^*}}\mathbb{E}_\varepsilon\left(\left|\sum_{i=1}^n\varepsilon_i\varphi(x_i)\right|\right)\\ &=\sqrt{2n}\sup_{\varphi\in B_{X^*}}\mathbb{E}_\varepsilon\left(\left|\varphi\left(\sum_{i=1}^n\varepsilon_i x_i\right)\right|\right) \leq \sqrt{2n}\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^n\varepsilon_i x_i\right\|\right) = \sqrt{2n}\,M_n.\end{aligned}$$

□

Ahora bien, en [AM83], utilizando el Teorema 3.3.3, se demuestra el siguiente teorema, que si bien no nos será necesario en lo que sigue, resulta de interés en términos de la discusión dada al inicio de esta sección.

**Teorema 3.3.7.** *Existen constantes $c_1, c_2 > 0$ tales que cada todo $0 < \varepsilon < 1$, si $X$ es un espacio normado cualquiera, $n \in \mathbb{N}$ y $x_1, \ldots, x_n \in X$ son vectores unitarios, entonces $X$ contiene un subespacio de dimensión $k$ $(1+\varepsilon)$-isomorfo a $l_\infty^k$ para todo $k$ que satisfaga $k \leq c_1 \exp\left(\frac{c_2 \ln(n)\ln(1+\varepsilon)}{\ln(M_n)}\right)$.*

Esto nos dice que si queremos encontrar subespacios "de una dimensión apreciable" $k$ cercanos a $l_\infty^k$ en $X$, entonces debemos encontrar vectores unitarios $x_1, \ldots, x_n \in X$ para los cuales $\ln(M_n)$ sea mucho menor a $\ln(n)$. Para dar un ejemplo, si tomamos como espacio normado a $X = l_p^N$ (¡o incluso $l_p$!), con $1 \leq p < \infty$, entonces puede verse que dados cualesquiera vectores unitarios $x_1, \ldots, x_n \in X$ se tiene que $M_n \geq C_p\, n^{1/r}$, con $r = \text{mín}(2,p)$ y $C_p > 0$ una constante que solo depende de $p$ (como podría imaginarse por su notación...). De manera que en este caso tendríamos $\ln(n)/\ln(M_n) \leq \widetilde{C_p}/r$, que es una constante si $p$ esta fijo. Es decir, en este caso no lograremos encontrar (al menos no por el Teorema 3.3.7[6]) subespacios "de una dimensión apreciable" $k$ cercanos a $l_\infty^k$.

Por otro lado, como se observa en [AM83], si encontráramos $n$ vectores unitarios $x_1, \ldots, x_n$ en un espacio normado $X$ tales que $\ln(M_n) \leq \ln(n)^\delta$ para algún $0 < \delta < 1$, entonces podríamos encontrar en $X$ un subespacio de dimensión $k = \left[c_1 \exp\left(c_2 \ln^{1-\delta}(n)\ln(1+\varepsilon)\right)\right]$. Esta observación llevó a los autores al siguiente teorema, que plantea una interesante dicotomía para los subespacios de un espacio normado $X$.

**Teorema 3.3.8** (Teorema 4.3 de [AM83])**.** *Dado $\varepsilon > 0$, existe una constante $c(\varepsilon) > 0$ tal que si $X$ es un espacio normado de dimensión $n$,* con $n$ suficientemente grande, *y $k = \exp\left(c(\varepsilon)\sqrt{\ln(n)}\right)$, entonces o bien existe en $X$ un subespacio $F$ $k$-dimensional tal que $d_{BM}(F, l_2^k) \leq 1+\varepsilon$, o bien existe en en $X$ un subespacio $F$ $k$-dimensional tal que $d_{BM}(F, l_\infty^k) \leq 1+\varepsilon$.*

Para una demostración más detallada que la del trabajo original en la que se encuentra este resultado, ver [AGM21, Capítulo 6].

## 3.4. Divisando el camino de Paouris y Valettas

Como dijimos en la Sección 3.1, en esta sección describiremos el camino llevado a cabo en [PV18] para demostrar el Teorema 3.1.3. Para una mejor lectura, derivaremos la mayor parte de las demostraciones de los resultados que discutiremos a continuación (así como de los enunciados y demostraciones de los lemas utilizados para hacer tales demostraciones) a los apéndices finales de este trabajo.

Podemos identificar en [PV18] tres *pasos claves* que llevan finalmente a obtener el Teorema 3.1.3. Organizamos la presentación de manera que queden, en lo posible, lo más claramente distinguidos.

[6] Ni deberíamos poder encontrar por cualquier otro, dado que $l_p$ tiene cotipo finito para todo $p < \infty$.

### 3.4.1. Primer paso clave - Uso de la desigualdad $L_1 - L_2$ de Talagrand

Para comenzar, asentemos un poco de notación, así como algunas observaciones, que utilizaremos en lo que sigue.

Dada una función Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$, consideraremos las constantes de Lipschitz de $f$ respecto a la norma euclídea y la norma infinito. Notaremos $b = b(f)$ a la constante respecto a la norma euclídea, de manera que

$$|f(x) - f(y)| \leq b \left\|x - y\right\|_2 \quad \forall\, x, y \in \mathbb{R}^n,$$

y $a = a(f)$ a la constante Lipschitz de $f$ respecto de la norma infinito, de manera que

$$|f(x) - f(y)| \leq a \left\|x - y\right\|_\infty \quad \forall\, x, y \in \mathbb{R}^n.$$

Observemos que en tal caso, las inecuaciones

$$|f(x) - f(y)| \leq a(f) \left\|x - y\right\|_\infty \leq a(f) \left\|x - y\right\|_2 \quad \text{y} \quad |f(x) - f(y)| \leq b(f) \left\|x - y\right\|_2 \leq \sqrt{n} b(f) \left\|x - y\right\|_\infty,$$

nos dicen que

$$b(f) \leq a(f) \leq \sqrt{n}\, b(f). \tag{3.3}$$

Por otro lado, notemos que si $\| \ \|$ es una norma en $\mathbb{R}^n$, dados $f$ Lipschitz con constante $L$ respecto de dicha norma, y $p \in \mathbb{R}^n$ tal que $f$ es diferenciable en $p$, para todo $v \in \mathbb{R}^n$ vale que

$$|\langle \nabla f(p), v \rangle| = \lim_{t \to 0} \left| \frac{f(p + tv) - f(p)}{t} \right| \leq L \|v\|. \tag{3.4}$$

De donde vemos, según lo mencionado en la Sección Norma dual y cuerpo polar de los Preliminares, que $\|\nabla f(p)\|_* \leq L$.

En particular nos queda que

$$\|\nabla f(p)\|_1 \leq a(f) \quad \text{y} \quad \|\nabla f(p)\|_2 \leq b(f).$$

Pues bien, estamos listos para enunciar el primero de los tres pasos claves mencionados anteriormente, que es la Proposición 3.4.1. En su demostración, dada en el Apéndice A, se utiliza la desigualdad $L_1 - L_2$ de Talagrand (o más precisamente la Proposición 3.2.5), y será esta la única instancia en la que se haga uso de la misma. Por lo que podríamos decir que es aquí donde *inyectaremos* dicha desigualdad a lo que sigue de aquí en adelante. Notar como, en vistas del Teorema 3.1.3, que es a lo que buscamos llegar en forma última en este capítulo, *nace aquí* un término lineal en $t$ en la cota (3.5) para la medida del conjunto en el que una función Lipschitz se desvía de su media en más de $t$.

**Proposición 3.4.1.** *Sean $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz $C^1$ y $A > 0$ tal que $\|\partial_i f\|_{L_1(\gamma_n)} \leq A$ para todo $1 \leq i \leq n$. Supongamos que $a(f)A < e^2 b(f)^2$, entonces*

1. *Si $\widetilde{f} = f - \mathbb{E}_{\gamma_n}(f)$ y $C_T$ es la constante de la desigualdad $L_1 - L_2$ de Talagrand, entonces se tiene que*

$$\operatorname{Var}_{\gamma_n}\left(e^{\lambda \widetilde{f}}\right) \leq \frac{C_T \lambda^2 b(f)^2}{1 + \ln\left(\dfrac{b(f)}{\sqrt{a(f)A}}\right)} \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right) \quad \textit{para cada } \lambda > 0.$$

2. *Para cada* $t > 0$

$$\gamma_n\left(\left|f - \mathbb{E}_{\gamma_n}(f)\right| > t\right) \leq 5\exp\left(-\sqrt{1+\ln\left(\frac{b(f)}{\sqrt{a(f)A}}\right)}\,\frac{t}{b}\right). \tag{3.5}$$

Considerando a la desigualdad (3.5) como nuestro punto de partida, veamos como continuamos a partir de allí.

### 3.4.2. Segundo paso clave - Concentración para ciertos espacios

El siguiente paso clave en [PV18] consiste en obtener el Teorema 3.1.3 para cierto tipo de espacios normados en particular. Para entender la idea de como se logra esto, notemos que en la Proposición 3.4.1 aparece la preocupante condición de que se satisfaga $a(f)A \leq e^2 b(f)^2$ como hipótesis. Veamos, a modo de observación, que esto en efecto ocurre si la función Lipschitz $f$ fuese invariante por permutaciones.

**Definición 3.4.2.** Decimos que una función $f : \mathbb{R}^m \to \mathbb{R}$ es invariante por permutaciones si para toda biyección $\sigma : \{1,\ldots,m\} \to \{1,\ldots,m\}$ se cumple que

$$f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(m)}\right) = f(x_1, x_2, \ldots, x_m).$$

**Proposición 3.4.3.** *Sea* $f : \mathbb{R}^m \to \mathbb{R}$ *Lipschitz e invariante por permutaciones. Entonces*

1. *Para cada* $1 \leq p \leq \infty$ *se tiene que* $\|\partial_i f\|_{L_p(\gamma_m)} = \left\|\partial_j f\right\|_{L_p(\gamma_m)}$ *si* $1 \leq i < j \leq m$.

2. $A = \dfrac{a(f)}{m}$ *satisface* $\max\limits_{1\leq i\leq m} \|\partial_i f\|_{L_1(\gamma_m)} \leq A$. *En particular* $\dfrac{b(f)}{\sqrt{a(f)A}} \geq 1$.

*Demostración.*

1. Dados $1 \leq i < j \leq m$, sea $\sigma$ la permutación de $\{1,\ldots,m\}$ que intercambia $i$ con $j$ y deja fijos todos los demás elementos de $\{1,\ldots,m\}$. Notemos $P_\sigma$ a la transformación ortogonal dada por $P_\sigma(x) = \left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(m)}\right)$.

   Si $f$ es diferenciable en un punto $z \in \mathbb{R}^m$, entonces

   $$\frac{\partial f}{\partial x_i}(z) = \frac{\partial (f \circ P_\sigma)}{\partial x_i}(z) = \nabla f(P_\sigma(z)).P_\sigma e_i = \nabla f(P_\sigma(z)).e_j = \frac{\partial f}{\partial x_j}(P_\sigma(z)). \tag{3.6}$$

   Al ser $f$ diferenciable en casi todo punto, (3.6) nos da la igualdad de las normas infinito de $\partial_i f$ y $\partial_j f$. A su vez, al ser $P_\sigma$ una transformación ortogonal, por la invariancia bajo tales transformaciones de la medida gaussiana tenemos de (3.6) que $\|\partial_i f\|_{L_p(\gamma_m)} = \left\|\partial_j f\right\|_{L_p(\gamma_m)}$ para todo $1 \leq p < \infty$.

2. Por el ítem anterior, basta con notar que, si $1 \leq j \leq m$,

   $$\left\|\partial_j f\right\|_{L_1(\gamma_m)} = \frac{1}{m}\sum_{j=1}^{m} \left\|\partial_j f\right\|_{L_1(\gamma_m)} = \frac{1}{m}\left\|\|\nabla f\|_1\right\|_{L_1(\gamma_m)} \leq \frac{1}{m} a(f).$$

Con lo que, si $A = a(f)/m$, recordando (3.3) nos queda

$$\frac{b(f)}{\sqrt{a(f)A}} \geq \sqrt{m}\,\frac{b(f)}{a(f)} \geq 1. \quad \square$$

A partir de lo visto en la proposición anterior, la siguiente observación, aunque sencilla, resulta *fundamental*.

**Observación 3.4.4.** Notemos que el ítem 2 de la Proposición 3.4.3 se cumple si $f$ es una función Lipschitz en $\mathbb{R}^m$ para la cual se cumple la tesis del ítem 1 en dicha proposición (más allá de que $f$ sea o no invariante por permutaciones).

Ahora bien, es claro que no toda función Lipschitz con la que trabajemos (normas en $\mathbb{R}^m$) será invariante por permutaciones... Sin embargo, como lo declara el importante lema que enunciamos a continuación, si $f$ es suave, podremos lograr "ajustarla" de manera que se satisfaga el ítem 1 de la Proposición 3.4.3.

Para enunciarlo, aclaremos que dado $\lambda \in \mathbb{S}^{m-1}$ notaremos también por $\Lambda = \Lambda_\lambda$ al operador diagonal definido por $\Lambda(x) = (\lambda_1 x_1, \ldots, \lambda_m x_m)$.

**Lema 3.4.5.** *Sean $f : \mathbb{R}^m \to \mathbb{R}$ una función Lipschitz, $\mathcal{C}^1$ salvo quizás en un subconjunto discreto $D \subseteq \mathbb{R}^m$, y $p \geq 1$. Entonces existe $\widetilde{\lambda} = (\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m)$ con $\big\|\widetilde{\lambda}\big\|_2 = 1$, tal que si $\widetilde{\Lambda} : \mathbb{R}^m \to \mathbb{R}^m$ es el operador diagonal dado por $\widetilde{\Lambda}(x) = (\widetilde{\lambda}_1 x_1, \ldots, \widetilde{\lambda}_m x_m)$, se tiene que*

$$\big\|\partial_i (f \circ \widetilde{\Lambda})\big\|_{L_p(\gamma_m)} = \big\|\partial_j (f \circ \widetilde{\Lambda})\big\|_{L_p(\gamma_m)} \qquad \text{para todo } 1 \leq i < j \leq m.$$

*Más aún, si $f$ no es constante en ningún subespacio de dimensión positiva generado por algún subconjunto de los vectores canónicos en $\mathbb{R}^m$, entonces se puede tomar $\widetilde{\lambda} \in \mathbb{S}^{m-1}$ de manera que $\widetilde{\lambda}_i > 0$ para todo $1 \leq i \leq m$.*

Interesantemente, en la demostración del anterior lema (que damos en el Apéndice B), se utiliza el Teorema de Borsuk-Ulam para conseguir la existencia del operador diagonal $\widetilde{\Lambda}$.

A partir del Lema 3.4.5, que nos permite "ajustar" una norma *suave* $\|\ \|$ en $\mathbb{R}^m$ de manera que se cumpla el ítem 1 de la Proposición 3.4.3, llegamos a obtener el Teorema 3.1.3 para "cierta clase de espacios". Más concretamente, obtenemos la siguiente proposición, que también demostraremos en el Apéndice B.

**Proposición 3.4.6.** *Sea $X = (\mathbb{R}^m, \|\ \|)$ un espacio normado. Entonces existe un isomorfismo lineal $T : \mathbb{R}^m \to \mathbb{R}^m$, tal que para cada $t > 0$*

$$\gamma_m\Big(\Big|\, \|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|) \Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \leq 5\exp\left(-K \ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right) t\right), \tag{3.7}$$

*con $K > 0$ constante absoluta (que podemos tomar como $K = 1/31$).*

Donde $\mathrm{dia}(X)$ es una constante dependiente del espacio normado $X = (\mathbb{R}^m, \| \, \|)$ que definiremos en el Apéndice B. Por ahora, mencionemos que $\mathrm{dia}(X)$ satisface que $1 \leq \mathrm{dia}(X) \leq \sqrt{\dim(X)}$, y que, por ejemplo, los espacios $X = l_p^m$ con $1 \leq p \leq \infty$ (y sus sumas directas) cumplen que $\mathrm{dia}(X) = 1$.

Así, la proposición 3.4.6 nos *da* el Teorema 3.1.3 para todos aquellos espacios con $\mathrm{dia}(X) \leq 2$ (por ejemplo), o incluso para aquellas familias de espacios para las cuales se cumple que $\mathrm{dia}(X) \leq \dim(X)^{1/2-\theta}$, para algún $0 < \theta < 1/2$ fijo.

