



Departamento de Matemática, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires

Geometría de Finsler de grupos con distancias bi-invariantes

Tesis presentada para optar por el título de Doctor de la Universidad de Buenos Aires en el área Ciencias Matemáticas.

Iván Rey

Director: Dr. Gabriel Larotonda.

Consejero de estudios: Dr. Daniel Carando.

Lugares de trabajo: F.C.E.N., Universidad de Buenos Aires.

Instituto Argentino de Matemática (I.A.M.).

Buenos Aires 24 de octubre del 2025

Geometría de Finsler de grupos con distancias bi-invariantes

El objetivo de esta tesis es estudiar la geometría de Finsler de grupos de Lie conexos G equipados con una distancia bi-invariante

$$\text{dist}(kg, kh) = \text{dist}(g, h) = \text{dist}(gk, hk) \quad \forall g, h, k \in G$$

En el capítulo 2 estudiamos las normas de Finsler Ad-invariantes sobre las álgebras de Lie $\mathfrak{u}_n$ y $\mathfrak{su}_n$, las cuales inducen distancias bi-invariantes sobre los grupos U_n y SU_n respectivamente. Demostramos que dichas normas equivalen, vía cortes mediante álgebras de Lie abelianas maximales (MAB), a normas simétricas en $\mathbb{R}^n$. Esto nos permite ver ejemplos y propiedades de las mismas con mayor facilidad. Exhibimos además una familia de normas de Finsler Ad-invariantes las cuales llamamos *normas orbitales*, las cuales probamos que nos sirven de normas de control sobre la familia de todas las normas Ad-invariantes. Sobre el final del mismo probamos que esencialmente todas las normas de Finsler Ad-invariantes que provienen de un producto interno son, en $\mathfrak{su}_n$, múltiplos de la norma de Frobenius.

En el capítulo 3 estudiamos los funcionales que realizan la norma de un elemento $0 \neq V \in \mathfrak{u}_n$, probando que si $\varphi(V) = |V|$ entonces

$$\varphi([X, [X, V]]) \leq 0 \quad \forall X \in \mathfrak{u}_n,$$

estudiamos además el caso de la igualdad llegando a la conclusión de que esta ocurre si y solo si $[X_C, N] = 0$, donde X_C es la parte co-diagonal de X respecto a V y N es tal que $\varphi(\cdot) = (\cdot|N)_F$ (el producto interno de Frobenius). Vemos además que en el caso de que la norma sea estrictamente convexa, la igualdad se cumple si y solo si $[X, V] = 0$. Introducimos también la noción de dual polar, diferenciación de normas y operadores disipativos, las cuales utilizamos para estudiar cuando vale la igualdad

$$|V + [X, [X, V]]| = |V|,$$

obteniendo que esta misma ocurre cuando V y $V + [X, [X, V]]$ pertenecen a la misma cara de la esfera.

En el capítulo 4 mostramos que la una distancia bi-invariante en U_n obtenida a partir de una norma de Finsler Ad-invariante en $\mathfrak{u}_n$ satisface la desigualdad $d(e^X, e^Y) \leq |Y - X|$ y estudiamos el caso de la igualdad. Introducimos en este mismo capítulo una noción de curvatura seccional basada en el trabajo de Milnor y probamos que para dicha noción vale $\text{sec}(X, Y) = 0$ si y solo si $[X, Y] = 0$ para el caso de normas estrictamente convexas. Estudiamos también el caso general donde obtenemos resultados conectados con las desigualdades probadas en el capítulo previo.

En el capítulo 5 generalizamos los anteriores a grupos de Lie probando que un grupo de Lie G admite una distancia bi-invariante si y solo si su álgebra de Lie es reductiva. Esto nos da una descomposición de G como el producto de un grupo compacto y uno abeliano, extendiendo así un resultado de Milnor en el cual él prueba lo mismo pero para el caso en el cual la distancia proviene de una métrica riemanniana en el grupo de Lie.

Palabras clave: norma de Finsler Ad-invariante, mayorización, funcional que realiza la norma, métrica bi-invariante, curvatura, métrica de Finsler, geodésica, grupo de Lie, norma unitariamente invariante.

Finsler geometry of groups with bi-invariant distances

The objective of this thesis is to study the Finsler geometry of connected Lie groups G equipped with a bi-invariant distance

$$d(kg, kh) = d(g, h) = d(gk, hk) \quad \forall g, h, k \in G.$$

In Chapter 2, we study Ad-invariant Finsler norms on the Lie algebras $\mathfrak{u}_n$ and $\mathfrak{su}_n$, which induce bi-invariant distances on the groups U_n and SU_n , respectively. We prove that these norms are equivalent, via cuts through maximal abelian Lie subalgebras (MABs), to symmetric norms on $\mathbb{R}^n$. This allows us to more easily see examples and properties of them. We also exhibit a family of Ad-invariant Finsler norms that we call *orbit norms*, which we prove serve as controlling norms over the family of all Ad-invariant norms. Toward the end of the chapter, we prove that essentially all Ad-invariant Finsler norms that arise from an inner product are, on $\mathfrak{su}$, multiples of the Frobenius norm.

In Chapter 3, we study the norming functionals for a nonzero element $V \in \mathfrak{u}_n$, proving that if $\varphi(V) = |V|$ then

$$\varphi([X, [X, V]]) \leq 0 \quad \forall X \in \mathfrak{u}_n,$$

and we also study the case of equality, reaching the conclusion that it holds if and only if $[X_C, N] = 0$, where X_C is the co-diagonal part of X with respect to V and N is such that $\varphi(\cdot) = (\cdot|N)_F$ (the Frobenius inner product). We also see that in the case where the norm is strictly convex, the equality holds if and only if $[X, V] = 0$. In this chapter, we also introduce the notion of dual polar, norm differentiation, and dissipative operators, which we use to study when the equality

$$|V + [X, [X, V]]| = |V|,$$

holds, obtaining that this occurs when V and $V + [X, [X, V]]$ belong to the same face of the sphere.

In Chapter 4, we show that a bi-invariant distance in U_n obtained from an Ad-invariant Finsler norm in $\mathfrak{u}_n$ satisfies the inequality $d(e^X, e^Y) \leq |Y - X|$ and we study the case of equality. We introduce a notion of sectional curvature based on Milnor's work and prove that for this notion, $\sec(X, Y) = 0$ if and only if $[X, Y] = 0$ in the case of strictly convex norms. We also study the general case, where we obtain results connected to the inequalities proved in the previous chapter.

In Chapter 5, we generalize the previous results to Lie groups proving that a Lie group G admits a bi-invariant distance if and only if its Lie algebra is reductive. This gives us a decomposition of G as the product of a compact group and an abelian one, thus extending a result of Milnor in which he proves the same but for the case where the distance comes from a Riemannian metric on the Lie group.

Keywords: Ad-invariant Finsler norm, majorization, norming functional, bi-invariant metric, curvature, Finsler metric, geodesic, Lie group, unitarily invariant norm.

Agradecimientos

A Gabriel, por enseñarme tanta matemática durante todos estos años poniéndole siempre buena onda y empujando para adelante incluso en los momentos más complicados de este trabajo. Realmente, además de ser un gran director, se mostró conmigo como la excelente persona que es. En serio, gracias.

A Dani, por ser mi consejero de estudios y estar siempre ahí cuando necesitaba una mano.

A la UBA y a todos los grandes docentes que tuve en esta institución.

Al CONICET por otorgarme una beca que me permitió hacer esta tesis.

A mi querida madre por acompañarme en todas mis decisiones y ser tan fuerte en todos los momentos difíciles que pasamos estos últimos años. El hecho de tenerte acá presente vale oro.

A mi viejo por transmitirme su tranquilidad y por su gigante ayuda cuando mi madre estaba mal.

A mi abuelita eterna de mi alma, que me pidió que me apure en licenciarme para estar presente y sigue hoy día a mi lado siendo una de las personas más importantes en mi vida.

Al resto de mi querida familia: mi hermano Valen, mi prima Lu, mi primo Fer, mi madrastra Patri, mis tíos Cynthia y Roberto. Los quiero mucho.

A mi hermano de doctorado Pepe por estar siempre al lado mío en esta travesía y a todos mis otros compañeros de la docencia y doctorado de la UBA.

A todos mis compañeros de trabajo de UTDT, que hacen de dicho espacio un lugar de pertenencia hermoso. En especial a Romi y Anto por recibirme ahí y mostrarme su hermosa calidad humana, y a Vicky por generar en gran parte este espacio grupal y estar siempre presente cuidando y ayudando al otro.

A mi hermanita y motorcito Emi que en estos últimos años pasó a ser una gran parte de mi vida. Gracias a todo el cariño y el apoyo que me transmitió, pude ganar confianza en mi mismo y darle cierre a este proyecto. Realmente estoy agradecido de tener una Memi en mi vida.

A mi hermanita Luji por producir la base de mi alimentación durante este doctorado.

A mis amigos del secundario Brian y Martín por toda la buena onda y los grandes asados de estos años.

A mis amigos del rock que me ayudaron a desconectar cuando lo necesitaba.

A Nanchu por haber estado a mi lado en una buena parte de este trayecto.

A Jesi y Andrés por cuidar de mi salud mental todos estos años.

A Natrash, esa ardilla roja que es, fue y será mi chamix de la vida siempre.

Índice general

1. Introducción	9
2. Preliminares	11
2.1. Normas de Finsler	11
2.2. Normas de Finsler Ad-invariantes en $\mathfrak{u}_n$	20
2.3. Cuerpos convexos y normas.	29
3. Funcionales, conmutadores y dual polar	37
3.1. Caracterización de funcionales que realizan la norma	37
3.2. Derivada de una norma de Finsler	48
3.3. El dual polar y sus propiedades	52
3.4. Operadores disipativos	64
3.5. El caso de la igualdad	66
4. Curvatura del Grupo Unitario	69
4.1. Un poco de Geometría Riemanniana	69
4.2. Grupos de Lie	76
4.3. Geodésicas y distancia en $\mathcal{U}_n$	80
4.4. Distancia y curvatura	87
4.5. Curvatura seccional	95
5. Álgebras de Lie con una norma de Finsler Ad-invariante	101
5.1. Propiedades básicas de la forma de Killing de un álgebra de Lie	101
5.2. Operadores Hermitianos, rango numérico y espectro	105
5.3. Generalización de los resultados	111
Bibliografía	125

Capítulo 1

Introducción

En el trabajo de Milnor titulado “Curvatures of left invariant metrics on Lie groups”[55] y publicado en 1976, él estudia entre otras cosas la geometría de un grupo de Lie G cuando este mismo posee una métrica riemanniana bi-invariante. Entre otras cosas, él prueba que en este caso, el grupo debe ser isomorfo a un subgrupo de $\mathbb{R}^k \times \mathcal{U}_n$, donde $\mathcal{U}_n$ es el grupo de matrices unitarias de $n \times n$ con coeficientes complejos y $k \in \mathbb{N}$. Esto motiva la pregunta natural de qué sucede en el caso de que la métrica no sea riemanniana y esta es la pregunta que motiva este trabajo. Dado el resultado enunciado previamente, el ambiente más razonable para estudiar este problema es el del grupo unitario $\mathcal{U}_n$, cuyo álgebra de Lie $\mathfrak{u}_n$ está formada por las matrices antihermitianas de $n \times n$ con coeficientes complejos. Se desprenden de esta forma dos áreas de estudio íntimamente relacionadas que son: el análisis matricial sobre las normas Ad-invariantes en $\mathfrak{u}_n$ y por otro lado, la geometría de $\mathcal{U}_n$ cuando consideramos la métrica invariante a izquierda obtenida a partir de una norma Ad-invariante en $\mathfrak{u}_n$. En este trabajo estudiamos el enfoque del análisis matricial en los capítulos 2 y 3, en los cuales incluimos los resultados originales publicados en [47], mientras que en los capítulos 4 y 5 nos centramos en la parte geométrica y los resultados publicados en [48].

Desde el punto de vista del análisis matricial, decimos que una norma en $M_n(\mathbb{C})$ es *fuertemente invariante* si cumple $\|UXV\| = \|X\|$ para todo $X \in M_n(\mathbb{C})$ y $U, V \in \mathcal{U}_n$. Estas normas fueron estudiadas y hay muchas publicaciones, tanto teóricas como numéricas, entorno a ellas. De hecho tienen muchas aplicaciones tanto a la teoría de la información cuántica (ver [33, 36, 57]) como al procesamiento de señales vía optimización convexa (ver [1, 9, 30, 64]). Por otro sobre las normas *débilmente invariantes* que son las que satisfacen $\|UXU^*\| = \|X\|$ para todo $X \in M_n(\mathbb{C})$ y $U \in \mathcal{U}_n$, no hay tantos trabajos realizados, pero podemos encontrar estudios sobre las mismas en [15], [45] y [46]. En esta tesis vamos a presentar resultados originales sobre estas normas vistas en $\mathfrak{u}_n$. Dado que en este contexto la acción de conjugación coincide con la acción adjunta del grupo de Lie $\mathcal{U}_n$ sobre su álgebra de Lie, vamos a llamar a estas normas Ad-invariantes. Decidimos trabajar con normas de Finsler Ad-invariantes, es decir, cumplen que $|\lambda X| = \lambda|X|$ para $\lambda > 0$ y no necesariamente para todo $\lambda \in \mathbb{R}$. Esto nos va a dar más generalidad y ejemplos, pero los resultados siguen teniendo interés en el caso de normas completamente homogéneas.

En el Capítulo 2 vamos a introducir en detalle la noción de norma de Finsler Ad-invariante en $\mathfrak{u}_n$ y daremos varios ejemplos. Uno de los más importantes que mostraremos es el de las normas orbitales (ver Proposición 2.2.11), las cuales son una versión de las normas de Hofer introducidas en [46] adaptadas al contexto de matrices. Estas, además de constituir un ejemplo de interés por si mismas, cumplen un rol

muy similar al de las normas Ky-Fan en el caso fuertemente invariante, con lo cual las proponemos como las normas de control en la Observación 2.2.20. Los otros resultados originales de este capítulo son la Proposición 2.3.4, el Teorema 2.3.6 y el Teorema 2.3.8, los cuales nos permiten caracterizar una norma a partir de mirar un corte de la misma con una subálgebra abeliana maximal de $\mathfrak{u}_n$ y $\mathfrak{su}_n$ respectivamente.

En el Capítulo 3 presentaremos los resultados más importantes publicados en [47]. Realizaremos un estudio detallado de los funcionales que realizan la norma de un cierto elemento llegando al Teorema 3.1.13, en el cual caracterizamos los X para los cuales $V + [X, [X, V]]$ pertenece al hiperplano afín que soporta a V en la esfera de la norma (extendiendo así un resultado que es evidente para la norma de Frobenius). Con demostraciones muy similares probamos que $|V + [X, V]| \geq |V|$ para todo $X \in \mathfrak{u}_n$ y cualquier norma de Finsler Ad-invariante en el Teorema 3.4.4. Finalmente estudiamos el caso de la igualdad en el Teorema 3.5.3.

En el capítulo 4, nos vamos a centrar en estudiar el concepto de curvatura seccional en el contexto de una métrica bi-invariante en $\mathcal{U}_n$ inducida por una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Dado que la noción de curvatura seccional definida en una variedad riemanniana depende fuertemente de tener un producto interno, vamos a basarnos nuevamente en Milnor en [54, pag.101], donde describe la curvatura seccional en forma intuitiva sin usar la métrica riemanniana y sólo utilizando distancias y normas. Esto nos lleva a la siguiente igualdad:

$$\langle R_p(x, y)y, x \rangle_p = 6\|y - x\|_p^2 \lim_{r \rightarrow 0^+} \frac{r\|y - x\|_p - \text{dist}(\exp_p(rx), \exp_p(ry))}{r^2 \text{dist}(\exp_p(rx), \exp_p(ry))},$$

la cual discutiremos y demostraremos en la sección 4.1, para luego en la sección 4.5 poder proponer definición para la curvatura seccional para nuestro caso, cuya buena definición es consecuencia del Teorema 4.4.4 (también incluido en [48]). Esta definición ya fue usada a nivel del estudio del signo (no del valor concreto) en varios trabajos, como por ejemplo en [6]. Daremos también un resultado original incluido en el Teorema 4.4.14, en el cual caracterizamos las secciones planas y finalmente aplicaremos los resultados obtenidos para estudiar los ángulos de los triángulos geodésicos en nuestra variedad en el Teorema 4.5.8.

En el Capítulo 5, vamos a trabajar en forma más general y estudiar el caso de un grupo de Lie real y finito dimensional G para el cual su álgebra de Lie admite una norma Ad-invariante. Probaremos los resultados más importantes de [48]: el Teorema 5.2.15, el cual nos dice que si estamos en esta situación entonces $\mathfrak{g}$ tiene que ser un álgebra de Lie reductiva (i.e. la suma de un álgebra de Lie semisimple con forma de Killing definida negativa y su centro) y el Teorema 5.2.16, el cual nos dice que en dicho caso $G = H \times K$ donde H es un subgrupo conexo y abeliano y K es un subgrupo compacto con centro finito.

Finalmente en la última sección de este capítulo, repasaremos cómo se puede volver a interpretar los resultados más importantes de este trabajo en el contexto de un grupo de Lie cuya álgebra de Lie es reductiva. Esto sería básicamente generalizar los resultados probados en $\mathfrak{u}_n$ y en $\mathfrak{su}_n$ a un álgebra de Lie reductiva con una norma de Finsler Ad-invariante. Las demostraciones son un poco más técnicas, ya que dependen de la construcción de un sistema de raíces asociado a una subálgebra de Cartan. Incluiremos las demostraciones que son realmente distintas al caso original de esta tesis, las demás pueden ser encontradas en [48].

Capítulo 2

Preliminares

2.1. Normas de Finsler

En esta sección vamos a fijar notación y probar algunos resultados conocidos para normas homogéneas en el caso de normas de Finsler. Además, incluiremos algunos ejemplos y mostraremos una visualización de cómo pensamos gráficamente a los funcionales que realizan la norma de un cierto elemento.

Definición 2.1.1. Sea E un $\mathbb{R}$ -espacio vectorial de dimensión n . Una norma de Finsler en E es una aplicación $|\cdot| : E \longrightarrow \mathbb{R}_{\geq 0}$ que satisface

- (1) $|X| = 0$ si y solo si $X = 0$ (no degenerada),
- (2) $|X + Y| \leq |X| + |Y|$ (subaditiva),
- (3) $|\lambda X| = \lambda |X|$ para todo $\lambda \in \mathbb{R}_{\geq 0}$ (positivamente homogénea).

Definición 2.1.2 (Estrictamente convexa). Decimos que una norma de Finsler en un espacio vectorial E es *estrictamente convexa* si para $X, Y \neq 0$, la igualdad

$$|X + Y| = |X| + |Y|$$

implica que existe una constante $c > 0$ de forma tal que $Y = cX$.

Es importante observar que toda norma completamente homogénea en E es en particular una norma de Finsler, para facilitar la lectura, notaremos como $|\cdot|$ a las normas de Finsler en E y como $\|\cdot\|$ a las normas en el sentido usual. Cuando trabajemos en E' , el dual de E , notaremos $\|\varphi\|$ a la norma de los funcionales, aclarando previamente si estamos trabajando con una norma usual o con una de Finsler. En ocasiones, donde el contexto sea claro, nos referiremos a las normas Finsler simplemente como normas.

Proposición 2.1.3. Sea E un $\mathbb{R}$ -espacio vectorial de dimensión n y sean $|\cdot|_1$ y $|\cdot|_2$ dos normas de Finsler en E . Entonces, existen constantes positivas C_1 y C_2 de forma tal que

$$C_1 |X|_2 \leq |X|_1 \leq C_2 |X|_2 \quad \forall X \in E,$$

es decir, todas las normas de Finsler son equivalentes.

Demostración. Sean $B = \{E_1, E_2, \dots, E_n\}$ una base de E y $X \in E$, escribimos $X = \sum_{k=1}^n \lambda_k E_k$ y consideramos la norma dada por $\|X\| = \max_k |\lambda_k|$. Notamos que las desigualdades del enunciado definen una relación de equivalencia en las normas de Finsler de E , luego basta probar que toda norma de Finsler $|\cdot|$ es equivalente a la norma $\|\cdot\|$ previamente definida. Comenzamos notando que la segunda desigualdad es consecuencia de elegir $C_2 = \sum_{k=1}^n \max\{|E_k|, |-E_k|\}$ y aplicar la desigualdad triangular:

$$\begin{aligned} |X| &\leq \sum_{k=1}^n |\lambda_k E_k| = \sum_{k=1}^n |\lambda_k| |\operatorname{sg}(\lambda_k) E_k| \leq \sum_{k=1}^n \|X\| |\operatorname{sg}(\lambda_k) E_k| = \left(\sum_{k=1}^n |\operatorname{sg}(\lambda_k) E_k| \right) \|X\| \\ &\leq \left(\sum_{k=1}^n \max\{|E_k|, |-E_k|\} \right) \|X\| = C_2 \|X\|. \end{aligned}$$

Para probar la otra desigualdad, sea $K = \{X \in E : \|X\| = 1\}$, como E es de dimensión finita K resulta compacto en $(E, \|\cdot\|)$. Consideramos la función $|\cdot| : K \rightarrow \mathbb{R}$ y notamos que si $X, Y \in E$ tenemos que $|X| - |Y| \leq |X - Y| \leq C_2 \|X - Y\|$, análogamente $|Y| - |X| \leq C_2 \|Y - X\| = C_2 \|X - Y\|$ y entonces $||X| - |Y|| \leq C_2 \|X - Y\|$. Esto implica que la norma $|\cdot|$ es continua sobre K , y en tal caso alcanza un mínimo que llamaremos C_1 . Si $X \in E$, entonces $X = 0$ ó $\frac{X}{\|X\|} \in K$. Si $X = 0$ no hay nada que probar, caso contrario, tenemos que $|\frac{X}{\|X\|}| \geq C_1$ de donde se sigue que $|X| \geq C_1 \|X\|$, lo cual finaliza la demostración. $\square$

Observación 2.1.4. Las desigualdades de la proposición previa nos dicen que la noción de convergencia dada por

$$X_k \rightarrow X \Leftrightarrow |X_k - X| \rightarrow 0,$$

define la topología usual de E . Además, si bien $|X_k - X|$ podría no coincidir con $\|X - X_k\|$, el hecho de que una converja a cero cuando k tiende a infinito implica el hecho de que la otra lo haga, ya que

$$|X_k - X| \leq C \|X_k - X\| = C \|X - X_k\| \leq C' |X - X_k|,$$

donde $\|\cdot\|$ es alguna norma en E y C, C' son ciertas constantes positivas.

Observación 2.1.5 (Homogeneizando una norma de Finsler). Si consideramos $|X|_- := |-X|$, lo recién observado es consecuencia de que $|\cdot|_-$ es una norma de Finsler si $|\cdot|$ lo es. Mas aún, define la misma norma que $|\cdot|$ si y solo si es una norma en el sentido usual. Por otro lado, es fácil de ver que si $|\cdot|$ es una norma de Finsler entonces $\|X\| := |X| + |X|_-$ define una norma completamente homogénea.

Observación 2.1.6. Dadas $|\cdot|_1$ y $|\cdot|_2$, dos normas de Finsler en E , tenemos que

$$1/R |X|_1 \leq |X|_2 \leq 1/r |X|_1 \quad \forall X \in E,$$

donde r es el radio de la mayor bola en la norma $|\cdot|_1$ que podemos meter adentro de la bola unitaria de la norma $|\cdot|_2$ y R es el menor radio de la bola en la norma $|\cdot|_2$ que contiene a la bola unitaria de la norma $|\cdot|_1$.

Definición 2.1.7 (Balanceado/simétrico). Si $K \subseteq E$ es un subconjunto, vamos a decir que es *balanceado* si $x \in K$ implica que $-x \in K$. Por otro lado si $E = \mathbb{R}^n$, diremos que el conjunto es *simétrico* si es cerrado bajo permutaciones de coordenadas. Diremos también que una norma es balanceada (resp. simétrica) si su bola unitaria es balanceada (resp. simétrica).

Observación 2.1.8. Sea $B_r = \{X \in E : |X| < r\}$ la bola de radio r centrada en $0 \in E$, claramente $B_r = rB_1$ y B_r es un abierto y convexo alrededor de $0 \in E$. Recíprocamente, si B es un conjunto abierto y convexo alrededor de $0 \in E$, su funcional de Minkowski dado por

$$\rho_B(X) = \inf\{r > 0 : X \in rB\}$$

es subaditivo y positivamente homogéneo. Si además pedimos que B no contenga semirrectas por el origen tendremos que ρ_B es una norma de Finsler (cumple $\rho_B(x) = 0$ si y solo si $x = 0$) y B resulta su bola unitaria.

Notamos que no estamos pidiendo que el conjunto sea balanceado, esta es la razón por la cual la obtenemos una norma de Finsler en lugar de una norma usual al utilizar el funcional de Minkowski.

Un ejemplo que utilizaremos de forma repetida en $\mathbb{R}^2$ es el que llamaremos “tostada”, que será el borde de la bola unitaria de norma de Finsler definida como $|(x, y)| = \|(x, y)\|_1 = |x| + |y|$ para (x, y) fuera del primer cuadrante y $|(x, y)| = \frac{x^2 + y^2}{x + y}$ para (x, y) en el primer cuadrante. Dado que el interior de la tostada es un conjunto simétrico pero no balanceado, la norma que obtenemos al considerar el funcional de Minkowski no va a ser completamente homogénea pero si va a ser simétrica. Gráficamente se ve de la siguiente forma:

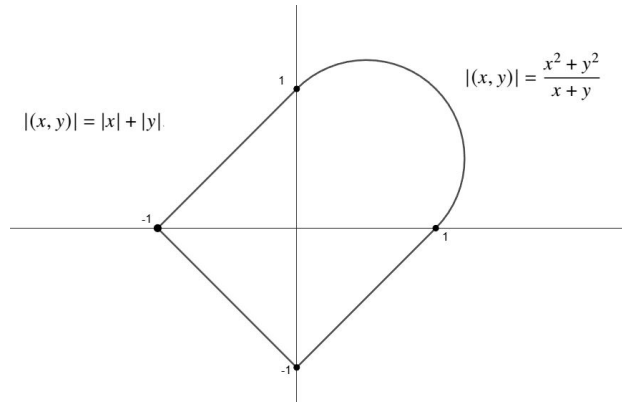


Figura 2.1: Tostada.

Proposición 2.1.9. Sea E un espacio vectorial de dimensión finita con norma de Finsler $|\cdot|$ y sea E' su espacio dual.

- (1) La aplicación dada por $\|\varphi\| = \max\{\varphi(X) : |X| \leq 1\}$ define una norma de Finsler en E' .
- (2) Dado $X \in E$, existe al menos un funcional $\varphi \in E'$ que verifica $\varphi(X) = |X|$ y $\|\varphi\| = 1$. Diremos que los funcionales con esta propiedad *realizan la norma* de X . En caso de que exista un único funcional que realiza la norma de X lo notaremos φ_X .

Demostración. (1) El número $\|\varphi\|$ está bien definido por la compacidad de la bola de E en la norma de Finsler, que es consecuencia de la Proposición 2.1.3 dado que estamos trabajando en un espacio vectorial de dimensión finita. La aplicación es subaditiva y positivamente homogénea por propiedades del máximo. El número $\|\varphi\|$ va a ser siempre mayor igual a cero, dado que siempre va a existir un $X \in E$ con $|X| \leq 1$ para el cual $\varphi(X) > 0$, esto se debe a que si $\varphi(X) < 0$ para algún X de norma menor o igual a 1, entonces $\varphi(-\lambda X) > 0$ para todo $\lambda > 0$ y podemos elegir λ de forma que $|\lambda X| \leq 1$.

Solo nos resta ver que es no degenerada, para eso supongamos que $\|\varphi\| = 0$, en tal caso $\varphi(X) \leq 0$ para todo $|X| \leq 1$. Recordamos que previamente observamos que todas las normas de Finsler son equivalentes (Proposición 2.1.3), sea $\|\cdot\|$ una norma usual en E y $r > 0$ de forma tal que $\varphi(X) \leq 0$ para todo X con $\|X\| \leq r$. Ahora si tenemos un elemento X con $\|X\| \leq r$ y $\varphi(X) < 0$ entonces $\| -X \| \leq r$ y $\varphi(-X) = -\varphi(X) > 0$ lo cual es absurdo. Debe ser entonces $\varphi(X) = 0$ para todo X con $\|X\| \leq r$ y en tal caso $\varphi = 0$.

- (2) Este último enunciado es consecuencia del Teorema de Hahn-Banach. Ver por ejemplo el libro de Rudin [61, Teorema 3.2] y sus corolarios.

□

Observación 2.1.10 (Funcionales que realizan la norma). Como vimos en la anterior proposición, dado un elemento no nulo X en un espacio normado de Finsler $(E, |\cdot|)$, existe al menos un funcional φ de norma 1 que cumple $\varphi(X) = |X|$. Estos funcionales se suelen llamar “norming functionals”, dado que no hay una traducción perfecta al español, hemos optado por llamarlos “funcionales que realizan la norma”. Siempre que los mencionemos de esa forma vamos a estar asumiendo que tienen norma 1.

Sea $\overline{B_1}$ la bola cerrada de la norma. Notamos que al ser un conjunto compacto y convexo, por el Teorema de Krein-Milman, es la cápsula convexa de sus puntos extremales. Recordamos además que $X \in K \subseteq E$ es *extremal en el sentido usual* si no pertenece al interior de ningún segmento que une dos puntos distintos del conjunto convexo K .

Definición 2.1.11. (Puntos extremales, caras, el cono generado por una cara y funcionales que exponen un elemento).

- (1) Decimos que $X \in E$ no nulo es *extremal* si $X/|X|$ es un punto extremal de $\overline{B_1}$ en el sentido usual. Notamos que una norma es estrictamente convexa si y solo si la bola de radio 1 de dicha norma es estrictamente convexa, Esto último sucede si y solo si todos los vectores no nulos son extremales.
- (2) Una *cara* F de la bola B_r del espacio normado $(E, |\cdot|)$ es la intersección de una bola cerrada $\overline{B_r}$ con el hiperplano determinado por un funcional de norma 1. Es decir, si $\varphi \in E'$, $\|\varphi\| = 1$, dicha cara es el conjunto

$$F_\varphi(r) = \overline{B_r} \cap \{X \in E : \varphi(X) = r\}.$$

En general vamos a omitir el numero r y F_φ va a hacer referencia a la cara que contiene cierto vector X , en tal caso $r = |X|$. Esta es la definición de cara que vamos a usar en este trabajo, algunos autores llaman a este concepto cara expuesta. Decimos que la cara es *maximal* si φ es un punto extremal del espacio dual.

- (3) El *cono generado* por la cara F_φ es $C_\varphi = \mathbb{R}_+ F_\varphi$. Este cono consiste exactamente de los $X \in E$ tales que $\varphi(X) = |X|$ para este funcional unitario φ . Notamos que si φ realiza la norma de un elemento, entonces realiza la norma de toda la semirrecta generada por el mismo, dado que para $\lambda > 0$ vale que $\varphi(\lambda X) = \lambda \varphi(X) = \lambda |X| = |\lambda X|$.
- (4) Decimos que un funcional φ *expone* a un elemento $X \neq 0$ si $F_\varphi = \{X\}$. Notamos que la norma va a ser estrictamente convexa si todo punto es expuesto por un único funcional que realiza la norma.

A continuación vamos a detallar como pensamos a los funcionales a nivel gráfico, esto si bien es algo evidente, ayuda mucho a la lectura y a los razonamientos que se desarrollan a lo largo de este trabajo.

Dado φ un funcional de norma 1, si consideramos el hiperplano dado por $\ker(\varphi)$ y miramos sus traslaciones obtendremos las curvas de nivel $\varphi = k$ para cada $k \in \mathbb{R}$.

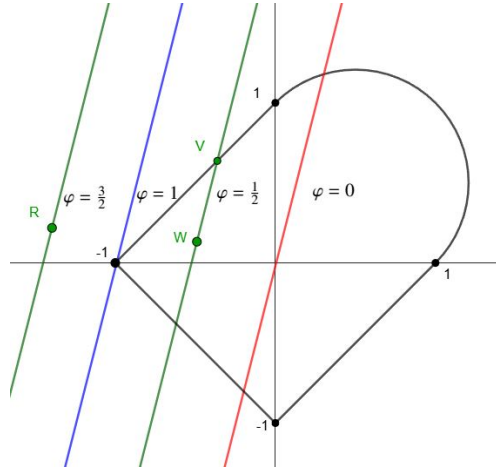


Figura 2.2: Un funcional sobre la tostada

En la Figura 2.2 de la bola unitaria de la tostada, podemos observar en rojo el núcleo del funcional φ mientras que en verde y en azul tendremos otras curvas de nivel. Supongamos primero que no sabemos que curvas de nivel son. Si miramos la primera curva de nivel a la izquierda de $\varphi = 0$ notamos que corta de forma transversal al borde de la bola, observando los puntos marcados como V y W tendremos que $\varphi(V) = \varphi(W)$, $|V| = 1$ y $|W| < 1$, este funcional no puede estar realizando la norma de V ya que en dicho caso tendríamos que $|V| = \varphi(V) = \varphi(W) \leq |W| < 1$. De hecho, esto nos dice que la curva de nivel corresponde a un valor $k < 1$, en este caso es $k = \frac{1}{2}$.

Por otro lado, si miramos la curva verde que está más a la izquierda, tendremos que debe ser una curva de nivel con $k > 1$, pues sino habría un elemento R con $|R| > 1$ y $\varphi(R) \leq 1$.

Por último, la curva de nivel marcada en azul, debe ser $\varphi = 1$, ya que sino llegaríamos a alguna de las contradicciones mencionadas previamente y luego φ realiza la norma de $(-1, 0)$.

Observación 2.1.12. Razonando en forma más general pero con las mismas cuentas, si $(E, |\cdot|)$ es un espacio normado de Finsler y φ es un funcional de norma 1, consideramos el hiperplano $\{\varphi = 1\}$. El funcional φ va a realizar la norma de los elementos de la bola unitaria que interseca al hiperplano si y solo si la bola unitaria de la norma queda de un solo lado de los semiespacios definidos por el hiperplano, más específicamente si sucede que $B_1 \subseteq \{X \in E : \varphi(X) \leq 1\}$ ó $B_1 \subseteq \{X \in E : \varphi(X) \geq 1\}$.

Observación 2.1.13. Toda cara está contenida en una cara maximal: dado que la bola cerrada del dual es un convexo compacto, por el Teorema de Krein-Milman, existen funcionales extremales $\{\varphi_i\}_{i=1, \dots, n}$ de norma 1, o sea puntos extremales de la bola dual, tal que φ es una combinación convexa de los mismos

$$\varphi = \sum_i \lambda_i \varphi_i, \quad \lambda_i \geq 0, \quad \sum_i \lambda_i = 1.$$

Es claro que si $\varphi(X) = |X|$ entonces $\varphi_i(X) = |X|$ para todo i . Luego si $X \in F_\varphi$, tiene que ser $X \in F_{\varphi_i}$ para todo i . Mas aún, F_φ es la intersección de todas las caras maximales que lo contienen.

Lema 2.1.14. Sean $(E, |\cdot|)$ un espacio normado de Finsler y $X, V_1, \dots, V_n \in E$.

- (1) La norma es estrictamente convexa si y solo si todas las caras son singletons, lo cual es equivalente a que para cada elemento no nulo exista un único funcional que realiza su norma. En tal caso, si $\varphi_X(Z) = |Z|$ entonces $Z = \lambda X$ con $\lambda = |Z|/|X|$.
- (2) $|V_1 + \dots + V_n| = |V_1| + \dots + |V_n|$ si y solo si $V_1, \dots, V_n$ pertenecen al cono generado por una cierta cara. En tal caso, todos los V_i pertenecen a la intersección de los conos generados por todas las caras que contienen a $\sum_i V_i$.
- (3) Si $\cap_\varphi C_\varphi = \{X + Y\}$ (donde la intersección la hacemos sobre todas las φ que realizan la norma de $X + Y$) y además $|X + Y| = |X| + |Y|$, entonces $|Y|X = |X|Y$.

Demostración. Para ver (1) supongamos primero que la norma es estrictamente convexa y F_φ una cara. Si $0 \neq X, Y \in F_\varphi$ entonces $\varphi(X) = |X| = |Y| = \varphi(Y)$ y luego

$$|X + Y| \geq \varphi(X + Y) = \varphi(X) + \varphi(Y) = |X| + |Y| \geq |X + Y|,$$

de donde se sigue que $|X + Y| = |X| + |Y|$ lo cual implica que $Y = cX$ para cierta constante $c > 0$ y aplicando φ a ambos lados obtenemos que $X = Y$, luego F_φ es un singleton. Por otro lado, suponiendo que toda cara es un singleton, supongamos que existen X, Y no nulos de forma tal que

$$|X + Y| = |X| + |Y|,$$

podemos considerar φ el único funcional que realiza la norma de $X + Y$ entonces

$$|X| + |Y| = |X + Y| = \varphi(X + Y) = \varphi(X) + \varphi(Y) \leq |X| + |Y|,$$

concluyéndose que $\varphi(X) + \varphi(Y) = |X| + |Y|$ y esto solo puede suceder si $\varphi(X) = |X|$ y $\varphi(Y) = |Y|$. Esto nos dice que un múltiplo positivo de X pertenece a F_φ y lo mismo para Y , como F_φ es un singleton, concluimos que existe $c > 0$ de forma tal que $Y = cX$ y luego la norma es estrictamente convexa.

Para ver (2) tenemos que hacer una cuenta muy similar pero con n elementos. Supongamos primero que $V_1, V_2, \dots, V_n$ están en el mismo cono C_φ , luego φ realiza la norma de todos estos elementos y entonces

$$|V_1| + \dots + |V_n| \geq |V_1 + \dots + V_n| \geq \varphi(V_1 + \dots + V_n) = |V_1| + \dots + |V_n|,$$

de donde se deduce la igualdad buscada.

Recíprocamente supongamos que $|V_1 + \dots + V_n| = |V_1| + \dots + |V_n|$, y sea φ un funcional que realiza la norma de la suma, es decir, $\varphi(\sum_i V_i) = |\sum_i V_i|$ entonces

$$|\sum_i V_i| = \varphi(\sum_i V_i) = \sum_i \varphi(V_i) \leq \sum_i |V_i| = |\sum_i V_i|,$$

como $\varphi(V_i) \leq |V_i|$ para todo i . Nuevamente tenemos que debe ser $\varphi(V_i) = |V_i|$ para cada i , lo cual finaliza la demostración de (2).

Para ver (3), usamos que si φ realiza la norma de $X + Y$ y $|X + Y| = |X| + |Y|$ entonces φ realiza la norma de X y de Y , lo cual probamos en (1). Se sigue entonces que $\alpha X, \beta Y \in \cap_{\varphi} C_{\varphi} = \{X + Y\}$ para ciertos $\alpha, \beta > 0$ luego por (1) tenemos que $X = |X| \frac{X+Y}{|X+Y|}$ e $Y = |Y| \frac{X+Y}{|X+Y|}$ y entonces $|Y|X = |X|Y$ lo cual finaliza la demostración. $\square$

Ejemplos 2.1.15. Tenemos muchos ejemplos de normas de Finsler: cualquier norma usual es en particular de Finsler y por la Observación 2.1.8 podemos construir ejemplos que sean normas en el sentido usual y otros que no lo sean, eligiendo convexos balanceados o no balanceados alrededor del cero como se indica en dicha observación. Vamos a construir dos ejemplos concretos (uno es el de la tostada, con el cual ya trabajamos) que nos van a servir mas adelante.

(1) *La Elipse Rotada*: consideramos la elipse rotada en el plano dada por

$$\frac{1}{a^2}(x+y)^2 + \frac{1}{b^2}(x-y)^2 = 1.$$

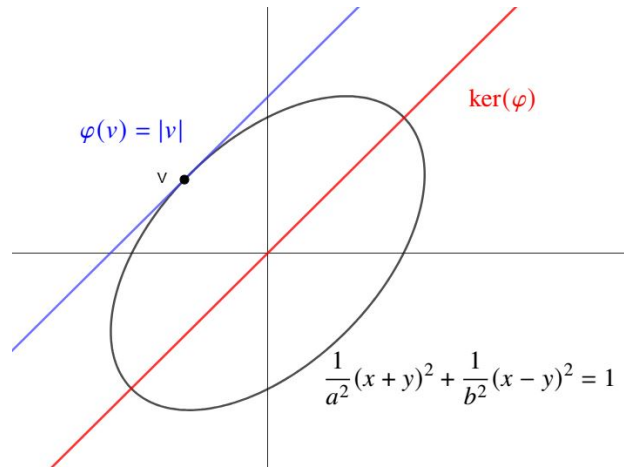


Figura 2.3: Elipse Rotada.

El interior de esta elipse es un cuerpo balanceado con borde suave y estrictamente convexo. Por la Observación 2.1.8, es la bola de cierta norma de Finsler en $\mathbb{R}^2$. El funcional de Mikowski para este cuerpo está dado por

$$F(x, y) = \sqrt{\frac{1}{a^2}(x+y)^2 + \frac{1}{b^2}(x-y)^2},$$

que es efectivamente una norma en $\mathbb{R}^2$. En la Figura 2.3 podemos observar que, por ser la norma estrictamente convexa, existe en cada punto un único funcional que realiza la norma. En este caso todas las caras son maximales y contienen un solo punto.

(2) *La Tostada*: consideramos el segmento lineal $y = x + 1$, que une los puntos $(-1, 0)$ y $(0, 1)$, el segmento $y = x - 1$ uniendo $(0, -1)$ y $(1, 0)$, y por último el segmento $y = -x - 1$ que une $(-1, 0)$ y $(0, -1)$. Cerramos esta caja con un arco de circunferencia

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

que une los puntos $(1, 0)$ y $(0, 1)$. La región encerrada por estas curvas es un cuerpo simétrico y convexo en $\mathbb{R}^2$ con forma de tostada. La fórmula para la norma de un vector está dada por

$$|(x, y)| = |x| + |y|.$$

si (x, y) pertenece al segundo, tercer o cuarto cuadrante.

Por otro lado, si $x, y \geq 0$, podemos describir los puntos de la bola como los puntos que cumplen $x^2 + y^2 = x + y$. Luego en el primer cuadrante la norma de Finsler está dada por

$$|(x, y)| = \frac{x^2 + y^2}{x + y}.$$

Gráficamente es la Figura 2.1. Vamos a agregar un gráfico indicando el comportamiento de los funcionales en los distintos puntos:

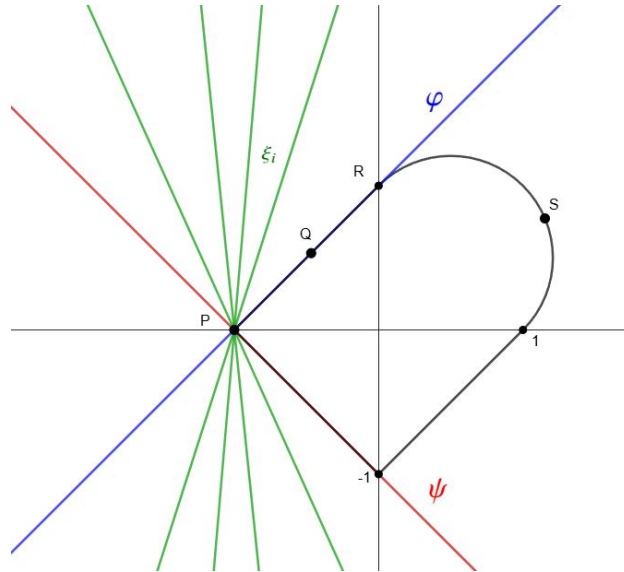


Figura 2.4: Tostada con funcionales.

Notamos en la Figura 2.4 que el funcional φ realiza la norma de los puntos P, Q, R , pero el comportamiento es muy distinto en cada uno de esos puntos. En Q y R , resulta ser el único funcional que realiza la norma, siendo R un punto que es una cara extremal mientras que Q es un punto que pertenece a una cara extremal más grande. Por otra parte en el punto P vemos que φ también realiza la norma y de hecho todos los funcionales ξ_i cuyos núcleos trasladados están representados en verde realizan la norma. Se puede observar

además que $\{P\}$ es una cara que es la intersección de las caras dadas por φ y ψ que son justamente las maximales que contienen $\{P\}$. En el punto S , por ser la norma estrictamente convexa en un entorno del mismo, tenemos que el comportamiento es idéntico al del ejemplo anterior.

Por otro lado notamos que el punto R si bien es extremal, no posee ningún funcional que lo expone, es decir no hay un funcional que realice la norma cuya cara contenga sólo el punto R . Si recordamos que antes mencionamos que una norma era estrictamente convexa si y solo si todos los puntos de la bola son extremales si y solo si es suave en todo punto si y solo si todos los puntos de la bola son expuestos por algún funcional (que resulta único). Este ejemplo nos muestra que en el caso no estrictamente convexo, pueden haber puntos extremales donde la norma es suave, pero puede que no sean expuestos por un funcional.

Observación 2.1.16. Dado E un espacio vectorial real de dimensión n y una base $B = \{E_1, E_2, \dots, E_n\}$, tenemos un único producto interno $(\cdot|\cdot)$ que hace de esta base una base ortonormal. A su vez, este producto interno nos da una isomorfismo lineal con E' a partir de asignar a cada $V \in E$ el funcional $\varphi_V(X) = (V|X)$.

Definición 2.1.17 (Norma dual). La dualidad de la observación previa nos permite definir la *norma dual* de $V \in E$ como

$$|V|' = \max\{(V|X) : |X| \leq 1\},$$

donde $(\cdot|\cdot)$ es un producto interno en E . Notamos que la norma claramente depende del producto interno elegido.

Proposición 2.1.18. La *norma dual* definida previamente, es efectivamente una norma de Finsler en E . Además, dado $V \in E$ y φ_V el funcional definido por V , tenemos que $\|\varphi_V\| = |V|'$.

Demostración. Sea $V \neq 0$, notamos que $|V|' \geq (V|\frac{V}{|V|}) = \frac{(V|V)}{|V|} > 0$, y si $|V|' = 0$ entonces $(V|V) = 0$, con lo cual $V = 0$. Dados $V, W \in E$ y $\lambda > 0$

$$\begin{aligned} |V + W|' &= \max\{(V + W|X) : |X| \leq 1\} = \max\{(V|X) + (W|X) : |X| \leq 1\} \\ &\leq \max\{(V|X) : |X| \leq 1\} + \max\{(W|X) : |X| \leq 1\} \\ &= |V|' + |W|'. \end{aligned}$$

Por otro lado

$$|\lambda V|' = \max\{(\lambda V|X) : |X| \leq 1\} = \max\{\lambda(V|X) : |X| \leq 1\} = \lambda \max\{(V|X) : |X| \leq 1\} = \lambda |V|',$$

lo cual prueba que la definición previa corresponde a una norma de Finsler.

Si $\varphi(X) = (V|X)$, entonces $\|\varphi\| = \max\{\varphi(X) : |X| \leq 1\} = \max\{(V|X) : |X| \leq 1\} = |V|'$. $\square$

Observación 2.1.19. Si bien en la proposición previa demostramos que la norma dual es una norma viendo las propiedades que cumple, esto también puede ser probado a partir de la igualdad $\|\varphi_V\| = |V|'$ y el isomorfismo lineal entre E y E' que nos da la representación de Riesz.

2.2. Normas de Finsler Ad-invariantes en $\mathfrak{u}_n$

En esta sección nos centraremos en un caso mas concreto, vamos a estudiar normas de Finsler en el espacio vectorial real de matrices antihermitianas. Los resultados originales publicados en [47] de esta sección están en la Proposición 2.2.21.

Comenzamos introduciendo algunas notaciones. Llamaremos $\mathcal{U}_n$ al grupo de matrices unitarias

$$\mathcal{U}_n = \mathcal{U}(n, \mathbb{C}) = \{U \in M_n(\mathbb{C}) : U^{-1} = U^*\}.$$

Notaremos como $\mathfrak{u}_n$ a su álgebra de Lie, la cual, la cual está formada por el conjunto de matrices antihermitianas de $n \times n$.

Definición 2.2.1. Una norma de Finsler $|\cdot|$ definida en $M_n(\mathbb{C})$ se dice *débilmente invariante* si cumple que $|UXU^*| = |X|$ para todo $X \in M_n(\mathbb{C})$ y $U \in \mathcal{U}_n$ y se dice *fuertemente invariante* si $|UXV| = |X|$ para todo $X \in M_n(\mathbb{C})$ y $U, V \in \mathcal{U}_n$.

Definición 2.2.2. Definimos para $U \in \mathcal{U}_n$, la aplicación $\text{Ad}_U : \mathfrak{u}_n \longrightarrow \mathfrak{u}_n$ como $\text{Ad}_U(X) = UXU^*$. Por otro lado para $X \in \mathfrak{u}_n$ definimos $\text{ad}_X : \mathfrak{u}_n \longrightarrow \mathfrak{u}_n$ como $\text{ad}_X(Y) = [X, Y] = XY - YX$.

Definición 2.2.3. Una norma de Finsler $|\cdot|$ en $\mathfrak{u}_n$ se dice Ad-invariante si $|\text{Ad}_U(X)| = |UXU^*| = |X|$ para todo $X \in \mathfrak{u}_n$ y todo $U \in \mathcal{U}_n$. Notamos que la noción de norma de Finsler Ad-invariante se corresponde con la de débilmente invariante (definida previamente en $M_n(\mathbb{C})$) cuando la pensamos en $\mathfrak{u}_n$.

Observación 2.2.4. Dado $X \in \mathfrak{u}_n$ sabemos que se puede escribir como $X = UDU^*$ con $U \in \mathcal{U}_n$ y $D \in \mathfrak{u}_n$ diagonal. Conjugando por matrices unitarias podemos permutar los elementos de la diagonal de D sin salirnos de la órbita de X bajo esta acción. Esta propiedad es sencilla pero fundamental en todo lo que respecta a normas de Finsler Ad-invariantes ya que nos va a dar la intuición geométrica de la simetría que tienen las bolas de estas norma. Además, nos dice que una norma de Finsler en $\mathfrak{u}_n$ aplicada a una matriz va a depender en definitiva de los autovalores de la misma.

Observación 2.2.5. Para $A, B \in M_n(\mathbb{C})$ la aplicación bilineal dada por

$$(A|B) := \text{Tr}(AB^*)$$

define un producto interno real, conocido como el producto interno de Hilbert-Schmidt. La norma de Frobenius $\|\cdot\|_F$ es la inducida por este producto interno. Por la ciclicidad de la traza, resulta que

$$\|UXU^*\|_F^2 = (UXU^*|UXU^*) = \text{Tr}(UXU^*UXU^*) = \text{Tr}(UXX^*U^*) = \text{Tr}(XX^*) = \|X\|_F^2,$$

luego $\|UXU^*\|_F = \|X\|_F$ para todo $X \in \mathfrak{u}_n$ y todo $U \in \mathcal{U}_n$, es decir, es una norma Ad-invariante. Mas aún, como veremos al final de la siguiente sección, es esencialmente la única norma Ad-invariante en $\mathfrak{u}_n$ inducida por un producto interno.

Observación 2.2.6. Si consideramos el producto interno de Frobenius, una base de $\mathfrak{u}_n$ para la cual este producto interno es ortonormal (ver Observación 2.1.16) es la siguiente:

$$B = \{E_{kk}, \frac{1}{2}(E_{l,m} - E_{m,l}), \frac{i}{2}(E_{l,m} + E_{m,l})\}$$

donde la notación es la usual de las matrices canónicas con $1 \leq k \leq n$ y $1 \leq l < m \leq 1$. Usaremos recurrentemente este producto interno y cuando hablemos de norma dual (ver 2.1.18), lo utilizaremos para definirla. Por otro lado, contando los elementos de la base en cuestión, la dimensión real de $\mathfrak{u}_n$ es

$$\dim_{\mathbb{R}}(\mathfrak{u}_n) = 2 \frac{n(n-1)}{2} + n = n^2 - n + n = n^2.$$

Proposición 2.2.7 (Dualidad de la traza). Sea $\mathfrak{u}_n'$ el dual de $\mathfrak{u}_n$ entonces:

Para cada $V \in \mathfrak{u}_n$ el elemento φ_V dado por $\varphi_V(X) = (V|X) = \text{Tr}(VX^*) = -\text{Tr}(VX)$ define un elemento en $\mathfrak{u}_n'$. Además dado $\varphi \in \mathfrak{u}_n'$, existe un único $V \in \mathfrak{u}_n$ que verifica $\varphi_V = \varphi$.

Demostración. Este resultado es una consecuencia directa de la Proposición 2.1.18 y la Observación 2.1.16. $\square$

Observación 2.2.8 (Ciclicidad de la traza). Sean $X, Y, V \in \mathfrak{u}_n$ entonces tenemos que

$$\begin{aligned} (V|[X, Y]) &= -\text{Tr}(V[X, Y]) = -\text{Tr}(VXY - VYX) = -\text{Tr}(XYV - X Y V) \\ &= -\text{Tr}(X(YV - VY)) = -\text{Tr}(YVX - Y X V) = -\text{Tr}(Y(VX - XV)) \end{aligned}$$

de donde se siguen las identidades

$$(V|[X, Y]) = (X|[Y, V]) = (Y|[V, X]).$$

Estas identidades tambien pueden ser escritas como $(V| \text{ad}_X(Y)) = -(\text{ad}_X(V)|Y) = (\text{ad}_Y(V)|X)$. Luego, la aplicación ad_X cumple $(\text{ad}_X)^* = -\text{ad}_X$, donde la adjunta es respecto al producto interno de Hilbert-Schmidt.

Ejemplo 2.2.9. A continuación enumeramos una lista de normas Ad-invariantes en $\mathfrak{u}_n$.

- (1) $\|X\|_1 = \text{Tr}|X|$ (norma de la traza).
- (2) $\|X\|_{\infty} = \max\{|\lambda| : \lambda \in \sigma(X)\}$ (norma espectral).
- (3) For $1 < p < \infty$, $\|X\|_p = (\text{Tr}|X|^p)^{\frac{1}{p}}$ (normas p).
- (4) $\|X\|_{(k)} = \sum_{1 \leq i \leq k} |\lambda_i|$, donde λ_i son los autovalores de X ordenados en forma decreciente (normas Ky-Fan).
- (5) $\|X\|_C = \max\{|\text{Tr}(CU^*XU)| : U \in \mathcal{U}_n\}$ para $C \in M_n(\mathbb{C})$ una matriz no escalar de traza no nula (normas C).

Tenemos que $\|\cdot\|_{(1)} = \|\cdot\|_{\infty}$, $\|\cdot\|_{(n)} = \|\cdot\|_1$, y $\|X\|_2 = \|X\|_F$ (la norma Frobenius). Todos estos ejemplos son de hecho normas (completamente homogéneas) y todas salvo la última son fuertemente invariantes; una buena referencia sobre el tema es el libro de Bhatia [15, Capítulo IV]. De hecho, la demostración detallada de que $\|\cdot\|_C$ es una norma puede ser encontrada en [15, Proposición IV.4.4]. Es fácil de probar que si C es una proyección unidimensional entonces la norma C es el radio numérico, en nuestro caso la norma espectral.

Observación 2.2.10. Notamos que, por la Observación 2.1.8, podemos identificar un cuerpo B convexo y abierto alrededor de $0 \in \mathfrak{u}_n$ que no contiene semirectas por el origen con una norma de Finsler en $\mathfrak{u}_n$ y viceversa. Si al cuerpo además le pedimos que sea Ad-invariante, es decir, que si $X \in B$ entonces $UXU^* \in B$ para todo $U \in \mathcal{U}_n$, entonces usando la misma identificación podemos construir una norma de Finsler Ad-invariante en $\mathfrak{u}_n$ de forma que su bola unitaria sea B . De hecho toda norma de Finsler Ad-invariante se obtiene unívocamente de esta forma, al considerar $B = \{X \in \mathfrak{u}_n : |X| < 1\}$.

A continuación vamos a construir una familia de normas llamadas *normas orbitales* que van a jugar un rol muy importante dentro de la teoría de normas de Finsler Ad-invariantes.

Definición 2.2.11. Llamamos SU_n al grupo de matrices unitarias de $n \times n$ con determinante 1. Identificamos su álgebra de Lie con la subálgebra de $\mathfrak{u}_n$ formada por las matrices de traza cero que notaremos $\mathfrak{su}_n$. Dada $C \in \mathfrak{su}_n$ no nula, definimos la *norma orbital* asociada a C como

$$|X|_{\mathcal{O}_C} = \max\{(UCU^*|X) : U \in \mathcal{U}_n\} + |\text{Tr}(X)|,$$

con $X \in \mathfrak{u}_n$. Notamos además que el término de la traza está para que efectivamente sea una norma en $\mathfrak{u}_n$, pero en $\mathfrak{su}_n$ este término va a ser siempre nulo. Esta es la razón por la cual muchas veces trabajaremos en $\mathfrak{su}_n$ en vez de en $\mathfrak{u}_n$, dado que es el espacio mas natural para estudiar estas normas.

