María José Pereira Sáez

# APLICACIÓN TRAZA, TRANSFORMACIÓN DE CAYLEY Y CATEGORÍA LS DE LOS GRUPOS DE LIE CLÁSICOS

113
2008
Publicaciones
del
Departamento
de Geometría y Topología
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

María José Pereira Sáez

# APLICACIÓN TRAZA, TRANSFORMACIÓN DE CAYLEY Y CATEGORÍA LS DE LOS GRUPOS DE LIE CLÁSICOS

Memoria realizada en el Departamento de Geometría y Topología de la Facultad de Matemáticas, bajo la dirección del Prof. Dr. Enrique Macías Virgós, para obtener el Diploma de Estudios Avanzados en Ciencias Matemáticas por la Universidad de Santiago de Compostela.

Se llevó a cabo su defensa el día 10 de septiembre de 2008 en la Facultad de Matemáticas de dicha universidad, obteniendo la máxima calificación de Sobresaliente (10).

**IMPRIME:** Imprenta Universitaria
Pavillón de Servicios
Campus Universitario

**ISBN:** 978-84-89390-30-0

**Dep. Leg.:** C 4704-2008

*A mi familia*

# Índice general

# Introducción

El objetivo principal de esta memoria es desarrollar una idea de A. Gómez Tato y E. Macías Virgós para calcular la categoría LS del grupo unitario $U(n)$ mediante la transformación de Cayley. Esto nos da una demostración más sencilla que la clásica de W. Singhof [39].

La categoría LS es un invariante homotópico que ha sido muy estudiado desde su introducción en 1935 por Lusternik y Schnirelmann [25]. En variedades compactas $\operatorname{cat}(M)+1$ nos da una cota inferior para el número de puntos críticos de cualquier función diferenciable. Pueden además darse cotas para la categoría LS en términos de invariantes cohomológicos como la longitud del cup producto. Al mismo tiempo, la categoría sirve también para acotar la Complejidad Topológica (TC) de una variedad [14].

Desafortunadamente, la categoría LS es muy difícil de calcular. Por ejemplo, aunque la categoría de $Sp(2)$ fue calculada por P. Schweitzer en 1965, hubo que esperar al año 2002 hasta obtener la de $Sp(3)$. Esta dificultad ha tratado de paliarse introduciendo aproximaciones algebraicas como la categoría LS racional en homotopía racional. En general son técnicas difíciles.

La transformación de Cayley es muy conocida para los grupos ortogonales clásicos [37]. Se trata de la aplicación

$$c(A) = \frac{\mathrm{id} - A}{\mathrm{id} + A}$$

que está definida en el abierto $\Omega$ de las matrices que no tengan al $-1$ como autovalor y envía un grupo de Lie ortogonal $G$ en su álgebra de Lie $\mathfrak{g}$.

Puede pensarse como una generalización de la proyección estereográfica. Entre sus propiedades más interesantes está la de ser involutiva, es decir, su inversa es ella misma. Esto nos permite traducir la contracción radial del espacio vectorial $\mathfrak{g}$ en una contracción del abierto $\Omega$. Trasladando adecuadamente este último obtenemos un recubrimiento abierto del grupo unitario $U(n)$, lo que nos permite demostrar que su categoría LS es $n$.

Al profundizar en este método, descubrimos que está íntimamente ligado a la teoría de Morse en los grupos de Lie. Recordemos que una función $f$ en una variedad $G$ es de Morse si sus puntos críticos son no degenerados (y en consecuencia aislados). Clásicamente, las funciones de Morse que suelen considerarse en una variedad (embebida en un espacio euclidiano) son las funciones "altura" y "distancia". Siguiendo a Dynnikov y Veselov [41] probamos que, salvo una constante, en un grupo de Lie estas funciones son siempre de la forma $f(X) = \Re Tr(BX)$, la parte real de la traza de $X$ módulo una matriz arbitraria $B$. La función que resulta de considerar $B = \mathrm{id}$ fue estudiada por Frankel [17]. Este caso supone una ventaja ya que la función es invariante por la acción adjunta, y por tanto basta estudiar los puntos críticos en un toro maximal. El inconveniente es que se trata más bien de una función de Bott-Samelson, es decir, no hay puntos críticos aislados sino subvariedades críticas.

En los trabajos de Dynnikov y Veselov [41] y Haibao Duan [12], sin embargo, se considera el caso en que $B$ es una matriz diagonal real positiva $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, con $\lambda_1 < \ldots < \lambda_n$, para la que $f$ es una función de Morse "perfecta", es decir, tiene el menor número posible, $\#\Sigma_f$, de puntos críticos. Recordemos que según el teorema clásico de Lusternik y Schnirelmann, $\mathrm{cat}\, G + 1 \leq \#\Sigma_f$ para cualquier función diferenciable (aunque no sea de Morse). Esto es así porque, para cada punto crítico, el flujo del gradiente de $f$ determina un abierto contráctil. Pues bien, resulta que ese flujo del gradiente está dado precisamente por la contracción asociada a la aplicación de Cayley.

# Contenido

El desarrollo de esta memoria es el siguiente:

Dedicamos el capítulo 1 a los preliminares. Comenzamos dando la definición de grupo $J$-ortogonal, donde $J$ es una matriz tal que $J^2 = \pm\mathrm{id}$ y calculamos su álgebra de Lie. En estos grupos será donde haremos la construcción de Cayley del capítulo 4. Casos particulares serán los grupos ortogonal $O(n)$, unitario $U(n)$ y simpléctico $Sp(n)$, que también denotaremos por $O_n(\mathbb{K})$, con $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ ó $\mathbb{H}$ respectivamente. Incluimos en este capítulo una sección (1.2) en la que estudiamos el álgebra lineal de las matrices sobre los cuaternios $\mathbb{H}$, pues la no conmutatividad proporciona algunas sorpresas (por ejemplo, no es cierto que para el producto de matrices se tenga $(AB)^t = B^t A^t$). En particular, para que las fórmulas habituales de la matriz asociada por columnas sigan funcionando, debe considerarse en $\mathbb{H}^n$ la estructura de $\mathbb{H}$-espacio vectorial por la derecha. Es especialmente importante discutir (teniendo en cuenta que no tenemos el determinante usual) la existencia y número de autovalores por la derecha y la diagonalización de matrices normales. Sería también de interés discutir más a fondo los autovalores por la izquierda, pero hemos encontrado muy pocos resultados en la literatura de este campo [22, 30]. La última sección contiene un breve resumen de la teoría de Morse siguiendo el libro de Milnor [35].

En el capítulo 2 estudiamos dos tipos de funciones de Morse en subvariedades del espacio euclidiano. En primer lugar, las funciones "distancia" dadas por $f_a(x) = |x - a|^2$ y probamos que son de Morse para casi todo $a$. Consideramos a continuación (sección 3.2) los grupos ortogonales $G = O_n(\mathbb{K})$ como subvariedades de $E = \mathcal{M}_{n\times n}(\mathbb{K})$, con lo que el producto escalar usual puede escribirse como $\langle M, N\rangle = \Re Tr M^* N$. Entonces, el tangente $T_I G$ y el ortogonal $\gamma_I G$ están formados respectivamente por las matrices simétricas (hermíticas) y antisimétricas (antihermíticas), mientras que en cualquier otro punto $X \in G$ basta trasladar por la isometría $L_X$. Esto nos servirá para calcular el gradiente de cualquier función.

A lo largo de la sección 2.3 seguimos el artículo de Haibao Duan [12]. Para obtener que $f_a$ es una función de Morse hay que suponer que $a$ es una matriz diagonal real positiva $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$, con $\lambda_1 < \cdots < \lambda_n$. Entonces probamos (proposición 2.3.1) que sus puntos críticos son las matrices diagonales $X = \mathrm{diag}(\varepsilon_1, \dots, \varepsilon_n)$, con $\varepsilon_k = \pm 1$, y que el Hessiano viene dado (lema 2.3.2) por

$$H_X(f_B)(U) = (1/2)X(BU - U^*B).$$

En la proposición 2.3.4 calculamos el índice de cada punto crítico.

En realidad las funciones "distancia" consideradas hasta ahora son, salvo una constante, funciones "altura". En efecto, éstas son necesariamente la parte real de la traza, $f_B(X) = \Re Tr(BX)$, para alguna matriz $B$. Siguiendo el artículo de Dynnikov y Veselov [41] probamos que el gradiente viene dado por $\mathrm{grad}_B(X) = B^* - XBX$ y que por tanto en 2.4.3 los puntos críticos son las matrices $X \in G$ tales que $BX$ es hermítica.

El caso $B = \mathrm{id}$ fue estudiado por Frankel en [17]. Esta vez se trata de una función de Bott-Samelson. Obtenemos también el Hessiano y el índice en un punto crítico para el caso en que $B$ es una matriz diagonal apropiada.

El capítulo 3 está dedicado a la categoría LS de un espacio topológico $X$. Este invariante, $\mathrm{cat}(X)$, nos da el menor número entero $n$ tal que existe un recubrimiento de $X$ por $n+1$ abiertos, cada uno de ellos contráctil en $X$.

En la sección 3.1 damos las definiciones básicas y recordamos el teorema clásico de Lusternik y Schnirelman (3.2.1): en una variedad compacta la categoría es siempre menor o igual que el número de puntos críticos menos 1 de cualquier función diferenciable $f\colon M \longrightarrow \mathbb{R}$.

Recordamos también la definición de la longitud del cup producto en cohomología pues nos sirve como cota inferior. Se tiene que $l.c.p.(X) \leq \mathrm{cat}(X)$.

Abordamos el estudio de la transformación de Cayley en el capítulo 4. Sea $E$ el espacio euclidiano $\mathcal{M}_{n\times n}(\mathbb{K})$ y consideremos el abierto $\Omega$ de las matrices $X$ tales que $\mathrm{id}+X$ es inversible. Entonces está bien definida $c\colon \Omega \to \Omega$ dada por $c(X) = (\mathrm{id}-X)/(\mathrm{id}+X)^{-1}$.

En la sección 4.2 damos las propiedades básicas de la aplicación $c$ y fórmulas tales como $c(A^{-1}) = -c(A)$ ó $c(A^*) = c(A)^*$ que nos permitirán hacer cálculos posteriormente. En particular probamos que $c^2 = \mathrm{id}$.

La sección 4.3 nos sirve para probar que la transformación de Cayley $c$ envía el grupo de Lie $G$ en su álgebra de Lie $\mathfrak{g}$ (y viceversa), siempre que no salgamos de su dominio de definición

$$\Omega = \{X \in \mathcal{M}_{n\times n}(\mathbb{K})\colon I + X \text{ es inversible}\}.$$

A continuación (teorema 4.3.4) probamos que cuando $G = O_n(\mathbb{K})$ se tiene $\mathfrak{g} \subset \Omega$, con lo que el abierto $O_n(\mathbb{K}) \cap \Omega$ es contráctil. Será este abierto el que usaremos para recubrir el grupo y calcular su categoría LS.

Esta sección se cierra con varios ejemplos.

A lo largo de la sección 4.4 hacemos el cálculo del desarrollo en serie de la transformación de Cayley y estudiamos su radio de convergencia. Para ello utilizamos la serie de Taylor en espacios de Banach tal como puede verse en [3]. Es necesario primero reescribir las fórmulas de derivación de un producto y un cociente en el caso no conmutativo (lemas 4.4.1 y 4.4.2) y finalmente obtenemos las series

$$c(H) = \mathrm{id} - 2H + 2H^2 - 2H^3 + \cdots, \quad \text{para } H \text{ próximo a 0 y}$$

$$c(\mathrm{id}+H) = -\frac{1}{2}H + \frac{1}{4}H^2 - \frac{1}{8}H^3 + \cdots, \quad \text{para } H \text{ próximo a 0.}$$

Para el radio de convergencia de series matriciales usamos la definición que da G. Choquet en [7].

Para finalizar este capítulo, en la sección 4.5 generalizamos la transformación de Cayley para tener una aplicación

$$c_A\colon \Omega(A) \to \Omega(A^*)$$

dada por

$$c_A(X) = (\mathrm{id} - A^*X)(A+X)^{-1}$$

donde

$$\Omega(A) = \{X\colon A + X \text{ es inversible}\}$$

y estudiamos sus propiedades, que son análogas al caso $A = \mathrm{id}$. Del mismo modo, probamos (teorema 4.5.4) que $\Omega(A) \cap O_n(\mathbb{K})$ es un abierto contráctil.

En el último capítulo vamos a aplicar todos los resultados anteriores. En primer lugar vemos que para resolver la ecuación del gradiente de la función "altura" $f_B$

$$X' = \frac{1}{2}(B - XBX) \tag{1}$$

podemos transformarla, gracias a la aplicación de Cayley, en la ecuación

$$Y' = -2(BY + YB). \tag{2}$$

Este proceso de paso de $G$ a $\mathfrak{g}$ es denominado "linealización" por Dynnikov y Veselov en [41].

La ecuación (2) se resuelve fácilmente como

$$Y(t) = e^{-tB} Y_0 e^{-tB}$$

y aplicando de nuevo $c$ obtenemos como solución de (1)

$$X(t) = (\sinh(At) + \cosh(At)X_0) \cdot (\cosh(At) + \sinh(At)X_0)^{-1}.$$

En la sección 5.2 mostramos cómo se puede calcular la categoría de $U(n)$ obteniendo explícitamente un recubrimiento por abiertos categóricos. La idea es como sigue: tomamos los abiertos $\Omega(A_0), \dots, \Omega(A_n)$, donde $A_k = z_k\mathrm{id}$, para diferentes números complejos $z_0, \dots, z_n$ con $|z_k| = 1$. Puesto que una matriz unitaria $A$ no puede tener más de $n$ autovalores distintos (y de módulo 1), alguna de las matrices $A_k + A$ debe ser inversible, es decir $A \in \Omega(A_k)$, con lo que hemos obtenido un recubrimiento por abiertos contráctiles. Por tanto $\mathrm{cat}(U(n)) \leq n$. Por otra parte, como la cohomología de $U(n)$ es $\Lambda(x_1, \dots, x_n)$, la longitud del cup producto es $n$, y por tanto $\mathrm{cat}\, U(n) = n$.

En el caso de $Sp(n)$ no podemos aplicar este razonamiento, pues sobre los cuaternios no es equivalente que la matriz $q\mathrm{id} + A$ sea inversible a que $q$ sea un autovalor de $A$.

Sin embargo, en el caso de $Sp(2)$ (una variedad de dimensión 10) todavía podemos probar explícitamente que los cuatro abiertos $\Omega(P_{ij})$ lo recubren, donde $P_{ij}$ es una matriz diagonal con $\pm 1$ en la diagonal. Por tanto $\mathrm{cat}\, Sp(2) \leq 3$. Que es exactamente 3 no puede deducirse de la longitud del cup producto (que es 2) pero sí de un resultado de Singhof y Schweitzer [39] que asegura que $\mathrm{cat}\, Sp(n) \geq n + 1$.

El caso $Sp(1)$ es trivial ya que $\mathrm{cat}\, S^3 = 1$.

Una aplicación de la "linealización" que vimos en la primera sección de este capítulo es un método para obtener la descomposición polar de una matriz [32] como explicitamos en la sección 5.4.

# Agradecimientos

La idea fundamental de este trabajo la debo al profesor E. Macías Virgós. Sin toda su dedicación y su labor de guía a través de ese mundo entre la Topología, el Álgebra y la Geometría que forman los grupos de Lie, probablemente esta memoria no habría salido a la luz. Especialmente importante ha sido la colaboración del profesor M. A. Fugarolas Villamarín en lo que se refiere a las series de Taylor sobre un álgebra normada y su radio de convergencia. Los últimos dibujos realizados con el Illustrator se los debo al profesor J. Barja Pérez.

No quería dejar de agradecer a Silvia Vilariño su paciencia para resolver los problemas con LateX, sobre todo a lo largo de este mes de agosto en que la facultad estaba desierta. Gracias también a María Pérez porque al pedirme que le explicase en qué estaba trabajando me exigía a mí comprenderlo mejor.

Antes de terminar me gustaría recordar a las personas concretas que, desde fuera del ámbito de la Facultad, han hecho posible que este trabajo llegara a buen puerto. En especial a mi madre, que sin querer llevarme por su mismo camino suscitó desde pequeña en mí el gusto por la belleza, el orden y la armonía de las Matemáticas. Y a mi padre, maestro en la tarea investigadora.

# Capítulo 1

# Preliminares

## 1.1. El grupo $O_n(J,\mathbb{K})$ de matrices J-ortogonales

En general una matriz $A$ es *J-ortogonal* si $AJA^* = J$. Tomamos $J$ tal que $J^2 = \pm \mathrm{id}$ y denotamos al conjunto de las matrices $J$-ortogonales por

$$O_n(J,\mathbb{K}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{K}) \text{ tales que } AJA^* = J\}$$

siendo $\mathbb{K}$ un álgebra asociativa con uno, $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ó $\mathbb{H}$.

**Lema 1.1.1** $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ *es J-ortogonal si y sólo si* $A^*$ *también lo es.*

*Demostración.–*
Al estar $A \in O_n(J,\mathbb{K})$ se tiene que $A^* = J^{-1}A^{-1}J$ luego

$$A^*JA = (J^{-1}A^{-1}J)JA = \pm J^{-1} = J.$$

Además, como $A = (A^*)^*$, queda probada la equivalencia.

*q.e.d.*

### 1.1.1. Álgebra de Lie

Calculamos su álgebra de Lie:

$$G = O_n(J,\mathbb{K}) \subset \mathcal{M}_{n\times n}(\mathbb{K})$$

subgrupo cerrado del grupo lineal general $GL(n,\mathbb{K})$. Así, por el teorema de Cartan, su álgebra de Lie vendrá dada por

$$\mathfrak{g} = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) : e^{tX} \in G\,, \forall\, t \in \mathbb{R}\}.$$

**Proposición 1.1.2** $\mathfrak{g} = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) :\ JX + X^*J = 0\}$.

*Demostración.–*
$X \in \mathrm{Lie}(G)$ si y sólo si $e^{tX} \in G\ \forall\, t \in \mathbb{R}$, es decir, si y sólo si $e^{tX}J(e^{tX})^* = J$ $\forall\, t \in \mathbb{R}$.

Ahora bien, la expresión $JX + X^*J = 0$ la podemos expresar como $JX = -X^*J$ o, lo que es lo mismo, $JXJ^{-1} = -X^*$. Esta condición es equivalente a

$$e^{tJXJ^{-1}} = e^{-tX^*}$$

si y sólo si

$$Je^{tX}J^{-1} = e^{-tX^*}$$

si y sólo si

$$Je^{tX}J^{-1}e^{tX^*} = \mathrm{id}$$

o, equivalentemente,

$$e^{tX}J^{-1}e^{tX^*} = J^{-1}.$$

Si $J^2 = \pm\mathrm{id}$, $J^{-1} = \pm J$ y por tanto $e^{tX}J(e^{tX})^* = J$.

*q.e.d.*

*Observación 1.—* Hemos utilizado que para cualquier matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$, $(e^A)^* = e^{A^*}$. Esto es es cierto para las matrices con coeficientes reales o complejos al ser ambos cuerpos conmutativos, lo comprobamos para las matrices cuaterniónicas.

Al escribir el desarrollo en serie de la exponencial de matrices observamos que lo único necesario para que $(e^A)^* = e^{A^*}$ es que $(A^n)^* = (A^*)^n$ pero si $\mathbb{K} = \mathbb{H}$, $(AB)^* = B^*A^*$.