Algo notable de la Proposición 3.4.6, que constituye nuestro segundo paso clave hacia el Teorema 3.1.3, es que para obtenerla se hará uso de una posición "no clásica", a diferencia de, por ejemplo, el uso de la posición de John en la Proposición 3.1.2 (ver Observación 1.1.6 (iv)). En efecto, Lema 3.4.5 mediante, pondremos a la bola unitaria de nuestro espacio normado en una posición cuya existencia estará dada por el teorema de Borsuk-Ulam, y que no necesariamente se encuentra caracterizada por alguna condición extremal (como ocurre en general con las posiciones más clásicas).

### 3.4.3. Tercer paso clave - Concentración para todos los espacios

Finalmente discutiremos aquí los dos últimos resultados, la Proposición 3.4.7 y el Teorema 3.4.8, relevantes en el *divisamiento* del camino seguido en [PV18] para demostrar el Teorema 3.1.3.

Podríamos decir que el último paso clave de este *camino* consiste en separar en casos (notar la similaridad con el Teorema 3.3.8 demostrado por Alon-Milman).

Por un lado, si el $k(X)$ de nuestro espacio normado (donde $k(X)$ es como en la Proposición 3.1.1) *es grande*, veremos en la demostración del Teorema 3.4.8 (dada en el Apéndice C) que no hay mucho por hacer para obtener el resultado buscado en este capítulo. Por otro lado, en caso de que nuestro espacio normado no posea un $k(X)$ grande, la Proposición 3.4.7 (que sigue a continuación) nos dirá que existe entonces en $X$ un subespacio de dimensión "apreciable" donde sí valdrá la desigualdad de concentración que estamos buscando (aquella que se enuncia en el Teorema 3.1.3). Para la demostración de la Proposición 3.4.7 (también dada en el Apéndice C) se utilizan la Proposición 3.4.6 enunciada en la sección anterior, junto con el Teorema 3.3.4, demostrado por Talagrand, que constituye una mejora sobre el Teorema 3.3.3 obtenido por Alon y Milman.

**Proposición 3.4.7.** *Sean $X = (\mathbb{R}^n, \| \, \|)$ un espacio normado tal que $B_X$ está en posición de John y $0 < \delta < 1/2$. Entonces al menos una de las siguientes condiciones se satisface:*

1. $k(X) \geq n^{1/2-\delta}$.

2. *Existen un subespacio $F \subseteq \mathbb{R}^n$ con $\dim(F) = m$ tal que $\sqrt{n}/128 \leq m \leq n/2$ y una base $\{v_1, \ldots, v_m\}$ de $F$ tal que si $(\Omega, \Sigma, \mu)$ es un espacio de probabilidad y $g_1, \ldots, g_m : \Omega \to \mathbb{R}$ son variables aleatorias gaussianas estándar e independientes, entonces la variable aleatoria gaussiana $F$-valuada $\Phi : \Omega \to F$ dada por*

$$\Phi(\omega) = \sum_{i=1}^{m} g_i(\omega) v_i$$

*satisface que*

$$\mu\Big(\big|\|\Phi\| - \mathbb{E}_\mu(\|\Phi\|)\big| > t\,\mathbb{E}_\mu(\|\Phi\|)\Big) \leq 6\exp(-C\delta t\ln(n)) \quad \forall 0 < t \leq 1,$$

*donde* $C > 0$ *es una constante absoluta, que podemos tomar como* $C = 1/62$.

Si bien el enunciado del Teorema 3.4.8 que sigue a continuación constituye el resultado principal de este capítulo, ya que es el Teorema 3.1.3 enunciado con las constantes concretas obtenidas durante la demostración, remarcamos que la importancia del mismo *dentro del camino* reside en su demostración. En efecto, allí se ve que, en caso de que el espacio $X$ no posea un $k(X)$ grande, podremos obtener la desigualdad de concentración buscada, valiéndonos de que dicha desigualdad es cierta en un subespacio de dimensión "apreciable" (que es lo que nos dice la Proposición 3.4.7).

**Teorema 3.4.8.** *Sea* $X = (\mathbb{R}^n, \|\ \|)$ *un espacio normado. Entonces existe un isomorfismo* $S: \mathbb{R}^n \to \mathbb{R}^n$ *tal que para cada* $0 < t \leq 1$ *se tiene que*

$$\gamma_n\Big(\big|\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|)\big| > t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\Big) \leq 14\exp\left(\frac{-1}{310}\,t\ln(n)\right).$$

Ahora, usando el Teorema 3.4.8 (en lugar de la Proposición 3.1.2) en la demostración del teorema de Dvoretzky dada en la página 50, llegamos a la siguiente versión del teorema de Dvoretzky.

**Teorema 3.4.9** (Teorema de Dvoretzky, versión [PV18]). *Sean* $X = (\mathbb{R}^n, \|\ \|)$ *un espacio normado y* $0 < \varepsilon < 1$. *Entonces, si*

$$k < \frac{\frac{\varepsilon}{2480}\ln(n) - \ln(14)}{\ln(1 + 16/\varepsilon)}$$

*existe en* $X$ *un subespacio* $F \subseteq X$ *de dimensión* $k$ *tal que* $d_{BM}(F, l_2^k) \leq 1 + \varepsilon$.

Una vez enunciado el Teorema 3.4.9, hagamos algunas observaciones y consideraciones finales acerca de este resultado.

**Observación 3.4.10.**

(I) Notemos que para que el anterior enunciado "nos diga algo" (en el sentido de que la cota superior para $k$ pueda ser un número positivo), debe ser $n \geq 14^{2480} \approx 2,5 \, . \, 10^{2842}$.

(II) A su vez, podemos expresar el Teorema 3.4.9 diciendo que basta con tomar

$$n(k,\varepsilon) > \exp\left(\frac{2480}{\varepsilon}\Big(k\ln(1 + 16/\varepsilon) + \ln(14)\Big)\right). \tag{3.8}$$

Esto representa una mejora sobre la cota obtenida para $n(k,\varepsilon)$ en (2.6) si $\varepsilon \lesssim 0,1$. Por ejemplo, si buscamos un subespacio de dimensión 100 a distancia 0,1 de $l_2^{100}$, en general encontraremos un tal subespacio en cualquier espacio normado de dimensión $n \geq 1,41 \, . \, 10^{5501352}$ (comparar con la cota obtenida en la Observación 2.4.2 del Capítulo 2).

Rescatando una expresión más sencilla de (3.8), nos queda que si $N(k,\varepsilon)$ es como en la página VIII de la Introducción, entonces

$$N(k,\varepsilon) \leq \exp\left(3720 \frac{k \ln(17/\varepsilon)}{\varepsilon}\right).$$

Equivalentemente, si $k(n,\varepsilon)$ esta definido como en la página x de la Introducción, entonces

$$k(n,\varepsilon) \geq \frac{\varepsilon}{3720 \ln(17/\varepsilon)} \ln(n).$$

# Apéndice A

# Uso de la desigualdad $L_1 - L_2$

En este Apéndice demostraremos la Proposición 3.4.1 de la Sección 3.4.1 Comencemos enunciado un lema que utilizaremos para la demostración de la Proposición 3.4.1.

**Lema A.0.1.** *Sea $\psi\colon (0,1) \to (0,+\infty)$ una función tal que para todo $0 < s < 1$ se tiene que*

$$(1-s^2)\psi(s) \leq \psi^2\left(\frac{s}{2}\right).$$

*Si además*

$$\lim_{s\to 0^+} \frac{\psi(s)-1}{s} = 0, \tag{A.1}$$

*entonces vale que*

$$\psi(s) \leq \frac{1}{(1-s^2)} \frac{1}{\left(1-\left(\frac{s}{2}\right)^2\right)^4} \quad \forall\, s \in (0,1).$$

La demostración del anterior lema puede darse siguiendo las mismas técnicas y pasos que en [BLM13, Lema 3.19], aplicando una pequeña variación en dicha cuenta.

Ahora si, demostremos la Proposición 3.4.1, que volvemos a enunciar

**Proposición 3.4.1.** *Sean $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz $\mathcal{C}^1$ y $A > 0$ tal que $\|\partial_i f\|_{L_1(\gamma_n)} \leq A$ para todo $1 \leq i \leq n$. Supongamos que $a(f)A < e^2 b(f)^2$, entonces*

1. *Si $\widetilde{f} = f - \mathbb{E}_{\gamma_n}(f)$ y $C_T$ es la constante de la desigualad $L_1 - L_2$ de Talagrand, entonces se tiene que*

$$\mathrm{Var}_{\gamma_n}\left(e^{\lambda \widetilde{f}}\right) \leq \frac{C_T \lambda^2 b(f)^2}{1+\ln\left(\frac{b(f)}{\sqrt{a(f)A}}\right)} \mathbb{E}_{\gamma_n}\left(e^{2\lambda\widetilde{f}}\right) \quad \textit{para cada } \lambda > 0.$$

2. *Para cada $t > 0$*

$$\gamma_n\left(\left|f - \mathbb{E}_{\gamma_n}(f)\right| > t\right) \leq 5\exp\left(-\sqrt{1+\ln\left(\frac{b(f)}{\sqrt{a(f)A}}\right)}\,\frac{t}{b}\right). \tag{3.5}$$

*Demostración.*

1. Fijemos $\lambda > 0$ y notemos $h(x) = e^{\lambda \widetilde{f}(x)}$. Entonces, al ser $f$ Lipschitz, es fácil ver que $h \in W^{2,1}(\gamma_n)$. En efecto, por un lado tenemos que $|\widetilde{f}(x)| \leq |\widetilde{f}(0)| + b\|x\|_2$, de donde $|h(x)| \leq C\, e^{b\,\lambda\,\|x\|_2} \in L_2(\gamma_n)$. Por otro lado, al ser

$$\nabla h = \lambda\, e^{\lambda \widetilde{f}}\, \nabla f$$

en cada punto en el que $f$ es diferenciable (lo cual ocurre en casi todo punto, por el teorema de Rademacher), resulta que

$$\|\nabla h(x)\|_2 \leq \lambda\, h(x)\, b(f) \tag{A.2}$$

en tales puntos, por lo que también tenemos $\|\nabla h(x)\|_2 \in L_2(\gamma_n)$.

Así, podemos aplicar la Proposición 3.2.5 para $h$, obteniendo

$$\mathrm{Var}_{\gamma_n}(h) \leq C_T \frac{\|\nabla h\|^2_{L_2(\gamma_n)}}{1 + \frac{1}{2}\ln\left(\dfrac{\sum_{j=1}^n \left\|\partial_j h\right\|^2_{L_2(\gamma_n)}}{\sum_{j=1}^n \left\|\partial_j h\right\|^2_{L_1(\gamma_n)}}\right)}. \tag{A.3}$$

Ahora bien, dado $r > 0$ fijo, observemos que la función

$$\xi(s) = \frac{s}{1 + \frac{1}{2}\ln\left(\dfrac{s}{r}\right)}$$

es creciente para $s > r\,e^{-1}$. En efecto, derivando encontramos que

$$\xi'(s) = \frac{1}{2}\,\frac{1 + \ln(s/r)}{\left(1 + \frac{1}{2}\ln(s/r)\right)^2} > 0 \quad \text{si } s/r > e^{-1}.$$

Como para cada $1 \leq j \leq n$ se tiene $\left\|\partial_j h\right\|^2_{L_2(\gamma_n)} \geq \left\|\partial_j h\right\|^2_{L_1(\gamma_n)}$, estamos bajo dichas condiciones en la expresión (A.3). Por lo que si aplicamos la cota que nos queda para $\|\nabla h\|^2_{L_2(\gamma_n)}$ tras elevar al cuadrado e integrar en (A.2), nos queda

$$\mathrm{Var}_{\gamma_n}(h) \leq C_T \frac{\lambda^2\, b(f)^2\, \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right)}{1 + \frac{1}{2}\ln\left(\dfrac{\lambda^2\, b(f)^2\, \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right)}{\sum_{j=1}^n \left\|\partial_j h\right\|^2_{L_1(\gamma_n)}}\right)}. \tag{A.4}$$

A partir de aquí, notemos que la expresión en (A.4) es creciente en $\sum_{j=1}^n \left\|\partial_j h\right\|^2_{L_1(\gamma_n)}$. Acotemos superiormente entonces esta cantidad.

Para $1 \leq j \leq n$, tenemos que $\partial_j h = \lambda e^{\lambda \widetilde{f}} \partial_j f$, por lo que, usando la desigualdad de Hölder

$$\begin{aligned}\left\|\partial_j h\right\|_{L_1(\gamma_n)}^2 &= \lambda^2 \mathbb{E}_{\gamma_n}\left(e^{\lambda \widetilde{f}} |\partial_j f|\right)^2 = \lambda^2 \mathbb{E}_{\gamma_n}\left(e^{\lambda \widetilde{f}} |\partial_j f|^{1/2} |\partial_j f|^{1/2}\right)^2 \\ &\leq \lambda^2 \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}} |\partial_j f|\right) \mathbb{E}_{\gamma_n}\left(|\partial_j f|\right) \leq \lambda^2 A \,\mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}} |\partial_j f|\right).\end{aligned}$$

Sumando sobre $j$ nos queda

$$\sum_{j=1}^{n} \left\|\partial_j h\right\|_{L_1(\gamma_n)}^2 \leq \lambda^2 A \,\mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}} \,\|\nabla f\|_1\right) \leq \lambda^2 A \,a(f) \,\mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right).$$

De esta manera, llegamos a que

$$\mathrm{Var}_{\gamma_n}(h) \leq C_T \frac{\lambda^2 b(f)^2 \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right)}{1 + \ln\left(\dfrac{b(f)}{\sqrt{a(f)A}}\right)}.$$

2. Vimos en el anterior ítem que

$$\mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right) - \mathbb{E}_{\gamma_n}\left(e^{\lambda \widetilde{f}}\right)^2 \leq C_T \frac{\lambda^2 b(f)^2 \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right)}{1 + \ln\left(\dfrac{b(f)}{\sqrt{a(f)A}}\right)} \quad \text{para cada } \lambda > 0. \tag{A.5}$$

Notemos como $\rho$ al número positivo tal que

$$\rho^2 = \frac{C_T \, b(f)^2}{1 + \ln\left(\dfrac{b(f)}{\sqrt{a(f)A}}\right)}.$$

De manera que (A.5) se exprese como

$$\mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right) - \mathbb{E}_{\gamma_n}\left(e^{\lambda \widetilde{f}}\right)^2 \leq \rho^2 \lambda^2 \mathbb{E}_{\gamma_n}\left(e^{2\lambda \widetilde{f}}\right).$$

Entonces, realizando el cambio de variables $s = \rho\,\lambda$, y considerando la función

$$\psi(s) = \mathbb{E}_{\gamma_n}\left(e^{2\frac{s}{\rho} \widetilde{f}}\right)$$

vemos que $\psi$ satisface la desigualdad funcional

$$(1 - s^2)\psi(s) \leq \psi^2\left(\frac{s}{2}\right) \quad \text{para } s \in (0,1) \text{ (ie. } \lambda \in (0, 1/\rho)).$$

Usando el Lema A.0.1 enunciado anteriormente, tendremos que entonces

$$\psi(s) \leq \frac{1}{(1-s^2)} \frac{1}{\left(1-\left(\frac{s}{2}\right)^2\right)^4} \quad \forall\, s \in (0,1). \tag{A.6}$$

Para esto, solo nos resta chequear que se cumpla sobre $\psi$ la condición

$$\lim_{s\to 0^+} \frac{\psi(s)-1}{s} = 0 \quad \text{requerida en el Lema A.0.1.}$$

Para ver esto, a su vez, chequeemos que se cumplen las condiciones de la Proposición 2.5.3, de manera que podamos escribir

$$\lim_{s\to 0^+} \frac{\psi(s)-1}{s} = \lim_{s\to 0^+} \frac{\psi(s)-\psi(0)}{s} = \mathbb{E}_{\gamma_n}\left(e^{2\frac{s}{\rho}\widetilde{f}}(2/\rho)\widetilde{f}\right)\Big|_{s=0} = \mathbb{E}_{\gamma_n}\left(\frac{2}{\rho}\widetilde{f}\right) = 0.$$

En efecto, sabemos que para todo $s \in [0,1)$ vale que $e^{2(s/\rho)\widetilde{f}} \in L_1(\gamma_n)$. También tenemos que para todo $x \in \mathbb{R}^n$ se cumple que $e^{2(s/\rho)\widetilde{f}}$ es derivable con respecto a $s$ (solo necesitaríamos tener que esto ocurre para casi todo $x \in \mathbb{R}^n$). Por último, si $s \in [0,1/2]$, usando que $u < e^u$ para todo $u \in \mathbb{R}$

$$\left|\frac{\partial\left(e^{2(s/\rho)\widetilde{f}}\right)}{\partial s}\right| = \left|e^{2(s/\rho)\widetilde{f}}(2/\rho)\widetilde{f}\right| \leq 2\,\frac{|\widetilde{f}|}{\rho}\,e^{\widetilde{f}/\rho} \leq 2e^{2\widetilde{f}/\rho} \in L_1(\gamma_n).$$

Así, podemos continuar a partir de (A.6).