Notamos que este máximo está claramente bien definido por la compacidad de $\mathcal{U}_n$ y la continuidad de las operaciones. Veremos en esta sección que se trata efectivamente de una norma de Finsler Ad-invariante en $\mathfrak{u}_n$.

A continuación vamos a desarrollar una versión del Teorema de Horn y Uhlmann sobre las cápsulas convexas de la órbita de una matriz autoadjunta [37]; las equivalencias para el caso de normas homogéneas pueden ser encontradas en el libro de Bhatia [15, Teorema IV.4.7]. Para comenzar necesitamos algunas notaciones:

Notaciones 2.2.12. Dado $X \in \mathfrak{u}_n$ existen $U \in \mathcal{U}_n$ y D una matriz diagonal de entradas puramente imaginarias de forma tal que $X = UDU^*$.

- Notaremos $(ix_k)_{k=1}^n$ a las entradas de la matriz D , es decir los autovalores de D . Observamos que cada x_k debe ser un número real.
- Para X de esta forma notaremos $x = (x_1, x_2, x_3, \dots, x_n)$ a la tira formada por la parte imaginaria de los autovalores de X .
- En general, notaremos como $x^\downarrow$ al reordenamiento decreciente del vector $x \in \mathbb{R}^n$, denotaremos también

$$x_i^\downarrow = (x^\downarrow)_i$$

para no cargar la notación.

- Dado $B \subseteq \mathfrak{u}_n$ vamos a notar como $\mathcal{O}(B)$ a la órbita coadjunta bajo la acción del grupo unitario, es decir

$$\mathcal{O}(B) = \{UXU^* : X \in B, U \in \mathcal{U}_n\}.$$

Si $C \in \mathfrak{u}_n$, escribimos $\mathcal{O}_C = \mathcal{O}(\{C\})$.

- Si Σ es un sub-conjunto de un espacio vectorial, notaremos $\text{conv}(\Sigma)$ a su cápsula convexa.

Definición 2.2.13. Sean $x, y \in \mathbb{R}^n$, diremos que y *mayoriza fuertemente* a x y notaremos $x < y$ si

- (1) $\sum_{k=1}^m x_k^\downarrow \leq \sum_{k=1}^m y_k^\downarrow$ para cada $m = 1, 2, \dots, n$.
- (2) $\sum_{k=1}^n x_k = \sum_{k=1}^n y_k$.

Vamos a probar una serie de equivalencias que nos serán útiles para entender las normas de Finsler Ad-invariantes y el papel que juegan las normas orbitales. La equivalencia de los primeros dos ítems de este resultado (Proposición 2.2.16) es una consecuencia del Teorema de Schur-Horn, así que primero vamos a dar un lema y recordar este teorema.

Lema 2.2.14. Sean $z, w \in \mathbb{R}^n$, entonces son equivalentes

- (1) z pertenece a la cápsula convexa del conjunto de todos los vectores que se obtienen permutando coordenadas de w .
- (2) $z < w$
- (3) Existe $A \in M_n(\mathbb{R})$ doblemente estocástica que cumple $z^\downarrow = Aw^\downarrow$. Recordamos que una matriz doblemente estocástica es una matriz de elementos no negativos para la cual la suma de cada una de sus filas y de cada una de sus columnas es exactamente 1.

Demostración. La demostración de este resultado puede ser encontrada en [4, Teorema 1.3]. $\square$

Teorema 2.2.15. (Schur-Horn) Dados $x, y \in \mathbb{R}^n$, existe una matriz hermitiana con diagonal x y cuyos autovalores son $y_1, y_2, \dots, y_n$ si y solo si $x < y$.

Demostración. La demostración de este resultado se puede encontrar en el Teorema 5 del paper de Horn [37], en el cual prueba una implicación y cita la otra probada previamente por Schur. $\square$

Proposición 2.2.16 (Mayorización y normas de Finsler). Sean $Z, W \in \mathfrak{u}_n$. Las siguientes afirmaciones son equivalentes:

- (1) $Z \in \text{conv}(\mathcal{O}_W)$, mas precisamente, existen (a lo sumo) $n+1$ matrices $U_i \in \mathcal{U}_n$ y $n+1$ números reales $\lambda_i \geq 0$ con $\sum_i \lambda_i = 1$ tales que $Z = \sum_{i=1}^{n+1} \lambda_i U_i W U_i^*$.
- (2) $z < w$
- (3) $|Z| \leq |W|$ para toda norma de Finsler Ad-invariante en $\mathfrak{u}_n$ y $\text{Tr}(Z) = \text{Tr}(W)$.
- (4) $\max_{U \in \mathcal{U}_n} (UCU^*|Z) \leq \max_{U \in \mathcal{U}_n} (UCU^*|W)$ para toda $C \in \mathfrak{su}_n$ no nula y $\text{Tr}(Z) = \text{Tr}(W)$.
- (5) $\max_{U \in \mathcal{U}_n} (UCU^*|Z) \leq \max_{U \in \mathcal{U}_n} (UCU^*|W)$ para toda $C \in \mathfrak{u}_n$.

Demostración. La primera observación que debemos hacer es que, si $Z \in \text{conv}(\mathcal{O}_W)$, entonces Z se puede escribir como una combinación convexa de m elementos de $\mathcal{O}_W$. Por otra parte, el Teorema de Caratheodory nos permite limitar ese m para que sea menor igual a $n+1$. Una demostración de este teorema se puede ver en [23, Proposición 2.6]. Habiendo aclarado esto, pasamos a las implicaciones.

Comenzaremos viendo que (1) y (2) son equivalentes. Supongamos que vale (2), por el Lema 2.2.14, z debe pertenecer a la cápsula convexa de las permutaciones del vector w . Esto nos dice que podemos escribir

$$z = \sum_{\sigma} \lambda_{\sigma} P_{\sigma}(w),$$

donde σ recorre todas las posibles permutaciones de n elementos, $P_{\sigma}(w) = (w_{\sigma(1)}, w_{\sigma(2)}, \dots, w_{\sigma(n)})$ y $\sum_{\sigma} \lambda_{\sigma} = 1$. Si llamamos D_w a la matriz diagonal cuyas entradas son las coordenadas del vector w y análogamente D_z a la asociada al vector z y recordamos que toda permutación de elementos de una matriz diagonal se puede obtener al conjugar por una unitaria, obtenemos que existen matrices $U_{\sigma} \in \mathcal{U}_n$ de forma tal que

$$D_z = \sum_{\sigma} \lambda_{\sigma} U_{\sigma} D_w U_{\sigma}^*.$$

Para terminar la demostración, escribimos $Z = iU_z D_z U_z^*$ y $W = iU_w D_w U_w^*$ para ciertas $U_z, U_w \in \mathcal{U}_n$. Multiplicando la igualdad previa por i y conjugando por U_z obtenemos

$$Z = \sum_{\sigma} \lambda_{\sigma} U_z U_{\sigma} U_w^* W U_w U_{\sigma}^* U_z^*,$$

que es a lo que queríamos probar.

Por otro lado, suponiendo que vale (1), y operando como hicimos previamente tenemos que

$$D_z = \sum_i \lambda_i U_i D_w U_i^*.$$

Como la matriz $U_i D_w U_i^*$ es una matriz hermitiana, por el Teorema 2.2.15 (Schur-Horn), su diagonal tiene que estar mayorizada por su tira de autovalores. Como esta tiene los mismos autovalores que W , entonces su diagonal está mayorizada por w y por el lema previo tenemos que su diagonal está en la cápsula convexa de las permutaciones del vector w . Se sigue que el vector z también tiene que estar en dicha cápsula, ya que es una combinación convexa de elementos que pertenecen a la misma, y luego aplicando nuevamente el Lema 2.2.14, concluimos que z está mayorizado por w , o sea vale (2).

Supongamos que vale (1) y veamos (3), aplicando norma en ambos miembros y la desigualdad triangular tenemos que

$$|Z| \leq \sum_{i=1}^{n+1} |\lambda_i U_i W U_i^*| = \sum_{i=1}^{n+1} \lambda_i |U_i W U_i^*| = \sum_{i=1}^{n+1} \lambda_i |W| = |W|,$$

donde en la anteúltima igualdad usamos que la norma es Ad-invariante y en la última que la suma de los λ_i es igual a 1. El hecho de que las trazas coincidan es una consecuencia inmediata de tomar traza en la escritura de (1) y de la ciclicidad de la misma. Concluimos entonces que vale (3).

Veamos que (5) implica (1). Supongamos que (1) no es verdadera, es decir $Z \notin \text{conv}(\mathcal{O}_W)$, por el Teorema de Hahn-Banach, existe un funcional φ que cumple

$$\varphi(Y) < r < \varphi(Z) \text{ para todo } Y \in \text{conv}(\mathcal{O}_W).$$

Representando a este funcional como $\varphi = (X| -)$ para cierto $X \in \mathfrak{u}_n$, y considerando $U \in \mathfrak{u}_n$ tenemos que $\varphi(Z) > r > \varphi(UWU^*)$ para todo $U \in \mathfrak{u}_n$. Tomando máximo sobre todo $U \in \mathcal{U}_n$ se sigue que

$$\varphi(Z) > \max_{U \in \mathcal{U}_n} (X|UWU^*) = \max_{U \in \mathcal{U}_n} (U^*XU|W) \geq \max_{U \in \mathcal{U}_n} (U^*XU|Z) \geq (X|Z) = \varphi(Z),$$

lo cual es absurdo. Concluimos que $Z \in \mathcal{O}_W$, es decir, que vale (1).

El hecho de que (3) implique (4), va a ser una consecuencia directa de que si $C \in \mathfrak{su}_n$ y es no nula, entonces

$$|X|_{\mathcal{O}_C} = \max\{(UCU^*|X) : U \in \mathcal{U}_n\} + |\text{Tr}(X)|,$$

define una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, lo cual demostraremos en la Proposición 2.2.21 y del hecho de que ambos elementos tienen la misma traza.

Supongamos por último que vale (4). Sea $C \in \mathfrak{u}_n$, si C es un múltiplo de la identidad, reemplazando $C = \lambda I$, la desigualdad se reduce a

$$-\lambda \text{Tr}(Z) \leq -\lambda \text{Tr}(W)$$

la cual es evidente pues $\text{Tr}(Z) = \text{Tr}(W)$.

Si C no es un múltiplo de la identidad, entonces $0 \neq C' = C - \frac{1}{n} \text{Tr}(C) \in \mathfrak{su}_n$. Aplicando (4) a C'

$$\begin{aligned} \max_{U \in \mathcal{U}_n} (UC'U^*|Z) &\leq \max_{U \in \mathcal{U}_n} (UC'U^*|W) \\ \max_{U \in \mathcal{U}_n} (UCU^*|Z) - \text{Tr}(C) \text{Tr}(Z) &\leq \max_{U \in \mathcal{U}_n} (UCU^*|W) - \text{Tr}(C) \text{Tr}(W) \\ \max_{U \in \mathcal{U}_n} (UCU^*|Z) &\leq \max_{U \in \mathcal{U}_n} (UCU^*|W) \end{aligned}$$

donde usamos que $\text{Tr}(Z) = \text{Tr}(W)$. □

Observación 2.2.17. Notamos que la equivalencia entre (1) y (2) fue demostrada sin utilizar las otras afirmaciones. Esto será importante ya que utilizaremos dicha equivalencia mas adelante para probar que las normas orbitales en el caso de que C es no nula y de traza cero son normas de Finsler. Dicho resultado fue el argumento usado para ver que (3) implica (4) en la demostración previa.

Observación 2.2.18. Es suficiente chequear la condición (3) de la proposición previa para normas estrictamente convexas. Esto se debe a que toda norma de Finsler Ad-invariante $|\cdot|$ puede ser aproximada por una estrictamente convexa y de las mismas características considerando para $\epsilon > 0$ la norma de Finsler Ad-invariante dada por $|X|_\epsilon = |X| + \epsilon \|X\|_F$. Por otra parte, es suficiente ver la condición (5) para matrices regulares en $\mathfrak{u}_n$ (i.e. con todos autovalores distintos), ya que las mismas forman un subconjunto denso en $\mathfrak{u}_n$.

Proposición 2.2.19. Con la notación de la Proposición 2.2.16. Si $Z \in \text{conv}(\mathcal{O}_W)$ y la igualdad vale para alguna norma de Finsler, entonces Z y todos los elementos de la forma $U_i W U_i^*$ caen en la misma cara de la esfera de esa norma. Más aún pertenecen a la intersección de todas las caras que contienen a Z . Si la norma es estrictamente convexa entonces $Z = UWU^*$ para algún $U \in \mathcal{U}_n$.

Demostración. Sea F_φ una cara de la esfera con $Z \in F_\varphi$, luego φ es un funcional que realiza la norma de Z . Por otro lado $Z = \sum_{i=1}^{n+1} \lambda_i U_i W U_i^*$ y aplicando el funcional a la igualdad tenemos

$$|Z| = \varphi(Z) = \sum_{i=1}^{n+1} \lambda_i \varphi(U_i W U_i^*) \leq \sum_{i=1}^{n+1} \lambda_i |U_i W U_i^*| = \sum_{i=1}^{n+1} \lambda_i |W| = |W| = |Z|$$

y luego se tiene que $\varphi(U_i W U_i^*) = |Z|$ y entonces $U_i W U_i^* \in F_\varphi$ para todo i . Si la norma es estrictamente convexa $F_\varphi = \{Z\}$ y entonces $Z = U_i W U_i^*$ para todo i y se sigue el resultado. $\square$

Observación 2.2.20. En la demostración de la Proposición 2.2.16, mencionamos que

$$|X|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (UCU^*|X| + |\text{Tr}(X)|,$$

define una norma cuando $0 \neq C \in \mathfrak{su}_n$ (lo cual probaremos a continuación). Si nos centramos en (4) implica (3) de dicha proposición, observamos que obtuvimos una familia de normas de Finsler Ad-invariantes que controlan a todas las normas de Finsler Ad-invariantes sobre $\mathfrak{su}_n$. Como podemos ver en el libro de Bhatia [15, Teorema IV.2.2], las normas Ky-Fan juegan un papel similar en el caso fuertemente invariante, y por ello proponemos a las orbitales como las normas de control para el caso débilmente invariante.

A continuación veremos que efectivamente las normas orbitales son normas de Finsler Ad-invariantes en $\mathfrak{su}_n$ y entonces, si sumamos el término del módulo de la traza, lo son en $\mathfrak{u}_n$.

Proposición 2.2.21. Sea $C \in \mathfrak{su}_n$ una matriz no nula, luego la norma orbital alrededor de C definida como

$$|X|_{\mathcal{O}_C} = \max\{(UCU^*|X|) : U \in \mathcal{U}_n\}$$

es una norma de Finsler Ad-invariante en $\mathfrak{su}_n$ y además se tiene que

- (1) La norma de X se realiza en un elemento de la forma UCU^* . Dicho elemento conmuta con X y luego se diagonalizan simultáneamente. Además vale la igualdad:

$$|X|_{\mathcal{O}_C} = (UCU^*|X|) = \sum_{i=1}^n c_i^\downarrow x_i^\downarrow.$$

Ver la Observación 2.2.12 para la notación.

- (2) $|C|_{\mathcal{O}_C} = -\text{Tr}(C^2) = \|C\|_F^2$. En particular, $|C|_{\mathcal{O}_C} = 1$ si y solo si $\|C\|_F = 1$.
- (3) Si $|C|_{\mathcal{O}_C} = 1$ y $X \in \mathcal{O}_C$ entonces $|X|_{\mathcal{O}_C} = 1$.
- (4) Si suponemos que $X \in \text{conv}(\mathcal{O}_C)$ entonces $|X|_{\mathcal{O}_C} \leq 1$, con igualdad si y solo si $X \in \mathcal{O}_C$.
- (5) La norma es completamente homogénea si y solo si el espectro de C es balanceado, es decir $\sigma(C) = -\sigma(C)$.
- (6) La extensión dada por $|X|_{\mathcal{O}_C} = \max\{(UCU^*|X|) : U \in \mathcal{U}_n\} + |\text{Tr}(X)|$ define una norma Ad-invariante en $\mathfrak{u}_n$.

Demostración. Sea $X \in \mathfrak{su}_n$, luego $\sum_j x_j = -i \operatorname{Tr}(X) = 0$. Si ordenamos los x_k de forma decreciente se sigue que $\sum_{k=1}^m x_k \geq 0$ para $m = 1, 2, \dots, n$ y como además la suma de todos es cero, tenemos que $\vec{0} < \vec{x}$. Por la proposición previa se deduce que $0 \in \operatorname{conv}(\mathcal{O}_X)$ para cualquier $X \in \mathfrak{su}_n$ y además que, para $C \in \mathfrak{su}_n$, vale que

$$\max\{(UCU^*|X) : U \in \mathcal{U}_n\} = \max\{-\operatorname{Tr}(CU^*XU) : U \in \mathcal{U}_n\} \geq 0.$$

Notamos que para $\lambda > 0$, $|\lambda X|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (UCU^*|\lambda X) = \lambda \max_{U \in \mathcal{U}_n} (UCU^*|X) = \lambda |X|_{\mathcal{O}_C}$ lo cual me dice que es positivamente homogénea. Por otro para $X, Y \in \mathfrak{su}_n$ tenemos que

$$|X + Y|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (UCU^*|X + Y) \leq \max_{U \in \mathcal{U}_n} (UCU^*|X) + \max_{U \in \mathcal{U}_n} (UCU^*|Y) = |X|_{\mathcal{O}_C} + |Y|_{\mathcal{O}_C}$$

lo cual prueba la desigualdad triangular.

Para ver que es Ad-invariante, simplemente consideramos $V \in \mathfrak{u}_n$ y calculamos

$$|VXV^*|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (UCU^*|VXV^*) = \max_{U \in \mathcal{U}_n} (V^*UCU^*V|X) = |X|_{\mathcal{O}_C}.$$

Solo resta ver que es no degenerada. Sea $\vec{c}$ la tira de autovalores de $-iC$. Como C no es un múltiplo de la identidad y $\operatorname{Tr}(C) = 0$, entonces existen $n - 1$ permutaciones $\sigma_k \in S_n$ tales que $\{\sigma_k(\vec{c})\}_{k=1, \dots, n-1}$ genera $\{v \in \mathbb{R}^n : \sum_{i=1}^n v_i = 0\}$ en $\mathbb{R}^n$ (ver por ejemplo [35]), luego la órbita de C genera $\mathfrak{su}_n$. Sabemos que $0 \in \operatorname{conv}(\mathcal{O}_C)$, afirmamos que es un punto interior (notar que esto nos dice en particular que el interior de $\operatorname{conv}(\mathcal{O}_C)$ es no vacío en $\mathfrak{su}_n$). Caso contrario, por el Teorema de separación de Hahn-Banach, existe un funcional unitario φ tal que $\varphi(0) = 0$ y $\varphi(\operatorname{conv}(\mathcal{O}_C)) \geq 0$. Para este funcional será en tal caso $\varphi(UCU^*) \geq 0$ para todo $U \in \mathcal{U}_n$, si consideramos dU la medida de Haar normalizada en $\mathcal{U}_n$ y $W = \int_{\mathcal{U}_n} UCU^* dU$, tendremos que $W \in \mathfrak{su}_n$ y $\operatorname{Ad}_U W = W$ para todo $U \in \mathcal{U}_n$. Esto es posible si y solo si $W = \lambda 1$ y luego debe ser $W = 0$ y entonces

$$0 = \varphi(W) = \int_{\mathcal{U}_n} \varphi(UCU^*) dU \geq 0,$$

lo cual ocurre si y solo si $\varphi(UCU^*) = 0$ para todo $U \in \mathcal{U}_n$. Esto último implica que $\operatorname{conv}(\mathcal{O}_C) \subset \ker \varphi$, lo cual es una contradicción y entonces 0 es un punto interior de $\operatorname{conv}(\mathcal{O}_C)$. Sea ahora $X \neq 0$ en $\mathfrak{su}_n$, multiplicando X por un escalar suficientemente pequeño, obtenemos un elemento en el interior de la cápsula convexa de la órbita de C , luego existe $\lambda > 0$ tal que $\lambda X = \sum_i \lambda_i U_i C U_i^*$. Si suponemos que $|X|_{\mathcal{O}_C} = 0$, luego $(UCU^*|X) \leq 0$ para todo $U \in \mathcal{U}_n$. Pero por otro lado

$$0 < \lambda \|X\|_F^2 = \sum_i \lambda_i (U_i C U_i^*|X) \leq 0,$$

lo cual es absurdo. Debe ser entonces $|X|_{\mathcal{O}_C} \neq 0$, lo cual finaliza la prueba de que nuestra norma es una norma de Finsler Ad-invariante en $\mathfrak{su}_n$.

Veamos (1), para ello, sean $Z \in \mathfrak{su}_n$ y $U \in \mathcal{U}_n$ tal que $(UCU^*|X) = |X|_{\mathcal{O}_C}$, es decir, realiza la norma de X . Consideramos la función $f(t) = (e^{tZ}UCU^*e^{-tZ}|X)$, que debe tener un mínimo en $t = 0$ y en tal caso $f'(0) = 0$. Calculando $f'(t)$ obtenemos

$$f'(t) = (e^{tZ}ZUCU^* - e^{tZ}UCU^*Z|X)$$

y luego $f'(0) = (ZUCU^* - UCU^*Z|X) = ([Z, UCU^*]|X) = (Z|[UCU^*, X]) = 0$.

Notar que usamos la Observación 2.2.8 para la anteúltima igualdad. Tomando $Z = [UCU^*, X]$ se sigue que $[UCU^*, X] = 0$ y podemos diagonalizar ambos elementos simultáneamente. Notando como $(c_k)_{k=1}^n$ a los autovalores de C , tenemos que $(UCU^*|X) = \sum_i c_{\sigma(i)} x_i$ para alguna permutación $\sigma \in S_n$, pero cualquiera de esas sumas esta dominada por el producto de ambas tiras de coeficientes ordenadas en orden decreciente (ver por ejemplo [15, Corolario II.4.4]). Lo que resta ver es que si efectivamente $|X|_{\mathcal{O}_C} = (UCU^*|X)$ entonces necesariamente la permutación debe ser tal que la tira $(c_{\sigma(i)})_i$ está ordenada de forma decreciente.

Sea $W \in \mathcal{U}_n$ de forma tal que WXW^* es diagonal con los autovalores ordenados de forma decreciente. Luego tenemos que

$$\begin{aligned} |X|_{\mathcal{O}_C} &= \max_{U \in \mathcal{U}_n} (UCU^*|X) = \max_{U \in \mathcal{U}_n} (WUCU^*W^*|WXW^*) = \\ &= \max_{V \in \mathcal{U}_n} (VCV^*|WXW^*) \geq \sum_i c_{\sigma(i)} x_i, \end{aligned}$$

donde $V \in \mathcal{U}_n$ esta elegida para que VCV^* sea la matriz que tiene en la diagonal la tira $(c_i^\downarrow)_i$. Habiendo visto entonces ambas desigualdades, se sigue que si $|X|_{\mathcal{O}_C} = (UCU^*|X)$ entonces $|X|_{\mathcal{O}_C} = \sum_i c_{\sigma(i)} x_i$.

Veamos (2), para ello notamos que

$$(UCU^*|C) \leq \|UCU^*\|_F \|C\|_F = \|C\|_F^2 = -\text{Tr}(C^2),$$

donde la desigualdad usada es la de Cauchy-Schwartz para el producto interno de la traza. Tomando máximo sobre todo $U \in \mathcal{U}_n$, se sigue que $|C|_{\mathcal{O}_C} \leq -\text{Tr}(C^2) = \|C\|_F^2$. Por otro lado, tomando $U = 1$, obtenemos que $(UCU^*|C) = (C|C) = -\text{Tr}(C^2)$ y luego $\|C\|_{\mathcal{O}_C} \geq -\text{Tr}(C^2) = \|C\|_F^2$, entonces

$$\|C\|_F^2 = -\text{Tr}(C^2) = |C|_{\mathcal{O}_C},$$

lo cual prueba (2). Para ver (3), consideramos $X \in \mathcal{O}_C$ donde $|C|_{\mathcal{O}_C} = 1$. Como $X \in \mathcal{O}_C$, debe existir V unitario de forma tal que $VXV^* = C$ y luego

$$|X|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (UCU^*|X) = \max_{U \in \mathcal{U}_n} (UCU^*|V^*CV) = \max_{U \in \mathcal{U}_n} (VUCU^*V^*|C) = |C|_{\mathcal{O}_C} = 1$$

Para ver (4), suponemos que $X \in \text{conv}(\mathcal{O}_C)$ y en tal caso tenemos que $X = \sum_{i=1}^{n+1} \lambda_i U_i C U_i^*$, donde $\sum_{i=1}^{n+1} \lambda_i = 1$ y los λ_i son todos números positivos, luego

$$(UCU^*| \sum_{i=1}^{n+1} \lambda_i U_i C U_i^*) = \sum_{i=1}^{n+1} \lambda_i (UCU^*|U_i C U_i^*) \leq \sum_{i=1}^{n+1} \lambda_i \|C\|_F \leq \sum_{i=1}^{n+1} \lambda_i = 1,$$

lo cual prueba que $|X|_{\mathcal{O}_C} \leq 1$.

Para concluir (4) solo resta ver que si $X \in \text{conv}(\mathcal{O}_C)$ y $|X|_{\mathcal{O}_C} = 1$ entonces $X \in \mathcal{O}_C$. Notamos que debe de existir un U_0 que realice la norma de X es decir $1 = (C|U_0^* X U_0)$. escribiendo a X como antes y usando nuevamente la desigualdad de Cauchy-Schwarz obtenemos que

$$1 = (C|U_0^* X U_0) = \sum_i \lambda_i (C|\tilde{U}_i C \tilde{U}_i^*) \leq \sum_i \lambda_i \|C\|_F^2 = 1,$$

donde $\tilde{U}_i = U_0 U_i$ para cada $1 \leq i \leq n+1$. Esto nos dice que $\sum_i \lambda_i (C|\tilde{U}_i C \tilde{U}_i^*) = \sum_i \lambda_i \|C\|_F^2$ y luego, al ser una combinación con elementos positivos, debe valer la igualdad $(C|\tilde{U}_i C \tilde{U}_i^*) = \|C\|_F \|\tilde{U}_i C \tilde{U}_i^*\|_F$ para cada $1 \leq i \leq n+1$. Esto implica que debe existir, para cada i , un número $t_i \geq 0$ de forma tal que $C = t_i \tilde{U}_i C \tilde{U}_i^*$, pero como ambos elementos tienen normal orbital sobre C igual a 1, debe de ser $t_i = 1$ para todo i , de donde se deduce que C conmuta con cada $\tilde{U}_i$. Concluimos que

$$U_0^* X U_0 = \sum_i \lambda_i \tilde{U}_i C \tilde{U}_i^* = C$$

y luego $X = U_0 C U_0^* \in \mathcal{O}_C$ lo cual finaliza (3).

Veamos (5), para ello asumamos primero que el espectro de C es balanceado. Si consideramos $X \in \mathfrak{u}_n$, U de forma tal que $(UCU^*|X) = |X|_{\mathcal{O}_C}$ y U_0 de forma tal que $-C = U_0 C U_0^*$ (notar que la existencia de U_0 es consecuencia de que cada autovalor de C tiene la misma multiplicidad algebraica que su opuesto, puesto que $\text{Tr}(C) = 0$ y C es balanceada), entonces tenemos que

$$|-X|_{\mathcal{O}_C} \geq (UU_0 C U_0^* U^*| -X) = (UCU^*|X) = |X|_{\mathcal{O}_C}.$$

Argumentando igual con $-X$ se deduce la otra desigualdad y luego $|X|_{\mathcal{O}_C} = |-X|_{\mathcal{O}_C}$ lo cual me dice que $|\cdot|$ es completamente homogénea.

Por último, si el espectro de C no es balanceado, entonces $-C \notin \mathcal{O}_C$, ya que sino tendríamos una contradicción en la igualdad $-C = UCU^*$. Asumamos que la norma es completamente homogénea, es decir que $|-X|_{\mathcal{O}_C} = |X|_{\mathcal{O}_C}$ para todo $X \in \mathfrak{u}_n$ y lleguemos a un absurdo. Recordamos que $(c_1, c_2, \dots, c_n)$ es la tira formada por la parte imaginaria de los autovalores de C ordenada en forma decreciente.

Si p es una proyección 1-dimensional, por la igualdad $-C = UCU^*$ y el ítem (1) de esta proposición tenemos que $c_1 = |ip|_{\mathcal{O}_C} = |-ip|_{\mathcal{O}_C} = |-ip|_{\mathcal{O}_{-C}} = -c_n$. Sean ahora p_1, p_2 proyecciones 1-dimensionales ortogonales, luego, usando nuevamente el ítem 1 de esta proposición:

$$c_1 + c_2 = |i(p_1 + p_2)|_{\mathcal{O}_C} = |-i(p_1 + p_2)|_{\mathcal{O}_C} = -c_{n-1} - c_n,$$

de donde se sigue que $c_2 = -c_{n-1}$. Procediendo de esta forma terminamos en que $c_{k+1} = 0$ (para n impar), o en que $c_k = -c_{k+1}$ (para n par). En cualquier caso, $\sigma(C) = -\sigma(C)$, lo cual es una contradicción. Luego la norma no puede ser homogénea.

Para ver (6), simplemente observamos que tenemos que probar que es no degenerada, todas las otras propiedades son consecuencia de las propiedades de la traza y de las propiedades vistas previamente. Para ello supongamos que $|X|_{\mathcal{O}_C} = 0$ para $X \in \mathfrak{u}_n$, en dicho caso tendremos que $\text{Tr}(X) = 0$ y además $\max_{U \in \mathfrak{u}_n} (UCU^*|X)$, luego $X \in \mathfrak{su}_n$ y $|X|_{\mathcal{O}_C} = 0$ y como vimos al principio de esta demostración, esto implica que $X = 0$. $\square$

Observación 2.2.22. Si asumimos $|C|_{\mathcal{O}_C} = 1$ y notamos B_1 a la bola cerrada de radio 1 de la norma orbital, el ítem (4) de la proposición previa nos dice que $\text{conv}(\mathcal{O}_C) \subseteq B_1$. En general no tienen porque coincidir estos dos conjuntos, como ejemplo podemos considerar en $\mathfrak{u}_3$ la matriz $C = \frac{i}{\sqrt{2}} \text{diag}(1, 0, -1)$ que cumple $|C| = 1$ y la matriz $X = i \frac{\sqrt{2}}{3} \text{diag}(2, -1, -1)$ que, utilizando el ítem (1) de la proposición previa, es inmediato ver que $|X|_{\mathcal{O}_C} = 1$. Notamos que X no puede pertenecer a $\text{conv}(\mathcal{O}_C)$, ya que por el ítem (4) de la proposición previa, si fuese así, debería ser $X \in \mathcal{O}_C$ por ser un elemento de la cápsula convexa de la órbita

de norma 1. Esto nos lleva a una contradicción ya que implicaría que existe una matriz unitaria $U \in \mathcal{U}_3$ de forma que $X = UCU^*$, pero X e Y tienen autovalores distintos. Luego en este caso $\text{conv}(\mathcal{O}_C) \subsetneq B_1$. Ver Proposición 3.3.10 para mas resultados que relacionan la cápsula convexa de la órbita con la bola unitaria de la norma.

2.3. Cuerpos convexos y normas.

En esta sección estudiaremos la relación entre las normas de Finsler Ad-invariantes en $\mathfrak{u}_n$ y las normas simétricas en $\mathbb{R}^n$ asociadas que se obtienen mirando los autovalores de las matrices. Este proceso es un poco más fino y será estudiado con detalle mediante la utilización del concepto de MAB definido en esta sección. Presentaremos los resultados originales publicados en [47], los cuales están dados en la Proposición 2.3.4 y en los teoremas 2.3.6 y 2.3.8. Finalmente caracterizaremos las normas de Finsler Ad-invariantes que provienen de un producto interno en la Proposición 2.3.10.

Definición 2.3.1. Sea $0 \in B \subset \mathbb{R}^n$, decimos que B es *simétrico* si para cada permutación $\sigma \in S_n$ y cada $(x_1, \dots, x_n) \in B$ tenemos que

$$\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \in B.$$

Si $F : \mathbb{R}^n \rightarrow \mathbb{R}$, decimos que F es *simétrica* si $F \circ \sigma = F$, donde σ debe de interpretarse como una permutación de variables. Notamos que F es simétrica si y solo si los conjuntos de nivel de F lo son.

Definición 2.3.2 (MAB). Decimos que $\mathfrak{M} \subseteq \mathfrak{u}_n$ es una MAB si es una subálgebra de Lie abeliana y maximal con dicha propiedad.

En el siguiente teorema caracterizaremos como son todas estas subálgebras abelianas.

Teorema 2.3.3. Sea $\mathfrak{M} \subseteq \mathfrak{u}_n$ una MAB entonces existe un operador unitario $U \in \mathcal{U}_n$ de forma tal que $U^* \mathfrak{M} U = \{X \in \mathfrak{u}_n : X \text{ es diagonal}\}$. En otras palabras, las subálgebras abelianas maximales se obtienen a partir de fijar una base de proyecciones ortonormales $\{p_1, p_2, \dots, p_n\}$ y considerar los elementos en $\mathfrak{u}_n$ diagonalizables en dicha base.

Demostración. Notamos primero que el conjunto $\mathfrak{D} = \{X \in \mathfrak{u}_n : X \text{ es diagonal}\}$ es claramente una subálgebra abeliana. Para ver que es maximal consideramos $A \in \mathfrak{u}_n$ que conmuta con todo elemento de $\mathfrak{D}$ y veamos que $A \in \mathfrak{D}$. Como $iE_{k,k} \in \mathfrak{D}$ para cada $1 \leq k \leq n$ (ver notación en la Observación 2.2.6), entonces $AiE_{k,k} = iE_{k,k}A$, en particular para $1 \leq k, l \leq n$ con $k \neq l$, debe ser $A_{l,k} = (A.E_{k,k})_{l,k} = (E_{k,k}.A)_{l,k} = 0$. Como k, l son arbitrarios, hemos probado que A debe ser diagonal y luego $\mathfrak{D}$ maximal.

Sea ahora $\mathfrak{M} \subseteq \mathfrak{u}_n$ una subálgebra maximal abeliana, consideramos una base de $\mathfrak{M}$ como espacio vectorial sobre $\mathbb{R}$, la cual notamos $\{X_1, X_2, \dots, X_n\}$. Al tener n elementos que conmutan entre si, es un resultado conocido que se diagonalizan en forma simultánea, luego existe $U \in \mathcal{U}_n$ de forma tal que $X_i = UD_iU^*$ para ciertas $D_i \in \mathfrak{D}$ y $1 \leq i \leq n$. Sea $X \in \mathfrak{M}$, escribiéndolo como combinación lineal de los elementos de la base tenemos que

$$X = \sum_{i=1}^n a_i X_i = \sum_{i=1}^n a_i UD_iU^* = U \left(\sum_{i=1}^n a_i D_i \right) U^*,$$

llamando $D = \sum_{i=1}^n a_i D_i$, obtenemos que $X = UDU^*$. Esto implica que todos los elementos de $\mathfrak{M}$ se diagonalizan a partir de U y entonces $U^* \mathfrak{M} U \subseteq \mathfrak{D}$. Pero como $\mathfrak{M}$ es maximal, es evidente que $U^* \mathfrak{M} U$ también lo es, y luego $U^* \mathfrak{M} U = \mathfrak{D}$. $\square$

Dada un álgebra abeliana maximal $\mathfrak{M} \subseteq \mathfrak{u}_n$ podemos considerar las proyecciones ortonormales $(p_k)_{k=1}^n$ que la generan e identificar un elemento de dicha MAB con un vector de $\mathbb{R}^n$ a partir de mirar sus coordenadas en dicha base. Mas específicamente, si $X \in \mathfrak{M}$, entonces $X = U \text{diag}(ix_1, ix_2, \dots, ix_n) U^*$ y la identificación esta dada por $X \in \mathfrak{M} \rightarrow (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Esta identificación nos da una isometría entre $\mathfrak{M}$ con el producto interno de Hilbert-Schmidt y $\mathbb{R}^n$ con el producto interno canónico, ya que si $X, Y \in \mathfrak{M}$ entonces existe $U \in \mathcal{U}_n$ de forma tal que $X = U \text{diag}(ix_1, ix_2, \dots, ix_n) U^*$ e $Y = U \text{diag}(iy_1, iy_2, \dots, iy_n) U^*$ y luego

$$(X|Y) = \text{Tr}(XY^*) = \text{Tr}(\text{diag}(ix_1, ix_2, \dots, ix_n) \text{diag}(-iy_1, -iy_2, \dots, -iy_n)) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n,$$

es decir el producto interno canónico en $\mathbb{R}^n$ de las identificaciones.

Notamos que está U no es única, pero la elegimos acorde a la base de proyecciones ortonormales que generan la MAB y el orden de los autovalores queda determinado unívocamente.

Proposición 2.3.4. (Relación entre normas Ad-invariantes en $\mathfrak{u}_n$ y normas simétricas en $\mathbb{R}^n$).

- (1) Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$ y sea B_1 su bola abierta unitaria. Para cada MAB $\mathfrak{M} \subset \mathfrak{u}_n$, consideramos $B_{\mathfrak{M}} = B_1 \cap \mathfrak{M}$. Si a $B_{\mathfrak{M}}$ lo pensamos como subconjunto de $\mathbb{R}^n$ mediante la identificación dada previamente, tenemos que $B_{\mathfrak{M}}$ es un abierto simétrico y convexo alrededor de $0 \in \mathbb{R}^n$ y no depende de la MAB elegida.
- (2) Sea $B \subset \mathbb{R}^n$ un abierto simétrico y convexo alrededor de 0. Si pensamos a X como matriz diagonal con entradas puramente imaginarias y consideramos

$$B_1 = \text{conv}(\mathcal{O}_B) = \left\{ \sum_i \lambda_i U_i X U_i^* : X \in B, U_i \in \mathcal{U}_n, \lambda_i \geq 0, \sum \lambda_i = 1 \right\}.$$

Entonces B_1 es un entorno abierto, Ad-invariante y convexo del $0 \in \mathfrak{u}_n$ y para cada MAB $\mathfrak{M}$, tenemos que $B_1 \cap \mathfrak{M} = B$. Notamos que nuevamente esta última igualdad tiene sentido en $\mathbb{R}^n$ mediante la identificación dada previamente.

Demostración. Veamos (1). Como B_1 es abierto y convexo, se sigue que $B_{\mathfrak{M}} = B_1 \cap \mathfrak{M} \subset \mathfrak{M} \simeq \mathbb{R}^n$ es un conjunto abierto y convexo. Sea $\mathfrak{M}_1$ otra MAB, luego existe una matriz unitaria U tal que $U \mathfrak{M} U^* = \mathfrak{M}_1$. Entonces

$$B_{\mathfrak{M}_1} = B_1 \cap \mathfrak{M}_1 = U B_1 U^* \cap U \mathfrak{M} U^* = U (B_1 \cap \mathfrak{M}) U^* = U B_{\mathfrak{M}} U^*,$$

esto prueba que $B_{\mathfrak{M}}$ y $B_{\mathfrak{M}_1}$ (vistos como subconjuntos de $\mathbb{R}^n$) son lo mismo, por otro lado el conjunto que obtenemos es claramente simétrico, lo cual finaliza la demostración del primer ítem.

Para ver (2), sea $B \subset \mathbb{R}^n$ un conjunto abierto, convexo y simétrico alrededor de 0. Notamos que es evidente de su definición que B_1 que es abierto, convexo, Ad-invariante y contiene la matriz nula. Lo que resta ver es que para cualquier MAB arbitraria $\mathfrak{M}$ resulta $B_1 \cap \mathfrak{M} = B$. Para ello, fijada $\mathfrak{M}$, pensamos a

$B \subseteq \mathfrak{M}$ (lo cual podemos hacer por su identificación con $\mathbb{R}^n$). Si $Z \in B$ entonces $Z \in \mathfrak{M}$ porque lo estamos representando ahí y $Z \in B_1$ por ser $B_1 = \text{conv}(\mathcal{O}_B)$. Por otro lado si $Z \in B_1 \cap \mathfrak{M}$, entonces podemos escribir

$$Z = \sum_i \lambda_i X_i$$

donde $X_i \in \mathcal{O}_B$, los λ_i son no negativos y suman 1.

Recordando la notación de 2.2.12, 2.2.13 y el Lema 2.2.14, tenemos que para $X \in B$ vale $x_i \prec x$ para todo i , luego existen matrices doblemente estocásticas A_i que cumplen $x_i^\downarrow = A_i x^\downarrow$. Si consideramos la matriz $A = \sum_i \lambda_i A_i$ es fácil ver que esta sigue siendo una matriz doblemente estocástica y por la misma construcción tenemos que $Ax^\downarrow = z^\downarrow$, invocando nuevamente al Lema 2.2.14 se concluye que z es una permutación de las coordenadas de x y por lo tanto Z pertenece a B_1 , que es lo que queríamos probar. $\square$

Observación 2.3.5. Si consideramos el plano

$$\pi = \{(x_1, \dots, x_n) : \sum_{i=1}^n x_i = 0\},$$

y nos restringimos al plano π tenemos una versión análoga a la proposición anterior en $\mathfrak{su}_n$ donde el papel que juega $\mathbb{R}^n$, ahora lo cumple este plano.

El siguiente teorema es una consecuencia inmediata de la proposición previa y nos va a permitir relacionar las normas de Finsler Ad-invariantes en $\mathfrak{u}_n$ con las simétricas en $\mathbb{R}^n$, mostrando la biyección que hay entre las mismas.

- Teorema 2.3.6.** (1) Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$ y sea B_1 su bola unitaria abierta. Si $\mathfrak{M}$ es una MAB podemos mirar $B = B_1 \cap \mathfrak{M}$ como subconjunto de $\mathbb{R}^n$ y considerar la norma dada por el funcional de Minkowsky (ver 2.1.8) asociada a esta bola. De esta forma obtenemos una norma simétrica en $\mathbb{R}^n$ cuya bola unitaria es B . Además la norma obtenida no depende de la MAB elegida.
- (2) Si tenemos una norma en $\mathbb{R}^n$ cuya bola unitaria es B , podemos pensar a $B \subseteq \mathfrak{M}$ para cierta MAB y si miramos a $B_1 = \text{conv}(\mathcal{O}_B)$, podemos definir (nuevamente usando el funcional de Minkowsky) una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Este procedimiento no depende de la MAB elegida.
- (3) Los procedimientos dados previamente son uno el inverso del otro, lo cual nos da una biyección entre las normas de Finsler Ad-invariantes de $\mathfrak{u}_n$ y las normas de Finsler simétricas en $\mathbb{R}^n$.
- (4) Si $X \in \mathfrak{u}_n$ y x es la tira de partes imaginarias de autovalores de X entonces $|X| = |x|$, donde la primer norma la pensamos en $\mathfrak{u}_n$ y la segunda en $\mathbb{R}^n$.

Demostración. (1) Comenzado con una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, si B_1 es una bola unitaria, por la Proposición 2.3.4 tenemos que B , visto como subconjunto de $\mathbb{R}^n$, es un abierto simétrico y convexo alrededor del 0 y además no depende de la MAB elegida. Considerando el funcional de Minkowsky, obtenemos una norma cuya bola unitaria es B y en tal caso esta norma resulta de Finsler y simétrica que no depende de la MAB, ya que su bola unitaria no lo hace.

- (2) Recíprocamente si comenzamos con una norma de Finsler simétrica en $\mathbb{R}^n$ y llamamos $B \subseteq \mathbb{R}^n$ a su bola unitaria, obtenemos un abierto simétrico y convexo alrededor de cero. Pensando a B dentro de una MAB $\mathfrak{M}$ y construyendo $B_1 = \text{conv}(\mathcal{O}_B)$, tenemos por la Proposición 2.3.4, que B_1 es un entorno abierto, Ad-invariante y convexo de $0 \in \mathfrak{u}_n$. Dado que $B_1 \cap \mathfrak{M}' = B$ sin importar la masa $\mathfrak{M}'$ que elijamos, este no depende de la MAB y por lo tanto la norma de Finsler Ad-invariante cuya bola es B_1 tampoco.
- (3) Comenzamos con una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, cuya bola unitaria abierta es B_1 , si aplicamos (1) obtenemos una norma de Finsler simétrica en $\mathbb{R}^n$ cuya bola unitaria es B . Ahora, si aplicamos (2) a esta construcción tendremos otra norma de Finsler Ad-invariante en $\mathfrak{u}_n$, pero esta norma tiene la misma bola unitaria que la original, luego debe ser la misma. En el otro sentido el razonamiento es análogo utilizando nuevamente que las construcciones no dependen de que MAB elegimos.
- (4) Esto es simplemente consecuencia de que ambas normas pueden ser recuperadas utilizando el funcional de Minkowsky.

□

Observación 2.3.7. Por la identificación que hemos hecho en el teorema previo, podemos pensar en una norma de Finsler Ad-invariante en $\mathfrak{su}_2$ y al cortarla con una MAB, trabajar con un subconjunto en $\mathbb{R}^2$. La traducción de que la norma sea Ad-invariante a nivel geométrico es que el interior de la bola que se obtiene al cortar con una MAB sea un conjunto simétrico (ver Observación 2.1.7) y esta es la razón por la cual en los Ejemplos 2.1.15 utilizamos figuras rotadas. Si partimos por ejemplo de una elipse sin rotar centrada en $(0, 0)$ su funcional de Minkowski nos va a dar una norma no simétrica y luego no será el corte de una norma Ad-invariante en $\mathfrak{su}_2$ con una MAB.

Teorema 2.3.8. (1) Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{su}_n$ y sea $B_1 \subseteq \mathfrak{su}_n$ su bola unitaria abierta. Si $\mathfrak{M}$ es una MAB en $\mathfrak{u}_n$, podemos mirar $B = B_1 \cap \mathfrak{M}$ como subconjunto de $\mathbb{R}^n \cap \pi$, donde

$$\pi = \{(x_1, \dots, x_n) : \sum_{i=1}^n x_i = 0\},$$

y considerar la norma dada por el funcional de Minkowsky en el espacio vectorial π (ver 2.1.8) asociada a esta bola. De esta forma obtenemos una norma simétrica en $\pi \simeq \mathbb{R}^{n-1}$ cuya bola unitaria es B . Además la norma obtenida no depende de la MAB elegida.

- (2) Si tenemos una norma en $\mathbb{R}^{n-1}$, podemos pensar a la bola unitaria de la norma sobre el plano π de $\mathbb{R}^n$ que mencionamos en (1). Ahora podemos repetir el procedimiento del teorema anterior y pensar a $B \subseteq \mathfrak{M}$ para cierta MAB. Si miramos a $B_1 = \text{conv}(\mathcal{O}_B)$, podemos definir, nuevamente usando el funcional de Minkowsky una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Este procedimiento no depende de la masa elegida.
- (3) Los procedimientos dados previamente son uno el inverso del otro, lo cual nos da una biyección entre las normas de Finsler Ad-invariantes de $\mathfrak{su}_n$ y las normas de Finsler simétricas sobre el plano π del primer ítem.

- (4) Si $X \in \mathfrak{su}_n$ y x es la tira de partes imaginarias de autovalores de X entonces $|X| = |x|$, donde la primer norma la pensamos en $\mathfrak{su}_n$ y la segunda en π .

Demostración. La demostración de este teorema es análoga a la del teorema anterior, ya que este resultado es simplemente una reescritura del previo en $\mathfrak{su}_n$. $\square$

Para finalizar este capítulo, como ya mencionamos previamente, vamos a caracterizar todas las normas de Finsler Ad-invariantes en $\mathfrak{u}_n$ que provienen de un producto interno. Comenzamos con una observación.

Observación 2.3.9. Sea $|\cdot|$ una norma de Finsler en un espacio vectorial E . Si dicha norma proviene de un producto interno $\langle \cdot, \cdot \rangle$, entonces es una norma totalmente homogénea, es decir, una norma en el sentido usual. Esto se debe a que si $\lambda \in \mathbb{R}$ entonces

$$|\lambda X|^2 = \langle \lambda X, \lambda X \rangle = \lambda^2 \langle X, X \rangle = \lambda^2 |X|^2,$$

y luego $|\lambda X| = |\lambda| |X|$.

Proposición 2.3.10. Sea $|\cdot|$ una norma de Finsler Ad-invariante que proviene de un producto interno en $\mathfrak{u}_n$. Si consideramos la descomposición de $\mathfrak{u}_n$ como $\mathfrak{u}_n = \text{gen}\{iI\} \oplus \mathfrak{su}_n$, donde I es la matriz identidad y $\mathfrak{su}_n = \{X \in \mathfrak{u}_n : \text{Tr}(X) = 0\}$, entonces existen escalares positivos a y b de forma tal que para cada $X = S + itI$ con $S \in \mathfrak{su}_n$ resulta

$$|X| = \sqrt{a \|S\|_F^2 + bt^2},$$

en particular, toda norma de Finsler Ad-invariante en $\mathfrak{su}_n$ que proviene de un producto interno es un múltiplo positivo de la norma de Frobenius.

Demostración. Sea $\langle \cdot, \cdot \rangle$ el producto interno del cual proviene nuestra norma. Utilizando la identidad de polarización

$$\langle X, Y \rangle = \frac{1}{4} (|X+Y|^2 - |X-Y|^2),$$

es evidente que el producto interno también es Ad-invariante, es decir

$$\langle UXU^*, UYU^* \rangle = \langle X, Y \rangle \quad \text{para todo } U \in \mathcal{U}_n.$$

Por otro lado, por el Teorema de Representación de Riesz, existe una transformación lineal $f : \mathfrak{u}_n \longrightarrow \mathfrak{u}_n$ definida positiva (estrictamente) que cumple

$$\langle X, Y \rangle = (f(X)|Y) \quad \text{para todo } X, Y \in \mathfrak{u}_n.$$

Veamos que si $X \in \mathfrak{u}_n$ y $U \in \mathcal{U}_n$ entonces $f(UXU^*) = Uf(X)U^*$. Para ello notamos que, para $Y \in \mathfrak{u}_n$,

$$(f(UXU^*)|UYU^*) = \langle UXU^*, UYU^* \rangle = \langle X, Y \rangle = (f(X)|Y) = (Uf(X)U^*|UYU^*),$$

donde la última igualdad es consecuencia de que la norma de Frobenius es Ad-invariante como vimos en la Observación 2.2.5 y por ende el producto interno $(\cdot|\cdot)$ también, como observamos previamente en esta demostración.

Restando ambos miembros de la igualdad previa obtenemos que

$$\begin{aligned} 0 &= (f(UXU^*) - Uf(X)U^*|UYU^*) = \text{Tr}((f(UXU^*) - Uf(X)U^*)UY^*U^*) \\ &= \text{Tr}(U^*(f(UXU^*) - Uf(X)U^*)UY^*) = (U^*(f(UXU^*) - Uf(X)U^*)U|Y), \end{aligned}$$

y como esto vale para todo $Y \in \mathfrak{u}_n$ entonces $U^*(f(UXU^*) - Uf(X)U^*)U = 0$, de donde se sigue de forma inmediata que

$$f(UXU^*) = Uf(X)U^* \quad (2.1)$$

Sea $\mathfrak{su}_n = \{X \in \mathfrak{u}_n : \text{Tr}(X) = 0\}$, como esto es un hiperplano de $\mathfrak{u}_n$ y $iI \notin \mathfrak{su}_n$, tenemos la descomposición $\mathfrak{u}_n = \text{gen}\{iI\} \oplus S$. Lo primero que queremos ver es que S es f -invariante. Recordamos que toda matriz A de $n \times n$ se puede descomponer como $A = A_h + A_{ah}$ donde $A_h = \frac{A+A^*}{2}$ (parte hermitiana) y $A_{ah} = \frac{A-A^*}{2}$ (parte antihermitiana), además esta descomposición es única como suma de una matriz hermitiana y una antihermitiana. Por otra una matriz X tiene traza cero si y solo si se puede escribir como un conmutador $X = [A, B]$ (este resultado fue demostrado en 1937 por Shoda [63] en cualquier cuerpo de característica cero). Como estamos trabajando con matrices de traza cero antihermitianas, es decir $X \in \mathfrak{u}_n$, queremos ver como queda esta descomposición en dicho caso, para ello, si $X \in \mathfrak{u}_n$

$$X = [A, B] = [A_h + A_{ah}, B_h + B_{ah}] = [A_h, B_h] + [A_{ah}, B_{ah}] + [A_h, B_{ah}] + [A_{ah}, B_h],$$

estudiando cada corchete, podemos ver fácilmente que $X_{ah} = [A_h, B_h] + [A_{ah}, B_{ah}]$ y $X_h = [A_h, B_{ah}] + [A_{ah}, B_h]$ y como en nuestro caso $X = X_{ah}$ tenemos entonces que

$$X = [A_h, B_h] + [A_{ah}, B_{ah}].$$

Luego, utilizando la linealidad de f , tenemos que

$$f(X) = f([A_h, B_h]) + f([A_{ah}, B_{ah}]) = -f([iA_h, iB_h]) + f([A_{ah}, B_{ah}]).$$

Finalmente, para ver que S es f -invariante, solo resta ver que si $X, Y \in \mathfrak{u}_n$ entonces $\text{Tr}(f([X, Y])) = 0$. Para ello notamos que $e^Y \in \mathcal{U}_n$ entonces, utilizando 2.1, se sigue que para todo $t \in \mathbb{R}$ vale

$$f(e^{tY} X e^{-tY}) = e^{tY} f(X) e^{-tY}$$

derivando esta expresión en $t = 0$ obtenemos que

$$f(YX - XY) = Yf(X) - f(X)Y,$$

es decir, $f([Y, X]) = [Y, f(X)]$. Como el miembro de la derecha tiene traza cero por ser un conmutador, podemos afirmar que $\text{Tr}(f([Y, X])) = 0$ para todo par de matrices $X, Y \in \mathfrak{u}_n$ y S resulta f invariante.

Sea ahora $V \in \mathfrak{su}_n$ un autovector de autovalor λ de f . Recordamos que como f surge de la Representación de Riesz de un producto interno, necesariamente $\lambda > 0$. Afirmamos que

$$\mathfrak{su}_n = \text{gen}\{UVU^* : U \in \mathcal{U}_n\}.$$

Es claro que S contiene al subespacio mencionado, ya que por la ciclicidad de la traza, $\text{Tr}(UVU^*) = 0$. Para ver la otra contención vamos a volver a utilizar un resultado usado en esta tesis y cuya demostración

puede ser encontrada en [35]. El resultado es el siguiente: si $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ es una tira de números no nula que suma cero, entonces existen $n - 1$ permutaciones de dicha tira de forma tal que cuando miramos el espacio generado por dichas permutaciones, obtenemos $\{x \in \mathbb{R}^n : x_1 + x_2 + \dots + x_n = 0\}$.

Aplicamos esto de la siguiente forma: sea $v = (v_1, v_2, \dots, v_n)$ la tira de autovalores de V . Primero vamos a considerar $U \in \mathcal{U}_n$ de forma tal que $UVU^* = \text{idiag}(v_1, v_2, \dots, v_n)$. Al conjugar por matrices unitarias, podemos obtener cualquier matriz diagonal cuya diagonal este formada por alguna permutación de los elementos de la diagonal de UVU^* . Sean entonces $U_1, U_2, \dots, U_{n-1} \in \mathcal{U}_n$ de forma tal que las diagonales de $D_i = U_i UVU^* U_i^*$ forman una base de $\{x \in \mathbb{R}^n : x_1 + x_2 + \dots + x_n = 0\}$. En dicho caso, toda matriz diagonal en $\mathfrak{su}_n$ se puede escribir como combinación lineal de las matrices $(D_i)_i$. Dado que toda matriz en $\mathfrak{su}_n$ es de la forma $\tilde{U} D \tilde{U}^*$ con $\tilde{U} \in \mathcal{U}_n$ podremos escribir a cualquier matriz en $\mathfrak{su}_n$ como combinación lineal de elementos en $\text{gen}\{UVU^* : U \in \mathcal{U}_n\}$, lo cual prueba la contención.