*Observación 2.—* Aunque en general el álgebra $\mathbb{K}$ no sea conmutativa, es fácil comprobar que podemos poner como condición de pertenencia a $\mathfrak{g}$ que $X$ verifique tanto $JX + X^*J = 0$ como $XJ + JX^* = 0$, utilizando en este último caso que $-X = J^{-1}X^*J^{-1}$.

De esta observación se sigue el siguiente lema:

**Lema 1.1.3** *Dada $X \in \mathcal{M}_{n\times n}(\mathbb{K})$, se verifica que $X \in \mathfrak{g}$ si y sólo si $X^* \in \mathfrak{g}$.*

### 1.1.2. Ejemplos

**Ejemplo 1** $J$=id.

- $O_n(J, \mathbb{R}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) : AA^t = \mathrm{id}\}$ grupo ortogonal de orden $n$.

- $O_n(J,\mathbb{C}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{C}) : AA^* = \mathrm{id}\}$ grupo unitario de orden $n$.
- $O_n(J,\mathbb{H}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{H}) : AA^* = \mathrm{id}\}$ grupo unitario simpléctico de orden $n$.

| | $\mathbb{R}$ | $\mathbb{C}$ | $\mathbb{H}$ |
|---|---|---|---|
| | $O(n)$ | $U(n)$ | $Sp(n)$ |
| det=1 | $SO(n)$ | $SU(n)$ | |

Obtenemos entonces

- $\mathfrak{o}(n) = \mathfrak{so}(n) = \{X \in \mathcal{M}_{n\times n}(\mathbb{R}) : X + X^t = 0\}$, matrices antisimétricas,
- $\mathfrak{u}(n) = \{X \in \mathcal{M}_{n\times n}(\mathbb{C}) : X + X^* = 0\}$, antihermíticas complejas,
- $\mathfrak{sp}(n) = \{X \in \mathcal{M}_{n\times n}(\mathbb{H}) : X + X^* = 0\}$, antihermíticas cuaterniónicas y
- $\mathfrak{su}(n) = \{X \in \mathcal{M}_{n\times n}(\mathbb{H}) : X + X^* = 0 \text{ y } Tr(X) = 0\}$, antihermíticas cuaterniónicas de traza cero (este último álgebra de Lie no es propiamente de un grupo $J$-ortogonal sino que corresponde a un subgrupo de Lie de un grupo $J$-ortogonal).

**Ejemplo 2** $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $J^2 = -\mathrm{id}$.

Se comprueba fácilmente que en este caso $O_2(J,\mathbb{R})$ es el grupo lineal especial de orden 2,

$$O_2(J,\mathbb{R}) = \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) : \det A = 1\} = SL(2,\mathbb{R})$$

y $\mathfrak{sl}(2,\mathbb{R})$ está formado por las matrices reales de orden dos que tienen traza nula.

**Ejemplo 3** $J = \left(\begin{array}{c|c} I_p & \\ \hline & -I_q \end{array}\right)$, $J^2 = \mathrm{id}$.

| $\mathbb{R}$ | $\mathbb{C}$ |
|---|---|
| $O(p,q)$ | $U(p,q)$ |

Sus álgebras de Lie están formadas por matrices antisimétricas o antihermíticas en las cajas y simétricas o hermíticas fuera de ellas.

**Ejemplo 4** $J = \left(\begin{array}{c|c} 0 & I_n \\ \hline -I_n & 0 \end{array}\right)$, $J^2 = -\mathrm{id}$.

Para esta matriz $J$ obtenemos como grupo $J$-ortogonal al grupo simpléctico real de orden $2n$.

$$O_{2n}(J,\mathbb{R}) = \{X \in \mathcal{M}_{2n}(\mathbb{R}) : X^t JX = J\} = Sp(2n,\mathbb{R})$$

y su álgebra de Lie es

$$\mathfrak{sp}(2n,\mathbb{R}) = \{X \in \mathcal{M}_{2n}(\mathbb{R}) : JX + X^t J = 0\}.$$

## 1.2. Matrices cuaterniónicas

Es bien sabido que el cuerpo de los cuaternios no es conmutativo. Esto lleva consigo que no podamos utilizar los resultados habituales en lo referente a la diagonalización de matrices. Hacemos a continuación un breve estudio de las propiedades que verifican en este ámbito las matrices simplécticas.

### 1.2.1. Autovalores por la derecha

Aun cuando en el caso cuaterniónico no podemos definir el determinante de la misma manera que sobre $\mathbb{R}$ y $\mathbb{C}$ se define el concepto de autovalor de manera similar a una de las caracterizaciones de los casos real y complejo.

**Definición 1.2.1** *Se dice que un cuaternio $q \in \mathbb{H}$ es un* autovalor por la derecha *de $A \in \mathcal{M}_{n\times n}(\mathbb{H})$ si existe un vector $v \in \mathbb{H}^n, v \neq 0$ tal que $Av = vq$.*

Donde $A$ puede considerarse como la matriz asociada a una aplicación $\mathbb{H}$-lineal (por la derecha), $\varphi : \mathbb{H}^n \longrightarrow \mathbb{H}^n$.

Recordemos que para poder usar la matriz asociada de la manera usual, es decir, por columnas, es necesario poner los escalares por la derecha .

**Lema 1.2.2** *Sea $A \in O_n(\mathbb{K})$ y $\lambda$ autovalor de $A$. Se verifica que $|\lambda| = 1$.*

*Demostración.–*

Si $A \in G$, $A$ conserva el producto escalar $\langle v, w\rangle = w^*v$ :

$$\langle Av, Aw\rangle = (Aw)^*Av = w^*A^*Av = w^*v.$$

Sea $v$ un autovector asociado al autovalor $\lambda$,

$$\langle v, v\rangle = \langle Av, Av\rangle = \langle v\lambda, v\lambda\rangle = \langle v, v\rangle|\lambda|^2$$

luego $|\lambda| = 1$.

*q.e.d.*

Sin embargo, con esta definición el conjunto de autovectores asociado a un autovalor $q$, $\Sigma(q)$, no es un subespacio vectorial pues, dados $v, v' \in \Sigma(q)$, la suma se mantiene dentro de $\Sigma(q)$ pero el producto por escalares no:

$$\varphi(v\lambda) = \varphi(v)\lambda = (vq)\lambda.$$

A pesar de esto tenemos el siguiente resultado:

**Proposición 1.2.3** *Sea $\varphi \in O_n(\mathbb{H})$. Si dos autovalores $\lambda, \lambda'$ de $\varphi$ no son conjugados, entonces los autovectores correspondientes son perpendiculares (para el producto hermítico) y, en consecuencia, son linealmente independientes.*

*Demostración.–*
Sean $v, v'$ dos vectores no nulos y $\lambda, \lambda'$ dos autovalores de $\varphi$ con $|\lambda| = |\lambda'| = 1$,

$$\begin{aligned} \varphi(v) &= v\lambda, \\ \varphi(v') &= v'\lambda'. \end{aligned}$$

Tenemos entonces

$$\langle v, v' \rangle = \langle \varphi(v), \varphi(v') \rangle = \langle v\lambda, v'\lambda' \rangle = \overline{\lambda} v^* v' \lambda' = \overline{\lambda} \langle v, v' \rangle \lambda'.$$

Si $\langle v, v' \rangle \neq 0$ entonces

$$\langle v, v' \rangle^{-1} \lambda \langle v, v' \rangle = \lambda',$$

es decir, $\lambda$ y $\lambda'$ serían conjugados.

*q.e.d.*

**Corolario 1.2.4** *$\varphi$ no puede tener más de $n$ autovalores no conjugados.*

### 1.2.2. Diagonalización

Probemos ahora que toda matriz del grupo unitario simpléctico puede diagonalizarse. Utilizaremos la siguiente generalización del lema de Schur probada por Brenner en [4, p.331] .

**Teorema 1.2.5** *Toda matriz cuaterniónica es triangularizable por una matriz unitaria.*

Es decir, dada $X \in \mathcal{M}_{n\times n}(\mathbb{H})$ existe una matriz $T \in \mathcal{M}_{n\times n}(\mathbb{H})$ triangular superior tal que $X = UTU^*$ con $U \in Sp(n)$.

**Teorema 1.2.6** *Toda matriz cuaterniónica unitaria es diagonalizable.*

*Demostración.–*
Sea $A \in Sp(n)$. Por el teorema anterior, $A = UTU^*$ con $T$ una matriz triangular superior y $U \in Sp(n)$. Despejando $T$ en la igualdad anterior, $T = U^*AU$ y así,

$$\begin{aligned} TT^* = (U^*AU)(U^*A^*U) = U^*AA^*U &= \\ U^*A^*UU^*A^*U = (U^*AU)^*(U^*AU) &= T^*T. \end{aligned}$$

Es decir, $T$ es una matriz normal. Además, como $T$ es triangular superior, de esta igualdad obtenemos para $i = 1, \ldots, n$ que

$$|t_{ii}|^2 = |t_{ii}|^2 + |t_{ii+1}|^2 + \ldots + |t_{in}|^2$$

por tanto, $t_{ij} = 0$ siempre que $j > i$, luego $T$ es diagonal.

*q.e.d.*

*Nota.*— La misma demostración vale tanto para el caso complejo como para la diagonalización de cualquier matriz normal.

Más aún, veremos a continuación una serie de propiedades de los cuaternios que nos permiten comprobar que las matrices simplécticas no sólo son diagonalizables sino que pueden diagonalizarse a una matriz compleja.

Sea $q \in \mathbb{H}$. La conjugación $i_q \colon \mathbb{H} \to \mathbb{H}$ conserva el módulo,

$$|i_q(h)| = |qhq^{-1}| = |q||h||q^{-1}| = |q||q^{-1}||h| = |h|.$$

En consecuencia, conserva el producto escalar usual de $\mathbb{R}^4$ (no así el producto hermítico de $\mathbb{H}$).

Como $i_q$ conserva los reales, también conserva $\langle 1\rangle^\perp = \langle i, j, k\rangle \cong \mathbb{R}^3$ luego $i_q$ es una isometría de $\mathbb{R}^3$. Es decir, tenemos una aplicación

$$q \in \mathbb{H} \mapsto i_q \in O(3).$$

Además, por conexión, la imagen está contenida en $SO(3)$. De hecho como la conjugación por $q$ y por $q/|q|$ es la misma, tenemos un morfismo de grupos de Lie (cubierta universal)

$$Sp(1) = S^3 \to SO(3) \tag{1.1}$$

cuyo núcleo es $\pm 1$. Para ver que es sobreyectivo puede usarse el hecho de que siempre podemos escribir un cuaternio $h$ de la forma

$$h = \Re(h) + |\Im(h)|\omega, \tag{1.2}$$

donde $\omega \in \langle i, j, k\rangle \cong \mathbb{R}^3$ tiene módulo 1, es decir $\omega \in S^2$. Si $|q| = 1$, tendremos

$$q = \cos\theta + |\operatorname{sen}\theta|\omega.$$

Entonces la rotación correspondiente a $i_q$ es la que tiene por eje de rotación $\langle\omega\rangle$ y por ángulo $2\theta$.

**Proposición 1.2.7** *Dos cuaternios son conjugados si y sólo si tienen el mismo módulo y la misma parte real.*

*Demostración.–*

Ya hemos visto que la conjugación conserva el módulo. Veamos que también conserva la parte real,

$$i_q(\Re(h)) = i_q(\frac{1}{2}(h + \overline{h})) = \frac{1}{2}(i_q(h) + \overline{i_q(h)}) = \Re(i_q(h)),$$

porque $q^{-1} = \overline{q}/|q|^2$ y $\overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}$.

Recíprocamente, por la fórmula (1.2), nos bastará probar que dos elementos $\omega, \omega' \in S^2$ son siempre conjugados, lo que se deduce de la sobreyectividad del morfismo (1.1).

*q.e.d.*

**Corolario 1.2.8** *Todo cuaternio $q \in \mathbb{H}$ es conjugado a un complejo $\Re(q) \pm i|\Im(q)|$.*

**Proposición 1.2.9** *Dado $q \in \mathbb{H}$ autovalor de $A \in \mathcal{M}_{n\times n}(\mathbb{H})$, cualquier conjugado $xqx^{-1}$ de $q$ es también un autovalor de $A$.*

*Demostración.–*

$$A(vx^{-1}) = (Av)x^{-1} = vqx^{-1} = (vx^{-1})xqx^{-1}.$$

*q.e.d.*

De esta proposición y del corolario 1.2.8 se deduce el resultado que buscábamos.

**Corolario 1.2.10** *Toda matriz simpléctica $A \in Sp(n)$ puede expresarse como $A = UDU^*$ donde $D = \operatorname{diag}(\lambda_1, \dots \lambda_n)$ con $\lambda_i \in \mathbb{C}$ y $|\lambda_i| = 1$.*

### 1.2.3. Determinante cuaterniónico

A pesar de no ser $\mathbb{H}$ conmutativo, ha habido varios intentos de extender el determinante a $\mathcal{M}_{n\times n}(\mathbb{H})$ con funcionales multiplicativos de $\mathbb{H}^{n\times n} \to \mathbb{H}$ que coincidan con el determinante para las matrices complejas. N. Cohen y S. De Leo los recogen en [9]. Se observa que no es posible generalizar a los cuaternios la noción de determinante complejo, sin embargo sí se puede generalizar el módulo del determinante, $|\det|$ (es decir, con valores reales). Se hace mediante el *determinante de Study*

$$\mathrm{Sdet}\colon \mathcal{M}_{n\times n}(\mathbb{H}) \to \mathbb{R}.$$

Para definirlo, consideramos cada matriz cuaterniónica como $M = A + jB$, descompuesta en matrices $A, B \in \mathcal{M}_{n\times n}(\mathbb{C})$. Si $M$ es la matriz asociada a una aplicación $\mathbb{H}$-lineal por la derecha $\mu\colon \mathbb{H}^n \to \mathbb{H}^n$ respecto de la base canónica $\{e_1, \ldots, e_n\}$, podemos considerar la misma $\mu$ como aplicación $\mathbb{C}$-lineal $\tilde{\mu}\colon \mathbb{C}^{2n} \to \mathbb{C}^{2n}$. Esta vez la matriz asociada a $\tilde{\mu}$ respecto de la base $\{e_1, \ldots, e_n, e_1 j, \ldots, e_n j\}$ es

$$\tilde{M} = \begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix}.$$

**Definición 1.2.11** $\mathrm{Sdet}(M) := |\det(\tilde{M})|^{\frac{1}{2}}$.

*Nota.—* Esta definición está motivada por el caso complejo. Si $A = P + iQ$ está descompuesta en matrices $P, Q \in \mathcal{M}_{n\times n}(\mathbb{R})$ y llamamos

$$\tilde{A} = \begin{pmatrix} P & -Q \\ Q & P \end{pmatrix}$$

entonces

$$\det_{\mathbb{R}}(\tilde{A}) = |\det_{\mathbb{C}}(A)|^2,$$

como se comprueba haciendo las siguientes transformaciones for filas y columnas

$$\tilde{A} \sim \begin{pmatrix} P + iQ & -Q \\ Q - iP & P \end{pmatrix} \sim \begin{pmatrix} P + iQ & -Q \\ 0 & P - iQ \end{pmatrix}.$$

**Proposición 1.2.12** *El determinante de Study verifica las siguientes propiedades:*

1. $\mathrm{Sdet}(M) \geq 0$*;*
2. $\mathrm{Sdet}(M) > 0$ *si y sólo si $M$ es inversible;*
3. $\mathrm{Sdet}(MN) = \mathrm{Sdet}(M)\,\mathrm{Sdet}(N)$*;*

*4.* $\mathrm{Sdet}(A) = |\det(A)|$ *si la matriz* $A$ *es compleja.*

De hecho puede probarse que es el único funcional que verifica estas propiedades (véase [9]). Además de las que acabamos de ver estas otras facilitan mucho su cálculo.

**Proposición 1.2.13** *Dada* $M$ *una matriz cuaterniónica de orden* $n$ *se tiene que:*

1. *Si* $E_{ij}$, $i \neq j$, *es una matriz elemental con todas las entradas cero excepto un* $1$ *en la posición* $(i, j)$, *entonces* $\mathrm{Sdet}(I + qE_{ij}) = 1$ $\forall q \in \mathbb{H}$*;*

2. $\mathrm{Sdet}(M)$ *no cambia si a una columna se le suma un múltiplo (por la derecha) de otra columna;*

3. $\mathrm{Sdet}(M)$ *no cambia si a una fila se le suma un múltiplo (por la izquierda) de otra fila;*

4. *Si* $M = (m_{ij})$ *es una matriz triangular superior entonces* $\mathrm{Sdet}(M) = |m_{11} \cdots m_{nn}|$*;*

5. *En particular, si* $M$ *es una matriz con autovalores (por la derecha)* $\lambda_1, \ldots, \lambda_n$ *entonces* $\mathrm{Sdet}(M) = |\lambda_1 \cdots \lambda_n|$*;*

6. $\mathrm{Sdet}(M^*) = \mathrm{Sdet}(M)$*;*

7. *En particular, si* $M$ *es simpléctica entonces* $\mathrm{Sdet}(M) = 1$.

### 1.2.4. Autovalores por la izquierda

La teoría de autovalores por la izquierda está muy poco desarrollada [22, 30]. Una de las dificultades es que dos matrices similares $A$ y $UAU^{-1}$ no tienen por qué tener los mismos autovalores. Aún así es interesante notar que si definimos los autovalores cuaterniónicos por la izquierda obtenemos más propiedades para los autovectores que al hacerlo por la derecha. Sin embargo, en este caso no es sencillo ver cuántos autovalores puede tener asociados una aplicación $\mathbb{H}$-lineal.

**Definición 1.2.14** *Sea* $\varphi$ *una aplicación* $\mathbb{H}$*-lineal. Se dice que* $q \in \mathbb{H}$ *es un* autovalor por la izquierda *de* $\varphi$ *si existe un* $w \neq 0$ *tal que* $\varphi(w) = qw$.

En este caso:

**Proposición 1.2.15** *El conjunto de los autovectores asociados a un autovalor por la izquierda* $q$, $\Sigma(q)$, *forman un subespacio vectorial.*

*Demostración.–*

Sean $\lambda$ un autovalor, $w, w' \in \Sigma(\lambda)$ y $q \in \mathbb{H}$. Se verifica entonces:

$$\varphi(wq) = \varphi(w)q = (\lambda w)q = \lambda(wq),$$
$$\varphi(w + w') = \varphi(w) + \varphi(w') = \lambda(w + w').$$

*q.e.d.*

**Proposición 1.2.16** *Los autovalores por la izquierda de una matriz simpléctica tienen módulo* 1*.*

*Demostración.–*

Sea $A \in Sp(n)$. Si $Av = qv$ entonces

$$0 < |v|^2 = \langle Av, Av\rangle = \langle qv, qv\rangle = v^*\overline{q}qv = v^*|q|^2v = |q|^2|v|^2.$$

*q.e.d.*

Sea $M$ una matriz simpléctica. Entonces es claro que $q$ es un autovalor por la izquierda de $M$ si y sólo si $M - qI$ es inversible. Por tanto,

**Proposición 1.2.17** *Los autovalores (por la izquierda) de $M$ son las raíces de la ecuación* $\mathrm{Sdet}(M - qI) = 0$*.*

En [22, 30] pueden verse dos métodos diferentes para resolver esta ecuación en el caso $2 \times 2$. De hecho, en [22] Huang y So prueban el siguiente resultado:

**Proposición 1.2.18** *Sea $M \in \mathcal{M}_2(\mathbb{H})$. En este caso se verifica que $M$ tiene o bien sólo dos autovalores por la izquierda (posiblemente iguales) o bien infinitos, pudiendo estar, en este segundo caso, en una, dos o infinitas clases de conjugación.*

En este mismo artículo prueban también el siguiente teorema:

**Teorema 1.2.19** *Dada $M \in \mathcal{M}_2(\mathbb{H})$, si el número de autovalores por la izquierda y por la derecha son finitos entonces los dos espectros coinciden.*

## 1.3. Sobre funciones de Morse

En topología diferencial las técnicas de la teoría de Morse son un camino, bastante directo, para analizar la topología de una variedad diferenciable estudiando determinadas funciones diferenciables sobre la variedad. Además, como puede verse en la primera parte de [35], la teoría de Morse nos permite encontrar la estructura celular de un $CW$- complejo y obtener también bastante información sobre su cohomología (véase [12]).

