María José Pereira Sáez

CATEGORÍA DE LUSTERNIK-SCHNIRELMANN Y FUNCIONES DE MORSE EN LOS ESPACIOS SIMÉTRICOS

121
2012
Publicaciones del Departamento de Geometría y Topología
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

María José Pereira Sáez

CATEGORÍA DE LUSTERNIK-SCHNIRELMANN Y FUNCIONES DE MORSE EN LOS ESPACIOS SIMÉTRICOS

Memoria realizada en el Departamento de Geometría y Topología de la Facultad de Matemáticas, bajo la dirección del Prof. Dr. Enrique Macías Virgós, para optar al grado de Doctor en Ciencias Matemáticas por la Universidad de Santiago de Compostela.

Se llevó a cabo su defensa el día 31 de enero de 2012 en la Facultad de Matemáticas de dicha universidad, obteniendo la calificación de Sobresaliente *cum laude.*

IMPRIME: Imprenta Universitaria
Pavillón de Servicios
Campus Universitario

ISBN: 978-84-89390-38-6

Dep. Leg.: C 1107-2012

A mis padres

Agradecimientos

Pienso que para muchos ésta es la parte de la tesis más difícil de escribir, no hay proposiciones, lemas ni teoremas; no hay que hacer ninguna demostración. Además, me parece que no es fácil percibir todas las personas que de algún modo han colaborado en esta tesis y es prácticamente imposible reflejar con fidelidad la realidad; yo haré lo que pueda. Vaya por delante mi agradecimiento a todos aquellos que de un modo u otro me han ayudado en su elaboración.

Gracias Quique por tu optimismo, por tu disponibilidad, por tu paciencia para leer y releer las infinitas versiones, por tus correcciones, por ponerme plazos, por tu empeño en hacerme "ver" las cosas, por todo lo que me has enseñado. Por presentarme a Profesores como Lucía Fernández o Daniel Tanré. Por las clases de Grupos de Lie, por dirigirme la tesis.

Antonio (Tato para los alumnos de Topología Algebraica), gracias por estar disponible y compartir ideas.

Mariquiña, gracias por toda la carrera y el doctorado"compartidos", por estar, por hablar, por transmitir humanidad en este ambiente.

De todos los *bolseiros* con los que he ido compartiendo sala, aunque sólo fuera por unos días, he aprendido algo, gracias; de modo especial a Silvia, Sandra, Miguel y Javi, aunque ahora estéis en la sala de "fuera".

A quien más debo fuera del ámbito de la facultad es a mis padres; ellos me enseñaron a pensar (y muchas otras cosas).

Gracias Puri por animarme a seguir siempre adelante, corriendo, nadando o viviendo; por estar al lado en gran parte de *a tese*, por estar dispuesta incluso a aprender un poco de Tex para ayudarme con algunas correcciones en el "pequeño Mac"(no será por lo poco que te gusta la informática).

Gracias Rosa por las correcciones que me hiciste, como siempre, tan precisas.

Gracias, en definitiva, a quien más se las debo; por todo.

La realización de esta Memoria (con la correspondiente participación en Congresos, Jornadas, Encuentros y similares) ha sido financiada por el Proyecto del MICINN MTM2008-05861.

Índice general

Introducción

Introducción

La categoría de Lusternik y Schnirelmann de un espacio topológico X es el menor entero n tal que X puede ser recubierto por $n+1$ abiertos contráctiles en X (abiertos categóricos). El primero en la lista de problemas de la teoría de invariantes homotópicos numéricos de T. Ganea es "calcular la categoría de variedades familiares: variedades de Stiefel, grupos de Lie, etc." [26]. La pregunta de Ganea es de 1970 y, sin embargo, aún no se ha podido responder completamente pues los avances son lentos y difíciles. La dificultad del cálculo directo ha tratado de paliarse introduciendo diferentes técnicas y aproximaciones algebraicas.

El objetivo inicial de la presente Memoria era calcular de modo sencillo la categoría de Lusternik-Schnirelmann de algunos grupos de Lie clásicos. Uno de los primeros métodos con los que nos encontramos al introducirnos en el estudio de este invariante topológico para dichos grupos fue el que utilizó W. Singhof en 1975 para calcular la categoría del grupo unitario $U(n)$ [76, 77]. Este autor obtiene un recubrimiento explícito formado por n abiertos categóricos. Considera abiertos del tipo $\Omega(z)$ formados por la matrices complejas $n \times n$ tales que $A - zI$ es inversible, donde $z \in \mathbb{C}$ es un complejo de norma 1 y prueba que sus componentes conexas son contráctiles utilizando la aplicación exponencial. Una adaptación de este método ha sido utilizada en 2008 por M. Mimura y K. Sugata [61] para calcular la categoría de los espacios simétricos $SU(n)/SO(n)$ y $SU(2n)/Sp(n)$.

La principal dificultad al tratar de extender el método de Singhof al grupo simpléctico $Sp(n)$ radica en que, en el contexto cuaterniónico, la condición "$A - \lambda I$ no es inversible" no guarda ninguna relación con los autovalores por la derecha de la matriz A. Un examen más detallado de esta cuestión muestra que hay que distinguir entre autovalores por la derecha y por la izquierda, y que la condición anterior es equivalente por definición a estos últimos. Otra dificultad es que la condición de ser inversible en el caso complejo depende del determinante, noción que, como veremos, es difícil de generalizar en el caso cuaterniónico. Este hecho ha sido el que nos movió a interesarnos en los autovalores por la izquierda de las matrices cuaterniónicas. Al comenzar a trabajar con el álgebra lineal cuaterniónica, observamos que el estudio de autovalores por la derecha está completamente establecido mientras que los autovalores por la izquierda están muy poco estudiados. No se sabe cuántos autovalores por la izquierda puede tener una matriz cuaterniónica de orden n ni cómo calcularlos. Los únicos resultados importantes conocidos hasta el momento son un teorema de 1985 de R.M.W. Wood [90], según el cual toda matriz cuaterniónica tiene al menos un autovalor por la izquierda, y el estudio de las matrices de orden dos que hicieron en el año 2001 L. Huang y W. So [33]. Estos autores prueban que una matriz cuaterniónica de orden dos puede tener uno, dos o infinitos autovalores por la izquierda. Nosotros daremos una nueva demostración basándonos en un método introducido en 2006 por De Leo *et al.* [54] para resolver ecuaciones cuadráticas unilaterales. Adaptando estos resultados a $Sp(2)$ hemos obtenido una caracterización de las matrices simplécticas de orden dos con infinitos autovalores. Haciendo uso de este resultado comprobamos que el método de Singhof no es eficaz para este grupo pues nunca podremos obtener un recubrimiento categórico formado por menos de cinco abiertos del tipo $\Omega(\lambda)$, aunque $\operatorname{cat}(Sp(2)) = 3$.