Usando la desigualdad de Markov llegamos a que si $t > 0$, entonces

$$\gamma_n\left(\widetilde{f} > t\right) = \gamma_n\left(e^{2(s/\rho)\widetilde{f}} > e^{2(s/\rho)t}\right) \leq \frac{\mathbb{E}_{\gamma_n}\left(e^{2\frac{s}{\rho}\widetilde{f}}\right)}{e^{2(s/\rho)t}} \leq \frac{1}{(1-s^2)\left(1-\left(\frac{s}{2}\right)^2\right)^4}\, e^{-2(s/\rho)t}.$$

Tomando $C_T = 3/2$ (ver Teorema 3.2.4), y evaluando esta última expresión en $s = \frac{1}{2}\sqrt{\frac{3}{2}}$ llegamos a que

$$\mathbb{E}_{\gamma_n}\left(f - \mathbb{E}_{\gamma_n}(f) > t\right) \leq \frac{5}{2}\exp\left(-\sqrt{1+\ln\left(\frac{b(f)}{\sqrt{a(f)A}}\right)}\,\frac{t}{b}\right).$$

Aplicando esto mismo para $-f$, y juntando ambos resultados, obtenemos

$$\mathbb{E}_{\gamma_n}\left(\left|f - \mathbb{E}_{\gamma_n}(f)\right| > t\right) \leq 5\exp\left(-\sqrt{1+\ln\left(\frac{b(f)}{\sqrt{a(f)A}}\right)}\,\frac{t}{b}\right),$$

que es lo que queríamos demostrar. □

**Observación A.0.2.**

(I) En el enunciado de la Proposición 3.4.1, notemos que $a(f)$ satisface la condición requerida sobre el número $A > 0$, ya que

$$\mathbb{E}_{\gamma_n}(|\partial_i f|) \leq \mathbb{E}_{\gamma_n}(\|\nabla f\|_1) \leq a(f) \quad \text{para todo } 1 \leq i \leq n.$$

Sin embargo, de utilizar $a(f)$ en forma directa en lugar de una cota más cuidadosa para $\max_{1\leq i\leq n} \|\partial_i f\|_{L_1(\gamma_n)}$, nos quedaría en el resultado la expresión

$$\sqrt{1 + \ln\left(\frac{b(f)}{a(f)}\right)},$$

que en vistas de lo observado en (3.3), no es muy alentadora.

(II) Comparando con la demostración realizada en el Capítulo 2 del Teorema 2.2.11, puede verse una similaridad en la técnica empleada en ambos casos, esto es, la acotación (obtenida de forma diferente en cada caso, eso sí) de la función $M(\lambda) = \mathbb{E}_{\gamma_n}\left(e^{\lambda(f - \mathbb{E}_{\gamma_n}(f))}\right)$[1], para luego, mediante la desigualdad de Markov, obtener una cota sobre $\mathbb{E}_{\gamma_n}\left(f - \mathbb{E}_{\gamma_n}(f) > t\right)$ escribiendo

$$\mathbb{E}_{\gamma_n}\left(f - \mathbb{E}_{\gamma_n}(f) > t\right) \leq M(\lambda)e^{-\lambda t}.$$

A su vez, una diferencia intrigante surge entre estas dos demostraciones. En la demostración del Teorema 2.2.11, buscamos el $\lambda_t > 0$ que minimizara (para cada $t > 0$) la expresión obtenida tras aplicar la desigualdad de Markov, $C(\lambda)e^{-\lambda t}$, donde $C(\lambda)$ representa la cota obtenida en aquella ocasión para $M(\lambda) = \mathbb{E}_{\gamma_n}\left(e^{\lambda(f - \mathbb{E}_{\gamma_n}(f))}\right)$. Esto, que resulta lo más natural y sensato, es conocido como el método de Cramér–Chernoff, el cual es muy fructífero en diversas situaciones (ver [BLM13]).

Ahora bien, por otro lado, en la demostración de la Proposición 3.4.1 no hicimos esto, sino que elegimos el valor de $s$ (ya que hicimos un cambio de variables de $\lambda$ por $s$ en esta demostración) de forma ciertamente arbitraria o caprichosa, buscando relativa sencillez en la expresión de la cota obtenida. ¿Por qué no buscar el $s$ que nos de la mejor cota? Pues bien, aún más curioso puede resultar el hecho de que si buscáramos el $s$ que minimice la expresión

$$\frac{4}{(1 - s^2)} e^{-2(s/\rho)t},$$

(donde por cuestiones cardíacas acotamos previamente $1/(1 - s^2/4)^4$ por 4 en $(0,1)$) encontraríamos a la siguiente expresión como la mejor cota superior

$$2\left(1 + \sqrt{1 + \theta}\right)\exp\left(-\frac{\theta}{1 + \sqrt{1 + \theta}}\right), \quad \text{con } \theta = \frac{-4}{C_T}\frac{t^2}{b^2}\left(1 + \ln\left(\frac{b}{\sqrt{Aa}}\right)\right). \tag{A.7}$$

[1]En lenguaje probabilístico, $M(\lambda)$ es la función generadora de momentos de $\widetilde{f} = f - \mathbb{E}_{\gamma_n}(f)$.

Es decir, *para $t$ pequeño* ¡nos volvería a aparecer una dependencia de orden cuadrático en $t$ en el exponente!

Sin embargo, debe notarse que, para $t$ grande, el orden de $t$ en el exponente de (A.7) sería lineal. Pues bien, así como lo hicimos en el Capítulo 2, más adelante utilizaremos estas desigualdades aplicadas no a $t$ (con $0 < t < 1$), sino a $\mathbb{E}_{\gamma_n}(\|x\|)\,t$, donde $\| \ \|$ es la norma de un espacio normado $X = (\mathbb{R}^n, \| \ \|)$, con la *esperanza* de que $k(X) = \mathbb{E}_{\gamma_n}(\|x\|)^2/b^2$ sea grande. En tal caso, para $0 < t < 1$ fijo, podemos esperar que el orden en $t$ sea lineal en el exponente de (A.7), con lo cual el resultado obtenido en la Proposición 3.4.1 "en busca de una expresión más sencilla" coincidiría asintóticamente con el de la mejor cota posible que nos hubiese sido posible conseguir tras los argumentos dados en la demostración de esa proposición.

# Apéndice B

# Concentración para ciertos espacios

En este Apéndice veremos la demostración de la Proposición 3.4.6 de la Sección 3.4.2 (así como las definiciones y los lemas necesarios para ello). Mencionemos que de los tres pasos claves mencionados en la Sección 3.4, este es el paso que requiere más trabajo.

## B.1. $L_p(\gamma_m)$-invariancia por permutaciones

Comenzamos dando una demostración del Lema 3.4.5, que enunciamos nuevamente

**Lema 3.4.5.** *Sean $f : \mathbb{R}^m \to \mathbb{R}$ una función Lipschitz, $\mathcal{C}^1$ salvo quizás en un subconjunto discreto $D \subseteq \mathbb{R}^m$, y $p \geq 1$. Entonces existe $\widetilde{\lambda} = (\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m)$ con $\left\|\widetilde{\lambda}\right\|_2 = 1$, tal que si $\widetilde{\Lambda} : \mathbb{R}^m \to \mathbb{R}^m$ es el operador diagonal dado por $\widetilde{\Lambda}(x) = (\widetilde{\lambda}_1 x_1, \ldots, \widetilde{\lambda}_m x_m)$, se tiene que*

$$\left\|\partial_i (f \circ \widetilde{\Lambda})\right\|_{L_p(\gamma_m)} = \left\|\partial_j (f \circ \widetilde{\Lambda})\right\|_{L_p(\gamma_m)} \quad \text{para todo } 1 \leq i < j \leq m.$$

*Más aún, si $f$ no es constante en ningún subespacio de dimensión positiva generado por algún subconjunto de los vectores canónicos en $\mathbb{R}^m$, entonces se puede tomar $\widetilde{\lambda} \in \mathbb{S}^{m-1}$ de manera que $\widetilde{\lambda}_i > 0$ para todo $1 \leq i \leq m$.*

Previo a la demostración, recordemos el teorema de Borsuk-Ulam, el cual nos dice que si $h : \mathbb{S}^{m-1} \to \mathbb{R}^{m-1}$ es una función continua sobre la esfera $(m-1)$ dimensional, entonces existe $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{S}^{m-1}$ tal que $h(\lambda) = h(-\lambda)$.

*Demostración del Lema 3.4.5.* Para cada $1 \leq j < m$ sea $h_j : \mathbb{S}^{m-1} \to \mathbb{R}$ dada por

$$h_j(\lambda) = \lambda_j \left\|\partial_j f \circ \Lambda\right\|_{L_p(\gamma_m)} - \lambda_{j+1} \left\|\partial_{j+1} f \circ \Lambda\right\|_{L_p(\gamma_m)}.$$

Notemos que por la invariancia de la medida gaussiana bajo transformaciones ortogonales, en particular bajo la transformación $-Id : z \mapsto -z$, tenemos que $h_j(-\lambda) = -h_j(\lambda)$. Es decir, $h_j$ es una función impar para cada $1 \leq j \leq m-1$.

Veamos ahora que cada $h_j$ es continua. Para esto, nos basta con ver que si $g$ (que haría las veces de $|\partial_i f|^p$ para algún $1 \leq i \leq m$) es una función continua en $\mathbb{R}^m \setminus D$, con $D \subseteq \mathbb{R}^m$ discreto, y acotada en $\mathbb{R}^m$, entonces la función $\xi : \mathbb{S}^{m-1} \to \mathbb{R}$ dada por

$$\xi(\lambda) = \int_{\mathbb{R}^m} g(\lambda_1 x_1, \ldots, \lambda_m x_m)\, d\gamma_m(x)$$

es continua. Pero en efecto, si $\lambda^k \to \lambda$ en $\mathbb{S}^{m-1}$, entonces $g(\Lambda^k(x)) \to g(\Lambda(x))$ para cada $x \in \mathbb{R}^m \setminus D$ al ser $g$ continua allí, y como por otro lado tenemos que $|g(\lambda_1 x_1, \ldots, \lambda_m x_m)| \leq \|g\|_\infty \in L_1(\gamma_m)$, entonces por el teorema de convergencia dominada se cumple que $\xi(\lambda^k) \to \xi(\lambda)$.

De esta manera, la función $h : \mathbb{S}^{m-1} \to \mathbb{R}^{m-1}$ resulta continua, y por lo tanto, por el teorema de Borsuk-Ulam, tenemos que existe $\widetilde{\lambda} \in \mathbb{S}^{m-1}$ tal que $h(\widetilde{\lambda}) = h(-\widetilde{\lambda})$. Pero como $h$ es impar, esto implica que $h(\widetilde{\lambda}) = 0$. Es decir

$$\widetilde{\lambda}_j \left\|\partial_j f \circ \widetilde{\Lambda}\right\|_{L_p(\gamma_m)} = \widetilde{\lambda}_{j+1} \left\|\partial_{j+1} f \circ \widetilde{\Lambda}\right\|_{L_p(\gamma_m)} \tag{B.1}$$

para todo $1 \leq j \leq m-1$. Tomando módulo en esta última igualdad obtenemos que

$$\left\|\partial_j (f \circ \widetilde{\Lambda})\right\|_{L_p(\gamma_m)} = \left\|\partial_{j+1} (f \circ \widetilde{\Lambda})\right\|_{L_p(\gamma_m)}$$

para todo $1 \leq j \leq m-1$.

Con respecto a la afirmación realizada al final del enunciado, comencemos notando que bajo la hipótesis allí realizada el $\widetilde{\lambda}$ obtenido no puede tener ninguna coordenada nula.

En efecto, si notamos $J = \left\{1 \leq j \leq m \,/\, \widetilde{\lambda}_j \neq 0\right\}$, entonces al ser $\widetilde{\lambda} \in \mathbb{S}^{m-1}$ debe ser $J \neq \emptyset$. Ahora, si $\widetilde{\lambda}$ tuviera alguna coordenada nula (es decir, si fuese $J^c \neq \emptyset$), entonces para cada $j \in J$ tendríamos que $\left\|\partial_j f \circ \widetilde{\Lambda}\right\|_{L_p(\gamma_m)} = 0$. Es decir, si $V = \langle e_i \rangle_{i \in J}$, usando la continuidad de las derivadas parciales, tendríamos que $\partial_j f = 0$ en $V$, para cada $j \in J$. Es decir, tendríamos que $f$ es constante en un subespacio formado por un subconjunto de los vectores canónicos. Así bajo la hipótesis de que esto último no ocurre, debe ser $J^c = \emptyset$.

De esta manera, teniendo en cuenta nuevamente (B.1), todas las coordenadas de $\widetilde{\lambda}$ deben tener el mismo signo, que podemos tomar positivo (ya que $h(\widetilde{\lambda}) = h(-\widetilde{\lambda}) = 0$). □

Teniendo en cuenta la Observación 3.4.4 y el Lema 3.4.5, obtenemos el siguiente lema.

**Lema B.1.1.** *Sea $f : \mathbb{R}^m \to \mathbb{R}$ una función Lipschitz, $\mathcal{C}^1$ salvo quizás en un conjunto discreto de $\mathbb{R}^m$, tal que no existe ningún subconjunto no vacío $I \subseteq \{1, \ldots, m\}$ para el que se cumpla que $f$ es constante en $\langle e_i \rangle_{i \in I}$, siendo $e_i$ el i-ésimo vector de la base canónica de $\mathbb{R}^m$. Entonces existe $\lambda \in \mathbb{S}^{m-1}$ tal que $\lambda_i > 0$ para todo $1 \leq i \leq m$ y $\left\|\partial_j (f \circ \Lambda)\right\|_{L_1(\gamma_m)} \leq \frac{1}{m} a(f \circ \Lambda)$ para todo $1 \leq j \leq m$.*

## B.2. Constantes de incondicionalidad y divergencia incondicional aleatoria

La Proposición 3.4.6, que es la que buscamos demostrar en este Apéndice, esta formulada en términos de la constante de divergencia incondicional aleatoria de un espacio normado $X = (\mathbb{R}^m, \| \, \|)$. Esta constante es “una versión más débil” que la de constante de incondicionalidad de un espacio $X$, $\operatorname{inc}(X)$, siendo esta última una noción más clásica. Comencemos entonces definiendo la constante de incondicionalidad de un espacio normado de dimensión finita.

**Definición B.2.1.** Sea $X = (\mathbb{R}^m, \| \, \|)$ un espacio normado.

1. Dada una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ de $X$, definimos inc($\mathcal{B}$), la constante de incondicionalidad de la base $\mathcal{B}$, como
$$\mathrm{inc}(\mathcal{B}) = \max_{\varepsilon_i = \pm 1} \sup_{(\alpha_1, \ldots, \alpha_m) \neq 0} \frac{\left\| \sum_{i=1}^m \varepsilon_i \, \alpha_i \, b_i \right\|}{\left\| \sum_{i=1}^m \alpha_i \, b_i \right\|}.$$
Es decir, inc($\mathcal{B}$) es la menor constante tal que, para cualesquiera escalares $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ se tiene que
$$\max_{(\varepsilon_1, \ldots, \varepsilon_m) \in \{-1,1\}^m} \left\| \sum_{i=1}^m \varepsilon_i \, \alpha_i \, b_i \right\| \leq \mathrm{inc}(\mathcal{B}) \left\| \sum_{i=1}^m \alpha_i \, b_i \right\|.$$

2. Definimos inc($X$), la constante de incondicionalidad de $X$, como
$$\mathrm{inc}(X) = \inf_{\mathcal{B} \text{ base de } X} \mathrm{inc}(\mathcal{B}). \tag{B.2}$$

Notemos que dada una base cualquiera $\mathcal{B}$ de $X$, $\mathrm{inc}(\mathcal{B}) \geq 1$. Por lo que a su vez tenemos $\mathrm{inc}(X) \geq 1$.