**Definición 1.3.1** *Sea $M$ una variedad diferenciable. Dada una función con valores reales, $f : M \longrightarrow \mathbb{R}$ diferenciable, se dice que un punto $p \in M$ es un* punto crítico *de $f$ si $f_* : T_pM \longrightarrow T_{f(p)}\mathbb{R}$ se anula. Las imágenes por $f$ de los puntos críticos se denominan* valores críticos.

*Si $p$ no es un punto crítico de $f$ se dice que $p$ es* regular.

Dado un punto crítico de $f$ podemos definir una forma bilineal simétrica $H_pf$ en el tangente, el *Hessiano de $f$ en $p$*. El índice de $H$ en un espacio vectorial $V$ está definido como la mayor dimensión de un subespacio de $V$ en el cual $H$ es definida negativa. Para referirnos al índice de $H_pf$ en el $T_pM$ hablaremos del *índice de $f$ en $p$*.

El índice de un punto crítico no degenerado $p$, es el número de autovalores negativos de la matriz Hessiana en $p$. Intuitivamente, la información que nos da es el número de direcciones independientes en un entorno de $p$ en las que $f$ decrece.

**Definición 1.3.2** *Un* punto crítico *$p$ es* no degenerado *si la Hessiana de $f$ en $p$ es no singular.*

Conviene señalar que los conceptos de punto crítico, índice y ser no degenerado son independientes de la elección de las coordenadas locales que tomemos. Estas nociones nos permiten recordar uno de los conceptos básicos del presente trabajo:

**Definición 1.3.3** *Una función $f : M \longrightarrow \mathbb{R}$ es una* función de Morse *si todos sus puntos críticos son no degenerados.*

**Lema 1.3.4** *Si $p \in M$ es un punto crítico no degenerado de una función diferenciable $f$, existen coordenadas locales $(x_1, \ldots, x_n)$ centradas en $p$ de manera que*

$$f(x_1, \ldots, x_n) = f(p) - x_1^2 - \cdots - x_r^2 + x_{r+1}^2 + \cdots + x_n^2,$$

*donde $r$ es el índice de $f$ en $p$.*

**Corolario 1.3.5** *Los puntos críticos de una función de Morse son aislados.*

Un ejemplo típico de funciones de Morse es el que se muestra a continuación:

**Ejemplo 5** *Función "altura" en el toro.*

Se observa que esta función tiene cuatro puntos críticos de índices $0, 1, 1$ y $2$.

Al considerar funciones de Morse en variedades diferenciables nos damos cuenta de que es interesante estudiar los siguientes "conjuntos de nivel":

$$M^a := \{x \in M : f(x) \leq a\}.$$

Para estos subconjuntos tenemos el siguiente lema:

**Lema 1.3.6** *Dados $a, b \in \mathbb{R}$ con $a < b$, si el subconjunto de la variedad formado por $\{x \in M : a \leq f(x) \leq b\}$ no contiene ningún punto crítico de $f$ en $M$, entonces $M^a$ es un retracto por deformación de $M^b$.*

Es decir, podemos pasar de un conjunto a otro mediante una aplicación homótopa a la identidad.

De este lema se deduce de manera inmediata un criterio de existencia. Si dos "conjuntos de nivel" $M^a$ y $M^b$ no son topológicamente equivalentes (no son homótopos) entonces podemos asegurar que entre $a$ y $b$ hay un valor crítico de $f$. Es decir, el conjunto $\{x \in M : a \leq f(x) \leq b\}$ contiene algún punto crítico.

La diferencia entre la teoría de Morse y la de Lusternik y Schnirelmann radica precisamente en el modo de estudiar las clases de conjuntos homótopos en una variedad. La primera utiliza los grupos de homología para contar el número de puntos críticos de una cierta función en $M$, mientras que la segunda busca directamente abiertos contráctiles. En principio, la teoría de Morse afina más que la de Lusternik y Schnirelmann pero exige que los puntos críticos sean no degenerados. Esto no siempre nos da los resultados más precisos, como veremos más adelante, en el ejemplo del toro de dimensión dos en $\mathbb{R}^3$, para el que podemos obtener funciones diferenciables con sólo tres puntos críticos, mientras que toda función de Morse tiene como mínimo cuatro puntos críticos.

### 1.3.1. Funciones de Bott-Morse

La teoría clásica de Morse considera sólo funciones cuyos puntos críticos son no degenerados y por tanto, aislados. En muchas situaciones, sin embargo, nos encontraremos que los puntos críticos forman subvariedades de $M$. Por ejemplo, si ponemos el toro en horizontal sobre un plano, entonces la función "altura" respecto del plano tendrá las "tapas" de arriba y de abajo como subvariedades críticas.

Debemos a Bott la extensión de la teoría de Morse a estas situaciones.

**Definición 1.3.7** *Se dice que una* variedad crítica $N$ *es* no degenerada *si para cualquier punto crítico* $p$ *de* $N$ *la Hessiana de* $f$ *restringida al espacio normal de* $N$ *es no singular.*

**Definición 1.3.8** *Una* función de Bott-Morse *es una función diferenciable cuyos puntos críticos forman una subvariedad cerrada y la Hessiana es no degenerada en la dirección normal.*

Equivalentemente, podemos decir que el núcleo del operador Hessiano en un punto crítico es el espacio tangente a la subvariedad crítica.

El índice lo pensamos ahora como un par $(i_-, i_+)$ donde $i_-$ es la dimensión de la variedad inestable en un punto de la variedad crítica, mientras que $i_+$ es $i_-$ más la dimensión de la variedad crítica.

# Capítulo 2

# Funciones de Morse en subvariedades de espacios euclidianos

En general una función de Morse nos da bastante información sobre la variedad en la que está definida. Al estudiar estas funciones la cuestión que nos planteamos consiste en cómo encontrar alguna función de Morse concreta en una variedad dada. En esta sección estudiaremos dos grandes tipos de estas funciones. Las funciones "distancia", para las que seguiremos principalmente el artículo [12] y las funciones "altura" [41].

Cualquier variedad diferenciable $n$-dimensional puede ser embebida diferenciablemente en un espacio euclidiano de dimensión menor que $2n+1$. Nos llega entonces con considerar subvariedades en un espacio euclidiano.

## 2.1. Funciones "distancia" en subvariedades de espacios euclidianos

Sea $M$ una subvariedad en un espacio euclidiano $E$. Para cada punto $a \in E$ consideramos una función "distancia" en la variedad:

$$\begin{aligned} f_a : M &\longrightarrow \mathbb{R} \\ x &\mapsto f_a(x) = \|x-a\|^2. \end{aligned}$$

Nos interesa identificar el conjunto de puntos críticos de esta función y buscar qué condiciones tiene que cumplir $a$ para que $f_a$ sea una función de Morse en la variedad $M$.

Denotamos por $\Sigma_a$ al conjunto de puntos críticos de $f_a$ y por $\gamma_x$ el plano normal a $M$ en $x$, es decir, $\gamma_x = \{v \in E \mid v \perp T_xM\}$.

**Lema 2.1.1** *Sea $f_a : M \longrightarrow \mathbb{R}$ una función "distancia", entonces:*
(i) $\Sigma_a = \{x \in M : a - x \in \gamma_x\}$.
(ii) *Para casi todo $a \in E$, $f_a$ es una función de Morse.*

Es decir, los puntos críticos de $f_a$ son aquellos puntos de $E$ tales que el vector que une $a$ y $x$, $x - a$, es perpendicular a la variedad $M$ en $x$.

*Demostración.–*
*(i)* Sea $\alpha(t)$ una curva tal que $\alpha(0) = x$ y $\alpha'(0) = v \in T_xM$. Si denotamos por $\mu(t) = \alpha(t) - a$, la expresión que obtenemos para la diferencial de $f_a$ en un punto $x$ es:

$$f_{*x}(v) = \frac{d}{dt}_{|t=0} f(\alpha(t)) = 2\langle \mu(0), \mu'(0) \rangle = 2\langle x - a, v \rangle.$$

Como

$$\langle \mathrm{grad}_x f, v \rangle = f_{*x}(v)$$

entonces

$$\mathrm{grad}_x f = 2(x - a) + w \in T_xM$$

donde $w \in \gamma_x$. Así

$$2(x - a) = \mathrm{grad}_x f - w$$

con $\mathrm{grad}_x f \in T_xM$ y $w \in \gamma_x$. Por tanto, el gradiente de $f_a$ en el punto $x$ es la proyección ortogonal de $2(x-a)$ sobre el tangente; luego $x$ es singular cuando y sólo cuando $x - a \in \gamma_x$.

*(ii)* Para la demostración de este apartado hemos seguido el argumento usado en [35, p. 32-38].

Sea $\nu M = \bigcup_{x \in M} \gamma_x$ el fibrado normal. Recordemos que un punto $x \in E$ es un foco de $M$ si hay algún punto $q$ de la variedad tal que $x = q + v$ con $v$ un vector perpendicular a $M$ en $q$ y $J_{(q,v)}\varphi$ es degenerado, con $\varphi\colon \nu M \longrightarrow E$ tal que a cada par $(q, v)$ le asocia el punto final $q+v$. Es decir, si y sólo si $(x-q) \in \gamma_q$ y el Jacobiano de $\psi$ en $(q, x - q)$ tiene núcleo no trivial.

Sea $\Lambda \subset E$ el conjunto de puntos focales de $M$. Probaremos primero que $f_a$ es una función de Morse si y sólo si $a \in E \setminus \Lambda$ y después que $\Lambda$ tiene medida nula en $E$.

Por el siguiente lema que prueba Milnor en [35, p. 36]:

> **Lema** *Un punto $q \in M$ es un punto crítico degenerado de $f_a$ si y sólo si $a$ es un foco de $M$ en $q$.*

Tenemos entonces que $f_a$ es de Morse cuando y sólo cuando $a$ no es un foco.

Para ver que $\Lambda$ tiene medida nula en $E$ utilizaremos el siguiente teorema de Sard probado en [19, p. 205].

> **Teorema** *Si $M_1$ y $M_2$ son dos variedades diferenciables con una base numerable de la misma dimensión y $f : M_1 \to M_2$ es de clase $\mathcal{C}^1$, entonces la imagen del conjunto de puntos críticos tiene medida nula en $M_2$.*

Ahora bien, el fibrado normal $\nu M$ es una variedad de la misma dimensión que $E$. Además, un punto de la variedad $x \in M$ es un foco si y sólo si es un valor crítico de la aplicación $\varphi$

$$\begin{aligned} \varphi : \nu M &\longrightarrow E \\ v \in \gamma_x &\mapsto x + v. \end{aligned}$$

Por tanto, el conjunto de focos de $M$ tiene medida nula.

*q.e.d.*

Después de este resultado general nos restringimos al estudio del espacio euclidiano $E = \mathcal{M}_{n\times n}(\mathbb{K}) = \mathbb{K}^{n^2}$ y $O_n(\mathrm{id}, \mathbb{K})$ como subvariedad que, como vimos en el primer ejemplo es:

$$O_n(\mathrm{id}, \mathbb{K}) = O_n(\mathbb{K}) = \begin{cases} O(n) & \text{si } \mathbb{K} = \mathbb{R} \\ U(n) & \text{si } \mathbb{K} = \mathbb{C} \\ Sp(n) & \text{si } \mathbb{K} = \mathbb{H} \end{cases}$$

## 2.2. Producto escalar en $\mathcal{M}_{n\times n}(\mathbb{K})$

Consideramos un elemento $u$ de la variedad $O_n(\mathrm{id}, \mathbb{K})$ como vector de $\mathbb{R}^k$ con $k = n^2, 2n^2$ ó $4n^2$ según $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ó $\mathbb{H}$ respectivamente. La norma euclidiana usual,

$$\|u\|^2 = \sum_{i,j=1}^{n} |u_{i,j}|^2$$

coincide con la norma de $u$ como matriz de $\mathcal{M}_n(\mathbb{K})$ :

$$\begin{aligned} u^* u &= \begin{pmatrix} \overline{x_{11}} & \cdots & \overline{x_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{x_{1n}} & \cdots & \overline{x_{nn}} \end{pmatrix} \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \ldots & x_{n,n} \end{pmatrix} \\ &= \begin{pmatrix} \overline{x_{11}}x_{11} + \overline{x_{21}}x_{21} + \cdots + \overline{x_{n1}}x_{n1} & & \\ & \ddots & \\ & & \overline{x_{1n}}x_{1n} + \cdots + \overline{x_{nn}}x_{n,n} \end{pmatrix} \end{aligned}$$

por tanto

$$\|u\|^2 = \sum_{i,j=1}^{n} |u_{i,j}|^2 = Tr(u^* u).$$

A partir de esta norma obtenemos ahora el producto escalar:

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

luego

$$\langle u, v\rangle = \frac{1}{2} Tr(u^*v + v^*u) = ReTr(u^*v).$$

### 2.2.1. Espacio tangente y ortogonal.

Estudiemos ahora cómo son el tangente y el ortogonal de estas variedades.

De ahora en adelante denotaremos por $G$ a la variedad $O_n(\mathbb{K})$.

**Lema 2.2.1** *Para $X \in G$,*
*(i) El espacio tangente a la variedad en $X$ está formado por:*

$$T_X G = \{U \in \mathcal{M}_{n\times n}(\mathbb{K}) : X^*U = -U^*X\}.$$

*(ii) En consecuencia, el espacio normal es:*

$$\gamma_X G = \{U \in \mathcal{M}_{n\times n}(\mathbb{K}) : X^*U = U^*X\}.$$

*Demostración.–*
*(i)* Como vimos en el ejemplo 1,

$$\mathfrak{g} = \{v \in \mathcal{M}_{n\times n}(\mathbb{K}) : v + v^* = 0\}$$

es el tangente en el neutro. En un punto $X$ cualquiera:

$$T_X G = L_{X^*} T_e G = L_{X^*} \mathfrak{g}.$$

Tenemos pues que $U \in T_X G$ si y sólo si $U = XV$ con $V$ tal que $V + V^* = 0$, es decir,

$$X^{-1}U = V$$

y como $X \in G$

$$X^*U = V$$

esto es,

$$X^*U + U^*X = 0.$$

*(ii)* Dado un $X \in G$ podemos descomponer el espacio de matrices como

$$\mathcal{M}_n(\mathbb{K}) = T_X G \oplus \gamma_X G,$$

es decir,

$$\mathcal{M}_{n\times n}(\mathbb{K}) = \{U \in \mathcal{M}_n(\mathbb{K}) : X^*U = -U^*X\} \oplus \{U \in \mathcal{M}_n(\mathbb{K}) : X^*U = U^*X\}.$$

Luego podemos expresar cualquier matriz como suma de una antihermítica y otra hermítica:

$$V = \frac{1}{2}(V - V^*) + \frac{1}{2}(V + V^*)$$

donde el primer sumando pertenece al $T_{\mathrm{id}}G$ y el segundo a $\gamma_{\mathrm{id}}G$. Además, se cumple que $\langle V - V^*, V + V^*\rangle = 0$, es decir, los sumandos son ortogonales.

Ahora basta ver que $(L_X)_{*\mathrm{id}} = L_X$ conserva el producto escalar. Para ello, al haber obtenido el producto en función de la norma, es suficiente comprobar que $L_X$ preserva la norma.

$$\|XV\|^2 = Tr((XV)^*XV) = Tr(V^*X^*XV) = Tr(V^*V) = \|V\|^2.$$

Luego $\gamma_X G$ es la traslación por la izquierda por $X$ de las matrices hermíticas.

*q.e.d.*

## 2.3. Funciones "distancia" en $O_n(\mathbb{K})$

Nos detendremos a partir de ahora el caso particular que interesa más para el presente trabajo, las funciones de Morse en los grupos de Lie ortogonales.

Sea la sucesión de números reales $0 < \lambda_1 < \ldots < \lambda_n$ y la matriz diagonal $B = \mathrm{diag}(\lambda_1 \ldots \lambda_n) \in E = \mathcal{M}_{n\times n}(\mathbb{K})$. Interesa considerar esta matriz para que los puntos críticos de la función "distancia" sean aislados.

Con la métrica que definimos en la sección 2.2 obtenemos la siguiente expresión de $f_B$ :

$$\begin{aligned}
f_B(X) &= \|X - B\|^2 \\
&= \langle X - B, X - B\rangle \\
&= \langle X, X\rangle - 2\langle X, B\rangle + \langle B, B\rangle \\
&= n + \sum_{i=1}^{n} \lambda_i^2 - 2\sum_{i=1}^{n} \lambda_i \Re(x_{ii}) \\
&= (n + \sum_{i=1}^{n} \lambda_i^2) - 2\Re Tr(BX).
\end{aligned}$$

Para una subsucesión $J = \{i_1, \dots, i_r\} \subseteq \{1, \dots, n\}$ denotamos por $\sigma_J$ a la matriz diagonal con $-1$ en los puestos señalados por el conjunto $J$ y 1 en el resto:

$$\sigma_J = \begin{pmatrix} \varepsilon_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \varepsilon_n \end{pmatrix} \text{ con } \varepsilon_k = \begin{cases} -1 & \text{si } k \in J, \\ 1 & \text{si } k \notin .J \end{cases}$$

Probamos ahora que $f_B$ es una función de Morse, hallamos sus puntos críticos y calculamos el índice.

### 2.3.1. Puntos críticos

**Proposición 2.3.1** *El conjunto de puntos críticos de $f_B$ viene dado por:*

$$\Sigma_B = \{\mathrm{id}\} \cup \{\sigma_J : J \subseteq I = \{1, \dots, n\}\}.$$

*Demostración.–*

Por el lema 2.1.1 $X$ es un punto crítico de $f_B$ si y sólo si $X - B \in \gamma_X$, es decir, $X^*(X - B)$ es una matriz hermítica,

$$X^*(X - B) = (X - B)^* X,$$

de donde, al estar $X \in G$, obtenemos la condición

$$X^* B = BX$$

pues $B^* = B$ por ser una matriz diagonal real. Como todos los $\lambda_i \neq 0$, $B$ es inversible así que podemos escribir

$$X = B^{-1} X^* B$$

lo que nos da que, si $i \leq j$, entonces

$$x_{ji} = (\frac{\lambda_i}{\lambda_j})\overline{x_{ij}}. \tag{2.1}$$

La condición $\lambda_i < \lambda_j$ podemos escribirla

$$c_{ij} = (1 - (\frac{\lambda_i}{\lambda_j})^2) > 0. \tag{2.2}$$

Por ser la matriz $X$ ortogonal, *i.e,* $X^*X = 0$, sus columnas forman una base ortonormal para el producto $\langle v, w\rangle = v^*w$. Además, como $X^*$ sigue siendo ortogonal, los conjugados de las filas de $X$ forman a su vez una base ortonormal.