Por otra parte, el trabajo de Huang y So es un estudio caso por caso de naturaleza algebraica y difícil de generalizar para matrices de orden $n \geq 3$. Aunque no es el tema central de esta Memoria, proponemos una aproximación diferente construyendo una función característica μ para las matrices cuaterniónicas tal que las raíces de la ecuación $\mu(\lambda) = 0$ son los autovalores por la izquierda de la matriz. En general, esta μ no es un polinomio sino una función racional a partir de la cual haremos un estudio topológico del espectro por la izquierda de las matrices de orden dos y parcialmente para las de orden tres, calculando su diferencial y aplicando la teoría del grado. En principio, este nuevo método se podría generalizar de modo más sencillo a órdenes superiores. Además, esta función característica verifica el teorema de Cayley-Hamilton, es decir, la extensión natural de μ a las matrices cumple que $\mu_A(A) = 0$.

Para demostrar que los conjuntos $\Omega(z) \subset U(n)$ son contráctiles, W. Singhof utilizaba la aplicación exponencial. Esto lleva consigo la dificultad de tener que escoger una rama del logaritmo para hacer la inversa. De hecho, ni siquiera queda claro si los abiertos son conexos, por lo que habla de sus componentes conexas. Asimismo, viendo la demostración de Mimura y Sugata [61], observamos que no hay una manera elegante de pasar a los espacios homogéneos más sencillos, que son los espacios simétricos. En el estudio que proponemos aquí, la idea para probar que ese tipo de abiertos son contráctiles consiste en utilizar la transformación de Cayley, procedimiento que puede generalizarse de modo natural a los espacios simétricos. Esta transformación es clásica y fue definida por Cayley en 1846. Está definida en el abierto $\Omega \subset \mathcal{M}_{n\times n}(\mathbb{K})$ formado por las matrices tales que $I + X$ es inversible y viene dada por $c(X) = \frac{I-X}{I+X}$. Además cumple que $c^2 = \mathrm{id}$. Nosotros la generalizamos a la transformación

$$c_A(X) = c(A^*X)A^*,$$

definida en el abierto $\Omega(A) \subset \mathcal{M}_{n\times n}(\mathbb{K})$ de las matrices X tales que $A + X$ es inversible. De este modo podemos recubrir el grupo por abiertos contráctiles. La principal propiedad de cada c_A será que establece un difeomorfismo entre $\Omega(A) \cap G$ y el espacio tangente $T_{A^*}G$. Con estas herramientas daremos una demostración más sencilla del resultado de Singhof para $U(n)$. Además, extenderemos estos resultados a los espacios simétricos comprobando que todas las construcciones son compatibles con el llamado modelo Cartan, y obteniendo así abiertos contráctiles en espacios simétricos. En particular esto nos permitirá calcular de manera sencilla la categoría de $U(n)/O(n)$ y $U(2n)/Sp(n)$.

Para dar cotas de la categoría de $Sp(n)$ recurriremos al estudio de las funciones altura. Recordemos que en una variedad compacta la categoría LS (más uno) es una cota inferior para el número de puntos críticos de cualquier función diferenciable. A grandes rasgos, esto se debe a que el flujo del gradiente determina, para cada punto crítico, un abierto categórico. En el contexto de los grupos de Lie y espacios simétricos las funciones que han sido más estudiadas son las de Bott-Morse. Esto se debe a que tienen una expresión sencilla, $h_X(A) = \Re \operatorname{Tr}(XA)$. En particular, H. Kadzisa y M. Mimura [50] trabajan con estas funciones en $SU(n)/SO(n)$ y $SU(2n)/Sp(n)$ para determinar la longitud en conos de estos espacios. Lo que nosotros hemos obtenido es una simplificación del estudio de las funciones de Bott-Morse. En efecto, obtenemos que el flujo del gradiente de estas funciones puede integrarse en los espacios simétricos utilizando otra vez la transformación de Cayley. En particular, esto nos permite dar cartas locales para el conjunto de puntos críticos, ya que dado un punto crítico A de una función altura h^M_X en el espacio simétrico M, la transformación de Cayley generalizada establece un difeomorfismo entre el núcleo del operador hessiano, que es un espacio vectorial al que llamaremos $S^M(A)$, y el conjunto de puntos críticos $\Sigma(h^M_X) \cap \Omega(A)$ en un entorno de A. De ahí nuestro interés en establecer una teoría lo más general posible de las funciones altura en el modelo de Cartan de un espacio simétrico ya que los resultados conocidos hasta el momento se refieren sólo a estudios caso por caso.

Aunque hemos obtenido muchos resultados, nuestro enfoque deja igualmente una serie de problemas interesantes abiertos. En el ámbito del álgebra lineal es necesario desarrollar una teoría general de autovalores por la izquierda y funciones características. En espacios simétricos, la teoría de la descomposición polar generalizada permitirá caracterizar de modo general las funciones altura que son de Morse. Por último, la continuación natural del trabajo será intentar aplicar nuestro método de Cayley para calcular la categoría LS de las Grassmannianas y variedades de Stiefel.

Contenidos

Hemos estructurado la memoria en dos grandes bloques. En el primero de ellos recogemos los preliminares que nos han parecido necesarios para poder establecer los resultados del resto del trabajo.

Preliminares

Álgebra lineal cuaterniónica. El Capítulo 1 lo dedicamos al álgebra lineal cuaterniónica por la derecha, recogiendo en él las principales herramientas que necesitaremos a lo largo de la Memoria.

Las dificultades mayores que encontramos en este ámbito son la ausencia de determinante y la distinción entre autovalores por la derecha y por la izquierda. Se dice que un cuaternio $q \in \mathbb{H}$ es un autovalor por la derecha de la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si existe algún $v \in \mathbb{H}^n, v \neq 0$ tal que $Mv = vq$. Su cálculo es sencillo, de hecho, los autovalores por la derecha de una matriz cuaterniónica M de orden n no son más que los cuaternios similares a los autovalores complejos de su forma compleja $c(M) \in \mathcal{M}_{2n\times 2n}(\mathbb{C})$. Así, el espectro de M por la derecha, $\sigma_r(M)$, estará formado a los sumo por n clases de similitud. Un problema es que los autovectores asociados a un autovalor por la derecha no forman un subespacio vectorial pero aún así, el comportamiento es análogo al caso complejo. J.L. Brenner generaliza el lema de Schur a $\mathcal{M}_{n\times n}(\mathbb{H})$ en [4] y así obtenemos que las matrices normales se pueden diagonalizar. Además, F. Zhang extiende al caso cuaterniónico algunas descomposiciones matriciales como son la descomposición polar y la SVD en [92]. Estos resultados nos serán de gran utilidad a la hora de caracterizar las funciones altura en un grupo de Lie que son de Morse y simplificar el estudio de estas mismas funciones restringidas al modelo de Cartan de un espacio simétrico clásico.