Si bien debe ser algo conocido, en el breve tiempo en el que nos lo planteamos, no logramos encontrar en la literatura o vislumbrar por cuenta propia si el ínfimo en (B.2) se alcanza[1]. Es decir, si para cada norma $\| \ \|$ en $\mathbb{R}^m$ efectivamente existe una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ en $X$ tal que dados $\alpha_1, \ldots, \alpha_m$ escalares cualesquiera valga
$$\max_{(\varepsilon_1, \ldots, \varepsilon_m) \in \{-1,1\}^m} \left\| \sum_{i=1}^m \varepsilon_i \, \alpha_i \, b_i \right\| \leq \mathrm{inc}(X) \left\| \sum_{i=1}^m \alpha_i \, b_i \right\|.$$

Por otro lado, decimos que $X$ es 1-incondicional si existe una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ en $X = (\mathbb{R}^m, \| \ \|)$ tal que $\mathrm{inc}(\mathcal{B}) = 1$, de manera que dados cualesquiera escalares $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ y cualesquiera signos $\varepsilon_1, \ldots, \varepsilon_m \in \{-1, 1\}$ se cumpla que
$$\left\| \sum_{i=1}^m \varepsilon_i \, \alpha_i \, b_i \right\| = \left\| \sum_{i=1}^m \alpha_i \, b_i \right\|^{2}. \tag{B.3}$$

Así, en principio tenemos que
$$\Big\{ \| \ \| \text{ norma en } \mathbb{R}^m \text{ 1-incondicional} \Big\} \subseteq \Big\{ \| \ \| \text{ norma en } \mathbb{R}^m \text{ tal que } \mathrm{inc}(X) = 1 \Big\}.$$

Ejemplos de espacios (o normas) 1-incondicionales son los espacios $l_p^m$, con $1 \leq p \leq \infty$, y sus sumas directas.

Es interesante notar que la cantidad inc($X$) tiene una interpretación *geométrica*, que demostramos a continuación.

---

[1] Tenemos razones para pensar tanto que esto ocurre, como que no. Por lo que seguramente estamos equivocados.

[2] Ya mencionamos esta noción en el comentario final de la Introducción.
Notemos a su vez que podríamos interpretar esto diciendo que la norma $\| \ \|$ posee al menos las $2^m$ isometrías dadas por $b_i \mapsto \pm b_i$.

**Proposición B.2.2.** *Sea $X = (\mathbb{R}^m, \|\ \|)$ un espacio normado. Entonces*

$$\operatorname{inc}(X) = \inf_{Y \in \mathcal{I}_m} d_{BM}(X,Y), \quad \text{donde } \mathcal{I}_m = \Big\{ Y = (\mathbb{R}^m, \|\ \|_Y) \text{ tal que } Y \text{ es 1-incondicional} \Big\}.$$

*Es decir,* $\operatorname{inc}(X)$ *(o su logaritmo natural) es la distancia en* $\mathcal{Q}_m$ *al espacio de todas las normas 1-incondicionales en* $\mathbb{R}^m$*.*

*Demostración.* Sea $\mathcal{B} = \{b_1, \dots, b_m\}$ una base de $\mathbb{R}^m$. Definamos en $\mathbb{R}^m$ la norma $\|\ \|_I$ dada por

$$\left\| \sum_{i=1}^m \alpha_i\, b_i \right\|_I = \max_{\varepsilon \in \{-1,1\}^m} \left\| \sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i \right\|.$$

Es claro que $\|\ \|_I$ es una norma 1-incondicional, y que para todo $x \in \mathbb{R}^m$ se tiene

$$\|x\| \le \|x\|_I \le \operatorname{inc}(\mathcal{B})\,\|x\|.$$

Luego, $d_{BM}(X, (\mathbb{R}^m, \|\ \|_I)) \le \operatorname{inc}(\mathcal{B})$. Y por lo tanto tenemos que

$$\inf_{Y \in \mathcal{I}_m} d_{BM}(X,Y) \le \operatorname{inc}(\mathcal{B}).$$

Así, tomando ínfimo sobre todas las bases en $\mathbb{R}^m$ llegamos a que

$$\inf_{Y \in \mathcal{I}_m} d_{BM}(X,Y) \le \operatorname{inc}(X).$$

Por otro lado, sea $Y = (\mathbb{R}^m, \|\ \|_Y)$ un espacio normado 1-incondicional. Tomemos entonces una base 1-incondicional de $Y$, $\mathcal{B}_Y = \{y_1, \dots, y_m\}$, y un isomorfismo $T : Y \to X$ tal que $\|T\| \left\|T^{-1}\right\| = d_{BM}(X,Y)$[3]. Consideremos la base $\mathcal{B}_X = \{x_i = T(y_i), 1 \le i \le m\}$ de $X$. Entonces tenemos que, dados $\alpha_1, \dots, \alpha_m \in \mathbb{R}$,

$$\begin{aligned}
\max_{\varepsilon \in \{-1,1\}^m} \left\| \sum_{i=1}^m \varepsilon_i\, \alpha_i\, x_i \right\|_X &= \max_{\varepsilon \in \{-1,1\}^m} \left\| T\left( \sum_{i=1}^m \varepsilon_i\, \alpha_i\, y_i \right) \right\|_X \le \|T\| \max_{\varepsilon \in \{-1,1\}^m} \left\| \sum_{i=1}^m \varepsilon_i\, \alpha_i\, y_i \right\|_Y \\
= \|T\| \left\| \sum_{i=1}^m \alpha_i\, y_i \right\|_Y &\le \|T\| \left\|T^{-1}\right\| \left\| \sum_{i=1}^m \alpha_i\, x_i \right\|_X = d_{BM}(X,Y) \left\| \sum_{i=1}^m \alpha_i\, x_i \right\|_X .
\end{aligned}$$

Por lo que $\operatorname{inc}(X) \le \operatorname{inc}(\mathcal{B}_X) \le d_{BM}(X,Y)$. De donde obtenemos $\operatorname{inc}(X) \le \inf_{Y \in \mathcal{I}_m} d_{BM}(X,Y)$. □

**Observación B.2.3.**

(I) Como $l_2^m$ es un espacio 1-incondicional y para todo espacio normado $X = (\mathbb{R}^m, \|\ \|)$ se satisface $d_{BM}(X, l_2^m) \le \sqrt{m}$ (ver Observación 1.1.6 (II)), tenemos según la proposición anterior que $\operatorname{inc}(X) \le \sqrt{m}$ para todo espacio normado $X = (\mathbb{R}^m, \|\ \|)$.

(II) Las constantes de incondicionalidad de espacios normados de dimensión finita surgen en el estudio de espacios de Banach infinito-dimensionales, ver [DJT95, Capítulo 17].

[3] Usamos aquí el hecho de que el ínfimo en la definición de la distancia de Banach-Mazur sí se alcanza (en el caso finito dimensional), aunque podríamos obviar esto para la presente demostración.

En [LT16] Lopez-Abad y Tradacete introducen, interesados en el estudio de espacios de Banach infinito-dimensionales, una noción más débil (en el contexto infinito dimensional) a la de incondicionalidad, que utilizaremos más adelante, y definimos a continuación en el contexto finito-dimensional.

**Definición B.2.4.** Sea $X = (\mathbb{R}^m, \| \;\|)$ un espacio normado.

1. Dada una base $\mathcal{B} = \{b_1, \dots, b_m\}$ de $X$, definimos $\mathrm{dia}(\mathcal{B})$, la constante de divergencia incondicional aleatoria de la base $\mathcal{B}$, como

$$\mathrm{dia}(\mathcal{B}) = \sup_{(\alpha_1,\dots,\alpha_m)\neq 0} \frac{\left\|\sum_{i=1}^m \alpha_i\, b_i\right\|}{\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i\right\|\right)}.$$

   Es decir, $\mathrm{dia}(\mathcal{B})$ es la menor constante tal que, para cualesquiera escalares $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ se tiene que

$$\left\|\sum_{i=1}^m \alpha_i\, b_i\right\| \leq \mathrm{dia}(\mathcal{B})\, \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i\right\|\right). \tag{B.4}$$

2. Definimos $\mathrm{dia}(X)$, la constante de divergencia incondicional aleatoria de $X$, como

$$\mathrm{dia}(X) = \inf_{\mathcal{B} \text{ base de } X} \mathrm{dia}(\mathcal{B}). \tag{B.5}$$

Notemos que para cada base $\mathcal{B}$ de $\mathbb{R}^m$ se tiene que $\mathrm{dia}(\mathcal{B}) \geq 1$, ya que de (B.4) vemos que

$$\max_{\tau_i=\pm1}\left\|\sum_{i=1}^m \tau_i\, \alpha_i\, b_i\right\| \leq \mathrm{dia}(\mathcal{B})\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i\right\|\right) \leq \mathrm{dia}(\mathcal{B}) \max_{\tau_i=\pm1}\left\|\sum_{i=1}^m \tau_i\, \alpha_i\, b_i\right\|.$$

Por lo que resulta $\mathrm{dia}(X) \geq 1$.

Por otro lado, mencionemos que al igual que con $\mathrm{inc}(X)$, ignoramos si el ínfimo en la definición de $\mathrm{dia}(X)$ se alcanza en general.

Veamos como se relacionan las constantes $\mathrm{inc}(X)$ y $\mathrm{dia}(X)$ de un espacio normado de dimensión finita.

**Proposición B.2.5.** *Sea $X = (\mathbb{R}^m, \| \;\|)$ un espacio normado. Entonces* $\mathrm{dia}(X) \leq \mathrm{inc}(X)$.
*En particular* $\mathrm{dia}(X) \leq \sqrt{m}$.

*Demostración.* Dada una base $\mathcal{B} = \{b_1, \dots, b_m\}$ de $X$, para cada $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m) \in \{-1, 1\}^m$ fijo, se tiene que dados $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ cualesquiera

$$\max_{\tau_i=\pm1}\left\|\sum_{i=1}^m \tau_i\, \alpha_i\, b_i\right\| = \max_{\tau_i=\pm1}\left\|\sum_{i=1}^m \tau_i\, \varepsilon_i\, \alpha_i\, b_i\right\| \leq \mathrm{inc}(\mathcal{B}) \left\|\sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i\right\|.$$

Por lo que tomando esperanza respecto de $\varepsilon \in \{-1, 1\}^m$ a ambos miembros obtenemos

$$\max_{\tau_i=\pm1}\left\|\sum_{i=1}^m \tau_i\, \alpha_i\, b_i\right\| \leq \mathrm{inc}(\mathcal{B})\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\, \alpha_i\, b_i\right\|\right).$$

En particular

$$\left\|\sum_{i=1}^{m}\alpha_i\, b_i\right\| \le \mathrm{inc}(\mathcal{B})\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^{m}\varepsilon_i\,\alpha_i\, b_i\right\|\right).$$

Por lo que $\mathrm{dia}(X) \le \mathrm{dia}(\mathcal{B}) \le \mathrm{inc}(\mathcal{B})$. A partir de aquí, tomando ínfimo sobre todas las bases en $X$ obtenemos el enunciado. □

## B.3. Concentración para espacios con dia no muy grande

Para demostrar la Proposición 3.4.6 utilizaremos, además de lo ya visto en este apéndice, tres lemas que enunciamos a continuación. Cada uno de ellos nos proveerá de una desigualdad (que en algunos casos utilizaremos en más de una ocasión). Sus demostraciones resultan de interés en tanto muestran como pueden "ponerse en acción" varios de los conceptos vistos en la Sección 2.2.

**Lema B.3.1.** *Sean $(\Omega,\Sigma,\mu)$ espacio de probabilidad y $g_1,\dots,g_m : \Omega \to \mathbb{R}$ variables aleatorias gaussianas estándar e independientes. Dados $v_1,\dots,v_m$ vectores en un espacio normado $X$, se tiene que*

$$M_m((v_i)_{1\le i\le m}) = \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^{m}\varepsilon_i\, v_i\right\|\right) \le \sqrt{\frac{\pi}{2}}\,\mathbb{E}_\mu\left(\left\|\sum_{i=1}^{m} g_i\, v_i\right\|\right).$$

*Demostración.* Comencemos notando que si $b\colon(\Omega_1,\mu_1)\to\mathbb{R}$ es una variable aleatoria con distribución Bernoulli 1/2 tomando los valores 1 y $-1$, y $g\colon(\Omega_2,\mu_2)\to\mathbb{R}$ es una variable aleatoria gaussiana estándar, entonces la variable aleatoria $b\cdot|g|\colon(\Omega_1,\mu_1)\times(\Omega_2,\mu_2)\to\mathbb{R}$, dada por $b\cdot|g|(\omega_1,\omega_2) = b(\omega_1)|g(\omega_2)|$ tiene distribución gaussiana estándar.

En efecto, dado $a\in\mathbb{R}$

$$(\mu_1\times\mu_2)(b\cdot|g|>a) = \mu_1(b=1)\,\mu_2(|g|>a)+\mu_1(b=-1)\,\mu_2(-|g|>a) = \frac{1}{2}\Big(\mu_2(|g|>a)+\mu_2(|g|<-a)\Big).$$

De donde, si $a\ge 0$, usando la simetría en la distribución de $g$, tenemos

$$(\mu_1\times\mu_2)(b\cdot|g|>a) = \mu_2(g>a) = \gamma_1((a,+\infty)).$$

Por otro lado, si $a<0$, nos queda

$$(\mu_1\times\mu_2)(b\cdot|g|>a) = \frac{1}{2}\Big(1+\mu_2(g\in(a,-a))\Big) = \mu_2(g\ge 0)+\mu_2(g\in(a,0)) = \mu_2(g>a) = \gamma_1((a,+\infty)).$$

A partir de esta observación, teniendo en cuenta la Proposición 2.2.5, y que $\mathbb{E}_\mu(|g|) = \sqrt{\frac{2}{\pi}}$ (ver (2.7) en el Lema 2.5.6), podemos escribir

$$\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^{m}\varepsilon_i\, v_i\right\|\right) = \sqrt{\frac{\pi}{2}}\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^{m}\varepsilon_i\,\mathbb{E}_\mu(|g_i|)\, v_i\right\|\right) = \sqrt{\frac{\pi}{2}}\mathbb{E}_\varepsilon\left(\left\|\mathbb{E}_\mu\left(\sum_{i=1}^{m}\varepsilon_i\,|g_i|\, v_i\right)\right\|\right)$$
$$\le\sqrt{\frac{\pi}{2}}\,\mathbb{E}_\varepsilon\,\mathbb{E}_\mu\left(\left\|\sum_{i=1}^{m}\varepsilon_i\,|g_i|\, v_i\right\|\right) = \sqrt{\frac{\pi}{2}}\mathbb{E}_\mu\left(\left\|\sum_{i=1}^{m} g_i\, v_i\right\|\right).$$

□

Veremos ahora que $k(X)$ esta acotado inferiormente para todo espacio normado $X = (\mathbb{R}^m, \|\ \|)$.