Si tomamos la primera columna de $X^*$ y la primera fila de $X$ obtenemos:

$$|\bar{x}_{11}|^2 + |\bar{x}_{12}|^2 + \cdots + |\bar{x}_{1n}|^2 = 1,$$

$$|x_{11}|^2 + |x_{21}|^2 + \cdots + |x_{n1}|^2 = 1. \tag{2.3}$$

Sustituyendo en la segunda expresión los $x_{ji}$ según la fórmula (2.1) nos queda

$$|x_{11}|^2 + (\frac{\lambda_1}{\lambda_2})^2|\bar{x}_{12}|^2 + \cdots + (\frac{\lambda_1}{\lambda_n})^2|\bar{x}_{1n}|^2 = 1. \tag{2.4}$$

Restando la ecuación (2.3) a (2.4) obtenemos:

$$c_{12}|\bar{x}_{12}|^2 + \cdots + c_{1n}|\bar{x}_{1n}|^2 = 0,$$

luego, por (2.2),

$$x_{12} = \cdots = x_{1n} = 0.$$

Por tanto $|x_{11}| = 1$. Además, la igualdad (2.1) nos muestra que todos los elementos de la diagonal son reales luego $x_{11} = \pm 1$.

Con un cálculo análogo para la fila y columna $j$-ésimas, con $j = 2, \ldots, n$, obtendríamos $x_{ij} = 0$ cuando $i \neq j$. Es decir, los puntos críticos son las matrices diagonales con todos sus elementos principales 1 ó $-1$.

*q.e.d.*

### 2.3.2. Operador Hessiano

El siguiente lema nos da una expresión del operador Hessiano de $f_B$ que utilizaremos para probar que sus puntos críticos son no degenerados, siendo así $f_B$ una función de Morse.

**Lema 2.3.2** *Dado un punto crítico $X \in G$ el operador Hessiano para $f_B$ en $X$ viene dado por:*

$$H_X(f_B)(U) = (1/2)X(BU - U^*B).$$

*Demostración.–*

De acuerdo con el lema 2.1.1 el campo gradiente en $X$ de $f_B$ es la proyección ortogonal de $2(X - B)$ al $T_X G$. En general, cualquier matriz puede descomponerse como vimos en la demostración del lema 2.2.1,

$$2(X - B) \in L_X(\mathcal{A}) \oplus L_X(\mathcal{H})$$

siendo $\mathcal{H}$ y $\mathcal{A}$ las matrices hermíticas y las antihermíticas respectivamente.

Denotamos por $P = (X - B)$. Como $X^{-1} = X^*$, $(L_X)^{-1}(P) = X^*P \in T_{\mathrm{id}}G$. Entonces

$$\begin{aligned} L_X[\frac{1}{2}(X^*P - P^*X) + (X^*P + P^*X)] &= \\ \frac{1}{2}[(XX^*P - XP^*X) + (XX^*P + XP^*X)] &= \\ \frac{1}{2}(P - XP^*X) + \frac{1}{2}(P + XP^*X), \end{aligned}$$

luego

$$\begin{aligned} \mathrm{grad}_X f_B &= \\ \frac{1}{2}[(X - B) - X(X - B)^*X] &= \\ \frac{1}{2}(-B + XBX). \end{aligned}$$

Para $U \in T_X G$

$$\begin{aligned} H_X(f_B)(U) &= \\ \lim_{t\to 0} \frac{1}{t}[\mathrm{grad}\ f_B(X + tU) - \mathrm{grad}\ f_B(X)] &= \\ \lim_{t\to 0} \frac{1}{2t}[-B + (X + tU)B(X + tU) - (-B + X)BX)] &= \\ \lim_{t\to 0} \frac{1}{2t}[XBX + tXBU + tUB + t^2UBU - XBX] &= \\ \frac{1}{2}(XBU + UBX). \end{aligned}$$

Ahora bien, como los vectores del tangente a $G$ en $X$ anticonmutan salvo la traspuesta de la conjugada, con $X$ y $X$ y $B$ conmutan por ser ambas matrices diagonales reales,

$$\begin{aligned} XBU + UBX &= \\ XBU + UXB &= \\ XBU - XU^*B &= \\ X(BU - U^*B) & \end{aligned}$$

esto es,

$$H_X(f_B)(U) = \frac{1}{2}X(BU - U^*B)$$

que, efectivamente, sigue estando en el $T_XG$ :

$$\begin{aligned} X^*X(BU - U^*B) &= \\ BU - U^*B &= \\ (U^*B - BU)^* &= \\ -(BU - U^*B)^*X^*X &= \\ -(X(BU - U^*B))^*X. & \end{aligned}$$

*q.e.d.*

*Observación.*— Según el artículo [12], (lema 3.4 pág 20)

$$H_X(f_B)(U) = (U^*B - BU^*)X.$$

De todas formas, probablemente sea un error tipográfico pues esta expresión ni siquiera está en el plano tangente. Además, después aparecen bien calculados tanto los autovalores como el índice.

**Teorema 2.3.3** *$f_B : G \longrightarrow \mathbb{R}$ es una función de Morse en $G$.*

*Demostración.–*
Por la proposición 2.1.1 ya sabemos quienes son los puntos críticos de $f_B$. Nos llega entonces con probar que son no degenerados.

Un conjunto natural de generadores del espacio de matrices de orden $n$ viene dado por

$$\{E_{s,t} : 1 \leq s \leq t \leq n\} \sqcup \{F_{s,t} : 1 \leq s < t \leq n\}$$

donde $E_{s,t}$ son las matrices simétricas de componentes:

$$e_{ij} = \begin{cases} 1 & \text{si } (i,j) = (s,t) \text{ ó } (i,j) = (t,s) \\ 0 & \text{en otro caso} \end{cases}$$

y $F_{s,t}$ las antisimétricas de componentes:

$$f_{ij} = \begin{cases} -1 & \text{si } (i,j) = (s,t) \\ 1 & \text{si } (i,j) = (t,s) \\ 0 & \text{en otro caso.} \end{cases}$$

Dado un punto crítico $X \in \Sigma_B, X = \sigma_J$, una base para el $T_X G$ es la trasladada de la anterior, es decir, aquellos $E_{s,t}$ y $F_{s,t}$ que conmutan con $X$:

$$\{E_{s,t} : (s,t) \in J \times S\} \sqcup \{F_{s,t} : (s,t) \in J \times J, S \times S, s < t\}.$$

Donde $S$ es el complementario de $J$ en $I$.

Aplicamos ahora la expresión obtenida en el lema 2.3.2 para el operador Hessiano a las matrices de la base $E_{s,t}$ y $F_{s,t}$, y resulta

$$H_X(f_B)(E_{s,t}) = (\lambda_t - \lambda_s)E_{s,t} \text{ si } (s,t) \in I \times J, \tag{2.5}$$

$$H_X(f_B)(F_{s,t}) = \begin{cases} -(\lambda_s + \lambda_t)F_{s,t} & \text{si } (s,t) \in J \times J \\ (\lambda_s + \lambda_t)F_{s,t} & \text{si } (s,t) \in S \times S. \end{cases} \tag{2.6}$$

Esto nos muestra que las matrices de la base escogida son precisamente los autovectores del operador Hessiano de $f_B$. Tenemos pues que la matriz asociada a $H_X(f_B)$ tiene rango máximo y por tanto, $f_B$ es una función de Morse.

*q.e.d.*

### 2.3.3. Índice

**Proposición 2.3.4** *El índice de $f_B$ en cada punto crítico viene dado por:*

$$Ind(\sigma_{i_1,\dots,i_r})f_B = dim_{\mathbb{R}}\mathbb{K} \cdot (i_1 + \cdots + i_r) - r.$$

*Demostración.–*

Veamos que el número de direcciones en las cuales el operador Hessiano es definido negativo es

$$Ind(\sigma_{i_1,\dots,i_r}) = \Sigma i_s - r.$$

Según las expresiones obtenidas para las imágenes de la base por el operador $H_{X_0}(f_B)$ tenemos que

$$Ind(\sigma_J) = \#\{(s,t) \in J \times J : s < t\} + \#\{(s,t) \in J \times S : s > t\},$$

donde el primer sumando viene de la primera ecuación de (2.6) y el segundo de cuando (2.5) es negativo. Entonces,

$$Ind(\sigma_J) = [(r-1) + (r-2) + \cdots + 1] + [(i_1 - 1) + (i_2 - 2) + (i_r - r)] = \sum_{k=1}^{r} i_k - r$$

por tanto, la dimensión real es

$$Ind(\sigma_J) = \dim_{\mathbb{R}} \mathbb{K}(i_1 + \cdots + i_r) - r.$$

*q.e.d.*

**Ejemplo 6** *Función "distancia" en $G = U(2)$.*

Dada $B \in \mathcal{M}_2(\mathbb{C}),\ B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ y $X \in U(2)$,

$$f_B(X) = 2 + \lambda_1^2 + \lambda_2^2 - 2(\lambda_1 \Re(x_{11}) + \lambda_2 \Re(x_{22})) = (2 + \lambda_1^2 + \lambda_2^2) - 2\Re Tr(BX).$$

Por la proposición 2.3.1 el conjunto de puntos críticos está formado por

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

con índices

$$\begin{aligned}
&Ind(id) = 0, \\
&Ind(-id) = Ind(\sigma_{\{1,2\}}) = 2(1+2) - 2 = 4, \\
&Ind \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = Ind(\sigma_1) = 2 \cdot 1 - 1 = 1, \\
&Ind \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = Ind(\sigma_2) = 2 \cdot 2 - 1 = 3.
\end{aligned}$$

## 2.4. Funciones "altura" en $O_n(\mathbb{K})$

Consideremos ahora otro tipo de funciones de Morse en los grupos de Lie clásicos, las funciones "altura".

Fácilmente se comprueba que estas funciones coinciden precisamente con las funciones "distancia" salvo una constante.

De hecho, si consideramos $h_B$ una función "altura" arbitraria en $G$ con el producto escalar usual, podemos afirmar que:

**Proposición 2.4.1** *Toda función "altura" $G$ es de la forma*

$$h_B(X) = \Re Tr(BX).$$

*Demostración.–*
Sea $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ un vector arbitrario no nulo. Podemos suponer que el hiperplano desde el que vamos a calcular la distancia pasa por el origen (en otro caso la función resultante diferiría sólo en una constante), de manera que el hiperplano es simplemente $B^\perp$. Supongamos también, sin pérdida de generalidad, que $\|B\| = 1$.

Tomamos un elemento del grupo $g \in G$ que puede expresarse como

$$g = h_B(g)B + v, \text{ con } v \in B^\perp,$$

así

$$\langle B, g\rangle = h_B(g)\langle B, B\rangle + 0 = h_B(g)$$

pues $\|B\| = 1$. Entonces, como el producto escalar con el que trabajamos es precisamente la parte real de la traza,

$$h_B(g) = \Re Tr(B^* g).$$

*q.e.d.*

### 2.4.1. Puntos críticos

**Lema 2.4.2** *Sea $h_B : G \longrightarrow \mathcal{M}$ una función "altura". El gradiente de $h_B$ en un punto $X \in \mathcal{M}$ viene dado por:*

$$\operatorname{grad} h_B(X) = B^* - XBX.$$

*Demostración.–*

$\mathrm{grad}_X h_B$ en $G$ es la proyección del gradiente de $h_B$ en $E$ en el espacio tangente $T_X G$. Ahora bien,

$$h_B(X) = \Re Tr(BX) = \langle B^*, X\rangle$$

de modo que, siguiendo el mismo procedimiento que en la demostración del lema 2.1.1 obtenemos

$$\mathrm{grad}_X h_B = B^* + W \text{ con } W \in \gamma_X G.$$

De nuevo, como las traslaciones son isometrías y en el tangente en el neutro tenemos la descomposición

$$X^{-1}B^* = (BX)^* = \frac{1}{2}[(BX)^* - BX] + \frac{1}{2}[(BX)^* + BX]$$

entonces

$$\mathrm{grad}_X h_B = X\frac{1}{2}[(BX)^* - BX] = \frac{1}{2}(B^* - XBX).$$

*q.e.d.*

**Proposición 2.4.3** *Los puntos críticos de $h_B$ son las matrices $X \in G$ que satisfacen $BX = (BX)^*$,*

$$\Sigma_B = \{X \in G : BX = (BX)^*\}.$$

*Demostración.–*

Por el lema 2.4.2 los puntos críticos de $h_B$ son aquellos $X \in G$ tales que

$$0 = \frac{1}{2}(B^* - XBX)$$

es decir, $B^* = XBX$. Además, como $X^* = X^{-1}$, entonces $X^*B^* = BX$, por tanto $BX = (BX)^*$.

*q.e.d.*

### 2.4.2. Caso $B = \mathrm{id}$

Consideramos en primer lugar el caso en el que $B$ es la matriz identidad, con $h_{\mathrm{id}}(X) = \Re TrX$. Este caso fue estudiado por Frankel en [17]. Al ser $h_{id}$ invariante por la acción adjunta, los puntos críticos ya no son aislados, sino que forman subvariedades, de hecho, son Grassmannianas. Es decir, $h_{id}$ no es estrictamente una función de Morse y para estudiarla necesitamos la teoría de Bott.

La particularidad de este caso consiste en que nos llega con calcular los puntos críticos en un toro maximal. En efecto, sea $T$ un toro maximal de $G$. Como la traza es invariante por conjugación, si $\mathrm{grad}_X f = 0$ para alguna $X \in G$, entonces $\mathrm{grad}_{BXB^{-1}} f = 0$ para cada $B$ de $G$. Luego, si $X$ es punto crítico todos los puntos de la órbita de $X$, $M_X$, también lo son. Por otro lado, sabemos que todas las órbitas cortan al toro maximal, es decir, cualquier elemento del grupo $G$ es conjugado a algún otro del toro maximal, por tanto, el conjunto de puntos críticos de $f$ está formado por las órbitas de los puntos críticos de $f$ en $T$.

Puede verse en [17] que los toros maximales de los grupos ortogonales pueden elegirse como sigue:

$$\begin{cases} T = \mathrm{diag}(1, R(\theta_1), \ldots, R(\theta_m)) \in \mathcal{M}_{2n+1}(\mathbb{R}), & \text{si } G = O_{2n+1}(\mathbb{R}); \\ T = \mathrm{diag}(R(\theta_1), \ldots, R(\theta_m)) \in \mathcal{M}_{2n}(\mathbb{R}), & \text{si } G = O_{2n}(\mathbb{R}); \\ T = \mathrm{diag}(z_1, \ldots, z_n) \in \mathcal{M}_n(\mathbb{C}) \text{ con } |z_i| = 1, & \text{si } G = U(n); \\ T = \mathrm{diag}(z_1, \ldots, z_n) \in \mathcal{M}_n(\mathbb{H}) \text{ con } |z_i| = 1, & \text{si } G = Sp(n). \end{cases}$$

Donde denotamos por $R(\theta_i)$ a la matriz de orden dos asociada a la rotación de ángulo $\theta$,

$$R(\theta_i) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Obtenemos así el siguiente resultado.

**Proposición 2.4.4** *Para la función "altura" $h_{\mathrm{id}}$ se verifica:*

(i) *Su conjunto de puntos críticos está formado por las involuciones,*

$$\Sigma_{\mathrm{id}} = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) : X^2 = \mathrm{id}\}.$$

(ii) *Si $S_V$ es la reflexión con respecto de la variedad estable $V$, con $\dim V = m$, el índice del punto crítico $S_V$ viene dado por:*

$$ind_{S_V} h_i d = \begin{cases} \frac{1}{2}m(m-1), & si\ \mathbb{K} = \mathbb{R}; \\ m^2, & si\ \mathbb{K} = \mathbb{C}; \\ m(2m+1), & si\ \mathbb{K} = \mathbb{H}. \end{cases}$$

*Demostración.–*

*(i)* Esta afirmación se deduce inmediatamente del hecho de que todas las matrices de $O_n(\mathbb{K})$ son normales y la proposición 2.4.3.

*(ii)* Puede verse en [17].

*q.e.d.*

**Ejemplo 7** *La función parte real de la traza en el toro 2-dimensional.*

Consideramos $h_{id}(X) = \Re Tr(X)$ en el toro $T^2 = [-\pi, \pi] \times [-\pi, \pi]$. En este caso tenemos un máximo local, de índice 2; dos puntos de silla, de índice 1 y un mínimo local, de índice 0.

En la siguiente figura aparece representada la gráfica de esta función.

### 2.4.3. Funciones “altura” para una matriz diagonal

Volviendo a tomar $B$ del tipo $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ con $\lambda_1 < \ldots < \lambda_n$, podemos obtener explícitamente el operador Hessiano de $h_B$ igual que para las funciones “distancia”.

El estudio de estas funciones que hacemos aquí se basa principalmente en el artículo de Dynnikov y Veselov [41].

**Lema 2.4.5** *Sea $X \in G$ un punto crítico y $U \in T_X G$, entonces*

$$H_X(h_B)(U) = \frac{1}{2}(BU^* - UB)X.$$

**Teorema 2.4.6** *Sea $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Entonces:*

1. *Los puntos críticos de $h_B$ son las matrices $\sigma_J$ de la forma $\sigma_J = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_n)$ con $\varepsilon_i = 1$ si $i \in J$ y $\varepsilon_i = -1$ si $i \notin J$.*

2. *$h_B$ es una función de Morse en el grupo $G$.*

3. *El índice de cada punto crítico $X$ es*

$$ind_X(h_B) = \dim_{\mathbb{R}}(\mathbb{K})(i_1 + \ldots i_r) - r.$$

*Demostración.–*

1. De acuerdo con la proposición 2.4.3 sus puntos críticos son las matrices $X$ tales que $BX = (BX)^*$. Procediendo de igual manera que en la demostración del lema 2.3.1 obtenemos el resultado.

2. Sea $X_0$ un punto crítico de $h_B$.
   Tomamos una carta local en el neutro y consideramos la aplicación exponencial que lleva una matriz $Z \in T_{\mathrm{id}}G$ en la matriz $X_0\exp(Z) \in T_{X_0}G$. Al hacer el desarrollo de Taylor de orden 2 de $h_B$ en $X_0$ con esta carta local obtenemos, para $X \sim X_0$:

$$\begin{aligned} h_B(X) - h_B(X_0) &= \\ \Re TrB(X - X_0) &= \\ \Re TrBX_0(expZ - \mathrm{id}) &= \\ \Re TrBX_0(Z + \frac{1}{2}B^2) + o(\|B\|^2) &= \\ \Re Tr(\sum_{i=1}^{n} \lambda_i \varepsilon_i z_{ii} + \frac{1}{2}\sum_{i=1}^{n} \lambda_i \varepsilon_i \sum_{i,j=1}^{n} z_{ij} z_{ji}) + o\|Z\|^2 &= \\ -\frac{1}{2}\sum_{i=1}^{n} \lambda_i \varepsilon_i |z_{ij}|^2 + o\|Z\|^2 & \end{aligned}$$

   ya que como $Z \in T_{\mathrm{id}}G,\ Z + Z^* = 0$, i.e., $z_{ij} = -\overline{z_{ji}}$ y $z_{ii} = 0$.

3. Por el lema 2.4.5 obtenemos el índice de $h_B$ del mismo modo que para las funciones “distancia”.

*q.e.d.*

Puede verse (por ejemplo, [41, p. 7]) que las funciones “altura” obtenidas para estas matrices diagonales son funciones de Morse “perfectas”, es decir, con el menor número de puntos críticos posibles.

# Capítulo 3

# Categoría LS

## 3.1. Nociones básicas

Un objetivo clásico en el estudio de cualquier objeto matemático ha sido siempre el de tratar de reducir dicho objeto a partes más simples y conocidas cuyo análisis nos permita conocer la estructura de nuestro objeto inicial. En la teoría de homotopía el objeto más simple es un punto, es decir, un conjunto contráctil. Sin embargo, si tenemos en cuenta cómo están situados estos conjuntos en cada variedad no nos interesará tanto que sean contráctiles en sí mismos como que lo sean en el espacio. Esto es precisamente lo que busca la categoría LS.