Ha habido varios intentos de extender la idea de determinante al ámbito cuaterniónico, tarea complicada debido a la no conmutatividad de los cuaternios [1]. De hecho, no es posible adaptar completamente este concepto a $\mathbb{H}$; si se trata de extender este funcional multiplicativo de modo que coincida con el determinante usual sobre $\mathbb{R}$ y $\mathbb{C}$ se observa que necesariamente tendrá que tomar valores reales. Nos detendremos especialmente en el *determinante de Study* y su relación con otros determinantes pues será una herramienta imprescindible para construir una función característica para las matrices cuaterniónicas y para probar de forma sencilla que los conjuntos de Cayley son abiertos. Se define como $\mathrm{Sdet}(M) = (\det_{\mathbb{C}}(c(M)))^{1/2}$, con lo que toma valores reales y extiende el módulo del determinante usual. Su comportamiento es bastante similar al del determinante complejo; enunciamos sus principales propiedades en la Subsección 1.2.1. Entre los posibles determinantes cuaterniónicos destacan los *quasideterminantes*, introducidos en 1992 por I.M. Gelfand y V.S. Retakh en [28]. En realidad, la teoría que establecen estos autores no se reduce a $\mathbb{H}$ sino que trabajan en el ámbito de matrices con entradas en un anillo de división. Como ejemplo de la relación entre la teoría de Gelfand y Retakh y el determinante de Study generalizamos en la Subsección 1.2.5 la llamada identidad de Jacobi.

Nuestro interés en el estudio del álgebra cuaterniónica se debe a que buscamos aplicar para el cálculo de la categoría LS del grupo simpléctico $Sp(n)$ el método que desarrolla W. Singhof en [76, 77] para calcular la categoría del grupo unitario $SU(n)$. Dicho método emplea abiertos del tipo $\Omega(z)$, $z \in \mathbb{C}$, formados por las matrices A tales que $A - zI$ es inversible. Sin embargo, en el caso cuaterniónico, la condición de que $A - qI$ sea inversible no guarda ninguna relación con los autovalores por la derecha, sino que es equivalente a que $q \in \mathbb{H}$ no sea un autovalor por la izquierda.

A diferencia del espectro por la derecha, se sabe muy poco del espectro por la izquierda de una matriz cuaterniónica arbitraria. Dedicamos la Sección 1.3 a los resultados conocidos en este ámbito. Diremos que un cuaternio $\lambda \in \mathbb{H}$ es un autovalor por la izquierda de $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si existe algún vector no nulo $v \in \mathbb{H}^n$ tal que $Av = \lambda v$. En 1985, R.M.W. Wood demostró, usando métodos homotópicos, que cualquier matriz cuaterniónica tiene al menos un autovalor en este sentido [90]. Esta vez, el conjunto de los autovectores asociados a un autovalor por la izquierda sí forman un subespacio vectorial. Además, el espectro por la izquierda $\sigma_l(M)$ es compacto, aunque no necesariamente finito, ya que en cuanto al número de autovalores tenemos un resultado de L. Huang y W. So de 2001 [33] según el cual una matriz $M \in \mathcal{M}_{2\times 2}(\mathbb{H})$ puede tener uno, dos o infinitos autovalores. Asímismo, uno de estos autores, W. So, hizo en 2005 un estudio caso por caso de las matrices 3×3 cuaterniónicas obteniendo para cada caso diferentes polinomios cuyas raíces son los autovalores por la izquierda [79].

Parece natural que para calcular autovalores cuaterniónicos sea necesario saber hallar las raíces de algún tipo de función característica. Pero ni siquiera esto es tarea fácil para los polinomios cuaterniónicos. Por ahora sólo se ha probado la extensión del teorema fundamental del álgebra a los polinomios cuaterniónicos bilaterales con un único término de mayor grado [14] y sabemos que no puede verificarse para cualquier polinomio cuaterniónico ya que, por ejemplo, $x^n a - ax^n - 1$ es un polinomio que no tiene ninguna raíz en $\mathbb{H}$. Recogemos en la Subsección 1.4.2 el método que desarrollaron L. Huang y W. So [34] para hallar las raíces de un polinomio cuaterniónico mónico unilateral del tipo $x^2 + bx + c$; lo utilizan para poder calcular los autovalores de las matrices 2×2.

La definición de autovalor por la izquierda es equivalente a que $\mathrm{Sdet}(M-\lambda I)=0$. Es precisamente esta propiedad la que nos dará la clave para construir una función característica μ y estudiar el espectro por la izquierda ya que al calcular la diferencial de la función característica obtenemos ecuaciones lineales cuaterniónicas. En el caso de orden 2 la diferencial es del tipo ecuación de Sylvester, $\alpha x + x\beta = \gamma$, por lo que recogemos en la Subsección 1.5.1 el modo de resolverla y en 1.5.2 el método de D. Janovská y G. Opfer [46] para discutirla. Generalizamos este método en la Subsección 1.5.3 para adaptarlo a la situación de la linealización de la función característica en orden 3.

Espacios simétricos. Daremos una visión general de los espacios simétricos en el Capítulo 2, deteniéndonos un poco más en la estructura del modelo de Cartan, en el cual construiremos abiertos categóricos.

Un espacio simétrico Riemanniano puede pensarse como una variedad diferenciable tal que en cada punto la inversión geodésica es una simetría. Desde un punto de vista más global, un espacio simétrico es un G-espacio homogéneo G/H para algún grupo de Lie G, tal que el estabilizador H de un punto es un subgrupo abierto del conjunto de puntos fijos de una involución $\sigma\colon G \to G$. Esta definición incluye tanto los espacios globalmente simétricos Riemannianos como los pseudo-Riemannianos y los dotados simplemente de una conexión afín.

Desde un punto de vista geométrico, los espacios localmente simétricos se caracterizan porque la conexión es sin torsión y la curvatura es paralela.

Si a cada "par simétrico" (G, K, σ) le asociamos el "par" $(\mathfrak{g}, \mathfrak{h}, s)$ formado por sus álgebras de Lie y el morfismo inducido tenemos el espacio globalmente simétrico descrito algebraicamente. En esta interpretación algebraica se basó E. Cartan para establecer la clasificación de los posibles espacios simétricos Riemannianos irreducibles compactos de tipo clásico que recogemos en la Tabla de la página 35.

Una herramienta que ha mostrado ser muy útil a la hora de estudiar funciones de Morse en espacios simétricos (veánse, por ejemplo, los trabajos de S. Ramanujam [69] o H. Kadzisa y M. Mimura [49]) es el modelo de Cartan. Nos permite embeber el espacio simétrico $G/K \cong M$ en el grupo G al interpretar la variedad $M = \{h \in G\colon\ h = g\sigma(g)^{-1}, g \in G\}$ como una subvariedad de G, de hecho la componente conexa del neutro de $N = \{g \in G\colon \sigma(g) = g^{-1}\}$. Este modo de trabajar con un espacio simétrico será imprescindible para establecer la carta local del Teorema 6.2.1.

Funciones de Morse. La teoría de Morse proporciona técnicas que permiten analizar de modo eficiente la estructura homotópica de una variedad diferenciable estudiando las propiedades de determinadas funciones diferenciables sobre esa variedad. Nos permite encontrar la estructura celular de un CW-complejo [60] y obtener información sobre su cohomología [12].