Notamos que si $X \in \mathfrak{su}_n$ entonces $X = \sum_{i=1}^k a_i U_i V U_i^*$ y luego

$$\begin{aligned} |X|^2 &= \langle X, X \rangle = (f(X), X) = (f(\sum_{i=1}^k a_i U_i V U_i^*), X) \\ &= (\sum_{i=1}^k a_i U_i f(V) U_i^*, X) = (\lambda \sum_{i=1}^k a_i U_i V U_i^*, X) = \lambda (\sum_{i=1}^k a_i U_i V U_i^*, X) = \lambda \|X\|_F^2, \end{aligned}$$

por lo tanto, $|X| = \sqrt{\lambda} \|X\|_F$ en los elementos de traza cero. Esto nos da la igualdad buscada en $\mathfrak{su}_n$ y nos dice además que f tiene un único autovalor, pues podríamos repetir el procedimiento con otro y llegar a la misma igualdad para todo $X \in \mathfrak{su}_n$, en definitiva, la función f restringida a $\mathfrak{su}_n$ es la multiplicación por $\sqrt{\lambda}$.

Sea $X = S + itI$ un elemento arbitrario de $\mathfrak{u}_n$, llamando $a = \sqrt{\lambda}$ y $b = \langle iI, iI \rangle$ obtenemos

$$\begin{aligned} |S + itI|^2 &= \langle S + itI, S + itI \rangle = \langle S, S \rangle + 2t \langle S, iI \rangle + t^2 \langle iI, iI \rangle \\ &= a \|S\|_F^2 + 2t(f(S)|iI) + bt^2 \\ &= a \|S\|_F^2 + 2ta(S|iI) + bt^2 \\ &= a \|S\|_F^2 + 2ta \text{Tr}(-iS) + bt^2 \\ &= a \|S\|_F^2 + bt^2, \end{aligned}$$

resultando entonces que $|S + itI| = \sqrt{a \|S\|_F^2 + bt^2}$, como se quería demostrar. $\square$

Capítulo 3

Funcionales, conmutadores y dual polar

3.1. Caracterización de funcionales que realizan la norma

Recordamos que si $X \in \mathfrak{u}_n$ y φ es un funcional que realiza la norma de X , entonces existe $N \in \mathfrak{u}_n$ de forma tal que $\varphi(\cdot) = (N|\cdot)$. Además, dado que $\|\varphi\| = 1$, por la Proposición 2.1.18 debe ser $|N|' = 1$. En esta sección vamos a estudiar las propiedades de esta matriz N que representa vía el producto interno al elemento φ del espacio dual. Los resultados originales más importantes de esta sección publicados en [47] son la Proposición 3.1.8, el Teorema 3.1.13 y el Teorema 3.1.18.

Comenzamos con una observación de índole general:

Observación 3.1.1. Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Si $|V| = 1$, para cada funcional φ que realiza la norma de V tenemos que existe $N \in \mathfrak{u}_n$ de *norma dual* igual a 1, i.e. $|N|' = 1$, tal que $\varphi(V) = (N|V) = 1$. Luego si $\varphi' = (V|\cdot)$, como $|V| = 1$, este funcional es de norma 1 en el espacio doble dual. Por otro lado $\varphi'(N) = (V|N) = |V| = 1 = |N|'$, luego φ' realiza la norma de N en el espacio dual. Sea $F_{\varphi'} \subset \mathfrak{u}_n$ la cara de la esfera en la norma dual dada por el funcional φ' . Claramente $\varphi = (N|\cdot)$ es el único funcional que realiza la norma de V si y solo si $F_{\varphi'} = \{N\}$ y además $F_{\varphi} = \{V\}$ si y solo si $\varphi' = (V|\cdot)$ es el único funcional que realiza la norma de N en el espacio dual.

Lema 3.1.2. (Naturalidad de la exponencial) Sean $X, Y \in \mathfrak{u}_n$ entonces el elemento e^Y es unitario y tenemos que

$$\text{Ad}_{e^Y} X = e^{\text{ad}_Y} X.$$

En particular, la expresión $\text{Ad}_{e^{sY}} X = e^{s \text{ad}_Y} X = \sum_{k=0}^{\infty} \frac{s^k \text{ad}_Y^k}{k!} X$, es una función analítica en $\mathbb{R}$.

Demostración. El hecho de que la exponencial de un elemento de $\mathfrak{u}_n$ nos da un elemento unitario es simplemente una consecuencia del desarrollo en serie de la exponencial y las propiedades de adjuntar. Para mostrar la igualdad consideramos $B(M_n(\mathbb{C}))$ el conjunto de operadores lineales acotados en $M_n(\mathbb{C})$ y la ecuación diferencial para $f : \mathbb{R} \rightarrow B(M_n(\mathbb{C}))$ dada por $f'(s) = f(s) \circ \text{ad}_Y$ con condición inicial $f(0) = id$.

Sabemos que esta ecuación tiene única solución, lo que queremos ver es que tanto $f(s) = \text{Ad}_{e^{sY}}$ como $g(s) = e^{s \text{ad}_Y}$ son soluciones. Luego será $f = g$ y evaluando en $s = 1$ tendremos la igualdad buscada como operadores.

Notamos que la condición inicial se cumple de forma evidente. Calculamos las derivadas de ambas expresiones:

$$g'(s) = \left(\sum_{k=0}^{\infty} \frac{s^k \operatorname{ad}_Y^k}{k!} \right)' = \sum_{k=0}^{\infty} \frac{k s^{k-1} \operatorname{ad}_Y^k}{k!} = \left(\sum_{k=1}^{\infty} \frac{s^{k-1} \operatorname{ad}_Y^{k-1}}{(k-1)!} \right) \circ \operatorname{ad}_Y = g \circ \operatorname{ad}_Y.$$

Por otro lado si consideramos $f(s)(X) = e^{sY} X e^{-sY}$ y derivamos

$$f'(s)X = (f(s)X)' = e^{sY} Y X e^{-sY} - e^{sY} X Y e^{-sY} = e^{sY} (YX - XY) e^{-sY} = f(s) \circ \operatorname{ad}_Y(X),$$

luego $f'(s) = g'(s)$ para todo $s \in \mathbb{R}$ y se sigue el resultado por lo observado previamente. $\square$

Observación 3.1.3. Normalmente no usaremos todo el desarrollo de la expresión $\operatorname{Ad}_{e^{sY}} X$, sino que lo utilizaremos hasta cierto término. Por ejemplo a orden 1 y orden 2 tenemos las siguientes expresiones

$$\operatorname{Ad}_{e^{sY}} X = e^{\operatorname{ad}_{sY}} X = X + s[Y, X] + o(s^2), \quad (3.1)$$

$$\operatorname{Ad}_{e^{sY}} X = e^{\operatorname{ad}_{sY}} X = X + s[Y, X] + \frac{s^2}{2}[Y, [Y, X]] + o(s^3). \quad (3.2)$$

Lema 3.1.4. Sean $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, $0 \neq X \in \mathfrak{u}_n$ y $\varphi = (N|\cdot|)$ un funcional que realiza la norma de X o de $-X$. Para $Y \in \mathfrak{u}_n$ tenemos que

- (1) $\varphi([X, Y]) = 0$ (equivalentemente $\varphi \circ \operatorname{ad} X \equiv 0$ en $\mathfrak{u}_n$) y N conmuta con X .
- (2) Si φ realiza la norma de X entonces $\varphi([Y, [Y, X]]) \leq 0$.

Demostración. Por (3.1) tenemos que

$$\operatorname{Ad}_{e^{sY}} X = X + s[Y, X] + o(s^2).$$

Luego, para cada funcional φ que realiza la norma de X , se sigue que

$$\varphi(s[Y, X] + o(s^2)) = \varphi(\operatorname{Ad}_{e^{sY}} X - X) = \varphi(\operatorname{Ad}_{e^{sY}} X) - \varphi(X) \leq |\operatorname{Ad}_{e^{sY}} X| - |X| = 0,$$

dividiendo por $s > 0$ y tomando $s \rightarrow 0^+$, obtenemos $\varphi([Y, X]) \leq 0$. Reemplazando Y por $-Y$, conseguimos la otra desigualdad y luego queda probado que $\varphi([Y, X]) = 0$. La prueba en el caso de que $-\varphi(X) = |-X|$ es análoga, finalizando así la demostración de la primera afirmación.

Por otro lado para cada $Z \in \mathfrak{u}_n$ tenemos que

$$0 = \varphi([X, Z]) = -\operatorname{Tr}(N[X, Z]) = (N|[X, Z]) = (Z|[N, X])$$

donde usamos las identidades cíclicas de la Observación 2.2.8. Tomando $Z = [N, X]$ obtenemos $\|[N, X]\|_F^2 = 0$, luego $[N, X] = 0$. Por otra parte, aplicando φ a (3.2) y usando que $\varphi([Y, X]) = 0$, tenemos que

$$\varphi(X) + \frac{s^2}{2}\varphi([Y, [Y, X]]) + o(s^3) = \varphi(\operatorname{Ad}_{e^{sY}} X) \leq |\operatorname{Ad}_{e^{sY}} X| = |X| = \varphi(X),$$

en tal caso $\frac{s^2}{2}\varphi([Y, [Y, X]]) + o(s^3) \leq 0$. Dividiendo en este caso por s^2 y tomando $s \rightarrow 0$, queda demostrada la segunda afirmación. $\square$

Observación 3.1.5. Dado que $g : s \mapsto \varphi(e^{s \operatorname{ad}_Y} X)$ es una función analítica y que tiene un máximo en $s = 0$. Una posibilidad es que todas las derivadas se anulen en $s = 0$; esto solo puede pasar si $\varphi(e^{s \operatorname{ad}_Y} X) \equiv g(0) = |X|$. Si este no fuese el caso, todas las derivadas van a dar cero hasta un cierto k par en el cual debe ser $g^{(k)}(0) < 0$. Como $g^{(m)}(0) = \varphi((\operatorname{ad}_Y)^m(X))$, lo que tenemos es que $\varphi((\operatorname{ad}_Y)^m(X)) = 0$ para todo $1 \leq m \leq k-1$ y $\varphi((\operatorname{ad}_Y)^k(X)) < 0$ para dicho k par.

Notación 3.1.6. En las siguientes proposiciones y observaciones, vamos a indexar los elementos como $v_1, v_2, \dots, v_F$, cuando escribamos $k \in F$, nos estaremos refiriendo a que k recorre los valores desde 1 hasta F . Nos permitiremos este leve abuso de notación para facilitar la lectura. Recordamos además que $\|\cdot\|_F$ es la norma de Frobenius, la cual ya utilizamos repetidamente a lo largo de este trabajo.

Definición 3.1.7 (Descomposición Diagonal/Codiagonal). Sea $0 \neq V \in \mathfrak{u}_n$, diagonalizando tenemos que $V = i \sum_{k \in F} v_k P_k$ con $v_1 > v_2 > \dots > v_F$, donde los autovalores están tomados todos distintos y de forma decreciente, por otro lado, P_k es la proyección ortogonal en el autoespacio asociado al autovalor v_k y claramente tenemos que $P_k P_j = 0$ para $k \neq j$. Para cada $X \in \mathfrak{u}_n$ podemos descomponer $X = X_D + X_C$, donde

$$X_D = \sum_k P_k X P_k = \sum_k X_k \quad \text{y} \quad X_C = \sum_{k \neq j} P_k X P_j = \sum_{k \neq j} X_{kj}.$$

Dichos elementos son conocidos como *parte diagonal por bloques* y *parte codiagonal por bloques* de X , las cuales están asociadas a la descomposición de V . Notar que bien podría ser que $[X_D, X_C] \neq 0$, sin embargo tenemos

$$X = X_D + X_C, \quad [X, V] = [X_C, V] \quad \text{y} \quad [X_D, V] = 0.$$

Sea $\varphi = (N|\cdot) = -\operatorname{Tr}(N\cdot)$ un funcional que realiza la norma de V . Dado que N conmuta con V por el Lema 3.1.4, tenemos que $N = i \sum_{k \in F} N_k$ con N_k autoadjunto y $N_k N_j = 0$ para $k \neq j$ (N es diagonal por bloques).

En el Lema 2.2.21.(1) estudiamos los funcionales que realizan la norma para el caso de normas orbitales. A continuación, vamos a estudiar estos funcionales en el caso general de una norma de Finsler Ad-invariante:

Proposición 3.1.8 (Caracterización de funcionales que realizan la norma). Sean $0 \neq V = i \sum_{k \in F} v_k P_k$, $N = i \sum_k N_k$ con $\varphi = (N|\cdot)$ un funcional que realiza la norma de V y F_φ la cara dada por φ .

- (1) Si $j > k$, cualquier autovalor de N_k es mayor o igual que todos los autovalores de N_j .
- (2) Si alguno de ellos es igual, luego permutando las proyecciones 1-dimensionales inducidas, obtenemos otra $\tilde{V} \in F_\varphi$.
- (3) Si $F_\varphi = \{V\}$, la desigualdad establecida en el primer ítem es estricta.
- (4) Dado $s \in F$, intercambiando proyecciones 1-dimensionales de N_s correspondientes a distintos autovalores, nos da otra ϕ que realiza la norma de V .
- (5) Si φ es el único funcional que realiza la norma de V , luego N es diagonal, i.e. $N_k = \sum_{k \in F} n_k P_k$, con $n_{k+1} \geq n_k$.

En particular si la norma es estrictamente convexa y suave, debe ser $N = \sum_{k \in F} n_k P_k$, con los n_k estrictamente decrecientes.

Demostración. Tenemos que

$$|V| = \varphi(V) = (N|V) = -\text{Tr}(NV) = -\text{Tr}(i \sum_s N_s i \sum_l v_l P_l) = \sum_s \sum_l v_l \text{Tr}(N_s P_l) = \sum_s v_s \text{Tr}(N_s). \quad (3.3)$$

Ahora, si $U \in \mathcal{U}_n$

$$-\text{Tr}(U^* NUV) = -\text{Tr}(NUVU^*) = \varphi(UVU^*) \leq |UVU^*| = |V| = -\text{Tr}(NV), \quad (3.4)$$

lo cual implica que si $\tilde{V} = UVU^*$ entonces $\varphi(\tilde{V}) \leq |\tilde{V}| = |V|$. Sea $\{e_1^s, \dots, e_{d(s)}^s\}$ una base ortonormal de $\text{Ran}(P_s)$ (en esta parte $d(s) = \dim(P_s)$), diagonalizando cada N_s tenemos que

$$N_s e_l^s = \lambda_l(N_s) e_l^s \quad l = 1, \dots, d(s)$$

para cada $s \in F$; podemos asumir además (eventualmente permutando los elementos de la base elegida) que los $\lambda_l(N_s)$ están en orden no creciente. Como $\sum_s P_s = 1$, si ponemos todas esas bases juntas en el orden correspondiente, obtenemos una base que diagonaliza a N y a V de forma simultánea. Sea $U \in \mathcal{U}_n$ la matriz unitaria que se obtiene al intercambiar $e_{d(k)}^k$ con e_1^j (el último autovector de N_k con el primer autovector de N_j), con $j > k$. Razonando como en (3.3)

$$|V| = \varphi(V) = -\text{Tr}(NV) = \sum_{s \in F} v_s \sum_{l=1}^{d(s)} \lambda_l(N_s),$$

mientras que $-\text{Tr}(U^* NUV) = -\text{Tr}(NUVU^*)$ es la misma suma intercambiando los dos autovalores de N considerados. Luego restando ambas expresiones y cancelando los términos repetidos, por (3.4) obtenemos:

$$v_j \lambda_{d(k)}(N_k) + v_k \lambda_1(N_j) \leq v_k \lambda_{d(k)}(N_k) + v_j \lambda_1(N_j).$$

De donde deducimos que $(v_k - v_j)(\lambda_{d(k)}(N_k) - \lambda_1(N_j)) \geq 0$, y como $v_k > v_j$ cuando $j > k$, se sigue la conclusión de la primera afirmación. Notamos que si $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$ tenemos que

$$\varphi(\tilde{V}) = -\text{Tr}(NUVU^*) = -\text{Tr}(U^* NUV) = -\text{Tr}(NV) = \varphi(V) = |V| = |UVU^*| = |\tilde{V}|,$$

luego $\tilde{V} \in F_\varphi$, lo cual prueba la segunda afirmación. Si la cara tiene un solo elemento y vale la igualdad $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$, luego $UVU^* = \tilde{V} = V$, pero esto no es posible dado que $v_k \neq v_j$. Se sigue que en ese caso debemos tener una desigualdad estricta, lo cual prueba la tercer afirmación. Si permutamos dos autoproyecciones de N_k , en tal caso existe $U \in \mathcal{U}_n$ de forma que $\text{Ran}(UN_k U^*) = \text{Ran}(N_k) = \text{Ran}(P_k)$ que actúa permutando en el bloque en cuestión y U es la identidad en los otros bloques, si consideramos $\phi = (UNU^*|\cdot)$ tenemos que, desarrollando de forma similar a (3.3)

$$\begin{aligned} \phi(V) &= \sum_{s \neq k} v_s \text{Tr}(N_s) + \sum_l v_l \text{Tr}(UN_k U^* P_l) = \sum_{s \neq k} v_s \text{Tr}(N_s) + v_k \text{Tr}(UN_k U^* P_k) \\ &= \sum_{s \neq k} v_s \text{Tr}(N_s) + v_k \text{Tr}(N_k) = (N|V) = \varphi(V) = |V|. \end{aligned}$$

Por otro lado para $Z \in \mathfrak{u}_n$

$$\phi(Z) = -\text{Tr}(UNU^* Z) = -\text{Tr}(NU^* ZU) = (N|UZU^*) \leq \|\varphi\| |UZU^*| = |Z|,$$

que combinado con la igualdad previa, prueba que $\|\phi\| = 1$ y obtenemos así un nuevo funcional que realiza la norma de V , lo cual prueba (4). Finalmente, si hay exactamente un funcional que realiza la norma de V , luego los autovalores de cada N_s deben ser iguales, ya que caso contrario podríamos armarnos otro funcional que realiza la norma como hicimos para probar (4). Concluimos que $N_s = n_s P_s$ para cada s . $\square$

Observación 3.1.9. Dado que N conmuta con V , y cualquier reordenamiento de autovectores dentro de cada bloque N_k es admisible, si es necesario, podemos asumir que los autovalores de N están todos ordenados en forma decreciente. Por otro lado, es claro que si $\lambda_{d(k)}(N_k)$ es el menor autovalor de N_k , luego

$$N_k \geq \lambda_{d(k)}(N_k) P_k.$$

De la misma forma, si $\lambda_1(N_j)$ es el mayor autovalor de N_j , tenemos

$$N_j \leq \lambda_1(N_j) P_j.$$

Ambas desigualdades son con respecto al orden parcial de matrices hermitianas,

$$X \geq Y \quad \text{sii} \quad \langle X\xi, \xi \rangle \geq \langle Y\xi, \xi \rangle \quad \forall \xi \in \mathbb{C}^n.$$

Observación 3.1.10. Si $[X, [X, Y]] = 0$ ó $[Y, [Y, X]] = 0$, luego $[X, Y] = 0$, dado que

$$\|[X, Y]\|_F^2 = -(Y|[X, [X, Y]]) = -([Y, [Y, X]]|X).$$

Para la norma de Frobenius, el único funcional que realiza la norma de $0 \neq V \in \mathfrak{u}_n$ está dado por $\varphi = (\frac{V}{\|V\|_F} | \cdot)$, luego $N_V = \frac{V}{\|V\|_F}$. Por lo dicho previamente, en el contexto de la norma de Frobenius tenemos que

$$\varphi_V([X, [X, V]]) = 0$$

implica que $[X, V] = 0$. Lo que vamos a hacer a continuación, es analizar esta condición para otras normas de Finsler Ad-invariantes. En particular vamos a probar que la afirmación previa, sigue siendo válida en el caso de normas estrictamente convexas.

Observación 3.1.11. Sea $V = i \sum_k v_k P_k$ con $v_1 > v_2 > \dots > v_F$, al igual que antes. Si $X = X_D + X_C$ es la descomposición de X en sus partes diagonal y codiagonal con respecto a V se tiene que X_D conmuta con V y luego $[X, V] = [X_C, V]$ (esto es consecuencia de que cada bloque de V es un múltiplo de la identidad). Por otro lado, si $\varphi = (N|\cdot)$ realiza la norma de V , por la identidad de Jacobi y (2.2.8) obtenemos

$$\begin{aligned} \varphi([X, [X, V]]) &= \varphi([X, [X_C, V]]) = (N|[X, X_C], V] + [X_C, [X, V]]) \\ &= -(N|[V, [X, X_C]]) + (N|[X_C, [X_C, V]]) \\ &= -([N, V]|[X, X_C]) + ([N, X_C]|[X_C, V]) = ([N, X_C]|[X_C, V]) \end{aligned}$$

donde usamos que N conmuta con V , en conclusión

$$\varphi([X, [X, V]]) = ([N, X_C]|[X_C, V]). \quad (3.5)$$

Observación 3.1.12. Para el siguiente teorema vamos a necesitar usar alguna propiedades del orden de matrices. Si $0 \leq A \leq B$ para A, B matrices hermitianas, es un resultado conocido que $0 \leq \sqrt{A} \leq \sqrt{B}$ y por otro lado es bien sabido que elevar al cuadrado no preserva el orden necesariamente. Sin embargo, en el caso particular que las matrices A y B conmuten, la propiedad es válida. Esto es ya que si $0 \leq A \leq B$ y $AB = BA$ entonces podemos escribir

$$B^2 - A^2 = (B - A)(B + A),$$

llamando $X = B - A$ e $Y = B + A$, es claro que X e Y son positivas y conmutan, luego se escriben como $X = UD_XU^*$ y $Y = UD_YU^*$ con D_X, D_Y matrices diagonales positivas y $U \in \mathcal{U}_n$. Tendremos entonces que $XY = UD_XD_YU^*$ que es claramente positiva ya que sus autovalores son el producto de los autovalores de X con los de Y . Se concluye entonces que $0 \leq A^2 \leq B^2$, lo cual usaremos repetidamente en el siguiente teorema.

Teorema 3.1.13 (El caso de la igualdad). Sean $X, V \in \mathfrak{u}_n$, escribimos en forma diagonalizada a V como $V = i \sum_k v_k P_k$ con $v_1 > \dots > v_F$ y sea $\varphi = (N | \cdot)$ un funcional que realiza la norma de V .

- (1) $\varphi([X, [X, V]]) = 0$ si y solo si $[X_C, N] = 0$
- (2) Si $\varphi([X, [X, V]]) = 0$ y $F_\varphi = \{V\}$ entonces $[X, V] = 0$.

Demostración. Si $[X_C, N] = 0$, por (3.5) tenemos que $\varphi([X, [X, V]]) = ([N, X_C] | [X_C, V]) = 0$. Supongamos ahora que $\varphi([X, [X, V]]) = 0$, consideramos inicialmente $Z = P_j X P_k + P_k X P_j \in \mathfrak{u}_n$ y notamos que $Z = Z_C$ si $k \neq j$. Luego, por el Lema 3.1.4 y la igualdad (3.5), tenemos que

$$\begin{aligned} \text{Tr}([Z, N][Z, V]) &= -\text{Tr}([N, Z][Z, V]) = ([N, Z] | [Z, V]) \\ &= ([N, Z_C] | [Z_C, V]) = \varphi([Z, [Z, V]]) \leq 0. \end{aligned} \quad (3.6)$$

Desarrollando $[Z, N]$ obtenemos

$$\begin{aligned} [Z, N] &= i(P_j X P_k + P_k X P_j) \sum_s N_s - i \sum_s N_s (P_j X P_k + P_k X P_j) \\ &= i(P_j X N_k + P_k X N_j - N_j X P_k - N_k X P_j). \end{aligned}$$

De la misma forma, pero trabajando con $[Z, V]$, se tiene

$$[Z, V] = i(P_j X v_k P_k + P_k X v_j P_j - v_j P_j X P_k - v_k P_k X P_j) = i(v_k - v_j)(P_j X P_k - P_k X P_j).$$

Usando las igualdades previas y la ciclicidad de la traza tenemos que

$$\begin{aligned} \text{Tr}([Z, N][Z, V]) &= -(v_k - v_j) \text{Tr}((P_j X N_k + P_k X N_j - N_j X P_k - N_k X P_j)(P_j X P_k - P_k X P_j)) \\ &= -(v_k - v_j) \text{Tr}(-P_j X N_k X P_j + P_k X N_j X P_k + N_j X P_k X P_j - N_k X P_j X P_k) \\ &= -2(v_k - v_j) \text{Tr}(N_j X P_k X - N_k X P_j X). \end{aligned} \quad (3.7)$$

Usando (3.6), se sigue que

$$2(v_k - v_j) \text{Tr}(N_j X P_k X - N_k X P_j X) \geq 0 \quad \forall k \neq j. \quad (3.8)$$

Ahora pasamos al caso general y calculamos $[X_C, N]$

$$\begin{aligned} [X_C, N] &= i \sum_{r \neq s} (P_r X P_s + P_s X P_r) \sum_l N_l - i \sum_l N_l \sum_{r \neq s} (P_r X P_s + P_s X P_r) \\ &= i \sum_{r \neq s} (P_r X N_s + P_s X N_r) - i \sum_{r \neq s} (N_r X P_s + N_s X P_r) \\ &= i \sum_{r \neq s} (P_r X N_s + P_s X N_r - N_r X P_s - N_s X P_r). \end{aligned}$$

Por otra parte desarrollando $[X_C, V]$

$$\begin{aligned} [X_C, V] &= i \left(\sum_{j \neq k} (P_k X v_j P_j + P_j X v_k P_k - v_k P_k X P_j - v_j P_j X P_k) \right) \\ &= i \sum_{k \neq j} (v_k - v_j) (P_j X P_k - P_k X P_j). \end{aligned}$$

Luego $\varphi([X, [X, V]]) = \text{Tr}([X_C, N][X_C, V])$ es igual a

$$\begin{aligned} &= - \sum_{k \neq j} \sum_{r \neq s} \text{Tr}((P_r X N_s + P_s X N_r - N_r X P_s - N_s X P_r)(v_k - v_j)(P_j X P_k - P_k X P_j)) \\ &= -4 \sum_{k \neq j} (v_k - v_j) \text{Tr}(N_j X P_k X - N_k X P_j X), \end{aligned}$$

donde en la última igualdad usamos que la expresión es no nula solo si $k = r$ y $j = s$ ó $k = s$ y $j = r$. Luego nos queda para cada $k \neq j$ el doble de la expresión (3.7). Por la desigualdad (3.8), esta expresión solo puede ser cero si $\text{Tr}(N_j X P_k X - N_k X P_j X) = 0$ para todo $k \neq j$. Supongamos primero que $j > k$, y en tal caso por la Proposición 3.1.8, tenemos que $\lambda_{d(k)}(N_k) \geq \lambda_1(N_j)$. Vamos a mostrar que estos números son iguales y que

$$P_j X P_k N = P_k X N_j = \lambda_1(N_j) P_k X P_j, \quad P_j X P_k N = P_j X N_k = \lambda_1(N_j) P_j X P_k. \quad (3.9)$$

Una vez probado esto, adjuntando a ambos lados obtenemos que:

$$N P_j X P_k = N_j X P_k = \lambda_1(N_j) P_j X P_k$$

$$N P_k X P_j = N_k X P_j = \lambda_1(N_j) P_k X P_j,$$

y entonces para $j > k$ se tiene

$$P_j X P_k N = P_j X N_k = \lambda_1(N_j) P_j X P_k, \quad N P_j X P_k = N_j X P_k = \lambda_1(N_j) P_j X P_k,$$

i.e. $[P_j X P_k, N] = 0$. Por simetría, el mismo argumento funciona para $j < k$ (usando $\lambda_1(N_k)$ en vez de $\lambda_1(N_j)$), y luego sumando sobre todos los $j \neq k$ tendremos que $[X_C, N] = 0$.

Para simplificar los cambios de signos va a ser conveniente trabajar con $X = i\tilde{X}$, que satisface $\tilde{X}^* = \tilde{X}$, de hecho vamos a cometer un abuso de notación y renombrarlo simplemente X (sin usar la tilde) para no sobrecargar la notación. Observamos que X_C conmuta con N si y solo si $\tilde{X}_C$ conmuta con N .

Supongamos primero que $j > k$. Sabemos que $\text{Tr}(N_j X P_k X) = \text{Tr}(N_k X P_j X)$ con $X^* = X$. Como $N_k \geq \lambda_{d(k)}(N_k) P_k$ y $\lambda_{d(k)}(N_k) \geq \lambda_1(N_j)$ se sigue que

$$P_j X N_k X P_j \geq \lambda_{d(k)}(N_k) P_j X P_k X P_j \geq \lambda_1(N_j) P_j X P_k X P_j,$$

donde en la primera igualdad usamos que conjugar preserva desigualdades de operadores y en la segunda desigualdad usamos que $P_j X P_k X P_j = (X P_j)^* X P_j \geq 0$. Por la monotonía y la ciclicidad de la traza tenemos que

$$\text{Tr}(P_j X N_k X) \geq \lambda_{d(k)}(N_k) \text{Tr}(P_j X P_k X) \geq \lambda_1(N_j) \text{Tr}(P_j X P_k X),$$

notamos además que $\|P_k X P_j\|_F^2 = \|P_j X P_k\|_F^2 = \text{Tr}(P_j X P_k X)$ y luego

$$\text{Tr}(P_j X N_k X) \geq \lambda_{d(k)}(N_k) \|P_j X P_k\|_F^2 \geq \lambda_1(N_j) \|P_j X P_k\|_F^2.$$

Análogamente partiendo ahora de $N_j \leq \lambda_1(N_j) P_j$ obtenemos por un lado que

$$P_k X N_j X P_k \leq \lambda_1(N_j) P_k X P_j X P_k \leq \lambda_{d(k)}(N_k) P_k X P_j X P_k$$

y tomando traza

$$\text{Tr}(P_k X N_j X) \leq \lambda_1(N_j) \|P_j X P_k\|_F^2 \leq \lambda_{d(k)}(N_k) \|P_j X P_k\|_F^2.$$

Usando la hipótesis y las desigualdades previamente demostradas

$$\lambda_1(N_j) \|P_k X P_j\|_F^2 \leq \lambda_{d(k)}(N_k) \|P_k X P_j\|_F^2 \leq \text{Tr}(P_j X N_k X) = \text{Tr}(P_k X N_j X) \leq \lambda_1(N_j) \|P_j X P_k\|_F^2.$$

Como $\|P_j X P_k\|_F^2 = \|P_k X P_j\|_F^2$ tenemos que las desigualdades deben ser todas igualdades.

$$\lambda_{d(k)}(N_k) \text{Tr}(P_k X P_j X) = \text{Tr}(P_j X N_k X) = \text{Tr}(P_k X N_j X) = \lambda_1(N_j) \text{Tr}(P_k X P_j X).$$

Si $\|P_k X P_j\|_F = 0$, luego $P_k X P_j = P_j X P_k = 0$ y (3.9) es claro. Si no es así, debe ser $\lambda_{d(k)}(N_k) = \lambda_1(N_j) = \lambda$. Mas aún, observamos que

$$A = P_j X N_k X P_j - \lambda P_j X P_k X P_j \geq 0$$

pero acabamos de probar que $\text{Tr}(A) = 0$. Luego $A = 0$, y entonces $P_j X N_k X P_j = \lambda P_j X P_k X P_j$. De la misma forma, como $P_k X N_j X P_k \leq \lambda P_k X P_j X P_k$, utilizando la igualdad de las trazas obtenidas se sigue que

$$P_j X N_k X P_j = \lambda P_j X P_k X P_j \quad \text{y} \quad P_k X N_j X P_k = \lambda P_k X P_j X P_k. \quad (3.10)$$

Supongamos primero que $\lambda \geq 0$, sea $N_j = N_j^+ - N_j^-$ donde $N_j^+, N_j^- \geq 0$ es la descomposición de Hahn de N_j . Como $N_j \leq \lambda P_j$ debe ser $0 \leq N_j^+ \leq \lambda P_j$ y luego por la observación previa $0 \leq (N_j^+)^2 \leq \lambda^2 P_j$. Ahora

$$\lambda P_k X P_j X P_k = P_k X N_j X P_k = P_k X N_j^+ X P_k - P_k X N_j^- X P_k \leq \lambda P_k X P_j X P_k - P_k X N_j^- X P_k,$$

luego $0 \leq P_k X N_j^- X P_k \leq 0$ y esto es solo posible si $P_k X N_j^- X P_k = 0$. Como $P_k X N_j^- X P_k = |\sqrt{N_j^-} X P_k|^2$, se sigue que $N_j^- X P_k = 0 = P_k X N_j^-$. Por otra parte

$$\|P_k X N_j^+\|_F^2 = \text{Tr}(P_k X (N_j^+)^2 X P_k) \leq \lambda^2 \text{Tr}(P_k X P_j X P_k)$$

entonces $\|P_k X N_j^+\|_F \leq \lambda \|P_j X P_k\|_F$. Esto nos da la siguiente igualdad en la desigualdad de Cauchy-Schwarz:

$$\begin{aligned} \lambda \|P_j X P_k\|_F^2 &= \text{Tr}(P_k X N_j X P_k) = \text{Tr}(P_k X N_j^+ X P_k) = (P_k X N_j^+ | P_k X P_j) \\ &\leq \|P_k X N_j^+\|_F \|P_k X P_j\|_F \leq \lambda \|P_k X P_j\|_F^2. \end{aligned}$$

Se sigue entonces que los vectores están alineados (ver por ejemplo [44]), i.e. $P_k X N_j^+ = \mu P_k X P_j$, y multiplicando por $P_j X P_k$ a derecha resulta $\mu = \lambda$. Obtenemos entonces que

$$P_k X N_j = P_k X N_j^+ - P_k X N_j^- = P_k X N_j^+ = \lambda P_k X P_j,$$

que es la primer identidad de (3.9). Para obtener la segunda identidad, notamos que $N_k \geq \lambda P_k \geq 0$ luego (nuevamente por la observación previa) $\sqrt{N_k} \geq \sqrt{\lambda} P_k \geq 0$ y entonces $P_j X \sqrt{\lambda} P_k X P_j \leq P_j X \sqrt{N_k} X P_j$. Por otra parte, usando (3.10), tenemos que

$$\|P_j X \sqrt{N_k}\|_F^2 = \text{Tr}(P_j X \sqrt{N_k} \sqrt{N_k} X P_j) = \text{Tr}(P_j X N_k X P_j) = \lambda \|P_k X P_j\|_F^2 = \lambda \|P_j X P_k\|_F^2,$$

lo cual nos da nuevamente una igualdad en la desigualdad de Cauchy-Schwarz :

$$\begin{aligned} \sqrt{\lambda} \|P_k X P_j\|_F^2 &= \sqrt{\lambda} \text{Tr}(P_j X P_k X P_j) \leq \text{Tr}(P_j X \sqrt{N_k} X P_j) \\ &= (P_j X \sqrt{N_k} | P_j X P_k) \leq \|P_j X \sqrt{N_k}\|_F \|P_j X P_k\|_F \leq \sqrt{\lambda} \|P_j X P_k\|_F^2. \end{aligned}$$

Obtenemos entonces que $P_j X \sqrt{N_k} = \sqrt{\lambda} P_j X P_k$ y luego

$$P_j X N_k = P_j X \sqrt{N_k} \sqrt{N_k} = \sqrt{\lambda} P_j X P_k \sqrt{N_k} = \sqrt{\lambda} P_j X \sqrt{N_k} = \lambda P_j X P_k,$$

que es la segunda identidad de (3.9). Esto finaliza el caso $\lambda \geq 0$. El caso $\lambda < 0$ es similar con algunas modificaciones: comenzamos observando que $N_j \leq \lambda P_j < 0$, luego $-N_j \geq 0$ y además, usando la monotonía de la raíz cuadrada, $\sqrt{-N_j} \geq \sqrt{-\lambda} P_j \geq 0$. Luego $P_k X \sqrt{-N_j} X P_k \geq P_k X P_j \sqrt{-\lambda} X P_k$ y

$$\begin{aligned} \sqrt{-\lambda} \|P_j X P_k\|_F^2 &\leq \text{Tr}(P_k X \sqrt{-N_j} X P_k) = (P_k X \sqrt{-N_j} | P_k X P_j) \\ &\leq \|P_k X \sqrt{-N_j}\|_F \|P_k X P_j\|_F \leq \sqrt{\text{Tr}(P_k X (-N_j) X P_k)} \|P_k X P_j\|_F \\ &= \sqrt{-\lambda \text{Tr}(P_k X P_j X P_k)} \|P_k X P_j\|_F = \sqrt{-\lambda} \|P_j X P_k\|_F^2. \end{aligned}$$

Nuevamente tenemos una igualdad en la desigualdad de Cauchy-Schwarz, la cual es solamente posible si $P_k X \sqrt{-N_j} = \sqrt{-\lambda} P_k X P_j$. Al igual que antes (multiplicando por $\sqrt{-N_j}$ a derecha) se sigue que $P_k X N_j = \lambda P_k X P_j$, lo cual establece la primer identidad en (3.9). Ahora, observamos que $N_k \geq \lambda P_k$ con $\lambda < 0$, luego si $N_k = N_k^+ - N_k^-$ es la descomposición de Hahn de N_k , debe ser $-N_k^- \geq \lambda P_k$ y por la observación previa $(N_k^-)^2 = (-N_k^-)^2 \leq \lambda^2 P_k$. Razonando como el otro caso vemos primero que $P_j X N_k^+ X P_j = 0$, i.e. $P_j X N_k^+ = N_k^+ X P_j = 0$. Entonces $P_j X N_k = P_j X N_k^-$ y

$$\|P_j X N_k\|_F^2 = \text{Tr}(P_j X N_k^2 X P_j) = \text{Tr}(P_j X (N_k^-)^2 X P_j) \leq \lambda^2 \text{Tr}(P_j X P_k X P_j) = \lambda^2 \|P_k X P_j\|_F^2.$$

Nuevamente tenemos una igualdad en la desigualdad de Cauchy-Schwarz

$$\begin{aligned} -\lambda \|P_k X P_j\|_F^2 &= \text{Tr}(-P_j X N_k X P_j) = (-P_j X N_k | P_j X P_k) \leq \|P_j X N_k\|_F \|P_k X P_j\|_F \\ &\leq |\lambda| \|P_k X P_j\|_F^2 = -\lambda \|P_k X P_j\|_F^2, \end{aligned}$$

lo que nos dice que $P_j X N_k = \lambda P_j X P_k$, la cual es la segunda igualdad de la identidad (3.9). Ahora que probamos que $[X_C, N] = 0$, vamos a ver (2). Recordamos que la Proposición 3.1.8 dice que cuando existe un único φ que realiza la norma de V debe ser N debe ser diagonal, i.e. $N = i \sum_k n_k P_k$. Pero la parte

diagonal de X (con respecto V) es en tal caso la parte diagonal de X respecto a N , o equivalentemente, $[X_D, N] = 0$. Luego en este caso $[X, N] = [X_C, N] + [X_D, N] = 0 + 0 = 0$. Por otro parte si $F_\varphi = \{V\}$, citando la misma proposición no podemos tener igualdades de la forma $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$, luego en cada paso debe ser $P_k X P_j = 0$, lo cual implica que $X_C = 0$ y entonces $X = X_D$. Esto último es equivalente a que X conmute con V . $\square$

Corolario 3.1.14 (Normas estrictamente convexas). Si la norma es estrictamente y φ es un funcional que realiza la norma de V tenemos que $\varphi([X, [X, V]]) = 0$ implica que $[X, V] = 0$.

Demostración. Como la norma es estrictamente convexa, tenemos que cada cara consta solo de un punto, luego $F_\varphi = \{V\}$. Si asumimos que $\varphi([X, [X, V]]) = 0$ y aplicamos el ítem (2) del Teorema 3.1.13 se sigue $[X, V] = 0$. $\square$

Observación 3.1.15. El Teorema 3.1.13 y su corolario van a ser muy útiles para probar resultados como el Teorema 3.5.3, del cual podremos concluir una interpretación geométrica más clara. Sin embargo si miramos un ejemplo en $\mathbb{R}^3$ usando como corchete el producto vectorial y como norma la euclídea tenemos que el único funcional que realiza la norma de V es $\varphi(W) = \frac{\langle W, V \rangle}{\|V\|}$. Por otro lado la condición $\varphi([X, [X, V]]) = 0$ se traduce en $\langle X, V \rangle = \|X\| \cdot \|V\|$, la cual implica una igualdad en la desigualdad de Cauchy-Schwartz y luego X y V deben ser paralelos, es decir $[X, V] = 0$.

A continuación veremos un contraejemplo de como falla el Corolario 3.1.14 cuando la norma no es estrictamente convexa.

Ejemplo 3.1.16. Vamos a mostrar dos matrices V y X , de forma tal que existe un único funcional que realiza la norma de V con la norma dada por $|X|_1 = \text{Tr}(|X|)$ (que no es estrictamente convexa) y este funcional va a verificar que $\varphi([X, [X, V]]) = 0$, pero que sin embargo $[X, V] \neq 0$.

Sean $V, X \in \mathfrak{u}_2$ las matrices dadas por $V = \begin{pmatrix} -i & 0 \\ 0 & -2i \end{pmatrix}$ y $X = \begin{pmatrix} -2i & \frac{1}{2} \\ -\frac{1}{2} & -3i \end{pmatrix}$, calculando obtenemos que

$$[X, V] = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix} \text{ y } [X, [X, V]] = \begin{pmatrix} -\frac{i}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{i}{2} \end{pmatrix}.$$

Notamos además que $|V| = 3$. Dado φ un funcional que realiza la norma de V , por el Lema de Riesz, existe una matriz $N \in \mathfrak{u}_2$ de forma tal que $\varphi(Y) = (N|Y|)$ para todo $Y \in \mathfrak{u}_2$. Por otro lado, por el Lema 3.1.4, esta matriz debe conmutar con V . Esto implica, usando el Lema 2.3.3, que N debe ser una matriz diagonal, es decir

$$N = \begin{pmatrix} ia & 0 \\ 0 & ib \end{pmatrix},$$

con $a, b \in \mathbb{R}$. En tal caso $\varphi(V) = -a - 2b$. Como estamos trabajando con matrices diagonales, fácilmente podemos pensar que estamos en $\mathbb{R}^2$ con la norma de $l_1(\mathbb{R}^2)$ y entonces la norma del funcional va a ser la norma en $l_\infty(\mathbb{R}^2)$ del vector (ia, ib) . Tenemos entonces que $\|\varphi\| = \max\{|a|, |b|\}$ y queremos que esto sea 1 al mismo tiempo que $\varphi(V) = -a - 2b = 3$. Es fácil ver que la única forma en que esto suceda es tomar $a = -1$ y $b = -1$. Resultando así que el único funcional que realiza la norma de V es el dado por

$\varphi\left(\begin{pmatrix} ia & c \\ c^* & ib \end{pmatrix}\right) = -a - b$. Como este funcional verifica que $\varphi([X, [X, V]]) = 0$, entonces cumple todo lo pedido y el ejemplo queda finalizado.

Observación 3.1.17. Si $\varphi([X, [X, V]]) = 0$ para ciertos X, V no nulos, y φ realiza la norma de V , podemos escribir a V como una combinación convexa de los puntos extremales de la esfera de radio $|V|$, i.e. $V = \sum_i \lambda_i V_i$, con $0 < \lambda_i \leq 1$ y $\sum_i \lambda_i = 1$. Como cada V_i debe estar en la misma cara que V se sigue que φ debe realizar la norma de cada V_i . Luego tenemos que

$$0 = \varphi([X, [X, V]]) = \sum_i \lambda_i \varphi([X, [X, V_i]]).$$

Esto implica que debe ser $\varphi([X, [X, V_i]]) = 0$ para cada V_i (notar que todos los términos son menores o iguales a cero por el Lema 3.1.4). Por otro lado si $\varphi = (N|\cdot)$ y trabajamos en el espacio dual, tenemos que $\varphi = \sum_i \varphi_i = \sum_i (N^i|\cdot)$ con φ_i extremales, y se deduce de forma análoga que $\varphi_i([X, [X, V]]) = 0$.

Teorema 3.1.18. Sea $V = i \sum_k v_k P_k \in \mathfrak{u}_n$ con $v_1 > v_2 > \dots > v_F$. Existe $\varphi_0 = (N_0|\cdot)$ un funcional que realiza la norma de V tal que N_0 es diagonal, i.e. $N_0 = i \sum_{k \in F} \lambda_k P_k$. Sea $X \in \mathfrak{u}_n$, luego $\varphi_0([X, [X, V]]) = 0$ si y solo si $[X, N_0] = 0$. En particular, si hay un único funcional que realiza la norma de V debe tener la propiedad antes mencionada.

Demostración. Sea $\varphi = (N|\cdot)$ un funcional que realiza la norma de V , luego $N = i \sum_{k \in F} N_k$ con $N_k^* = N_k = N_k P_k$ para todo $k \in F$. Al igual que en la demostración de la Proposición 3.1.8, elegimos una base ortonormal $\{e_1^k, \dots, e_{d(k)}^k\}$ de $\text{Ran}(P_k)$ en la cual N_k es diagonal, es decir

$$N_k e_l^k = \lambda_l(N_k) e_l^k \quad l = 1, \dots, d(k)$$

para cada $k \in F$; juntando todas estas bases en el orden dado, obtenemos una base que diagonaliza N y V de forma simultánea. Para cada $k \in F$ y para cada $l, s \in F_k = \{1, \dots, d(k)\}$, sea $U_{k,ls}$ la matriz unitaria que intercambia e_l^k con e_s^k . Claramente

$$U_{k,ls} P_j U_{k,ls}^* = P_j \quad \forall j, k \in F,$$

y luego $U_{k,ls} V U_{k,ls}^* = V$. Observamos que para cada k , existen exactamente $C(k) = \binom{d(k)}{2}$ formas de elegir $l \neq s$ de cada conjunto F_k . Notamos $|F|$ al cardinal de F y consideramos

$$N_0 = \frac{1}{|F|} \sum_{k \in F} \frac{1}{C(k)} \sum_{l \neq s \in F_k} U_{k,ls} N U_{k,ls}^*.$$

Afirmamos que este elemento N_0 tiene todas las propiedades deseadas. Notamos primero que al considerar todas los reordenamientos de los elementos en la diagonal de cada N_k , cada nuevo bloque de la matriz N_0 va a ser un bloque diagonal formado por la media aritmética de los autovalores de N_k en su diagonal, es decir $N_0 = i \sum_k \lambda_k P_k$ con

$$\lambda_k = \frac{1}{d(k)} \sum_{l=1}^{d(k)} \lambda_l(N_k).$$

Por otro lado

$$|N_0|' \leq \frac{1}{|F|} \sum_{k \in F} \frac{1}{C(k)} \sum_{l \neq s \in F_k} |N|' = \frac{1}{|F|} |F| \frac{1}{C(k)} C(k) = 1$$

y además

$$(N_0|V) = \frac{1}{|F|} \sum_{k \in F} \frac{1}{C(k)} \sum_{l \neq s \in F_k} (N|U_{k,ls}^* V U_{k,ls}) = \frac{1}{|F|} \sum_{k \in F} \frac{1}{C(k)} \sum_{l \neq s \in F_k} |V| = |V|,$$

luego $|N_0|' = 1$ y $\varphi_0 = (N_0|\cdot)$ realiza la norma de V . Si $X \in \mathfrak{u}_n$ tenemos que

$$\varphi_0([X, [X, V]]) = (N_0|[X, [X, V]]) = ([N_0, X][X, V]),$$

de donde se sigue que si $[X, N_0] = 0$ entonces $\varphi_0([X, [X, V]]) = 0$. Por otro lado, por el Teorema 3.1.13, si $\varphi_0([X, [X, V]]) = 0$ entonces $[X_C, N_0] = 0$, pero como cada bloque de N_0 es diagonal también tenemos que $[X_D, N_0] = 0$ y luego $[X, N_0] = 0$. $\square$

Observación 3.1.19. Sean V y N_0 como en el teorema anterior. Notamos que los valores de λ_k para los cuales vale la igualdad $N_0 = i \sum_{k \in F} \lambda_k P_k$, se obtienen al hacer promedios de los autovalores de los bloques de N a los que notamos N_k . Como todos los autovalores de N_k son mayores o iguales a los de N_{k+1} (por la Proposición 3.1.8) entonces la media aritmética de los autovalores de N_k también va a ser mayor o igual a la de N_{k+1} , es decir, $\lambda_k \geq \lambda_{k+1}$. Si para algún k valiese $\lambda_k = \lambda_{k+1}$, tendríamos que $\lambda_l(N_k) = \lambda_s(N_{k+1}) = \lambda$ para todo $l \in \{1, \dots, d(k)\}$, $s \in \{1, \dots, d(k+1)\}$ y en dicho caso podríamos escribir $N_k + N_{k+1} = \lambda(P_k + P_{k+1})$. Siguiendo la demostración del Teorema 3.1.13, obtendríamos que si $X \in \mathfrak{u}_n$ es tal que $\varphi_0([X, [X, V]]) = 0$ entonces debería ser $P_k X P_{k+1} = 0$.

3.2. Derivada de una norma de Finsler

En esta sección estudiaremos las derivadas direccionales laterales de una norma de Finsler en un espacio vectorial. Sobre el final de la misma, mencionaremos algunos resultados generales sobre diferenciación de normas que necesitaremos usar y además ayudan a la comprensión de las mismas.

Observación 3.2.1. Sea $h : \mathbb{R} \rightarrow \mathbb{R}$ una función continua y convexa, entonces el límite

$$\lim_{t \rightarrow 0^+} \frac{h(x+t) - h(x)}{t}$$

existe (en el sentido que es un número real o $-\infty$). Esto es una consecuencia de que el cociente incremental decrece cuando $t \rightarrow 0^+$. Para ver esto, tomamos $0 < s < t$ y observamos que

$$h(x+s) = h\left(\frac{s}{t}(x+t) + \left(1 - \frac{s}{t}\right)x\right) \leq \frac{s}{t}h(x+t) + \left(1 - \frac{s}{t}\right)h(x),$$

luego

$$\frac{h(x+s) - h(x)}{s} \leq \frac{h(x+t) - h(x)}{t}, \quad (3.11)$$

con un razonamiento análogo se llega a que

$$\frac{h(x-t) - h(x)}{-t} \leq \frac{h(x-s) - h(x)}{-s}.$$

Proposición 3.2.2. Dado $(E, |\cdot|)$ un espacio normado con una norma de Finsler y $X, Y \in E$ no nulos. Consideramos la función $h(t) = |X + tY|$. En tal caso tenemos que

(1) Si $0 < s < t$ entonces

$$\frac{|X - tY| - |X|}{-t} \leq \frac{|X - sY| - |X|}{-s} \leq \frac{|X + sY| - |X|}{s} \leq \frac{|X + tY| - |X|}{t} \quad (3.12)$$

(2) La función h tiene derivadas laterales finitas en el origen, a las cuales notaremos $h'_-(X, Y)$ y $h'_+(X, Y)$. Las pensaremos como las derivadas laterales de la norma en X en la dirección de Y . Valen además las desigualdades $-|Y| \leq h'_-(X, Y) \leq |-Y|$ y $-|Y| \leq h'_+(X, Y) \leq |Y|$.

Demostración. Notamos que h es una función continua y convexa en $\mathbb{R}$, luego por la observación previa tenemos que, para $0 < s < t$, valen las desigualdades

$$\begin{aligned} \frac{|X + sY| - |X|}{s} &\leq \frac{|X + tY| - |X|}{t}, \\ \frac{|X - tY| - |X|}{-t} &\leq \frac{|X - sY| - |X|}{-s}. \end{aligned}$$

Por otro lado tenemos que $2|X| = |X + sY + X - sY| \leq |X + sY| + |X - sY|$, acomodando y dividiendo por $s > 0$ de ambos lados obtenemos

$$\frac{|X - sY| - |X|}{-s} \leq \frac{|X + sY| - |X|}{s},$$

de donde se sigue la desigualdad que nos faltaba probar de la primera afirmación.

Para la segunda afirmación, simplemente notamos que si $t > 0$, entonces $\frac{|X + tY| - |X|}{t} \leq \frac{t|Y|}{t} = |Y|$ y por otro lado $\frac{|X + tY| - |X|}{t} \geq \frac{|X| - |tY| - |X|}{t} = -|Y|$, de donde se deduce

$$-|Y| \leq h'_+(X, Y) \leq |Y|,$$

la demostración de las otras desigualdades es análoga. $\square$

A continuación vamos a caracterizar las derivadas laterales de la norma en términos de funcionales que realizan la norma de un elemento $X \in E$.

Definición 3.2.3. Para $0 \neq X \in E$, notamos $\mathcal{N}_X = \{\varphi \in E' : \|\varphi\| = 1 \text{ y } \varphi(X) = |X|\}$, como consecuencia del Teorema de Hahn-Banach, tenemos que este es un subconjunto no vacío de E' y además si $\varphi \in \mathcal{N}_X$ entonces $\varphi \in \mathcal{N}_{\lambda X}$ para todo $\lambda > 0$.

Proposición 3.2.4. Sean $(E, |\cdot|)$ un espacio normado de Finsler y $X, Y \in E$ con $X \neq 0$. Si $h(t) = |X + tY|$, entonces

$$h'(0^+) = h'_+(X, Y) = \lim_{t \rightarrow 0^+} \frac{|X + tY| - |X|}{t} = \max_{\varphi \in \mathcal{N}_X} \varphi(Y).$$

Demostración. Notamos que para cada $\varphi \in \mathcal{N}_X$ tenemos que

$$\frac{|X + tY| - |X|}{t} \geq \frac{\varphi(X + tY) - \varphi(X)}{t} = \varphi(Y).$$

Tomando límite se sigue que $h'(0^+) \geq \max_{\varphi \in \mathcal{N}_X} \varphi(Y)$. Lo único que resta probar es que hay al menos un elemento $\varphi \in \mathcal{N}_X$ que realiza la igualdad.

Si X e Y son múltiplos, se sigue la igualdad calculando el límite con $Y = \lambda X$ de forma inmediata (separando el caso $\lambda > 0$ y el caso $\lambda < 0$). Supongamos entonces que X e Y son linealmente independientes y sea $\psi : \text{gen}\{X, Y\} \rightarrow \mathbb{R}$ el funcional lineal dado por $\psi(aX + bY) = a|X| + bh'(0^+)$, para $a, b \in \mathbb{R}$. Notamos que $\psi(X) = |X|$ y $\psi(Y) = h'(0^+)$. Afirmamos que $\psi(aX + bY) \leq |aX + bY|$, lo cual es claro si $a = 0$ ó $b = 0$. Recordamos que por la desigualdad (3.12), h es una función decreciente. Supongamos primero que $a > 0$ y $b > 0$, luego

$$\begin{aligned} a|X| + bh'(0^+) &= a|X| + b \lim_{t \rightarrow 0^+} \frac{|X + tY| - |X|}{t} \leq a|X| + b \frac{|X + \frac{b}{a}Y| - |X|}{\frac{b}{a}} \\ &= a|X| + a|X + \frac{b}{a}Y| - a|X| = |aX + bY|. \end{aligned}$$

Si $a < 0$ y $b < 0$, volvemos a usar la desigualdad (3.12) dos veces

$$\begin{aligned} a|X| + bh'(0^+) &\leq a|X| + bh'(0^-) = a|X| + b \lim_{t \rightarrow 0^-} \frac{|X + tY| - |X|}{t} \leq a|X| + b \frac{|X - Y| - |X|}{-1} \\ &= a|X| + b|Y| - b|X - Y| = |bY - bX| - |-(a + b)X| \\ &\leq |bY - bX + (a + b)X| = |aX + bY|. \end{aligned}$$

Si $a < 0$ y $b > 0$ luego

$$\begin{aligned} a|X| + bh'(0^+) &= a|X| + b \lim_{t \rightarrow 0^+} \frac{|X + tY| - |X|}{t} \leq a|X| + b \frac{|X - \frac{b}{a}Y| - |X|}{-\frac{b}{a}} \\ &= 2a|X| + | -aX + bY| = | -aX + bY| - | -2aX| \\ &\leq | -aX + bY + 2aX| = |aX + bY|. \end{aligned}$$

Si $a > 0$ y $b < 0$ luego

$$\begin{aligned} a|X| + bh'(0^+) &\leq a|X| + b \frac{|X - \frac{b}{a}Y| - |X|}{-\frac{b}{a}} = 2a|X| - a|X - \frac{b}{a}Y| \\ &= |2aX| - |aX - bY| \leq |2aX - aX + bY| = |aX + bY|. \end{aligned}$$

Concluimos que $\psi(aX + bY) \leq |aX + bY|$ para todo $a, b \in \mathbb{R}$, luego $\|\psi\| = 1$. Por el Teorema de Hahn-Banach existe $\varphi \in E'$ que extiende a ψ con la misma norma; en particular $\varphi(X) = |X|$ y entonces $\varphi \in \mathcal{N}_X$. Además $\varphi(Y) = h'(0^+)$, lo cual finaliza la demostración. $\square$

Observación 3.2.5. Notamos que intercambiando Y por $-Y$ obtenemos que

$$\lim_{t \rightarrow 0^+} \frac{|X - tY| - |X|}{t} = \max_{\varphi \in \mathcal{N}_X} \varphi(-Y) = - \min_{\varphi \in \mathcal{N}_X} \varphi(Y),$$

de donde se sigue que

$$h_-(X, Y) = \lim_{t \rightarrow 0^+} \frac{|X - tY| - |X|}{-t} = \min_{\varphi \in \mathcal{N}_X} \varphi(Y)$$

Observación 3.2.6. Si la norma proviene de un producto interno y $h(t) = |X + tY| = \langle X + tY, X + tY \rangle^{1/2}$ entonces $h'(0) = \frac{\langle Y, X \rangle}{\langle X, X \rangle} = \varphi(Y)$. Donde φ es el único funcional que realiza la norma dado por el Teorema de Representación de Riesz.