El estudio de las funciones de Morse y sus puntos críticos sobre determinadas variedades nos servirá ahora para estudiar la categoría LS de algunos grupos de Lie. Antes de meternos con el cálculo de dicho invariante topológico recordemos algunas nociones básicas sobre la categoría LS.

**Definición 3.1.1** *Dado un espacio $X$ diremos que un abierto $U$ de $X$ es* categórico *si es contráctil en $X$. Un recubrimiento $\{U_i\}$ de $X$ formado por abiertos de este tipo es un* recubrimiento categórico.

**Definición 3.1.2** *La* categoría de Lusternik-Schnirelmann *o* categoría LS de un espacio X *es el menor entero $n$ tal que existe un recubrimiento categórico de $X$ formado por $n+1$ abiertos, $U_0, \ldots, U_n$. La denotaremos por $cat(X)$.*

*Si no existe un recubrimiento de este tipo diremos que la categoría de $X$ es infinito.*

*Análogamente, se define la* categoría de un subespacio $A \subset X$, $\mathrm{cat}_X(A)$, *como el menor entero $n$ tal que existe $n+1$ abiertos contráctiles en $X$ que cubren $A$.*

En el caso de que los abiertos sean contráctiles en sí mismos tendríamos la denominada *categoría fuerte,* $scat(X)$, que tiene la desventaja de no ser un invariante homotópico, como puede comprobarse en el ejemplo de Fox [16]. Evidentemente, $cat(X) \leq scat(X)$.

Además de esta definición de categoría hay bastantes variantes dependiendo del tipo de espacio topológico que se considere y de la condición que se le exija a los abiertos (véase, por ejemplo, [25]). Algunas de ellas son la categoría punteada, cuando se consideran espacios topológicos con un punto; la categoría débilmente punteada; para un espacio fibrado, la categoría seccional y unas cuantas variantes más que no interesan para el presente trabajo.

Podemos afirmar que en un cierto sentido la categoría es subaditiva. Si $X = Y \cup Z$ donde $Y$ y $Z$ son subespacios abiertos de $X$, entonces

$$cat(X) \leq catY + catZ + 1.$$

Luego, la categoría de un $CW-$complejo será siempre menor o igual que el número de dimensiones en las que el complejo tiene celdas más una unidad.

**Proposición 3.1.3** *Para una variedad diferenciable $M$ conexa,*

$$\mathrm{cat}(M) \leq \dim M + 1.$$

Aunque en general no es fácil obtener la categoría, si trabajamos sobre grupos de Lie tenemos una cota superior en función de sus toros maximales:

**Proposición 3.1.4** *Sea $G$ un grupo de Lie compacto y conexo y $j\colon T \longrightarrow G$ la inclusión de un toro maximal en el grupo. Entonces*

$$\mathrm{cat}(G) \leq (\mathrm{cat}(j) + 1)[\frac{1}{2}(\dim G - \mathrm{rang}\ G) + 1] - 1.$$

### 3.1.1. Una cota inferior.

Generalmente es difícil obtener la categoría de una cierta variedad pero tenemos algunas cotas que nos la delimitan.

Sabemos que la cohomología reducida es una $\mathbb{K}$-álgebra graduada con el cup producto

$$H^*(X;\mathbb{K}) \otimes H^*(X;\mathbb{K}) \longrightarrow H^*(X;\mathbb{K})$$

donde a su vez $H^*(X;\mathbb{K}) \otimes H^*(X;\mathbb{K})$ es también una $\mathbb{K}$-álgebra graduada con la multiplicación

$$(u_1 \otimes v_1) \cdot (u_2 \otimes v_2) = (-1)|v_1||u_2|u_1u_2 \otimes v_1v_2$$

siendo $|v_1|, |u_2|$ los grados de homología correspondientes.

**Definición 3.1.5** *La longitud del producto no trivial más largo en $H^*(X;\mathbb{K})$ es la* longitud del cup producto (l.c.p.).

Debemos a Berstein y Ganea la siguiente proposición que nos da una cota inferior para la categoría.

**Proposición 3.1.6** *La $\mathbb{K}$-longitud del cup producto de un espacio topológico $X$ es menor o igual que la categoría de todo el espacio:*

$$l.c.p._{\mathbb{K}}(X) \leq cat(X).$$

En el caso de los grupos ortogonales puede verse, por ejemplo en [34, p. 119], que sus cohomologías viene dadas por:

$$\begin{aligned} H^*(U(n)) &= \Lambda(e_1, e_3, \ldots, e_{2n-1}), \\ H^*(Sp(n)) &= \Lambda(e_3, e_7, \ldots, e_{4n-1}). \end{aligned}$$

De manera que la longitud del cup producto correspondiente es:

$$\begin{aligned} l.c.p.\ (U(n)) &= n \\ l.c.p.\ (Sp(n)) &= n. \end{aligned}$$

## 3.2. Categoría LS y puntos críticos.

Enunciamos a continuación uno de los teoremas principales en la teoría de la categoría LS, el teorema de Lusternik y Schnirelmann (1934).

**Teorema 3.2.1** *Sea $M$ una variedad compacta y $f\colon M \to \mathbb{R}$ una función diferenciable sobre $M$. Entonces, el número de puntos críticos de $f$ es mayor o igual que $cat(M)+1$. Es decir, si denotamos por $\Sigma_f$ al conjunto de puntos críticos de $f$:*

$$cat(M) + 1 \leq \#\Sigma_f.$$

Este resultado se basa en que cada punto crítico determina un abierto contráctil mediante el flujo gradiente [25]. Veremos más adelante que en nuestro caso esto corresponde a la transformación de Cayley.

De aquí surge también nuestro interés en los puntos críticos pues buscamos nuevos métodos para calcular la categoría LS de los grupos de Lie clásicos.

*Nota.*— En el libro [10, p. 7] los autores insiten en que el teorema es válido no sólo para funciones de Morse sino para cualquier función diferenciable. Este resultado es muy interesante puesto que, por ejemplo, en el toro $T^2$ toda función de Morse tiene al menos cuatro puntos críticos y sin embargo puede encontrarse una función diferenciable con tres puntos críticos. Como sabemos que $l.c.p\ (T^2) = 2$ obtenemos que $cat(T^2) = 2$.

En la siguiente figura podemos observar la gráfica de dicha función.

# Capítulo 4

# La transformación de Cayley

## 4.1. Definición

Para pasar de un grupo de Lie $J$-ortogonal a su álgebra de Lie y viceversa vamos a utilizar la aplicación de Cayley:

$$c : \mathfrak{g} \cap \Omega \longrightarrow G \cap \Omega$$
$$X \mapsto c(X) = \frac{\mathrm{id} - X}{\mathrm{id} + X}$$

donde

$$\Omega = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) \text{ tales que } \mathrm{id} + X \text{ inversible}\}.$$

Nótese que tiene sentido escribir la expresión de $c$ como un "cociente" de matrices porque

$$(\mathrm{id} - X)(\mathrm{id} + X)^{-1} = (\mathrm{id} + X)^{-1}(\mathrm{id} - X).$$

En efecto, dadas $M, N \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que $\exists N^{-1}$, si $NM = MN$ entonces $M = N^{-1}MN$ y así $MN^{-1} = N^{-1}M$. Es obvio que

$$(\mathrm{id} - X)(\mathrm{id} + X) = (\mathrm{id} + X)(\mathrm{id} - X).$$

**Proposición 4.1.1** *La aplicación $c : \Omega \longrightarrow \Omega$ está bien definida.*

*Demostración.–*

Sea $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz tal que $\mathrm{id} + X$ es inversible, tenemos que ver que $c(X) \in \Omega$, es decir, que $\mathrm{id} + c(X)$ es inversible.

$$\begin{aligned}
(\mathrm{id} + c(X))(\mathrm{id} + X) &= \\
\mathrm{id} + X + c(X)(\mathrm{id} + X) &= \\
(\mathrm{id} + X) + (\mathrm{id} - X)(\mathrm{id} + X)^{-1}(\mathrm{id} + X) &= 2\mathrm{id}
\end{aligned}$$

luego

$$(\mathrm{id} + c(X))^{-1} = \frac{1}{2}(\mathrm{id} + X)$$

ya que, en general, dadas dos matrices cuadradas $A, B$ tales que $AB = \mathrm{id}$ entonces $BA = \mathrm{id}$, es decir, $B$ es la inversa de $A$.

*q.e.d.*

## 4.2. Propiedades

Estudiamos en este apartado algunas propiedades de la transformación de Cayley que nos serán útiles más adelante.

**Propiedad 4.2.1** *Si $A \in \Omega$ es inversible entonces $A^{-1} \in \Omega$ y $c(A^{-1}) = -c(A)$.*

*Demostración.–*
Comprobamos que $A^{-1} \in \Omega$ hallando la inversa de $(\mathrm{id} + A^{-1})$. Puesto que $A \in \Omega$, existe $(\mathrm{id} + A)^{-1}$. Buscamos ahora una matriz $E \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que $E(\mathrm{id} + A^{-1}) = \mathrm{id}$:

$$E + EA^{-1} = \mathrm{id}$$

implica que

$$A = EA + E$$

es decir,

$$A = E(A + \mathrm{id})$$

y por tanto,

$$A(\mathrm{id} + A)^{-1} = E.$$

Para comprobar la segunda afirmación veamos que $c(A^{-1}) + c(A) = 0$ :

$$\begin{aligned}
c(A) + c(A^{-1}) &= \\
(\mathrm{id} - A)(\mathrm{id} + A)^{-1} + (\mathrm{id} - A^{-1})(\mathrm{id} + A^{-1})^{-1} &= \\
(\mathrm{id} - A)(\mathrm{id} + A)^{-1} + (\mathrm{id} - A^{-1})A(\mathrm{id} + A)^{-1} &= \\
[(\mathrm{id} - A) + (\mathrm{id} - A^{-1})A](\mathrm{id} + A)^{-1} &= 0.
\end{aligned}$$

*q.e.d.*

**Propiedad 4.2.2** *Dada $A \in \Omega$, entonces $A^* \in \Omega$ y $c(A^*) = c(A)^*$.*

*Demostración.–*

Puesto que $(\mathrm{id}+A)^* = (\mathrm{id}+A^*)$ tenemos que $(\mathrm{id}+A^*)^{-1} = [(\mathrm{id}+A)^{-1}]^*$:

$$\begin{aligned} c(A^*) &= \\ (\mathrm{id}-A^*)(\mathrm{id}+A^*)^{-1} &= \\ (\mathrm{id}-A)^*[(\mathrm{id}+A)^{-1}]^* &= \\ [(\mathrm{id}+A)^{-1}(\mathrm{id}-A)]^* &= \\ [(\mathrm{id}-A)(\mathrm{id}+A)^{-1}]^* &= c(A)^*. \end{aligned}$$

*q.e.d.*

**Propiedad 4.2.3** *Si $X \in \Omega$ es tal que $-X \in \Omega$ entonces $c(-X) = c(X)^{-1}$.*

*Demostración.–*

$$c(X)c(-X) = [(\mathrm{id}-X)(\mathrm{id}+X)^{-1}][(\mathrm{id}+X)(\mathrm{id}-X)^{-1}] = \mathrm{id}.$$

*q.e.d.*

**Propiedad 4.2.4** *La aplicación $c$ es involutiva, es decir, $c^{-1} = c$.*

*Demostración.–*

Sea $X \in \Omega$,

$$c^2(X) = c(c(X)) = \frac{\mathrm{id}-c(X)}{\mathrm{id}+c(X)}.$$

Por un lado tenemos

$$\begin{aligned} \mathrm{id}-c(X) &= \\ \mathrm{id}-(\mathrm{id}-X)(\mathrm{id}+X)^{-1} &= \\ (\mathrm{id}+X)(\mathrm{id}+X)^{-1}-(\mathrm{id}-X)(\mathrm{id}+X)^{-1} &= \\ [(\mathrm{id}+X)-(\mathrm{id}-X)](\mathrm{id}+X)^{-1} &= \\ 2X(\mathrm{id}+X)^{-1} & \end{aligned}$$

Por otro lado, al comprobar que $c : \Omega \longrightarrow \Omega$ está bien definida obtuvimos

$$(\mathrm{id}+c(X))^{-1} = \frac{1}{2}(\mathrm{id}+X)$$

luego

$$c^2(X) = 2X[(\mathrm{id}+X)^{-1}]\frac{1}{2}(\mathrm{id}+X) = X$$

$\forall X \in \Omega$, i.e. $c^2 = \mathrm{id} : \Omega \longrightarrow \Omega$. *q.e.d.*

**Propiedad 4.2.5** *Sean $B \in \mathcal{M}_{n\times n}$ inversible y $A \in \Omega$, entonces $BAB^{-1} \in \Omega$ y $c(BAB^{-1}) = Bc(A)B^{-1}$.*

*Demostración.–*
$BAB^{-1} \in \Omega$ pues $\mathrm{id} + BAB^{-1} = B(\mathrm{id} + A)B^{-1}$ es inversible.

Además,

$$\begin{aligned} c(BAB^{-1}) &= \\ (\mathrm{id} - BAB^{-1})(\mathrm{id} + BAB^{-1}) &= \\ B(\mathrm{id} - A)B^{-1}(B(\mathrm{id} + A)B^{-1})^{-1} &= \\ B(\mathrm{id} - A)B^{-1}B(\mathrm{id} + A)^{-1}B^{-1} &= \\ Bc(A)B^{-1}. & \end{aligned}$$

*q.e.d.*

Nótese que cuando $B^{-1} = B^*$, como ocurre por ejemplo en $U(n)$, tendremos $c(BAB^*) = Bc(A)B^*$.

**Proposición 4.2.6** *Si $\lambda$ es autovalor de $A \in \Omega$ entonces $\frac{1-\lambda}{1+\lambda}$ es autovalor de $c(A)$.*

*Demostración.–*
Sean $\lambda_1, \ldots, \lambda_n$ los autovalores de $A$. Sabemos que necesariamente $\lambda_i \neq -1$ para todo $i = 1, \ldots, n$. Si $\lambda = -1$ fuese autovalor existiría un $v \neq 0$ tal que $Av = -v$, es decir, $Av + v = 0$, esto es, $(A + \mathrm{id})v = 0$ y $v \neq 0$, luego $A + \mathrm{id}$ no tendría inversa, pero esto contradice $A \in \Omega$.

Sea $\lambda$ un autovalor de $A$. Existe un $v \in \mathbb{K}^n, v \neq 0$ tal que $Av = v\lambda$, tenemos entonces

$$(\mathrm{id} + A)v = v + Av = v(1 + \lambda)$$

por tanto,

$$c(A)v = \frac{\mathrm{id} - A}{\mathrm{id} + A}v = \frac{v - Av}{1 + \lambda} = v\frac{1 - \lambda}{1 + \lambda}.$$

*q.e.d.*

Otro modo de probar esta proposición sería utilizando que las matrices de $G$ son diagonalizables. Si $\mathbb{K} = \mathbb{C}$, $G$ es el grupo unitario de manera que cualquier $A \in G$ es diagonalizable (véase, por ejemplo, [34, p. 24-26]). En la sección 1.2.2 vimos que si la matriz está en el grupo unitario simpléctico también puede diagonalizarse. Con este resultado podemos probar la proposición 4.2.6 como vemos a continuación.

Sean $\lambda_1, \ldots, \lambda_n$ los autovalores de $A$. Como $A$ es diagonalizable,

$$A = P \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} P^*$$

luego, por la propiedad 4.2.5

$$c(A) = Pc \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} P^*.$$

Ahora bien

$$c \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} = \begin{pmatrix} \frac{1-\lambda_1}{1+\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \frac{1-\lambda_n}{1+\lambda_n} \end{pmatrix}$$

## 4.3. Restricción de la transformación de Cayley

Veamos ahora que la aplicación de Cayley $c \colon \Omega \longrightarrow \Omega$ envía matrices $J$-hermíticas en $J$-ortogonales y viceversa.

**Proposición 4.3.1** $c : \mathfrak{g} \cap \Omega \longrightarrow G \cap \Omega$ *está bien definida.*

*Demostración.–*
Sea $X \in \mathfrak{g}$, *i.e,* $XJ + JX^* = 0$, entonces

$$-X = JX^*J^{-1}$$

así por las propiedades 4.2.2 y 4.2.5

$$c(-X) = Jc(X^*)J^{-1}$$

o, equivalentemente,

$$c(-X)J = Jc(X^*)$$

y utilizando la propiedad 4.2.3 obtenemos que $c(X)$ es también $J$-ortogonal:

$$c(X)Jc(X^*) = J.$$

*q.e.d.*

**Proposición 4.3.2** $c : G \cap \Omega \longrightarrow \mathfrak{g} \cap \Omega$ *está bien definida.*

*Demostración.–*
Dada $A \in G \cap \Omega$, $AJA^* = J$ de donde $JA^*J^{-1} = A^{-1}$ luego, utilizando la propiedad 4.2.4,

$$-c(A) = c(A^{-1}) = c(JA^*J^{-1})$$

como $J^2 = \pm\mathrm{id}$

$$JA^*J^{-1} = J^{-1}A^*J$$

y, por la propiedad 4.2.5,

$$-c(A) = J^{-1}c(A^*)J$$

esto es, aplicando la propiedad 4.2.2

$$-Jc(A) = c(A)^*J$$

por tanto, $c(A) \in \mathfrak{g} \cap \Omega$.

*q.e.d.*

**Proposición 4.3.3** *Sea* $X \in \mathfrak{o}_n(\mathbb{K})$*, entonces* $X$ *no tiene autovalores reales no nulos.*

*Demostración.–*
Supongamos que existe $t \in \mathbb{R}$ tal que $Xt = vt$ para algún $v \neq 0$. Entonces $v^*Xv = v^*vt = t|v|^2$ es un número real y, en consecuencia

$$v^*Xv = (v^*Xv)^* = v^*X^*v = v^*(-X)v = -v^*Xv$$

por tanto $v^*Xv$ tiene que anularse. Es decir, $t|v|^2 = 0$ luego $t = 0$.

*q.e.d.*

**Teorema 4.3.4** *El abierto* $O_n(\mathbb{K}) \cap \Omega$ *es contráctil.*

*Demostración.–*
En efecto, el álgebra de Lie $\mathfrak{o}_n(\mathbb{K}) \subset \Omega$ y la aplicación de Cayley

$$c\colon O_n(\mathbb{K}) \cap \Omega \to \mathfrak{o}_n(\mathbb{K})$$

es un difeomorfismo.

*q.e.d.*

*Nota.*— Por el teorema 4.3.4 sabemos que $\Omega \cap O_n(\mathbb{K})$ es difeomorfo a $\mathfrak{o}_n(\mathbb{K})$ y por tanto contráctil. La contracción viene dada explícitamente por

$$(A,t) \mapsto c((1-t)c(A)),\ t \in [0,1].$$

En el caso de las matrices complejas (o cuaterniónicas) si diagonalizamos $A = U\,\mathrm{diag}(\lambda_1,\ldots,\lambda_n)U^*$, $\lambda_k \neq -1$, como $c(A) = Uc(D)U^*$ sólo queda calcular

$$c(D) = \mathrm{diag}\left((1-\lambda_k)/(1+\lambda_k)\right)$$

y

$$c((1-t)c(D)) = \mathrm{diag}((t+(2-t)\lambda_k)/((2-t)+t\lambda_k)).$$

Nótese que cada complejo $\lambda_k \neq -1$ se mueve hacia 1 siguiendo la proyección estereográfica de $S^1 - \{-1\}$.