Como puede verse en los trabajos de T. Frankel [23] y S. Ramanujam [69], las funciones de este tipo que se han considerado clásicamente en los grupos de Lie y espacios homogéneos han sido las funciones *altura* o *distancia*. Consideremos h la función altura en relación a un hiperplano tanto en un grupo de Lie G inmerso en un espacio euclídeo como en el modelo de Cartan de un espacio simétrico M embebido en G. Puede probarse que, salvo desplazamientos, h es la parte real de la traza, $h_X(A) = \Re\,\mathrm{Tr}(XA)$. Calcularemos el gradiente $\operatorname{grad} h_X$; esto nos permitirá afirmar que un punto $A \in G$ es crítico para la función altura en el grupo h_X^G si y sólo si $X^* = AXA$. También obtendremos el hessiano para este tipo de puntos.

Veremos que la transformación de Cayley permite dar una carta local del conjunto de puntos críticos a partir del núcleo del hessiano de estas funciones. Esto es importante porque el conjunto de puntos críticos está relacionado con la categoría LS de la variedad. De hecho, M. Reeken [71] prueba que $\operatorname{cat} M \leq \operatorname{cat} \Sigma$, donde Σ denota el conjunto de puntos críticos de una función diferenciable en la variedad M.

Resultados

Transformación de Cayley

Al comienzo del Capítulo 4 recogemos brevemente la definición y principales propiedades de la transformación de Cayley.

La transformación de Cayley clásica c está definida en el abierto Ω formado por la matrices cuyo espectro no contiene al -1. La introdujo A. Cayley en 1846 [5] como un modo de expresar una transformación ortogonal en coordenadas antisimétricas. Viene dada por

$$c(X) = \frac{I - X}{I + X}$$

y es involutiva. Será de utilidad para pasar de un grupo de Lie G a su álgebra de Lie $\mathfrak{g}$ y viceversa.

En principio, c se define en los grupos ortogonales clásicos $O(n)$, $U(n)$ y $Sp(n)$. Denotaremos por $O_n(\mathbb{K})$ a estos grupos, formados por las matrices tales que $AA^* = I$ con $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ó $\mathbb{H}$. Su álgebra de Lie está formada por las matrices antihermíticas (antisimétricas). Es sencillo comprobar que estas matrices no tienen autovalores reales no nulos.

Introducimos a continuación una generalización de la aplicación de Cayley que es apropiada para obtener un recubrimiento por abiertos categóricos de los grupos de Lie ortogonales. Viene dada por $c_A(X) = c(A^*X)A^* = A^*c(XA^*)$ y está definida en el abierto $\Omega(A) \subset \mathcal{M}_{n\times n}(\mathbb{K})$ de las matrices X tales que $A + X$ es inversible. Obtenemos para esta generalización propiedades análogas a las de la transformación clásica c (que corresponde al caso $A = I$). En particular veremos que la inversa de la transformación de Cayley c_A centrada en A es $(c_A)^{-1} = c_A^*$. Esta propiedad resulta particularmente interesante porque a partir de ella daremos el Teorema 4.3.2 por el que c_A establece un difeomorfismo entre el abierto $\Omega(A) \cap G$ y el espacio tangente $T_{A^*}G$.

En el método que proponemos en el Capítulo 8 para calcular la categoría LS, siguiendo a W. Singhof [76, 77], utilizaremos para recubrir el grupo abiertos del tipo $\Omega(\lambda)$ formados por las matrices A tales que $A - \lambda I$ es inversible. La transformación de Cayley generalizada c_A nos permitirá probar que estos abiertos son contráctiles en los grupos de Lie ortogonales.

Por otro lado, veremos que c_A nos permite linealizar el flujo del gradiente de las funciones altura h_X tanto en los grupos de Lie como en los espacios simétricos. Una consecuencia es que permite dar una carta local para el conjunto de puntos críticos modelada por el núcleo del Hessiano (ver Subsección 6.2) .

En orden a obtener la categoría LS de los espacios simétricos nos planteamos si es posible aplicar esta misma técnica en el ámbito de estos espacios. Pasamos entonces a demostrar que es posible extender la transformación de Cayley generalizada a los espacios simétricos, para lo que es esencial el Lema 4.4.1, que establece que para todo elemento A de un grupo G en el que esté definido un automorfismo involutivo, que es la restricción de un morfismo de $\mathbb{R}$-álgebras, σ, se verifica que

$$c_{\sigma(A)} \circ \sigma = \sigma \circ c_A.$$

Teniendo en cuenta que el modelo de Cartan de un espacio simétrico del tipo G/K viene dado por $M = \{h \in G \colon h = g\sigma(g)^{-1}, g \in G\}$ y que el autorfismo σ es $\mathbb{R}$-lineal para todos los espacios simétricos irreducibles simplemente conexos de la clasificación de É. Cartan, en la Subsección 4.4.2 utilizaremos el Lema 4.4.1 para probar que los abiertos $\Omega(A) \cap M$ en el modelo de Cartan M de un espacio simétrico son contráctiles.

Funciones de Morse

En el Capítulo 6 analizaremos el comportamiento de las funciones altura h_X^M en el modelo de Cartan de un espacio simétrico M. S. Ramanujam estableció en [69] un estudio caso por caso de la función h_I en determinados espacios homogéneos. Sin embargo, el Lema 4.4.1 nos permitirá estudiar al mismo tiempo todas las funciones altura en los espacios simétricos.

Describiremos en primer lugar el espacio tangente $T_A M$ y calcularemos el gradiente $\operatorname{grad} h_X^M$. La dificultad a la hora de obtener el gradiente de estas funciones altura es que, salvo en algunos casos particulares como los que recogen A.P. Dynnikov e I.A. Veselov en [84], el gradiente de h_X^M no es simplemente la restricción a M del gradiente de la función en el grupo h_X^G. De hecho, en la Subsección 6.2.4 desarrollaremos un ejemplo de una función para la que $\Sigma(h_X^G) \cap M = \emptyset$ y sin embargo, esa misma función altura en el modelo de Cartan sí tiene puntos críticos, $\Sigma(h_X^M) \neq \emptyset$.

Obtendremos que

$$(\operatorname{grad} h_X^M)(A) = \frac{1}{4}\left(\widehat{X} - A\sigma(\widehat{X})A\right),$$

donde $\widehat{X} = X^* + \sigma(X)$, expresión que nos permite caracterizar los puntos críticos y comprobar que para un espacio simétrico $M \cong G/K$, $A \in M$ es un punto crítico de la función altura en el modelo de Cartan h_X^M si y solo si es un punto crítico de $h^G_{\sigma(\widehat{X})}$ (ver Teorema 6.1.7).

Calcularemos el operador Hessiano de h_X^M en la Subsección 6.1.3 y veremos que la transfomación de Cayley generalizada c_A nos permite "linealizar" el flujo del gradiente de h_X^M. Para grupos de Lie y para el caso particular $A = I$ esta idea aparece en K. Volchenko y A. Kozachko [86]. En concreto, es posible integrar la ecuación diferencial del gradiente,

$$4\alpha' = \widehat{X} - \alpha\sigma(\widehat{X})\alpha,$$

convirtiéndola, mediante Cayley, en una ecuación diferencial

$$\beta' = (-1/4)\left(\beta\widehat{X}A^* + A^*\widehat{X}\beta\right)$$

en el álgebra de Lie, que es posible resolver explícitamente. Lo haremos en la Subsección 3.3.4.