**Lema B.3.2.** *Sea $X = (\mathbb{R}^m, \|\ \|)$ un espacio normado. Entonces*

$$k(X) = \frac{\left(\mathbb{E}_{\gamma_m}(\|x\|)\right)^2}{b^2} \geq \frac{2}{\pi},$$

*donde $b = \left\| Id : l_2^m \to (\mathbb{R}^m, \|\ \|)\right\|$.*

*Demostración.* Comencemos observando que

$$b = \left\| Id^* : (\mathbb{R}^m, \|\ \|)^* \to l_2^m \right\|.$$

Más concretamente, si $\varphi \in X^*$ esta dada por $\varphi(x) = \langle x, v_\varphi \rangle$, con $v_\varphi \in \mathbb{R}^m$ y $\langle\ ,\ \rangle$ el producto interno de $l_2^m$, entonces tenemos que $Id^*(\varphi) = v_\varphi$, por lo que

$$b = \sup_{\varphi \in B_{X^*}} \|Id^*(\varphi)\|_2 = \sup_{\varphi \in B_{X^*}} \left\|v_\varphi\right\|_2.$$

Ahora, para cada $\varphi \in B_{X^*}$, usando el Lema 2.2.1, y que si $g$ es una variable aleatoria gaussiana estándar entonces $\mathbb{E}(|g|) = \sqrt{2/\pi}$, tenemos que

$$\mathbb{E}_{\gamma_m}(\|x\|) \geq \mathbb{E}_{\gamma_m}(|\varphi(x)|) = \mathbb{E}_{\gamma_m}\left(|\langle x, v_\varphi \rangle|\right) = \left\|v_\varphi\right\|_2 \mathbb{E}_{\gamma_m}\left(\left|\langle x, \frac{v_\varphi}{\left\|v_\varphi\right\|_2}\rangle\right|\right) = \left\|v_\varphi\right\|_2 \mathbb{E}(|g|) = \left\|v_\varphi\right\|_2 \sqrt{\frac{2}{\pi}}.$$

Así, tomando supremo sobre $\varphi \in B_{X^*}$, nos queda el enunciado. □

**Observación B.3.3.** También se cumple que $k(X) \leq \dim(X)$ (ver [Pis89, página 43]).

El siguiente lema resulta de interés, en tanto que nos dice que la cota vista en la Proposición 3.1.1 (que a su vez se desprende del Teorema 2.2.11) es "de un orden asintótico correcto" (salvo constantes en el exponente) *para valores grandes* de $t$.

**Lema B.3.4.** *Sea $X = (\mathbb{R}^m, \|\ \|)$ un espacio normado. Entonces*

$$\gamma_m\left(\|x\| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\right) \geq \exp(-4\,k(X)\,t^2) \quad \forall\, t \geq 1.$$

*Demostración.* Usando la misma notación y argumentos que en la demostración del Lema B.3.2, tenemos que si $\varphi \in B_{X^*}$, entonces

$$\begin{aligned}\gamma_m\left(\|x\| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\right) &\geq \gamma_m\left(|\varphi(x)| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\right) = \gamma_m\left(|\langle x, v_\varphi\rangle| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\right)\\ &= \gamma_m\left(\left|\langle x, \frac{v_\varphi}{\left\|v_\varphi\right\|_2}\rangle\right| \geq (1+t)\,\frac{\mathbb{E}_{\gamma_m}(\|x\|)}{\left\|v_\varphi\right\|_2}\right) = \gamma_1\left(|x_1| > (1+t)\frac{\mathbb{E}_{\gamma_m}(\|x\|)}{\left\|v_\varphi\right\|_2}\right).\end{aligned}$$

Tomando supremo en $\varphi \in B_{X^*}$ en la última expresión, llegamos a que

$$\gamma_m\left(\|x\| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\right) \geq \gamma_1\left(|x_1| > (1+t)\sqrt{k(X)}\right).$$

Valiéndonos ahora de las cotas para la distribución gaussiana vistas en el Lema 2.2.8, tenemos que

$$\begin{aligned}
&\gamma_m\Big(\|x\| \geq (1+t)\,\mathbb{E}_{\gamma_m}(\|x\|)\Big) \geq \gamma_1\Big(|x_1| > (1+t)\sqrt{k(X)}\Big) = 2\,\gamma_1\Big(x_1 > (1+t)\sqrt{k(X)}\Big)\\
\geq &\sqrt{\frac{2}{\pi}}\,\frac{(1+t)\sqrt{k(X)}}{1+(1+t)^2k(X)}\exp\left(-\frac{(1+t)^2k(X)}{2}\right) \geq \exp\Big(-(1+t)^2k(X)\Big)\\
\geq &\exp\Big(-(2t)^2k(X)\Big) = \exp\Big(-4t^2\,k(X)\Big).
\end{aligned}$$

Donde hemos usado que $(1+t)\sqrt{k(X)} \geq 2\sqrt{2/\pi} > \sqrt{2}$, y la desigualdad

$$\sqrt{\frac{2}{\pi}}\frac{s}{1+s^2} \geq \exp(-s^2/2) \quad \forall s \geq \sqrt{2},$$

la cual puede deducirse viendo que la función

$$\xi(s) = \sqrt{\frac{2}{\pi}}\,\frac{s\,\exp(s^2/2)}{1+s^2}$$

es creciente en $\mathbb{R}$ y es mayor a 1 en $s = \sqrt{2}$. $\square$

Ahora sí, veremos a continuación la demostración de la Proposición 3.4.6. Para un mejor entendimiento de la misma, es conveniente realizar algunas observaciones previas a la demostración. La primera, es que para poder utilizar tanto el Lema 3.4.5 como la Proposición 3.4.1, supondremos inicialmente que la norma con la que trabajamos es suave. Luego de haber visto lo deseado para normas suaves, para poder obtener el resultado en general, necesitaremos aproximar normas arbitrarias por normas suaves. En [Sch13, Sección 3.4] se muestra como esto es posible (aclaramos ante la posible duda, que donde dice *función soporte* en dicha fuente, podemos leer sin escrúpulos la palabra *norma*). Más precisamente usaremos que dados $0 < \delta < 1$ y una norma $\|\ \|$ en $\mathbb{R}^m$, existe una norma $\|\ \|_\delta$ $\mathcal{C}^\infty$ en $\mathbb{R}^m \setminus \{0\}$ tal que $\|x\|_\delta \leq \|x\| \leq (1+\delta)\|x\|_\delta$ para todo $x \in \mathbb{R}^m$.

La segunda observación, es que esta demostración es larga y un tanto tediosa, pero también está aquí uno de los tres puntos claves mencionados en la Sección 3.4 para "desbloquear" el Teorema 3.1.3.

**Proposición 3.4.6.** *Sea $X = (\mathbb{R}^m, \|\ \|)$ un espacio normado. Entonces existe un isomorfismo lineal $T : \mathbb{R}^m \to \mathbb{R}^m$, tal que para cada $t > 0$*

$$\gamma_m\Big(\Big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \leq 5\exp\left(-K\,\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right)t\right), \tag{3.7}$$

*con $K > 0$ constante absoluta (que podemos tomar como $K = 1/31$).*

*Demostración.* Supondremos inicialmente que la norma $\|\ \|$ es $\mathcal{C}^1$ en $\mathbb{R}^m \setminus \{0\}$.

Sea $c > 1$ fijo (a determinar más adelante), y tomemos una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ de $X$ tal que $\mathrm{dia}(\mathcal{B}) < c\,\mathrm{dia}(X)$. De manera que

$$\left\|\sum_{i=1}^m \alpha_i\,b_i\right\| \leq c\,\mathrm{dia}(X)\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\,\alpha_i\,b_i\right\|\right) \quad \text{para cualesquiera } \alpha_1, \ldots, \alpha_m \in \mathbb{R}.$$

Dada la naturaleza del enunciado ("dada una norma $\| \; \|$ en $\mathbb{R}^m$ existe un isomorfismo $T$ en $\mathbb{R}^m$ tal que $\|T(x)\|$ satisface..."), podemos considerar un isomorfismo $T_1 : \mathbb{R}^m \to \mathbb{R}^m$ tal que $T_1(e_i) = b_i$ para cada $1 \leq i \leq m$ (donde $e_i$ es el $i$-ésimo vector de la base canónica), y buscar demostrar lo afirmado para la norma $p : \mathbb{R}^m \to \mathbb{R}$, isométrica a $\| \; \|$, dada por

$$p(\alpha_1, \dots, \alpha_m) = \left\| \sum_{i=1}^{m} \alpha_i b_i \right\| = \|T_1(\alpha_1, \dots, \alpha_m)\|.$$

Ya que, de ver lo afirmado en el enunciado para la norma $p$, tendremos lo afirmado para $\| \; \|$, con $T_1 \circ T$ como el isomorfismo que cumple lo dicho en el enunciado.

Sin riesgo de confusión, tras lo observado anteriormente, en lo que sigue notaremos nuevamente por $\| \; \|$ a la norma $p$ (gajes del oficio).

Entonces tenemos que

$$\|(\alpha_1, \dots, \alpha_m)\| \leq c \operatorname{dia}(X) \mathbb{E}_\varepsilon \left( \|(\varepsilon_1 \alpha_1, \dots, \varepsilon_m \alpha_m)\| \right) \quad \forall (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m.$$

Sea ahora $\Lambda$ el operador diagonal del que nos provee el Lema 3.4.5 con respecto a la función Lipschitz $f(x) = \|x\|$ (notemos que $\Lambda$ depende de $c$, que igualmente dejaremos fijo). Sean

$$a_\Lambda = \|\Lambda : l_\infty^m \to X\| \quad \text{y} \quad b_\Lambda = \left\|\Lambda : l_2^m \to X\right\|$$

las constantes Lipschitz de la función $\|\Lambda(x)\|$ respecto de la norma infinito y la norma euclídea respectivamente. Notemos que al ser $\Lambda$ un operador positivo, $\|\Lambda(x)\|$ es una norma en $\mathbb{R}^m$. De esta manera, como le corresponde, ponemos $k_\Lambda = \dfrac{\mathbb{E}_{\gamma_m}(\|\Lambda(x)\|)^2}{b_{\Lambda^2}}$.

Queremos llegar a ver que

$$\gamma_m\left( \Big| \|\Lambda(x)\| - \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \Big| > t\, \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \right) \leq 5 \exp\left( -K \ln\left( \frac{e\sqrt{m}}{c \operatorname{dia}(X)} \right) t \right). \tag{B.6}$$

Usando, como no podía ser de otra manera, la Proposición 3.4.1 junto con el Lema B.1.1, tenemos que

$$\gamma_m\left( \Big| \|\Lambda(x)\| - \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \Big| > t\, \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \right) \leq 5 \exp\left( -\sqrt{\ln\left( \frac{e\sqrt{m}\, b_\Lambda}{a_\Lambda} \right)} \sqrt{k_\Lambda}\, t \right). \tag{B.7}$$

Ahora bien, si fuese $\sqrt{k_\Lambda} \geq \sqrt{2/\pi} \dfrac{\sqrt{m}}{c \operatorname{dia}(X)}$, entonces, usando que $\ln(s) \leq s/e$, tendríamos

$$\sqrt{k_\Lambda} \geq \sqrt{\frac{2}{\pi}} \frac{\sqrt{m}}{c \operatorname{dia}(X)} = \sqrt{\frac{2}{\pi} \frac{1}{e} \frac{e\sqrt{m}}{c \operatorname{dia}(X)}} \geq \sqrt{\frac{2}{\pi} \ln\left( \frac{e\sqrt{m}}{c \operatorname{dia}(X)} \right)},$$

y como

$$\ln\left( \frac{e\sqrt{m}\, b_\Lambda}{a_\Lambda} \right) \geq 1,$$

tendríamos (B.6) con $K = \sqrt{\frac{2}{\pi}}$.

Veamos ahora como obtener (B.6) si se cumple

$$\sqrt{k_\Lambda} < \sqrt{2/\pi}\,\frac{\sqrt{m}}{c\,\mathrm{dia}(X)}. \tag{B.8}$$

En tal caso, podemos acotar superiormente $a_\Lambda$, mediante el uso del Lema B.3.1 (recordando que las funciones coordenadas $x_1, \ldots, x_m$ son variables aleatorias gaussianas estándar e independientes en $(\mathbb{R}^m, \gamma_m)$), de la siguiente manera.[4]

$$a_\Lambda = \max_{\varepsilon_i = \pm 1} \left\| \sum_{i=1}^m \varepsilon_i\, \lambda_i\, e_i \right\| \leq c\,\mathrm{dia}(X)\,\mathbb{E}_\varepsilon \left( \left\| \sum_{i=1}^m \varepsilon_i\, \lambda_i\, e_i \right\| \right) \leq c\,\mathrm{dia}(X) \sqrt{\frac{\pi}{2}}\,\mathbb{E}_{\gamma_m}(\|\Lambda(x)\|).$$

Luego, a partir de (B.7) nos queda

$$\gamma_m\Big( \big|\, \|\Lambda(x)\| - \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \big| > t\,\mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \Big) \leq 5\,\exp\left( -\sqrt{\ln\left( \frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)} \right)}\,\sqrt{k_\Lambda}\,t \right). \tag{B.9}$$

Para obtener (B.6), queremos ver entonces que

$$k_\Lambda \ln\left( \frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)} \right) \geq \widetilde{C}\,\ln^2\left( \frac{e\sqrt{m}}{c\,\mathrm{dia}(X)} \right) \tag{B.10}$$

para alguna constante $\widetilde{C} > 0$ (y en tal caso será $K = \sqrt{\widetilde{C}}$). Para eso, haremos muchas cuentas.

Usando (B.9) y el Lema B.3.4 para la norma $\|\Lambda(x)\|$, si $t \geq 1$ podemos escribir

$$\begin{aligned} \exp\left(-4\,k_\Lambda\,t^2\right) &\leq \gamma_m\Big( \|\Lambda(x)\| - \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \geq t\,\mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \Big) \\ &\leq \gamma_m\Big( \big|\, \|\Lambda(x)\| - \mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \big| \geq t\,\mathbb{E}_{\gamma_m}(\|\Lambda(x)\|) \Big) \\ &\leq 5\,\exp\left( -\sqrt{\ln\left( \frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)} \right)}\,\sqrt{k_\Lambda}\,t \right). \end{aligned}$$

De donde, tomando logaritmo a ambos miembros y evaluando en $t = 1$

$$-4k_\Lambda \leq \ln(5) - \sqrt{\ln\left( \frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)} \right)}\,\sqrt{k_\Lambda},$$

o equivalentemente

$$\sqrt{\ln\left( \frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)} \right)}\,\sqrt{k_\Lambda} \leq \ln(5) + 4k_\Lambda.$$

[4] Donde usamos también que por el teorema de Krein-Milman, la norma de un operador saliendo de $l_\infty^m$ siempre se alcanza en alguno de los puntos extremales de la bola $B_\infty^m$. Es decir, en alguna “de las esquinas” del hipercubo $[-1, 1]^m$.

Usando el Lema B.3.2, nos queda

$$\sqrt{\ln\left(\frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)}\right)} \leq \frac{\ln(5)}{\sqrt{k_\Lambda}} + 4\sqrt{k_\Lambda} \leq \sqrt{\frac{\pi}{2}}\ln(5) + 4\sqrt{k_\Lambda} \leq \frac{\pi}{2}\ln(5)\sqrt{k_\Lambda} + 4\sqrt{k_\Lambda} = R\sqrt{k_\Lambda}, \tag{B.11}$$

siendo $R = (\ln(5)\pi/2 + 4)$.

Elevando al cuadrado en (B.11), y usando nuevamente que $\ln(s) \leq s/e$, tenemos que

$$\begin{aligned}
&\ln\left(\frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)}\right) \leq R^2 k_\Lambda \\
&\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right) - \frac{1}{2}\ln\left(\frac{\pi}{2}k_\Lambda\right) \leq R^2 k_\Lambda \\
&\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right) \leq R^2 k_\Lambda + \frac{1}{2}\ln\left(\frac{\pi}{2}k_\Lambda\right) \leq R^2 k_\Lambda + \frac{\pi}{4e}k_\Lambda = \left(R^2 + \frac{\pi}{4e}\right)k_\Lambda.
\end{aligned} \tag{B.12}$$

Supongamos ahora que

$$\ln\left(\sqrt{\frac{\pi}{2}k_\Lambda}\right) < \frac{1}{2}\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right).$$

Entonces, usando (B.12)

$$k_\Lambda \ln\left(\frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)}\right) = k_\Lambda\left(\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right) - \ln\left(\sqrt{\frac{\pi}{2}k_\Lambda}\right)\right) \geq k_\Lambda\frac{1}{2}\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right) \geq \widetilde{C}\ln^2\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right),$$

con $\widetilde{C} = \dfrac{2e}{\pi + 4eR^2}$.