## 4.3.1. Ejemplos

**Ejemplo 8** *Consideramos el grupo unitario unidimensional:*

$$U(1) = \{z \in \mathbb{C} \mid zz^* = 1\} = S^1.$$

Su álgebra de Lie es:

$$\mathfrak{u}(1) = \{z \in \mathbb{C} \mid z + \overline{z} = 0\}.$$

Puesto que

$$z + z^* = z + \overline{z} = 2\Re(z)$$

su álgebra de Lie es la recta imaginaria, $\mathfrak{u}(1) =< i >_{\mathbb{R}}$. Además, en este caso,

$$\Omega = \mathbb{C} \setminus \{x \in \mathbb{C} : 1 + x = 0\} = \mathbb{C} \setminus \{-1\}$$

y por tanto

$$\mathfrak{u}(1) \cap \Omega = \mathfrak{u}(1).$$

La aplicación de Cayley es precisamente la proyección estereográfica de la circunferencia desde el punto $z = -1$:

$$c : \mathbb{S}^1 \setminus \{-1\} \longrightarrow < i >_{\mathbb{R}}$$

dada por

$$z \mapsto c(z) = \frac{1-z}{1+z} = \frac{\overline{z}-z}{|1+z|^2} = \frac{-it}{1+s}$$

donde $z = s + it$. Es decir, $c$ lleva los puntos de la circunferencia (excepto $z = -1$) en la recta. Su inversa tiene esta misma expresión.

**Ejemplo 9** $U(2) = \{A \in \mathcal{M}_2(\mathbb{C}) : AA^* = \mathrm{id}\}$.

Las matrices de $U(2)$ son de la forma $\begin{pmatrix} \alpha & -\lambda\overline{\beta} \\ \beta & \lambda\overline{\alpha} \end{pmatrix}$ con $\alpha, \beta \in \mathbb{C}$ tales que $|\alpha|^2 + |\beta|^2 = 1$ y $|\lambda| = 1$.

Sean $\lambda_1, \lambda_2$ autovalores de $A$. Como $A \in U(2)$, A es diagonalizable,

$$A = P \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} P^*$$

Así, para que $A \in \Omega$ tiene que cumplirse que $(1 + \lambda_1)(1 + \lambda_2) \neq 0$, es decir, $A$ no puede tener a $\lambda = -1$ por autovalor.

Utilizando ahora la propiedad 4.2.6 obtenemos que los autovalores de $c(A)$ son $\frac{1-\lambda}{1+\lambda}$.

**Ejemplo 10** *Consideremos* $Sp(1) = \{q \in \mathbb{H} : |q| = 1\} = S^3$.

Como grupo $Sp(1)$ es precisamente $SU(2)$. Su álgebra de Lie es

$$\mathfrak{sp}(1) = \{w \in \mathbb{H} : w + w^* = 0\} =< i, j, k > .$$

En este caso el abierto $\Omega$ está formado por todos los cuaternios excepto $q = -1$. Observamos entonces que la aplicación de Cayley es de nuevo la proyección estereográfica de $S^3 \setminus \{-1\}$ en $\mathbb{R}^3$:

$$c\colon \mathbb{S}^3 \setminus \{-1\} \longrightarrow < i, j, k >= \mathbb{R}^3$$

$$q \mapsto c(z) = \frac{1-q}{1+q}$$

Pero

$$\frac{1-q}{1+q} = \frac{(1-q)(1+\overline{q})}{(1+q)(1+\overline{q})} = \frac{\overline{q}-q}{|1+q|^2}$$

es decir, expresando $q$ como $q = t + xi + yj + zk$ con $t, x, y, z \in \mathbb{R}$,

$$c(q) = \frac{-2(x,y,z)}{2(1+t)}.$$

*Nota.*— Consideremos el grupo $SO(3)$ de rotaciones de $\mathbb{R}^3$ o, equivalentemente, las rotaciones de la 2-esfera $S^2$. Mediante la proyección estereográfica estas rotaciones pasan a transformaciones lineales de variable compleja de la forma

$$z \mapsto \frac{\alpha z - \overline{\beta}}{\beta z + \overline{\alpha}}, \text{ con } |\alpha|^2 + |\beta|^2 = 1.$$

El grupo de matrices complejas de orden 2 determinadas por estas transformaciones (salvo el signo) forman el grupo $SU(2)$. Tenemos pues que $Sp(1) \cong SU(2)$ y el isomorfismo de grupos de Lie es

$$Sp(1) \longrightarrow SU(2)$$

$$\alpha + j\beta \mapsto \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}, \ \alpha, \beta \in \mathbb{C}.$$

Por otro lado, $SU(2)$ es un subgrupo de $U(2)$ y en el grupo unitario tenemos la aplicación de Cayley entre él y su álgebra de Lie $c\colon \Omega \cap U(2) \to \mathfrak{u}(2)$. Sin embargo, $c(SU(2)) \nsubseteq \mathfrak{su}(2)$ pues $c$ no lleva las matrices de determinante 1 en matrices de traza 0. En efecto, evaluemos $c$ sobre un elemento $A \in SU(2)$. Tenemos

$$(\mathrm{id} + A) = \begin{pmatrix} 1+\alpha & -\overline{\beta} \\ \beta & 1+\overline{\alpha} \end{pmatrix},$$

luego

$$\det(\mathrm{id} + A) = 2(1 + \Re(\alpha)).$$

Por tanto, como

$$2(1+\Re(\alpha))c(A) = \begin{pmatrix} 1-\alpha & -\overline{\beta} \\ \beta & 1-\overline{\alpha} \end{pmatrix} \begin{pmatrix} 1+\overline{\alpha} & \beta \\ -\overline{\beta} & 1+\alpha \end{pmatrix} =$$

$$\begin{pmatrix} 1+\overline{\alpha}-\alpha-|\alpha|^2+|\beta|^2 & -\alpha\overline{\beta}-\overline{\beta}\alpha \\ \beta\overline{\alpha}+\overline{\alpha}\beta & |\beta|^2+1+\alpha-\overline{\alpha}-|\alpha|^2 \end{pmatrix}$$

queda

$$Tr[c(A)] = \frac{2+2|\beta|^2-2|\alpha|^2}{2(1+\Re(\alpha))} = \frac{2|\beta|^2}{1+\Re(\alpha)}.$$

Concluimos pues que el isomorfismo $Sp(1) \cong SU(2)$ no conmuta con la aplicación de Cayley.

$$\begin{tikzcd}
U(2) & SU(2) \arrow[l, hook'] \arrow[r, "\cong", "\varphi"'] \arrow[d, "\not c"] & Sp(1) \arrow[d, "c"] \\
\mathfrak{u}(2) \arrow[from=u, "c"] & \mathfrak{su}(2) \arrow[l, hook'] \arrow[r, "\cong", "\varphi_*"'] & \mathfrak{sp}(1)
\end{tikzcd}$$

## 4.4. Desarrollo en serie de la aplicación de Cayley

En este apartado haremos una construcción análoga a la serie de Taylor para una función $f : \mathbb{R} \longrightarrow \mathbb{R}$, es decir, considerando $c$ como función de una variable sobre $\mathcal{M} = \mathcal{M}_{n\times n}(\mathbb{K})$ y no como de $n^2$ variables sobre $\mathbb{K}$ [5].

Consideramos $c : \Omega \longrightarrow \Omega$ diferenciable, donde $\Omega \subset \mathcal{M}$ es un abierto, así:

$$c'_X = c_{*X} : T_X\Omega \longrightarrow T_{c(X)}\Omega,$$

donde ambos espacios tangentes son el espacio vectorial $\mathcal{M}$, es decir,

$$c' : \Omega \longrightarrow \mathcal{L}(\mathcal{M}, \mathcal{M}).$$

Ahora tiene sentido hacer las diferenciales sucesivas:

$$\begin{aligned} &c'' : \Omega \longrightarrow \mathcal{L}(\mathcal{M}, \mathcal{L}(\mathcal{M}, \mathcal{M})) = Bil(\mathcal{M}, \mathcal{M}) = \mathcal{L}_2(\mathcal{M}, \mathcal{M}) \\ &c^{k)} : \Omega \longrightarrow \mathcal{L}_k(\mathcal{M}, \mathcal{M}) \end{aligned}$$

Hacemos entonces el desarrollo en serie de $c$ en $X = 0$ para $H \sim 0$ :

$$c(H) = c(0) + c'_0(H) + \frac{1}{2}c''_0(H,H) + \frac{1}{3!}c'''_0(H,H,H) + \cdots$$

y de $c^{-1} = c$ en $X = \text{id}$ para $H \sim 0$:

$$c(\text{id} + H) = c(\text{id}) + c'_{\text{id}}(H) + \frac{1}{2}c''_{\text{id}}(H,H) + \frac{1}{3!}c'''_{\text{id}}(H,H,H) + \cdots$$

Observemos que se conservan las reglas de derivación del producto y del cociente:

**Lema 4.4.1** *Sean $A, B : \mathbb{R} \longrightarrow \mathcal{M}$ diferenciables. Entonces*

$$\frac{d}{dt}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t).$$

*Demostración.–*

$$\begin{aligned} \frac{d}{dt}[A(t)B(t)] &= \\ \lim_{h\to 0}\frac{1}{h}[A(t+h)B(t+h) - A(t)B(t)] &= \\ \lim_{h\to 0}\frac{1}{h}[A(t+h)B(t+h) - A(t)B(t+h) + A(t)B(t+h) - A(t)B(t)] &= \\ \lim_{h\to 0}\frac{1}{h}[(A(t+h) - A(t))B(t+h) + A(t)(B(t+h)) - B(t))] &= \\ A'(t)\lim_{h\to 0}B(t+h) + A(t)B'(t) &= \\ A'(t)B(t) + A(t)B'(t). \end{aligned}$$

**Lema 4.4.2** *Sea $B : \mathbb{R} \longrightarrow GL$, entonces*

$$\frac{d}{dt}B(t)^{-1} = -B(t)^{-1}B'(t)B(t)^{-1}.$$

*Demostración.–*

$$\begin{aligned}
\frac{d}{dt}B(t)^{-1} &= \\
\lim_{h\to 0}\frac{1}{h}[B(t+h)^{-1} - B(t)^{-1}] &= \\
\lim_{h\to 0}\frac{1}{h}[B(t+h)^{-1}B(t)B(t)^{-1} - B(t+h)^{-1}B(t+h)B(t)^{-1}] &= \\
\lim_{h\to 0}\frac{1}{h}[B(t+h)^{-1}(B(t)B(t)^{-1} - B(t+h)B(t)^{-1})] &= \\
\lim_{h\to 0}\frac{1}{h}[B(t+h)^{-1}(B(t) - B(t+h))B(t)^{-1}] &= \\
\lim_{h\to 0}B(t+h)^{-1}(\frac{B(t)-B(t+h)}{h})B(t)^{-1} &= \\
\lim_{h\to 0}B(t+h)^{-1}[-B'(t)]B(t)^{-1} &= \\
-B(t)^{-1}B'(t)B(t)^{-1}.
\end{aligned}$$

*q.e.d.*

Obtenemos ahora la "serie de Taylor":

**Proposición 4.4.3** *Sea $c : \mathfrak{g}\cap\Omega \to G$ y $H \sim 0$, entonces*

$$\begin{aligned}
c(H) &= \mathrm{id} - 2H + \frac{1}{2!}(4H^2) + \frac{1}{3!}(-12H^3) + \cdots \\
&= \mathrm{id} - 2H + 2H^2 - 2H^3 + \cdots
\end{aligned}$$

*Demostración.–*

- $c(0) = \mathrm{id}$.

- $c_{*X}(H) = \frac{d}{dt}|_{t=0}\ c(X+tH)$,

$$\begin{aligned} c_{*X}(H) &= \\ \frac{d}{dt}|_{t=0}[(\mathrm{id}-X-tH)(\mathrm{id}+X+tH)^{-1}] &= \\ [-H(\mathrm{id}+X+tH)^{-1}-(\mathrm{id}-X-tH) &\times \\ (\mathrm{id}+X+tH)^{-1}H(\mathrm{id}+X+tH)^{-1}]|_{t=0} &= \\ -H(\mathrm{id}+tX)^{-1}-(\mathrm{id}-X)(\mathrm{id}+X)^{-1}H(\mathrm{id}+X)^{-1} &= \\ [-\mathrm{id}-(\mathrm{id}-X)(\mathrm{id}+X)^{-1}]H(\mathrm{id}+X)^{-1} &= \\ -(\mathrm{id}+c(X))H(\mathrm{id}+X)^{-1}. & \end{aligned}$$

En particular: $c_{*0}(H) = -2H$.

- Calculamos la "derivada" segunda:

$$\begin{aligned} c'_{*X}(H) &= \\ \frac{d}{dt}|_{t=0}c'(X+tH) &= \\ \frac{d}{dt}|_{t=0}[-(\mathrm{id}+c(X+tH))H(\mathrm{id}+X+tH)^{-1}]. & \end{aligned}$$

Denotamos por

$$A(t) = -(\mathrm{id}+c(X+tH))H = -H - c(X+tH)H$$

y

$$B(t) = (\mathrm{id}+X+tH),$$

así

$$c'_{*X}(H) = [A(t)B(t)^{-1}]' = A'(t)B(t)^{-1} - A(t)B(t)^{-1}B'(t)B(t)^{-1},$$

es decir

$$\begin{aligned} c'_{*X}(H) &= \\ [(\mathrm{id}+c(X))H(\mathrm{id}+X)^{-1}H(\mathrm{id}+X+tH)^{-1}] &= \\ [(-H-c(X+tH)H)(\mathrm{id}+X+tH)^{-1}H(\mathrm{id}+X+tH)^{-1}]|_{t=0} &= \\ (\mathrm{id}+c(X))H(\mathrm{id}+X)^{-1}H(\mathrm{id}+X)^{-1} &+ \\ (H+c(X)H))H(\mathrm{id}+X)^{-1}H(\mathrm{id}+X)^{-1} &= \\ 2(\mathrm{id}+c(X))H(\mathrm{id}+X)^{-1}H(\mathrm{id}+X)^{-1}. & \end{aligned}$$

En particular: $c'_{*0}(H) = 4H^2$.

- Hallamos ahora la tercera "derivada":

$$\begin{aligned} c'''_X(H) = c''_{*X}(H) &= \\ \frac{d}{dt}|_{t=0}\ c''(X+tH) &= \\ \frac{d}{dt}|_{t=0}[2(\mathrm{id}+c(X+tH))H(\mathrm{id}+X+tH)^{-1}H(\mathrm{id}+X+tH)^{-1}]. \end{aligned}$$

De nuevo denotamos por

$$D(t) = 2(\mathrm{id}+c(X+tH))H(\mathrm{id}+X+tH)^{-1}H = -2A(t)H$$

luego

$$D'(t) = -2A'(t)H.$$

Teníamos

$$B(t) = (\mathrm{id}+X+tH),$$

por tanto

$$\begin{aligned} c''_{*X}(H) &= \\ \frac{d}{dt}|_{t=0}[D(t)B(t)^{-1}] &= \\ [D'(t)B(t)^{-1} - D(t)B(t)^{-1}B'(t)B(t)^{-1}]\ |_{t=0} &= \\ [-2c'_{*X}(H)H(\mathrm{id}+X+tH)^{-1} &+ \\ -(-2)c_{*X}(H)H(\mathrm{id}+X+tH)^{-1}H(\mathrm{id}+X+tH)^{-1}]\ |_{t=0} &= \\ -2c'_{*X}(H)H(\mathrm{id}+X)^{-1} + 2c_{*X}(H)H(\mathrm{id}+X)^{-1}H(\mathrm{id}+X)^{-1}. \end{aligned}$$

En particular: $c''_{*0}(H) = -2c'_{*0}(H)H + 2c_{*0}(H)H^2 = -12H^3$.

Obtenemos así la expresión de la "serie de Taylor" de $c$ que buscábamos.

*q.e.d.*

Análogamente:

**Proposición 4.4.4** *Sea* $c : G \to \mathfrak{g} \cap \Omega$ *y* $H \sim 0$, *entonces*

$$\begin{aligned} c(\mathrm{id}+H) &= -\frac{1}{2}H + \frac{1}{2!}(\frac{1}{2}H^2) - \frac{1}{3!}(\frac{3}{4}H^3) + \cdots \\ &= -\frac{1}{2}H + \frac{1}{4}H^2 - \frac{1}{8}H^3 + \cdots \end{aligned}$$

*Demostración.–*

$$\begin{aligned} c(\mathrm{id}) &= 0, \\ c_{*\mathrm{id}}(H) &= -\frac{1}{2}H, \\ c'_{*\mathrm{id}}(H) &= \frac{1}{2}H^2, \\ c''_{*\mathrm{id}}(H) &= -2c'_{*\mathrm{id}}(H)H\frac{1}{2}\mathrm{id} + 2c_{*\mathrm{id}}(H)\frac{1}{4}idH^2 = -\frac{3}{4}H^3. \end{aligned}$$

*q.e.d.*

Estudiamos ahora la convergencia de esta serie. Aunque $\mathcal{M}_{n\times n}(\mathbb{H})$ no es un álgebra sobre los cuaternios es un álgebra unitaria sobre el cuerpo de los reales de manera que para hallar el radio de convergencia de ambas series matriciales podemos utilizar la definición de radio que da G. Choquet en [7, p. 247].

**Definición 4.4.5** *Sea* $a_0 + a_1X + \ldots + a_nX^n + \ldots$ *una serie formal en* $X$ *con coeficientes en* $\mathbb{K}$. *Denominaremos* radio de convergencia *de esta serie al elemento* $\rho \in \mathbb{R}^+$ *definido por:*

$$\frac{1}{\rho} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$$

**Proposición 4.4.6** *Sea* $H \in \Omega$ *con* $\|H\| < 1$ *para alguna norma submultiplicativa. Entonces,*

$$c(H) = \mathrm{id} + \sum_{n=1}^{\infty}(-1)^n 2H^n.$$

**Proposición 4.4.7** *Sea* $H \in \Omega$ *con* $\|H\| < 2$ *para alguna norma submutiplicativa, se verifica entonces*

$$c(\mathrm{id}+H) = \sum_{n=1}^{\infty}(-1)^n \frac{1}{2^n}H^n.$$

En efecto, para la proposición 4.4.6:

$$\frac{1}{\rho} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} \sqrt[n]{2} = 1.$$

En el caso de 4.4.7:

$$\frac{1}{\rho} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{2^n}^{\frac{1}{n}} = \frac{1}{2}.$$

*Nota.*—Si $\rho(A)$ el radio espectral de la matriz $A$, $\rho(A) = \max_{\lambda_i \in Sp(A)} |\lambda_i|$. Para cualquier norma submultiplicativa en $\mathcal{M}_{n\times n}(\mathbb{K})$ se verifica que $\rho(A) \leq \|A\|$ (véase el teorema 1.4.3 de [8, p. 31]).

**Corolario 4.4.8** *Si* $\mathrm{id} + B$ *es singular entonces* $\|B\| = 1$.

En efecto, si $\mathrm{id} + B$ singular $B$ tiene a $\lambda = -1$ por autovalor, luego $\rho(B) \geq 1$.

Este corolario nos garantiza que los elementos del álgebra de Lie para los que converge la serie de la proposición 4.4.6 están en $\Omega$. Sin embargo, es interesante observar que si $B$ es ortogonal $\rho(B) = 1$ y por tanto $\|B\| \geq 1$. Es coherente entonces obtener para la serie en el grupo un radio de convergencia superior a la unidad.