Definición 3.2.7. Sea $(E, |\cdot|)$ un espacio con una norma de Finsler y sea $F(X) = |X|$, supongamos además que tenemos un producto interno $\langle \cdot, \cdot \rangle$ que hace de E un espacio métrico completo (notar que no estamos pidiendo que el producto interno induzca la norma de Finsler).

- (1) Si $0 \neq X \in E$, cada elemento $\varphi \in \mathcal{N}_X$, se puede representar como $\varphi(\cdot) = \langle N_\varphi, \cdot \rangle$. Al conjunto $\{N_\varphi : \varphi \in \mathcal{N}_X\}$ lo llamaremos subgradiente de F en X y lo notaremos como $\partial_X(F)$ o simplemente ∂_X cuando sea claro de que norma estamos hablando.
- (2) Sea $0 \neq X \in E$, si existe un único funcional que realiza la norma de X (en particular si la norma es estrictamente convexa, ver 2.1.14), tendremos que $\mathcal{N}_X = \{\varphi\}$. Por el Lema de Riesz podemos representar a este funcional como $\varphi(\cdot) = \langle N, \cdot \rangle$. A este elemento N lo llamaremos *gradiente de F en X* y lo notaremos $N = N_X = \text{grad}(F)(X)$. En este caso $\partial_X = \{\text{grad}(F)(X)\}$.

A continuación daremos algunos resultados de diferenciación, para ello necesitamos algunas definiciones.

Definición 3.2.8. Sea $(E, |\cdot|)$ un espacio vectorial con una norma de Finsler.

- (1) Decimos que $|\cdot| : E \rightarrow \mathbb{R}_{\geq 0}$ es diferenciable en el sentido de Gâteaux en un punto $X \in E$ no nulo, si para todo $Y \in E$ existe el límite $\lim_{t \rightarrow 0} \frac{|X+tY| - |X|}{t}$.
- (2) Decimos que $|\cdot| : E \rightarrow \mathbb{R}_{\geq 0}$ es diferenciable en el sentido de Fréchet en $X_0 \neq 0$ si existe un operador lineal $T : E \rightarrow \mathbb{R}$ de forma tal que $\lim_{X \rightarrow X_0} \frac{|X| - |X_0| - T(X)}{|X - X_0|} = 0$. En dicho caso T es único y lo llamamos la diferencial de la norma en x_0 .

Notamos que en general se pueden dar las mismas definiciones para operadores acotados entre espacios normados y que la diferenciación en el sentido de Fréchet es más fuerte que en el sentido de Gâteaux. Los siguientes resultados van a ilustrar que sucede en el caso específico de normas.

Observación 3.2.9. (Convexidad y suavidad).

- (1) La norma $|\cdot|$ es Gâteaux diferenciable en $x \neq 0$ si y solo si es Fréchet diferenciable (esto es una consecuencia del Lema de Šmulian, ver [29, Lema 8.4] por ejemplo). Por lo tanto simplemente vamos a decir que la norma es *suave* cuando esto sucede para cualquier $X \neq 0$; en tal caso la norma es más aún C^1 fuera de $X = 0$ [29, Corolario 8.5].
- (2) La norma es suave si y solo si la norma dual es estrictamente convexa, y esto sucede si y solo si hay un único funcional que realiza la norma de cada $X \neq 0$ [29, Lema 8.4 e ítem 8.12].
- (3) La norma es estrictamente convexa si y solo si es uniformemente convexa (esto se sigue de las propiedades previas y del hecho de que la norma es equivalente a la norma del doble dual o puede ser probado de forma directa utilizando la compacidad de la esfera).

3.3. El dual polar y sus propiedades

En esta sección exhibiremos varios ejemplos de normas de Finsler Ad-invariantes y sus normas duales. Además, en la Proposición 3.3.10, daremos ciertas propiedades publicadas en [47] para el caso de las normas orbitales.

Definición 3.3.1 (Dual polar). Sea K un conjunto convexo en un espacio vectorial real con producto interno $(E, \langle \cdot, \cdot \rangle)$. El *dual polar* de K es el conjunto

$$K^\circ = \{X \in E : \langle X, Y \rangle \leq 1 \quad \forall Y \in K\}.$$

Notamos que $K \subseteq (K^\circ)^\circ$, de hecho tenemos que $(K^\circ)^\circ$ es igual a la clausura de la cápsula convexa de $K \cup \{0\}$. Si K es cerrado y $0 \in K$, como estamos asumiendo que K es convexo, tenemos que $K = (K^\circ)^\circ$. Todos estos resultados tienen demostraciones elementales y las mismas se encuentran en el sexto capítulo del libro de Brøndsted [20].

Proposición 3.3.2. Sean $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, B_1 su bola unitaria, $\mathfrak{M}$ una MAB y $B = B_1 \cap \mathfrak{M} \subseteq \mathbb{R}^n$ (ver Proposición 2.3.4). Son equivalentes

- (1) La norma $|\cdot|$ es estrictamente convexa.
- (2) La norma inducida en $\mathbb{R}^n$ (ver Teorema 2.3.6) cuya bola unitaria es B es estrictamente convexa.
- (3) El conjunto ∂B (la frontera topológica de B) es estrictamente convexo.

Demostración. La equivalencia de (2) y (3) es un resultado conocido que se suele usar muchas veces como definición de norma estrictamente convexa. Lo que nos interesa es ver la equivalencia con (1). Esto va a ser una consecuencia del Teorema 2.3.6, que nos da la unicidad de la norma de Finsler simétrica inducida en $\mathbb{R}^n$ por una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Supongamos entonces (1) y veamos (2), tenemos entonces $x, y \in \mathbb{R}^n$ que cumplen $|x + y| = |x| + |y|$ y queremos ver que x e y son múltiplos. Para ello los pensamos dentro de $\mathfrak{M}$ como elementos X e Y y tenemos $|X + Y| = |X| + |Y|$ lo cual nos dice que son múltiplos y entonces x e y también lo son, esto prueba que (1) implica (2).

Por otra parte supongamos (2) y veamos (1), para ello, consideramos $X, Y \in \mathfrak{u}_n$ que verifican $|X + Y| = |X| + |Y|$ y veamos que X e Y son múltiplos. Vamos a notar $\lambda(X)$ a la tira de partes imaginarias de autovalores de X ordenada en forma decreciente, análogamente $\lambda(Y)$ y $\lambda(X + Y)$. Tenemos en tal caso que: $|X + Y| = |\lambda(X + Y)|$, $|X| = |\lambda(X)|$ y $|Y| = |\lambda(Y)|$, donde la norma en el caso de las tiras de partes imaginarias de autovalores es la norma inducida en $\mathbb{R}^n$. En tal caso lo que tenemos es que:

$$|\lambda(X + Y)| = |\lambda(X)| + |\lambda(Y)|.$$

Aplicando el Teorema de Weyl (ver [5, Teorema 5.1.6]) tenemos que $\lambda(X + Y) < \lambda(X) + \lambda(Y) = \lambda(W)$ para cierta $W \in \mathcal{U}_n$. Por la Proposición 2.2.16, obtenemos que $|\lambda(X + Y)| \leq |\lambda(X) + \lambda(Y)|$. Se sigue entonces la siguiente cadena de desigualdades:

$$|X + Y| = |\lambda(X + Y)| \leq |\lambda(X) + \lambda(Y)| \leq |\lambda(X)| + |\lambda(Y)| = |X| + |Y| = |X + Y|,$$

esto implica que las desigualdades son todas igualdades y entonces utilizando que la norma en $\mathbb{R}^n$ es estrictamente convexa tenemos que $\lambda(X) = s\lambda(Y)$ para cierto $s > 0$. Notamos que esto nos dice además que $\lambda(X + Y) = \lambda(X) + \lambda(Y)$, por ser la norma en $\mathfrak{u}_n$ estrictamente convexa.

Vamos a denotar $\lambda(X) = (\lambda_1(X), \lambda_2(X), \dots, \lambda_n(X))$ y análogamente para $\lambda(Y)$. Sea v_1 un autovector de $X + Y$ de norma igual a 1 y autovalor $\lambda_1(X + Y)$. Usando el producto interno de Frobenius tenemos que

$$((X + Y)v_1 | v_1) = \lambda_1(X + Y) = \lambda_1(X) + \lambda_1(Y),$$

de donde se sigue que

$$\lambda_1(X) + \lambda_1(Y) = ((X + Y)v_1 | v_1) = (Xv_1 | v_1) + (Yv_1 | v_1) \leq \lambda_1(X) + \lambda_1(Y)$$

Por el principio mín/máx (ver [4, Teorema 2.3.1] y [4, Teorema 2.3.3]) tenemos que $(Xv_1 | v_1) \leq \lambda_1(X)$ y $(Yv_1 | v_1) \leq \lambda_1(Y)$. Utilizando el caso de igualdad de dicho principio se sigue que $(Xv_1 | v_1) = \lambda_1(X)$ y $(Yv_1 | v_1) = \lambda_1(Y)$. De esta forma tenemos un autovector común de X e Y asociado al primer autovalor, si proseguimos así podemos diagonalizar X e Y de forma simultánea y llegamos a que están en la misma MAB y entonces como $\lambda(X) = s\lambda(Y)$ para cierto $s > 0$, se sigue que $X = sY$ para cierto $s > 0$ lo cual concluye la demostración. $\square$

Observación 3.3.3. Sea $X \in \mathfrak{u}_n$ con $X \neq 0$. Sabemos que la norma va a ser diferenciable en X si y solo si es Fréchet diferenciable en X (ver 3.2.9). Esto va a suceder si existe la derivada de la norma en X en cada dirección $Y \in E$ y por 3.2.5 ocurrirá si y solo si

$$\max_{\varphi \in N_X} \varphi(Y) = \min_{\varphi \in N_X} \varphi(Y).$$

Si queremos que valga esta igualdad para todo $Y \in \mathfrak{u}_n$ debería existir un único funcional que realice la norma de X , luego tenemos que la norma es suave (donde con esto nos referimos a diferenciable salvo en 0) si y solo si es estrictamente convexa (ver 2.1.14). Por otra parte esto va a suceder si y solo si la norma inducida en $\mathbb{R}$ es estrictamente convexa lo cual se deduce de la construcción realizada en el Teorema 2.3.6. Si volvemos a aplicar el razonamiento en $\mathbb{R}$ para la norma inducida, tenemos que va a ser estrictamente convexa si y solo si es suave y eso es si y solo si ∂B es suave.

En conclusión: para ver si una norma es suave en $\mathfrak{u}_n$, alcanza estudiar si la frontera de la bola unitaria de la norma inducida es suave (o equivalentemente estrictamente convexa) en $\mathbb{R}^n$.

Definición 3.3.4 (Normas autoduales). Decimos que una norma en $\mathfrak{u}_n$ es *autodual* si B_1 es homotéticamente equivalente a una rotación de B_1° .

Observación 3.3.5. Por las observaciones dadas en esta sección, alcanza con estudiar el dual polar de un conjunto convexo n -dimensional $B \subset \mathbb{R}^n$, obtenido a partir de cortar la bola unitaria $B_1 \subseteq \mathfrak{u}_n$ de la norma con una MAB arbitraria. Dado que son conjuntos convexos, se comportan bien con respecto a las operaciones de interior y de clausura, luego vamos a abusar un poco la notación y llamar indistintamente B, B_1 a las bolas y a sus correspondientes clausuras, entendiéndolo de cual se trata según el contexto.

Observación 3.3.6. Uno podría dar una definición más estricta y decir que una norma es autodual si $B_1^\circ = B_1$, pero es fácil ver que el único cuerpo convexo, cerrado y que coincide con su dual polar en $\mathbb{R}^n$ es la bola unitaria euclídea, lo cual hace que la definición no tenga interés.

Antes de comenzar con los ejemplos notamos que, por la biyección dada en el Teorema 2.3.6 y la Observación 3.3.5, una norma de Finsler Ad-invariante en u_n va a ser autodual si y solo si su norma de Finsler simétrica inducida en $\mathbb{R}^n$ es autodual. Por otra parte, los elementos de norma dual igual a 1 en u_n están identificados vía la norma de Frobenius con los funcionales de norma 1. Por otra parte la correspondencia con $\mathbb{R}^n$ mediante el corte de u_n con una MAB es un isomorfismo que preserva producto interno cuando miramos el producto interno canónico de $\mathbb{R}^n$.

Necesitamos para nuestros ejemplos la siguiente proposición:

Proposición 3.3.7. Sean $|\cdot|$ una norma de Finsler estrictamente convexa (lo cual vimos que es lo mismo que suave) en $\mathbb{R}^n$ y $F : \mathbb{R}^n \rightarrow \mathbb{R}$ la función dada por $F(X) = |X|$. Para $X \in \mathbb{R}^n$ con $|X| = 1$ consideramos el funcional

$$\varphi(V) = \langle \text{grad}(F)(X), V \rangle,$$

donde el producto interno es el usual de $\mathbb{R}^n$. Entonces $\|\varphi\| = 1$ (equivalentemente $|\text{grad}(f)(X)|' = 1$) y además la norma de φ se realiza en X .

Demostración. Para $t > 0$ tenemos que $F(tX) = |tX| = t|X|$. Derivando en $t = 1$ obtenemos $DF_X(X) = |X|$, como $DF_X(V) = \langle \text{grad}(F)(X), V \rangle$, se sigue que $\langle \text{grad}(F)(X), X \rangle = |X|$. Para ver que el funcional φ realiza la norma de X solo nos resta ver que $\|\varphi\| = 1$. Sea $V \in \mathbb{R}^n$ de norma 1, entonces

$$F(X + tV) - F(X) = |X + tV| - |X| \leq |X| + t|V| - |X| = t,$$

luego

$$\langle \text{grad}(F)(X), V \rangle = DF_X(V) = \lim_{t \rightarrow 0^+} \frac{F(X + tV) - F(X)}{t} \leq 1.$$

Se sigue que $\|\varphi\| = 1$, finalizando así la demostración. $\square$

Ejemplos 3.3.8. En los siguientes ejemplos estudiaremos las normas duales de distintas normas de Finsler Ad-invariantes. Daremos ejemplos de normas autoduales y normas que no lo son. Además, exhibiremos los funcionales que realizan la norma de ciertos elementos en cada caso. Por lo trabajado previamente, en vez de estudiar normas en u_n vamos a suponer que ya están cortadas con una MAB y las pensaremos en $\mathbb{R}^n$.

(1) Supongamos que $1/p + 1/q = 1$, incluyendo el caso de que $q = \infty$ cuando $p = 1$ y vice-versa. Recordamos que $(l_p)' = (l_q)$.

- Para $p = 2$ tenemos que la norma es autodual, esto es exactamente lo mismo que decir que la norma de Frobenius es autodual en u_n .
- Si $1 < p < \infty$ con $p \neq q$ tenemos que la bola de la norma p no es una dilatación conjunta con una rotación de la bola de norma q y por lo tanto no son autoduales. Sabemos que estas normas son estrictamente convexas y es fácil de probar que el único funcional que realiza la norma de $0 \neq x \in \mathbb{R}^n$ es el dado por $\varphi(z) = \langle \frac{1}{\|x\|_p^{p-1}} (sg(x_1)x_1^{p-1}, sg(x_2)x_2^{p-1}, \dots, sg(x_n)x_n^{p-1}), z \rangle$.
- Si $n = 2$ tenemos que la norma 1 y la norma ∞ son autoduales. De hecho sabemos que la bola de la norma 1 y la norma ∞ se pueden obtener una a partir de la otra mediante una rotación y como $(l_1)^* = l_\infty$ tendremos que la norma 1 es autodual. Vamos a pensar quienes son los funcionales que realizan la norma asociados a cada elemento de la cáscara de la bola unitaria en la norma

1. Evidentemente, obtendremos la norma infinito, pero esta forma de razonar nos ayudará a entender que pasa en casos no tan conocidos como en el ejemplo de la tostada y de la norma 1 en dimensiones mayores.

Comenzamos con un lado de la cáscara de la bola unitaria de la norma 1, vamos a tomar el dado por $L = \{(x, y) \in \mathbb{R}^2 : x + y = 1, x > 0, y > 0\}$ (con los otros la construcción es análoga). La norma en estos puntos es la función $F(x, y) = x + y$, siendo la misma una función suave, podemos aplicar la Proposición 3.3.7 y obtenemos $\text{grad}(F)(x, y) = (1, 1)$. Esto nos dice que el único funcional que realiza la norma de un elemento de L es, independientemente del elemento, el dado por $\varphi(x, y) = x + y$. Cuando lo pensamos como elemento de l_∞ obtenemos el vector $(1, 1)$, es decir todo el lado L nos da un solo punto de la cáscara de la bola unitaria de la norma dual. Lo mismo sucede con los otros lados, los cuales nos dan los vértices $(1, -1)$, $(-1, 1)$, $(-1, -1)$.

Ahora nos interesa pensar cuales son los funcionales que realizan la norma de $P = (0, 1)$ y tienen norma dual igual a 1. Sabemos que tienen que ser de la forma $\varphi(x, y) = ax + by$ y cumplir simultáneamente que $\langle (a, b), (0, 1) \rangle = 1$ y $\max\{|a|, |b|\} = 1$. En definitiva obtenemos los elementos de la forma $\varphi(x, y) = a + y$ con $-1 \leq a \leq 1$. Al pensarlos en l_∞ son los vectores de la forma $(a, 1)$ con $-1 \leq a \leq 1$.

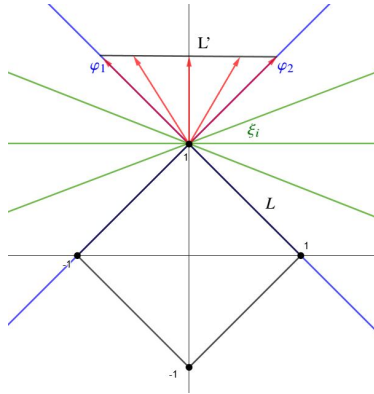


Figura 3.1: Norma 1.

En la Figura 3.1 podemos observar la bola unitaria de la norma $\|\cdot\|_1$. Cuando pensamos en lado L obtenemos el vector perpendicular de norma infinito igual a 1, es decir el $(1, 1)$ como mencionamos antes. Si nos centramos en el punto $P = (0, 1)$ de la misma figura, tenemos representados en verde todos los núcleos trasladados de los funcionales ξ_i que realizan la norma de dicho punto y en azul los funcionales que representan las caras maximales φ_1 y φ_2 dados por $\varphi_1(x, y) = x + y$ y $\varphi_2(x, y) = -x + y$. Para ver la bola dual, simplemente tenemos que pensar en los vectores perpendiculares a los núcleos trasladados de dichos funcionales y normalizarlos para que tengan norma infinito igual a 1, obtenemos así los vectores rojos marcados en el gráfico. Si juntamos todos sus puntos finales obtenemos L' que, al trasladarlo al origen, es la cara superior de la bola unitaria de la norma $\|\cdot\|_\infty$.

Si comenzamos con la norma infinito el procedimiento es análogo y obtenemos lo norma 1 como su norma dual.

- En el caso $n \geq 3$ las norma 1 e infinito se van a comportar de forma distinta, por su puesto que seguirán siendo una el dual de la otra, pero no van a ser autoduales. Para ejemplificar esto vamos a desarrollar el caso $n = 3$ para la norma infinito. Razonando de forma análoga al ítem previo en $\mathbb{R}^3$ podemos ver que cada vértice del cubo $\|(x, y, z)\|_\infty = 1$ nos va a dar una cara del octaedro $\|(x, y, z)\|_1 = 1$, cada cara de $\|(x, y, z)\|_\infty = 1$ nos va a dar un vértice de $\|(x, y, z)\|_1 = 1$ y cada lado de $\|(x, y, z)\|_\infty = 1$ nos dará un lado de $\|(x, y, z)\|_1 = 1$. Realizando este proceso, a partir del cubo inicial obtenemos el octaedro $|x| + |y| + |z| = 1$ el cual no se puede obtener mediante rotaciones y homotecias a partir de un cubo.

Este proceso que en $\mathbb{R}^3$ asocia mediante el dual a un vértice una cara, a un lado otro lado y a una cara un vértice se puede generalizar para entender los objetos de mayor dimensión. En este caso obtendremos que cada cara de dimension k del polígono del que partimos nos da una cara de dimensión $n - k$ del polígono de la norma dual.

- (2) Recordamos el Ejemplo 2.1.15 de la elipse rotada. La norma cuya bola unitaria es esta elipse está dada por

$$F(x, y) = \sqrt{\frac{1}{a^2}(x+y)^2 + \frac{1}{b^2}(x-y)^2}.$$

Como es estrictamente convexa, sabemos que el único funcional de norma 1 que realiza la norma de $p = (x, y)$ viene dado por

$$\varphi_p(v) = \langle \text{grad}(F)(p), v \rangle$$

luego para cada $(x, y) \in \mathbb{R}^2$ queremos entender quien es $\text{grad}(F)(x, y)$ y ver geométricamente que representa cuando recorremos todos los (x, y) con $F(x, y) = 1$.

Para ello vamos a considerar la parametrización de la esfera unitaria de esta norma dada por

$$\alpha(t) = \left(\frac{a \cos(t) + b \sin(t)}{2}, \frac{a \cos(t) - b \sin(t)}{2} \right)$$

la cual surge del cambio de variables

$$x + y = a \cos(t)$$

$$x - y = b \sin(t).$$

En tal caso tenemos que

$$\text{grad}(F)(\alpha(t)) = \left(\frac{1}{a} \cos(t) + \frac{1}{b} \sin(t), \frac{1}{a} \cos(t) - \frac{1}{b} \sin(t) \right).$$

Si consideramos la norma cuya esfera unitaria es la elipse rotada de radios $\frac{2}{a}$ y $\frac{2}{b}$, es inmediato ver que es la imagen de la curva $\text{grad}(F)(\alpha(t))$, en otras palabras la norma dual de F es

$$F'(x, y) = \sqrt{\frac{a^2}{4}(x+y)^2 + \frac{b^2}{4}(x-y)^2}.$$

Veamos cuando es autodual, en el caso $a = b$ es evidente, ya que ambas normas tienen como bola unitaria una circunferencia centrada en el origen.

Para el caso general, notamos que la elipse original y la dual tienen los mismos ejes. Luego para que la norma sea autodual debe ocurrir una de las dos siguientes posibilidades:

Posibilidad 1: no rotar y solo multiplicar por un escalar, para ello debe existir $k > 0$ de forma tal que $\alpha(t) = k\beta(t)$ donde $\beta(t) = \text{grad}(F)(\alpha(t))$ y $t \in [0, 2\pi)$. Haciendo esta cuenta se deduce que $a^2 = b^2 = 2k$ entonces $a = b$ y volvemos al caso antes mencionado.

Posibilidad 2: rotar la elipse en una magnitud de $\frac{\pi}{2}$, intercambiando los ejes de la elipse. En dicho caso debería existir $k > 0$ que cumpla $\alpha(t) = k\gamma(t)$ donde $\gamma(t) = \left(-\frac{1}{a} \cos(t) + \frac{1}{b} \sin(t), \frac{1}{a} \cos(t) + \frac{1}{b} \sin(t)\right)$ y nuevamente llegamos a la misma conclusión.

En resumen, la norma va a ser autodual si y solo si $a = b$.

- (3) Para el ejemplo de la tostada dado en 2.1.15 queremos realizar un procedimiento similar. Si llamamos $F(x, y)$ a la norma cuya bola unitaria es la tostada tenemos que

$$F'(x, y) = \|(x, y)\|_{\infty}$$

para (x, y) fuera del primer cuadrante, ya que se comporta igual que la norma 1 en dicho caso.

Notamos que a diferencia de usar la norma 1 completa, los puntos $(0, 1)$ y $(1, 0)$ tienen un único funcional que realiza la norma, los cuales están asociados al vector $(1, 1)$ en ambos casos. Finalmente lo que obtenemos al considerar los elementos con norma dual igual a 1 (sin incluir aún el primer cuadrante) son los segmentos que unen al $(-1, -1)$ con el $(1, -1)$ y con el $(-1, 1)$ incluyendo los bordes.

Por otro lado si consideramos (x, y) en el primer cuadrante

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

y en tal caso la curva $F(x, y) = 1$ está parametrizada por

$$\alpha(t) = \left(\frac{\sqrt{2}}{2} \cos(t) + \frac{1}{2}, \frac{\sqrt{2}}{2} \sin(t) + \frac{1}{2}\right) \quad \text{con} \quad -\frac{\pi}{4} < t < \frac{3\pi}{4}.$$

Luego $\text{grad}(F)(\alpha(t)) = \left(\sqrt{2} \cos(t), \sqrt{2} \sin(t)\right)$, parametrizando así el arco de la circunferencia de radio $\sqrt{2}$ centrada en el origen que une al punto $(1, -1)$ con el $(-1, 1)$. En otras palabras si (x, y) pertenece al primer cuadrante, tenemos que

$$F'(x, y) = \sqrt{\frac{x^2 + y^2}{2}}.$$

Finalmente obtenemos el borde de la norma dual de la tostada que es representado en la siguiente figura, la cual se encuentra superpuesta con la bola unitaria de la tostada (en azul):

- (4) La norma dual de las normas Ky-Fan k está dada por $\|X\|'_{(k)} = \max \left\{ \frac{1}{k} \|X\|_1, \|X\|_{\infty} \right\}$, como se puede ver en [66, Teorema 1]. Los funcionales que realizan la norma de un elemento dado para la norma

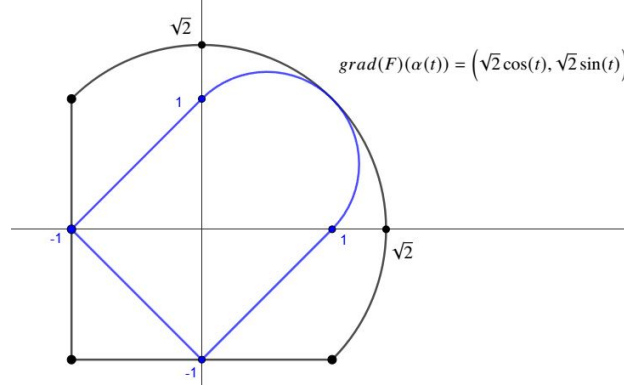


Figura 3.2: Bola unitaria de la tostada y de su norma dual.

Ky-Fan $\|\cdot\|_{(k)}$ fueron completamente caracterizados por Watson en [66, 67] usando el subgradiente de la norma definido en 3.2.7. Para agregar un ejemplo, podemos fijar un vector $a = (a_1, a_2, \dots, a_n)$ y notamos que el funcional dado por

$$\varphi(x) = \langle (sg(a_1), sg(a_2), \dots, sg(a_k), 0, \dots, 0), x \rangle$$

realiza la norma del elemento a , donde $sg(a_i)$ es el signo de a_i o cero si $a_i = 0$.

Como las normas orbitales juegan un papel privilegiado en nuestra construcción vamos a dar algunas de sus propiedades con detalle. Comenzamos con un lema:

Lema 3.3.9. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial real con producto interno. Sean $A, B \subseteq E$ entonces

- (1) $A \subseteq B \Rightarrow A^\circ \supseteq B^\circ$.
- (2) $A^\circ = (\text{conv } A)^\circ$.

Demostración. Sea $x \in B^\circ$ entonces $\langle x, y \rangle \leq 1$ para todo $y \in B$ y como $A \subseteq B$ entonces tenemos la desigualdad para todo $y \in A$ lo cual nos lleva a que $x \in A^\circ$ lo cual concluye (1).

Para ver (2) notamos que $A \subseteq \text{conv}(A)$, luego por (1) tenemos que $A^\circ \supseteq (\text{conv } A)^\circ$, con lo cual solo resta ver la otra contención. Sea $x \in A^\circ$, consideramos $y \in \text{conv } A$ y lo escribimos como $y = \sum_i t_i y_i$ con $y_i \in A$, $t_i \geq 0$ y $\sum_i t_i = 1$. En tal caso

$$\langle x, y \rangle = \langle x, \sum_i t_i y_i \rangle = \sum_i t_i \langle x, y_i \rangle \leq \sum_i t_i = 1,$$

donde usamos que $\langle x, y_i \rangle \leq 1$ para todo i por ser $x \in A^\circ$. Luego hemos probado que $\langle x, y \rangle \leq 1$ para todo $y \in \text{conv}(A)$ y entonces $x \in (\text{conv}(A))^\circ$. $\square$

Proposición 3.3.10 (Propiedades de las Normas Orbitales). Dada $C \in \mathfrak{su}_n$ no nula, consideramos la norma orbital dada por

$$|X|_{\mathcal{O}_C} = \max_{U \in \mathfrak{u}_n} (UCU^*|X)$$

que como ya vimos en la Proposición 2.2.21, resulta una norma de Finsler Ad-invariante en $\mathfrak{su}_n$. Notaremos además

- $B_1 = \{X \in \mathfrak{su}_n : |X|_{\mathcal{O}_C} \leq 1\}$ (bola unitaria de la norma).
- $B_1^\circ = \{X \in \mathfrak{su}_n : (X|Y) \leq 1 \text{ para todo } Y \in \mathfrak{su}_n \text{ con } |Y|_{\mathcal{O}_C} \leq 1\}$ (dual polar).
- $B'_1 = \{X \in \mathfrak{su}_n : |X|'_{\mathcal{O}_C} \leq 1\}$ (bola unitaria de la norma dual).

En tal caso valen las siguientes afirmaciones:

- (1) $B_1^\circ = B'_1 = \text{conv}(\mathcal{O}_C)$. Equivalentemente $B_1 = \text{conv}(\mathcal{O}_C)^\circ$, en particular, el máximo puede ser tomado sobre toda la órbita o sobre la cápsula convexa de la misma:

$$|X|_{\mathcal{O}_C} = \max_{V \in \mathcal{O}_C} (X|V) = \max_{Z \in \text{conv}(\mathcal{O}_C)} (X|Z) \quad \text{para todo } X \in \mathfrak{u}_n.$$

- (2) Si π el plano dado por

$$\pi = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \sum_i x_i = 0\},$$

y $\mathfrak{M}$ una MAB arbitraria, entonces el conjunto $B_1 \cap \mathfrak{M}$ visto dentro de π , está formado por la cápsula convexa del vector cuyas entradas son las partes imaginarias de los autovalores de C .

- (3) Si $\|C\|_F = 1$ entonces $|X|'_{\mathcal{O}_C} \geq |X|_{\mathcal{O}_C}$ para cualquier $X \in \mathfrak{su}_n$.
- (4) Si $\|C\|_F = 1$ tenemos que $\mathcal{O}_C = \partial B_1 \cap \partial B_1^\circ$.
- (5) Si $\|C\|_F = 1$ y $X \in \mathfrak{u}_n$ es distinto de cero, entonces existe un funcional de la forma $\varphi = (U_0 C U_0^* | \cdot)$ que realiza la norma de X .

Demostración. (1) La igualdad $B_1^\circ = B'_1$ se obtiene de forma inmediata mirando las definiciones. Veamos que $B_1^\circ = \text{conv}(\mathcal{O}_C)$, para ello notamos que si $X \in \mathcal{O}_C$ entonces $X = V C V^*$ para cierto $V \in \mathcal{U}_n$, tomando $Y \in \mathfrak{u}_n$ con $|Y|_{\mathcal{O}_C} \leq 1$ se sigue que $\max_{U \in \mathcal{U}_n} (U C U^* | Y) \leq 1$ y luego $(V C V^* | Y) \leq 1$ concluyéndose que $X \in B_1^\circ$. Por la misma definición de B_1° , se sigue que es cerrado por combinaciones convexas y luego tenemos la inclusión $\text{conv}(\mathcal{O}_C) \subseteq B_1^\circ$.

Por otro lado supongamos que $X \in B_1^\circ$, en tal caso $(Y|X) \leq 1$ para todo $|Y|_{\mathcal{O}_C} = 1$, reemplazando Y por $\frac{Y}{|Y|_{\mathcal{O}_C}}$ y multiplicando ambos lados de la desigualdad por $|Y|_{\mathcal{O}_C}$ obtenemos que

$$(Y|X) \leq |Y|_{\mathcal{O}_C} = \max_{U \in \mathcal{U}_n} (U Y U^* | C) \quad \text{para todo } Y \in \mathfrak{su}_n,$$

entonces

$$\max_{U \in \mathcal{U}_n} (U Y U^* | X) \leq \max_{U \in \mathcal{U}_n} (U Y U^* | C),$$

como además X y C tienen traza cero, utilizando (4) de la Proposición 2.2.16, obtenemos que $X \in \text{conv}(\mathcal{O}_C)$ que es lo que queríamos probar.

Ahora que sabemos que $B_1^\circ = B'_1 = \text{conv}(\mathcal{O}_C)$, tomando nuevamente dual polar obtenemos que $(B_1^\circ)^\circ = (\text{conv}(\mathcal{O}_C))^\circ$ y como B_1 es un cerrado convexo que contiene al cero, se deduce que $B_1 = (\text{conv}(\mathcal{O}_C))^\circ$.

Para ver la igualdad del enunciado simplemente tendremos que ver que $\max_{Z \in \text{conv}(\mathcal{O}_C)} (X|Z) \leq \max_{V \in \mathcal{O}_C} (X|V)$.

Para ello si $Z \in \text{conv}(\mathcal{O}_C)$ entonces $Z = \sum_i t_i V_i$ con $V_i \in \mathcal{O}_C$, $t_i \geq 0$ para todo i y $\sum_i t_i = 1$. Entonces $(X|Z) = \sum_i t_i (X|V_i) \leq (X|V_k)$ para cierto k que maximiza el producto $(X|V_i)$, como $V_k \in \mathcal{O}_C$, se sigue la desigualdad.

- (2) Primero consideramos una MAB arbitraria. Al cortar el conjunto con dicha MAB y mirarlo dentro de $\mathbb{R}^n$ se sigue que el mismo no depende de la MAB elegida, esto es consecuencia del Teorema 2.3.8. Dicho esto, simplemente podemos elegir $\mathfrak{M}$ de forma que C sea diagonal en la base asociada y usando la Observación 2.3.5 tenemos el resultado enunciado.
- (3) Con la normalización $\|C\|_F = 1$ tenemos que $\text{conv}(\mathcal{O}_C) \subset B_1$ (Proposición 2.2.21), en tal caso $B'_1 \subset B_1$ luego $|X|'_{\mathcal{O}_C} \geq |X|_{\mathcal{O}_C}$ para cualquier $X \in \mathfrak{u}_n$.
- (4) Para ver este resultado usaremos repetidamente la Proposición 2.2.21.3. Notamos que por dicha proposición $|C|_{\mathcal{O}_C} = 1$ ya que $\|C\|_F = 1$. Veamos ahora la doble inclusión: si $|X|_{\mathcal{O}_C} = |X|'_{\mathcal{O}_C} = 1$, como $B'_1 = \text{conv}(\mathcal{O}_C)$ y $|X|'_{\mathcal{O}_C} = 1$ entonces $X \in \text{conv}(\mathcal{O}_C)$ y por la misma proposición $X \in \mathcal{O}_C$. Por otro lado si $X \in \mathcal{O}_C$ entonces $|X|_{\mathcal{O}_C} = 1$ (también por la proposición citada) y como $B_1 = \text{conv}(\mathcal{O}_C)$ tenemos que $|X|_{\mathcal{O}_C} \leq 1$, recurriendo al ítem previo de la proposición que estamos demostrando se sigue que debe ser mayor igual a 1 y por lo tanto igual a 1.
- (5) Sabemos que $|X|_{\mathcal{O}_C} = (U_0 C U_0^* | X)$ para cierto U_0 , lo único que tenemos que observar es que el funcional mencionado tiene norma 1 y esto se sigue del hecho de que $|U_0 C U_0^*|_{\mathcal{O}_C} = |C|_{\mathcal{O}_C} = 1$.

□

En lo que queda de la sección nos dedicaremos a estudiar las normas duales de las normas orbitales en $\mathfrak{su}_n$. En general tendremos que estas normas no van a ser autoduales, salvo casos específicos y en dimensiones chicas. Para ello estudiaremos primero estos casos para los cuales nos valdremos del ítem (2) de la proposición previa y el plano π será el de dicha proposición. Por otro lado, utilizaremos este enunciado para caracterizar B_1° y el primer ítem de la proposición previa para obtener $B_1 = (\text{conv}(\mathcal{O}_C))^\circ$.

Un resultado geométrico que vamos a utilizar es el hecho de que los polígonos regulares en $\mathbb{R}^2$ que tienen el origen en su interior son autoduales. Esto no es válido en mayores dimensiones, de hecho ya vimos que el dual del cubo en $\mathbb{R}^3$ dado por $\|(x, y, z)\|_\infty = 1$ es el octaedro $\|(x, y, z)\|_1 = 1$. Algunas referencias clásicas para el estudio de estos temas son los libros de Coxeter [25] y Brøndsted [20].

Proposición 3.3.11. Sea $C \in \mathfrak{su}_2$ no nula, la norma orbital $|\cdot|_C$ es autodual.

Demostración. Como $C \in \mathfrak{su}_n$ entonces la tira de partes imaginarias de autovalores de C va a ser de la forma $v = (a, -a)$ para cierto $a \in \mathbb{R}$ no nulo. Pensando este elemento dentro de $\pi \subseteq \mathbb{R}^2$ y si considerando sus permutaciones, obtendremos el conjunto de dos elementos dado por $\{(a, -a), (-a, a)\}$. Al mirar su cápsula convexa tendremos el segmento cerrado que une a $v = (a, -a)$ con $w = (-a, a)$, es decir $B_1^\circ = \text{conv}(\mathcal{O}_C) = \{(-t, t) : |t| \leq |a|\}$. Por el Lema 3.3.9, nos bastará mirar el dual polar de la órbita de C para obtener B_1 . Sea $(x, -x) \in \pi$, tenemos que $(x, -x) \in (\text{conv}(\mathcal{O}_C))^\circ$ si y solo si $\langle (x, -x), (a, -a) \rangle \leq 1$ y $\langle (x, -x), (-a, a) \rangle \leq 1$. Utilizando las dos ecuaciones previas obtenemos que esto sucede si y solo si

$$-\frac{1}{2|a|} \leq x \leq \frac{1}{2|a|},$$

de donde se sigue de forma inmediata que $B_1 = \{(-t, t) : |t| \leq \frac{1}{2|a|}\}$. En tal caso tenemos que $2|a|^2 B_1 = B_1^\circ$ y la norma resulta autodual. $\square$

Proposición 3.3.12. Si $C \in \mathfrak{su}_3$ es no nula y tiene un autovalor repetido o bien 0 es autovalor, entonces la norma orbital $|\cdot|_C$ es autodual.

Demostración. Consideramos primero el caso en el cual C tiene un autovalor repetido. Considerando nuevamente la tira de partes imaginarias de autovalores de C , obtendremos un vector de la forma $v = (2a, -a, -a)$ para cierto $a \in \mathbb{R}$ no nulo. Al mirar las permutaciones de este vector obtenemos tres vectores, los dados por: $v = (2a, -a, -a)$, $w = (-a, 2a, -a)$ y $z = (-a, -a, 2a)$, luego la cápsula convexa va a ser el triángulo sobre π formado por dichos vectores. Como la distancia entre cualquier par de dichos elementos es $3\sqrt{2}|a|$, se trata de un triángulo equilátero. Concluimos entonces que la bola dual unitaria es el triángulo equilátero con dichos vértices, el cual claramente está centrado en el origen.

Es un resultado conocido que el dual de un triángulo equilátero centrado en el origen es a su vez otro triángulo equilátero centrado en el origen, luego, tenemos que B_1 y B_1° difieren en una rotación y una homotecia, por lo tanto la norma es autodual.

El caso que nos resta estudiar es si la tira de las partes imaginarias de los autovalores de C es de la forma $(a, 0, -a)$ con $a \neq 0$. En dicho caso tenemos $3! = 6$ permutaciones y la figura que forman es un hexágono regular. Nuevamente, es sabido que el dual polar de un hexágono regular vuelve a ser un hexágono regular y se concluye, de igual forma que antes, que la norma es autodual. $\square$

En general la norma orbital no va a ser autodual y esto está íntimamente relacionado con la simetría de la tira de autovalores de la matriz C con la que comencemos. Para ilustrar esto daremos un ejemplo en $\mathfrak{su}_3$ en el cual la norma no es autodual:

Ejemplo 3.3.13. Comenzamos con la matriz $C = \text{diag}(3i, -2i, -i)$ y al igual que antes pensamos en el vector que induce en $\pi \subseteq \mathbb{R}^3$ dado por $v = (3, -2, -1)$. Al mirar las $3! = 6$ obtenemos un hexágono sobre el plano π . Los vértices de este hexágono son

$$\begin{aligned} v_1 &= (3, -2, -1) & v_2 &= (3, -1, -2) & v_3 &= (-1, 3, -2) \\ v_4 &= (-1, -2, 3) & v_5 &= (-2, -1, 3) & v_6 &= (-2, 3, -1). \end{aligned}$$

La bola dual va a ser la cápsula convexa del conjunto generado por estos vectores y estará contenida en el plano π . Dado que estamos trabajando con una figura plana, para comprender mejor lo que está sucediendo vamos a rotar el plano π en el plano $z = 0$, de esta forma podremos pensarlo de la misma manera en la cual trabajamos en $\mathbb{R}^2$, para ello aplicaremos la matriz

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

a cada permutación de v , obteniendo de esta forma seis vectores con su última coordenada nula. Al

proyectarlos sobre el plano $z = 0$ obtenemos los vectores (que nombraremos igual para aliviar la notación)

$$\begin{aligned} v_1 &= \left(\frac{5}{\sqrt{2}}, \frac{3}{\sqrt{6}} \right) & v_2 &= \left(\frac{4}{\sqrt{2}}, \frac{6}{\sqrt{6}} \right) & v_3 &= \left(\frac{-4}{\sqrt{2}}, \frac{6}{\sqrt{6}} \right) \\ v_4 &= \left(\frac{-5}{\sqrt{2}}, \frac{3}{\sqrt{6}} \right) & v_5 &= \left(\frac{-1}{\sqrt{2}}, \frac{-9}{\sqrt{6}} \right) & v_6 &= \left(\frac{1}{\sqrt{2}}, \frac{-9}{\sqrt{6}} \right). \end{aligned}$$

Podemos ver el hexágono no regular que forman en la siguiente figura:

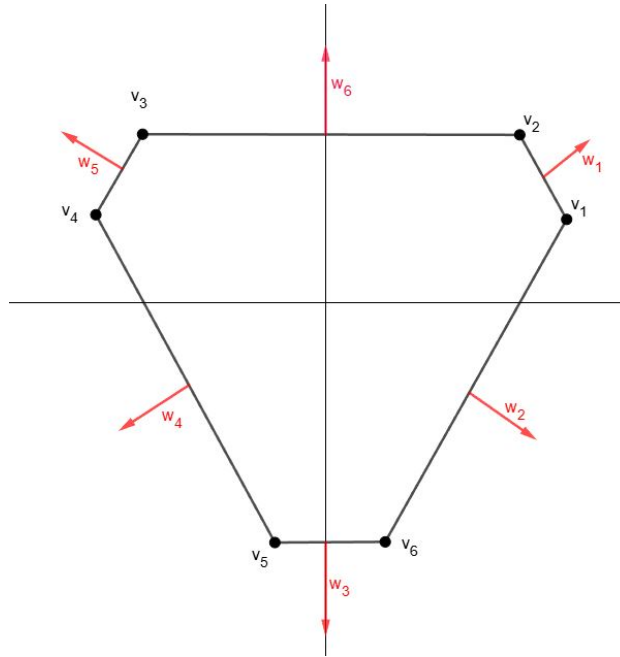


Figura 3.3: Borde de la bola dual.

Este hexágono es el borde de la bola unitaria de la norma dual. Ahora queremos calcular la bola de la norma original, para esto usaremos que cada lado nos induce un vértice del nuevo hexágono al mirar vectores normales a dichos lados, haciendo esto y normalizando (dado que nos importa si coinciden salvo rotaciones y homotecias) obtendremos los vértices de la bola original. Notar que calcular la dirección es sencillo y obtenemos los vectores de norma euclídea igual a 1 dados por

$$\begin{aligned} w_1 &= \left(\frac{3}{2\sqrt{3}}, \frac{1}{2} \right) & w_2 &= \left(\frac{\sqrt{3}}{2}, -\frac{1}{2} \right) & w_3 &= (0, -1) \\ w_4 &= \left(\frac{-3}{2\sqrt{3}}, -\frac{1}{2} \right) & w_5 &= \left(\frac{-\sqrt{3}}{2}, \frac{1}{2} \right) & w_6 &= (0, 1). \end{aligned}$$

Ahora necesitamos pensarlos en el plano $z = 0$, aplicarles $U^{-1} = U^t$ para verlos en el plano π , re-normalizar para que tengan todos la misma norma orbital y luego volver a llevarlos a $\mathbb{R}^2$ aplicando U . Comenzamos por llevarlos al plano π , obtenemos así ciertos vectores que nombraremos igual (abusando nuevamente de la notación ya que son los mismos pensados en el plano π). Estos vectores son:

$$\begin{aligned}
w_1 &= \frac{1}{\sqrt{6}}(2, -1, -1) & w_2 &= \frac{1}{\sqrt{6}}(1, -2, 1) & w_3 &= \frac{1}{\sqrt{6}}(-1, -1, 2) \\
w_4 &= \frac{1}{\sqrt{6}}(-2, 1, 1) & w_5 &= \frac{1}{\sqrt{6}}(-1, 2, -1) & w_6 &= \frac{1}{\sqrt{6}}(1, 1, -2).
\end{aligned}$$

Observamos que el hexágono que forman no puede ser obtenido a partir de mirar las permutaciones de uno de sus vértices, esto ya nos dice que la norma no va a ser autodual, igual prosigamos para verlo gráficamente.

Necesitamos calcular su norma orbital, es decir la norma orbital de las matrices que inducen y para ello usaremos la Proposición 2.2.21, obteniendo que tenemos que multiplicar la tira ordenada $(3, -1, -2)$ con las tiras ordenadas de cada vector y así obtenemos que

$$|w_1|_C = |w_3|_C = |w_5|_C = 9\sqrt{6},$$

$$|w_2|_C = |w_4|_C = |w_6|_C = 6\sqrt{6},$$

con lo cual si multiplicamos a w_2, w_4 y w_6 por el factor $\frac{3}{2}$ tendrán todos la misma norma orbital. Luego, como multiplicar por la matriz U y proyectar al plano $z = 0$ conmuta con multiplicar por el factor $\frac{3}{2}$, volvemos a mirar a los w_i en $\mathbb{R}^2$ y multiplicamos por dicho factor a los que necesitamos obteniendo así los vértices de la bola de la norma orbital. Estos son:

$$\begin{aligned}
z_1 &= \left(\frac{3}{2\sqrt{3}}, \frac{1}{2}\right) & z_2 &= \left(\frac{3\sqrt{3}}{4}, -\frac{3}{4}\right) & z_3 &= (0, -1) \\
z_4 &= \left(\frac{-9}{4\sqrt{3}}, -\frac{3}{4}\right) & z_5 &= \left(\frac{-3}{2\sqrt{3}}, \frac{1}{2}\right) & z_6 &= \left(0, \frac{3}{2}\right).
\end{aligned}$$

Gráficamente podemos observar la bola de la norma orbital de radio $9\sqrt{6}$:

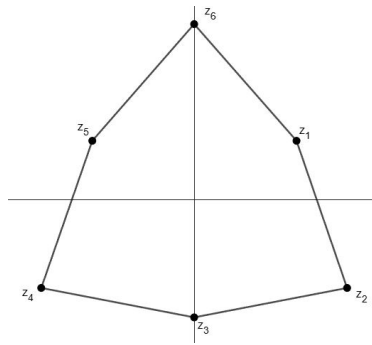


Figura 3.4: Borde de la bola original.

Como podemos ver gráfica o analíticamente, la bola original tiene distintos ángulos que la bola dual, esto nos vuelve a mostrar que la norma no es autodual. Otro invariante que podemos tomar es el máximo del cociente de lados opuestos, el cual es mayor en la bola dual que en la original.

Para mayores dimensiones el tema es mucho más complicado, dependerá de cada caso como mostramos previamente en §13. El hecho de que los polítopos regulares no tienen porque ser autoduales en dimensiones mayores a dos hace más difícil encontrar ejemplos donde la norma sea autodual, incluso si partimos de una matriz con espectro simple y balanceado.

3.4. Operadores disipativos

En esta sección estudiaremos, mediante la teoría de operadores disipativos, la relación entre $|V|$ con $|V + t[X, [X, V]]|$ para $X, V \in \mathfrak{u}_n$ y $t \in \mathbb{R}$. A pesar de que los resultados son un poco técnicos, la motivación de los mismos va a ser clara en el Capítulo 4 de este trabajo, más específicamente en las secciones 4.4 y 4.5, donde se ve la importancia de esta relación al estudiar la distancia entre e^{tV} y e^{tX} . En esta situación podemos pensar a la expresión $V + t[X, [X, V]]$ como el término de segundo orden del desarrollo en serie de la distancia antes mencionada. El Teorema 3.4.4 y sus consecuencias son resultados originales publicados en [47].

Recordamos que si $(E, \|\cdot\|)$ es un espacio de Banach, un operador $T : E \rightarrow E$ se dice *disipativo* si, para cada $\xi \in E$ y cada funcional $\varphi \in E'$ que realiza su norma, vale la desigualdad $\operatorname{Re}(\varphi(T\xi)) \leq 0$. Del Teorema de Hille-Yosida y la teoría de operadores disipativos (ver [58, Capítulo 1] y [42]) se sigue el siguiente teorema:

Teorema 3.4.1. Sean $(E, \|\cdot\|)$ un espacio de Banach y $T : E \rightarrow E$ un operador lineal continuo, son equivalentes:

- (1) Para cada $\xi \in E$ existe un funcional $\varphi \in E'$ que realiza su norma y que verifica $\operatorname{Re}(\varphi(T\xi)) \leq 0$.
- (2) $\|e^{sT}\| \leq 1$ para todo $s \geq 0$.
- (3) $1 - sT$ es inversible y expansivo para cada $s \geq 0$, i.e. $\|\xi - sT\xi\| \geq \|\xi\|$.

Observación 3.4.2 (Complexificación y la Norma de Taylor). Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, definimos para $A, B \in \mathfrak{u}_n$

$$|A + iB|_T = \sup_{t \in [0, 2\pi]} |A \cos t - B \sin t|$$

como la *Norma de Taylor* de $A + iB$ en la complexificación $M_n(\mathbb{C}) = \mathfrak{u}_n \oplus i\mathfrak{u}_n$ de $\mathfrak{u}_n$ asociada a la norma $|\cdot|$. Esta es una norma de Finsler en $M_n(\mathbb{C})$ que extiende a la norma dada en $\mathfrak{u}_n$. Es fácil de ver que $|A + iB|_T = |A - iB|_T$ y que $|U(A + iB)U^*|_T = |A + iB|_T$ para cualquier $U \in \mathcal{U}_n$, en tal caso resulta una norma débilmente invariante. Otros ejemplos de complexificaciones pueden verse por ejemplo en [56].

Teorema 3.4.3. Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Para cada $X \in \mathfrak{u}_n$, los operadores ad_X y ad_X^2 son disipativos con respecto a dicha norma, en particular $1 + s\operatorname{ad}_X$ es inversible y expansivo para todo $s \in \mathbb{R}$ y $1 - s\operatorname{ad}_X^2$ es inversible y expansivo para $s \geq 0$.

Demostración. Comenzamos pensando el operador lineal ad_X de forma que sea $\mathbb{C}$ lineal en $M_n(\mathbb{C})$ de la única forma posible, es decir $\operatorname{ad}_X(A + iB) = \operatorname{ad}_X(A) + i\operatorname{ad}_X(B)$. En tal caso se sigue que

$$|e^{s\operatorname{ad}_X}(A + iB)|_T = |\operatorname{Ad}_{e^{sX}}(A + iB)|_T = |e^{sX}(A + iB)e^{-sX}|_T = |A + iB|_T.$$

Esto nos dice que $\|e^{s \operatorname{ad}_X}\| \leq 1$ para todo $s \in \mathbb{R}$, y por el Teorema 3.4.1 se sigue que ad_X resulta disipativo en $M_n(\mathbb{C})$ con la complexificación de la norma. Dado que la restricción a $\mathfrak{u}_n$ de la norma de Taylor es la norma Ad-invariante con la que comenzamos, se sigue que la norma como operador de $e^{s \operatorname{ad}_X}$ también es menor o igual a 1 visto en $\mathfrak{u}_n$ y por el Teorema 3.4.1 concluimos que $1 - s \operatorname{ad}_X$ resulta también disipativo en $(\mathfrak{u}_n, |\cdot|)$. Razonando de la misma manera se sigue que $s \operatorname{ad}_X$ es disipativo para todo $s \in \mathbb{R}$, de donde concluimos utilizando nuevamente el Teorema 3.4.1 que $1 + s \operatorname{ad}_X$ es inversible y expansivo para todo $s \in \mathbb{R}$.

Por otro lado podemos escribir $1 - s^2 \operatorname{ad}_X^2 = (1 - s \operatorname{ad}_X)(1 + s \operatorname{ad}_X)$, como ambos factores son expansivos e inversibles, también lo es el producto y entonces ad_X^2 es disipativo. Se sigue del Teorema 3.4.1 que $1 - s \operatorname{ad}_X^2$ es inversible y expansivo para $s \geq 0$. $\square$

Teorema 3.4.4. Sean $X, V \in \mathfrak{u}_n$ y $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Si $0 \leq s \leq s'$ tenemos que

$$|V| \leq |V - s[X, [X, V]]| \leq |V - s'[X, [X, V]]|,$$

por otro lado si $t'/t \geq 1$ entonces

$$|V| \leq |V + t[X, V]| \leq |V + t'[X, V]|.$$

Demostración. Sea $s \geq 0$, por el teorema previo $1 - s \operatorname{ad}_X^2$ es expansivo e inversible, se sigue entonces que $\|(1 - s \operatorname{ad}_X^2)^{-1}\| \leq 1$. Sean $\xi = A + iB \in (\mathfrak{u}_n \oplus i\mathfrak{u}_n, |\cdot|_T)$ y φ un funcional que realiza la norma de ξ , luego

$$\operatorname{Re} \left(\varphi \left(((1 - s \operatorname{ad}_X^2)^{-1} - 1)\xi \right) \right) = \operatorname{Re} \left(\varphi \left((1 - s \operatorname{ad}_X^2)^{-1}\xi \right) \right) - |\xi|_T \leq |\xi|_T - |\xi|_T = 0,$$

luego por el Teorema 3.4.1 $(1 - s \operatorname{ad}_X^2)^{-1} - 1 = s \operatorname{ad}_X^2(1 - s \operatorname{ad}_X^2)^{-1}$ es disipativo y como todo múltiplo positivo de un disipativo también lo es, resulta que $\operatorname{ad}_X^2(1 - s \operatorname{ad}_X^2)^{-1}$ es disipativo. Luego si $0 \leq s \leq s'$, el operador

$$A = (1 - s' \operatorname{ad}_X^2)(1 - s \operatorname{ad}_X^2)^{-1} = 1 - (s' - s) \operatorname{ad}_X^2(1 - s \operatorname{ad}_X^2)^{-1}$$

es expansivo e inversible y luego su inversa es una contracción, por lo tanto

$$\begin{aligned} |(1 - s \operatorname{ad}_X^2)V| &= |(1 - s \operatorname{ad}_X^2)(1 - s' \operatorname{ad}_X^2)^{-1}(1 - s' \operatorname{ad}_X^2)V| \\ &= |A^{-1}(1 - s' \operatorname{ad}_X^2)V| \leq |(1 - s' \operatorname{ad}_X^2)V|, \end{aligned}$$

lo cual prueba la primera afirmación. La segunda afirmación se prueba de forma análoga a la primera. $\square$

Observación 3.4.5 (Mayorización). Usando la Proposición 2.2.16 podemos reformular el teorema previo en términos de mayorización. Si $\lambda(A) \subset \mathbb{R}^n$ denota la tira de las partes imaginarias de los autovalores de $A \in \mathfrak{u}_n$, para $V, X \in \mathfrak{u}_n$ y $s'/s \geq 1$ tenemos que

$$\lambda(V) < \lambda(V + s[X, V]) < \lambda(V + s'[X, V])$$

y tenemos un resultado análogo para $-[X, [X, V]]$.