Por otro lado, si

$$\ln\left(\sqrt{\frac{\pi}{2}k_\Lambda}\right) \geq \frac{1}{2}\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right),$$

entonces, exponenciando a ambos miembros tenemos que

$$\frac{\pi}{2}k_\Lambda \geq \frac{e\sqrt{m}}{c\,\mathrm{dia}(X)} = \left(\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right)^{1/2}\right)^2 \geq \left(e\ln\left(\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right)^{1/2}\right)\right)^2 = \frac{e^2}{4}\ln^2\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right).$$

Y como, por (B.8)

$$\ln\left(\frac{e\sqrt{2m}}{\sqrt{\pi}\sqrt{k_\Lambda}\,c\,\mathrm{dia}(X)}\right) \geq 1,$$

también llegamos a (B.10) en este caso, con $\widetilde{C} = \dfrac{e^2}{2\pi}$.

Luego, tenemos (B.6), con

$$K = \mathrm{mín}\left(\sqrt{\frac{2}{\pi}}, \left(\frac{2e}{\pi + 4e\big(\ln(5)\pi/2 + 4\big)^2}\right)^{1/2}, \frac{e}{\sqrt{2\pi}}\right) \geq 1/10.$$

Así, tenemos por ahora que si $\|\ \|$ es una norma suave, dado $c > 1$, existe un isomorfismo lineal $T = T_c$ en $\mathbb{R}^m$ tal que

$$\gamma_m\Big(\Big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \leq 5\exp\left(-\frac{1}{10}\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right)t\right).$$

Aplicando lo anterior para $c = e^{1/11}$ y notando que

$$\frac{1}{10}\ln\left(\frac{e\sqrt{m}}{c\,\mathrm{dia}(X)}\right)t = \frac{1}{10}\left(\frac{10}{11} + \ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right)t > \frac{1}{11}\left(1 + \ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right)t = \frac{1}{11}\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right)t,$$

llegamos a que, dada una norma suave en $\mathbb{R}^m$, existe un isomorfismo lineal $T : \mathbb{R}^m \to \mathbb{R}^m$ tal que

$$\gamma_m\Big(\Big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \leq 5\exp\left(-\frac{1}{11}\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right)t\right).$$

Sea ahora $\|\ \|$ una norma cualquiera en $\mathbb{R}^m$. Entonces, como lo mencionamos previo a esta demostración, dado $0 < \delta < 1$ tenemos que existe una norma $\|\ \|_\delta$ $\mathcal{C}^\infty$ en $\mathbb{R}^m \setminus \{0\}$ tal que

$$\|x\|_\delta \leq \|x\| \leq (1+\delta)\,\|x\|_\delta \quad \forall\, x \in \mathbb{R}^m.$$

Notemos $X_\delta = (\mathbb{R}^m, \|\ \|_\delta)$. Veamos que $\mathrm{dia}(X_\delta) \leq (1+\delta)\,\mathrm{dia}(X)$. Tomando una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ de $\mathbb{R}^m$, vemos que en efecto

$$\left\|\sum_{i=1}^m \alpha_i\, b_i\right\|_\delta \leq \left\|\sum_{i=1}^m \alpha_i\, b_i\right\| \leq \mathrm{dia}(\mathcal{B}, X)\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\,\alpha_i\, b_i\right\|\right) \leq (1+\delta)\,\mathrm{dia}(\mathcal{B}, X)\,\mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i\,\alpha_i\, b_i\right\|_\delta\right).$$

De donde $\mathrm{dia}(\mathcal{B}, X_\delta) \leq (1+\delta)\,\mathrm{dia}(\mathcal{B}, X)$. Tomando ínfimo sobre las bases en $\mathbb{R}^m$ vemos que $\mathrm{dia}(X_\delta) \leq (1+\delta)\,\mathrm{dia}(X)$.

Fijemos entonces $0 < \delta < \dfrac{1}{3 + \ln(\sqrt{m}/\mathrm{dia}(X))}$ (¡lo mantendremos *fijo* siempre!). Entonces, por lo visto anteriormente, tenemos que existe un isomorfismo lineal $T = T_\delta$ en $\mathbb{R}^m$ tal que, para todo $t > 0$

$$\gamma_m\Big(\Big|\,\|T(x)\|_\delta - \mathbb{E}_{\gamma_m}(\|T(x)\|_\delta)\Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|_\delta)\Big) \leq 5\exp\left(-\frac{1}{11}\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right)t\right).$$

Delegamos en quien lee el redescubrimiento de que si $t > 3\delta$ (y al cumplirse por definición que $\delta < 1/3$), se tiene que

$$\Big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\Big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|) \implies \Big|\,\|T(x)\|_\delta - \mathbb{E}_{\gamma_m}(\|T(x)\|_\delta)\Big| > \frac{t}{2}\,\mathbb{E}_{\gamma_m}(\|T(x)\|_\delta).$$

Así, si $t > 3\delta$ podemos escribir

$$\begin{aligned}
&\gamma_m\Big(\big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \\
&\leq \gamma_m\Big(\big|\,\|T(x)\|_\delta - \mathbb{E}_{\gamma_m}(\|T(x)\|_\delta)\big| > \frac{t}{2}\,\mathbb{E}_{\gamma_m}(\|T(x)\|_\delta)\Big) \\
&\leq 5\exp\left(-\frac{1}{22}\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X_\delta)}\right)\right) \leq 5\exp\left(-\frac{1}{22}\ln\left(\frac{e\sqrt{m}}{(1+\delta)\mathrm{dia}(X)}\right)t\right).
\end{aligned}$$

Y como

$$\begin{aligned}
\frac{1}{22}\ln\left(\frac{e\sqrt{m}}{(1+\delta)\mathrm{dia}(X)}\right) &> \frac{1}{22}\ln\left(\frac{e\sqrt{m}}{(4/3)\mathrm{dia}(X)}\right) = \frac{1}{22}\left(1-\ln(4/3)+\ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right) \\
&> \frac{(1-\ln(4/3))}{22}\left(1+\ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right) > \frac{1}{31}\left(1+\ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right),
\end{aligned}$$

tenemos que, si $t > 3\delta$, entonces

$$\gamma_m\Big(\big|\,\|T(x)\| - \mathbb{E}_{\gamma_m}(\|T(x)\|)\big| > t\,\mathbb{E}_{\gamma_m}(\|T(x)\|)\Big) \leq 5\exp\left(-\frac{1}{31}\ln\left(\frac{e\sqrt{m}}{\mathrm{dia}(X)}\right)t\right).$$

Por otro lado, si $0 < t \leq 3\delta < \dfrac{3}{3+\ln(\sqrt{m}/\mathrm{dia}(X))}$, entonces

$$5\exp\left(-\frac{1}{31}\left(1+\ln\left(\frac{\sqrt{m}}{\mathrm{dia}(X)}\right)\right)t\right) > 5\exp(-3/31) > 1,$$

por lo que la desigualdad (3.7) se cumple trivialmente. □

**Observación B.3.5.** Tras una necesaria pausa y respiración, es interesante observar lo siguiente. Como mencionamos en la Introducción, Bourgain y Lindenstrauss demostraron en [BL89] que si un espacio $X = (\mathbb{R}^m, \|\ \|)$ tiene una norma simétrica, esto es, una norma para la cual existe una base $b_1, \ldots, b_m$ tal que

$$\left\|\sum_{j=1}^m a_j\, b_j\right\| = \left\|\sum_{j=1}^m \pm a_{\sigma(j)}\, b_j\right\| \quad \forall\, \sigma \in S_m,$$

entonces se pueden encontrar en dicho espacio subespacios $(1+\varepsilon)$-isomorfos al espacio euclídeo de dimensión $k = \left[c_1 \frac{\ln(m)}{|\ln(c_2\varepsilon)|}\right]$.

Ahora bien, el hecho de tener una base simétrica es, como se ve, la conjunción de tener una base que sea tanto 1-incondicional como "invariante por permutaciones" (noción que no hemos definido en verdad, pero creemos se adivina con facilidad). Resulta interesante notar entonces que en los argumentos de la Proposición 3.4.6, el Lema 3.4.5 nos da "la invariancia por permutaciones" (al menos en el sentido en el que la necesitamos, o en que la usamos), y la constante de divergencia incondicional aleatoria nos da la incondicionalidad (al menos en el sentido en que la necesitamos, o en que la usamos).

# Apéndice C

# Concentración (para todos los espacios)

En este apéndice demostraremos los resultados asentados en la Sección 3.4.3.
Comencemos entonces enunciando nuevamente y demostrando la Proposición 3.4.7

**Proposición 3.4.7.** *Sean $X = (\mathbb{R}^n, \|\ \|)$ un espacio normado tal que $B_X$ está en posición de John y $0 < \delta < 1/2$. Entonces al menos una de las siguientes condiciones se satisface:*

1. *$k(X) \geq n^{1/2-\delta}$.*
2. *Existen un subespacio $F \subseteq \mathbb{R}^n$ con $\dim(F) = m$ tal que $\sqrt{n}/128 \leq m \leq n/2$ y una base $\{v_1, \ldots, v_m\}$ de $F$ tal que si $(\Omega, \Sigma, \mu)$ es un espacio de probabilidad y $g_1, \ldots, g_m \colon \Omega \to \mathbb{R}$ son variables aleatorias gaussianas estándar e independientes, entonces la variable aleatoria gaussiana $F$-valuada $\Phi \colon \Omega \to F$ dada por*

$$\Phi(\omega) = \sum_{i=1}^{m} g_i(\omega) v_i$$

*satisface que*

$$\mu\Big(\Big|\|\Phi\| - \mathbb{E}_\mu(\|\Phi\|)\Big| > t\,\mathbb{E}_\mu(\|\Phi\|)\Big) \leq 6\exp\left(-C\delta\, t\ln(n)\right) \quad \forall 0 < t \leq 1,$$

*donde $C > 0$ es una constante absoluta, que podemos tomar como $C = 1/62$.*

*Demostración.* Fijemos $0 < \delta < 1/2$, y supongamos que $k(X) < n^{1/2-\delta}$. Como $B_X$ esta en posición de John, podemos tomar la base ortonormal $\{u_1, \ldots, u_n\}$ de $\mathbb{R}^n$ de la cual nos provee el Lema 3.1.5. Entonces, teniendo en cuenta que $\|4u_i\| \geq 1$ para todo $1 \leq i \leq n$, por el Teorema 3.3.4 tenemos que existe $A \subseteq \{1, \ldots, n\}$, con $|A| \geq n/(32\tau)$ tal que

$$\frac{1}{2}\max_{i\in A}|\alpha_i| \leq \left\|\sum_{i\in A}\alpha_i\, 4u_i\right\| \leq 4\widetilde{M}_n \max_{i\in A}|\alpha_i|,$$

donde $\tau = \|T_1 : l_\infty^n \to X\|$, con $T_1$ definido por $T_1(e_i) = 4u_i$ (siendo $e_i$ el $i$-ésimo vector de la base canónica de $\mathbb{R}^n$), y $\widetilde{M}_n = \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^n \varepsilon_i\, 4u_i\right\|\right)$. Por lo que, si notamos $M_n = \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^n \varepsilon_i\, u_i\right\|\right)$, tenemos que

$$\frac{1}{8}\max_{i\in A}|\alpha_i| \leq \left\|\sum_{i\in A}\alpha_i\, u_i\right\| \leq 4M_n \max_{i\in A}|\alpha_i|.$$

Así, en caso de ser $|A| \leq n/2$, tomamos $F = \langle u_i \rangle_{i \in A}$ y $m = |A|$, de manera que $d_{BM}(F, l_\infty^m) \leq 32 M_n$. Por otro lado, en caso de ser $|A| > n/2$, tomamos $B \subseteq A$ con $m = |B| = [n/2]$ y definimos $F = \langle u_i \rangle_{i \in B}$, de manera que se sigue satisfaciendo $d_{BM}(F, l_\infty^m) \leq 32 M_n$.

Veamos ahora que $\tau \leq 4\sqrt{n}$, con lo que será $m = |A| \geq \sqrt{n}/128$. En efecto, al estar $B_X$ en posición de John tenemos que $b = 1$ (es decir $\|x\| \leq \|x\|_2$ para todo $x \in \mathbb{R}^n$), y por lo tanto

$$\|T_1(\alpha_1, \ldots, \alpha_n)\| \leq \|T_1(\alpha_1, \ldots, \alpha_n)\|_2 = \left\| \sum_{i=1}^n \alpha_i \, 4 u_i \right\|_2 = \left\| \sum_{i=1}^n \alpha_i \, 4 e_i \right\|_2 \leq 4\sqrt{n} \, \|(\alpha_1, \ldots, \alpha_n)\|_\infty .$$

Por otro lado, usando el Lema B.3.1 y la invariancia de la medida gaussiana por transformaciones ortogonales, tenemos que

$$M_n = \mathbb{E}_\varepsilon \left( \left\| \sum_{i=1}^n \varepsilon_i \, u_i \right\| \right) \leq \sqrt{\frac{\pi}{2}} \, \mathbb{E}_{\gamma_n} \left( \left\| \sum_{i=1}^n x_i \, u_i \right\| \right) = \sqrt{\frac{\pi}{2}} \, \mathbb{E}_{\gamma_n} \left( \left\| \sum_{i=1}^n x_i \, e_i \right\| \right) = \sqrt{\frac{\pi}{2}} \sqrt{k(X)}.$$

De aquí, al ser $l_\infty^m$ un espacio 1-incondicional, tenemos que

$$\mathrm{dia}(F) \leq \mathrm{inc}(F) \leq d_{BM}(F, l_\infty^m) \leq 32 \, M_n \leq 32 \sqrt{\frac{\pi}{2}} \sqrt{k(X)}.$$

Ahora supongamos, por facilidad en la escritura y sin perdida de generalidad, que $A = \{1, \ldots, m\}$ (en efecto, podemos lograr esto tras un reindexamiento de los vectores $(u_i)_{1 \leq i \leq n}$). Consideremos entonces en $\mathbb{R}^m$ la norma

$$\|(\alpha_1, \ldots, \alpha_m)\|_I = \left\| \sum_{i=1}^m \alpha_i \, u_i \right\|.$$

Entonces $(\mathbb{R}^m, \| \, \|_I)$ resulta un espacio isométrico a $(F, \| \, \|)$, mediante la isometría $H : (\mathbb{R}^m, \| \, \|_I) \to (F, \| \, \|)$ dada por $H(e_i) = u_i$ para $1 \leq i \leq m$. Así, aplicando la Proposición 3.4.6 vista en la sección anterior a $(\mathbb{R}^m, \| \, \|_I)$, tenemos que existe un isomorfismo $T : \mathbb{R}^m \to \mathbb{R}^m$ tal que

$$\gamma_m \Big( \big| \, \|T(x)\|_I - \mathbb{E}_{\gamma_m}(\|T(x)\|_I) \big| > t \, \mathbb{E}_{\gamma_m}(\|T(x)\|_I) \Big) \leq 5 \exp\left( \frac{-1}{31} \ln\left( \frac{e\sqrt{m}}{\mathrm{dia}(F)} \right) t \right).$$

Llamando $v_i = H(T(e_i)) \in F$ para cada $1 \leq i \leq m$, tenemos que la base $\{v_1, \ldots, v_m\}$ de $F$ cumple con que dados un espacio de probabilidad $(\Omega, \Sigma, \mu)$ y $g_1, \ldots, g_m : \Omega \to \mathbb{R}$ variables aleatorias gaussianas estándar e independientes, la variable aleatoria $\xi_1 : \Omega \to \mathbb{R}$, dada por

$$\xi_1(\omega) = \left\| \sum_{i=1}^m g_i(\omega) \, v_i \right\|$$

tiene la misma distribución que la variable aleatoria $\xi_2 : (\mathbb{R}^m, \gamma_m) \to \mathbb{R}$, dada por

$$\xi_2(x_1, \ldots, x_m) = \left\| \sum_{i=1}^m x_i \, v_i \right\| = \left\| \sum_{i=1}^m x_i \, T(e_i) \right\|_I = \left\| T\left( \sum_{i=1}^m x_i \, e_i \right) \right\|_I .$$

Luego,

$$
\begin{aligned}
&\mu\left(\left|\left\|\sum_{i=1}^{m} g_i\, v_i\right\| - \mathbb{E}_{\mu}\left(\left\|\sum_{i=1}^{m} g_i\, v_i\right\|\right)\right| > \mathbb{E}_{\mu}\left(\left\|\sum_{i=1}^{m} g_i\, v_i\right\|\right) t\right) \leq 5\, \exp\left(\frac{-1}{31}\left(1+\frac{1}{2}\ln\left(\frac{m}{\mathrm{dia}(F)^2}\right)\right) t\right)\\
\leq& 5\, \exp\left(\frac{-1}{31}\left(1+\frac{1}{2}\ln\left(\frac{n^{1/2}/128}{2^9 \pi\, k(X)}\right)\right) t\right) \leq 5\, \exp\left(\frac{-1}{31}\left(1+\frac{1}{2}\ln\left(\frac{n^{\delta}}{2^{16}\pi}\right)\right) t\right)\\
=& 5\, \exp\left(\frac{-1}{31}\left(1-\frac{1}{2}\ln\left(2^{16}\pi\right)+\frac{\delta}{2}\ln(n)\right) t\right) = 5\, \exp\left(\frac{1}{31}\left(\frac{1}{2}\ln\left(2^{16}\pi\right)-1\right) t\right) \exp\left(\frac{-\delta}{62}\ln(n)\, t\right)\\
\leq& 6 \exp\left(\frac{-\delta}{62}\ln(n)\, t\right) \qquad \forall\, 0 < t \leq 1.
\end{aligned}
$$

□

Comentemos que, si bien la cota superior sobre la dimensión $m$ en la anterior proposición puede parecer un tanto “artificial”, la realizamos en vistas de que nos será necesaria para la demostración del procurado Teorema 3.1.3.