## 4.5. Una generalización

Para nuestro propósito más interesante que la curva que pasa por un punto es el abierto radial que sale de cada punto crítico, de ahí la generalización de la transformación de Cayley que hacemos a continuación. Podemos así modificar la definición de $c$ para tener un recubrimiento abierto del grupo $O_n(\mathbb{K})$. Todas las propiedades que veremos a continuación se prueban de manera ánaloga al caso ya visto en el que $A = \mathrm{id}$.

Sea $A \in O_n(\mathbb{K})$. Definimos el abierto $\Omega(A) \subset \mathcal{M}_{n\times n}(\mathbb{K})$ como el conjunto formado por las matrices $X$ tales que $A + X$ es inversible.

**Definición 4.5.1** *La* transformación de Cayley centrada en $A$ *se define como la aplicación*

$$c_A \colon \Omega(A) \to \Omega(A^*)$$

*dada por*

$$c_A(X) = (\mathrm{id} - A^*X)(A + X)^{-1}.$$

Esta aplicación esta bien definida ya que, como veremos en la siguiente proposición, si $X \in \Omega(A)$ entonces $c_A(X) \in \Omega(A^*)$.

**Proposición 4.5.2** *Si $X \in \Omega(A)$ entonces,*

1. $(A+X)^{-1}(I - XA^*) = (I - A^*X)(A+X)^{-1} = c_A(X)$;
2. *La matriz* $A^* + c_A(X)$ *tiene como matriz inversa* $(1/2)(A+X)$;
3. $c_A(\Omega(A)) \subset \Omega(A^*)$;
4. $(c_A)^{-1} = c_{A^*}$.

Se tienen las siguientes propiedades:

**Proposición 4.5.3** *Sea $X \in \Omega(A)$.*

1. $X^* \in \Omega(A^*)$ *y* $c_{A^*}(X^*) = c_A(X)^*$;
2. *Para cualquier matriz* $U \in O_n(\mathbb{K})$ *se tiene que* $\Omega(UAU^*) = U\Omega(A)U^*$ *y*
$$c_{UAU^*}(UXU^*) = Uc_A(X)U^*;$$
3. *Si la matriz $X$ es inversible entonces* $X^{-1} \in \Omega(A^*)$ *y*
$$c_{A^*}(X^{-1}) = -Ac_A(X)A.$$

*Demostración.–*
Únicamente probaremos la tercera afirmación pues las demostraciones de lo demás son análogas a las del caso $A = \mathrm{id}$.

En primer lugar buscamos una matriz $E \in \mathcal{M}_{n\times n}(\mathbb{K})$ inversa de $A^* + X^{-1}$. Se verifica entonces que

$$\begin{aligned} (XA^* + \mathrm{id}) &= E \\ (XA^* + \mathrm{id})AA^*E &= X \\ (X + A)A^*E &= X. \end{aligned}$$

Hemos obtenido
$$(A^* + X^{-1}) = A(X^*A^{-1})X.$$

Podemos entonces hacer actuar a $c_{A^*}$ sobre $X^{-1}$ de manera que:

$$\begin{aligned}
c_{A^*}(X^{-1}) &= \\
(A^* + X^{-1})^{-1}(\mathrm{id} - X^{-1}A) &= \\
A(X+A)^{-1}X(\mathrm{id} - X^{-1}A) &= \\
-A(X+A)^{-1}(A - X) &= \\
-A(X+A)^{-1}(\mathrm{id} - XA^*)A &= \\
-Ac_A(X)A.
\end{aligned}$$

*q.e.d.*

A partir de estas propiedades y de la proposición 4.3.3 obtenemos el siguiente interesante resultado.

**Teorema 4.5.4** *La aplicación $c_A$ lleva difeomórficamente $\Omega(A) \cap O_n(\mathbb{K})$ en el espacio vectorial real $T_{A^*}G$ de las matrices $X$ tales que $AX + (AX)^* = 0$.*

*En consecuencia, $\Omega(A) \cap O_n(\mathbb{K})$ es un abierto contráctil.*

*Demostración.–*

En efecto, sea $X \in \Omega(A) \cap O_n(\mathbb{K})$, es decir, $X$ tal que $XX^* = \mathrm{id}$ y de manera que $A + X$ tenga inversa. Entonces

$$Ac_A(X) + (Ac_A(X))^* = 0$$

si y sólo si

$$Ac_A(X) = -c_A(X)^*A^*,$$

esto es

$$Ac_A(X)A = c_A(X)^*$$

pero por la proposición 4.5.3, apartado 3,

$$Ac_A(X)A = -c_{A^*}(X^{-1}) = -c_{A^*}(X^*)$$

por ser $X \in O_n(\mathbb{K})$. Luego, por la primera afirmación de 4.5.3

$$Ac_A(X)A = -c_A(X^*).$$

Además, $\Omega(A^*) \cap T_{A^*}G = T_{A^*}G$, es decir, si $AY + Y^*A^* = 0$ entonces $A^* + Y$ es inversible. De hecho, si $A^* + Y$ no fuese inversible existiría un $\omega \neq 0$ tal que $Y\omega = -A^*\omega$ luego $AY\omega = -\omega$ y por tanto $AY$ tendría a $-1$ como autovalor real. Sin embargo, $AY$ es antisimétrica de manera que no puede tener autovalores reales no nulos (proposición 4.3.3).

*q.e.d.*

*Nota.*— De hecho $\Omega(A) = L_{A^*}(\Omega(I)) = R_{A^*}(\Omega(I))$, y como

$$T_{A^*}(G) = L_{A^*}(\mathfrak{o}_n(\mathbb{K})) = R_{A^*}(\mathfrak{o}_n(\mathbb{K})),$$

tenemos una relación entre $c_A$ y la aplicación de Cayley $c = c_I$ dada por

$$c_A(M) = c_I(A^*M)A^*,$$

es decir, $c_A = R_{A^*} \circ c_I \circ L_{A^*}$.

$$\begin{tikzcd}
\Omega(\mathrm{id}) \arrow[r, "c_{\mathrm{id}}"] & \mathfrak{g} = T_{\mathrm{id}}G \arrow[d, "R_A"] \\
\Omega(A) \arrow[u, "L_{A^*}"] \arrow[r, "c_A"] & T_A G
\end{tikzcd}$$

### 4.5.1. Líneas de flujo para la contracción de Cayley generalizada

Sea $A \in G$. La contracción radial que sale de este punto podemos parametrizarla como $c_{A^*}(e^t X)$ para algún $X$ del tangente $T_{A^*}G$, es decir $AX = -X^*A^*$. Como $AX \in T_{id}G$ es antisimétrica, es normal y por tanto diagonalizable, $AX = UDU^*$. Entonces

$$\phi(t) = c_{A^*}(e^t X) = c_{A^*}(e^t A^* UDU^*) = (\mathrm{id} - e^t UDU^*)(A^* + e^t A^* UDU^*)^{-1},$$

esto es

$$\phi(t) = c_I(e^t UDU^*)A = Uc_I(e^t D)U^*A.$$

Veamos donde termina esta línea. Sea $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Primero calculamos

$$\lim_{t\to\infty} c_I(e^t D) = \lim_{t\to\infty} \mathrm{diag}((1 - e^t\lambda_k)/(1 + e^t\lambda_k)) = -I,$$

por tanto

$$\lim_{t\to\infty} \phi(t) = U(-I)U^*A = -A.$$

Es decir, los radios de la contracción van de $A$ a $-A$.

# Capítulo 5

# Aplicaciones

## 5.1. "Linealización" del gradiente de una función de Morse

Veamos a continuación una propiedad interesante de las funciones "altura" que nos permite integrar su flujo de un modo mucho más sencillo de lo que cabría esperar. De hecho, comprobaremos que el flujo $X(t)$ del gradiente $X' = \frac{1}{2}(B - XBX)$ lo podemos pasar mediante la aplicación de Cayley a otro flujo en el álgebra de Lie y allí integrarlo. Utilizando la expresión de Dynnikov y Veselov denominamos "linealización" a este proceso.

**Lema 5.1.1** *Una solución de*

$$X' = \frac{1}{2}(B - XBX) \tag{5.1}$$

*es la imagen por la transformación de Cayley de la solución de la ecuación diferencial:*

$$c'(X) = -2(Bc(X) + c(X)B). \tag{5.2}$$

*Demostración.–*

$$\begin{aligned} c'(X) &= \\ [(\mathrm{id} - X)(\mathrm{id} + X)^{-1}]' &= \\ -X'(\mathrm{id} + X)^{-1} + (\mathrm{id} - X)[-(\mathrm{id} + X)^{-1}X'(\mathrm{id} + X)^{-1}] &= \\ -X'(\mathrm{id} + X)^{-1} - c(X)X'(\mathrm{id} + X)^{-1} &= \\ -(\mathrm{id} + c(X))X'(\mathrm{id} + X)^{-1} &= \\ -\frac{1}{2}(\mathrm{id} + c(X))(B - XBX)(\mathrm{id} + X)^{-1}. & \end{aligned}$$

Ahora bien, como $(\mathrm{id}+X)^{-1} = 2(\mathrm{id}+c(X))$ tenemos

$$\begin{aligned} -(\mathrm{id}+c(X))(B-XBX)(\mathrm{id}+c(X)) &= \\ -(\mathrm{id}+c(X))[B-(\mathrm{id}+c(X))^{-1}(\mathrm{id}-c(X))\times & \\ B(\mathrm{id}-c(X))(\mathrm{id}+c(X))^{-1}](\mathrm{id}+c(X)) &= \\ [(\mathrm{id}+c(X))B(\mathrm{id}+c(X))-(\mathrm{id}-c(X))B(\mathrm{id}-c(X))] &= \\ -2(Bc(X)+c(X)B). & \end{aligned}$$

*q.e.d.*

**Proposición 5.1.2** *La ecuación (5.1) puede resolverse explícitamente para la condición inicial $X(0) = X_0$. La solución es de la forma:*

$$X(t) = (\sinh(tA) + \cosh(tA)X_0) \cdot (\cosh(tA) + \sinh(tA)X_0)^{-1}.$$

*Demostración.–*

Dado $X_0 \in \Omega$ la curva integral en el grupo que pasa por $X_0$ no es más que la imagen por $c$ de la solución de la ecuación (5.2). La solución en el álgebra de Lie es

$$Y(t) = e^{-tB}Y_0e^{-tB}, \text{ con } Y_0 = c(X_0)$$

luego

$$\begin{aligned} X(t) = c(Y(t)) = c(e^{-tB}Y_0e^{-tB}) &= \\ (\mathrm{id}-e^{-tB}Y_0e^{-tB})(\mathrm{id}+e^{-tB}Y_0e^{-tB})^{-1} &= \\ (e^{tB}-e^{-tB}Y_0)e^{-tB}e^{tB}(e^{tB}+e^{-tB}Y_0)^{-1} &= \\ (e^{tA}(\mathrm{id}+X_0)-e^{-tA}(\mathrm{id}-X_0))(\mathrm{id}+X_0)^{-1} &\times \\ (\mathrm{id}+X_0)(e^{tA}(\mathrm{id}+X_0)+e^{-tA}(\mathrm{id}-X_0))^{-1} &= \\ [(e^{tA}-e^{-tA})+(e^{tA}-e^{-tA})X_0]\cdot[(e^{tA}+e^{-tA})+(e^{tA}-e^{-tA})X_0]^{-1} &= \\ [2\sinh(tA)+2\cosh(tA)X_0]\cdot[2\cosh(tA)+2\sinh(tA)X_0]^{-1}. & \end{aligned}$$

**Ejemplo 11** $B = \mathrm{id}$

$$h_{\mathrm{id}}(X) = \Re Tr(X)$$

y el flujo gradiente es

$$\mathrm{grad}_X h = \frac{1}{2}(\mathrm{id}-X^2).$$

En la siguiente figura podemos observar las líneas de flujo y el campo gradiente de esta función.

Integrar este campo consiste en hallar una curva $\alpha(t)$ tal que

$$\alpha'(t) = \mathrm{grad}_{\alpha(t)} h. \tag{5.3}$$

Para hacernos una idea de la solución pasamos al caso real

$$\alpha'(t) = \frac{1}{2}(1 - \alpha(t)^2) \tag{5.4}$$

y resolvemos en $\mathbb{R}$ esta ecuación diferencial.

$$\begin{aligned}
2\frac{d\alpha(t)}{1-\alpha^2(t)} &= dt \\
\left(\frac{1}{1-\alpha(t)} + \frac{1}{1+\alpha(t)}\right)d\alpha(t) &= dt \\
\ln\frac{1-\alpha(t)}{1+\alpha(t)} &= -t - cte \\
\frac{1-\alpha(t)}{1+\alpha(t)} &= ke^{-t} \\
\alpha(t) &= \frac{1-ke^{-t}}{1+ke^{-t}}.
\end{aligned}$$

Es decir, $\alpha(t) = c(Ke^{-t})$ es la solución general de (5.4).

Volviendo al caso matricial veamos que análogamente la curva

$$\alpha(t) = c(Ke^{-t}) = \frac{\mathrm{id} - Ke^{-t}}{\mathrm{id} + Ke^{-t}}$$

es la solución general de (5.3). Además, si le imponemos la condición inicial $\alpha(0) = X_0$ tenemos que $K = c(X_0)$. Es decir, la solución particular es entonces

$$\alpha(t) = c(c(X_0)e^{-t}).$$

En efecto, según la proposición 5.1.2 la solución es $Y(t) = e^{-t\mathrm{id}}Y_0e^{-t\mathrm{id}}$ pero como $Y_0$ y $e^{-t\mathrm{id}}$ conmutan, $Y(t) = e^{-2t\mathrm{id}}Y_0 = e^{-2t}Y_0$.

Si utilizásemos el lema 5.1.1 obtendríamos esta misma solución más rápidamente. De hecho, según este lema la solución de (5.3) es la imagen por $c$ de la solución de $c'(X) = -2c(X)$. Al pasar esta ecuación diferencial al caso real obtenemos inmediatamente que $\alpha(t) = ke^{-2t}$.

### 5.1.1. Líneas de flujo de la función de Morse con $B \neq \mathrm{id}$

Estudiemos ahora las líneas de flujo $\psi(t) = c_{\mathrm{id}}(e^{-Bt}Xe^{-Bt})$ que pasan por un cierto punto $P \in \Omega(\mathrm{id}) \cap G$ de la función $\Re Tr(BP)$. La matriz $B$ es diagonal real positiva, con autovalores ordenados en orden creciente. Para $t = 0$ sabemos que $X = c_{\mathrm{id}}(P)$. La curva $e^{-tB}$ tiende a cero cuando $t \to \infty$. Por tanto $\psi(+\infty) = \lim_{t\to+\infty} \psi(t) = c_{\mathrm{id}}(0) = \mathrm{id}$.

Por otra parte, $\psi(-\infty) = \lim_{t\to-\infty} \psi(t) = -\mathrm{id}$. Para probarlo usamos la regla de L'Hopital:

$$\begin{aligned} \lim_{t\to-\infty} (\mathrm{id} - e^{-tB}Xe^{-tB})(\mathrm{id} + e^{-tB}Xe^{-tB})^{-1} &= \\ \lim_{t\to-\infty} \left(-(-Be^{-tB}X)e^{-tB} - e^{-tB}X(-Be^{-tB})\right) \times & \\ \left(-Be^{-tB}Xe^{-tB} + e^{-tB}X(-Be^{-tB})\right)^{-1} &= -\mathrm{id} \end{aligned}$$

pues $B$ y $e^{-tB}$ conmutan.

Es decir, las líneas de flujo de los puntos de $\Omega(I)$ van de $I$ a $-I$.

## 5.2. Aplicación al cálculo de la categoría de $U(n)$

Al desarrollar la generalización de la aplicación de Cayley obtuvimos los $\Omega(A)$, abiertos contráctiles de $O_n(\mathbb{K})$ centrados en $A$. Estos abiertos los utilizamos ahora para calcular la categoría de $U(n)$ de un modo mucho más sencillo que el método clásico.

### 5.2.1. Demostración clásica

Reproducimos aquí el cálculo clásico de la categoría de $U(n)$ y $SU(n)$ de Singhof siguiendo el libro [10]. Como se verá a continuación este cálculo es mucho más complejo que el nuestro. Utiliza la exponencial para hacer las contracciones en el álgebra de Lie, con el inconveniente de que esta aplicación no tiene inversa global, lo que hace que tengamos que andar con bastante cuidado para poder volver al grupo. De hecho, quizá lo más complejo de la demostración sea precisamente el modo de escoger la rama de logaritmo apropiada.

Por cohomología tenemos que $\mathrm{cat}(SU(n)) \geq n-1$ y $\mathrm{cat}(U(n)) \geq n$. Probaremos que $n-1$ y $n$ sirven también como cota superior. En realidad, nos llega con probar que $\mathrm{cat}(SU(n)) \leq n-1$ ya que $U(n)$ es difeomorfo a $S^1 \times SU(n)$ y $\mathrm{cat}\, S^1 = 1$.

**Teorema 5.2.1** $\mathrm{cat}\, SU(n) \leq n-1$.

*Demostración.–*

Recordemos que si $A \in SU(n)$, $\det A = 1$ y $AA^* = id$. Sabemos también que estas matrices son diagonalizables vía otra matriz unitaria $U$, $UAU^{-1} = D$.

Busquemos pues un recubrimiento por abiertos contráctiles. Sean $\xi_1, \ldots, \xi_n$ $n$ complejos de módulo $|\xi_i| = 1$ tales que $\xi_1 \cdot \xi_2 \cdots \xi_n \neq 1$. Con esta condición conseguimos que no puedan aparecer todos a la vez como autovalores de una misma matriz $T \in SU(n)$ (porque si $T \in SU(n)$, $1 = \det T = \det D = \xi_1 \ldots \xi_n$). Tenemos entonces una partición de $SU(n)$ formada por los conjuntos

$$A_i = \{T \in SU(n) : \xi_i \text{ no es autovalor de T}\}.$$

Los $A_i$ son abiertos de $SU(n)$ pero en principio no tiene por qué ser conexos. Consideramos $B \subset A_i$ una de sus componentes conexas. Si probamos que $B$ es contráctil en $SU(n)$, tendremos que cada una de las componentes conexas de $A_i$ es contráctil en $SU(n)$, luego, como $SU(n)$ es conexo por caminos, cada $A_i$ sería un abierto categórico.