Observación 3.4.6 (Ortogonalidad de Birkhoff). El resultado previo se puede interpretar en el siguiente sentido: para todo $X, V \in \mathfrak{u}_n$ y para toda norma de Finsler Ad-invariante en $\mathfrak{u}_n$, el vector V es ortogonal en el sentido de Birkhoff [16] al subespacio S generado por $[X, V]$ i.e.

$$|V| = \inf_{s \in \mathbb{R}} |V - s[X, V]| = \operatorname{dist}(V, S).$$

3.5. El caso de la igualdad

En esta sección, utilizando el criterio de la igualdad desarrollado en el Teorema 3.1.13, vamos a estudiar el caso $|V| = |V + [X, V]|$ dando equivalencias que involucran funcionales que realizan la norma y nos dan una noción más geométrica de la igualdad en términos de caras de la esfera. El siguiente teorema es original y está publicado en [47].

Teorema 3.5.1. Sean $X, V \in \mathfrak{u}_n$ y $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Son equivalentes:

- (1) $|V + [X, V]| = |V|$
- (2) Existe un funcional $\psi = (N|\cdot)$ que realiza la norma de la matriz $V + [X, V]$ y cumple que $[N, X_C] = 0$.
- (3) V y $V + [X, V]$ pertenecen a la misma cara de la esfera.
- (4) $V + [X, V]$ pertenece a todas las caras de la esfera a las cuales pertenece V .

Demostración. Supongamos (1) entonces por el Teorema 3.4.4, entonces para $0 < h < 1$ tenemos que

$$|V| = |V + [X, V]| \geq |V + (1 - h)[X, V]| \geq |V|.$$

Luego la función $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(s) = |V + s[X, V]|$ es constante en $[0, 1]$ y en tal caso calculando su derivada lateral en 1 por izquierda y aplicando la Observación 3.2.5 tenemos que:

$$0 = f'(1^-) = \lim_{h \rightarrow 0^+} \frac{|V + (1 - h)[X, V]| - |V + [X, V]|}{-h} = \min_{\psi \in \mathcal{N}_{V+[X, V]}} ([X, V]).$$

En tal caso existe $\psi = (N|\cdot)$ que realiza la norma de $V + [X, V]$ y que verifica $\psi([X, V]) = 0$. Por otro lado

$$|V| = |V + [X, V]| = (N|V + [X, V]) = (N|V) + (N|[X, V]) = (N|V) = \psi(V),$$

se deduce entonces que ψ también realiza la norma de V . Por otro lado, como ψ realiza la norma de $V + [X, V]$, tenemos que, por el Lema 3.1.4 $\psi \circ \text{ad}_V + \psi \circ \text{ad}_{[X, V]} = 0$ y entonces

$$\psi([X, [X, V]]) = -\psi \circ \text{ad}_{[X, V]}(X) = \psi \circ \text{ad}_V(X) = \psi([V, X]) = 0,$$

y por el Teorema 3.1.13 debe ser $[N, X_C] = 0$, lo cual demuestra (2). Si la norma es suave en V , entonces debe ser $F_\psi(V) = \{V\}$ y en tal caso por el Teorema 3.1.13 debe ser $[N, X] = 0$.

Supongamos que vale (2), entonces

$$\begin{aligned} |V| &\geq \psi(V) = (N|V) = (N|V) + ([N, X_C]|V) = (N|V) + (N|[X_C, V]) \\ &= (N|V + [X, V]) = \psi(V + [X, V]) = |V + [X, V]| \geq |V|, \end{aligned}$$

luego se deduce (1) de forma inmediata y también (3) ya que ψ realiza la norma simultáneamente de V y de $V + [X, V]$, lo cual nos dice que ambos elementos pertenecen a la misma cara de la esfera.

Si vale (3) utilizando un funcional que realiza la norma de V y $V + [X, V]$ y razonando igual que partiendo de (2), es inmediato de que $|V + [X, V]| = |V|$, luego si $V \in F_\varphi$ tenemos que $\varphi([X, V]) = 0$ y entonces

$$|V| = \varphi(V) = \varphi(V + [X, V]) \leq |V + [X, V]| = |V|.$$

Luego $V + [X, V] \in F_\varphi$, de donde se sigue (4). Si vale (4), nuevamente tomando un funcional que realice la norma de ambos elementos simultáneamente y aplicando la misma cadena de desigualdades que en (2) se sigue (1). $\square$

Observación 3.5.2. Las afirmaciones del teorema anterior siguen siendo validas si se reemplaza $[X, V]$ con $-[X, [X, V]]$, mediante una demostración análoga. Si la norma es suave en V , en la condición (2) podemos reemplazar $[N, X_C] = 0$ por $[N, X] = 0$. Si V es tal que $F_\varphi = \{V\}$ entonces la igualdad ocurre si y solo si $[X, V] = 0$.

Observación 3.5.3. (Interpretación geométrica) Notamos que el Teorema y la Observación previa nos dan una equivalencia en términos de funcionales de cuando no nos salimos de la misma cara de la esfera cuando nos paramos en V y nos movemos con el conmutador $[X, V]$ ó $-[X, [X, V]]$.

Corolario 3.5.4. Sean $X, V \in \mathfrak{u}_n$, sea $\lambda \in \mathbb{R}$. Luego

- (1) Si los autovalores de V son todos de la forma $i\lambda$ para cierto $0 \neq \lambda \in \mathbb{R}$, luego $\|V + [X, V]\|_\infty > \|V\|_\infty$ salvo que $[X, V] = 0$.
- (2) Si $V = i\lambda P$ con P un proyector unidimensional, luego $\|V + [X, V]\|_1 > \|V\|_1$ salvo que $[X, V] = 0$.
- (3) Sea $1 < k < n$ y V como en cualquiera de ambos ítems previos. Luego $\|X + [X, V]\|_{(k)} > \|V\|_{(k)}$ salvo que $[X, V] = 0$.

Podemos dar enunciados similares reemplazando $[X, V]$ por $-[X, [X, V]]$.

Demostración. En el primer caso estamos describiendo las matrices que, cuando pensamos en su tira de parte imaginaria de autovalores en $\mathbb{R}^n$, obtenemos los vértices de la bola de radio $|\lambda|$ para la $\|\cdot\|_\infty$, si $n = 2$ serían los vértices de un cuadrado, en $n = 3$ de un cubo, etc. En uno de estos puntos (y sólo en estos) existe φ un funcional que realiza la norma en este punto y en ningún otro de dicha bola (es decir un funcional que los expone), luego la conclusión se sigue del Teorema 3.5.3. Para el segundo caso estamos viendo lo mismo pero con la norma 1. En el último caso se razona de la misma manera, utilizando la caracterización de los puntos extremales de la bola para la norma de Ky-Fan, la cual puede encontrarse en [31, Teorema 4]. $\square$

Capítulo 4

Curvatura del Grupo Unitario

En este capítulo estudiaremos la geometría de $\mathcal{U}_n$, cuando consideramos métrica invariante a izquierda inducida por una norma de Finsler Ad-invariante. En particular nos va a interesar el concepto de curvatura seccional de este grupo. Recordamos que si M es una variedad diferenciable con una métrica riemanniana, se define la curvatura seccional en p asociada al plano $\pi = \text{gen}(x, y) \subseteq T_p M$ como

$$\text{sec}(x, y) = \frac{\langle R_p(x, y)y, x \rangle_p}{A^2(x, y)},$$

donde A^2 es el cuadrado del área del paralelogramo generado por x, y y R el tensor de curvatura. En el caso de no tener una métrica riemanniana y tener una de Finsler, lo que vamos a hacer es basarnos en la igualdad

$$\langle R_p(x, y)y, x \rangle_p = 6\|y - x\|^2 \lim_{r \rightarrow 0^+} \frac{r\|y - x\|_p - \text{dist}(\exp_p(rx), \exp_p(ry))}{r^2 \text{dist}(\exp_p(rx), \exp_p(ry))},$$

donde $p \in M$ y $x, y \in T_p M$. Dicha igualdad fue dada por Milnor en [54]. Esta estrategia fue usada en [12] y en [41] para estudiar el signo de la expresión $\langle R(x, y)y, x \rangle$ en distintos contextos. En este trabajo la usaremos para estudiar el grupo unitario y daremos una definición de curvatura seccional que puede ser usada con más generalidad.

4.1. Un poco de Geometría Riemanniana

Vamos a comenzar recordando algunas definiciones clásicas que necesitamos de geometría riemanniana, para fijar notación y comprender que nos dice la curvatura seccional en términos locales del comportamiento de las geodésicas. Para los resultados que siguen (M, g) va a ser una variedad riemanniana de dimensión finita, la cual supondremos suave, es decir C^∞ . Mas detalles sobre los postulados enunciados se pueden ver en cualquier libro de introducción a la geometría riemanniana, como por ejemplo el libro de Lee [49]. Vamos a notar como es usual ∇ a la conexión de Levi-Civita de la variedad riemanniana que es la única conexión que preserva la métrica y es libre de torsión. Mas específicamente, si X, Y, Z son campos en M y notamos $Xg(Y, Z)$ a la derivada del campo $g(Y, Z)$ en la dirección de X , tenemos que

- $Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ (preserva la métrica).

- $\nabla_X Y - \nabla_Y X = [X, Y]$ (es libre de torsión).

Asociado a dicha conexión, recordamos la definición del tensor de curvatura

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$$

y algunas de sus propiedades:

- R es lineal en cada coordenada.
- $R(X, Y)Z = -R(Y, X)Z$.
- $R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$.
- $\langle R(X, Y)Z, W \rangle = -\langle R(X, Y)W, Z \rangle$.
- $\langle R(X, Y)Z, W \rangle = \langle R(Z, W)X, Y \rangle$.

Observación 4.1.1. Si $\gamma : I \rightarrow M$ es una curva suave, un campo suave sobre dicha curva es una aplicación $X : I \rightarrow TM$ suave, que verifica $X(t) \in T_{\gamma(t)}M$ para todo $t \in I$. Recordamos además que la conexión de Levi-Civita induce una derivada covariante a lo largo de la curva γ , que a cada campo vectorial X a lo largo de la misma le asigna otro campo que notaremos $\frac{D}{dt}X$. Dados X, Y dos campos a lo largo de γ y f una función suave a valores reales, esta aplicación satisface

- $\frac{D}{dt}(X + Y) = \frac{D}{dt}X + \frac{D}{dt}Y$.
- $\frac{D}{dt}(fX) = \frac{df}{dt}X + f\frac{D}{dt}X$.
- Si X está inducido por un campo definido globalmente $\tilde{X}$, es decir $X(t) = \tilde{X}(\gamma(t))$, entonces $\frac{D}{dt}X = \nabla_{\dot{\gamma}}\tilde{X}$.

Usaremos indistintamente las notaciones $\frac{D}{dt}$ y $\nabla_{\dot{\gamma}}$ para la derivada covariante sobre campos suaves a lo largo de γ . Recordamos que $\dot{\gamma}$ se puede pensar como un campo a lo largo de γ a partir de $\dot{\gamma}(\gamma(t)) = \dot{\gamma}(t)$ y que dados dos campos a lo largo de γ podemos calcular

$$\frac{d}{dt}\langle X(t), Y(t) \rangle_{\gamma(t)} = \langle \nabla_{\dot{\gamma}}X(t), Y(t) \rangle_{\gamma(t)} + \langle X(t), \nabla_{\dot{\gamma}}Y(t) \rangle_{\gamma(t)}$$

Definición 4.1.2 (Geodésicas). Dada (M, ∇) una variedad riemanniana, decimos que una curva γ es una geodésica si tiene aceleración nula, es decir si $\nabla_{\dot{\gamma}}\gamma \equiv 0$.

Observación 4.1.3. Dado $p \in M$ y $v \in T_pM$, existe una única geodésica $\gamma : (-\epsilon, \epsilon) \rightarrow M$ que satisface $\gamma(0) = p$ y $\dot{\gamma}(0) = v$. Notaremos a esta curva $\gamma_{p,v}$.

Definición 4.1.4 (Exponencial Riemanniana). Definimos el conjunto

$$\mathcal{E} = \{(p, v) \in TM : \gamma_{p,v} \text{ está definida en un intervalo que contiene a } [0, 1]\},$$

y la exponencial riemanniana como $\exp : \mathcal{E} \rightarrow M$ dada por $\exp(p, v) = \gamma_{p,v}(1)$.

Fijado $p \in M$ y considerando el conjunto $\mathcal{E}_p = \{v \in T_pM : \gamma_{p,v} \text{ está definida en un intervalo que contiene a } [0, 1]\}$ tenemos la aplicación $\exp_p : \mathcal{E}_p \subseteq T_pM \rightarrow M$ dada por $\exp_p(v) = \gamma_{p,v}(1)$.

Proposición 4.1.5. (Propiedades de la exponencial).

- (1) $\mathcal{E}$ es un abierto de TM que contiene la sección nula, y $\mathcal{E}_p$ es estrellado con respecto a 0.
- (2) Para cada $(p, v) \in T_p M$, se tiene que $\gamma_{p,v}(t) = \exp_p(tv)$, donde la igualdad vale cuando alguno de los lados está definido (equivalentemente ambos).
- (3) El mapa exponencial es suave.
- (4) Si $c \in \mathbb{R}$, tenemos que $\exp_{p,cv}(t) = \exp_{p,v}(ct)$, equivalentemente $\gamma_{p,v}(ct) = \gamma_{p,cv}(t)$.

Demostración. La prueba de estos enunciados puede ser encontrada en [49, Proposición 5.7] y en [49, Lema 5.8]. $\square$

Lema 4.1.6. Dado $p \in M$ existen $\mathcal{U} \subseteq T_p M$ entorno del origen y $\mathcal{V} \subseteq M$ entorno de p , de forma tal que $\exp_p : \mathcal{U} \rightarrow \mathcal{V}$ es un difeomorfismo.

Demostración. Ver [49, Lema 5.10]. $\square$

Los resultados mencionados a continuación pueden ser encontrados en [49, Capítulo 10].

Definición 4.1.7 (Campo de Jacobi). Sea γ una geodésica de la variedad (M, g) , y J un campo sobre M . Decimos que J es un *campo de Jacobi* si satisface

$$\frac{D^2}{dt^2} J(t) + R(J(t), \dot{\gamma}(t))\dot{\gamma}(t) = 0. \quad (4.1)$$

Observación 4.1.8. Si γ es una geodésica definida en $I \subseteq \mathbb{R}$ y $a \in I$ es tal que $\gamma(a) = p \in M$, dados $X, Y \in T_p M$ existe un único campo de Jacobi que satisface las condiciones $J(a) = X$ y $D_t J(a) = Y$.

Observación 4.1.9. (Notaciones) Sean M, N variedades diferenciables y sea $f : M \rightarrow N$ una función suave, para $p \in M$ notaremos como $f_{*,p}$ a su diferencial $f_{*,p} : T_p M \rightarrow T_{f(p)} N$. Usaremos también la notación Df_p cuando trabajemos en un ejemplo concreto como $\mathcal{U}_n$.

Observación 4.1.10. Sean $\gamma : [a, b] \rightarrow M$ un segmento geodésico y $\Gamma : (-\epsilon, \epsilon) \rightarrow M$ una variación por geodésicas de γ , es decir $\Gamma_s(t)$ es geodésica para cada s fijo, entonces el campo $\frac{d}{ds}|_{s=0} \Gamma(s, t)$ es un campo de Jacobi a lo largo de γ . Más aún, todo campo de Jacobi puede ser construido de esta forma para cierta variación de geodésicas.

Sean $p \in M$ y γ una geodésica con $p = \gamma(0)$. Si tomamos entornos de $0 \in T_p M$ y de $p \in M$ de forma tal que $\exp_p : \mathcal{U} \rightarrow \mathcal{V}$ sea un difeomorfismo, tenemos que $\gamma(t) = \exp_p(tx)$ para cierto $x \in T_p M$. Considerando la función $\Gamma(s, t) = \exp_p(t(x + sy))$ para s suficientemente chico e $y \in T_p M$, tenemos una variación por geodésicas de γ . Luego por la observación previa el campo $\eta(t) = \frac{d}{ds}|_{s=0} \Gamma(s, t)$ es un campo de Jacobi. Aplicando la regla de la cadena obtenemos que

$$\eta(t) = (\exp_p)_{*,tx}(ty) = tg(t),$$

donde $g(t) = (\exp_p)_{*,tx}(y)$.

Notamos que $\eta(0) = 0$ y que $\dot{\eta}(0) = y \in T_p M$, luego $\nabla_{\dot{\gamma}}(0) = y$.

El siguiente lema nos servirá para establecer la relación entre la curvatura de la variedad y las geodésicas de la métrica. Esta relación es la motivación de lo que va a ser desarrollado en las siguientes secciones de este capítulo. Para establecer una conexión entre el tensor de curvatura y nuestro límite en cuestión, vamos a considerar la función $f(t) = \|\eta(t)\|_{\gamma(t)} - t\|y\|_{\gamma(0)=p}$.

Observación 4.1.11. Notamos que la función f es suave en todo t en el cual $\eta(t) \neq 0$, por estar trabajando en una variedad C^∞

Notación 4.1.12. Con el fin de aliviar la lectura, en el siguiente lema, usaremos la notación $|\cdot|$ para la norma y dejaremos de aclarar en que punto estamos considerando las normas y los productos internos. Luego retomaremos a la notación usual de esta tesis indicando por $\|\cdot\|$ a las normas y $\langle \cdot, \cdot \rangle$ a las normas de Finsler.

Lema 4.1.13. Sea $f(t) = \|\eta(t)\|_{\gamma(t)} - t\|y\|_{\gamma(0)=p}$, la función que definimos previamente, si t es tal que $\eta(t) \neq 0$, entonces

$$\begin{aligned} f'(t) &= \frac{1}{\|\eta\|} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle - \|y\| \\ f''(t) &= -\frac{1}{\|\eta\|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle + \frac{1}{\|\eta\|} |\nabla_{\dot{\gamma}} \eta|^2 + \frac{1}{\|\eta\|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle. \end{aligned}$$

Además en $t = 0$ vale $f(0) = f'(0) = f''(0) = 0$ y $f'''(0) = -\frac{\langle R(x,y)y,x \rangle}{\|y\|}$.

Demostración. Calculamos la primer derivada en t , en el caso que $\eta(t) \neq 0$:

$$\begin{aligned} f'(t) &= \frac{d}{dt} (\|\eta\| - t\|y\|) = \frac{d}{dt} \langle \eta, \eta \rangle^{\frac{1}{2}} - \|y\| = \frac{1}{2\|\eta\|} 2\langle \eta, \nabla_{\dot{\gamma}} \eta \rangle - \|y\| \\ &= \frac{1}{\|\eta\|} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle - \|y\|. \end{aligned}$$

En $t = 0$, como $\eta(0) = 0$, la tenemos que calcular por definición. Recordamos que $\eta(t) = tg(t)$ donde $g(t) = \exp_{*,tx}(y)$ y entonces

$$\lim_{t \rightarrow 0} \frac{\eta(t)}{t} = \lim_{t \rightarrow 0} (g(t) + tg'(t)) = \exp_{*,0}(y) = y, \quad (4.2)$$

donde aplicamos la regla de L'Hopital. Luego se sigue que

$$f'(0) = \lim_{t \rightarrow 0} \frac{f(t)}{t} = \lim_{t \rightarrow 0} \left| \frac{\eta(t)}{t} \right| - \|y\| = \lim_{t \rightarrow 0} \left| \frac{\eta(t)}{t} \right| - \|y\| = \|y\| - \|y\| = 0.$$

Para calcular la segunda derivada en los puntos en los cuales $\eta(t) \neq 0$ simplemente usamos la regla del producto para la derivada de un producto interno. La cuenta es directa al igual que en la primer derivada y se obtiene

$$f''(t) = -\frac{1}{\|\eta\|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2 + \frac{1}{\|\eta\|} |\nabla_{\dot{\gamma}} \eta|^2 + \frac{1}{\|\eta\|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle.$$

Para calcular $f''(0)$, debemos calcular el límite cuando t tiende a cero de $f''(t)$. Previamente a esto probaremos que

$$\lim_{t \rightarrow 0} \frac{1}{\|\eta\|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle = 0.$$

Usando la ecuación 4.1 y el resultado del límite 4.2 obtenemos

$$\frac{1}{|\eta|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle = \frac{1}{|\eta|} \langle \eta, R(\dot{\gamma}, \eta) \dot{\gamma} \rangle = \frac{1}{|\frac{\eta}{t}|} \left\langle \frac{\eta}{t}, R(\dot{\gamma}, \eta) \dot{\gamma} \right\rangle \longrightarrow \frac{1}{|y|} \langle y, R(x, 0)x \rangle = 0.$$

Luego para probar que la segunda derivada tiende a cero, tenemos que ver que

$$\lim_{t \rightarrow 0} -\frac{1}{|\eta|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2 + \frac{1}{|\eta|} |\nabla_{\dot{\gamma}} \eta|^2 = 0,$$

ya que hemos probado que el otro término tiende a cero. Reacomodando la expresión tenemos que

$$\begin{aligned} -\frac{1}{|\eta|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2 + \frac{1}{|\eta|} |\nabla_{\dot{\gamma}} \eta|^2 &= \frac{1}{|\eta|^3} \left(|\eta|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2 \right) \\ &= \frac{1}{|\frac{\eta}{t}|^3} \frac{|\eta|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t^3}. \end{aligned}$$

Nuevamente, por el resultado del límite 4.2 tenemos que ver que $\lim_{t \rightarrow 0} \frac{|\eta|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t^3} = 0$.

Aplicando la regla de L'Hopital, la identidad de Jacobi y operando:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{|\eta|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t^3} &= \lim_{t \rightarrow 0} \frac{2|\eta|^2 \langle \nabla_{\dot{\gamma}} \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle - 2 \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{3t^2} \\ &= \lim_{t \rightarrow 0} \frac{2}{3} |\eta| \left| \frac{\eta}{t} \right| \left\langle \nabla_{\dot{\gamma}} \eta, R\left(\frac{\eta}{t}, \dot{\gamma}\right) \dot{\gamma} \right\rangle - \lim_{t \rightarrow 0} \frac{2 \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{3t^2} \\ &= 0 - \frac{2}{3} \lim_{t \rightarrow 0} \frac{|\eta| \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle \frac{1}{|\eta|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{t^2} \\ &= -\frac{2}{3} \lim_{t \rightarrow 0} \left| \frac{\eta}{t} \right| \left\langle \frac{\eta}{t}, \nabla_{\dot{\gamma}} \eta \right\rangle \frac{1}{|\eta|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle = 0. \end{aligned}$$

En la última igualdad usamos el límite que habíamos calculado al comienzo del cálculo de la derivada segunda. Concluimos que

$$\lim_{t \rightarrow 0} f''(t) = 0,$$

luego usando que $f'(0) = 0$ y la regla de L'Hopital se sigue que $f''(0) = 0$.

Para finalizar la demostración, debemos calcular la derivada tercera de f en cero, tenemos que

$$f'''(0) = \lim_{t \rightarrow 0} \frac{f''(t) - f''(0)}{t} = \lim_{t \rightarrow 0} \frac{f''(t)}{t}.$$

luego

$$f'''(0) = \lim_{t \rightarrow 0} \frac{-\frac{1}{|\eta|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2 + \frac{1}{|\eta|} |\nabla_{\dot{\gamma}} \eta|^2 + \frac{1}{|\eta|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{t}.$$

Repartiendo el denominador, obtenemos tres términos. El del medio ya vimos que tiende a cero, luego resta calcular los otros dos límites que llamaremos L_1 y L_2 . En tal caso será $f'''(0) = L_1 + L_2$ donde:

$$L_1 = \lim_{t \rightarrow 0} \frac{\frac{1}{|\eta|} \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{t}.$$

$$L_2 = \lim_{t \rightarrow 0} \frac{-\frac{1}{|y|^3} \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t}.$$

Calculamos L_1 , usando la misma estrategia que en la prueba de que $f''(0) = 0$

$$L_1 = \lim_{t \rightarrow 0} \frac{1}{t|y|} \langle \eta, R(\dot{\gamma}, \eta) \dot{\gamma} \rangle = \lim_{t \rightarrow 0} \frac{1}{|\frac{y}{t}|} \left\langle \frac{\eta}{t}, R(\dot{\gamma}, \frac{\eta}{t}) \dot{\gamma} \right\rangle = \frac{\langle y, R(x, y)x \rangle}{|y|},$$

con lo cual

$$L_1 = \frac{\langle y, R(x, y)x \rangle}{|y|}.$$

Usando el mismo reordenamiento que en la prueba de la derivada segunda y aplicando la regla de L'Hopital se sigue que

$$\begin{aligned} L_2 &= \lim_{t \rightarrow 0} \frac{1}{|\frac{y}{t}|^3} \frac{|y|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t^4} = \frac{1}{|y|^3} \lim_{t \rightarrow 0} \frac{|y|^2 |\nabla_{\dot{\gamma}} \eta|^2 - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle^2}{t^4} \\ &= \frac{1}{|y|^3} \lim_{t \rightarrow 0} \frac{2|y|^2 \langle \nabla_{\dot{\gamma}} \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle - 2 \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle \langle \eta, \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \eta \rangle}{4t^3} \\ &= \frac{1}{2|y|^3} \lim_{t \rightarrow 0} \frac{|y|^2 \langle \nabla_{\dot{\gamma}} \eta, R(\dot{\gamma}, \eta) \dot{\gamma} \rangle - \langle \eta, \nabla_{\dot{\gamma}} \eta \rangle \langle \eta, R(\dot{\gamma}, \eta) \dot{\gamma} \rangle}{t^3} \\ &= \frac{1}{2|y|^3} \lim_{t \rightarrow 0} \left| \frac{\eta}{t} \right|^2 \langle \nabla_{\dot{\gamma}} \eta, R(\dot{\gamma}, \frac{\eta}{t}) \dot{\gamma} \rangle - \frac{1}{2|y|^3} \lim_{t \rightarrow 0} \left\langle \frac{\eta}{t}, \nabla_{\dot{\gamma}} \eta \right\rangle \left\langle \frac{\eta}{t}, R(\dot{\gamma}, \frac{\eta}{t}) \dot{\gamma} \right\rangle \\ &= \frac{1}{2|y|^3} \langle y, y \rangle \langle y, R(x, y)x \rangle - \frac{1}{2|y|^3} \langle y, y \rangle \langle y, R(x, y)x \rangle = 0, \end{aligned}$$

donde estamos usando que $\nabla_{\dot{\gamma}} \eta \rightarrow y$. En conclusion, tenemos que $L_2 = 0$ y luego

$$f'''(0) = L_1 = \frac{\langle y, R(x, y)x \rangle}{|y|} = -\frac{\langle R(x, y)y, x \rangle}{|y|},$$

donde en el último paso usamos que el producto interno es real y por lo tanto simétrico y las propiedades del tensor de curvatura mencionadas al comienzo de esta sección. $\square$

El siguiente lema nos va a ser útil para probar la fórmula de Milnor de la curvatura, la demostración del mismo puede ser encontrada con detalle en las notas de Meyer [53].

Lema 4.1.14. Sea (M, g) una variedad riemanniana de dimensión finita, $p \in M$ y $x, y \in T_p M$, entonces

$$\text{dist}(\exp_p(tx), \exp_p(ty)) = t\|y - x\|_p - \frac{t^3}{6} \frac{\langle R_p(x, y)y, x \rangle_p}{\|y - x\|_p} + O(t^4).$$

Proposición 4.1.15 (Fórmula de Milnor). Sea (M, g) una variedad riemanniana de dimensión finita y sean $x, y \in T_p M$. Entonces

$$\langle R_p(x, y)y, x \rangle_p = 6\|y - x\|_p^2 \lim_{t \rightarrow 0^+} \frac{t\|y - x\|_p - \text{dist}(\exp_p(tx), \exp_p(ty))}{t^2 \text{dist}(\exp_p(tx), \exp_p(ty))}.$$

Demostración. Comenzamos utilizando el lema previo para desarrollar el límite del enunciado:

$$\begin{aligned}
 \lim_{t \rightarrow 0^+} \frac{t\|y-x\|_p - \text{dist}(\exp_p(tx), \exp_p(ty))}{t^2 \text{dist}(\exp_p(tx), \exp_p(ty))} &= \lim_{t \rightarrow 0^+} \frac{\frac{t^3}{6} \frac{\langle R_p(x,y)y,x \rangle}{\|y-x\|_p} + O(t^4)}{t^3\|y-x\|_p - \frac{t^5}{6} \frac{\langle R_p(x,y)y,x \rangle_p}{\|y-x\|_p} + O(t^6)} \\
 &= \lim_{t \rightarrow 0^+} \frac{\frac{1}{6} \frac{\langle R_p(x,y)y,x \rangle}{\|y-x\|_p} + O(t)}{\|y-x\|_p - \frac{t^2}{6} \frac{\langle R_p(x,y)y,x \rangle_p}{\|y-x\|_p} + O(t^3)} \\
 &= \frac{\langle R_p(x,y)y,x \rangle_p}{6\|y-x\|_p^2},
 \end{aligned}$$

multiplicando a ambos lados por $6\|y-x\|_p^2$ obtenemos la fórmula de Milnor. $\square$

Habiendo probado estos resultados daremos el resultado principal de esta sección introductoria:

Teorema 4.1.16. Sea M una variedad con una métrica riemanniana, son equivalentes

- (1) Para todo $p \in M$ y $x, y \in T_p M$ vale que $\text{dist}(\exp_p(tx), \exp_p(ty)) \leq \|tx - ty\|_p$ para $t \geq 0$ suficientemente chico (donde con d nos referimos a la distancia dada por el ínfimo longitudes de curvas suaves a trozos en M).
- (2) Para todo $p \in M$ y $x, y \in T_p M$ se cumple que $\text{sec}(x, y) = \frac{\langle R_p(x,y)y,x \rangle_p}{A^2(x,y)} \geq 0$. En otras palabras, M es de curvatura seccional no negativa.

Demostración. Supongamos primero (1) y apliquemos la fórmula de Milnor dada en la introducción de este capítulo, en tal caso

$$\langle R_p(x, y)y, x \rangle_p = 6\|y-x\|_p^2 \lim_{r \rightarrow 0^+} \frac{r\|y-x\|_p - \text{dist}(\exp_p(rx), \exp_p(ry))}{r^2 \text{dist}(\exp_p(rx), \exp_p(ry))}$$

y este límite es mayor o igual a cero por (1), se sigue entonces que $\text{sec}(x, y) \geq 0$ y luego vale (2).

Si asumimos que vale (2), observamos que la función definida en el Lema 4.1.13 cumple $f(0) = f'(0) = f''(0) = 0$ y $f'''(0) \leq 0$, entonces tenemos que $f(t) \leq 0$ para $t \geq 0$ suficientemente pequeño, es decir

$$\|(\exp_p)_{*,tx}(ty)\|_{\exp_p(tx)} \leq \|ty\|_p$$

para t suficientemente chico. Si dividimos por t , llamamos $v = tx$ y $u = y$, se sigue que

$$\|(\exp_p)_{*,v}(u)\|_{\exp_p(v)} \leq \|u\|_p$$

si v está suficientemente cerca de $0 \in T_p M$.

Habiendo demostrado esto, consideramos $\mathcal{U}$ un abierto de $0 \in T_p M$ y $\mathcal{V}$ un abierto de p en M de forma tal que $\exp : \mathcal{U} \rightarrow \mathcal{V}$ sea un difeomorfismo y que además se cumpla que para $v \in \mathcal{U}$ valga

$$\|(\exp_p)_{*,v}(u)\|_{\exp_p(v)} \leq \|u\|_p,$$

si $u \in T_p M$. Podemos asumir además, eventualmente achicando los abiertos que $\mathcal{U}$ es convexo.

Sean ahora $x, y \in T_p M$ y t suficientemente chico, de manera tal que $tx, ty \in \mathcal{U}$. Si Γ es el segmento que une a tx con ty en $\mathcal{U}$ entonces

$$\begin{aligned} \text{dist}(\exp_p(tx), \exp_p(ty)) &\leq \text{Long}(\exp_p(\Gamma)) = \int_0^1 \|(\exp_p)_*\dot{\Gamma}\|_{\exp(\Gamma)} \leq \\ &\leq \int_0^1 \|\dot{\Gamma}\|_{\Gamma} = \text{Long}(\Gamma) = \|ty - tx\|_p, \end{aligned}$$

concluyéndose así que vale (1). □

4.2. Grupos de Lie

En esta sección, además de fijar notación, vamos a discutir la relación entre la exponencial riemanniana en un grupo de Lie y la exponencial del grupo de Lie. Veremos que en el caso de que la métrica riemanniana sea provenga de una norma Ad-invariante las exponenciales en $p = 0$ van a coincidir, lo cual va a ser la motivación para la definición de curvatura que daremos en el caso no riemanniano. Discutiremos además la cuasi distancia inducida por la métrica y daremos sus propiedades generales. Daremos las definiciones en un grupo de Lie general, ya que si bien en las secciones siguientes a esta trabajaremos sobre el grupo unitario $\mathcal{U}_n$ para seguir el eje de esta tesis, en el capítulo siguiente retomaremos el caso abstracto.

Para la presentación de generalidades de grupos de Lie nos basaremos en el libro de Lee [50], pero los resultados no demostrados de esta sección pueden encontrarse en básicamente cualquier libro que trate el tema.

Definición 4.2.1. Un grupo de Lie es un grupo G con una estructura de variedad suave (para nosotros será C^∞) para el cual las operaciones de multiplicación $m : G \times G \rightarrow G$ e inversión $i : G \rightarrow G$ son suaves. Al elemento neutro del grupo lo notaremos $1 \in G$ y al álgebra de Lie de G la notaremos como $\text{Lie}(G) = T_1 G = \mathfrak{g}$.

Observación 4.2.2. Si bien muchas de las definiciones e incluso de los resultados se pueden dar en forma más general, vamos a asumir que la dimensión de G es finita ya que los resultados originales de este trabajo son todos en este contexto.

Para $g \in G$, las aplicaciones $l_g, r_g, c_g : G \rightarrow G$ dadas por

$$l_g : h \rightarrow gh \quad r_g : h \rightarrow hg \quad c_g : h \rightarrow ghg^{-1}$$

resultan difeomorfismos y l_g conmuta con r_h para $g, h \in G$, lo cual implica que sus diferenciales también lo hacen, lo cual usaremos más adelante. Notaremos como $GL(\mathfrak{g})$ al grupo de isomorfismos lineales de $\mathfrak{g}$. Por otro lado, como c_g es un difeomorfismo, su diferencial $\text{Ad}_g := (c_g)_{*,1}$ pertenece a $GL(\mathfrak{g})$, esto nos da la *representación adjunta* que es la aplicación $\text{Ad} : G \rightarrow GL(\mathfrak{g})$ dada por $\text{Ad}(g) = \text{Ad}_g = (c_g)_{*,1}$.

Definición 4.2.3 (Campos invariantes a izquierda). Decimos que un campo $X : G \rightarrow TG$ es invariante a izquierda si para todo $g, h \in G$ satisface

$$X(l_h(g)) = (l_h)_{*,g} X(g).$$

Al conjunto de estos campos lo notaremos $\mathfrak{X}(G)$.

Proposición 4.2.4. Si X es un campo invariante a izquierda en G , entonces X queda determinado por lo que vale en 1. Además, dado $v \in \mathfrak{g}$ la aplicación $X_v(g) = (l_g)_* v$ define un campo invariante a izquierda.

Como consecuencia de la proposición previa tenemos que todo campo invariante a izquierda es suave y de la forma X_v para cierto $v \in \mathfrak{g}$ y tenemos la biyección usual entre $\mathfrak{g}$ y $\mathfrak{X}(G)$ dada por $v \rightarrow X_v$ con inversa $X_v \rightarrow X_v(1)$. Por otro lado, es un hecho conocido que el corchete de dos campos invariantes a izquierda resulta nuevamente un campo invariante a izquierda y que como consecuencia de este hecho podemos definir el corchete en $\mathfrak{g}$ como

$$[v, w] := [X_v, X_w](1),$$

obteniendo así una estructura de álgebra de Lie en $\mathfrak{g}$.

Definición 4.2.5 (Exponencial). Vamos a definir la exponencial $\exp : \mathfrak{g} \rightarrow G$ como $\exp(v) = \alpha(1)$ donde $\alpha(t)$ es la única solución de la ecuación diferencial $\alpha' = X_v(\alpha)$ con condiciones iniciales $\alpha(0) = 1$.

Proposición 4.2.6. (1) La ecuación diferencial que define a la exponencial tiene solución global. Lo cual nos permite definir $\exp(tv)$ para todo $t \in \mathbb{R}$.

(2) Si $f : G \rightarrow H$ es un morfismo de grupos de Lie entonces $f_{*,1} : \mathfrak{g} \rightarrow \mathfrak{h}$ es un morfismo de álgebras de Lie y se tienen los siguientes el siguiente diagramas conmutativos:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{f_{*,1}} & \mathfrak{h} \\ \downarrow \exp_G & & \downarrow \exp_H \\ G & \xrightarrow{f} & H \end{array} \qquad \begin{array}{ccc} \mathfrak{g} & \xrightarrow{\text{ad}} & \text{End}(\mathfrak{g}) \\ \downarrow \exp_G & & \downarrow \exp_H \\ G & \xrightarrow{\text{Ad}} & GL(\mathfrak{g}) \end{array}$$

donde notamos que el segundo diagrama es un caso particular del primero definiendo $\text{ad} = \text{Ad}_{*,1} = \text{Ad}_{*,1}$.

(3) Si $v, w \in \mathfrak{g}$ entonces $\text{ad}_v(w) = [v, w]$.

Observación 4.2.7. En el caso de un grupo de Lie G inmerso en $GL(\mathbb{C})$ con el corchete dado por $[v, w] = vw - wv$, tenemos que la exponencial coincide con la exponencial usual dada por la serie. Por otro lado notamos que el segundo diagrama conmutativo del segundo ítem de la proposición anterior lo probamos para este caso en el Lema 3.1.2.

Si G es un grupo de Lie, notaremos TG a su fibrado tangente, el cual puede ser identificado con $G \times \mathfrak{g}$. Cuando hablemos de una métrica en el grupo en cuestión nos estaremos refiriendo a una métrica de Finsler o una métrica riemanniana sobre TG , por otro lado al trabajar en $\mathfrak{g}$ las llamaremos normas y notaremos $|\cdot|$. Si la norma proviene de una métrica o induce una métrica, valdrá la igualdad $|\cdot| = |\cdot|_1$ (donde 1 es el neutro del grupo). Cuando trabajemos en el grupo como espacio métrico hablaremos en términos de distancias.

Definición 4.2.8 (Métrica bi-invariante). Dado un grupo G con una métrica riemanniana decimos que:

- (1) La métrica es invariante a izquierda si para cada $g \in G$ la aplicación l_g es una isometría, es decir, dados $h \in G$ y $v, w \in T_h G$

$$\langle v, w \rangle_h = \langle (l_g)_* v, (l_g)_* w \rangle_{gh}.$$

- (2) La métrica es invariante a derecha si para cada $g \in G$ la aplicación r_g es una isometría, es decir, dados $h \in G$ y $v, w \in T_h G$

$$\langle v, w \rangle_h = \langle (r_g)_* v, (r_g)_* w \rangle_{gh}.$$

- (3) La métrica es bi-invariante si lo es a izquierda y a derecha.

En el paper de Milnor [55], él demostró que si tenemos una métrica bi-invariante entonces la conexión de Levi-Civita de la métrica vista en el álgebra de Lie es $\nabla_x = \text{ad } x$; de aquí se desprende que las geodésicas de la métrica coinciden con traslaciones de los grupos a un parámetro (una demostración completa de esta afirmación puede verse en el libro de Price [60, Capítulo 4, Sección 3]). Este resultado nos dice que en este contexto la exponencial riemanniana coincide con la exponencial del grupo de Lie y nos permitirá utilizar la fórmula para la curvatura seccional con la exponencial del grupo.

Esto no es siempre cierto, en [7, Teorema 2.14] podemos ver un ejemplo sobre el grupo de Lie de matrices inversibles con la métrica riemanniana inducida por el producto interno de la traza, en donde tenemos que las geodésicas no son en general grupos a un parámetro y en tal caso la exponencial riemanniana no coincide con la exponencial como grupo de Lie. Más específicamente, si consideramos $g_0 \in G = GL_n(\mathbb{C})$ y $v_0 \in \text{Lie}(GL_n(\mathbb{C})) = M_n(\mathbb{C})$ entonces la única geodésica de la conexión con punto inicial g_0 y velocidad inicial $g_0 v_0$ está dada por

$$g(t) = g_0 e^{t v_0^*} e^{t(v_0 - v_0^*)}$$

la cual no es un grupo a un parámetro si v_0 no conmuta con v_0^* .

Definición 4.2.9. Sea G un grupo de Lie con una norma de Finsler $|\cdot|$, decimos que dicha norma es Ad-invariante si cumple que la aplicación Ad_g es una isometría para cada $g \in G$, es decir $|\text{Ad}_g(v)| = |v|$ para todo $v \in \mathfrak{g}$.

Notamos que la definición anterior es una generalización de la Definición 2.2.1.

Observación 4.2.10 (Métricas bi-invariantes). Notamos que, como $\text{Ad}_g = (c_g)_* = (l_g)_* (r_{g^{-1}})_*$, entonces toda métrica bi-invariante en G induce una norma Ad-invariante en $\mathfrak{g}$. Por otro lado si tenemos una norma Ad-invariante en un grupo de Lie $\mathfrak{g}$, podemos definir la métrica $|X|_g = |(l_{g^{-1}})_* X|$ para $X \in \mathfrak{g}$. Notamos que es una métrica bi-invariante ya que si $g, h \in G$ y $X \in T_h G$ entonces

$$|(l_g)_* X|_{gh} = |(l_h^{-1})_* (l_g^{-1})_* (l_g)_* X| = |(l_h^{-1})_* X| = |X|_h.$$

Por otro lado si $X \in T_g G$

$$\begin{aligned} |(r_h)_* X|_{gh} &= |(l_h^{-1})_* (l_g^{-1})_* (r_h)_* X| = |(l_h^{-1})_* (r_h)_* (l_g^{-1})_* X| \\ &= |\text{Ad}_{h^{-1}}(l_g^{-1})_* X| = |(l_g^{-1})_* X| = |X|_g, \end{aligned}$$

donde en el penúltimo paso usamos que la norma es Ad-invariante. En conclusión tenemos que toda métrica bi-invariante en TG induce una norma Ad-invariante en $\mathfrak{g}$ y que a partir de una norma Ad-invariante podemos construirnos una métrica bi-invariante. Como toda métrica bi-invariante es en particular invariante a izquierda, tenemos que este proceso es reversible y las métricas bi-invariantes en TG están identificadas con las normas Ad-invariantes en $\mathfrak{g}$.

Por la observación anterior tenemos que en el caso riemanniano, donde la métrica viene inducida por una norma Ad-invariante, vale la fórmula de la curvatura de Milnor:

$$\langle R(x, y)y, x \rangle = 6\|y - x\|^2 \lim_{r \rightarrow 0^+} \frac{r\|y - x\| - \text{dist}(\exp(rx), \exp(ry))}{r^2 \text{dist}(\exp(rx), \exp(ry))},$$

donde estamos pensando la exponencial del grupo en vez de la riemanniana (ya que coinciden) y estamos trasladando todo a $\mathfrak{g} = T_1 G$. Esto es fundamental y la motivación para la definición que vamos a dar en el contexto no riemanniano.

A continuación vamos a dar algunos resultados sobre la distancia en G inducida por una norma de Finsler Ad-invariante en $\mathfrak{g}$. Muchas veces la llamaremos distancia, pero como vamos a ver en breve, va a ser una cuasi distancia (no va a ser simétrica).

Observamos además que hasta ahora no le hemos pedido regularidad a la métrica de Finsler, pero la aplicación $(g, X) \mapsto |X|_g = |(l_{g^{-1}})_* X|$ va a ser siempre continua y en tal caso integrable.

Definición 4.2.11 (Caminos rectificables y longitudes). Decimos que una curva $\alpha : [a, b] \rightarrow G$ es *rectificable* si es suave en casi todo punto para alguna carta de G y la aplicación $t \mapsto |\alpha'(t)|_{\alpha(t)}$ es Lebesgue integrable. Para una curva suave a trozos o rectificable definimos su longitud como

$$\text{Long}(\alpha) = \int_a^b |\alpha'(t)|_{\alpha(t)} dt = \int_a^b |(l_{\alpha^{-1}(t)})_* \alpha'(t)| dt = \int_a^b |(l_{\alpha^{-1}})_* \alpha'|.$$

En general nos interesará trabajar en la componente conexa de la identidad, así que no perdemos nada en suponer que G es conexo, lo cual haremos de aquí en adelante.

Definición 4.2.12. Sean $g, h \in G$, definimos la distancia entre g y h como

$$\text{dist}(g, h) = \inf\{\text{Long}(\alpha) : \alpha : [a, b] \rightarrow G \text{ es rectificable}, \alpha(a) = g, \alpha(b) = h\},$$

notamos que como G es un grupo conexo y localmente arco-conexo (por ser una variedad) resultará ser arco-conexo y luego el conjunto dónde tomamos el ínfimo es no vacío. Como el mismo está acotado inferiormente por cero, el ínfimo queda bien definido.

Observación 4.2.13. La distancia definida previamente cumple todos los axiomas de distancia salvo quizás la simetría, luego no es una distancia en el sentido usual sino que es una cuasi distancia. Va a ser una distancia si y solo si la norma es completamente homogénea.

Proposición 4.2.14. La distancia definida previamente es bi-invariante, es decir

$$\text{dist}(kg, kh) = \text{dist}(g, h) = \text{dist}(gk, hk) \quad \forall g, h, k \in G.$$

Además, si la norma es completamente homogénea, vale que

$$\text{dist}(g, h) = \text{dist}(g^{-1}, h^{-1}).$$

Demostración. Comenzamos tomando una curva rectificable $\alpha : [a, b] \rightarrow G$ que verifique $\alpha(a) = g$ y $\alpha(b) = h$ y vamos a ver que las curvas $\beta(t) = k.\alpha(t)$ y $\gamma(t) = \alpha(t).k$, que también son rectificables y unen a los puntos kg con kh y gk con hk respectivamente, tienen la misma longitud que α . El resultado se seguirá entonces del hecho de que el conjunto de longitudes que se puede obtener es el mismo y el ínfimo por lo tanto también. Veamos lo dicho:

$$\begin{aligned} \text{Long}(\beta) &= \int_a^b |\beta'|_{\beta} = \int_a^b |(k\alpha)'|_{k\alpha} = \int_a^b |(l_k)_{*,\alpha}\alpha'|_{k\alpha} = \int_a^b |(l_{\alpha^{-1}})_{*,\alpha}(l_{k^{-1}})_{*,k}\alpha(l_k)_{*,\alpha}\alpha'| \\ &= \int_a^b |(l_{\alpha^{-1}})_{*,\alpha}\alpha'| = \int_a^b |\alpha'|_{\alpha} = \text{Long}(\alpha). \end{aligned}$$

Notamos que en este caso no hemos usado que la norma es Ad-invariante, lo cual usaremos en el siguiente caso

$$\begin{aligned} \text{Long}(\gamma) &= \int_a^b |\gamma'|_{\gamma} = \int_a^b |(\alpha k)'|_{\alpha k} = \int_a^b |(r_k)_{*,\alpha}\alpha'|_{\alpha k} = \int_a^b |(l_{k^{-1}})_{*,k}(l_{\alpha^{-1}})_{*,\alpha k}(r_k)_{*,\alpha}\alpha'| \\ &= \int_a^b |(l_{k^{-1}})_{*,k}(r_k)_{*,1}(l_{\alpha^{-1}})_{*,\alpha}\alpha'| = \int_a^b |\text{Ad}_{k^{-1}}(l_{\alpha^{-1}})_{*,\alpha}\alpha'| = \int_a^b |(l_{\alpha^{-1}})_{*,\alpha}\alpha'| \\ &= \int_a^b |\alpha'|_{\alpha} = \text{Long}(\alpha). \end{aligned}$$

Para la siguiente parte consideramos α como antes y la curva α^{-1} . Lo que queremos ver es que tienen la misma longitud, luego el argumento se sigue como en el caso anterior. Recordamos que en grupos de Lie vale que $(\alpha^{-1})' = -(r_{\alpha^{-1}})_{*,1}(l_{\alpha^{-1}})_{*,\alpha}\alpha'$, luego calculamos

$$\begin{aligned} \text{Long}(\alpha^{-1}) &= \int_a^b |(\alpha^{-1})'|_{\alpha^{-1}} = \int_a^b |-(r_{\alpha^{-1}})_{*,1}(l_{\alpha^{-1}})_{*,\alpha}\alpha'|_{\alpha^{-1}} = \int_a^b |(r_{\alpha^{-1}})_{*,1}(l_{\alpha^{-1}})_{*,\alpha}\alpha'|_{\alpha^{-1}} \\ &= \int_a^b |(l_{\alpha})_{*,\alpha^{-1}}(r_{\alpha^{-1}})_{*,1}(l_{\alpha^{-1}})_{*,\alpha}\alpha'| = \int_a^b |\text{Ad}_{\alpha}(l_{\alpha^{-1}})_{*,\alpha}\alpha'| = \int_a^b |\alpha'|_{\alpha} = \text{Long}(\alpha), \end{aligned}$$

donde usamos que la norma es homogénea para sacar el signo negativo y la Ad-invariancia posteriormente. $\square$

Más detalles sobre distancias asimétricas o cuasi distancias pueden ser encontrados en [51].

4.3. Geodésicas y distancia en $\mathcal{U}_n$

En esta sección vamos a recordar ciertas propiedades de los caminos rectificables del grupo unitario y vamos a establecer dos hechos importantes: todas las métricas son equivalentes en esta configuración y todo

mapa corto en $\mathfrak{u}_n$ induce un mapa corto en $\mathcal{U}_n$ mediante la exponencial de matrices. Además presentaremos algunas propiedades importantes de la distancia en $\mathcal{U}_n$ inducida por una norma Ad-invariante en $\mathfrak{u}_n$. Los resultados originales de esta sección publicados en [48] son el Teorema 4.3.6 y el Lema 4.3.8, siendo el segundo un resultado mas bien técnico pero que nos va a ser útil para los resultados principales de las siguientes secciones.

Comenzamos recordando ciertos resultados sobre la exponencial del grupo de Lie $\mathcal{U}_n$. Sea $\|\cdot\|_\infty$ la norma espectral en $M_n(\mathbb{C})$, denotamos

$$B_r^\infty = \{X \in M_n(\mathbb{C}) : \|X\|_\infty < r\},$$

y a su clausura como $\overline{B_r^\infty}$.

Por otro lado notaremos $BCH(X, Y)$ a la serie de corchetes definida para X e Y suficientemente pequeños como:

$$BCH(X, Y) = \exp^{-1}(\exp(X) \exp(Y)).$$

Proposición 4.3.1 (Mapa exponencial en $\mathcal{U}_n$). El mapa exponencial de $\mathcal{U}_n$ es la restricción del mapa exponencial de $M_n(\mathbb{C})$, con lo cual puede ser calculado como la serie de potencias usual $\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}$.

Sea $B = B_\pi^\infty \cap \mathfrak{u}_n$ y $\overline{B}$ su clausura.

- (1) El mapa exponencial $\exp : \mathfrak{u}_n \rightarrow \mathcal{U}_n$ es sobreyectivo y su restricción $\exp|_{\overline{B}}$ también lo es. La restricción $\exp|_B$, donde $B = B_\pi^\infty \cap \mathfrak{u}_n$, es un difeomorfismo con $\mathcal{V} = \{U \in \mathcal{U}_n : \|U - 1\|_\infty < 2\}$, el cual es un abierto denso del grupo $\mathcal{U}_n$. Además

$$\frac{2}{\pi} \|Z\|_\infty \leq \|1 - e^Z\|_\infty = 2 \sin(\|Z\|_\infty/2) \leq \|Z\|_\infty \quad \text{para } \|Z\|_\infty \leq \pi.$$

- (2) Si $X, Y \in \mathfrak{u}_n$ entonces

$$D \exp_X(Y) = e^X \int_0^1 \text{Ad}_{e^{-sX}} Y \, ds = e^X \int_0^1 e^{-s \text{ad} Y} X \, ds = e^X F(\text{ad} X)Y = (G(\text{ad} X)Y)e^X$$

donde F, G son los mapas enteros dados por $F(\lambda) = \frac{1-e^{-\lambda}}{\lambda}$, $G(\lambda) = e^\lambda F(\lambda)$.

- (3) Si $\|1 - e^{tX} e^{sY}\|_\infty < 2$ luego

$$BCH(tX, sY) = tX + sY + \frac{st}{2} [X, Y] + \frac{1}{12} \left(t^2 s [X, [X, Y]] + t s^2 [Y, [Y, X]] \right) + \text{términos de mayor orden}$$

los cuales son los primeros términos de la serie de Baker-Campbell-Hausdorff de (tX, sY) . En particular

$$BCH(tX, tY) = t(X + Y) + \frac{t^2}{2} [X, Y] + \frac{t^3}{12} ([X, [X, Y]] + [Y, [Y, X]]) + O(t^4).$$

- (4) Si $\|X\|_\infty + \|Y\|_\infty < \pi$, luego $BCH(X, Y)$ es una serie de polinomios homogéneos convergentes en $\text{Lie}(X, Y)$ (el álgebra de Lie cerrada generada por X e Y).

Demostración. Comenzamos por el primer ítem, sea $U \in \mathcal{U}_n$, entonces $U = VDV^*$ para cierta matriz diagonal $D = (d_1, d_2, \dots, d_n)$ y $V \in \mathcal{U}_n$. Consideramos la matriz $\log(D) = (\log(d_1), \log(d_2), \dots, \log(d_n))$, donde $\log$ es una rama del logaritmo complejo bien definida en los puntos donde queremos evaluar, como los números tienen módulo 1 entonces $\log(d_k) = \ln|d_k| + i \arg(d_k) = i \arg(d_k)$ donde $\arg$ es el argumento elegido para la rama del logaritmo. De esta forma $\log(D) \in \mathfrak{u}_n$ y $\exp(V \log(D) V^*) = V \exp(\log(D)) V^* = VDV^* = U$, lo cual prueba que la exponencial es sobreyectiva. Como no es necesario usar la misma rama del logaritmo para cada autovalor, siempre podremos elegir cada coordenada de $\log(D)$ de forma que $\arg(d_k) \in [-\pi, \pi]$, lo cual nos prueba que $\exp|_B$ sigue siendo sobreyectiva. Dado que el conjunto $\mathcal{V}$ es el conjunto de matrices unitarias que no tienen a -1 como autovalor, podemos aplicar la rama principal del logaritmo en dicho conjunto y obtener todas las matrices antihermitianas que no tienen a $i\pi$ como autovalor. Aclarado esto y vista la construcción anterior es fácil ver que esta rama es la inversa de la exponencial y entonces $\exp : B \rightarrow \mathcal{V}$ es biyectiva. Será un difeomorfismo por ser analítica con inversa analítica. La densidad de $\mathcal{V} \subseteq \mathcal{U}_n$ es una consecuencia inmediata de que $\mathcal{U}_n = \{U \in \mathcal{U}_n : \|U - 1\| \leq 2\}$.

Veamos las desigualdades enunciadas en (1):

Por un lado tenemos que, como $1 - e^Z$ es normal al ser unitario, su norma coincide con su radio espectral y además sus autovalores son de la forma $1 - e^{it}$ donde it es un autovalor de Z (esto se puede ver simplemente haciendo la cuenta con matrices o pensando en que es un cálculo funcional de una función continua en un elemento normal), en conclusión

$$\|1 - e^Z\|_\infty = \max_{it \in \sigma(Z)} |1 - e^{it}|,$$

donde el máximo esta tomado sobre los $t \in \mathbb{R}$ de forma que it es autovalor de Z . Desarrollando el módulo y utilizando identidades trigonométricas básicas obtenemos que $|1 - e^{it}| = 2|\sin(t/2)|$, como la función $\sin(t/2)$ es creciente en $[-\pi, \pi]$ y Z es normal, tenemos que para t de forma que it es autovalor de Z vale

$$-\sin(\|Z\|_\infty/2) \leq \sin(t/2) \leq \sin(\|Z\|_\infty/2),$$

de donde deducimos que

$$\|1 - e^Z\|_\infty = \max_{it \in \sigma(Z)} |1 - e^{it}| = 2\sin(\|Z\|_\infty/2) \leq \|Z\|.$$

Por otro lado, se puede ver que $|2\sin(t/2)| \geq \frac{2}{\pi}|t|$ y razonando de forma similar llegamos a que

$$\|1 - e^Z\|_\infty \geq \frac{2}{\pi}\|Z\|_\infty.$$

El segundo ítem es un resultado clásico de análisis matricial y puede ser encontrado por ejemplo en [34, Teorema 1.7, Capítulo II]

Para la tercer afirmación utilizamos (1), recordando que si $X, Y \in B$ son tales que $\exp(X) \cdot \exp(Y) \in \mathcal{V}$ tenemos bien definida la serie BCH en (X, Y) como $BCH(X, Y) = \log(\exp(X) \cdot \exp(Y))$. Los coeficientes de esta serie conocida como Fórmula de Dynkin fueron estudiados por Dynkin en 1947 [28]. Nosotros solo necesitaremos los primeros términos de dicha fórmula.