Veamos ahora un último lema antes de demostrar dicho teorema.

**Lema C.0.1.** *Sea $\|\ \|$ una norma en $\mathbb{R}^n$. Entonces vale que*

$$\left(\mathbb{E}_{\gamma_n}(\|x\|)\right)^2 \geq \frac{1}{14}\mathbb{E}_{\gamma_n}\left(\|x\|^2\right).$$

*En particular, se tiene que*

$$k(l_2^n) = \left(\mathbb{E}_{\gamma_n}(\|x\|_2)\right)^2 \geq \frac{n}{14}.$$

*Demostración.* Por homogeneidad, podemos suponer que $\mathbb{E}_{\gamma_n}(\|x\|) = 1$. Observemos que en tal caso, como sabemos por el Lema B.3.2 que $k(X) \geq 2/\pi$, deberá ser $b \leq \sqrt{\pi/2}$, donde $b$ es la menor constante tal que $\|x\| \leq b\|x\|_2$. Así, escribamos

$$
\begin{aligned}
\mathbb{E}_{\gamma_n}\left(\|x\|^2\right) = \int_0^{+\infty} 2t\, \gamma_n(\|x\| > t)\, dt \leq 1 + \int_1^{+\infty} 2t\, \gamma_n(\|x\| > t)\, dt\\
= 1 + \int_0^{+\infty} 2(s+1)\gamma_n\left(\|x\| - \mathbb{E}_{\gamma_n}(\|x\|) > s\right) ds. \qquad \text{(C.1)}
\end{aligned}
$$

Donde usamos la sustitución $t = s+1$. Recordemos ahora que en la demostración del Teorema 2.2.11 vimos que, si $f : \mathbb{R}^n \to \mathbb{R}$ es Lipschitz con constante $b$ (respecto de la norma euclídea), entonces

$$\gamma_n\left(f - \mathbb{E}_{\gamma_n}(f) > s\right) \leq \exp\left(\frac{-2}{\pi^2}\frac{s^2}{b^2}\right). \qquad \text{(C.2)}$$

Por lo que, usando esto junto con que $b \leq \sqrt{\pi/2}$ (según lo observado más arriba) en (C.1), nos queda

$$\mathbb{E}_{\gamma_n}\left(\|x\|^2\right) \leq 1 + \int_0^{+\infty} 2(s+1)\exp\left(\frac{-4}{\pi^3}s^2\right)ds = 1 + 2\int_0^{+\infty}\exp\left(\frac{-4}{\pi^3}s^2\right)ds + \int_0^{+\infty} 2s\exp\left(\frac{-4}{\pi^3}s^2\right)ds$$
$$= 1 + 2\frac{1}{2}\sqrt{\frac{\pi^3}{4}\pi} + \int_0^{+\infty}\exp\left(\frac{-4}{\pi^3}v\right)dv = 1 + \frac{\pi^2}{2} + \frac{\pi^3}{4} \leq 14.$$

Luego, si $\| \; \|$ es una norma cualquiera en $\mathbb{R}^n$,

$$\mathbb{E}_{\gamma_n}\left(\|x\|^2\right) \leq 14\left(\mathbb{E}_{\gamma_n}(\|x\|)\right)^2,$$

que es lo que queríamos ver.

En particular, si consideramos la norma euclídea en $\mathbb{R}^n$, usando la notación previa al enunciado del Lema 2.5.6, y el cálculo para $\gamma(p)$ realizado en la demostración de dicho lema, nos queda que

$$k(l_2^n) = \left(\mathbb{E}_{\gamma_n}(\|x\|_2)\right)^2 \geq \frac{1}{14}\mathbb{E}_{\gamma_n}\left(\|x\|_2^2\right) = \frac{1}{14}n\,\gamma(2)^2 = \frac{1}{14}n\frac{2}{\sqrt{\pi}}\Gamma(3/2) = \frac{n}{14}. \qquad \square$$

**Observación C.0.2.**

(I) Mediante un cálculo directo, y usando el teorema de Stirling, puede verse que

$$\frac{k(l_2^n)}{n} = \frac{\left(\mathbb{E}_{\gamma_n}(\|x\|_2)\right)^2}{n} \underset{n\to+\infty}{\longrightarrow} 1.$$

(II) De forma análoga a la expuesta en el Lema C.0.1 puede verse que en general, para cada $1 \leq p < \infty$ existe una constante $K(p) > 0$ tal que

$$\mathbb{E}_{\gamma_n}(\|x\|) \leq \left(\mathbb{E}_{\gamma_n}(\|x\|^p)\right)^{1/p} \leq K(p)\mathbb{E}_{\gamma_n}(\|x\|).$$

Notar que dicha constante es independiente de la norma $\| \; \|$. Ver, por ejemplo, [Pis89, Capítulo 4].

Demostremos ahora el Teorema 3.4.8, cuyo enunciado no es más que el Teorema 3.1.3, con las constantes concretas obtenidas en la demostración.

**Teorema 3.4.8.** *Sea $X = (\mathbb{R}^n, \| \; \|)$ un espacio normado. Entonces existe un isomorfismo $S : \mathbb{R}^n \to \mathbb{R}^n$ tal que para cada $0 < t \leq 1$ se tiene que*

$$\gamma_n\left(\left|\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|)\right| > t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right) \leq 14\exp\left(\frac{-1}{310}t\ln(n)\right).$$

*Demostración.* Dada la naturaleza del enunciado, podemos suponer que $B_X$ esta en posición de John.

Supongamos primero que $k(X) < n^{1/10}$, de manera que considerando $\delta = 2/5$ en la Proposición 3.4.7, tenemos que existen un subespacio $F \subseteq X$ de dimensión $m$, con $\sqrt{n}/128 \leq m \leq n/2$ y una base $\{v_1, \ldots, v_m\}$ de $F$ tal que

$$\mu\left(\left|\left\|\sum_{i=1}^m g_i\, v_i\right\| - \mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, v_i\right\|\right)\right| > \mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, v_i\right\|\right) t\right) \leq 6\exp\left(-\frac{1}{62}\frac{2}{5}\, t\ln(n)\right) \tag{C.3}$$
$$= 6\exp\left(-\frac{1}{155}\, t\ln(n)\right) \quad \forall\, 0 < t \leq 1,$$

donde $(\Omega, \Sigma, \mu)$ es un espacio de probabilidad, y $g_1, \ldots, g_n : \Omega \to \mathbb{R}$ son variables aleatorias gaussianas estándar e independientes.

Sea ahora $\{u_1, \ldots, u_n\}$ la base ortonormal de $l_2^n$ utilizada en la Proposición 3.4.7, arreglada de manera tal que $F = \langle u_i \rangle_{1 \leq i \leq m}$.

Afirmamos entonces que el isomorfismo $S : \mathbb{R}^n \to \mathbb{R}^n$ dado por $S(u_i) = v_i$ para $1 \leq i \leq m$ y $S(u_i) = \alpha u_i$ para $m < i \leq n$, con

$$\alpha = \frac{\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, v_i\right\|\right)}{\mathbb{E}_\mu\left(\left\|\sum_{i=m+1}^n g_i\, u_i\right\|_2\right)\ln(n)},$$

satisface lo enunciado.

Notemos que, llamando $T : F \to F$ al isomorfismo que cumple $T(u_i) = v_i$ para cada $1 \leq i \leq m$, tenemos que

$$S(f + f^\perp) = T(f) + \alpha f^\perp$$

para cada $f \in F, f^\perp \in F^\perp$.

Veamos ahora que

$$\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right) \leq \mathbb{E}_\mu\left(\left\|\sum_{i=1}^n g_i\, S(u_i)\right\|\right) \leq \left(1 + \frac{1}{\ln(n)}\right)\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right). \tag{C.4}$$

En efecto, usando la invariancia de la medida gaussiana en $\mathbb{R}^n$ por transformaciones ortogonales, vemos que

$$\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right) \leq \frac{1}{2}\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i) + \sum_{i=m+1}^n g_i\, S(u_i)\right\| + \left\|\sum_{i=1}^m g_i\, S(u_i) + \sum_{i=m+1}^n -g_i\, S(u_i)\right\|\right)$$
$$= \frac{1}{2}\left(\mathbb{E}_{\gamma_n}\left(\left\|\sum_{i=1}^m x_i\, S(u_i) + \sum_{i=m+1}^n x_i\, S(u_i)\right\|\right) + \mathbb{E}_{\gamma_n}\left(\left\|\sum_{i=1}^m x_i\, S(u_i) + \sum_{i=m+1}^n -x_i\, S(u_i)\right\|\right)\right)$$
$$= \mathbb{E}_{\gamma_n}\left(\left\|\sum_{i=1}^m x_i\, S(u_i) + \sum_{i=m+1}^n x_i\, S(u_i)\right\|\right) = \mathbb{E}_\mu\left(\left\|\sum_{i=1}^n g_i\, S(u_i)\right\|\right).$$

Por otro lado, usando la desigualdad triangular y que $\|z\| \leq \|z\|_2$ para todo $z \in \mathbb{R}^n$, al estar $B_X$ en posición de John, tenemos que

$$\begin{aligned}
\mathbb{E}_\mu\left(\left\|\sum_{i=1}^n g_i\, S(u_i)\right\|\right) &\leq \mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right) + \alpha \mathbb{E}_\mu\left(\left\|\sum_{i=m+1}^n g_i\, u_i\right\|\right) \\
=&\mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right)\left(1 + \frac{\mathbb{E}_\mu\left(\left\|\sum_{i=m+1}^n g_i\, u_i\right\|\right)}{\mathbb{E}_\mu\left(\left\|\sum_{i=m+1}^n g_i\, u_i\right\|_2\right)\ln(n)}\right) \leq \mathbb{E}_\mu\left(\left\|\sum_{i=1}^m g_i\, S(u_i)\right\|\right)\left(1 + \frac{1}{\ln(n)}\right).
\end{aligned}$$

Así, tenemos (C.4).

Para acotar $\gamma_n\left(\left|\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|)\right| > t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right)$, supondremos primero que $t > 6/\ln(n)$, y acotaremos por separado

$$\underbrace{\gamma_n\left(\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|) > t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right)}_{(i)} \quad \text{y} \quad \underbrace{\gamma_n\left(\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|) < -t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right)}_{(ii)}.$$

Comencemos acotando $(i)$. Por brevedad en la notación, dado un punto $x \in \mathbb{R}^n$ escribimos $f \in F$ y $f^\perp \in F^\perp$ por $P_F(x)$, la proyección ortogonal de $x$ sobre $F$, y $P_{F^\perp}(x)$, la proyección ortogonal de $x$ sobre $F^\perp$, respectivamente.

$$\begin{aligned}
(i) =&\gamma_n\left(\|S(x)\| > (1+t)\mathbb{E}_{\gamma_n}(\|S(x)\|)\right) \leq \gamma_n\left(\|T(f)\| + \alpha\left\|f^\perp\right\| > (1+t)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right) \\
\leq&\gamma_n\left(\|T(f)\| + \alpha\left\|f^\perp\right\|_2 > (1+t)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right) \\
=&\gamma_n\bigg(\left\{\|T(f)\| > (1+t)\mathbb{E}_{\gamma_n}(\|T(f)\|) - \alpha\left\|f^\perp\right\|_2\right\} \cap \left\{\left\|f^\perp\right\|_2 \leq 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right\}\bigg) \\
+&\gamma_n\bigg(\left\{\|T(f)\| > (1+t)\mathbb{E}_{\gamma_n}(\|T(f)\|) - \alpha\left\|f^\perp\right\|_2\right\} \cap \left\{\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right\}\bigg) \\
\leq&\underbrace{\gamma_n\left(\|T(f)\| > (1+t)\mathbb{E}_{\gamma_n}(\|T(f)\|) - 2\alpha\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right)}_{(*)} + \underbrace{\gamma_n\left(\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right)}_{(**)}
\end{aligned}$$

Acotemos $(*)$ y $(**)$ por separado. Notemos para eso que mediante el Lema 2.2.1 tenemos que $\xi_1 : (\mathbb{R}^n, \gamma_n) \to \mathbb{R}$ y $\xi_2 : (\Omega, \mu) \to \mathbb{R}$ dadas por

$$\xi_1(x) = \left\|f^\perp\right\|_2 = \left\|\sum_{i=m+1}^n \langle x, u_i\rangle\, u_i\right\|_2 \quad \text{y} \quad \xi_2(\omega) = \left\|\sum_{i=m+1}^n g_i(\omega)\, u_i\right\|_2$$

tienen la misma distribución, y que $\zeta_1 : (\mathbb{R}^n, \gamma_n) \to \mathbb{R}$ y $\zeta_2 : (\Omega, \mu) \to \mathbb{R}$ dadas por

$$\zeta_1(x) = \|T(f)\| = \left\|T\left(\sum_{i=1}^m \langle x, u_i\rangle\, u_i\right)\right\| = \left\|\sum_{i=1}^m \langle x, u_i\rangle\, v_i\right\| \quad \text{y} \quad \zeta_2(\omega) = \left\|\sum_{i=1}^m g_i(\omega)\, v_i\right\| \tag{C.5}$$

son igualmente distribuidas. Así, usando que $t > 6/\ln(n)$ y (C.3) nos queda que

$$\begin{aligned}
(*) =&\gamma_n\left(\|T(f)\| > \left(1 + t - \frac{2}{\ln(n)}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right) \leq \gamma_n\left(\|T(f)\| > \left(1 + \frac{2}{3}t\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right) \\
\leq& 6\exp\left(-\frac{1}{155}\,\frac{2}{3}t\ln(n)\right) = 6\exp\left(-\frac{2}{465}\,t\ln(n)\right).
\end{aligned}$$

Por otro lado, usando que $m \leq n/2$, la desigualdad (C.2), el Lema C.0.1 y notando como $z = (z_1 \ldots, z_{n-m})$ a los puntos de $\mathbb{R}^{n-m}$ tenemos que

$$\begin{aligned}
(**) =& \gamma_n\left(\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right) = \mu\left(\left\|\sum_{i=m+1}^{n} g_i(\omega)\,u_i\right\|_2 > 2\mathbb{E}_\mu\left(\left\|\sum_{i=m+1}^{n} g_i(\omega)\,u_i\right\|_2\right)\right)\\
=& \gamma_{n-m}\left(\|z\|_2 > 2\mathbb{E}_{\gamma_{n-m}}(\|z\|_2)\right) = \gamma_{n-m}\left(\|z\|_2 - \mathbb{E}_{\gamma_{n-m}}(\|z\|_2) > \mathbb{E}_{\gamma_{n-m}}(\|z\|_2)\right)\\
\leq& \exp\left(-\frac{2}{\pi^2}\,k\,(l_2^{n-m})\,1^2\right) \leq \exp\left(-\frac{2}{\pi^2}\,\frac{n-m}{14}\right) \leq \exp\left(-\frac{1}{14\pi^2}\,n\right) \leq \exp\left(-\frac{2}{465}\,t\ln(n)\right),
\end{aligned}$$

donde la última desigualdad vale si $t \leq \frac{465}{28\pi^2}\frac{n}{\ln(n)}$. Pero como esta última expresión es mayor a 1 para todo $n \geq 2$, tenemos que

$$(**) = \gamma_n\left(\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right) \leq \exp\left(-\frac{2}{465}\,t\ln(n)\right) \quad \forall\, 0 < t \leq 1,\, n \geq 2.$$

Luego, llegamos a que

$$(i) = \gamma_n\left(\|S(x)\| > (1+t)\mathbb{E}_{\gamma_n}(\|S(x)\|)\right) \leq 7\exp\left(-\frac{2}{465}\,t\ln(n)\right).$$

Ahora acotemos $(ii)$.

$$\begin{aligned}
(ii) =& \gamma_n\left(\|S(x)\| < (1-t)\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right) \leq \gamma_n\left(\|T(f)\| - \alpha\left\|f^\perp\right\| < (1-t)\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\right)\\
\leq& \gamma_n\left(\|T(f)\| - \alpha\left\|f^\perp\right\|_2 < (1-t)\left(1+\frac{1}{\ln(n)}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right)\\
=& \gamma_n\left(\|T(f)\| < (1-t)\left(1+\frac{1}{\ln(n)}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|) + \frac{\|f^\perp\|_2}{\ln(n)\,\mathbb{E}_{\gamma_n}(\|f^\perp\|_2)}\mathbb{E}_{\gamma_n}(\|T(f)\|)\right)\\
\leq& \gamma_n\left(\left\{\|T(f)\| < \left((1-t)\left(1+\frac{1}{\ln(n)}\right)+\frac{2}{\ln(n)}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right\} \cap \left\{\left\|f^\perp\right\|_2 \leq 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right\}\right)\\
&+ \gamma_n\left(\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right)\\
\leq& \underbrace{\gamma_n\left(\|T(f)\| < \left((1-t)\left(1+\frac{1}{\ln(n)}\right)+\frac{2}{\ln(n)}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right)}_{(\star)} + \underbrace{\gamma_n\left(\left\|f^\perp\right\|_2 > 2\mathbb{E}_{\gamma_n}\left(\left\|f^\perp\right\|_2\right)\right)}_{(\star\star)}.
\end{aligned}$$

Ya acotamos anteriormente $(\star\star)$ (que en su momento supo ser $(**)$), por lo que solo nos resta por acotar $(\star)$.