Todos estos resultados unidos a que el automorfismo σ es compatible con la transformación de Cayley generalizada c_A nos permitirán dar cartas locales $\Sigma(h_X^M) \cap \Omega_M(A)$ para el conjunto de puntos críticos, modeladas por el núcleo del Hessiano

$$S^M(A) = \{\beta_0 \in T_A M \colon A^*\widehat{X}\beta_0^* + \beta_0^*\widehat{X}A^* = 0\}.$$

Con el objeto de establecer en la Sección 5.2 la caracterización en un grupo de Lie de las funciones altura que son de Morse, comprobaremos en 5.1 que para una función altura h_X se tiene

$$\Sigma(h_X^G) = \Sigma(h_S^G)U^*,$$

siendo $X = US$ la descomposición polar de X. Asímismo, si tenemos la descomposición en valores singulares $X = UDV^*$, se verifica que

$$\Sigma(h_X^G) = V^* \Sigma(h_D^G) U.$$

Un resultado análogo se prueba para la diagonalización de matrices. Además, estas descomposiciones no sólo transforman difeomórficamente el conjunto de puntos críticos sino que conservan el carácter no degenerado. Es por ello que los Lemas 5.1.1 y 5.1.3 y el Teorema 5.1.2 nos permiten reducir el estudio de las funciones altura h_X al caso en el que la matriz X es diagonal real no negativa.

Podemos establecer entonces el Teorema 5.1.5 que nos da la estructura del conjunto de puntos críticos de cualquier función altura. Se prueba que para una matriz arbitraria $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que $0 < t_1 < \cdots < t_k$ son los autovalores reales no nulos de XX^* (valores singulares de X) con multiplicidades $n_1, \ldots, n_k$, y n_0 la dimensión de su núcleo, se tiene

$$\Sigma(h_X^G) \cong O_{n_0}(\mathbb{K}) \times \Sigma(n_1) \times \cdots \times \Sigma(n_k)$$

donde $n_0 + n_1 + \cdots + n_k = n$. Los espacios $\Sigma(n)$ son uniones disjuntas de Grassmannianas $\Sigma(n) = G_0^n \sqcup G_1^n \sqcup \ldots \sqcup G_n^n$, con $G_p^n = \frac{O(n)}{O(p)\times O(n-p)}$. Una consecuencia de este Teorema es que una función altura en un grupo de Lie G, h_X^G es de Morse si y sólo si XX^* tiene todos sus autovalores distintos y no nulos, resultado que es conocido.

Nos gustaría poder obtener también en el ámbito de los espacios simétricos una caracterización similar de las funciones altura que son de Morse. Sin embargo, el método de reducción a las diagonales que establecimos en el caso de grupos de Lie no se puede extender tal cual. El problema está en que, como el flujo del gradiente en el modelo de Cartan M no es simplemente la restricción a M del flujo en el grupo G, al hacer la descomposición polar (o la SVD) de una matriz $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ ya no tenemos por qué permanecer dentro del modelo de Cartan. La idea consistiría en construir alguna descomposición polar generalizada del tipo de la que proponen autores como J.D. Lawson [56] o la de H.Z. Munthe-Kaas *et al.* [63]. En concreto, estos autores consideran la descomposición polar usual $X = US$ de una matriz como el resultado de tomar el subgrupo $K = U(n)$ en un grupo más amplio, por ejemplo $G = GL(n, \mathbb{C})$, y pensar en las matrices simétricas como una subvariedad transversa, exactamente el modelo de Cartan del espacio simétrico G/K. Sin embargo, esta teoría no está suficientemente desarrollada como para garantizar en el caso general una descomposición $X = US$ donde se cumpliese que $\sigma(U) = U$ y $\sigma(S) = S^{-1}$, lo que nos permitiría avanzar.

Autovalores por la izquierda

Como ya hemos dicho, la posible generalización al caso simpléctico del método de Singhof para calcular la categoría LS del grupo unitario requiere de un estudio de los autovalores por la izquierda de las matrices cuaterniónicas. Recogemos en el Capítulo 7 los resultados que hemos obtenido en este tema.

En primer lugar, probaremos en la Subsección 7.1.1 de un modo nuevo el resultado de Huang y So que establece el posible número de autovalores por la izquierda de una matriz 2×2. La demostración de Huang y So está basada en las fórmulas recogidas en la Subsección 1.4.2 para hallar las raíces de un tipo concreto de polinomio de grado dos, fórmulas establecidas previamente por estos mismos autores. La prueba que desarrollamos nosotros utiliza otro método para resolver ecuaciones polinómicas unilaterales, introducido por de Leo *et al.* en [54], que reduce el problema a encontrar los autovalores por la derecha de una matriz asociada a la ecuación a la que llaman matriz compañera; lo explicamos brevemente en esa misma Subsección.

Utilizamos estos resultados previos para obtener en 7.1.2 una caracterización de las matrices simplécticas 2×2 cuyo espectro por la izquierda es infinito. Tendremos que los únicos elementos de $Sp(2)$ con infinitos autovalores son producto $A = L_q \circ R_\theta$ de una traslación izquierda por un cuaternio unitario q y de una rotación de ángulo θ tal que $\operatorname{sen} \theta \neq 0$. Asimismo proporcionamos una expresión explícita para los autovalores por la izquierda de estas matrices. A partir de esta expresión podemos comprobar que dados cuatro cuaternios de norma 1 siempre hay matrices en $Sp(2)$ cuyo espectro por la izquierda los contiene. Será precisamente este resultado el que nos permitirá mostrar en el Capítulo 8 que el método de Singhof ya no proporciona la categoría LS en el caso del grupo $Sp(2)$.

Como se puede observar en la Sección 1.3, los únicos resultados sistemáticos sobre el espectro por la izquierda conocidos hasta ahora son de naturaleza algebraica y difícilmente generalizables a órdenes superiores a tres. El propio Wood afirma en [90] que "desafortunadamente, no parece haber una función determinante adecuada en el caso cuaterniónico para reducir el problema de los autovalores al teorema fundamental (del álgebra)". Proponemos en la Sección 7.2 utilizar el determinante de Study para construir una función característica en el ámbito cuaterniónico. Antes de establecer esta teoría comentamos brevemente los principales intentos que ha habido de extender la idea de polinomio característico a $\mathcal{M}_{n\times n}(\mathbb{H})$. El único caso que podría considerarse completamente satisfactorio es el de las matrices hermíticas, pues tienen todos sus autovalores reales y, por tanto, conmutan con los coeficientes de cualquier polinomio con coeficientes en $\mathbb{H}$.