La cuarta afirmación fue probada en [22, Teorema 3.2]. □

Observación 4.3.2 (Grupos a un parámetro, segmentos). Sean $U, V \in \mathcal{U}_n$, por el ítem (1) de la Proposición 4.3.1 existe $Z \in \mathfrak{u}_n$ con $\|Z\|_\infty \leq \pi$ que cumple $e^Z = V^{-1}U$. A las curvas de la forma $\delta(t) = Ve^{tZ}$ las vamos a llamar *segmentos*, en este caso dicho segmento une a V con U en $\mathcal{U}_n$. Observamos que si $\|Z\|_\infty < \pi$ (equivalentemente si $\|V - U\|_\infty < 2$), existe exactamente un único segmento que une a V con U en $\mathcal{U}_n$, esto es consecuencia del primer ítem de la Proposición 4.3.1. Como $\delta(t)^{-1}\delta'(t) = Z$, tenemos que $|\delta'(t)|_{\delta(t)} = |Z|$ para cualquier norma de Finsler y luego

$$\text{dist}(U, V) \leq \text{Long}(\delta) = |Z|.$$

El siguiente teorema sobre la caracterización de geodésicas para normas de Finsler Ad-invariantes en $\mathfrak{u}_n$ fue probado en [46, Sección 5]:

Teorema 4.3.3 (Minimalidad y distancia). Sea $\gamma : [a, b] \rightarrow \mathcal{U}_n$ con $\gamma(a) = U$. Luego

- (1) Si $W \in \mathcal{U}_n$ y $Z \in \mathfrak{u}_n$ es tal que $\|Z\|_\infty \leq \pi$ y $W = Ue^Z$. Entonces $\text{dist}(U, W) = |Z|$ y el camino $\delta(t) = Ue^{tZ}$ es un camino que minimiza la distancia.
- (2) Si $\|1 - e^Xe^Y\|_\infty < 2$ luego $\text{dist}(1, e^Xe^Y) = |BCH(X, Y)| = |X + Y + \frac{1}{2}[X, Y] + \dots|$. En particular, si X, Y conmutan, entonces $\text{dist}(e^X, e^Y) = |Y - X|$.
- (3) Si $\gamma(b) = W$ y $\text{Long}(\gamma) = |Z|$, para cada funcional φ que realiza la norma de Z tenemos que $\gamma^{-1}\gamma'$ pertenece a la cara F_φ . En particular si $\cap_{\varphi \in \mathcal{N}_Z} F_\varphi = \{Z\}$, entonces γ es una reparametrización de δ .
- (4) Si γ es como en el ítem previo y la norma es estrictamente convexa, entonces γ es una reparametrización continua de δ .

Proposición 4.3.4 (Equivalencia de métricas). Sea $(\mathcal{U}_n, \text{dist})$, donde la distancia viene inducida por una norma de Finsler Ad-invariante en $\mathfrak{u}_n$. Sean $U, V \in \mathcal{U}_n$ y r, R números positivos como en la Observación 2.1.6. Luego

$$\frac{2}{\pi r} \text{dist}(U, V) \leq \|U - V\|_\infty \leq R \text{dist}(U, V)$$

Además, $(\mathcal{U}_n, \text{dist})$ es un cuasi espacio métrico compacto.

Demostración. Sea $B \in \mathfrak{u}_n$ tal que $\|B\|_\infty \leq \pi$ y $e^B = U^*V$. Luego

$$\|U - V\|_\infty = \|1 - U^*V\|_\infty = \|1 - e^B\|_\infty \leq \|B\|_\infty \leq R|B| = R \text{dist}(U, V),$$

por la Proposición 4.3.1, la Observación 2.1.6 y el teorema previo. Análogamente

$$\|U - V\|_\infty = \|1 - e^B\|_\infty \geq \frac{2}{\pi} \|B\|_\infty \geq \frac{2}{\pi} \frac{1}{r} \text{dist}(U, V).$$

La última afirmación es consecuencia de la compacidad del grupo unitario con la norma espectral y las desigualdades probadas. $\square$

Lema 4.3.5 (Exponencial y geodésicas). Si $\Gamma : [a, b] \rightarrow (\mathfrak{u}_n, |\cdot|)$ es un camino corto que une 0 con Z en $\mathfrak{u}_n$ y $\|Z\|_\infty \leq \pi$, luego $\gamma = e^\Gamma$ es un camino corto en $(\mathcal{U}_n, \text{dist})$ que une a 1 con e^Z en $\mathfrak{u}_n$. Más aún, para cada funcional φ que realiza la norma de Z tenemos que $\gamma_t^{-1}\gamma'_t \subset F_\varphi$.

Demostración. Notamos que

$$\begin{aligned} |\gamma'_t|_{\gamma_t} &= |\gamma_t^{-1} D \exp_{\Gamma_t} \Gamma'_t| = \left| \int_0^1 e^{-s \operatorname{ad}_{\Gamma_t}} \Gamma'_t ds \right| \leq \int_0^1 |e^{-s \operatorname{ad}_{\Gamma_t}} \Gamma'_t| ds = \\ &= \int_0^1 |\operatorname{Ad}_{-\gamma_t}(\Gamma'_t)| ds = \int_0^1 |\Gamma'_t| ds = |\Gamma_b - \Gamma_a| = |Z| \end{aligned}$$

donde utilizamos que la norma es Ad-invariante, la fórmula de la diferencial de la exponencial dada en 4.3.1 y que Γ es corta (ver [45, Lema 4.21]). Esto, combinado con el Teorema 4.3.3, nos muestra que γ es corta. Sea ahora φ un funcional que realiza la norma de Z , y asumamos que Γ está parametrizada a velocidad constante. Con una cuenta similar tenemos que $\varphi(\gamma_t^{-1} \gamma'_t) = \varphi(\Gamma'_t) = |\Gamma'_t| = |Z|$, donde en la segunda desigualdad volvemos a utilizar el Lema 4.21 de [45]. Luego

$$|Z| = \varphi(\gamma_t^{-1} \gamma'_t) \leq |\gamma_t^{-1} \gamma'_t| \leq |Z|$$

entonces $\gamma_t^{-1} \gamma'_t \subset F_\varphi$. Dicha contención es claramente invariante por reparametrizaciones que preservan la orientación, con lo cual la hipótesis de que Γ está parametrizada a velocidad constante no quita generalidad. $\square$

Teorema 4.3.6. Sean $X, Y \in \mathfrak{u}_n$, entonces para toda norma de Finsler Ad-invariante en $\mathfrak{u}_n$ tenemos que

$$\operatorname{dist}(e^X, e^Y) \leq |Y - X|. \quad (4.3)$$

Además, asumiendo la normalización $\|[V, W]\| \leq 2\|V\|_\infty\|W\|$ para la norma de Finsler, si tenemos que $\|X\|_\infty, \|Y\|_\infty \leq \pi/2$ entonces

$$\operatorname{dist}(e^X, e^Y) \geq \frac{2}{\pi} |Y - X|.$$

Demostración. Para la primer desigualdad consideramos la curva $\gamma(t) = e^{tY} e^{-tX}$, que une a 1 con $e^Y e^{-X}$ en $\mathcal{U}_n$. Derivando obtenemos que $\gamma'(t) = e^{tY} (Y - X) e^{-tX}$, utilizando que la norma es Ad-invariante obtenemos que

$$\operatorname{Long}(\gamma) = \int_0^1 |\gamma^{-1}(t) \gamma'(t)| dt = \int_0^1 |e^{tX} (Y - X) e^{-tX}| dt = \int_0^1 |Y - X| dt = |Y - X|,$$

entonces $\operatorname{dist}(e^X, e^Y) = \operatorname{dist}(1, e^Y e^{-X}) \leq L(\gamma) = |Y - X|$.

Supongamos ahora que $\|X\|_\infty, \|Y\|_\infty \leq \pi/2$, por la Proposición 4.3.1 tenemos que $\phi = \log : \mathcal{V} \rightarrow B_\pi^\infty \cap \mathfrak{u}_n$ es una carta de la variedad $\mathcal{U}_n$ alrededor de la identidad. Vamos a considerar $U(t)$ como una curva definida para $t \in (-\epsilon, \epsilon)$ que cumple $U(0) = U$ y su derivada $\dot{U}(t)$ que cumple $\dot{U}(0) = \dot{U}$. No vamos a escribir la dependencia de t por facilidad de notación. Aclarado esto, partiendo de $e^{\phi(U)} = U$ para $U \in \mathcal{V}$ y derivando, obtenemos

$$D \exp_{\phi(U)} D\phi_U \dot{U} = \dot{U}.$$

Si $U = e^Z \in \mathcal{V}$ y $\dot{U} = UW \in T_u \mathcal{U}_n = \mathfrak{u}_n$, utilizando la Observación 4.3.1, tenemos que

$$e^Z F(\operatorname{ad}_Z) D\phi_{e^Z}(\dot{U}) = D \exp_Z D\phi_{e^Z}(\dot{U}) = \dot{U} = e^Z W,$$

y entonces $D\phi_{e^Z}(\dot{U}) = F(\text{ad } Z)^{-1}W$. Sea $0 < M < \pi$, asumamos que $\|Z\|_\infty < M$ y en tal caso notamos que por la normalización de la hipótesis tenemos que $|(\text{ad } Z)^j W| \leq 2^j \|Z\|_\infty |W|$. Luego con la misma demostración que en [6, Proposición 4.6] se sigue que

$$|D\phi_{e^Z}(\dot{U})| = |F(\text{ad } Z)^{-1}W| \leq g(\|Z\|_\infty)|W| < g(M) |\dot{U}|_U$$

para $g(t) = \frac{t}{\sin(t)}$. Por otro lado de $|D \exp_Z \dot{Z}| = |\int_0^1 e^{s \text{ad } Z} \dot{Z} ds| \leq |\dot{Z}|$ obtenemos

$$|\dot{U}|_U = |D \exp_Z D\phi_{e^Z}(\dot{U})| \leq |D\phi_{e^Z}(\dot{U})|.$$

Luego la métrica de Finsler es compatible en el entorno de la identidad dado por $U = e^Z$, $\|Z\|_\infty < M$ (ver [43, Def. 9.1.3] ó [65, Def. 12.19]): para cualquier $\dot{U} \in T_U \mathcal{U}_n$ tenemos que

$$|\dot{U}|_U \leq |D\phi_U(\dot{U})| \leq g(M) |\dot{U}|_U.$$

De donde se sigue que

$$\text{dist}(U, V) \leq |\phi(V) - \phi(U)| \leq g(M) \text{dist}(U, V)$$

para cada $U = e^X, V = e^Y$ con $\|X\|_\infty, \|Y\|_\infty < M/2$ (ver [43, Lema 9.1.8] o la demostración de [65, Proposición 12.22]). Haciendo tender $M \rightarrow \pi/2$ obtenemos $|Y - X| \leq \frac{\pi}{2} \text{dist}(e^X, e^Y)$, lo cual demuestra la segunda desigualdad del enunciado. $\square$

Observación 4.3.7. Para $r, t \in \mathbb{R}$ podemos escribir:

$$\text{dist}(e^{rX}, e^{rY}) = \text{dist}(1, e^{(t-1)rX} e^{rY} e^{-trX}) = \text{dist}(1, e^W e^{-trX}) = \text{dist}(1, e^Z) = |Z|,$$

donde

$$W = BCH((t-1)rX, rY) \quad \text{y} \quad Z = BCH(W, -trX). \quad (4.4)$$

Para un valor de $t \in \mathbb{R}$ fijo, estas expresiones están bien definidas para valores de $r \in \mathbb{R}$ lo suficientemente chicos. Donde escribamos Z estaremos pensando $Z = Z_t(r)$.

Lema 4.3.8. Con la notación de la observación previa, para r lo suficientemente chico, tenemos que:

$$\text{dist}(e^{rX}, e^{rY}) = |r(Y - X) + \frac{r^2}{2}(2t-1)[X, Y] + \frac{r^3}{12}(6t^2 - 6t + 1)[X, [X, Y]] + \frac{r^3}{12}[Y, [X, Y]] + O(r^4)|.$$

En particular si $t = 1$ obtenemos:

$$\text{dist}(e^{rX}, e^{rY}) = |r(Y - X) + \frac{r^2}{2}[X, Y] + \frac{r^3}{12}[X + Y, [X, Y]] + O(r^4)|,$$

y si $t = 1/2$

$$\text{dist}(e^{rX}, e^{rY}) = |r(Y - X) + \frac{1}{48}r^3[X + Y, [X, Y]] + \frac{1}{16}r^3[Y - X, [X, Y]] + O(r^4)|.$$

Demostración. Recordamos que

$$BCH(U, V) = U + V + \frac{1}{2}[U, V] + \frac{1}{12}[U - V, [U, V]] + \text{términos de mayor orden.}$$

Comenzamos calculando W (ver 4.4):

$$\begin{aligned} W &= BCH((t-1)rX, rY) \\ &= (t-1)rX + rY + \frac{1}{2}[(t-1)rX, rY] + \frac{1}{12}[(t-1)rX - rY, [(t-1)rX, rY]] + O(r^4) \\ &= r((t-1)X + Y) + \frac{r^2}{2}(t-1)[X, Y] + \frac{r^3}{12}(t-1)^2[X, [X, Y]] - \frac{r^3}{12}(t-1)[Y, [X, Y]] + O(r^4). \end{aligned}$$

Por otro lado escribimos

$$BCH(W, -rtX) = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A - B, [A, B]] + O(r^4),$$

donde $A = W$ y $B = -rtX$. Vamos a hacer esta cuenta en varios pasos agrupando los términos que aparecen multiplicados por r^4 en una misma expresión de la forma $O(r^4)$. Comenzamos por $A + B$

$$A + B = r(Y - X) + \frac{r^2}{2}(t-1)[X, Y] + \frac{r^3}{12}(t-1)^2[X, [X, Y]] - \frac{r^3}{12}(t-1)[Y, [X, Y]] + O(r^4).$$

Para calcular $[A, B]$, escribimos:

$$\begin{aligned} [A, B] &= [r((t-1)X + Y) + \frac{r^2}{2}(t-1)[X, Y], -rtX] + O(r^4) \\ &= -r^2[(t-1)X + Y, tX] - \frac{r^3}{2}(t-1)[[X, Y], -rtX] + O(r^4) \\ &= tr^2[X, Y] + \frac{r^3}{2}(t^2 - t)[X, [X, Y]] + O(r^4). \end{aligned}$$

Para calcular $[A - B, [A, B]]$, escribimos

$$\begin{aligned} [A - B, [A, B]] &= [rtX - rX + rY + rtX, tr^2[X, Y]] + O(r^4) \\ &= r^3[2tX - X + Y, t[X, Y]] + O(r^4) \\ &= tr^3[2tX - X + Y, [X, Y]] + O(r^4) \\ &= (2t^2 - t)r^3[X, [X, Y]] + tr^3[Y, [X, Y]] + O(r^4). \end{aligned}$$

Finalmente calculamos

$$BCH(W, -rtX) = r(Y - X) + \frac{r^2}{2}(2t - 1)[X, Y] + \alpha r^3[X, [X, Y]] + \beta r^3[Y, [X, Y]],$$

donde

$$\begin{aligned} \alpha &= \frac{1}{12}(t-1)^2 + \frac{1}{4}(t^2 - t) + \frac{1}{12}(2t^2 - t) = \frac{1}{12}(6t^2 - 6t + 1), \\ \beta &= -\frac{1}{12}(t-1) + \frac{1}{12}t = \frac{1}{12}. \end{aligned}$$

En conclusión tenemos que

$$Z = r(Y - X) + \frac{r^2}{2}(2t - 1)[X, Y] + \left(\frac{t^2}{2} - \frac{t}{2} + \frac{1}{12}\right)r^3[X, [X, Y]] + \left(-\frac{t^2}{12} + \frac{t}{4} - \frac{1}{12}\right)r^3[Y, [X, Y]].$$

□

4.4. Distancia y curvatura

En esta sección discutiremos la curvatura en el contexto de $\mathfrak{u}_n$. Como mencionamos al inicio del capítulo y en la sección 4.2, tenemos que en base a Milnor, para un grupo de Lie con una métrica riemanniana bi-invariante vale la siguiente igualdad

$$\langle R(X, Y)Y, X \rangle = 6\|Y - X\|^2 \lim_{r \rightarrow 0^+} \frac{r\|Y - X\| - \text{dist}(e^{rX}, e^{rY})}{r^2 \text{dist}(e^{rX}, e^{rY})},$$

donde $X, Y \in \mathfrak{g}$ el álgebra de Lie de G , la exponencial es la del grupo que coincide en $1 \in G$ con la riemanniana y el producto interno está hecho en $1 \in G$. La idea es usar esta igualdad como motivación para definir la curvatura seccional en el caso de trabajar en el grupo $\mathcal{U}_n$ con una norma de Finsler Ad-invariante en $\mathfrak{u}_n$ y la métrica que induce. Los resultados de esta sección publicados en [48] son el Teorema 4.4.4, el cual nos permite dar una buena definición de curvatura seccional en nuestro contexto y el Teorema 4.4.14, en el cual se estudia que sucede cuando la curvatura seccional es nula.

Observación 4.4.1. Si $X, Y \in \mathfrak{u}_n$ y φ es un funcional que realiza la norma de $Y - X$, por el Lema 3.1.4 tenemos que

$$\begin{aligned} 0 &\geq \varphi([X, [X, Y - X]]) = \varphi([X, [X, Y]]) = \varphi([-X, [-X, Y]]) = \varphi([Y - X - Y, [Y, X]]) \\ &= \varphi([Y - X, [Y, X]]) - \varphi([Y, [Y, X]]) = -\varphi([Y, [Y, X]]), \end{aligned}$$

luego $\varphi([Y, [Y, X]]) \geq 0$ y además $\varphi([X, [X, Y]]) = 0$ si y solo si $\varphi([Y, [Y, X]]) = 0$.

Observación 4.4.2. A partir de la igualdad $\text{dist}(e^{rX}, e^{rY}) = |r(Y - X) + O(r^2)|$ del Lema 4.3.8, dividiendo por $r > 0$ y tomando límite cuando $r \rightarrow 0^+$ tenemos que

$$\lim_{r \rightarrow 0^+} \frac{d(e^{rX}, e^{rY})}{r} = \lim_{r \rightarrow 0^+} |Y - X + O(r)| = |Y - X|.$$

Definición 4.4.3 (Curvatura seccional a lo largo de dos rayos). Sean $X, Y \in \mathfrak{u}_n$ con $X \neq Y$ definimos la curvatura seccional de $\mathcal{U}_n$ en la identidad a lo largo de X e Y como

$$S(X, Y) = 6|Y - X| \lim_{r \rightarrow 0^+} \frac{r|Y - X| - \text{dist}(e^{rX}, e^{rY})}{r^3}.$$

Notamos que, por la observación previa, $S(X, Y)$ es exactamente el límite de la introducción de esta sección.

Teorema 4.4.4. Sean $X, Y \in \mathfrak{u}_n$ con $X \neq Y$, para la función $S(X, Y)$ definida previamente tenemos que

$$\begin{aligned} S(X, Y) &= 6|Y - X| \lim_{r \rightarrow 0^+} \frac{|Y - X| - |(Y - X) + \frac{1}{48}r^2[X + Y, [X, Y]]|}{r^2} \\ &= -\frac{|Y - X|}{4} \max_{\varphi \in N_{Y-X}} \varphi([Y, [Y, Y - X]]) = -\frac{|Y - X|}{4} \max_{\varphi \in N_{Y-X}} \varphi([X, [X, Y - X]]) \geq 0. \end{aligned}$$

Demostración. Notamos que por el Teorema 4.3.6 tenemos que para $r > 0$ suficientemente pequeño $d(e^{rX}, e^{rY}) \leq |rX - rY|$. Aplicando el Lema 4.3.8 para el caso $t = 1/2$ se sigue que:

$$\begin{aligned}
\lim_{r \rightarrow 0^+} \frac{r|Y - X| - d(e^{rX}, e^{rY})}{r^3} &= \lim_{r \rightarrow 0^+} \frac{r|Y - X| - |Z(r)|}{r^3} \\
&= \lim_{r \rightarrow 0^+} \frac{r|Y - X| - |r(Y - X) + \frac{1}{48}r^3[X + Y, [X, Y]] + \frac{1}{16}r^3[Y - X, [X, Y]] + O(r^4)|}{r^3} \\
&= \lim_{r \rightarrow 0^+} \frac{|Y - X| - |(Y - X) + \frac{1}{48}r^2[X + Y, [X, Y]] + \frac{1}{16}r^2[Y - X, [X, Y]] + O(r^3)|}{r^2} \\
&= \lim_{r \rightarrow 0^+} \frac{|Y - X| - |(Y - X) + \frac{1}{48}r^2[X + Y, [X, Y]] + \frac{1}{16}r^2[Y - X, [X, Y]]|}{r^2} \\
&= - \max_{\varphi \in N_{Y-X}} \varphi\left(\frac{1}{48}[X + Y, [X, Y]] + \frac{1}{16}[Y - X, [X, Y]]\right),
\end{aligned}$$

donde en la última igualdad usamos que $r^2 > 0$ y aplicamos 3.2.5. Como φ realiza la norma de $Y - X$ entonces, por el Lema 3.1.4, tenemos que $\varphi \circ \text{ad}_{Y-X} = 0$ y entonces el segundo término que está dentro del máximo desaparece. Se sigue que

$$\begin{aligned}
S(X, Y) &= -\frac{6|Y - X|}{48} \max_{\varphi \in N_{Y-X}} \varphi([2X + Y - X, [X, Y]]) = -\frac{|Y - X|}{8} \max_{\varphi \in N_{Y-X}} \varphi([2X, [X, Y - X]]) \\
&= -\frac{|Y - X|}{4} \max_{\varphi \in N_{Y-X}} \varphi([X, [X, Y - X]]).
\end{aligned}$$

Operando de forma similar obtenemos

$$S(X, Y) = -\frac{|Y - X|}{4} \max_{\varphi \in N_{Y-X}} \varphi([Y, [Y, Y - X]]).$$

Por otro lado

$$\lim_{r \rightarrow 0^+} \frac{|Y - X| - |(Y - X) + \frac{1}{48}r^2[X + Y, [X, Y]]|}{r^2} = - \max_{\varphi \in N_{Y-X}} \varphi\left(\frac{1}{48}[X + Y, [X, Y]]\right) = \frac{1}{6|Y - X|} S(X, Y),$$

donde usamos nuevamente la Observación 3.2.5. Para concluir, notamos que $S(X, Y) \geq 0$ como consecuencia inmediata de la Observación 4.4.1 y la igualdad demostrada. $\square$

Como consecuencia del Teorema 3.1.13 y el Corolario 3.1.14 tenemos que:

Corolario 4.4.5. Si la norma es estrictamente convexa, entonces $[X, Y] = 0$ si y solo si $S(X, Y) = 0$.

Demostración. Si $[X, Y] = 0$ se sigue que $BCH(sX, -sY) = s(Y - X)$ y entonces $\text{dist}(e^{sX}, e^{sY}) = |sX - sY|$, luego el numerador del límite que define a $S(X, Y)$ es idénticamente nulo, con lo cual $S(X, Y) = 0$. Notamos que esto sucede aún si la norma no es estrictamente convexa. Por otro lado si $S(X, Y) = 0$ entonces $\varphi([X, [X, Y]]) = 0$ para la única φ que realiza la norma de $Y - X$ (la unicidad es consecuencia de que la norma es estrictamente convexa, ver Lema 2.1.14). Por la Observación 4.4.1, tenemos que $\varphi([Y - X, [Y - X, X]]) = 0$ y aplicando el Corolario 3.1.14 obtenemos que $[Y - X, X] = 0$, con lo cual $[X, Y] = 0$. $\square$

En conclusión, para normas estrictamente convexas, tenemos que $S(X, Y) = 0$ implica que el plano generado por X, Y induce una distancia plana en $\mathcal{U}_n$, es decir

$$\text{dist}(e^{s_1 X + t_1 Y}, e^{s_2 X + t_2 Y}) = \text{dist}(1, e^{(s_2 - s_1)X + (t_2 - t_1)Y}) = |(s_2 - s_1)X + (t_2 - t_1)Y|,$$

siempre y cuando $\|(s_2 - s_1)X + (t_2 - t_1)Y\|_\infty < \pi$.

Si la norma no es necesariamente estrictamente convexa pero X e Y conmutan, tenemos que la distancia es localmente plana (ver el tercer ítem de 4.3.3). Para redondear el caso estrictamente convexo vamos a citar el siguiente teorema (para una demostración del mismo ver [45, Teorema 4.17]):

Teorema 4.4.6. Si la norma es estrictamente convexa y $\text{dist}(e^X, e^Y) = |Y - X|$ entonces $[X, Y] = 0$.

A continuación vamos a estudiar el caso general, es decir si la norma no es necesariamente estrictamente convexa. Comenzamos con las siguientes observaciones:

Observación 4.4.7. Dados $X, Y \in \mathfrak{u}_n$ vamos a notar $\text{Lie}(X, Y)$ al álgebra de Lie cerrada generada por X, Y , es decir, la subálgebra de Lie cerrada más chica que contiene a los elementos X e Y . Notamos que si Z conmuta con X e Y , entonces va a conmutar con cualquier elemento de $\text{Lie}(X, Y)$, esto es una consecuencia de la linealidad del corchete y la Identidad de Jacobi.

Observación 4.4.8. Si existe una cara de la esfera a la cual pertenecen X e Y simultáneamente, entonces la distancia es plana. Para ver esto, sabemos que existe un funcional φ que realiza la norma de ambos elementos, es decir $\varphi(X) = |X|$ y $\varphi(Y) = |Y|$, si representamos ese funcional como $\varphi = (N|\cdot)$ entonces N conmuta con X e Y . Para X, Y suficientemente chicos en norma escribimos

$$BCH(X + Y, -X) = Y + [X, f] + [Y, g]$$

donde $f, g \in \text{Lie}(X, Y)$. Esto lo podemos hacer fácilmente usando la fórmula de Dynkin [28] y agrupando todos los corchetes que empiezan con X y todos los que empiezan con Y . Se sigue que

$$\begin{aligned} |Y| &= |X + Y - X| \geq \text{dist}(1, e^{X+Y} e^{-X}) = |BCH(X + Y, -X)| \geq \varphi(BCH(X + Y, -X)) \\ &= \varphi(X + Y - X) + \varphi([X, f]) + \varphi([Y, g]) \\ &= |Y| + (N|[X, f]) + (N|[Y, g]) = |Y| + ([N, f]|X) + ([N, g]|Y) = |Y| + 0 = |Y|, \end{aligned}$$

donde usamos que ad_N es antihermitiana con respecto al producto interno de Frobenius y la observación previa. Luego tenemos que

$$\text{dist}(e^X, e^{X+Y}) = \text{dist}(1, e^{X+Y} e^{-X}) = |Y|.$$

Pero tenemos algo más fuerte, si F_φ es la cara de la esfera dada por el funcional unitario φ y C_φ es el cono generado por dicha cara (ver Definición 2.1.11), la misma cuenta nos dice que

$$\text{dist}(e^X, e^{X+Y}) = |Y| \quad \forall X, Y \in C_\varphi$$

siempre y cuando sea $\|X + Y\|_\infty + \|X\|_\infty < \pi$ (Proposición 4.3.1), luego $\exp(C_\varphi) \subset \mathcal{U}_n$ es (localmente) plano. Más aún, el mismo argumento muestra que el álgebra de Lie generada por la cara en cuestión es plana, en el siguiente sentido

$$\text{dist}(e^L, e^{L+Y}) = |Y| \quad \forall Y \in C_\varphi, \quad \forall L \in \text{Lie}(F_\varphi),$$

ya que $\varphi \circ \text{ad}_L = 0$. Nuevamente, asumiendo que $\|L + Y\|_\infty + \|L\|_\infty < \pi$. En tal caso la igualdad vale si L es lo suficientemente pequeño y $\|Y\|_\infty < \pi$.

Haciendo un cambio de variables de lo último observado, obtenemos el siguiente corolario:

Corolario 4.4.9. Si $V, W \in \mathfrak{u}_n$ son tales que $W - V \in C_\varphi$ y $\|L + V\|_\infty + \|L + W\|_\infty < \pi$ entonces

$$\text{dist}(e^{L+V}, e^{L+W}) = |W - V| \quad \forall L \in \text{Lie}(F_\varphi).$$

La esfera de una norma puede tener caras de distintas dimensiones. Vamos a estudiar ciertas propiedades de las caras de dimensión máxima, es decir de las que tienen dimensión como variedad con borde igual a $n^2 - 1$ (recordar que $\dim(\mathfrak{u}_n) = n^2$).

Definición 4.4.10 (Interior de una cara de dimensión máxima). Sean $|\cdot|$ una norma de Finsler en $\mathfrak{u}_n$ y φ un funcional de norma 1. Si $F_\varphi(r)$ es una cara de dimensión máxima, decimos que $V \in F_\varphi(r)^\circ$ si existe $\epsilon > 0$ de forma tal que

$$B_\epsilon(V) \cap \{X \in \mathfrak{u}_n : \varphi(X) = r\} \subseteq F_\varphi(r),$$

es decir estamos definiendo su interior como el interior topológico pensando la cara dentro del hiperplano $\{X \in \mathfrak{u}_n : \varphi(X) = r\}$. Notamos que, como la norma no es necesariamente homogénea, debemos ser específicos en que significa la bola centrada en V de radio epsilon, esto será:

$$B_\epsilon(V) = \{X \in \mathfrak{u}_n : |X - V| < \epsilon\}.$$

Lema 4.4.11. Sean $|\cdot|$ una norma de Finsler en $\mathfrak{u}_n$ y φ un funcional de norma 1. Si $F_\varphi(r)$ es una cara de dimensión máxima, son equivalentes:

- (1) $V \in F_\varphi(r)^\circ$.
- (2) Existe $\epsilon > 0$ de forma tal que si $Z \in \ker(\varphi)$ y $|Z| < \epsilon$ entonces $V + Z \in F_\varphi(r)$.

Demostración. Supongamos primero que vale (1), luego existe $\epsilon > 0$ de forma tal que

$$B_\epsilon(V) \cap \{X \in \mathfrak{u}_n : \varphi(X) = r\} \subseteq F_\varphi(r),$$

si $Z \in \ker(\varphi)$ y $|Z| < \epsilon$ entonces $|V + Z - V| = |Z| < \epsilon$ y $\varphi(V + Z) = \varphi(V) + \varphi(Z) = r$, lo cual me dice que $V + Z \in B_\epsilon(V) \cap \{X \in \mathfrak{u}_n : \varphi(X) = r\}$ y entonces $V + Z \in F_\varphi(r)$. Por otro lado si partimos de que vale (2), consideramos un elemento arbitrario en $B_\epsilon(V) \cap \{X \in \mathfrak{u}_n : \varphi(X) = r\}$ al que llamaremos $V + W$, luego tendremos por un lado que $|W| < \epsilon$ y por otro lado, usando que $\varphi(V) = r$, que $\varphi(Z) = 0$. Se sigue entonces que $V + Z \in F_\varphi(r)$ y resulta $V \in F_\varphi(r)^\circ$. $\square$

Lema 4.4.12. Sea $|\cdot|$ una norma de Finsler en $\mathfrak{u}_n$, si $V \in \mathfrak{u}_n$ pertenece al interior de una cara de dimensión máxima de la bola entonces existe un único funcional que realiza su norma.

Demostración. Sea φ un funcional de norma 1 de forma tal que $F_\varphi(r)$ es de dimensión máxima y $V \in F_\varphi(r)^\circ$. Por el lema previo sabemos que existe $\epsilon > 0$ de forma tal que si $|Z| < \epsilon$ y $Z \in \ker(\varphi)$ entonces $\varphi(V + Z) = |V|$. Luego, si ψ es otro funcional que realiza la norma de V , tenemos que

$$|V| + \psi(Z) = \psi(V) + \psi(Z) = \psi(Z + V) \leq |V + Z| = \varphi(V + Z) = |V|,$$

de donde se sigue que $\psi(Z) \leq 0$ para todo $Z \in \ker(\varphi)$ con $|Z| < \epsilon$. Si tomamos un elemento arbitrario $W \in \ker(\varphi)$, sabemos que existe $c > 0$ de forma tal que $|cW| < \epsilon$ y luego $\psi(cW) \leq 0$, lo cual implica que $\psi(W) \leq 0$ para todo $W \in \ker(\varphi)$. Esto último solo puede suceder si $\psi(W) = 0$ para todo $W \in \ker(\varphi)$, entonces los funcionales tienen el mismo núcleo y luego $\varphi = \psi$. $\square$

Proposición 4.4.13. Si F_φ es una cara de dimensión máxima en $\mathfrak{u}_n$, es decir de dimensión $n^2 - 1$, luego para cualquier V en el interior de dicha cara con $\|V\|_\infty < \pi$ y $Z \in \ker \varphi$ suficientemente pequeño, tenemos que $s^{-1}BCH(Z + sV, -Z) \in F_\varphi$ y

$$\text{dist}(e^Z, e^{Z+sV}) = s|V| = s|Z + V|,$$

para cualquier $s \in [0, 1]$. Cuando decimos Z suficientemente pequeño, nos estamos refiriendo a que $Z + V \in F_\varphi$ y además $\|Z\|_\infty < \frac{\pi - \|V\|_\infty}{2}$. Luego reemplazando Z por sZ tenemos que

$$s|V| = \text{dist}(e^{sZ}, e^{s(Z+V)})$$

Demostración. Sea V un elemento en el interior de la cara F_φ , si $Z \in \ker \varphi$ es suficientemente pequeño tenemos que $V + Z \in F_\varphi$ (ver Lema 4.4.11) y en tal caso $|V + Z| = \varphi(V + Z) = |V|$. Aplicando el Lema 3.1.4 tenemos que $\varphi \circ \text{ad}_V = 0$ y $\varphi \circ \text{ad}_V + Z = 0$, luego se sigue que $\varphi \circ \text{ad}_Z = 0$ entonces $\varphi([Z, [Z, V]]) = 0$ y como $\varphi = (N|\cdot|)$ es el único funcional que realiza la norma de V (ver Lema 4.4.12), podemos aplicar el Teorema 3.1.18 y obtenemos que $[Z, N] = 0$ (recordamos además que $[V, N] = 0$ por el Lema 3.1.4). Notamos que para $s \in [0, 1]$ tenemos que

$$\|Z + sV\|_\infty + \|Z\|_\infty \leq 2\|Z\|_\infty + \|V\|_\infty < \pi,$$

luego por la Proposición 4.3.1, la serie BCH de $e^{Z+sV}e^{-Z}$ es convergente. Esta serie comienza con sV y por la Fórmula de Dynkin (ver [28]) el resto de la serie está formada por combinaciones lineales de conmutadores que comienzan con Z o con $Z + sV$, por lo probado previamente, deben pertenecer al núcleo de φ . Luego para $s \in [0, 1]$ tenemos que

$$s|V| \geq |BCH(Z + sV, -Z)| \geq \varphi(BCH(Z + sV, -Z)) = \varphi(sV) + 0 = s|V|,$$

de donde se deduce que

$$\text{dist}(e^Z, e^{Z+sV}) = BCH(Z + sV, -Z) = s|V| = s|Z + V|$$

y también que $s^{-1}BCH(Z + sV, -Z) \in F_\varphi$. $\square$

Teorema 4.4.14 (Secciones planas). Dados $X, Y \in \mathfrak{u}_n$ y $|\cdot|$ una norma de Finsler Ad-invariante, consideramos las siguientes afirmaciones:

- (1) $\varphi([X, [X, Y]]) = 0$ para todo funcional φ que realiza la norma de $Y - X$.
- (2) Para $s > 0$ suficientemente pequeño, el camino $s^{-1}BCH(sY, -sX)$ cae dentro de una misma cara de la esfera de radio $r = |Y - X|$ que contiene a $Y - X$.
- (3) Para $s \geq 0$ suficientemente pequeño, $\text{dist}(e^{sX}, e^{sY}) = s|Y - X|$.

$$(4) S(X, Y) = 0.$$

$$(5) \varphi([X, [X, Y]]) = 0 \text{ para algun funcional } \varphi \text{ que realiza la norma de } Y - X.$$

En tal caso, las afirmaciones (1), (2) y (3) son equivalentes e implican (4) y (5), las cuales son equivalentes entre si.

Además, tenemos que:

- a) Si existe un único funcional que realiza la norma de $Y - X$ (por ejemplo si la norma es suave), entonces todas las condiciones son equivalentes.
- b) Si en (5) tenemos que $F_\varphi = \{Y - X\}$ (por ejemplo en el caso de que la norma es estrictamente convexa), luego todas las condiciones son equivalentes a que $[X, Y] = 0$.

Demostración. (1) $\Rightarrow$ (2) : supongamos que vale (1), es decir $\varphi([X, [X, Y]]) = 0$ para todo funcional φ que realiza la norma de $Y - X$. Elegimos $\varphi_0 = (N_0|\cdot)$ como en el Teorema 3.1.18, por la hipótesis del enunciado y el teorema mencionado previamente tendremos que $[X, N_0] = 0$. Además, dado que φ_0 realiza la norma de $Y - X$, se sigue que $[Y - X, N_0] = 0$ (ver Lema 3.1.4), lo cual implica que también debe ser $[Y, N_0] = 0$ y en tal caso N_0 conmuta con cualquier elemento de $Lie(X, Y)$ (Observación 4.4.7). Desarrollando la serie de $BCH(sY, -sX)$ podemos escribir

$$BCH(sY, -sX) = s(Y - X) + \sum_k [W_k, Z_k]$$

para ciertos $W_k, Z_k \in Lie(X, Y) = Lie(X, Y - X)$. Aplicando el Teorema 4.3.6, se sigue que

$$\begin{aligned} s|Y - X| &= |sY - sX| \geq |BCH(sY, -sX)| \geq \varphi_0(BCH(sY, -sX)) = s|Y - X| + \sum_k (N_0|[W_k, Z_k]) \\ &= s|Y - X| + \sum_k ([N_0, W_k]|Z_k) = s|Y - X| + 0 = s|Y - X|. \end{aligned}$$

Para s suficientemente pequeño concluimos que

$$s\varphi_0(Y - X) = s|Y - X| = |BCH(sY, -sX)| = \varphi_0(BCH(sY, -sX)),$$

dividiendo por s es claro que $s^{-1}(BCH(sY, -sX))$ pertenece a la cara F_{φ_0} de la esfera de radio $|Y - X|$.

(2) $\Rightarrow$ (3): si asumimos (2), tendremos que existe un funcional φ que realiza la norma de $Y - X$ para el cual se verifica que

$$\varphi(s^{-1}BCH(sY, -sX)) = |s^{-1}BCH(sY, -sX)| = |Y - X| = \varphi(Y - X).$$

En particular tenemos que $\varphi(BCH(sY, -sX)) = s|Y - X|$, pero como $\varphi(BCH(sY, -sX)) \leq |BCH(sY, -sX)| = \text{dist}(e^{sX}, e^{sY})$, se sigue la desigualdad $s|Y - X| \leq \text{dist}(e^{sX}, e^{sY})$, como la otra desigualdad vale siempre por el Teorema 4.3.6, concluimos que vale (3).

(3) $\Rightarrow$ (2): consideramos la curva $\gamma(s) = e^{sY}e^{-sX} = e^{BCH(sY, -sX)}$ para $s \in [0, s_0]$, donde este es el intervalo en el cual vale (3). Notamos que γ es una curva que une a 1 con e^Z donde $Z = BCH(s_0Y, -s_0X)$ y su longitud es:

$$\begin{aligned} \text{Long}(\gamma) &= \int_0^{s_0} |\dot{\gamma}(s)|_{\gamma(s)} ds = \int_0^{s_0} |\gamma^{-1}(s) \dot{\gamma}(s)| ds = \int_0^{s_0} |e^{sX} e^{-sY} (e^{sY} Y e^{-sX} - e^{sY} X e^{-sX})| ds \\ &= \int_0^{s_0} |e^{sX} (Y - X) e^{-sX}| ds = \int_0^{s_0} |\text{Ad}_{e^{sX}}(Y - X)| ds = \int_0^{s_0} |Y - X| ds = s_0 |Y - X|. \end{aligned}$$

Luego, usando la hipótesis de (3) y lo visto previamente tenemos que

$$|Z| = |BCH(s_0 Y, -s_0 X)| = \text{dist}(e^{s_0 X}, e^{s_0 Y}) = s_0 |Y - X| = \text{Long}(\gamma),$$

de donde se sigue que γ es una geodésica uniendo 1 y e^Z .

Si consideramos $\Gamma(s) = BCH(sY, -sX)$, tenemos que $e^{\Gamma(s)} = \gamma(s)$ y en tal caso, por el Teorema 4.22 de [45], existe φ un funcional unitario que cumple

$$\varphi(BCH(sY, -sX)) = |BCH(sY, -sX)| \quad \forall s \in [0, s_0].$$

Utilizando nuevamente la hipótesis de (3) tenemos que

$$\varphi(BCH(sY, -sX)) = |BCH(sY, -sX)| = \text{dist}(e^{sX}, e^{sY}) = s|Y - X|$$

para $s \in [0, s_0]$, lo cual prueba (2).

(3) $\Rightarrow$ (1): considerando la misma curva que en (3) $\Rightarrow$ (2) y aplicando el Teorema 4.22 de [45] tenemos la igualdad

$$\varphi(BCH(sY, -sX)) = |BCH(sY, -sX)| = \text{dist}(e^{sX}, e^{sY}) = s|Y - X|$$

para cualquier funcional φ que realiza la norma de $Y - X$. Derivando en ambos lados lateralmente tres veces y evaluando en $s = 0$ la igualdad $\varphi(BCH(sY, -sX)) = s|Y - X|$ obtenemos que:

$$\begin{aligned} 0 &= \frac{1}{2} (\varphi([X, [X, Y]]) - \varphi([Y, [Y, X]])) = \frac{1}{2} (\varphi([X, [X, Y]]) - \varphi([Y - X, [Y, X]]) - \varphi([X, [Y, X]])) \\ &= \frac{1}{2} (\varphi([X, [X, Y]]) + \varphi([X, [X, Y]])) = \varphi([X, [X, Y]]), \end{aligned}$$

donde usamos el desarrollo en serie dado en la Proposición 4.3.1 y el Lema 3.1.4. Se concluye entonces que $\varphi([X, [X, Y]]) = 0$ para todo funcional que realiza la norma de $Y - X$.

(3) $\Rightarrow$ (4): asumiendo (3) tendremos, mirando el numerador del límite de la definición de $S(X, Y)$, que debe ser $S(X, Y) = 0$.

(4) $\Leftrightarrow$ (5): por el Teorema 4.4.4, tenemos que $S(X, Y) = 0$ si y solo si $\min_{\varphi \in N_{Y-X}} \varphi([X, [X, Y]]) = 0$ y este mismo teorema nos dice que este mínimo es mayor o igual a cero. Se sigue entonces que $S(X, Y) = 0$ si y solo si existe un funcional que realiza la norma de $Y - X$ para el cual $\varphi([X, [X, Y]]) = 0$.

Además, si existe un único funcional que realiza la norma de $Y - X$ entonces (1) y (5) son equivalentes, por lo cual todas las afirmaciones lo son. Si por otra parte tenemos $F_\varphi = \{Y - X\}$ y vale (5) aplicando el Teorema 3.1.13 obtenemos que X, Y conmutan y si X, Y conmutan claramente vale (5). $\square$

A continuación vamos a dar un ejemplo que muestra que las afirmaciones (1), (2) y (3) del Teorema previo no son equivalentes a las afirmaciones (4) y (5).

Ejemplo 4.4.15. Consideramos $\mathfrak{u}_3$ con la norma espectral y los elementos dados por

$$X = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} i & 1 & 0 \\ -1 & i & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Notamos que $Y - X = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y en tal caso los funcionales dados por $\varphi = (Z|\cdot)$ y $\varphi_0 = (Z_0|\cdot)$ donde

$$Z = i \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad Z_0 = i \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

cumplen $\varphi(Y - X) = 1$ y $\varphi_0(Y - X) = 1$. Como $\|Y - X\|_\infty = 1$ y $\|Z\|_1 = \|Z_0\|_1 = 1$ ambos funcionales realizan la norma de $Y - X$. Calculando $[X, [X, Y]]$ obtenemos:

$$[X, [X, Y]] = \begin{pmatrix} 0 & 0 & -i \\ 0 & -2i & 0 \\ -i & 0 & 2i \end{pmatrix},$$

evaluando es fácil de ver que $\varphi([X, [X, Y]]) = 0$ mientras que $\varphi_0([X, [X, Y]]) = -1$.

La construcción del ejemplo fue pensada tomando a $Y - X$ como un elemento de la bola unitaria de la norma espectral que no es un vértice ni está en una cara maximal, ya que en esos casos tendríamos que todas las afirmaciones del Teorema 4.4.14 son equivalentes. El funcional φ_0 es justamente uno de los funcionales adaptados que se describen en el Teorema 3.1.18.

En el ejemplo previo se debería poder observar que no existe $s_0 > 0$ de forma tal que valga la igualdad $\text{dist}(e^{sX}, e^{sY}) = s|Y - X|$ para $s \in [0, s_0]$. Los desarrollos en serie involucrados hacen de esto algo difícil de calcular, el siguiente gráfico muestra con un nivel de precisión bastante grande como se comporta la diferencia entre ambos términos para el caso del ejemplo previo.

Los cálculos para la Figura 4.4 fueron realizados en Sage. Para aproximar la serie de Baker se usaron 500 términos de la misma y se evaluó en 1000 valores de $s \in [0, \frac{\pi}{20|X|+20|Y|}]$.

4.5. Curvatura seccional

En esta sección vamos a discutir nuestra noción de curvatura asociada al límite $S(X, Y)$. Los resultados originales de esta sección (publicados en [48]) son el Teorema 4.5.4 y el Teorema 4.5.8.

Observación 4.5.1 (El caso riemanniano). Si $|\cdot| = |\cdot|_F$ es la norma de Frobenius, la cual es esencialmente la única norma Ad-invariante en $\mathfrak{u}_n$ que proviene de un producto interno (ver Proposición 2.3.10), tenemos que dados $X, Y \in \mathfrak{u}_n$ existe un único funcional que realiza la norma de $Y - X$. Este viene dado por

$$\varphi(V) = \frac{(V|Y - X)}{\|Y - X\|_F} \quad \text{para } V \in \mathfrak{u}_n,$$

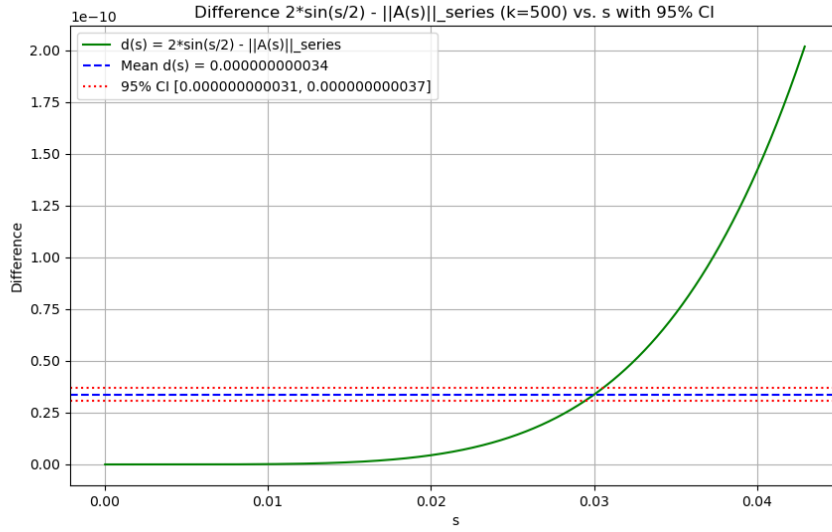


Figura 4.1: Diferencia entre $s|y - x|$ y $\text{dist}(e^{sX}, e^{sy})$.

luego aplicando el Teorema 4.4.4 tenemos que

$$\begin{aligned} S(X, Y) &= -\frac{\|Y - X\|_F}{4} \varphi([X, [X, Y]]) = -\frac{1}{4}([X, [X, Y]]|Y - X) \\ &= \frac{1}{4}([X, Y]|[X, Y - X]) = \frac{1}{4}([X, Y]|[X, Y]) = \frac{1}{4}\|[X, Y]\|_F^2, \end{aligned}$$

en conclusión

$$S(X, Y) = \frac{1}{4}\|[X, Y]\|_F^2.$$

En general, cuando pensamos en grupos de Lie con métricas inducidas por normas Ad-invariantes, como puede ser el caso riemanniano con la norma de Frobenius o cualquier otro caso, la curvatura seccional se puede obtener de la siguiente forma: si R es el tensor de curvatura asociado a la conexión de Levi-Civita de la métrica, $X, Y \in \mathfrak{u}_n$ y A es el área del paralelogramo formado por X e Y entonces

$$\sec(X, Y) = \frac{(R(X, Y)Y|X)}{A^2(X, Y)} = \frac{S(X, Y)}{A^2(X, Y)}.$$

El numerador de esta expresión lo caracterizamos en el Teorema 4.4.4 y al comienzo de esta sección explicitamos como queda en el caso de la métrica dada por la norma de Frobenius. El denominador es más complicado, ya que no tenemos una noción de área en el caso de que la norma no provenga de un producto interno, de esta forma la curvatura seccional asociada al plano $\pi = \text{gen}\{X, Y\}$ va a depender de los generadores del plano en cuestión. Para eliminar dicha dependencia vamos a proponer la siguiente definición:

Definición 4.5.2 (Curvatura seccional asociada a un plano $\pi \subset \mathfrak{u}_n$). Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{u}_n$, sea π un plano en $\mathfrak{u}_n$, definimos

$$\sec(\pi) = \max_{X, Y \in \pi, |X|=|Y|=1} S(X, X+Y) = 1/4 \max_{|X|=|Y|=1} \min_{\varphi \in N_Y} \varphi([X, [Y, X]]),$$

donde la segunda igualdad es consecuencia del Teorema 4.4.4. Notamos que por dicho teorema, este número es no negativo.

Observación 4.5.3. Si X, Y son vectores de norma 1, es fácil ver en el caso riemanniano que $S(X, X + Y) = S(X, Y)$ y $A^2(X, X + Y) = A^2(X, Y)$. Además mirar el máximo sobre todos los pares de vectores de norma 1 coincide con el máximo sobre los que además forman un conjunto ortogonal. Esto hace evidente la fórmula

$$\sec(\pi) = \max_{X, Y \in \pi, |X|=|Y|=1} S(X, X + Y)$$

para el caso riemanniano. Utilizamos $S(X, X + Y)$ en vez de $S(X, Y)$, simplemente por una conveniencia de notación.

Teorema 4.5.4. Dado un plano $\pi \subset \mathfrak{u}_n$ tenemos que

- (1) $\sec(\pi) = 0$ si y solo si $S(X, Y) = 0$ para todo $X, Y \in \pi$.
- (2) Con la normalización $|[X, Y]| \leq 2|X||Y|$ tenemos que $0 \leq \sec(\pi) \leq 1$.

Demostración. Si $S(X, Y) = 0$ para todo $X, Y \in \pi$, por la misma definición de la curvatura seccional se sigue de forma inmediata que $\sec(\pi) = 0$. Por otro lado, si $\sec(\pi) = 0$, entonces

$$\max_{|X|=|Y|=1} \min_{\varphi \in N_Y} \varphi([X, [Y, X]]) = 0.$$

Como $\varphi([X, [Y, X]]) \geq 0$ para todo $\varphi \in N_Y$ (ver Lema 3.1.4), luego tiene que ser $\min_{\varphi \in N_Y} \varphi([X, [Y, X]]) = 0$ para todo $|X| = |Y| = 1$. Sean $X, Y \neq 0$ in π , afirmamos que hay un funcional que realiza la norma de Y y verifica $\varphi([X, [X, Y]]) = 0$, si fuese $|X| = |Y| = 1$ es claro por la hipótesis de $S(X, Y) = 0$. Para el caso general, basta considerar φ que cumpla $\varphi([X_0, [X_0, Y_0]]) = 0$ donde $X_0 = \frac{X}{|X|}$ e $Y_0 = \frac{Y}{|Y|}$ y es fácil ver, usando la linealidad del corchete en cada coordenada y la linealidad de φ que el mismo funcional funciona para X e Y .

Aplicando el Teorema 4.4.4, tenemos que

$$S(X, X + Y) = -\frac{|Y|}{4} \max_{\varphi \in N_Y} \varphi([X, [X, X + Y]]) = \frac{|Y|}{4} \min_{\varphi \in N_Y} \varphi([X, [Y, X]]) = 0.$$

Como X e Y son arbitrarios, llamando $U = X$ y $V = X + Y$, tenemos que $S(U, V) = 0$ que era lo que queríamos probar. Para probar la segunda afirmación, por un lado ya vimos que $\sec(\pi) \geq 0$ y por otro lado con la normalización elegida tenemos que

$$\begin{aligned} \sec(\pi) &= \frac{1}{4} \max_{|U|=|V|=1} \min_{\varphi \in N_V} \varphi([U, [V, U]]) \leq \frac{1}{4} \max_{|U|=|V|=1} |[U, [V, U]]| \\ &\leq \frac{1}{4} \max_{|U|=|V|=1} 4|U||V||U| = 1. \end{aligned} \quad \square$$

Observación 4.5.5. La normalización $|[X, Y]| \leq 2|X||Y|$ se puede obtener simplemente utilizando la continuidad de la forma bilineal dada por el corchete y reescalando la norma.

Para concluir este capítulo daremos algunas observaciones y consecuencias de lo probado en el caso riemanniano utilizando la norma de Frobenius. En este caso se cumple con la normalización del segundo ítem del Teorema 4.5.4 como veremos en el siguiente lema:

Lema 4.5.6. Dados $X, Y \in \mathfrak{u}_n$ entonces

- (1) $\|X^2\|_F \leq \|X\|_F^2$.
- (2) $\|XY\|_F \leq \|X\|_F \|Y\|_F$
- (3) $\|[X, Y]\|_F \leq 2\|X\|_F \|Y\|_F$

Demostración. Si $X = 0$ es evidente el primer ítem es evidente, caso contrario, sea $\lambda_1, \lambda_2, \dots, \lambda_n$ su tira de autovalores. Entonces tenemos que

$$\|X\|_F^2 = \text{Tr}(XX^*) = \text{Tr}(X^*X) = \text{Tr}(|X|^2) = \sum_{i=1}^n |\lambda_i|^2,$$

por otro lado

$$\|X^2\|_F^2 = \text{Tr}(X^2(X^2)^*) = \text{Tr}(X^4) = \sum_{i=1}^n \lambda_i^4 \leq \left(\sum_{i=1}^n |\lambda_i|^2 \right)^2 = \|X\|_F^4,$$

de donde se sigue que $\|X^2\|_F \leq \|X\|_F^2$. Para ver el segundo ítem calculamos

$$\|XY\|_F^2 = \text{Tr}(XY(XY)^*) = \text{Tr}(X^2Y^2) = (X^2|Y^2) \leq \|X^2\|_F \|Y^2\|_F \leq \|X\|_F^2 \|Y\|_F^2,$$

donde utilizamos la desigualdad de Cauchy-Schwarz y el ítem previo. Tomando raíz de ambos lados se concluye (2). Finalmente para (3)

$$\|[X, Y]\|_F = \|XY - YX\|_F \leq \|XY\|_F + \|YX\|_F \leq \|X\|_F \|Y\|_F + \|Y\|_F \|X\|_F = 2\|X\|_F \|Y\|_F.$$

□

Observación 4.5.7. Por el Lema previo y el Teorema 4.5.4, tenemos que para el caso de $\mathfrak{u}_n$ con la norma de Frobenius vale que $0 \leq \sec(\pi) \leq 1$ para cualquier plano π .