Para esto, notemos que al ser $t > 6/\ln(n) > 6/(\ln(n)+2)$, entonces vale que

$$(1-t)\left(1+\frac{1}{\ln(n)}\right)+\frac{2}{\ln(n)} < 1 - \frac{t}{2}.$$

Por lo que, usando nuevamente (C.3) y que $\zeta_1$ y $\zeta_2$ definidas en (C.5) son igualmente distribuidas, vemos que

$$(\star) \leq \gamma_n\left(\|T(f)\| < \left(1-\frac{t}{2}\right)\mathbb{E}_{\gamma_n}(\|T(f)\|)\right) \leq 6\exp\left(-\frac{1}{155}\,\frac{1}{2}\,t\ln(n)\right) = 6\exp\left(\frac{-1}{310}\,t\ln(n)\right).$$

De esta manera, llegamos a que

$$
\begin{aligned}
(ii) = \gamma_n\Big(\|S(x)\| < (1-t)\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\Big) &\leq 6\exp\left(\frac{-1}{310}\,t\ln(n)\right) + \exp\left(\frac{-2}{465}\,t\ln(n)\right) \\
&\leq 7\exp\left(\frac{-1}{310}\,t\ln(n)\right).
\end{aligned}
$$

Finalmente, tenemos que si $t > 6/\ln(n)$, entonces

$$
\begin{aligned}
\gamma_n\Big(\Big|\|S(x)\| - \mathbb{E}_{\gamma_n}(\|S(x)\|)\Big| > t\,\mathbb{E}_{\gamma_n}(\|S(x)\|)\Big) \leq \quad (i) \quad + \quad (ii) \\
\leq 7\exp\left(-\frac{2}{465}\,t\ln(n)\right) + 7\exp\left(\frac{-1}{310}\,t\ln(n)\right) \leq 14\exp\left(\frac{-1}{310}\,t\ln(n)\right).
\end{aligned}
$$

Ahora bien, si $t \leq 6/\ln(n)$, entonces

$$
14\exp\left(\frac{-1}{310}\,t\ln(n)\right) \geq 14\exp\left(\frac{-6}{310}\right) > 1,
$$

por lo que la desigualdad buscada vale trivialmente.

Supongamos ahora que $k(X) \geq n^{1/10}$. En tal caso, tomamos como isomorfismo $S : \mathbb{R}^n \to \mathbb{R}^n$ a la identidad, y usando la Proposición 3.1.1 vemos que

$$
\gamma_n\Big(\Big|\|x\| - \mathbb{E}_{\gamma_n}(\|x\|)\Big| > t\,\mathbb{E}_{\gamma_n}(\|x\|)\Big) \leq 2\exp\left(-\frac{2}{\pi^2}t^2 k(X)\right) \leq 2\exp\left(-\frac{2}{\pi^2}t^2 n^{1/10}\right) \leq 14\exp\left(\frac{-1}{310}\,t\ln(n)\right),
$$

donde la última desigualdad vale si

$$
0 \leq \ln(7) - \frac{1}{310}\,t\ln(n) + \frac{2}{\pi^2}\,t^2\,n^{1/10}.
$$

A su vez, esta última desigualdad vale si

$$
0 \leq -\frac{1}{310}\,t\ln(n) + \frac{2}{\pi^2}\,t^2\,n^{1/10} \iff \frac{\pi^2}{620}\,\frac{\ln(n)}{n^{1/10}} \leq t.
$$

Ahora, si $t < \frac{\pi^2}{620}\frac{\ln(n)}{n^{1/10}}$, entonces la desigualdad buscada también vale trivialmente, ya que

$$
14\exp\left(\frac{-1}{310}\,t\ln(n)\right) \geq 14\exp\left(\frac{-1}{310}\,\frac{\pi^2}{620}\,\frac{\ln(n)}{n^{1/10}}\,\ln(n)\right) \geq 14\exp\left(\frac{-1}{310}\,\frac{\pi^2}{620}\,\frac{400}{e^2}\right) > 13 > 1,
$$

donde usamos que $\frac{\ln^2(n)}{n^{1/10}} \leq 400/e^2$ si $n \geq 1$. □

**Observación C.0.3.** Sobre la demostración anterior, destaquemos que en el caso "$k(X)$ grande" (o un poco más precisamente, en el caso en que $k(X)$ sea del orden de $\dim(X)^\theta$ para algún $\theta > 0$ fijo) podemos tomar como posición de $B_X$ para que se cumpla el Teorema 3.1.3 a la posición de John. Sin embargo, en el caso $k(X)$ chico (digamos por ejemplo, $k(X)$ del orden de $\ln(\dim(X))$) utilizamos una posición de $B_X$ que, en última instancia, esta ligada al Lema 3.4.5 y no esta caracterizada, en principio, por ninguna condición extremal (comparar con la Proposición 3.1.2, la cual siempre puede obtenerse poniendo a $B_X$ en posición de John).

# Bibliografía


[AK16] F. Albiac y N.J. Kalton. *Topics in Banach Space Theory*. Graduate Texts in Mathematics. Springer International Publishing, 2016.

[AM83] Noga Alon y Vitali Milman. «Embedding of $l_\infty^k$ in finite dimensional Banach spaces». En: *Israel Journal of Mathematics* 45 (1983), págs. 265-280.

[AGM15] Shiri Artstein-Avidan, Apostolos Giannopoulos y V. D. Milman. «Asymptotic Geometric Analysis, Part I». En: 2015.

[AGM21] Shiri Artstein-Avidan, Apostolos Giannopoulos y Vitali D. Milman. *Asymptotic geometric analysis. Part II*. Vol. 261. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, [2021] ©2021, págs. xxxvii+645.

[Ban32] Stefan Banach. *Théorie des opérations linéaires*. French. 1932. URL: http://eudml.org/doc/268537.

[Bir35] Garrett Birkhoff. «Integration of functions with values in a Banach space». En: *Transactions of the American Mathematical Society* 38 (1935), págs. 357-378. URL: https://api.semanticscholar.org/CorpusID:122611526.

[BLM13] Stéphane Boucheron, Gábor Lugosi y Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[BL89] J. Bourgain y J. Lindenstrauss. «Almost euclidean sections in spaces with a symmetric basis». En: *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1987–88*. Ed. por Joram Lindenstrauss y Vitali D. Milman. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, págs. 278-288.

[BS88] Jean Bourgain y Stanislaw J. Szarek. «The Banach-Mazur distance to the cube and the Dvoretzky-Rogers factorization». En: *Israel Journal of Mathematics* 62 (1988), págs. 169-180.

[Cha14] Sourav Chatterjee. *Superconcentration and related topics*. Springer Monographs in Mathematics. Springer, Cham, 2014, págs. x+156.

[CL12] Dario Cordero-Erausquin y Michel Ledoux. «Hypercontractive Measures, Talagrand's Inequality, and Influences». En: *Geometric Aspects of Functional Analysis: Israel Seminar 2006–2010*. Ed. por Bo'az Klartag, Shahar Mendelson y Vitali D. Milman. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, págs. 169-189.

[DU77] J. Diestel y J.J. Uhl. *Vector Measures*. Mathematical surveys and monographs. American Mathematical Society, 1977.

[DJT95] Joe Diestel, Hans Jarchow y Andrew Tonge. *Absolutely summing operators*. Vol. 43. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1995, págs. xvi+474.

[Dir37] PG Lejeune Dirichlet. «Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält». En: *Abhandlungen der Königlich Preussischen Akademie der Wissenschaften* 45 (1837), pág. 81.

[Dvo59] A Dvoretzky. «A theorem on convex bodies and applications to Banach spaces». En: *Proceedings of the National Academy of Sciences of the United States of America* 45.2 (feb. de 1959), págs. 223-226. URL: https://europepmc.org/articles/PMC222539.

[Dvo61] A. Dvoretzky. «Some results on convex bodies and Banach spaces». English. En: (1961).

[DR50] A. Dvoretzky y C. Ambrose Rogers. «Absolute and Unconditional Convergence in Normed Linear Spaces.» En: *Proceedings of the National Academy of Sciences of the United States of America* 36 3 (1950), págs. 192-197.

[Eva15] L.C. Evans. *Measure Theory and Fine Properties of Functions, Revised Edition*. Textbooks in Mathematics. CRC Press, 2015.

[Fig72] T. Figiel. «Some remarks on Dvoretzky's theorem on almost spherical sections of convex bodies». En: *Colloquium Mathematicae* 24.2 (1972), págs. 241-252. URL: http://eudml.org/doc/264051.

[Fig76] T. Figiel. «A short proof of Dvoretzky's theorem on almost spherical sections of convex bodies». En: *Compositio Mathematica* 33.3 (1976), págs. 297-301.

[FLM77] T. Figiel, J. Lindenstrauss y V. D. Milman. «The dimension of almost spherical sections of convex bodies». En: *Acta Mathematica* 139 (1977), págs. 53-94. URL: https://doi.org/10.1007/BF02392234.

[Gor85] Y. Gordon. «Some inequalities for Gaussian processes and applications». En: *Israel Journal of Mathematics* 50 (1985), págs. 265-289.

[Gro56] A. Grothendieck. «Sur certaines classes de suites dans les espaces de Banach et le théorème de Dvoretzky-Rogers». French. En: *Bol. Soc. Mat. São Paulo* 8 (1956), págs. 81-110.

[Gro55] Alexandre Grothendieck. *Produits tensoriels topologiques et espaces nucléaires*. Vol. 16. American Mathematical Society Providence, 1955.

[Joh48] Fritz John. «Extremum problems with inequalities as subsidiary conditions». En: *Studies and Essays : Courant Anniversary Volume*. John Wiley & Sons : Interscience Division, 1948, págs. 187-204.

[KN23] Bo'az Klartag y Tomer Novikov. *A weak version of the $\varepsilon$-Dvoretzky conjecture for normed spaces*. 2023. arXiv: 2206.13273 [math.FA].

[Lin64] Joram Lindenstrauss. «Extension of compact operators». En: *Mem. Amer. Math. Soc.* 48 (1964), pág. 112.

[LT16] J. Lopez-Abad y P. Tradacete. «Bases of random unconditional convergence in Banach spaces». En: *Trans. Amer. Math. Soc.* 368.12 (2016), págs. 9001-9032. URL: https://doi.org/10.1090/tran/6636.

[MP76] Bernard Maurey y Gilles Pisier. «Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach». En: *Studia Mathematica* 58 (1976), págs. 45-90. URL: https://api.semanticscholar.org/CorpusID:117227369.

[Mil71] V. D. Milman. «New proof of the theorem of A. Dvoretzky on intersections of convex bodies». En: *Functional Analysis and Its Applications* 5 (1971), págs. 288-295.

[Mil88] V. D. Milman. «A few observations on the connections between local theory and some other fields». En: *Geometric Aspects of Functional Analysis*. Ed. por Joram Lindenstrauss y Vitali D. Milman. Berlin, Heidelberg: Springer Berlin Heidelberg, 1988, págs. 283-289.

[MS86] Vitali D. Milman y Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*. Vol. 1200. Lecture Notes in Mathematics. With an appendix by M. Gromov. Springer-Verlag, Berlin, 1986, págs. viii+156. ISBN: 3-540-16769-2.

[PV18] Grigoris Paouris y Petros Valettas. «Dichotomies, structure, and concentration in normed spaces». En: *Advances in Mathematics* 332 (2018), págs. 438-464. URL: https://www.sciencedirect.com/science/article/pii/S000187081830197X.

[Pis89] Gilles Pisier. *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge Tracts in Mathematics. Cambridge University Press, 1989.

[Rie67] Bernhard Riemann. *Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe*. ger. Göttingen: Dieterich, 1867. URL: http://eudml.org/doc/203787.

[Rui06] José Carlos Rodríguez Ruiz. «Integración en espacios de Banach». En: 2006. URL: https://webs.um.es/beca/Investigacion/TesisJose.pdf.

[Rus82] Lucio Russo. «An approximate zero-one law». En: *Z. Wahrsch. Verw. Gebiete* 61.1 (1982), págs. 129-139. URL: https://doi.org/10.1007/BF00537230.

[Sch89] Gideon Schechtman. «A remark concerning the dependence on $\varepsilon$ in Dvoretzky's theorem». En: 1989, págs. 274-277.

[Sch06] Gideon Schechtman. «Two observations regarding embedding subsets of Euclidean spaces in normed spaces». En: *Advances in Mathematics* 200.1 (2006), págs. 125-135. URL: https://www.sciencedirect.com/science/article/pii/S0001870804003652.

[Sch13] Rolf Schneider. *Convex Bodies: The Brunn–Minkowski Theory*. 2.ª ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2013.

[Sch21] J. Schur. «Über lineare Transformationen in der Theorie der unendlichen Reihen.» En: *Journal für die reine und angewandte Mathematik* 1921.151 (1921), págs. 79-111.

[SBR02] Ageev Sergei, Semeon Bogatyi y Dušan Repovš. «The Banach-Mazur compactum is the Alexandroff compactification of a Hilbert cube manifold». En: (sep. de 2002). DOI: 10.48550/arXiv.math/0209361.

[Tal93] Michel Talagrand. «Regularity of Infinitely Divisible Processes». En: *The Annals of Probability* 21.1 (1993), págs. 362-432. URL: https://doi.org/10.1214/aop/1176989409.

[Tal94] Michel Talagrand. «On Russo's Approximate Zero-One Law». En: *The Annals of Probability* 22.3 (1994), págs. 1576-1587. URL: https://doi.org/10.1214/aop/1176988612.

[Tal95] Michel Talagrand. «Embedding of $l_\infty^k$ and a Theorem of Alon and Milman». En: *Geometric Aspects of Functional Analysis*. Ed. por J. Lindenstrauss y V. Milman. Basel: Birkhäuser Basel, 1995, págs. 289-293.

[Tom89] Nicole Tomczak-Jaegermann. «Banach-Mazur distances and finite-dimensional operator ideals». En: 1989. URL: https://sites.ualberta.ca/~ntj/bm_book/index.html.