Recordemos que un cuaternio λ es un autovalor por la izquierda de una matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si y sólo si la matriz $M - \lambda I$ es singular o, equivalentemente,

$\mathrm{Sdet}(M - \lambda I) = 0$, donde Sdet es el determinante de Study. Diremos que una aplicación $\mu\colon \mathbb{H} \to \mathbb{H}$ es una función característica para la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si, salvo una constante, $|\mu(\lambda)| = \mathrm{Sdet}(M - \lambda I)$. De este modo, λ estará en el espectro por la izquierda de M si y sólo si la función μ se anula en λ. Esta idea de función característica generaliza el polinomio característico usual en el caso complejo.

Para las matrices 2×2 se obtiene una función característica polinómica que coincide con la que propone Wood [90] (ver Def. 7.2.6). Por otro lado, los expresiones explícitas obtenidas para orden 3 concuerdan completamente con resultados anteriores de So [79].

Probamos que si una matriz A de orden 3 tiene alguna entrada nula fuera de la diagonal, entonces podemos escoger una función característica polinómica μ de grado tres. Pero en el caso general esta μ es una función racional con un punto de discontinuidad al que llamaremos *polo*. El modo de obtener estas funciones características se recoge en la Subsección 7.2.4.

A partir de esta μ y utilizando la teoría del grado realizamos en las Subsecciones 7.3.3 y 7.3.4 un estudio del espectro por la izquierda. Para ello recogemos muy brevemente en 7.3.1 las principales ideas sobre el grado topológico de una aplicación continua y establecemos en 7.3.2 las reglas de derivación en el ámbito no conmutativo. La idea fundamental consiste en calcular la diferencial de la función característica μ y, una vez linealizada, estudiar su rango. Pensamos que esta manera de abordar el estudio de los autovalores será más fácil de generalizar para las matrices de orden $n > 3$.

En el caso 2×2 comprobaremos que μ puede extenderse de manera continua a toda la esfera S^4. Así, tendremos que μ y el monomio λ^2 son homótopas y por tanto tienen el mismo grado topológico, dos. Podemos afirmar entonces que para una matriz genérica $A \in \mathcal{M}_{2\times 2}(\mathbb{H})$ el espectro por la izquierda $\sigma_l(M)$ está formado exactamente por dos autovalores. Caracterizamos también los casos especiales en los que hay uno o infinitos autovalores según el rango de la diferencial μ_*; nuestros resultados coinciden con los de Huang y So [33] pero tienen una interpretación geométrica clara, con lo que el estudio para las matrices 2×2 queda completamente cerrado con este nuevo método.

El método que recogimos en la Subsección 1.5.3 nos permitirá ahora calcular el rango de la diferencial $\mu_{*\lambda}$ para cada uno de los autovalores de una matriz de orden 3. A lo largo de 7.3.4 calcularemos cuáles son las ecuaciones que se obtienen al linealizar, tanto en el caso polinómino como en el no polinómico, y daremos algunos ejemplos de cálculo.

Aunque no guarda relación directa con la categoría LS, una cuestión que nos planteamos de manera natural al estudiar la función característica μ fue la posibilidad de extender el teorema de Cayley-Hamilton. Aparentemente, ninguna de las pruebas desarrolladas en el ámbito conmutativo puede adaptarse al caso cuaterniónico. Además, no hay una extensión obvia de una función cualquiera con variable en $\mathbb{H}$ a las matrices $\mathcal{M}_{n\times n}(\mathbb{H})$. Sin embargo, como nuestra función característica es un polinomio o una función racional, sí podemos extenderla de modo natural. Comprobaremos que con la extensión en este sentido de las funciones definidas en Def. 7.2.6 y 7.2.9 se verifica el teorema de Cayley-Hamilton, es decir, $\mu_A(A) = 0$. Cierran el Capítulo 7 las demostraciones correspondientes.

Categoría LS

Los resultados recogidos en los bloques anteriores los aplicaremos en el Capítulo 8 para discutir un nuevo método de obtener recubrimientos categóricos explícitos de algunos espacios simétricos y grupos de Lie clásicos; pensamos que podría adaptarse también a las variedades de Stiefel.

La categoría de Lusternik-Schnirelmann de un espacio topológico X fue introducida en 1934 en el contexto del cálculo de variaciones [57]. Se define como $\operatorname{cat} X = k$ si $k+1$ es el menor número de abiertos contráctiles en X (llamados abiertos categóricos) que se necesitan para recubrir X. La definición de abierto categórico y categoría de un espacio así como sus primeras propiedades quedan recogidas en la Sección 8.1.

Este invariante homotópico ha sido intensamente estudiado desde diferentes puntos de vista, como puede verse, por ejemplo en los resúmenes que hace I. M. James en [43] y [44] o en el libro de O. Cornea *et al.* [8]. En la actualidad es un campo de investigación muy activo, debido a su interés y aplicaciones en temas tan diversos como pueden ser sistemas dinámicos, topología simpléctica, teoría de foliaciones o complejidad topológica y robótica.

En una variedad diferenciable, $\operatorname{cat} X + 1$ es una cota inferior para el número de puntos críticos de cualquier función real, pues el gradiente permite asociar a cada punto crítico un abierto categórico. Por otra parte, una cota inferior (de Eilenberg) para la categoría LS es la longitud del cup producto en cohomología reducida (ver Prop. 8.2.4). Pero pese a estas cotas, el cálculo de la categoría LS es, en general, muy difícil. Como recordamos al principio de la introducción, el primero de la famosa lista de problemas que propone T. Ganea [26] es precisamente calcular la categoría LS de los grupos de Lie compactos y las variedades de Stiefel. Este tema sigue siendo central y se aborda mediante el uso de invariantes homotópicos, acotaciones cohomológicas y otras herramientas propias del ámbito de la topología algebraica. Por ejemplo, recientemente H. Kadzisa y M. Mimura [50] han usado el gradiente de funciones de Bott-Morse para determinar la longitud en conos, que es una cota superior para la categoría LS. Otros autores como N. Iwase, I.M. James o W. Singhof se han ocupado del problema, cuya importancia ha estimulado el desarrollo de herramientas topológicas y algebraicas.

En la Sección 8.4 hacemos un elenco de los avances más significativos en el cálculo de la categoría de grupos de Lie y espacios homogéneos. El método que nos inspira es el que estableció W. Singhof en 1975 [76]. Calculó la categoría de $SU(n)$, dando un recubrimiento explícito por n abiertos categóricos cuya contractibilidad probó usando la aplicación exponencial. Para los grupos ortogonales $SO(n)$ sólo se conoce la categoría de los de baja dimensión. I.M. James y W. Singhof calcularon en 1999 la categoría de $SO(5)$ [45] y N. Iwase, M. Mimura y T. Nishimoto llegan en 2005 hasta $SO(9)$ [42]. La de $SO(10)$ está calculada también por N. Iwase, K. Kikuchi y M. Miyauchien [38] mientras que la categoría de $Spin(9)$ aparece en [39].