Antes de enunciar el siguiente teorema vamos a recordar la construcción de triángulos de comparación, en vista de la observación previa vamos a hacer los triángulos de comparación en la esfera unitaria de la dimensión de $\mathcal{U}_n$, que recordamos es n^2 . Vamos a notar dicha esfera como $S = S^{n^2} \subseteq \mathbb{R}^{n^2+1}$, los vértices del triángulo en $\mathcal{U}_n$ los vamos a notar con las letras a, b, c , mientras que los vértices del triángulo de comparación los notaremos a', b', c' . Asumiendo que hay un único segmento geodésico uniendo a y b , notaremos a este segmento como ab y a su longitud como $|ab|$. Si para $a, b, c \in \mathcal{U}_n$ existen únicos segmentos geodésicos que unen a cada par de dichos puntos, podemos hablar del triángulo geodésico que es el conjunto formado por dichos segmentos y lo notaremos $\Delta(a, b, c) \subseteq \mathcal{U}_n$. Bajo las mismas condiciones, notaremos como $\Delta(a', b', c') \subseteq S$ a un triángulo geodésico en S . En $\mathcal{U}_n$ vamos a tomar la métrica riemanniana inducida por la norma de Frobenius, mientras que en S la métrica riemanniana usual de la esfera.

A continuación vamos a describir dos formas distintas de armar un triángulo de comparación:

- Construcción 1: para esta construcción nos basaremos en el libro de Burago [21] más específicamente en la sección 6.5.2. Como $\mathcal{U}_n$ es un grupo de Lie, mediante una isometría, podemos suponer que $a = 1$. Por otro lado, dado que la esfera es una variedad de curvatura constante, podemos pensar en armar el

triángulo de comparación con $a' = N$ (el polo norte de la esfera). Consideramos la exponencial del grupo de Lie $\mathfrak{u}_n$

$$\exp_{\mathfrak{u}_n} : B_\pi^\infty(\mathfrak{u}_n) \longrightarrow \mathcal{V} = \{U \in \mathcal{U}_n : \|U - 1\|_\infty < 2\},$$

sabemos que es un difeomorfismo por la Proposición 4.3.1 y además coincide con la exponencial riemanniana por la Observación 4.2.10 y los comentarios desarrollados a continuación de la misma. Por otro lado si identificamos el espacio tangente a la esfera con $\mathbb{R}^{n^2}$, tenemos la exponencial riemanniana de la esfera en 0 que es un difeomorfismo con la esfera sin el polo sur, el cual notamos como P , pensada de la siguiente forma:

$$\exp_S := \exp_S|_0 : B_\pi^\infty(\mathbb{R}^{n^2}) \longrightarrow S - \{P\}.$$

Sea ahora una isometría I entre $\mathfrak{u}_n$ y $T_N(S) \simeq \mathbb{R}^{n^2}$, la podemos pensar restringida y obtener la isometría $I : B_\pi^\infty(\mathfrak{u}_n) \longrightarrow B_\pi^\infty(\mathbb{R}^{n^2})$. Consideramos $\sigma := \exp_S \circ I \circ \exp_{\mathfrak{u}_n}^{-1}$ y notamos que σ transforma geodésicas por $1 \in \mathcal{U}_n$ en geodésicas por $N \in S$, además por la misma construcción de σ tenemos que $d_1\sigma$ es una isometría y luego σ preserva ángulos de vectores salientes a partir de 1. En tal caso si tenemos un triángulo geodésico $\Delta(a, b, c) \subseteq \mathcal{V} \subseteq \mathcal{U}_n$ y llamamos $b' = \sigma(b)$ y $c' = \sigma(c)$, obtendremos tres puntos en la esfera a' , b' , y c' , como partimos de un triángulo geodésico y σ preserva ángulos entonces los vértices b' y c' se tienen que poder unir con una única geodésica minimizante de $S - \{P\}$ (en realidad si el ángulo a partir de a de los vectores es π y tanto b' como c' pertenecen al ecuador hay dos geodésicas minimizantes, pero los triángulos formados en este caso son isométricos), formando así un triángulo geodésico $\Delta(a', b', c')$ en la esfera. Por construcción se va a cumplir que $|ab| = |a'b'|$, $|ac| = |a'c'|$ y el ángulo formado por ab y ac es el mismo que el formado por $a'b'$ y $a'c'$, lo que nos interesa comparar es $|bc|$ y $|b'c'|$.

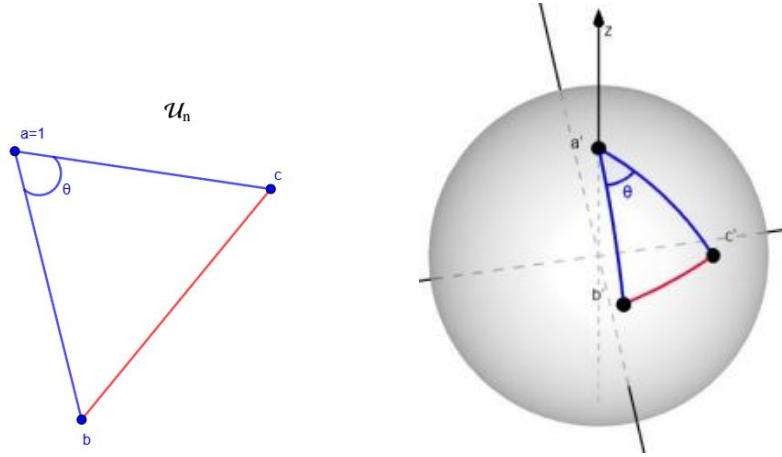


Figura 4.2: Construcción 1

En la Figura 4.2 podemos ver en azul los lados y el ángulo que son preservados en la construcción, mientras que en rojo el lado que queremos comparar.

- Construcción 2: esta construcción es similar a la primera, salvo que pediremos que los tres lados sean iguales y lo que haremos es comparar los ángulos. Para poder hacer esto nos basaremos nuevamente

en el libro de Burango [21], más específicamente en el capítulo 9.1. Para poder armar el triángulo de comparación en este caso necesitamos que $|ab| + |ac| + |bc| < 2\pi$, de esta forma va a existir dicho triángulo y será único salvo isometrías.

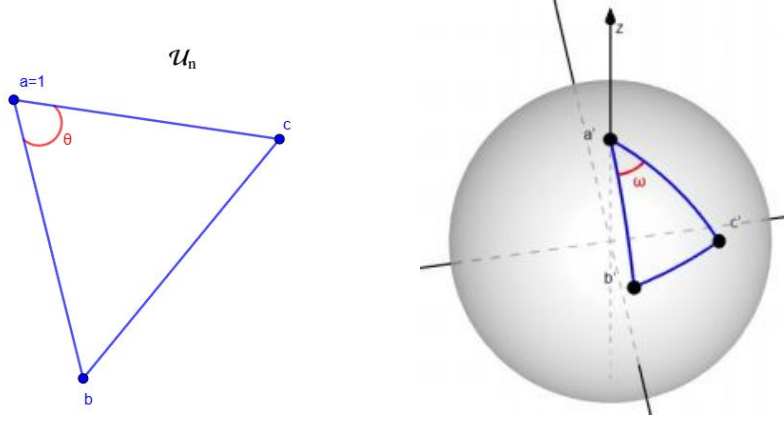


Figura 4.3: Construcción 2

En la figura 4.3, vemos nuevamente en azul lo preservado en la construcción y en rojo lo que queremos comparar (en este caso el ángulo).

Teorema 4.5.8. Sea $\Delta(a, b, c)$ un triángulo geodésico en $\mathcal{U}_n$.

- (1) Si $\Delta(a, b, c) \subseteq \mathcal{V} = \{U \in \mathcal{U}_n : \|U - 1\|_\infty < 2\}$ y realizamos el triángulo de comparación a partir de la Construcción 1 dada previamente entonces $\text{dist}(b, c) \leq \text{dist}(b', c')$.
- (2) Si $\Delta(a, b, c)$ es tal que $\text{dist}(a, b) + \text{dist}(a, c) + \text{dist}(b, c) < 2\pi$ el ángulo del triángulo $\Delta(a, b, c)$ al que en la Figura 4.3 llamamos θ es menor o igual que el ángulo del triángulo de comparación, al cual llamamos ω en la figura mencionada.

Demostración. (1) Usando que la curvatura está acotada superiormente por 1 y el Teorema de Comparación de Cartan-Alexandrov-Toponogov [21, Teorema 6.5.6], al realizar la primer construcción se deduce el resultado, es decir que $\text{dist}(b, c) \leq \text{dist}(b', c')$.

- (2) Para el triángulo geodésico del enunciado vamos a armar los dos triángulos de comparación de las construcciones previas. Llamaremos $\Delta_1(a_1, b'_1, c'_1)$ al realizado con el primer método de comparación y $\Delta_2(a_2, b_2, c_2)$ al del segundo método. Notamos que podemos tomar $a_1 = a_2 = N$, el polo norte de la esfera y tenemos por un lado que en Δ_1 el ángulo formado por los segmentos geodésicos a_1b_1 y a_1c_1 es exactamente θ por la misma construcción, mientras que en Δ_2 el ángulo formado por los segmentos a_2b_2 y a_2c_2 es ω . Tenemos en tal caso que comparar dos triángulos en la esfera que poseen dos lados de la misma longitud y un tercer lado que es menor o igual en el caso de Δ_2 , ya que $|b_2c_2| = |bc| = \text{dist}(b, c) \leq \text{dist}(b_1, c_1) = |b_1c_1|$, por el primer ítem de este teorema. Como Δ_2 tiene dos de los lados iguales a Δ_1 y el lado opuesto al lado que nos interesa mas chico, entonces el ángulo ω va a tener que ser necesariamente mas chico que θ , lo cual finaliza la demostración.

□

Observación 4.5.9. Podemos hacer los mismos razonamientos y construcciones comparando con el espacio de curvatura constante nula, es decir, el plano. En este caso las conclusiones que obtendríamos son las inversas, mas específicamente: si hacemos una construcción como la primera obtendríamos que el lado a comparar va a ser mas chico en el triángulo de comparación que en $\mathcal{U}_n$ y si hacemos la segunda obtendríamos que el ángulo también va a ser mas chico en el triángulo de comparación que en el de $\mathcal{U}_n$.

Corolario 4.5.10. Sea $\Delta(a, b, c)$ un triángulo geodésico en u_n con $|ab| + |ac| + |bc| < 2\pi$ entonces la suma de sus ángulos interiores es un número S con $\pi \leq S \leq 3\pi$.

Demostración. Por el segundo ítem del Teorema 4.5.8 tenemos que al armar un triángulo de comparación en la esfera con lados de la misma longitud que los lados de $\Delta(a, b, c)$, los ángulos interiores no se agrandan. Dado que la suma de los ángulos interiores de un triángulo geodésico en la esfera es a lo sumo 3π obtenemos que $S \leq 3\pi$. La desigualdad estricta se obtiene del hecho que estamos pidiendo que $|ab| + |ac| + |bc| < 2\pi$ y en tal caso el triángulo de comparación no puede ser una geodésica completa, como estos son los únicos triángulos en la esfera para los cuales la suma de los ángulos interiores es 3π , se sigue que $S < 3\pi$. Para la desigualdad $\pi \leq S$ simplemente hay que observar que en el plano la suma de los ángulos interiores del triángulo es π y tener en cuenta la observación previa a este corolario. $\square$

Capítulo 5

Álgebras de Lie con una norma de Finsler Ad-invariante

En este capítulo vamos a estudiar más en profundidad las propiedades que tiene que tener un álgebra de Lie $\mathfrak{g}$ para admitir una norma Ad-invariante. Para las definiciones y notaciones vamos a referir al lector a la Sección 4.2 del Capítulo 4 de este trabajo. Recordamos además que en todo este capítulo trabajaremos con álgebras de Lie reales de dimensión finita, por más que muchas de las definiciones pueden ser dadas de forma mas general. Además siempre que hablemos de un grupo de Lie G , vamos a estar asumiendo que dicho grupo es conexo.

En la primer sección de este capítulo nos centraremos en resultados generales de álgebras de Lie reales y su forma de Killing, mientras que en la segunda y en la tercera estudiaremos el caso en que el álgebra de Lie admite una norma Ad-invariante e incluiremos resultados originales publicados en [48].

5.1. Propiedades básicas de la forma de Killing de un álgebra de Lie

Una de las diferencias más importantes que encontramos al trabajar en un álgebra de Lie $\mathfrak{g}$ y $\mathfrak{u}_n$, es que en $\mathfrak{u}_n$ tenemos definido el producto interno de Frobenius, lo cual nos permite definir ciertas herramientas que imitan el caso riemanniano. En el caso de un álgebra de Lie real abstracta, el papel que jugaba antes el producto interno de Frobenius lo va a ocupar el opuesto de la forma de Killing del álgebra de Lie. La traducción no será exacta, pero nos será suficiente para entender gran parte de la geometría.

Comenzaremos definiendo la forma de Killing en el caso de un álgebra de Lie real y en la siguiente sección nos vamos a centrar en que sucede en el caso de que la misma admita una norma Ad-invariante.

Definición 5.1.1. Definimos la forma de Killing de $\mathfrak{g}$ como la aplicación $\mathfrak{B} : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$ dada por

$$\mathfrak{B}(X, Y) = \text{Tr}(\text{ad}_X \circ \text{ad}_Y).$$

Proposición 5.1.2. Sea $\mathfrak{g}$ un álgebra de Lie y $\mathfrak{B}$ su forma de Killing entonces:

- (1) $\mathfrak{B}$ es bilineal y simétrica.
- (2) Si $\varphi : \mathfrak{g} \longrightarrow \mathfrak{g}$ es un automorfismo y $X, Y \in \mathfrak{g}$ entonces $\mathfrak{B}(\varphi(X), \varphi(Y)) = \mathfrak{B}(X, Y)$. En particular como para $g \in \mathfrak{g}$ resulta que Ad_g es un automorfismo, se sigue que $\mathfrak{B}(\text{Ad}_g(X), \text{Ad}_g(Y)) = \mathfrak{B}(X, Y)$.

Demostración. Para ver (1) notamos que la simetría de $\mathfrak{B}$ es una consecuencia inmediata de la ciclicidad de la traza. La bilinealidad se sigue de forma inmediata de la linealidad de la traza y la bilinealidad del corchete de Lie. Por otro lado para probar (2) consideramos φ un automorfismo y entonces para todo $X, Y \in \mathfrak{g}$ tenemos que

$$\text{ad}_{\varphi(X)} \varphi(Y) = \varphi \circ \text{ad}_X(Y),$$

dado que es un morfismo de álgebras de Lie. Esto se puede reescribir como

$$\text{ad}_{\varphi(X)} = \varphi \circ \text{ad}_X \circ \varphi^{-1},$$

entonces

$$\begin{aligned} \mathfrak{B}(\varphi(X), \varphi(Y)) &= \text{Tr}(\text{ad}_{\varphi(X)} \circ \text{ad}_{\varphi(Y)}) = \text{Tr}(\varphi \circ \text{ad}_X \circ \varphi^{-1} \circ \varphi \circ \text{ad}_Y \circ \varphi^{-1}) \\ &= \text{Tr}(\varphi \circ \text{ad}_X \circ \text{ad}_Y \circ \varphi^{-1}) = \text{Tr}(\text{ad}_X \circ \text{ad}_Y) = \mathfrak{B}(X, Y). \end{aligned}$$

□

Lema 5.1.3. Sean $x, y, z \in \mathfrak{g}$ entonces $\mathfrak{B}([x, y], z) = \mathfrak{B}(x, [y, z])$.

Demostración. Recordamos que por la identidad de Jacobi vale la siguiente igualdad

$$\text{ad}_{[x, y]} = \text{ad}_x \circ \text{ad}_y - \text{ad}_y \circ \text{ad}_x,$$

luego aplicando esta igualdad y la ciclicidad de la traza obtenemos

$$\begin{aligned} \mathfrak{B}([x, y], z) &= \text{Tr}(\text{ad}_{[x, y]} \circ \text{ad}_z) = \text{Tr}(\text{ad}_x \circ \text{ad}_y \circ \text{ad}_z - \text{ad}_y \circ \text{ad}_x \circ \text{ad}_z) \\ &= \text{Tr}(\text{ad}_x \circ \text{ad}_y \circ \text{ad}_z) - \text{Tr}(\text{ad}_x \circ \text{ad}_z \circ \text{ad}_y) \\ &= \text{Tr}(\text{ad}_y \circ \text{ad}_x \circ \text{ad}_z) - \text{Tr}(\text{ad}_x \circ \text{ad}_z \circ \text{ad}_y) \\ &= \text{Tr}(\text{ad}_x \circ \text{ad}_y \circ \text{ad}_z - \text{ad}_x \circ \text{ad}_z \circ \text{ad}_y) \\ &= \text{Tr}(\text{ad}_x \circ \text{ad}_{[y, z]}) \\ &= \mathfrak{B}(x, [y, z]) \end{aligned}$$

□

Definición 5.1.4 (Álgebras de Lie solubles, nilpotentes y abelianas). Sea $\mathfrak{g}$ un álgebra de Lie, decimos que

(1) $\mathfrak{g}$ es *soluble* si la serie derivada

$$\mathfrak{g} \supseteq [\mathfrak{g}, \mathfrak{g}] \supseteq [[\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]] \supseteq \cdots,$$

mas formalmente definida como $\mathfrak{g}_0 = \mathfrak{g}$ y $\mathfrak{g}_k = [\mathfrak{g}_{k-1}, \mathfrak{g}_{k-1}]$ con $k \in \mathbb{N}$, es eventualmente nula.

(2) $\mathfrak{g}$ es *nilpotente* si la serie central descendente

$$\mathfrak{g} \supseteq [\mathfrak{g}, \mathfrak{g}] \supseteq [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \supseteq [\mathfrak{g}, [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]]] \supseteq \cdots,$$

mas formalmente definida por $\mathfrak{g}_0 = \mathfrak{g}$ y $\mathfrak{g}_k = [\mathfrak{g}, \mathfrak{g}_{k-1}]$ con $k \in \mathbb{N}$, es eventualmente nula,

(3) $\mathfrak{g}$ abeliana si $\text{ad}_x = 0$ para todo $x \in \mathfrak{g}$.

Definición 5.1.5. Decimos que un subespacio $I \subseteq \mathfrak{g}$ es un *ideal* si $[\mathfrak{g}, I] \subseteq I$. Notamos además que un ideal es a su vez un álgebra de Lie, por lo cual diremos que I es soluble, nilpotente o abeliano si lo es como álgebra de Lie.

Observación 5.1.6. Todo ideal abeliano es nilpotente y todo ideal nilpotente es soluble.

Proposición 5.1.7. (Criterio de Solubilidad de Cartan) Sea $\mathfrak{g}$ un álgebra de Lie de dimensión finita y $I \subseteq \mathfrak{g}$ un ideal entonces I es soluble si y solo si $\mathfrak{B}(v, z) = 0$ para todo $v \in [I, I]$ y para todo $z \in \mathfrak{g}$.

Demostración. Una demostración de este resultado se puede ver en [38, pág. 20]. $\square$

Proposición 5.1.8. (Teorema de Engel) Dada $\mathfrak{g}$ un álgebra de Lie real de dimensión finita, un ideal $I \subseteq \mathfrak{g}$ es nilpotente si y solo si ad_x es un endomorfismo nilpotente para todo $x \in \mathfrak{g}$.

Demostración. Una demostración de este resultado se puede ver en [38, pág. 12]. $\square$

Definición 5.1.9. (Ideal de Killing, centro y radical). Sea $\mathfrak{g}$ un álgebra de Lie y $\mathfrak{B}$ su forma de Killing definimos:

- (1) el *centro* de $\mathfrak{g}$ como $Z(\mathfrak{g}) = \{x \in \mathfrak{g} : \text{ad}_x = 0\}$,
- (2) el *ideal de Killing* como $\mathfrak{g}^\perp = \{x \in \mathfrak{g} : \mathfrak{B}(x, y) = 0 \quad \forall y \in \mathfrak{g}\}$,
- (3) el *radical* de $\mathfrak{g}$ como su ideal soluble maximal.

Observación 5.1.10. Es claro que el centro es un ideal bilátero de $\mathfrak{g}$, por otro lado el ideal de Killing es un ideal como consecuencia del Lema 5.1.3 y el radical es un ideal por la misma definición del mismo.

Observación 5.1.11. Tenemos la siguiente contención entre los ideales definidos previamente

$$Z(\mathfrak{g}) \subset \mathfrak{g}^\perp \subset \text{Rad}(\mathfrak{g}), \quad (5.1)$$

la primer contención es inmediata y la segunda es consecuencia del Criterio de Solubilidad de Cartan (Proposición 5.1.7).

Ejemplo 5.1.12. Los siguientes ejemplos muestran que las contenciones de la observación anterior pueden ser estrictas.

- (1) Para la primer contención observamos que si $\mathfrak{g}$ es un álgebra de Lie nilpotente entonces la serie central decreciente es eventualmente nula, con lo cual para $x, y \in \mathfrak{g}$ el endomorfismo $\text{ad}_x \circ \text{ad}_y$ va a ser nilpotente y por ende tener traza nula. Concluimos que si el álgebra de Lie es nilpotente tendremos que $\mathfrak{B}(x, y) = 0$ para todo $x, y \in \mathfrak{g}$, luego cualquier álgebra de Lie nilpotente no abeliana nos servirá como ejemplo. Uno clásico es el álgebra de Heisenberg que se construye como $\mathfrak{h}_3 = \text{gen}\{x, y, z\}$ con las relaciones $[x, y] = z$, $[x, z] = 0$ y $[y, z] = 0$, en este caso tendremos que $Z(\mathfrak{h}_3) = \text{gen}\{x, y\}$ pero como es nilpotente $\mathfrak{h}_3^\perp = \mathfrak{h}_3$ y entonces $Z(\mathfrak{h}_3) \subsetneq \mathfrak{h}_3^\perp$.

- (2) Para la segunda contención consideramos $\mathfrak{g} = \text{gen}\{x, y\}$ con la relación $[x, y] = y$. Luego tenemos que el álgebra es soluble ya que

$$[[\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]] = [\mathbb{R}y, \mathbb{R}y] = 0,$$

por otro lado $\text{ad}_x^2(x) = 0$ y $\text{ad}_x^2(y) = y$, entonces los autovalores de ad_x^2 son 1 y 0 (ambos de multiplicidad simple por ser $\mathfrak{g}$ de dimensión 2), en tal caso $\mathfrak{B}(x, x) = \text{Tr}(\text{ad}_x^2) = 0 + 1 = 1 \neq 0$. Tenemos entonces que $\mathfrak{g}^\perp \subsetneq \text{Rad}(\mathfrak{g})$.

Definición 5.1.13 (Álgebra de Lie simple y semisimple). Sea $\mathfrak{g}$ un álgebra de Lie decimos que $\mathfrak{g}$ es *simple* si no es abeliana y no tiene ideales propios distintos de cero. Por otro lado decimos que es *semisimple* si su único ideal conmutativo es $I = \{0\}$.

Teorema 5.1.14. Sea $\mathfrak{g}$ un álgebra de Lie de dimensión finita, son equivalentes

- (1) El álgebra de Lie $\mathfrak{g}$ es semisimple.
- (2) $\text{Rad}(\mathfrak{g}) = \{0\}$.
- (3) Podemos descomponer a $\mathfrak{g}$ como $\mathfrak{g} = \sum_{k=1}^n I_k$ donde I_k es un ideal que cuando lo pensamos como álgebra de Lie es simple. En este caso tendremos además que todo ideal simple de $\mathfrak{g}$ es alguno de los I_k y que la forma de Killing en cada I_k es la restricción de la forma de Killing en $\mathfrak{g}$.
- (4) La forma de Killing de $\mathfrak{g}$ es no degenerada, es decir, $\mathfrak{B}(x, x) \neq 0$ para todo $x \in \mathfrak{g}$ no nulo.

Demostración. Ver [38, pág. 22,23] y [18, Capítulo I, §6.1]. □

Definición 5.1.15 (Representaciones). Sea $\mathfrak{g}$ un álgebra de Lie real, una representación de $\mathfrak{g}$ en un espacio vectorial real V es un morfismo de álgebras de Lie

$$\rho : \mathfrak{g} \longrightarrow \mathfrak{gl}(V) = \text{Lie}(GL(V)),$$

además decimos que:

- (1) Un subespacio $S \subseteq V$ es invariante bajo la representación si $\rho(x)S \subseteq S$ para todo $x \in \mathfrak{g}$.
- (2) La representación es irreducible si V no tiene subespacios invariantes propios no nulos.
- (3) La representación es semisimple si es suma de representaciones irreducibles.

Definición 5.1.16 (Álgebra de Lie reductiva). Sea $\mathfrak{g}$ un álgebra de Lie, decimos que $\mathfrak{g}$ es *reductiva* si la representación adjunta

$$\text{ad} : \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$$

es semisimple.

Proposición 5.1.17 (Equivalencias de reductiva). Sea $\mathfrak{g}$ un álgebra de Lie, son equivalentes

- (1) $\mathfrak{g}$ es un álgebra de Lie reductiva.
- (2) $\mathfrak{g}$ se descompone como $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{a}$, donde $\mathfrak{s}$ es una subálgebra de Lie de $\mathfrak{g}$ semisimple y $\mathfrak{a}$ es una subálgebra de Lie abeliana.

(3) Vale la igualdad $Z(\mathfrak{g}) = \mathfrak{g}^\perp = \text{Rad}(\mathfrak{g})$.

Demostración. La demostración de este resultado puede encontrarse en [18, Capítulo I, §6.4]. $\square$

Observación 5.1.18. Siempre tenemos las contenciones de ideales de la Observación 5.1.11, por la proposición previa un álgebra de Lie es reductiva cuando esas contenciones son igualdades. Esto nos dice que los ejemplos dados en 5.1.12 son álgebras de Lie no reductivas.

Teorema 5.1.19. (Levi-Malcev) Sea $\mathfrak{g}$ un álgebra de Lie entonces existe $\mathfrak{s}$ una subálgebra semisimple de forma tal que

$$\mathfrak{g} = \text{Rad}(\mathfrak{g}) \oplus \mathfrak{s}.$$

Demostración. Ver [18, Capítulo I, §6.8]. $\square$

Definición 5.1.20 (Álgebra de Lie compacta). Si $\mathfrak{g}$ es un álgebra de Lie, decimos que es *compacta* si su forma de Killing es definida negativa. Notamos que si es compacta es en particular semisimple.

5.2. Operadores Hermitianos, rango numérico y espectro

En esta sección vamos a estudiar las propiedades de un álgebra de Lie real que admite una norma de Finsler Ad-invariante y vamos a ver que necesariamente tiene que ser un álgebra de Lie reductiva. Si bien muchas de las definiciones y los resultados de este capítulo se pueden dar en el contexto de espacios de Banach, vamos a asumir que el álgebra de Lie $\mathfrak{g}$ va a ser siempre real y de dimensión finita. Los resultados originales de esta sección son los teoremas 5.2.15, 5.2.16 y 5.2.19, publicados en [48].

El siguiente resultado es una generalización del Lema 3.1.4 dado en u_n y su demostración es análoga:

Lema 5.2.1. Sea G un grupo de Lie conexo y $\mathfrak{g}$ su álgebra de Lie. Supongamos que $\mathfrak{g}$ admite una métrica de Finsler Ad-invariante $|\cdot|$. Sea $v \neq 0$ y $\varphi \in \mathfrak{g}^*$ un funcional de norma 1 que realiza la norma de v . Luego, para cada $x \in \mathfrak{g}$ tenemos que

$$(1) \quad \varphi([v, x]) = 0 \quad (\text{i.e. } \varphi \circ \text{ad}_v = 0).$$

$$(2) \quad \varphi([x, [x, v]]) \leq 0.$$

Observación 5.2.2. Si $|\cdot|$ es una norma de Finsler en un espacio vectorial real, entonces $\|x\| = |x| + |-x|$ define una norma en el sentido usual, es decir, completamente homogénea. Si la norma de Finsler original es Ad-invariante, entonces la norma $\|\cdot\|$ también lo es y es topológicamente equivalente a la original (ver Observación 2.1.5). Para lo que nos interesa estudiar en este capítulo podemos asumir que estamos trabajando con una norma en el sentido usual.

Definición 5.2.3. (Complexificación) Si $\mathfrak{g}$ es un álgebra de Lie real, definimos su complexificación como $\mathfrak{g}^{\mathbb{C}} = \mathfrak{g} \oplus i\mathfrak{g}$. Esta complexificación tiene una estructura de álgebra de Lie compleja donde, para $x, x', y, y' \in \mathfrak{g}$ y $\alpha, \beta \in \mathbb{R}$, las operaciones y el coquete de Lie vienen dados por:

$$(x + iy) + (x' + iy') = (x + x') + i(y + y')$$

$$(\alpha + i\beta)(x + iy) = (\alpha x - \beta y) + i(\alpha y + \beta x)$$

$$[x + iy, x' + iy'] = [x, x'] - [y, y'] + i([x, y'] + [x', y]).$$

A continuación vamos a definir la norma de Taylor en la complexificación de $\mathfrak{g}$ de la misma forma que lo hicimos en la Observación 3.4.2:

Definición 5.2.4 (Norma de Taylor). Sea $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{g}$, definimos su norma de Taylor asociada en $\mathfrak{g}^{\mathbb{C}} = \mathfrak{g} \oplus i\mathfrak{g}$ como

$$|x + iy|_T = \sup_{t \in [0, 2\pi]} |x \cos t - y \sin t|.$$

Observación 5.2.5. Si tenemos un operador lineal $T : \mathfrak{g} \longrightarrow \mathfrak{g}$ definimos su extensión compleja como el operador $T^{\mathbb{C}} : \mathfrak{g}^{\mathbb{C}} \longrightarrow \mathfrak{g}^{\mathbb{C}}$ dado por $T^{\mathbb{C}}(x + iy) = T(x) + iT(y)$. Es fácil de ver que este operador sigue siendo lineal y acotado cuando en $\mathfrak{g}^{\mathbb{C}}$ consideramos la topología inducida por la norma de Taylor.

Observación 5.2.6. Si consideramos una norma de Finsler Ad-invariante $\mathfrak{g}$, la norma de Taylor que induce en $\mathfrak{g}^{\mathbb{C}}$ es claramente una norma de Finsler, además

$$\begin{aligned} |\text{Ad}_g(x + iy)|_T &= |\text{Ad}_g(x) + i \text{Ad}_g(y)|_T = \sup_{t \in [0, 2\pi]} |\text{Ad}_g(x) \cos t - \text{Ad}_g(y) \sin t| \\ &= \sup_{t \in [0, 2\pi]} |\text{Ad}_g(x \cos t - y \sin t)| = \sup_{t \in [0, 2\pi]} |x \cos t - y \sin t| \\ &= |x + iy|_T, \end{aligned}$$

es decir esta norma es Ad_g invariante. Notar que cuando escribimos $\text{Ad}_g(x + iy)$ estamos pensando en $\text{Ad}_g^{\mathbb{C}}$ y que además g pertenece al álgebra de Lie original y no a la complexificación. No tiene porque ser la norma de Taylor $\text{Ad}_g^{\mathbb{C}}$ -invariante.

Definición 5.2.7. Si $T : \mathfrak{g} \longrightarrow \mathfrak{g}$ es un operador lineal definimos el espectro de T como

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda \text{ id no es inversible}\},$$

es decir como el espectro de $T^{\mathbb{C}}$. Por otro lado definimos el espectro real de T como

$$\sigma_{\mathbb{R}}(T) = \{t \in \mathbb{R} : T - t \text{ id no es inversible}\}.$$

Proposición 5.2.8. Si $T : \mathfrak{g} \longrightarrow \mathfrak{g}$ es un operador lineal entonces

$$\sigma_{\mathbb{R}}(T) = \sigma(T^{\mathbb{C}}) \cap \mathbb{R}.$$

Demostración. Si $\lambda \in \mathbb{R}$, para $x, y \in \mathfrak{g}$, podemos escribir

$$(T^{\mathbb{C}} - \lambda I_{\mathfrak{g}^{\mathbb{C}}})(x + iy) = (T - \lambda I_{\mathfrak{g}})(x) + i(T - \lambda I_{\mathfrak{g}})(y),$$

donde $I_{\mathfrak{g}}$ es la identidad en $\mathfrak{g}$, mientras que $I_{\mathfrak{g}^{\mathbb{C}}}$ es la identidad en la complexificación. Si tomamos un operador arbitrario en la complexificación y lo escribimos como $A + iB$ donde A y B los podemos pensar como operadores en $\mathfrak{g}$ tenemos que la composición $(T - \lambda I_{\mathfrak{g}^{\mathbb{C}}}) \circ (A + iB)$ está dada por

$$(T - \lambda I_{\mathfrak{g}})(A - B) + i(T - \lambda I_{\mathfrak{g}})(A + B). \quad (5.2)$$

Notamos entonces que si $(T - \lambda I_g)$ es inversible, tomando $A = \frac{1}{2}(T - \lambda I_g)^{-1}$ y $B = -A$, obtenemos que $A + iB$ es la inversa de $(T - \lambda I_{\mathbb{C}})$ lo cual prueba la contención $\sigma_{\mathbb{R}}(T) \supseteq \sigma(T^{\mathbb{C}}) \cap \mathbb{R}$. Por otro lado si $(T - \lambda I_g)$ tiene inversa, escribiendo $(T - \lambda I_g)^{-1} = A + iB$, tenemos que $A - B$ debe ser la inversa de $T - \lambda I_g$, lo cual prueba la otra contención.

Observamos que como estamos trabajando en dimensión finita nos alcanzó probar que los operadores eran inversibles a derecha, pero con un argumento similar al de la fórmula 5.2, podemos ver que vale la igualdad sin usar argumentos de dimensión. $\square$

Vamos a dar la siguiente definición para operadores Hermitianos en espacios normados:

Definición 5.2.9 (Operadores Hermitianos). Sea $(X, \|\cdot\|)$ un espacio normado complejo, decimos que un operador lineal $A : X \rightarrow X$ es Hermitiano si para todo $z \in X$ y para todo $s \in \mathbb{R}$ se cumple la siguiente igualdad:

$$\|e^{isA}z\| = \|z\|.$$

Observación 5.2.10. Si $(X, \langle \cdot, \cdot \rangle)$ es un espacio vectorial complejo con producto interno y A es un operador lineal hermitiano en el sentido usual (i.e. $A^* = A$) entonces para $s \in \mathbb{R}$ y $z \in X$ tenemos que

$$\begin{aligned} \|e^{isA}z\|^2 &= \langle e^{isA}z, e^{isA}z \rangle = \langle z, (e^{isA})^* e^{isA}z \rangle = \langle z, e^{-isA^*} e^{isA}z \rangle \\ &= \langle z, e^{-isA} e^{isA}z \rangle = \langle z, z \rangle = \|z\|^2, \end{aligned}$$

luego es hermitiano según nuestra definición para espacios complejos normados.

Definición 5.2.11 (Rango numérico espacial). Sea $(X, \|\cdot\|)$ un espacio normado. Para $z \in X$ no nulo definimos

$$\mathcal{N}_z = \{\varphi \in X^* : \|\varphi\| = 1, \varphi(z) = |z|\}.$$

Por el Teorema de Hahn-Banach este conjunto es no vacío, además es cerrado y convexo. Sea A un operador acotado en X , definimos el *rango numérico espacial* de A como el conjunto $V(A) \subset \mathbb{C}$ dado por

$$V(A) = \{\varphi(Az) : \varphi \in \mathcal{N}_z, |z| = 1\}.$$

Observación 5.2.12. El conjunto $V(A)$ es conexo y si $\sigma(A)$ es el espectro de A tenemos que

$$\sigma(A) \subset \overline{V(A)},$$

es decir el espectro está contenido en la clausura del rango numérico espacial. Una demostración de estos resultados puede encontrarse en [17, pág 88 y 102].

Definición 5.2.13 (Radio espectral). Si A es un operador lineal acotado en un espacio de Banach, definimos su *radio espectral* como

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

A continuación daremos una proposición que es un resultado usual en operadores Hermitianos en espacios de Hilbert, pero también sigue valiendo cuando trabajamos en espacios de Banach:

Proposición 5.2.14. Si $(X, \|\cdot\|)$ es un espacio de Banach complejo y A es un operador hermitiano entonces

$$(1) \operatorname{conv}(\sigma(A)) = \overline{\operatorname{conv}(V(A))}.$$

$$(2) \|A\| = \rho(A).$$

Demostración. El primer ítem está demostrado en [17, pág. 53 y 86], para el segundo ítem ver [62]. $\square$

Teorema 5.2.15. Si $\mathfrak{g}$ es un álgebra de Lie real de dimensión finita que admite una norma de Finsler Ad-invariante entonces valen las siguientes afirmaciones:

- (1) $\sigma(\operatorname{ad} x) \subset i\mathbb{R}$ para todo $x \in \mathfrak{g}$.
- (2) $\mathfrak{B}$ es semidefinida negativa en $\mathfrak{g}$.
- (3) Todo elemento nilpotente de $\mathfrak{g}$ es central.
- (4) $Z(\mathfrak{g}) = \mathfrak{g}^\perp = \{x \in \mathfrak{g} : \mathfrak{B}(x, x) = 0\} = \operatorname{Rad}(\mathfrak{g})$ (i.e. $\mathfrak{g}$ es reductiva).

Demostración. Por la Observación 5.2.2, $\mathfrak{g}$ admite una norma Ad-invariante completamente homogénea, a la cual vamos a denotar para esta demostración $|\cdot|$.

Sea $x \in \mathfrak{g}$, pensándolo como un elemento de $\mathfrak{g}^\mathbb{C}$, sabemos que existe un funcional unitario φ en $\mathfrak{g}^\mathbb{C}$ que verifica $\varphi(x) = |x|_T = |x|$. Repitiendo la demostración del Lema 5.2.1 (que la hicimos en el Lema 3.1.4, pero es idéntica) para el caso complejo se sigue que para todo $z \in \mathfrak{g}^\mathbb{C}$ vale la igualdad $\operatorname{Re} \varphi(\operatorname{ad}_x(z)) = 0$. Esto prueba que $V(A) \subset i\mathbb{R}$, luego la primer afirmación se sigue de la Observación 5.2.12.

Como $\sigma(\operatorname{ad} x) \subset i\mathbb{R}$ para todo $x \in \mathfrak{g}$, luego $\sigma(\operatorname{ad}^2 x) \subset (-\infty, 0]$ y entonces $\mathfrak{B}(x, x) = \operatorname{Tr}(\operatorname{ad}^2 x) \in (-\infty, 0]$ de donde se sigue que $\mathfrak{B}$ es semidefinida negativa.

Para ver la tercer afirmación, consideraremos $x \in \mathfrak{g}$. Como la norma de Taylor 5.2.4 es Ad_g invariante, tenemos que $i \operatorname{ad}_x$ es un operador hermitiano en $(\mathfrak{g}^\mathbb{C}, |\cdot|_T)$ ya que para $s \in \mathbb{R}$ y $z \in \mathfrak{g}^\mathbb{C}$ se tiene

$$|e^{isi \operatorname{ad}_x} z|_T = |\operatorname{Ad}_{-sx}(z)|_T = |z|_T,$$

luego por la Proposición 5.2.14 tenemos que

$$\|\operatorname{ad}_x\| = \|i \operatorname{ad}_x\| = \rho(i \operatorname{ad}_x) = \rho(\operatorname{ad}_x).$$

Si x es nilpotente entonces existe $k \in \mathbb{N}$ de forma tal que $\operatorname{ad}_x^k = 0$, lo cual implica que el único autovalor de ad_x es cero y entonces $\|\operatorname{ad}_x\| = \rho(\operatorname{ad}_x) = 0$, con lo cual $\operatorname{ad}_x = 0$ y luego $x \in Z(\mathfrak{g})$, finalizando así la demostración de (3).

Para probar la cuarta afirmación recordamos las inclusiones de la Observación 5.1.11, es decir

$$Z(\mathfrak{g}) \subseteq \mathfrak{g}^\perp \subseteq \operatorname{Rad}(\mathfrak{g}).$$

Para ver que la primer contención es una igualdad tomamos $x \in \mathfrak{g}^\perp$, en tal caso $\mathfrak{B}(x, y) = 0$ para todo $y \in \mathfrak{g}$ y en particular $\operatorname{Tr}(\operatorname{ad}_x^2) = \mathfrak{B}(x, x) = 0$. Como $\sigma(\operatorname{ad}_x^2) \subseteq (-\infty, 0]$ y tiene traza cero debe ser $\sigma(\operatorname{ad}_x^2) = \{0\}$, lo cual me dice que ad_x^2 es nilpotente y en tal caso también lo es ad_x . Por el ítem previo tenemos que $x \in Z(\mathfrak{g})$. Esto prueba que $Z(\mathfrak{g}) = \mathfrak{g}^\perp$. Hasta ahora tenemos que

$$Z(\mathfrak{g}) = \mathfrak{g}^\perp \subseteq \operatorname{Rad}(\mathfrak{g}).$$

Veamos que la ultima contención es una igualdad, para ello supongamos que no, en tal caso existe $x \in \text{Rad}(\mathfrak{g})$ con $x \notin Z(\mathfrak{g})$, luego por lo visto previamente debe de ser $\sigma(\text{ad } x) \neq \{0\}$, lo cual implica que existen $r \neq 0$ e $y \neq 0$ de forma tal que

$$[x, [x, y]] = \text{ad}_x^2(y) = -r^2 y.$$

Como el radical es un ideal, debe ser $[x, y] \in \text{Rad}(\mathfrak{g})$, pero usando que $[\text{Rad}(\mathfrak{g}), \text{Rad}(\mathfrak{g})] \subset \mathfrak{g}^\perp$ (ver [18, Capítulo I, §5.5]), tenemos que $[x, [x, y]] \in \mathfrak{g}^\perp = Z(\mathfrak{g})$. Luego, como $r \neq 0$ e $y \in Z(\mathfrak{g})$ entonces $[x, y] = 0$ lo cual implica que $y = 0$, llegando así a una contradicción. $\square$

Teorema 5.2.16. Si $\mathfrak{g}$ admite una norma de Finsler Ad-invariante, luego $\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{s}$, donde $\mathfrak{s}$ es semisimple y $\mathfrak{B}(x, x) < 0$ para todo $x \in \mathfrak{s}$ no nulo. Además, $G = H \times K$, donde H es el grupo conexo y abeliano dado por $H = \exp(Z(\mathfrak{g}))$ y K es el grupo compacto con centro finito dado por $K = \exp(\mathfrak{s})$.

Demostración. En el teorema previo vimos que $\mathfrak{g}$ es reductiva, luego aplicando el Teorema de Levi-Malcev (Teorema 5.1.19) para el caso reductivo tenemos que

$$\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{s},$$

donde $\mathfrak{s}$ es una subálgebra semisimple de $\mathfrak{g}$. Por otro lado, mirando el mapa exponencial

$$\exp : Z(\mathfrak{g}) \oplus \mathfrak{s} \longrightarrow G,$$

tenemos que si $z + s \in Z(\mathfrak{g}) \oplus \mathfrak{s}$ entonces $\exp(z + s) = \exp(z) \exp(s)$ ya que $z \in Z(\mathfrak{g})$. Como G es conexo, sabemos que la exponencial es sobreyectiva y por lo anterior se sigue que

$$G = \exp(Z(\mathfrak{g}) \oplus \mathfrak{s}) = \exp(Z(\mathfrak{g})) \times \exp(\mathfrak{s}) = H \times K.$$

Como el álgebra de Lie de K es *compacta*, es decir su forma de Killing es no degenerada, el grupo K es compacto con centro finito por el Teorema de Weyl (Ver [19] Capítulo IX, §1.4). $\square$

Observación 5.2.17 (Grupos abstractos y sumas ortogonales). Sea $\mathfrak{g}$ un álgebra de Lie real de dimensión finita y supongamos que admite una norma de Finsler Ad-invariante. En tal caso podemos descomponer a $\mathfrak{g}$ como

$$\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{s},$$

con $\mathfrak{s}$ semisimple ya que $\mathfrak{g}$ es reductiva por el Teorema 5.2.15. Consideremos una base $\{e_i\}_{i=1, \dots, n}$ de $Z(\mathfrak{g})$ y definamos un producto interno declarando que la base en consideración es ortonormal. Para $z + s \in \mathfrak{g}$ y $z' + s' \in \mathfrak{g}$ vamos a definir ahora

$$\langle z + s, z' + s' \rangle_{\mathfrak{g}} = \langle z, z' \rangle_Z - \mathfrak{B}(s, s'),$$

obteniendo así una forma bilineal. Notamos por otro lado que

$$\langle z + s, z + s \rangle_{\mathfrak{g}} = \langle z, z \rangle_Z - \mathfrak{B}(s, s) \geq 0,$$

ya que $\mathfrak{B}$ es semidefinida negativa por el Teorema 5.2.15, luego la forma bilineal definida es semidefinida positiva. Por otra parte será también no degenerada ya que si

$$\langle z + s, z + s \rangle_{\mathfrak{g}} = \langle z, z \rangle_Z - \mathfrak{B}(s, s) = 0,$$

ambos sumandos deben ser 0, lo cual implica que $z = 0$ por ser $\langle \cdot, \cdot \rangle$ un producto interno y $\mathfrak{B}$ restringido a $\mathfrak{s}$ no degenerada (ver el Teorema 5.1.14). En conclusión, tenemos definido un producto interno en $\mathfrak{g}$. Además para $g \in G$, la representación Ad_g actúa de forma trivial en el centro y es invariante en la parte semisimple, luego

$$\begin{aligned} \langle \text{Ad}_g(z + s), \text{Ad}_g(z' + s') \rangle_g &= \langle z + \text{Ad}_g(s), z' + \text{Ad}_g(s') \rangle_g \\ &= \langle z, z' \rangle_g + \langle \text{Ad}_g(s), \text{Ad}_g(s') \rangle_g \\ &= \langle z, z' \rangle_Z - \mathfrak{B}(s, s') \\ &= \langle z + s, z' + s' \rangle_g, \end{aligned}$$

se sigue entonces que el producto interno es Ad-invariante. Notar que lo que hemos probado, en virtud del Teorema 5.2.15, el Teorema 5.2.16 y la Observación 4.2.10, es lo siguiente:

Teorema 5.2.18. Sea G un grupo de Lie conexo de dimensión finita y $\mathfrak{g}$ su álgebra de Lie, entonces

- (1) El álgebra de Lie $\mathfrak{g}$ admite una norma de Finsler Ad-invariante si y solo si G admite una métrica riemanniana bi-invariante.
- (2) En las condiciones de (1), tenemos que G se descompone como $G = H \times K$ con H un subgrupo abeliano y K un subgrupo compacto con centro finito.

Por otro lado si comenzamos desde el grupo de Lie con una distancia bi-invariante y basándonos en los resultados Berestovskii [10], tenemos el siguiente resultado:

Teorema 5.2.19. Si (G, dist) es un grupo topológico conexo, localmente compacto, localmente contractible con una cuasi distancia bi-invariante, entonces $G = H \times K$ con H un grupo de Lie conexo abeliano y K es un grupo de Lie conexo, compacto y con centro discreto. Además si definimos para $v \in \mathfrak{g}$ la aplicación

$$|v| = \lim_{t \rightarrow 0^+} \frac{\text{dist}(1, \exp(tv))}{t},$$

entonces tenemos una norma de Finsler Ad-invariante en $\mathfrak{g}$ y la cuasi distancia que induce en G mediante el ínfimo de longitud de caminos coincide con dist (la cuasi distancia con la que comenzamos). Por otro lado, si partimos de una distancia, tenemos el mismo resultado pero la norma va a estar definida como

$$\|v\| = \lim_{t \rightarrow 0} \frac{\text{dist}(1, \exp(tv))}{t}$$

y va a resultar completamente homogénea.

Demostración. Adaptando [10, Teorema 7] al caso de métricas no reversibles (lo cual es inmediato), obtenemos que si G es localmente compacto, localmente contractible y admite una cuasi distancia bi-invariante, entonces G es un grupo de Lie con una norma de Finsler Ad-invariante. La descomposición se sigue entonces del Teorema 5.2.16. La fórmula para la norma también fue probada en [10, Lema 9], considerando el límite con $t > 0$ para el caso de una cuasi distancia. $\square$

Observación 5.2.20. Las hipótesis del teorema previo son evidentes en un grupo de Lie conexo de dimensión finita. Luego en este caso tenemos que el primer ítem del Teorema 5.2.18 es equivalente a pedir que G admita una cuasi distancia bi-invariante.

Como consecuencia del Teorema 5.2.18, al tener una métrica riemanniana, podemos comparar los resultados de descomposición dados por Milnor en [55]. En particular el [55, Lema 7.5] nos dice que G es isomorfo al producto de un grupo compacto y una copia de $\mathbb{R}^m$ para cierto $m \in \mathbb{N}$. En el ejemplo que veremos a continuación fijaremos un grupo de Lie cuya álgebra de Lie admite una norma de Finsler Ad-invariante y veremos que la descomposición dada por Milnor y la que estamos presentando en este trabajo no es la misma:

Ejemplo 5.2.21. En [55, Lema 7.5] la descomposición viene dada por $G = \mathbb{R}^m \times K'$ en donde $\mathbb{R}^m \simeq Z(\mathfrak{g})$ y $\text{Lie}(K') = \mathfrak{s}$ (la parte semisimple). Para su construcción, él considera el revestimiento universal $\pi : \tilde{G} \rightarrow G$, si notamos $\Pi = \ker(\pi)$ y $\pi : \tilde{G} \rightarrow G$ a la proyección a la parte abeliana, lo que él hace es mirar $V = \text{gen}\{pr_1(\Pi)\}$ y su complemento ortogonal $V^\perp \subset \mathbb{R}^n$ obteniendo que

$$G \simeq V^\perp \times \frac{(K' \times V)}{\Pi} = \mathbb{R}^n \times K'',$$

donde K'' es compacto. Para llevar esto a un ejemplo, consideremos el grupo dado por $G = S^1 \times SO(n)$, cuya álgebra de Lie es $\mathfrak{g} = \mathbb{R} \times \mathfrak{so}_n$, en nuestra descripción dada en el Teorema 5.2.16 tendremos que $H = S^1$ y $K = SO(n)$, mientras que si miramos lo que hace Milnor tendremos que $\tilde{G} = \mathbb{R} \times Spin(n)$, $\Pi = \mathbb{Z} \times \mathbb{Z}_2$ y $V^\perp = \{0\}$, de donde se sigue que $K'' = G$ y $m = 0$. Esto muestra que las descomposiciones no coinciden.

5.3. Generalización de los resultados

En el Teorema 5.2.15 probamos que toda álgebra de Lie real de dimensión finita con una norma de Finsler Ad-invariante se puede descomponer como $\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{s}$, con $\mathfrak{s}$ semisimple. Además, como la representación adjunta actúa de forma fiel en $\mathfrak{s}$, podemos identificar a $\mathfrak{s}$ como una subálgebra de Lie de $\mathfrak{so}_n(\mathbb{R}) \subseteq \mathfrak{su}_n(\mathbb{C})$, donde $n = \dim(\mathfrak{s})$. La motivación original de trabajar en $\mathfrak{u}_n$ y en $\mathfrak{su}_n$ fue el hecho previamente mencionado, el cuál es una consecuencia inmediata de los resultados de Milnor [55] en el caso de que tengamos un producto interno Ad-invariante en $\mathfrak{g}$. Por el Teorema 5.2.18, que se deduce del Teorema 5.2.15, probamos que vale también si comenzamos con una norma que no necesariamente proviene de un producto interno. Sin embargo, no todos los resultados publicados en [47] se pueden deducir del hecho de estar trabajando en una subálgebra de Lie de $\mathfrak{su}_n$. En esta sección vamos a desarrollar algunos de los resultados publicados en [48] que son similares a los dados en $\mathfrak{u}_n$, pero trabajando intrínsecamente con las herramientas de un álgebra de Lie reductible.

Con este fin, vamos a tener que recurrir al concepto de sistema de raíces asociado a una subálgebra de Cartan, en tal caso necesitaremos trabajar con un álgebra de Lie y su complexificación (ver 5.2.3). Haremos esto siguiendo la construcción clásica en el caso de álgebras de Lie semisimples que se puede encontrar en el libro de Hall [32, Capítulos 7 y 8]. Como vamos a trabajar con un álgebra de Lie real y compacta (i.e. semisimple con forma de Killing definida negativa) la construcción se puede pensar como un caso particular del caso semisimple, sin embargo, vamos a incluir la construcción en detalle en este trabajo y será incluso un poco mas intuitiva ya que partiremos de un álgebra de Lie compacta y no necesitaremos usar la “forma real compacta”. Los resultados originales publicados en [48] son el Lema 5.3.33, el Teorema 5.3.34, el cual es la generalización del Teorema 3.1.13 y el Teorema 5.3.39 que generaliza el Teorema 4.4.14.

Observación 5.3.1. Si $\mathfrak{f}$ es un álgebra de Lie compacta y $\mathfrak{g}$ su complexificación entonces

$$\mathfrak{B}_{\mathfrak{g}}(x + iy, x' + iy') = \mathfrak{B}_{\mathfrak{g}}(x, x') - \mathfrak{B}_{\mathfrak{g}}(y, y') + i(\mathfrak{B}_{\mathfrak{g}}(x, y') + \mathfrak{B}_{\mathfrak{g}}(x', y)),$$

donde $\mathfrak{B}_{\mathfrak{g}}$ denota la forma de Killing de $\mathfrak{g}$. Esto nos dice que la forma de Killing de $\mathfrak{g}$ restringida a $\mathfrak{f}$ es la forma de Killing de $\mathfrak{f}$. Tomando un elemento arbitrario $x + iy \in \mathfrak{g}$, si suponemos que

$$\mathfrak{B}_{\mathfrak{g}}(x + iy, x' + iy') = 0 \quad \forall x' + iy' \in \mathfrak{g},$$

y elegimos adecuadamente x' e y' es fácil ver que $x = y = 0$. Esto nos dice que si $\mathfrak{f}$ es compacta, entonces su complexificación es semisimple, ya que $\mathfrak{B}_{\mathfrak{g}}$ es no degenerada.

Por lo observado previamente, notaremos como $\mathfrak{B}$ a la forma de Killing de $\mathfrak{g}$ y de $\mathfrak{f}$ indistintamente.

En general, si $\mathfrak{g}$ es un álgebra de Lie compleja semisimple, siempre existe un álgebra de Lie compacta $\mathfrak{f}$ de forma tal que $\mathfrak{g} \simeq \mathfrak{f} \oplus i\mathfrak{f}$. Este álgebra de Lie se suele llamar *forma real compacta* de $\mathfrak{g}$. Nosotros partiremos del álgebra de Lie compacta, ya que nos interesa estudiar la geometría de $\mathfrak{f}$. Su complexificación simplemente será utilizada como una herramienta para poder hablar de un sistema de raíces en $\mathfrak{f}$.

Definición 5.3.2 (Subálgebra de Cartan). Si $\mathfrak{g}$ es un álgebra de Lie semisimple compleja y $\mathfrak{c} \subseteq \mathfrak{g}$ es un subespacio complejo, diremos que $\mathfrak{c}$ es una subálgebra de Cartan si

- (1) Para todo $c_1, c_2 \in \mathfrak{c}$ vale que $[c_1, c_2] = 0$.
- (2) Dado $x \in \mathfrak{g}$, si $[c, x] = 0$ para todo $c \in \mathfrak{c}$ debe ser $x \in \mathfrak{c}$.
- (3) Para todo $c \in \mathfrak{c}$ el morfismo ad_c es diagonalizable.

Observación 5.3.3. Las primeras dos condiciones nos dicen que $\mathfrak{c}$ es una subálgebra de Lie abeliana y maximal con dicha propiedad. Por otro lado, la condición (3) nos dice que todo morfismo ad_c es diagonalizable y como $\mathfrak{c}$ es abeliana por (1), se sigue que los morfismos ad_c son diagonalizables en forma simultánea.

Observación 5.3.4. Esta definición es adecuada para el caso complejo y semisimple. Una definición mas general se puede encontrar en [19, Capítulo VII, §2], junto con los resultados necesarios para probar la equivalencia de las definiciones en nuestro caso.

Un ejemplo usual es $\mathfrak{sl}_n = \text{Lie}(SL_n(\mathbb{C}))$, el cual es un álgebra de Lie semisimple y una subálgebra de Cartan es la formada por las matrices diagonales con traza nula.

Definición 5.3.5. Sea $\mathfrak{f}$ un álgebra de Lie real y $\mathfrak{g}$ su complexificación, definimos el adjunto de $x + iy \in \mathfrak{g}$ como

$$(x + iy)^* = -x + iy.$$

Observación 5.3.6. En vista de nuestro trabajo en $\mathfrak{u}_n$, podemos pensar en $x + iy \in \mathfrak{u}_n^{\mathbb{C}}$, en dicho caso, si $x^* = \bar{x}^t$ (transpuesta y conjugada), tenemos que

$$(x + iy)^* = x^* - iy^* = -x + iy,$$

lo cual motiva nuestra definición.