Sea

$$\begin{aligned} \phi : B &\longrightarrow \mathcal{H}(n) \\ T &\mapsto -i\log(T) \end{aligned}$$

donde $\mathcal{H}(n)$ denota el conjunto de las matrices hermíticas de orden $n$. Escogemos la rama de logaritmo que pasa por $\xi_i$ (véase [10, p. 274-5]) para los detalles de esta aplicación). Si denotamos por $\lambda_i$ a los autovalores de $T \in B$ tenemos que

$$\begin{aligned} \exp(i\phi(T)) &= \\ \exp(\log(T)) &= \\ U^{-1}\operatorname{diag}(\exp(\log(\lambda_i)))U &= \\ U^{-1}\operatorname{diag}(\lambda_i)U &= T, \end{aligned}$$

pues $\log(\lambda_i) = \ln(\lambda_i) + i\arg(\lambda_i)$ y $|\lambda_i| = 1$. Entonces,

$$1 = \det(T) = \det(\exp i\phi(T)) = \exp(iTr(\phi(T))),$$

por tanto, $Tr(\phi(T)) = 2\pi k$ para algún entero $k$. Ahora bien, $\phi$ y la traza son aplicaciones continuas y acabamos de ver que esta última toma valores aislados. Como $B$ es conexo, la traza tiene que tomar un valor constante $2\pi k_0$ para algún $k_0 \in \mathbb{Z}$. Definimos así la matriz constante hermítica $X_0 \in \mathcal{H}(n)$,

$$X_0 = \frac{2\pi k_0}{n}\mathrm{id}_n.$$

Veamos que $B$ puede contraerse a $\exp(iX_0)$. Para poder contraer usamos el hecho de que las matrices hermíticas $\mathcal{H}(n)$ forman un $\mathbb{R}$-espacio vectorial de manera que podemos contraer por rectas. Utilizamos depués la exponencial para pasar la homotopía lineal a $SU(n)$. Definimos la homotopía

$$\begin{aligned} H : B \times I &\longrightarrow SU(n) \\ (T,s) &\mapsto H(T,s) = \exp[i((1-s)\phi(T) + sX_0)]. \end{aligned}$$

Es obvio que

$$\begin{aligned} H(T,0) &= \exp(i\phi(T)) = T \text{ y} \\ H(T,1) &= \exp iX_0 \text{ constante.} \end{aligned}$$

Comprobemos ahora que esta homotopía está bien definida, es decir, $H(T,s) \in SU(n)$ para todo $s \in I$. En efecto, las matrices antihermíticas $\mathcal{A}_n = i\mathcal{H}_n$ forman el álgebra de Lie de $U(n)$. Además,

$$\begin{aligned} \det(H(T,s)) &= \\ \det(\exp(i((1-s)\phi(T) + sX_0))) &= \\ \exp(Tr(i((1-s)\phi(T) + sX_0))), \end{aligned}$$

pero

$$\begin{aligned} Tr(i((1-s)\phi(T) + sX_0)) &= \\ Tr(i(1-s)\phi(T)) + Tr(is\frac{2\pi k}{n}\mathrm{id}_n) &= \\ i(1-s)Tr(\phi(T)) + is\frac{2\pi k}{n}Tr(\mathrm{id}_n) &= \\ 2\pi ki(1-s) + is2\pi k &= 2\pi ki. \end{aligned}$$

Entonces,

$$\det(H(T,s)) = \exp(2\pi ki) = 1.$$

Obtenemos así que los $\{A_i\}_{1\leq i\leq n}$ forman un recubrimiento por abiertos contráctiles de $SU(n)$.

*q.e.d.*

### 5.2.2. Demostración mediante la transformación de Cayley

En cada punto de $U(n)$ tenemos un abierto contráctil difeomorfo al tangente (teorema 4.5.4). Todos estos abiertos nos permiten recubrir el grupo unitario. Así, si calculamos el número de abiertos con los que podemos recubrir todo el grupo obtendremos una cota superior para la categoría.

Sea $M \in U(n)$ una matriz unitaria y $A_k$ una matriz diagonal. Recordemos que $M \in \Omega(A_k)$ si y sólo si $A_k + M$ tiene inversa. Escogemos en primer lugar la matriz diagonal con todos sus elementos distinguidos iguales,

$$A = \mathrm{diag}(z, \ldots, z) = z\mathrm{id}.$$

Ahora bien, si $\lambda_1, \ldots, \lambda_n$ son los autovalores de $M$, entonces al diagonalizarla obtenemos:

$$A + M = U(z + \lambda_1, \ldots, z + \lambda_n)U^*$$

de manera que $A + M$ es inversible si y sólo si $\lambda_i \neq z, \forall i = 1, \ldots, n$.

Encontrar el recubrimiento es ahora una tarea sencilla. Como cada matriz tiene a lo sumo $n$ autovalores distintos bastará tomar $n+1$ complejos distintos $z_0, z_1, \ldots, z_n$, con $|z_i| = 1$. Si denotamos por

$$A_k = \mathrm{diag}(z_k, \ldots, z_k) = z_k\mathrm{id}$$

tenemos que, dada $M \in U(n)$, siempre hay algún $z_k$ que no es autovalor de $M$. Luego existirá algún $k$ para el cual $M \in \Omega(A_k)$. Es decir, $\{\Omega(A_i)\}_{0 \leq i \leq n}$ es un recubrimiento de $U(n)$ por abiertos contráctiles. Por tanto,

$$\mathrm{cat}\, U(n) \leq n.$$

Por otro lado, $H(U(n)) = \bigwedge(x_1, x_3, \ldots x_{2n-1})$ (véase [10, p. 273]). El cup producto más largo que podemos hacer es $x_1 \wedge x_3 \wedge \ldots \wedge x_{2n-1} \in H^{n^2}$, es decir, la longitud de los cup divisores de cero de estos grupos es exactamente $n$ de manera que

$$\mathrm{cat}\, U(n) = n.$$

## 5.3. Aplicación al cálculo de la categoría de $Sp(2)$

Recordemos que $Sp(n)$ es el grupo simpléctico de orden $n$ integrado por las matrices de la forma

$$Sp(n) = \{A \in \mathcal{M}_{n\times n}(\mathbb{H}) : AA^* = \mathrm{id}\}$$

y tiene dimensión $\dim Sp(n) = n(2n+1)$.

El argumento utilizado en el apartado anterior para calcular la categoría del grupo unitario no puede extenderse al grupo simpléctico $Sp(n)$. Habitualmente las siguientes condiciones son equivalentes:

1. $\mathrm{id}\xi + X$ no es inversible.
2. $\xi$ es un autovalor de $X$.
3. $\xi$ aparece como elemento de la matriz diagonal en la diagonalización de $X$.

Sin embargo, en los cuaternios la aplicación que a cada $\omega$ le asocia $\omega\xi$ no es lineal por la derecha, de manera que las condiciones (2) y (3) siguen siendo equivalentes pero no guardan ninguna relación con (1).

*Nota.—* La aplicación que lleva cada $\omega$ en $\xi\omega$ sí es lineal. Con la definición de autovalor como los cuaternios $\xi$ tales que $X\omega = \xi\omega$ sí se tiene la equivalencia entre las igualdades (1) y (2) (véanse propiedades 1.2.17 y 1.2.12). Pero en principio, con esta caracterización, no tenemos porqué tener diagonalización y ni siquiera sabemos cuántos autovalores de este tipo tiene una aplicación lineal o si los conjugados de un autovalor siguen siendo autovalores.

Tomamos en $Sp(2)$ los abiertos $\Omega(\mathrm{id})$ y $\Omega(-\mathrm{id})$. La unión de ambos recubre todo el grupo excepto la órbita $UP_1U^*$ de la matriz diagonal $P_1 = \mathrm{diag}(-1, 1)$. Veamos qué más abiertos necesitamos para recubrir $Sp(n)$.

**Teorema 5.3.1** $\mathrm{scatSp}(2) \leq 3$.

*Demostración.–*

En primer lugar comprobemos que $\Omega(\mathrm{id}) \cup \Omega(-\mathrm{id})$ cubren todo el grupo menos la órbita de $P_1$. En efecto, sea $X \in Sp(2)$; $X \in \Omega(\mathrm{id})$ si y sólo si $X + \mathrm{id}$ es inversible. Pero como $X$ puede diagonalizarse a una cierta matriz compleja $D$, $X = UDU^*$, que $X + \mathrm{id}$ tenga inversa es equivalente a que la tenga $D + \mathrm{id}$. Ahora bien, si $D = \mathrm{diag}(d_{11}, d_{22})$, $\mathrm{id} + D$ es inversible si y sólo si $d_{ii} \neq -1$ para $i = 1, 2$. Si ambos elementos fuesen $d_{11} = d_{22} = -1$, entonces $D + (-\mathrm{id})$ sería inversible, $X \in \Omega(-\mathrm{id})$. Así, los únicos elementos del grupo que nos quedan por cubrir son los de la órbita de $P_1$.

Sea ahora $P_2 = -P_1$. Para poder afirmar que $\{\Omega(\mathrm{id}), \Omega(-\mathrm{id}), \Omega(P_1), \Omega(P_2)\}$ es un recubrimiento por abiertos contráctiles de $Sp(2)$ lo único que necesitamos probar es que dada cualquier $A = (a_{ij}) = UP_1U^*$, alguna de las matrices $X_k = P_k + A$, $k = 1, 2$ es inversible.

Al ser $P_1$ una matriz idempotente lo serán también todas las de su órbita de manera que,

$$X_1^2 = 2\mathrm{id} + P_1 A + A P_1 = 2\,\mathrm{diag}(1 - a_{11}, 1 + a_{22}).$$

Esta matriz cuaterniónica, como es diagonal, es inversible si y sólo si no se anulan ninguno de los elementos de la diagonal. Es decir, $X_1^2$ tiene inversa (y por tanto $X_1$) si y sólo si $a_{11} \neq 1$ y $a_{22} \neq -1$. Al mismo tiempo, como $A$ es una matriz ortogonal, sus filas y sus columnas son ortonormales luego, si $a_{11} = 1$, necesariamente $a_{12} = a_{21} = 0$ y $a_{22} = \pm 1$. Al ser también $A$ de la órbita de $P_1$ ya tenemos que $a_{22} = -1$, es decir, $A = P_2$. Hemos obtenido entonces que $\Omega(P_1)$ cubre todos los elementos de la órbita de $P_1$ excepto $P_2$, que está en $\Omega(P_2)$. Queda así recubierto el grupo por estos cuatro abiertos contráctiles.

*q.e.d.*

Este resultado nos permite decir cuánto vale la categoría de $Sp(2)$ pues utilizando el método de Schweitzer, W. Singhof obtuvo el siguiente teorema [39, p. 147].

**Teorema 5.3.2** $\mathrm{cat}\, Sp(n) \geq n + 1$ *para* $n \geq 2$.

Nótese que el texto de Singhof [39, II] dice $\mathrm{cat}\, Sp(n) \geq n+2$ pero define la categoría como el menor número de abiertos categóricos que recubren toda la variedad.

Obsérvese que al ser la categoría fuerte una cota superior de la categoría en este caso coinciden ambas, $\mathrm{scat} Sp(2) = \mathrm{cat}\, Sp(2)$.

**Teorema 5.3.3** $\mathrm{cat}\, Sp(2) = 3$.

Es interesante notar que si buscamos el valor exacto de la categoría de $Sp(2)$, la cota inferior que nos da la longitud del cup producto no es suficiente. En efecto, la cohomología de los grupos simplécticos viene dada por [34, p. 119]

$$H(Sp(2)) = \Lambda(x_3, x_7, \ldots, x_{4n-1})$$

de manera que el cup producto no nulo más largo tiene $n$ factores, es decir,

$$\mathrm{l.c.p.} Sp(n) = n.$$

En el caso que nos ocupa, como $\dim Sp(2) = 10$, tendríamos $\Lambda(x_3, x_7)$ con $H^0 = \mathbb{R}$, $H^3 =< x_3 >$, $H^7 =< x_7 >$ y $H^{10} =< x_3, x_7 >$ y en los demás grados de cohomología no aparece nada.

## 5.4. Aplicación a la descomposición polar de una matriz

A raíz de la solución de la ecuación del flujo gradiente de las funciones “altura” obtenemos un resultado interesante que nos permite encontrar la descomposición polar de una matriz no regular.

Recordemos que [6, p. 15]:

**Proposición 5.4.1** *Cualquier matriz regular $A$ puede descomponerse de manera única como un producto*

$$A = PQ$$

*con $P$ una matriz hermítica definida positiva y $Q$ una matriz unitaria.*

Dynnikov y Veselov [41] prueban que:

**Proposición 5.4.2** *El flujo gradiente de la función “altura” $h_B$ en el grupo $G$ resuelve el problema de la descomposición polar de la matriz $B$.*

En efecto, sea $B = PQ$ la descomposición polar de $B$. Entonces, para cualquier matriz $Y \in G$ puede probarse:

$$\Re Tr(PY) \leq \Re Tr(P)$$

y entonces

$$\begin{aligned} h_B(X) = \Re Tr(BX) = \Re Tr(PQX) \leq \\ \Re Tr(P) = h_B(Q^*) = \Re Tr(BQ^*), \end{aligned}$$

luego $Q^* = X$ es el máximo de la función $h_B$ en $G$.

Puede verse también (con la expresión del índice que aparece en 2.4.6) que los otros puntos críticos no son máximos locales. De hecho, para casi todo punto inicial $X(0) = X_0 \in G$ la solución de la ecuación diferencial del gradiente de la proposición 2.4.2 verificará que $\lim_{t\to\infty} X(t) = Q^*$.

# Bibliografía


[1] A. Ambrosetti, Critical Point Theory, a short survey. *Preprint.*

[2] H. Aslaksen, Quaternionic Determinants,*The Mathematical Intelligencer* 18 (1996), no. 3, 56-65.

[3] F.F. Bonsall y J. Duncan, *Complete Normed Algebras.* Springer-Verlag, Berlin Heidelberg NY, 1973.

[4] J.L. Brenner, Matrices of Quaternions, *Pacific Journal Math.* Volume 1, n. 3 (1951), 329-335.

[5] H. Cartan, *Teoría elemental de funciones analíticas de una o varias variables complejas* (1 ed. española). Madrid: Selecciones Científicas, 1968.

[6] C. Chevalley, *Theory of Lie groups I* (8 ed.). Princeton: Princeton University Press, VIII, 1970.

[7] G. Choquet, *Cours dánalyse. Tome II: Topologie. Espaces topologiques et espaces métriques. Fonctions numériques. Espaces vectoriels topologiques* (2 ed). Paris: Masson et Cie, Éditeurs, XI, 1969.

[8] P.G. Ciarlet, *Introducción á análise numérica matricial e á optimización.* Santiago de Compostela: Servicio de Publicacións da Universidade de Santiago de Compostela, 1999.

[9] N. Cohen y S. De Leo, The quaternionic determinant, *J. Linear Algebra* 7 (2000), 100-111, electronic only.

[10] O. Cornea, G. Lupton, J. Oprea y D. Tanré, *Lusternik-Schnirelmann category.* Mathematical Surveys and Monographs, 103. Providence, RI: American Mathematical Society XVII, 2003.

[11] J. Dieudonné, *Eléments dÁnalyse.* Tome 3, Paris: Gauthier-Villars, 1970.

[12] H. Duan, Morse functions and Cohomology of Homogeneous Spaces, *Topics in Cohomological Studies of Algebraic Varieties.* B. Basel (ed), p. 105-133, Spinger-Verlag, 2006.

[13] S. Eilenberg e I. Niven. The fundamental theorem of algebra for quaternions, *Bulletin of the American Mathematical Society,* 52:246248, (1944).

[14] M. Farber, Topological Complexity of motion planning, *Discrete and Computational Geometry,* 29 (2003), no. 2, 211-221.

[15] T.M. Flett; *Differential analysis : differentiation, differential equations and differential inequalities.* Cambridge: Cambridge University Press, 1980.

[16] R. H. Fox, On the Lusternik-Schnirelmann category, *Ann. of Math.,* II, 42 (1941), 333-370.

[17] T. Frankel, Critical Submanifolds of the Classical Groups and Stiefel Manifolds, *Differential & Combinatorial Topology,* ed. S. S. Cairus, Princeton, p. 37-53, 1965.

[18] W. Greub, S. Halperin y R. Vanstone, *Connections, curvature, and cohomology. Vol. III: Cohomology of principal bundles and homogeneous spaces.* Pure and Applied Mathematics, 47-III. New York-San Francisco-London: Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers XXI, 1976.

[19] V. Guillemin y A. Pollack, *Differential topology.* Englewood Cliffs, N.J.: Prentice-Hall, Inc. XVI, 1974.

[20] P. Halmos, *Espacios vectoriales finito-dimensionales.* México, D.F. : Compañía Editorial Continental, 1965.

[21] A. Hatcher, *Algebraic topology.* Cambridge: Cambridge University Press XII, 2002.

[22] L. Huang y W. So, On left eigenvalues of quaternionic matrix, *Linear Algebra and its Applications* 323 (2001) 105-116.

[23] N. Jacobson, *Lectures in abstract algebra II: Linear algebra* (2 ed.). Graduate Texts in Mathematics 31. New York - Heidelberg - Berlin: Springer-Verlag XII, 1975.

[24] I. M. James. Lusternik-Schnierlmann Category, *Handbook of algebraic topology.* I.M. James, (ed.). Amsterdam: North-Holland X, p. 1293-1310, 1995.

[25] I. M. James, On category in the sense of Lusternik-Schnirelmann, *Topology*, 17 (1978), 331-348.

[26] S. Kobayashi y K. Nomizu, K. *Foundations of differential geometry. Vol. I & II.* New York-London-Sydney: Interscience Publishers a division of John Wiley and Sons, 1969.

[27] S. Lang, *Algebraic structures.* Reading, Mass.-Palo Alto-London-Don Mills, Ontario: Addison-Wesley Publishing Company VIII, 1967.

[28] J.M. Lee, *Introduction to smooth manifolds.* Graduate Texts in Mathematics, 218. New York: Springer XVII, 2002.

[29] S. De Leo, G. Scolarici y L. Solombrino, Quaternionic eigenvalue problem, *Journal of Mathematical Physics* 43 (2002), 5815-5829.

[30] S. De Leo, G. Ducati y V. Leonardi, Zeros of Unilateral Quaternionic Polynomials, *Electronic Journal of Linear ALgebra* 15 (2006), 297-313.

[31] I. Madsen, J. Tornehave, *From calculus to cohomology: de Rham cohomology and characteristic classes.* Cambridge: Cambridge University Press VII, 1997.

[32] L.M. Machado, F. Silva Leite, Optimization on Quadratic Matrix Lie Groups, *preprint.*

[33] M. Mimura, Homotopy theory of Lie Groups, *Handbook of algebraic topology.* I.M. James, (ed.). Amsterdam: North-Holland X, p. 951-991, 1995.

[34] M. Mimura, H. Toda, *Topology of Lie groups, I and II.* Translations of Mathematical Monographs, 91. Providence, RI: American Mathematical Society IV, 1991.

[35] J. W. Milnor, *Morse theory. Based on lecture notes by M. Spivak and R. Wells.* Annals of Mathematics Studies, No.51. Princeton, N.J.: Princeton University Press VI, 1963.

[36] J. W. Milnor, *Topology from a Differential Viewpoint.* University of Virginia Press, 1965.

[37] M. Postnikov, *Lie groups and Lie algebras.* Moscow : URSS, Lectures in Geometry 5, 1994.

[38] P. Schweitzer, Secondary Cohomology Operations Induced by the Diagonal Mapping, *Topology* 3 (1964-65), 337-335 .

[39] W. Singhof, On the Lusternik-Schnirelmann Category of Lie Groups I, II. *Math. Z.* 145 (1975), 11-116, 151 (1976), 143-148.

[40] J. Sotomayor, *Licoes de equacoes diferenciais ordinárias.* Projeto Euclides, 11. Rio de Janeiro: Instituto de Matematica Pura e Aplicada, CNPq. XVI, 1979.

[41] A.P. Veselov, I.A. Dynnikov, Integrable gradient flows and Morse theory. *St. Petersbg. Math. J.* 8, No.3, 429-446 (1997); translation from *Algebra Anal.* **8**, No.3 (1996), 78-103.

[42] A. H. Wallace, *An introduction to algebraic topology.* (International Series of Monographs in Pure and Applied Mathematics, Vol. 1.) London-New York-Paris: Pergamon Press VII, 1957.

[43] F. W. Warner, *Foundations of differentiable manifolds and Lie groups.* Glenview, Illinois-London: Scott, Foresman & Comp, 1971.

[44] S. Willard, *General topology.* Addison-Wesley Series in Mathematics. Reading, Mass. etc.: Addison-Wesley Publishing Company XII, 1970.

[45] H. Samelson, *Notes on Lie algebras.* 2nd. ed. Universitext. New York: Springer-Verlag XII, 1990.