En el caso simpléctico, el problema se presenta aún más difícil. La categoría de $Sp(3)$ la calcularon por primera vez en 2002 L. Fernández-Suárez, A. Gómez-Tato, J. Strom y D. Tanré en [20]; de manera independiente, N. Iwase y M. Mimura [40] llegaron al mismo resultado con técnicas distintas y probaron que cat $Sp(n) \geq n+2$ si $n \geq 3$. Es de destacar que desde el cálculo de la categoría de $Sp(2)$ por P. Schweitzer en 1965 [74] no se había producido ningún otro avance significativo.

Entre los resultdos recientes relativos a la categoría LS de los espacios homogéneos señalamos el de J. Korbaš [53] para algunas Grasmannianas $O(n)/O(k) \times O(n-k)$; el trabajo de T. Fukaya [24] sobre $SO(n+3)/SO(n) \times SO(3)$ y el de T. Nishimoto [66] sobre la categoría LS de las variedades de Stiefel $O(n)/O(n-k)$ que cumplan que $n \geq 2k$. A.L.Mare [59] ha calculado la categoría LS de algunas variedades bandera como $Sp(n)/\left(U(n_1) \times U(n_k)\right)$, $n_1 + \cdots + n_k = n$. Si nos centramos en los espacios simétricos compactos es de destacar el trabajo de M. Mimura y K. Sugata [61], en el que prueban por el método de Singhof que $\mathrm{cat}(U(n)/O(n)) = n$ y $\mathrm{cat}(U(2n)/Sp(n)) = n$. Otros casos se han resuelto usando la longitud del cup producto, por ser variedades kahlerianas. La estructura celular de estos espacios fue estudiada por H. Kadzisa y M. Mimura en 2008 [49]; en 2011 los mismos autores la usan para obtener la longitud en conos de algunos casos ya conocidos [50].

Recogemos al final de esta sección una tabla de Mimura y Sugata [61] con la categoría LS de casi todos los espacios simétricos de tipo clásico.

Veremos en la Sección 8.5 que la transformación de Cayley generalizada nos permite obtener de un modo sencillo la categoría de $U(n)$ usando los mismo abiertos que Singhof.

En cuanto a la categoría de los espacios simétricos $U(n)/O(n)$ y $U(2n)/Sp(n)$ el método que propusimos para $U(n)$ es también efectivo para estos espacios. La demostración sigue en parte el trabajo de Mimura y Sugata [61] en el que calculan la categoría extendiendo el método de Singhof y utilizando el modelo de Cartan de ambos cocientes. Pero cuando comprueban que la contracción no se sale del modelo de Cartan aparecen las mismas complicaciones técnicas que en Singhof. Además, en el caso de $U(2n)/Sp(n)$ parece haber una imprecisión en su demostración, ya que la variedad M es conexa (modelo de Cartan) mientras que $N = \{g \in G \colon \sigma(g) = g^{-1}\}$ no lo es. De nuevo, utilizando el modelo de Cartan y la transformación de Cayley generalizada obtenemos rápidamente en las Subsecciones 8.6.1 y 8.7.1 abiertos categóricos que recubren ambos espacios simétricos. Este método, más sencillo, tiene la ventaja de que no necesitamos construir una contracción en cada caso sino que, por el Corolario 4.4.3, tenemos que los abiertos del tipo $\Omega_M(A)$ son contráctiles en el modelo de Cartan de cualquier espacio simétrico clásico.

Estudiamos a continuación el grupo simpléctico de orden dos. Recordemos que P. Schweitzer había obtenido en 1965 [74] la categoría de este grupo, $\operatorname{cat} Sp(2) = 3$, pero no proporcionó un recubrimiento explícito. La teoría de Morse estudiada en los Capítulos 3 y 6 nos permitirá ahora obtener cuatro abiertos categóricos para este grupo. En concreto, en la Subsección 8.8.1 veremos que el conjunto de puntos críticos de la función de Morse h_X^G con $X = \operatorname{diag}(1, 2)$ es $\Sigma(h_X^G) = \{\pm I, \pm P\}$, donde $P = \operatorname{diag}(-1, 1)$, y que los cuatro abiertos asociados a estos puntos nos dan un recubrimiento por abiertos contráctiles. Este resultado completa la demostración abstracta dada por Schweitzer.

En un principio, pensamos que el método de autovalores de Singhof podría extenderse a los grupos simplécticos. Sin embargo, veremos en la Subsección 8.8.2 que ni siquiera en el caso de $Sp(2)$ nos da la categoría. En efecto, los abiertos de Singhof están formados por las matrices tales que $A - \lambda I$ es inversible y, como sabemos, esta condición es equivalente a que el cuaternio λ no sea un autovalor por la izquierda de la matriz A. La caracterización de las matrices simplécticas con infinitos autovalores (ver Teorema 7.1.7) así como la expresión obtenida en la Proposición 7.1.8 para los autovalores de este tipo de matrices nos permitieron dar en la Subsección 7.1.2 un método para construir matrices cuyo espectro contenga unos autovalores fijados. Así, comprobaremos (Teorema 8.8.2) que nunca cuatro abiertos asociados a autovalores pueden recubrir todo $Sp(2)$. Por otra parte, en 8.8.4 damos cinco abiertos contráctiles de este tipo que sí recubren el grupo.

Los resultados obtenidos en los capítulos anteriores serán de utilidad también para estudiar la categoría de algunos espacios simétricos con una métrica Kähler. La cota inferior que se usa en este caso es la longitud del cup producto [61], ya que la existencia de una forma simpléctica no degenerada hace que sea $l.c.p = \dim X/2$. Por otra parte, la cota superior es también $\dim X/2$ si el espacio es simplemente conexo. Por ejemplo, consideremos $Sp(2)/U(2)$. Este espacio simétrico es difeomorfo a la Grassmaniana real $SO(5)/(SO(2) \times SO(3))$, de dimensión 6, con lo que tenemos que $\operatorname{cat}(Sp(2)/U(2)) = 3$. En 8.10.1 obtendremos que $\{\Omega_M(\pm\mathbf{i}I), \Omega_M(\pm E)\}$, donde $E = \operatorname{diag}(\mathbf{i}, -\mathbf{i})$, forma un recubrimiento categórico explícito del modelo de Cartan de $Sp(2)/U(2)$.

En el caso de $Sp(3)/U(3)$, que también posee una estructura hermítica, la categoría es $\operatorname{cat} Sp(3)/U(3) = 3(3+1)/2 = 6$. Esta vez sólo podemos proporcionar un recubrimiento por ocho abiertos contráctiles que recubren el modelo de Cartan de este espacio simétrico. Tanto en este caso como en el anterior la comprobación de que se tiene un recubrimiento utiliza las expresiones para el determinante de Study de las matrices de orden 2 y 3 obtenidas en 7.2.3 y 7.2.4.