Observación 5.3.7. Notamos que si $\mathfrak{f}$ es un álgebra de Lie compacta y $\mathfrak{g}$ su complexificación, la aplicación dada por $-\mathfrak{B}_{\mathfrak{f}}$ define un producto interno real en $\mathfrak{f}$, pero la aplicación $-\mathfrak{B}_{\mathfrak{g}}$ no define un producto interno complejo en $\mathfrak{g}$. Esto es debido a que la forma de Killing es lineal y no sesquilineal, sin embargo podemos extender el producto interno real que tenemos que $\mathfrak{f}$ a $\mathfrak{g}$ complexificándolo como se muestra en la siguiente proposición.

Proposición 5.3.8. Si $\mathfrak{f}$ es un álgebra de Lie real compacta, la aplicación dada por

$$\langle x + iy, x' + iy' \rangle := -\mathfrak{B}(x, x') + \mathfrak{B}(y, y') + i(\mathfrak{B}(x, y') - \mathfrak{B}(x', y)),$$

define un producto interno en $\mathfrak{g}$ por ser la extensión sesqui-lineal del producto interno en $\mathfrak{f}$ dado por el opuesto de la forma de Killing. Además tenemos que:

- (1) Si $x \in \mathfrak{f}$, $y, z \in \mathfrak{g}$ entonces $\langle \text{ad}_x(y), z \rangle = -\langle y, \text{ad}_x(z) \rangle$.
- (2) Si $x, y, z \in \mathfrak{g}$ entonces $\langle \text{ad}_x(y), z \rangle = \langle y, \text{ad}_{x^*}(z) \rangle$ i.e. $(\text{ad}_x)^* = \text{ad}_{x^*}$.

Demostración. El hecho de que la aplicación sea un producto interno en $\mathfrak{g}$ es una consecuencia inmediata de que $-\mathfrak{B}$ es un producto interno en $\mathfrak{f}$ por ser esta un álgebra de Lie compacta. Para los siguientes ítems notamos primero que si $x, y, z \in \mathfrak{f}$ entonces por el Lema 5.1.3

$$\langle [x, y], z \rangle = -\mathfrak{B}([x, y], z) = \mathfrak{B}([y, x], z) = \mathfrak{B}(y, [x, z]) = -\langle y, [x, z] \rangle.$$

Si $x, y, z \in \mathfrak{g}$, escribiendo $x = x_1 + ix_2$ e $y = y_1 + iy_2$, tenemos que

$$\begin{aligned} \langle [x, y], z \rangle &= \langle [x_1, y_1], z \rangle - \langle [x_2, y_2], z \rangle + i\langle [x_1, y_2], z \rangle + i\langle [x_2, y_1], z \rangle \\ &= -\langle y_1, [x_1, z] \rangle + \langle y_2, [x_2, z] \rangle - i\langle y_2, [x_1, z] \rangle - i\langle y_1, [x_2, z] \rangle \\ &= \langle y_1, [-x_1 + ix_2, z] \rangle + \langle iy_2, [-x_1 + ix_2, z] \rangle \\ &= \langle y, [x^*, z] \rangle, \end{aligned}$$

esto prueba (3) y como $x^* = -x$ para $x \in \mathfrak{f}$ se sigue de inmediato (2) como un caso particular. □

Proposición 5.3.9. Sean $\mathfrak{f}$ un álgebra de Lie real compacta y $\mathfrak{h} \subseteq \mathfrak{f}$ una subálgebra de Lie abeliana maximal entonces $\mathfrak{c} = \mathfrak{h} \oplus i\mathfrak{h}$ es una subálgebra de Cartan de $\mathfrak{g} := \mathfrak{f} \oplus i\mathfrak{f}$.

Demostración. Es evidente que $\mathfrak{c}$ va a ser un subespacio complejo de $\mathfrak{g}$ y (1) se sigue de forma inmediata de como se define el corchete en la complexificación y del hecho de que $\mathfrak{h}$ es abeliana. Para ver (2), sea $x + iy \in \mathfrak{g}$ de forma tal que $[x + iy, h_1 + ih_2] = 0$ para todo $h_1, h_2 \in \mathfrak{h}$, en particular para todo $h \in \mathfrak{h}$ tendremos que $0 = [x + iy, h] = [x, h] + i[y, h]$ y entonces $0 = [x, h] = [y, h]$ por la unicidad de la descomposición de elementos en $\mathfrak{g}$. Por la maximalidad de $\mathfrak{h}$ se sigue que $x, y \in \mathfrak{h}$ y entonces $x + iy \in \mathfrak{c}$. Finalmente para ver que ad_x es diagonalizable para $x \in \mathfrak{c}$, notamos que $x = x_1 + ix_2$ con $x_1, x_2 \in \mathfrak{h}$, por el ítem (2) de la proposición anterior ad_{x_1} y ad_{x_2} son operadores antihermitianos y en particular diagonalizables, como además $[x_1, x_2] = 0$ por estar ambos en $\mathfrak{h}$, se sigue que se diagonalizan en forma simultánea y entonces ad_x es diagonalizable. □

Observación 5.3.10. En la demostración de la proposición anterior queda en manifiesto la necesidad de trabajar en álgebras de Lie complejas para poder diagonalizar operadores antihermitianos.

Observación 5.3.11. Si $\mathfrak{h} \subseteq \mathfrak{k}$ es una subálgebra abeliana maximal, muchas veces nos referiremos a ella como una subálgebra de Cartan real de $\mathfrak{k}$. Esto tiene sentido en virtud de la definición mas general de subálgebra de Cartan que puede ser encontrada en [19, Capítulo VII, §2].

Observación 5.3.12. Dado que nuestro interés está en el álgebra de Lie real compacta, siempre que trabajemos con $\mathfrak{g}$ un álgebra compleja semisimple vamos a pensar que esta es de la forma $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$ con $\mathfrak{k}$ compacta real. Si $\mathfrak{c} \subseteq \mathfrak{g}$ es una subálgebra de Cartan, vamos pensarla como $\mathfrak{c} = \mathfrak{h} \oplus i\mathfrak{h}$ con $\mathfrak{h} \subseteq \mathfrak{k}$ abeliana maximal (ver Proposición 5.3.9).

Definición 5.3.13 (Raíz relativa a una subálgebra de Cartan). Sea $\mathfrak{k}$ un álgebra de Lie real compacta y $\mathfrak{c} \subseteq \mathfrak{g}$ una subálgebra de Cartan. Diremos que un funcional no nulo $\alpha : \mathfrak{c} \rightarrow \mathbb{C}$ es una *raíz relativa al álgebra de Cartan* $\mathfrak{c}$ si existe $x \in \mathfrak{g}$ no nulo de forma que

$$\text{ad}_c(x) = \alpha(c)x$$

para todo $c \in \mathfrak{c}$. Notaremos como Δ al conjunto de todas las raíces asociadas a una subálgebra de Cartan $\mathfrak{c}$.

Observación 5.3.14. Si $\alpha \in \mathfrak{c}^*$ es una raíz y consideramos el producto interno de la Proposición 5.3.8, notaremos H_α al único elemento en $\mathfrak{c}$ que satisface

$$\alpha(c) = \langle c, H_\alpha \rangle,$$

para todo $c \in \mathfrak{c}$.

Proposición 5.3.15. Si α es una raíz y H_α es como en la observación previa, entonces $H_\alpha \in i\mathfrak{h}$, donde recordamos que $\mathfrak{c} = \mathfrak{h} + i\mathfrak{h}$.

Demostración. Si α es una raíz entonces existe $x \in \mathfrak{g}$ no nulo para el cual $\text{ad}_c(x) = \alpha(c)x = \langle c, H_\alpha \rangle x$ para todo $c \in \mathfrak{c}$. En particular podemos tomar $c = h \in \mathfrak{h}$, por la Proposición 5.3.8, ad_h debe ser antihermitiano y luego $\langle h, H_\alpha \rangle$ imaginario puro. La única forma que esto suceda, en virtud de la definición de nuestro producto interno, es que $H_\alpha \in i\mathfrak{h}$. $\square$

Observación 5.3.16. Por la proposición previa podemos escribir $H_\alpha = ih_\alpha$ para un único $h_\alpha \in \mathfrak{h}$. Esto nos dice que cuando pensamos en la restricción α a $\mathfrak{h}$, este es un funcional del dual de $\mathfrak{h}$. En otras palabras, mediante la representación de Riesz, podemos pensar a las raíces como elementos del espacio real $i\mathfrak{h}$ o, si nos quedamos con los de la forma h_α , como elementos de $\mathfrak{h}$.

Definición 5.3.17 (Espacio asociado a una raíz). Sea $\mathfrak{g}$ un álgebra de Lie semisimple y $\mathfrak{c} \subseteq \mathfrak{g}$ una subálgebra de Cartan como en la Proposición 5.3.9. Si α es una raíz relativa a $\mathfrak{c}$ definimos el conjunto $\mathfrak{g}_\alpha$ como

$$\mathfrak{g}_\alpha := \{x \in \mathfrak{g} : [c, x] = \alpha(c)x \text{ para todo } c \in \mathfrak{c}\},$$

a este conjunto lo llamaremos *espacio asociado a la raíz* α . En caso que α no sea una raíz y no sea el funcional nulo, notaremos de la misma forma a este conjunto pero será $\mathfrak{g}_\alpha = \{0\}$ y en el caso de que $\alpha = 0$ tendremos $\mathfrak{g}_0 = \mathfrak{c}$.

A continuación mencionaremos algunos resultados importantes en la teoría de raíces de álgebras de Lie complejas semisimples, las demostraciones de los mismos pueden ser encontradas en el Capítulo 7 del libro de Hall [32], mas específicamente en las Proposiciones 7.16, 7.17, 7.29, el Lema 7.22 y los Teoremas 7.23 y 7.26.

Observación 5.3.18. Sea $\mathfrak{g}$ un álgebra de Lie semisimple y $\mathfrak{c} \subseteq \mathfrak{g}$ una subálgebra de Cartan. Si Δ es el conjunto de todas las raíces relativas a dicha álgebra de Cartan entonces:

- (1) Tenemos una descomposición como suma directa de *subespacios complejos* dada por

$$\mathfrak{g} = \mathfrak{c} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_{\alpha}.$$

- (2) Si α, β son raíces entonces $[\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}] \subseteq \mathfrak{g}_{\alpha+\beta}$.
- (3) Los elementos asociados a las raíces que notamos h_{α} generan $\mathfrak{c}$ como espacio vectorial complejo.
- (4) Si α es una raíz, los únicos múltiplos de α que son raíces son α y $-\alpha$.
- (5) Si α es una raíz entonces el espacio $\mathfrak{g}_{\alpha}$ tiene dimensión 1.
- (6) Si α es una raíz entonces el morfismo $s_{\alpha} : \mathfrak{c} \longrightarrow \mathfrak{c}$ dado por

$$s_{\alpha}(c) = c - 2 \frac{\alpha(c)}{\alpha(h_{\alpha})} h_{\alpha}$$

preserva Δ .

- (7) Si $\alpha, \beta \in \Delta$ entonces $2 \frac{\alpha(h_{\beta})}{\alpha(h_{\alpha})} \in \mathbb{Z}$.

A continuación daremos las definiciones y propiedades básicas de un sistema de raíces en un espacio vectorial real. Por las proposiciones previas se puede verificar como a partir de nuestra definición de raíces en un álgebra semisimple compleja $\mathfrak{g}$ de la forma $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$ con $\mathfrak{k}$ compacta, podemos construir un sistema de raíces reales en $\mathfrak{k}$ que cumple la siguiente definición.

Definición 5.3.19. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial real de dimensión finita con producto interno. Decimos que Δ es un *sistema de raíces reales* si se cumple que

- (1) El conjunto Δ genera E .
- (2) Si $x \in \Delta$ entonces los únicos múltiplos de x en Δ son x y $-x$.
- (3) Si $x, y \in \Delta$ entonces

$$y - 2 \frac{\langle y, x \rangle}{\langle x, x \rangle} x \in \Delta.$$

- (4) Si $x, y \in \Delta$ entonces $2 \frac{\langle y, x \rangle}{\langle x, x \rangle} \in \mathbb{Z}$.

Definición 5.3.20. Si $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial real de dimensión finita con producto interno y Δ es un sistema de raíces, decimos que $\Pi \subseteq \Delta$ es una *base asociada al sistema de raíces* (o simplemente base) si

- (1) El conjunto Π es una base de E .

- (2) Toda raíz $x \in \Delta$ se puede escribir como una combinación lineal de elementos de Δ solo utilizando escalares enteros y de forma tal que todo elemento se puede escribir utilizando escalares mayores iguales a cero o bien menores o iguales a cero.

Observación 5.3.21. Fijada una base Π vamos a notar como Δ^+ (resp. Δ^-) a los elementos de Δ que se escriben como combinación lineal de elementos de Π solo utilizando escalares no negativos (resp. no positivos). Como cero nunca es una raíz, tenemos la unión disjunta:

$$\Delta = \Delta^+ \cup \Delta^-.$$

Para volver a nuestro caso de interés vamos a citar el Teorema 6.11 del libro de Knapp [40]. Con este resultado vamos a construir un sistema de raíces reales en nuestro álgebra de Lie compacta a partir del sistema de raíces asociado a un subálgebra de Cartan de la complexificación.

Teorema 5.3.22. Sean $\mathfrak{k}$ un álgebra de Lie compacta, $\mathfrak{g}$ su complexificación y $\mathfrak{c} \subseteq \mathfrak{g}$ una subálgebra de Cartan de forma tal que $\mathfrak{c} = \mathfrak{h} \oplus i\mathfrak{h}$ con $\mathfrak{h} \subseteq \mathfrak{k}$ abeliana maximal. Si Δ es un sistema de raíces en $\mathfrak{g}$ asociado a $\mathfrak{c}$ entonces podemos escribir a $\mathfrak{k}$ como

$$\mathfrak{k} = \sum_{\alpha \in \Delta} \mathbb{R}iH_{\alpha} + \sum_{\alpha \in \Delta} \mathbb{R}(x_{\alpha} - x_{-\alpha}) + \sum_{\alpha \in \Delta} \mathbb{R}i(x_{\alpha} + x_{-\alpha}),$$

donde $x_{\alpha} \in \mathfrak{g}_{\alpha}$, son elementos que satisfacen las siguientes propiedades

- (1) $[x_{\alpha}, x_{-\alpha}] = H_{\alpha}$.
- (2) $[x_{\alpha}, x_{\beta}] = 0$ si $\alpha + \beta \neq 0$ y $\alpha + \beta \notin \Delta$.
- (3) $[x_{\alpha}, x_{\beta}] = n_{\alpha, \beta} x_{\alpha + \beta}$, con $n_{\alpha, \beta} \in \mathbb{R}$ que satisfacen $n_{\alpha, \beta} = -n_{-\alpha, -\beta}$.

Además este elección puede ser hecha de forma tal que $x_{-\alpha} = (x_{\alpha})^*$.

Observación 5.3.23. Recordamos que $ih_{\alpha} = H_{\alpha}$ y si llamamos

$$u_{\alpha} = \frac{1}{\sqrt{2}}(x_{\alpha} - x_{-\alpha}) \quad v_{\alpha} = \frac{1}{\sqrt{2}}i(x_{\alpha} - x_{-\alpha}),$$

como estos elementos generan $\mathfrak{k}$ como espacio vectorial real por el teorema anterior y cumplen las propiedades enunciadas en la Observación 5.3.18, tenemos que forman un sistema de raíces en $\mathfrak{k}$ como $\mathbb{R}$ -espacio vectorial. En tal caso podemos quedarnos con una base Π y descomponer a Δ como $\Delta = \Delta^+ \cup \Delta^-$. Tenemos entonces la siguiente igualdad

$$\mathfrak{k} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta^+} \mathbb{R}u_{\alpha} \oplus \bigoplus_{\alpha \in \Delta^+} \mathbb{R}v_{\alpha} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta^+} Z_{\alpha},$$

donde las sumas directas son de subespacios ortogonales para el producto interno dado por el opuesto de la forma de Killing y $Z_{\alpha} = \mathbb{R}u_{\alpha} \oplus \mathbb{R}v_{\alpha}$. Además, se deducen del teorema previo las igualdades

$$[h, u_{\alpha}] = \alpha(h)v_{\alpha} \quad [h, v_{\alpha}] = -\alpha(h)u_{\alpha} \quad [u_{\alpha}, v_{\alpha}] = h_{\alpha}. \quad (5.3)$$

Notamos que los elementos de la forma h_{α} con $\alpha \in \Pi$ generan el álgebra abeliana maximal $\mathfrak{h} \subseteq \mathfrak{k}$.

Para $x, y \in \mathfrak{k}$ vamos a notar como $x \otimes y$ a la aplicación lineal dada por $(x \otimes y)z = \langle z, y \rangle x$. Con esta notación tenemos además que para $h \in \mathfrak{h}$

$$\text{ad}_h = \sum_{\alpha \in \Delta^+} \alpha(h)(v_\alpha \otimes u_\alpha - u_\alpha \otimes v_\alpha) = \sum_{\alpha \in \Delta^+} \alpha(h)T_\alpha, \quad (5.4)$$

donde $T_\alpha = (v_\alpha \otimes u_\alpha - u_\alpha \otimes v_\alpha)$. Esta igualdad se puede verificar fácilmente viendo que ambos términos coinciden en todos los elementos de $\mathfrak{k}$, los de la forma u_α y los de la forma v_α . Observamos además que si $\alpha \neq \beta$ entonces $T_\alpha T_\beta = 0$ y

$$T_\alpha^2 = -(u_\alpha \otimes u_\alpha + v_\alpha \otimes v_\alpha) = -P_\alpha.$$

Resulta así que P_α es una proyección ortogonal en un espacio 2-dimensional en $(\mathfrak{k}, \langle \cdot, \cdot \rangle)$. En tal caso $\text{ad}_h^2 = -\sum_{\alpha} \alpha(h)^2 P_\alpha$ y los autovalores de ad_h son $\{\pm i\alpha(h) : \alpha \in \Delta_+\}$.

Observación 5.3.24. Dado que ya hemos probado la existencia de un sistema de raíces reales, volveremos a trabajar con un álgebra de Lie reductiva real, es decir, volveremos a tener en cuenta el centro. Para aliviar notación, llamaremos $\mathfrak{g}$ a dicha álgebra real y será $\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{k}$. Cuando estemos trabajando con un producto interno, este será el dado por el opuesto de la forma de Killing extendido al centro del álgebra, el cual fue construido en la Observación 5.2.17. Nos referiremos a este producto interno como el de Frobenius, ya que en el contexto del álgebra de Lie de las matrices unitarias coincide con el mismo.

Por otra parte, una raíz es un elemento $\alpha \in \mathfrak{k}^*$, pero podemos pensarla en $\mathfrak{g}^*$ simplemente extendiendo el funcional para que valga cero en el centro del álgebra.

A continuación generalizaremos los resultados que fueron obtenidos en capítulos previos en $\mathfrak{u}_n$ y $\mathfrak{su}_n$ para el contexto de un álgebra de Lie reductible. Estos resultados son originales y fueron publicados en [48].

Lema 5.3.25 (Funcionales que realizan la norma). Sea $|\cdot|$ una norma de Finsler Ad-invariante en un álgebra de Lie reductiva $\mathfrak{g}$. Sean $v \in \mathfrak{g}$ no nulo y $\varphi \in \mathfrak{g}^*$ un funcional que realiza la norma de v , entonces:

- (1) $\varphi([x, v]) = 0 \quad \forall x \in \mathfrak{g}$.
- (2) $\varphi([x, [x, v]]) \leq 0 \quad \forall x \in \mathfrak{g}$.
- (3) Existe $z \in \mathfrak{g}$ que cumple $\varphi(x) = \langle z, x \rangle = (\text{ad}_z | \text{ad}_x)$ para todo $x \in \mathfrak{k}$, donde $\langle \cdot, \cdot \rangle$ es el producto interno dado por el opuesto de la forma de Killing extendido como hicimos al final de la sección previa (ver Observación 5.2.17) y $(\cdot | \cdot)$ es el de Frobenius.
- (4) Si $g \in G$ y consideramos $\psi = \langle \text{Ad}_g z, \cdot \rangle$ entonces $\|\psi\| = \|\varphi\|$, donde φ es el funcional del ítem previo.
- (5) Si z es como en (3) entonces $[z, v] = 0$.

Demostración. Notamos que los ítems (1), (2) y (5) de este lema se prueban igual que en el Lema 3.1.4. La existencia de z es una consecuencia inmediata de que el opuesto de la forma de Killing define un producto interno real en un álgebra de Lie compacta y de que $\langle z, x \rangle = -\mathfrak{B}(z, x) = -\text{Tr}(\text{ad}_z \text{ad}_x^*) = \text{Tr}(\text{ad}_z \text{ad}_x) = (\text{ad}_z | \text{ad}_x)$, donde usamos que ad_x es antihermitiano. Para ver (4) notamos que

$$\psi(w) = \langle \text{Ad}_g z, w \rangle = \langle z, \text{Ad}_g^{-1} w \rangle = \varphi(\text{Ad}_g^{-1} w) \leq \|\varphi\| |\text{Ad}_g^{-1} w| = \|\varphi\| |w|$$

donde usamos que la norma es Ad-invariante en $\mathfrak{g}$, luego $\|\psi\| \leq \|\varphi\|$. Con una demostración similar obtenemos la otra desigualdad. $\square$

A continuación daremos un lema que usaremos repetidamente en esta sección, para una demostración del mismo, ver §2,3.Proposición 10 en [19].

Lema 5.3.26. Sean $\mathfrak{f}$ un álgebra de Lie real compacta y $v, z \in \mathfrak{f}$ tales que $[v, z] = 0$ entonces existe $\mathfrak{h} \subseteq \mathfrak{f}$, una subálgebra abeliana maximal, que contiene a v y a z .

Definición 5.3.27. Sean $\varphi = \langle z, \cdot \rangle \in \mathfrak{f}^*$ un funcional que realiza la norma de cierto $v \in \mathfrak{f}$ como en el Lema 5.3.25, $\mathfrak{h}$ un álgebra abeliana maximal que contiene a v y a z (cuya existencia la garantiza el lema previo) y Δ^+ el conjunto de raíces positivas, definimos

$$\text{supp}(v) = \{\alpha \in \Delta^+ : \alpha(v) \neq 0\},$$

es decir, el conjunto de raíces que soportan a v , de forma análoga definimos $\text{supp}(z)$. Observamos que

$$\text{supp}(v)^c = \{\alpha \in \Delta^+ : \alpha(v) = 0\} = \{\alpha \in \Delta^+ : [v, u_\alpha] = 0 = [v, v_\alpha]\}$$

por (5.3). Vamos a considerar el subespacio $S_v = \bigoplus_{\alpha \in \text{supp}(v)} Z_\alpha$ y denotar como $P_v : \mathfrak{f} \rightarrow S_v$ a la proyección ortogonal en S_v .

Observación 5.3.28 (Funcionales que realizan la norma y subálgebras de Cartan). Sean $z, v \in \mathfrak{f}$ con $[v, z] = 0$, consideramos una subálgebra de Cartan $\mathfrak{h} \subset \mathfrak{f}$ que contenga a z y a v . Por (5.4) y el Lema 5.3.25 tenemos que

$$\langle z, v \rangle = (\text{ad}_z | \text{ad}_v) = -\text{Tr}(\text{ad}_z \circ \text{ad}_v) = \sum_{\alpha} \alpha(z) \alpha(v) \text{Tr}(P_\alpha) = 2 \sum_{\alpha} \alpha(z) \alpha(v).$$

Tomamos $w \in \mathfrak{f}$ y lo escribimos como $w = w_0 + \sum_{\alpha} a_{\alpha} u_{\alpha} + b_{\alpha} v_{\alpha}$ con $w_0 \in \mathfrak{h}$ y entonces por (5.3)

$$[z, w] = \sum_{\alpha \in \text{supp}(z)} \alpha(z) (a_{\alpha} v_{\alpha} - b_{\alpha} u_{\alpha}) \quad [w, v] = \sum_{\alpha \in \text{supp}(v)} \alpha(v) (-a_{\alpha} v_{\alpha} + b_{\alpha} u_{\alpha}).$$

En tal caso usando en Lema 5.3.25, si φ realiza la norma de v , se sigue que

$$-\sum_{\alpha \in \text{supp}(v) \cap \text{supp}(z)} \alpha(z) \alpha(v) (a_{\alpha}^2 + b_{\alpha}^2) = \langle [z, w], [w, v] \rangle = \varphi([w, [w, v]]) \leq 0. \quad (5.5)$$

Tomando $w = u_{\alpha}$, se sigue que si $\varphi = \langle z, \cdot \rangle$ realiza la norma de v , entonces

$$\alpha(z) \alpha(v) \geq 0 \quad \text{para todo } \alpha \in \Delta^+. \quad (5.6)$$

Lema 5.3.29. Sean $|\cdot|$ una norma de Finsler Ad-invariante en $\mathfrak{g}$, $v, z \in \mathfrak{g}$, $\varphi = \langle z, \cdot \rangle \in \mathfrak{g}^*$ un funcional que realiza la norma de $v \in \mathfrak{g}$ y $F_{\varphi} \subset \mathfrak{g}$ la cara asociada al funcional φ . Si escribimos $v = v_0 + v_1, z = z_0 + z_1 \in Z(\mathfrak{g}) \oplus \mathfrak{s}$ y consideramos una subálgebra de Cartan $\mathfrak{h} \subseteq \mathfrak{f}$ que contiene a v_1 y a z_1 y una raíz $\alpha \in \Delta^+$. Entonces

- (1) Si $\alpha \in \text{supp}(v_1)^c \cap \text{supp}(z_1)$ entonces existe $\psi \neq \varphi$ que realiza la norma de v .
- (2) Si $\alpha \in \text{supp}(v_1) \cap \text{supp}(z_1)^c$ entonces F_{φ} no es un singleton.

Demostración. Sea $\ker \alpha = \{h_\alpha\}^\perp \subset \mathfrak{h}$, consideramos la descomposición

$$\mathfrak{k} = \mathbb{R}h_\alpha \oplus \ker \alpha \oplus Z_\alpha \oplus \bigoplus_{\beta \neq \alpha} Z_\beta$$

donde cada sumando es ortogonal respecto a la forma de Killing y la suma de los dos primeros subespacios es $\mathfrak{h}$. Sea $|\alpha|^2 = \alpha(h_\alpha)$, escribimos:

$$v = v_0 + \frac{\alpha(v)}{|\alpha|^2} h_\alpha + v_\perp \quad z = z_0 + \frac{\alpha(z)}{|\alpha|^2} h_\alpha + z_\perp. \quad (5.7)$$

donde $h_\alpha \perp v_\perp \in \mathfrak{h}$ y $h_\alpha \perp z_\perp \in \mathfrak{h}$. Luego

$$|v| = \langle z, v \rangle = \langle v_0, z_0 \rangle_Z + \frac{\alpha(v)\alpha(z)}{|\alpha|^2} + \langle v_\perp, z_\perp \rangle.$$

Por el Lema B.1 de [52] existe $g \in K$ (el grupo de Lie conexo semisimple que integra $\mathfrak{k}$) tal que $z' = \text{Ad}_g z_1$ verifica:

- $\alpha(z') = 0$,
- z_1 y z' tienen la misma componente en $\ker \alpha$,
- las componentes de z_1 y z' en $\bigoplus_{\beta \neq \alpha} Z_\beta$ tienen la misma norma.

Luego tenemos que:

$$z'' = \text{Ad}_g(z) = z_0 + z' = z_0 + 0 + z_\perp + a_\alpha u_\alpha + b_\alpha v_\alpha + 0 \in Z(\mathfrak{g}) \oplus \mathbb{R}h_\alpha \oplus \ker \alpha \oplus Z_\alpha \oplus \bigoplus_{\beta \neq \alpha} Z_\beta, \quad (5.8)$$

dado que la componente de z en $\bigoplus Z_\alpha$ era nula. Observamos que si $\psi = \langle z'', \cdot \rangle$, luego $\psi = \varphi \circ \text{Ad}_{g^{-1}}$ entonces $\|\psi\| = \|\varphi\| = 1$ por el Lema 5.3.25, y

$$\psi(v) = \langle z'', v \rangle = \langle z_0, v_0 \rangle_Z + 0 + \langle v_\perp, z_\perp \rangle = \langle z, v \rangle = |v|,$$

ya que $\alpha(v) = 0$. En conclusión, ψ realiza la norma de v . Como $\alpha(z) \neq 0$, tenemos que $z \neq z''$ y luego $\psi \neq \varphi$. Para ver (2) vamos a hacer lo mismo pero ahora con v : existe $g \in K$ de forma tal que, si $v' = \text{Ad}_g v$ (en tal caso $|v'| = |v|$) se verifica

$$v' = v_0 + 0 + v_1 + x_\alpha u_\alpha + y_\alpha v_\alpha + 0.$$

Entonces

$$\varphi(v') = \langle z, v' \rangle = \langle z_0, v_0 \rangle_Z + \langle v_\perp, z_\perp \rangle = |v| = |v'|$$

dado que $\alpha(z) = 0$, lo cual prueba que $\{v, v'\} \subset F_\varphi$. Como $\alpha(v) \neq 0$, tenemos que $v \neq v'$ lo cual prueba (2) y finaliza la demostración. $\square$

Definición 5.3.30. Dado un funcional $\varphi = \langle z_0 + z_1, \cdot \rangle \in \mathfrak{g}^*$ que realiza la norma de $v = v_0 + v_1$, $\mathfrak{h}$ una subálgebra de Cartan de $\mathfrak{k}$ que contiene a v_1 y a z_1 y Δ^+ un sistema de raíces asociado. Decimos que φ es *adaptado* a v si para cada $\alpha \in \Delta^+$ tenemos que $\alpha(v) = 0$ implica que $\alpha(z) = 0$.

Observación 5.3.31. Decimos que un elemento v es *regular* si su centralizador definido como

$$\mathfrak{g}^x = \{y \in \mathfrak{g} : [x, y] = 0\},$$

es de dimensión minimal con respecto a las dimensiones de todos los centralizadores de elementos del álgebra. Si bien no vamos a necesitar usar elementos regulares para nuestros resultados, aportan a la comprensión del tema. Si construimos una subálgebra de Cartan como en la definición anterior tenemos que si v es regular hay un único sistema de raíces asociado y $\text{supp}(v) = \Delta^+$ (de hecho esto es equivalente a ser regular). En tal caso se puede ver que todo funcional que realiza la norma de v es adaptado a v . Las demostraciones de estos resultados pueden ser encontradas en [19].

Observación 5.3.32. En el caso del grupo unitario establecimos la existencia de ciertos funcionales especiales (ver Teorema 3.1.18). Estos funcionales que realizan la norma de un vector $v \in \mathfrak{u}_n$ son exactamente los funcionales adaptados. Dicha construcción fue hecha a partir de promediar elementos y este promedio pudo ser hecho por la siguiente propiedad: dadas dos matrices diagonales $D_1, D_2 \in \mathfrak{u}_n$ que tienen los mismos elementos en la diagonal (pero eventualmente permutados) existe una matriz unitaria $U \in \mathcal{U}_n$ para la cual $D_1 = \text{Ad}_U D_2$. Si consideramos $\mathfrak{h} \subseteq \mathfrak{u}_n$ la subálgebra de Cartan formada por las matrices diagonales, tenemos que un sistema de raíces positivas asociado es el dado por las matrices de la forma $iE_{k,k} - iE_{l,l}$, con $k < l$. Lo que podemos ver en forma general es que: vamos a poder construir los funcionales adaptados con argumentos similares a los del Teorema 3.1.13 si el grupo G actúa sobre un sistema de raíces mediante la acción adjunta de forma transitiva.

Dado que en muchos casos tenemos que las raíces asociadas a un álgebra de Cartan no tienen la misma longitud, este método no va a poder ser utilizado en forma general. Para visualizar este fenómeno consideramos el grupo de Lie $SO(5)$ de matrices de 5×5 ortogonales de determinante 1 cuyo álgebra de Lie es $\mathfrak{so}_5$ formada por las matrices de 5×5 antisimétricas. Dicha álgebra de Lie es compacta, por lo cual el opuesto de la forma de Killing define un producto interno Ad-invariante en la misma. Una álgebra de Cartan $\mathfrak{h}$ en $\mathfrak{so}_5$ es la generada por las matrices $H_1 = E_{1,2} - E_{2,1}$ y $H_2 = E_{3,4} - E_{4,3}$, si ϵ_1, ϵ_2 es la base dual de $\mathfrak{h}$ tenemos un sistema de raíces dado por:

$$\Delta = \{\pm\epsilon_1, \pm\epsilon_2, \pm(\epsilon_1 + \epsilon_2), \pm(\epsilon_1 - \epsilon_2)\},$$

en este caso tenemos las raíces $\pm\epsilon_1, \pm\epsilon_2$ son cortas y $\pm(\epsilon_1 + \epsilon_2), \pm(\epsilon_1 - \epsilon_2)$ son raíces largas. Si tomamos la norma inducida por el opuesto de la forma de Killing tendremos que las raíces cortas miden 1 y las largas miden 2, como la norma es Ad-invariante, tendremos que el grupo $SO(5)$ no puede actuar de forma transitiva sobre las raíces.

Si bien no podemos aplicar las mismas ideas del Teorema 3.1.18, podemos probar la existencia de funcionales adaptados con otra estrategia, la cual se muestra en el siguiente lema:

Lema 5.3.33. Para cada $v \neq 0$ en $\mathfrak{g}$ existe al menos un funcional $\varphi_v = \langle z_v, \cdot \rangle$ adaptado a v . Si la norma es suave en v , el único funcional que realiza la norma de v debe ser adaptado.

Demostración. Consideramos el conjunto $C_v = \{w \in \mathfrak{g} : \varphi = \langle w, \cdot \rangle \text{ realiza la norma de } v\}$, este conjunto es compacto (ya que es cerrado y acotado), no vacío por Hahn-Banach y claramente convexo. Sea $z_v \in C_v$ un elemento minimal para la norma de Frobenius i.e. $\|z_v\|_F = \sqrt{\langle z_v, z_v \rangle} \leq \|w\|_F$ para todo $w \in C_v$.

La existencia de este elemento se debe a que C_v es compacto y convexo, además $z_v \neq 0$ ya que realiza la norma de v . Este elemento es además único en C_v ya que si existieran dos elementos $z_v, z'_v \in C_v$ de norma de Frobenius minimal en C_v , su media aritmética también pertenecería a C_v y tendría norma de Frobenius estrictamente menor (ya que la norma de Frobenius es estrictamente convexa) lo cual contradice la minimalidad.

Escribimos $v = v_0 + v_1, z_v = z_0 + z_1 \in Z(\mathfrak{g}) \oplus \mathfrak{k}$ y sea $\mathfrak{h}$ un álgebra abeliana maximal que contiene a v_1, z_1 . Consideramos Δ^+ un sistema de raíces positivas con respecto a dicha álgebra de Cartan y afirmamos que $\varphi = \langle z_v, \cdot \rangle$ es un funcional adaptado a v .

Sea $\alpha \in \Delta_+$ tal que $\alpha(v) = 0$, en tal caso será $\alpha(v_1) = 0$ (dado que $\alpha(v_0) = 0$ por definición) escribimos a v y z_v como en (5.7); luego $v_1 = v_\perp$ y afirmamos que $z_1 = z_\perp$. Si no fuese así, tendríamos que $\alpha \in \text{supp}(v_1)^c \cap \text{supp}(z_1)$ y luego, como en la demostración del Lema 5.3.29, existe $g \in G$ tal que

$$\text{Ad}_g(z_v) = z_0 + z_\perp + w_\alpha,$$

con $w_\alpha \in Z_\alpha$. Sea $\psi = \langle \text{Ad}_g z, \cdot \rangle$, entonces ψ también realiza la norma de v . Como la norma de Frobenius es Ad-invariante, se sigue que $\|\text{Ad}_g z_v\|_F = \|z_v\|_F$ y por la unicidad vista previamente, tenemos que:

$$z_0 + z_\perp + w_\alpha = \text{Ad}_g z_v = z_v = z_0 + \frac{\alpha(z_1)}{|\alpha|^2} + z_\perp.$$

Esto es posible solo si $w_\alpha = \alpha(z_1) = 0$, lo cual es una contradicción. Se sigue que $z_1 = z_\perp$ y luego $\alpha(z_1) = \alpha(z) = 0$, lo cual prueba la primera parte de la afirmación. Si la norma es suave, el conjunto C_v es un singleton y luego $C_v = \{z_v\}$. En tal caso, el único funcional φ que realiza la norma de v , debe ser adaptado a v . $\square$

Teorema 5.3.34. Sean $v = v_0 + v_1, z = z_0 + z_1 \in \mathfrak{g} = Z(\mathfrak{g}) \oplus \mathfrak{k}$, $\varphi = \langle z, \cdot \rangle$ un funcional que realiza la norma de v y $\mathfrak{h}$ una subálgebra de Cartan que contiene a v_1, z_1 . Luego $\varphi([x, [x, v]]) = 0$ si y solo si $[P_{v_1}x, z] = 0$ (ver Definición 5.3.27). En este caso

- (1) Si φ es un funcional adaptado a v entonces $[x, z] = 0$.
- (2) Si existe un único funcional que realiza la norma de v entonces $[x, z] = 0$.
- (3) Si $F_\varphi = \{v\}$ entonces $[x, v] = 0$.

Demostración. Para que la demostración sea menos técnica vamos a asumir que $x \in \mathfrak{k}$, dado que todos los enunciados involucran conmutadores, se puede demostrar de la misma forma el caso general.

Comenzamos utilizando el Teorema 5.3.22 y la observación que está a continuación del mismo para escribir

$$x = x_k + \sum_{\alpha \in \Delta^+} x_\alpha u_\alpha + y_\alpha v_\alpha,$$

en tal caso al aplicar P_{v_1} obtenemos

$$P_{v_1}x = \sum_{\sigma \in \text{supp}(v)} x_\sigma u_\sigma + y_\sigma v_\sigma,$$

notamos además que el mismo procedimiento que hicimos con x lo podemos hacer con z . Observamos que:

$$[P_{v_1}x, z] = \sum_{\alpha \in \text{supp}(v)} x_\alpha [u_\alpha, z] + y_\alpha [v_\alpha, z] = \sum_{\alpha \in \text{supp}(v)} \alpha(z)(y_\alpha u_\alpha - x_\alpha v_\alpha),$$

donde usamos (5.3), luego

$$\| [P_{v_1}x, z] \|_F^2 = \sum_{\alpha \in \text{supp}(v)} \alpha(z)^2 (x_\alpha^2 + y_\alpha^2), \quad (5.9)$$

entonces $[P_{v_1}x, z] = 0$ si y solo si $\alpha(z)(x_\alpha^2 + y_\alpha^2) = 0$ para todo $\alpha \in \text{supp}(v)$. Por la ecuación (5.5), tenemos que

$$- \sum_{\alpha \in \text{supp}(v) \cap \text{supp}(z)} \alpha(z)\alpha(v)(x_\alpha^2 + y_\alpha^2) = \varphi([x, [x, v]]).$$

Luego, la condición $[P_{v_1}x, z] = 0$ implica que $\varphi([x, [x, v]]) = 0$. Por otro lado si $\varphi([x, [x, v]]) = 0$ entonces

$$0 = \sum_{\alpha} \alpha(z)\alpha(v)(x_\alpha^2 + y_\alpha^2),$$

como $\alpha(v)\alpha(z) \geq 0$ para todo α por la ecuación (5.6), debe ser

$$\alpha(v)\alpha(z)(x_\alpha^2 + y_\alpha^2) = 0 \quad \forall \alpha \in \Delta^+. \quad (5.10)$$

Para los $\alpha \in \text{supp}(v)$ podemos cancelar el factor $\alpha(v)$ y obtenemos que

$$\alpha(z)(x_\alpha^2 + y_\alpha^2) = 0 \quad \forall \alpha \in \text{supp}(v),$$

mirando la ecuación (5.9), se sigue que $[P_{v_1}x, z] = 0$.

Para ver (1) observamos que

$$\| [x, z] \|_F^2 = \sum_{\alpha \in \text{supp}(z)} \alpha(z)^2 (x_\alpha^2 + y_\alpha^2).$$

Si $\alpha(v) \neq 0$ entonces por (5.10) tenemos que $\alpha(z)(x_\alpha^2 + y_\alpha^2) = 0$, y entonces

$$\| [x, z] \|_F^2 = \sum_{\alpha \in \text{supp}(v)^c \cap \text{supp}(z)} \alpha(z)^2 (x_\alpha^2 + y_\alpha^2). \quad (5.11)$$

Si φ es un funcional adaptado a v , esta intersección es vacía y entonces $[x, z] = 0$. El ítem (2) se sigue de forma inmediata del Lema previo. Veamos (3), para ello suponemos que $F_\varphi = \{v\}$ y escribimos

$$\| [x, v] \|_F^2 = \sum_{\alpha \in \text{supp}(v)} \alpha(v)^2 (x_\alpha^2 + y_\alpha^2).$$

Sea $\alpha \in \Delta^+$, si $\alpha(v) = 0$, el término asociado a α no va a estar en la suma. Si $\alpha(v) \neq 0$, debe ser $\alpha(z) \neq 0$ (ya que sino por el Lema 5.3.29 la cara no sería un singleton), luego por la ecuación (5.10) tenemos que $x_\alpha^2 + y_\alpha^2 = 0$ y entonces el término que involucra a α va a dar cero en la suma. Se concluye que $[x, v] = 0$. $\square$

Observación 5.3.35. Notamos que el teorema previo es la versión generalizada del Teorema 3.1.13. Decidimos incluir ambas demostraciones en este trabajo ya que se puede observar como en el grupo $\mathcal{U}_n$ la demostración es mas elemental (pero mas larga), mientras que en la versión abstracta es claramente mas corta, pero utiliza toda la maquinaria del sistema de raíces asociado a una subálgebra de Cartan.

A continuación vamos a incluir los resultados relacionados con la curvatura seccional en un álgebra de Lie compacta con una norma de Finsler Ad-invariante. Dado que las demostraciones son análogas a las hechas en $\mathfrak{u}_n$ y pueden encontrarse en [48], decidimos no incluirlas en esta sección.

Definición 5.3.36. Sean $x, y \in \mathfrak{g}$, definimos

$$S(x, y) = 6|y - x|^2 \lim_{r \rightarrow 0^+} \frac{r|y - x| - \text{dist}(e^{rx}, e^{ry})}{r^2 \text{dist}(e^{rx}, e^{ry})}.$$

Proposición 5.3.37. Sean $x, y \in \mathfrak{g}$, entonces

$$\begin{aligned} \frac{1}{6}|y - x|S(x, y) &= \lim_{s \rightarrow 0^+} \frac{|y - x| - |(y - x) + \frac{1}{48}s[x + y, [x, y]]|}{s} \\ &= -\frac{1}{4} \max_{\varphi \in N_{y-x}} \varphi([y, [y, y - x]]) = -\frac{1}{4} \max_{\varphi \in N_{y-x}} \varphi([x, [x, y - x]]) \geq 0. \end{aligned}$$

Teorema 5.3.38. Si la distancia en G es estrictamente convexa, entonces $S(x, y) = 0$ implica $[x, y] = 0$.

Teorema 5.3.39 (Secciones planas). Sean $x, y \in \mathfrak{g}$. Consideramos las siguientes afirmaciones:

- (1) $\varphi([x, [x, y]]) = 0$ para todo φ que realiza la norma de $y - x$.
- (2) Para $s > 0$ suficientemente chico, el camino $s^{-1}BCH(sy, -sx)$ está dentro de cierta cara de la esfera F_φ de radio $|y - x|$ que contiene a $y - x$.
- (3) Para $s \geq 0$ suficientemente pequeño, $\text{dist}(e^{sx}, e^{sy}) = s|y - x|$.
- (4) $S(x, y) = 0$.
- (5) $\varphi([x, [x, y]]) = 0$ para cierto φ que realiza la norma de $y - x$.

Luego (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Rightarrow$ (4) $\Leftrightarrow$ (5). Además

- a) Si existe un único funcional que realiza la norma de $y - x$ (i.e. la norma es suave en $y - x$), luego todas las condiciones son equivalentes.
- b) Si en (4) tenemos que φ es un funcional adaptado a $y - x$, todas las condiciones son equivalentes.
- c) Si en (4) tenemos que $F_\varphi = \{y - x\}$ (en particular si la norma es estrictamente convexa), luego todas las condiciones son equivalentes a $[x, y] = 0$.

Corolario 5.3.40. Sean $x, y \in \mathfrak{g}$ tales que $y - x$ es regular entonces son equivalentes:

- (1) $S(x, y) = 0$.
- (2) $\text{dist}(e^{tx}, e^{ty}) = |ty - tx|$ para $t \geq 0$ suficientemente chico.

Observación 5.3.41. El objetivo para la continuación de este trabajo es estudiar espacios homogéneos de grupos con distancias bi-invariantes. Más específicamente, si G es grupo con una métrica de Finsler Ad-invariante y K es un subgrupo de G que actúa sobre G , de forma invariante a izquierda, podemos considerar la variedad $M \simeq G/K$. Si asumimos que la métrica que se considera en G es riemanniana, en [45] fue probado que una curva $\gamma \subseteq G$ es corta cuando la proyectamos en M si y solo si realiza la distancia entre fibras de elementos del cociente. La idea, mediante los resultados de curvatura probados en [48], es probar un resultado similar sin la hipótesis de que la métrica sea riemanniana.

Bibliografía

- [1] P.-A. Absil, R. Mahony, R. Sepulchre: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton (2008).
- [2] Ado, Igor D. (1949): The representation of Lie algebras by matrices, American Mathematical Society Translations, 1949 (2): 21
- [3] Alexandrino, Marcos M.; Bettiol, Renato G.: Lie groups and geometric aspects of isometric actions. Springer, Cham, 2015. x+213 pp. ISBN: 978-3-319-16612-4; 978-3-319-16613-1
- [4] T. Ando: Majorization, doubly stochastic matrices, and comparison of eigenvalues. Linear Algebra Appl. 118 (1989), 163–248.
- [5] Antezana J., Stojanoff D.: Análisis Matricial. Cursos y seminarios de matemática- Serie B (2009), Departamento de Matemática, Facultad de Ciencias Exactas, UBA.
- [6] E. Andruchow, G. Larotonda, L. Recht: Finsler geometry and actions of the p -Schatten unitary groups. Tran. Amer. Math. Soc. 362 (2010) no. 1, 319–344.
- [7] E. Andruchow, G. Larotonda, L. Recht, A. Varela: The left invariant metric in the general linear group. J. Geom. Phys. 86 (2014), 241–257.
- [8] E. Andruchow, L. Recht: Sectional curvature and commutation of pairs of selfadjoint operators. J. Operator Theory 55 (2006) no. 2, 225–238.
- [9] J. J. Benedetto, M. Fickus: Finite Normalized Tight Frames. Advances in Computational Mathematics 18, 357–385 (2003).
- [10] V. N. Berestovskii: Homogeneous manifolds with an intrinsic metric. I. (Russian) ; translated from Sibirsk. Mat. Zh. 29 (1988), no. 6, 17–29 Siberian Math. J. 29 (1988), no. 6, 887–897 (1989).
- [11] R. Bhatia and J. A. R. Holbrook: A softer, stronger Lidskii theorem. Proc. Indian Acad. Sci. Math. Sci. 99 (1989), no. 1, 75–83
- [12] E. Andruchow, L. Recht: Sectional curvature and commutation of pairs of selfadjoint operators. J. Operator Theory 55 (2006) no. 2, 225–238.
- [13] J. Antezana, G. Larotonda, A. Varela: Optimal paths for symmetric actions in the unitary group. Comm. Math. Phys. 328 (2014), no. 2, 481–497.

- [14] C. J. Atkin: The Finsler geometry of groups of isometries of Hilbert space. *J. Austral. Math. Soc. Ser. A* 42 (1987) no. 2, 196–222.
- [15] R. Bhatia: Matrix analysis. Graduate Texts in Mathematics 169. Springer-Verlag, New York, 1997.
- [16] Birkhoff, Garrett: Orthogonality in linear metric spaces. *Duke Math. J.* 1 (1935), no. 2, 169–172.
- [17] F.F. Bonsall, J. Duncan: Numerical ranges of operators on normed spaces and of elements of normed algebras. London Mathematical Society Lecture Note Series, 2. Cambridge University Press, London-New York, 1971.
- [18] N. Bourbaki: Lie groups and Lie algebras. Chapters 1-3. Translated from the French. Reprint of the 1989 English translation. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1998.
- [19] N. Bourbaki: Lie groups and Lie algebras. Chapters 7-9. Translated from the 1975 and 1982 French originals by Andrew Pressley. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2005.
- [20] A. Brøndsted: An introduction to the theory of convex polytopes. Translated from the English by K. I. Oskolkov. Translation edited and with a preface by B. S. Kashin. “Mir”, Moscow, 1988. 240 pp
- [21] D. Burago, Y. Burago, S. Ivanov: A course in metric geometry. Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, RI, 2001.
- [22] F. Casas, A. Murua: An efficient algorithm for computing the Baker-Campbell-Hausdorff series and some of its applications. *Journal of Mathematical Physics* 50, 033513, 2009.
- [23] Clarke, Francis: Functional analysis, calculus of variations and optimal control. Graduate Texts in Mathematics, 264. Springer, London, 2013.
- [24] C. Conde, G. Larotonda: Manifolds of semi-negative curvature. *Proc. Lond. Math. Soc.* (3) 100 (2010), no. 3, 670–704.
- [25] Coxeter, H.S.M. (1973): Regular Polytopes (3ed.). Dover Publications. 321p. ISBN 0-486-61480-8.
- [26] C. E. Durán, L. E. Mata-Lorenzo, L. Recht: Metric geometry in homogeneous spaces of the unitary group of a C^* -algebra. I. Minimal curves. *Adv. Math.* 184 (2004), no. 2, 342–366.
- [27] C. E. Durán, L. E. Mata-Lorenzo, L. Recht: Metric geometry in homogeneous spaces of the unitary group of a C^* -algebra. II. Geodesics joining fixed endpoints. *Integral Equations Operator Theory* 53 (2005), no. 1, 33–50.
- [28] Dynkin, Eugene Borisovich (1947): Calculation of the coefficients in the Campbell–Hausdorff formula. *Doklady Akademii Nauk SSSR* (in Russian). 57: 323–326.
- [29] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant, V. Zizler: Functional analysis and infinite-dimensional geometry. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 8. Springer-Verlag, New York, 2001.

- [30] M. Fickus, D. G. Mixon, J.C. Tremain: Steiner equiangular tight frames. *Linear Algebra and its Applications* **436** no. 5 (2012) 1014–1027.
- [31] R. Grone, M. Marcus: Isometries of matrix algebras. *J. Algebra* 47 (1977), no. 1, 180–189.
- [32] Hall, Brian: Lie groups, Lie algebras, and representations. An elementary introduction. Second edition. *Graduate Texts in Mathematics*, 222. Springer, Cham, 2015. xiv+449 pp. ISBN: 978-3-319-13466-6; 978-3-319-13467-3
- [33] M. Hayashi: Quantum information theory. Mathematical foundation. Second edition. *Graduate Texts in Physics*. Springer-Verlag, Berlin, 2017.
- [34] Helgason, Sigurdur: Differential geometry, Lie groups, and symmetric spaces. Corrected reprint of the 1978 original. *Graduate Studies in Mathematics*, 34. American Mathematical Society, Providence, RI, 2001. xxvi+641 pp. ISBN: 0-8218-2848-7
- [35] R. Hemasinha: Permutation bases. *International Journal of Mathematical Education in Science and Technology* 25 (1994), no.1, 103–111.
- [36] A. S. Holevo: Bounds for generalized uncertainty of the shift parameter. *Probability theory and mathematical statistics* (Tbilisi, 1982), 243–251, *Lecture Notes in Math.*, 1021, Springer, Berlin, 1983.
- [37] A. Horn: Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* 76 (1954), 620–630.
- [38] J.E. Humphreys: Introduction to Lie algebras and representation theory. Second printing, revised. *Graduate Texts in Mathematics*, 9. Springer-Verlag, New York-Berlin, 1978.
- [39] A. Jensen: Self-polar polytopes. *Polytopes and discrete geometry*, 101–124, *Contemp. Math.*, 764, Amer. Math. Soc., Providence RI, 2021.
- [40] A. W. Knap: Lie groups beyond an introduction. Second edition. *Progress in Mathematics*, 140. Birkhäuser Boston, Inc., Boston, MA, 2002.
- [41] G. Larotonda: Nonpositive Curvature: a Geometrical approach to Hilbert-Schmidt Operators. *Differential Geom. Appl.* 25 (2007), no. 6, 679–700.
- [42] G. Larotonda: Norm inequalities in operator ideals. *J. Funct. Anal.* 255 (2008), no. 11, 3208–3228.
- [43] G. Larotonda: Estructuras geométricas para las variedades de Banach. *Colección Ciencia Innovación y Tecnología*, Editorial Universidad Nacional de General Sarmiento.
- [44] G. Larotonda: The case of equality in Hölder’s inequality for matrices and operators. *Math. Proc. R. Ir. Acad.* 118A (2018), no. 1, 1–4.
- [45] G. Larotonda: Metric geometry of infinite-dimensional Lie groups and their homogeneous spaces. *Forum Math.* 31 (2019), no. 6, 1567–1605.

- [46] G. Larotonda, M. Miglioli: Hofer's metric for compact Lie groups. Preprint arXiv:1907.09843, submitted (2020).
- [47] G. Larotonda, I. Rey: Weakly invariant norms: geometry of spheres in the space of skew-Hermitian matrices. *Linear Algebra Appl.* 678 (2023), 136–168.
- [48] G. Larotonda, I. Rey: Lie groups with a bi-invariant distance (pre-print).
- [49] Lee, J.M.: *Riemannian Manifolds: An Introduction to Curvature*, Springer-Verlag, New York (1997).
- [50] Lee, Jeffrey M.: *Manifolds and differential geometry*. Graduate Studies in Mathematics, 107. American Mathematical Society, Providence, RI, 2009.
- [51] A. C. G. Mennucci: On asymmetric distances. *Anal. Geom. Metr. Spaces* 1 (2013), 200–231.
- [52] J. Malkoun, N. Nahlus: Commutators and Cartan subalgebras in Lie algebras of compact semisimple Lie groups. *J. Lie Theory* 27 (2017), no. 4, 1027–1032.
- [53] Meyer, Wolfgang: *Lecture Notes on Differential Geometry*. Lecture notes, College on Differential Geometry, ICTP, Trieste, 1989.
Link: [Toponogov's theorem and applications \(pdf\)](#).
- [54] J. Milnor: Morse theory. Based on lecture notes by M. Spivak and R. Wells. *Annals of Mathematical Studies* **51**, Princeton University Press, Princeton, N.J. 1963.
- [55] J. Milnor: Curvatures of left invariant metrics on Lie groups. *Advances in Math.* 21 (1976), no. 3, 293–329.
- [56] G. Muñoz, Y. Sarantopoulos, A. Tonge: Complexifications of real Banach spaces, polynomials and multilinear maps. *Studia Math.* 134 (1999), no. 1, 1–33.
- [57] M. A. Nielsen, I. L. Chuang: *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.
- [58] A. Pazy: *Semigroups of linear operators and applications to partial differential equations*. Applied Mathematical Sciences, 44. Springer-Verlag, New York, 1983.
- [59] H. Porta, L. Recht: Minimality of geodesics in Grassmann manifolds. *Proc. Amer. Math. Soc.* 100 (1987) no. 3, 464–466.
- [60] Price, John F.: *Lie groups and compact groups*. London Mathematical Society Lecture Note Series, No. 25. Cambridge University Press, Cambridge-New York-Melbourne, 1977.
- [61] W. Rudin: *Functional analysis*. Second edition. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991.
- [62] A. M. Sinclair: The norm of a Hermitian element in a Banach algebra. *Proc. Amer. Math. Soc.* 28, 446–450 (1971).

- [63] Shoda, Kenjiro: Einige Sätze über Matrizen. (German) [[Some theorems on matrices]] Jpn. J. Math. 13 (1937), no. 3, 361–365
- [64] J. A. Tropp: Just relax: convex programming methods for identifying sparse signals in noise. IEEE Transactions on Information Theory, vol. 52, no. 3, pp. 1030–1051.
- [65] H. Upmeyer: Symmetric Banach manifolds and Jordan C^* -algebras. North-Holland Mathematics Studies, 104. Notas de Matemática [Mathematical Notes], 96. North-Holland Publishing Co., Amsterdam, 1985.
- [66] G. A. Watson: Linear best approximation using a class of polyhedral norms. Numer. Algorithms 2 (1992), no. 3-4, 321–335.
- [67] G. A. Watson: On matrix approximation problems with Ky Fan k norms. Algorithms for approximation, III (Oxford, 1992). Numer. Algorithms 5 (1993), no. 1-4, 263–272.
- [68] K. Zietak: On the characterization of the extremal points of the unit sphere of matrices. Linear Algebra Appl. 106 (1988), 57–75.