Otro resultado importante de este Capítulo es el siguiente. Supongamos que en $Sp(n)$ (análogamente en otros grupos ortogonales) tenemos una función altura h_X. Sabemos que las componentes conexas del conjunto crítico son de la forma

$$\Sigma[i_1, \dots, i_k] = Sp(n_0) \times G^{n_k}_{i_k} \times \dots \times G^{n_1}_{i_1}, \quad 0 \le i_j \le n_j,$$

donde G^p_q denota la grasmmanniana

$$Sp(p)/(Sp(q) \times Sp(p-q)).$$

Podemos entonces utilizar un refinamiento de la acotación del número de puntos críticos debida a Rudyak [73] para concluir que

$$\operatorname{cat} Sp(n) + 1 \le \sum (\operatorname{cat}_G \Sigma[i_1, \dots, i_k] + 1).$$

Este resultado generaliza un teorema reciente (2009) de M. Hunziker y M.R. Sepanski [35] que relaciona la categoría de $Sp(n)$ con la de las órbitas de la acción adjunta. En efecto, éste se deduce del nuestro tomando el caso particular de la función h_I.

Al final de la Memoria recogemos otras posibles aplicaciones, fuera del ámbito de la categoría LS, de los resultados obtenidos. La primera de ellas (Subsecc. 8.11.1), establece un método sencillo para calcular los autovalores por la izquierda de las matrices simplécticas de orden 2 usando la transformación de Cayley; proporciona también los autovectores asociados. La segunda, establece la forma de las matrices 2×2 antihermíticas con infinitos autovalores (ver 8.11.1).

Particularmente importante sería poder aplicar nuestro estudio de las funciones de Bott-Morse a la descomposición polar generalizada de la Subsección 8.11.2. En efecto, en los grupos de Lie es bien conocido que la parte unitaria U de la descomposición polar $X = HU$ corresponde al valor mínimo en el grupo unitario de la función altura h_X. De este modo, varios métodos numéricos basados en el "descenso del gradiente" sirven para calcular U dada X. En nuestro caso, al resolver explícitamente las ecuaciones del gradiente parece posible generalizar ese método a cualquier espacio simétrico. Por otra parte, sería de gran utilidad obtener resultados generales para una descomposición de este tipo compatible con el modelo de Cartan, puesto que, conforme a lo estudiado en el Capítulo 6, nos permitiría dar una caracterización general de las funciones altura en un espacio simétrico que son de Morse.

Parte I

Preliminares

Capítulo 1

Álgebra lineal cuaterniónica

El álgebra lineal sobre los cuaternios no es una simple generalización del caso complejo. Si se considera la estructura de $\mathbb{H}$-espacio vectorial por la derecha hay muchas analogías. Por ejemplo, podemos extender descomposiciones clásicas como son la descomposición en valores singulares o la descomposición polar. Sin embargo, la no conmutatividad de los cuaternios hace que surjan algunas dificultades; de entre ellas, las más señaladas son la ausencia de determinante y el hecho de que los autovectores asociados a un autovalor por la derecha no forman un subespacio vectorial.

Los autovalores por la derecha están muy estudiados [2] y su cálculo es sencillo. Dada una matriz cuaterniónica M diremos que un cuaternio $q \in \mathbb{H}$ es un autovalor por la derecha de M si existe algún vector no nulo $v \in \mathbb{H}^n$ tal que $Mv = vq$. El comportamiento de este tipo de autovalores es análogo al caso complejo; de hecho, podemos diagonalizar matrices normales.

En nuestro caso, el estudio de los autovalores cuaterniónicos se debe a la intención de extender al grupo simpléctico $Sp(n)$ el método que utiliza W. Singhof [76, 77] para calcular la categoría LS del grupo unitario $U(n)$. Este método utiliza abiertos contráctiles formados por matrices A que verifican una condición del tipo $A - \lambda I$ inversible. Sin embargo, esta condición no guarda ninguna relación con los autovalores por la derecha sino con los autovalores por la izquierda. Un cuaternio λ es autovalor por la izquierda de $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si y solo si $M - \lambda I$ es no inversible.

A diferencia del espectro por la derecha, el espectro por la izquierda no es sencillo de estudiar y, por ahora, se tienen muy pocos resultados. En este caso, los autovectores sí forman un subespacio vectorial. Su estudio es complicado y, como veremos, el comportamiento de los autovalores por la izquierda es muy diferente al caso complejo. En 1985 R. M .W. Wood probó que el espectro por la izquierda de cualquier matriz cuaterniónica es no vacío [90]. Esta demostración utiliza métodos homotópicos y por el momento no se conoce ninguna prueba más sencilla. En cuanto al número de autovalores, solo tenemos un resultado de L. Huang y W. So según el cual una matriz cuaterniónica de orden 2 puede tener uno, dos o infinitos autovalores por la izquierda [33]. Para órdenes superiores ni siquiera se sabe si una matriz de orden n puede tener un espectro finito formado por más de n elementos. Otra de las dificultades es que el espectro por la izquierda no es invariante por semejanza, es decir, las matrices M y UMU^* pueden tener autovalores distintos. Además, los espectros por la izquierda y por la derecha, en principio, no guardan ninguna relación, a no ser que la matriz o los autovalores sean reales.

En este Capítulo recogemos las principales nociones del álgebra lineal (por la derecha) cuaterniónica. La mayor parte de estos resultados serán necesarios a lo largo de la Memoria. En otros casos servirán para hacer ver la dificultad del estudio de los autovalores por la izquierda.

1.1. Resultados básicos

1.1.1. Cuaternios

Los cuaternios $\mathbb{H}$ forman un álgebra no conmutativa sobre $\mathbb{R}$, asociativa y con uno. Es un anillo de división

$$\mathbb{H} = \{t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \colon t, x, y, z \in \mathbb{R}\},$$

donde

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

Si $q = t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \in \mathbb{H}$ denotaremos

$$\begin{aligned} \Re(q) &= t, \\ \Im(q) &= x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \end{aligned}$$

las partes real e imaginaria de q respectivamente. El conjugado de q es

$$\overline{q} = t - x\mathbf{i} - y\mathbf{j} - z\mathbf{k}.$$

Definición 1.1.1. Se dice que $q, q' \in \mathbb{H}$ son *similares* si existe algún cuaternio no nulo u tal que $q' = uqu^{-1}$.

Teorema 1.1.2. *Dos cuaternios son similares si y solo si tienen la misma norma y la misma parte real. En particular,*

1. *todo cuaternio q es similar a su conjugado $\overline{q}$;*
2. *todo cuaternio $q \in \mathbb{H}$ es similar a un complejo, $\Re(q) \pm |\Im(q)|\mathbf{i}$.*

Denotaremos por $\mathbb{H}_0 \cong \mathbb{R}^3$ el espacio vectorial real de los cuaternios con parte real nula. Aquí el producto escalar viene dado por $\langle \xi, \xi' \rangle = -\Re(\xi\xi')$. Una base ortonormal para este producto es $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Si $\xi \in \mathbb{H}_0$ tenemos que $\overline{\xi} = -\xi$ y $\xi^2 = -|\xi|^2$. El conjunto de los vectores de $\mathbb{H}_0$ con norma 1 coincide con el de los cuaternios similares a la unidad imaginaria $\mathbf{i}$.

1.1.2. El espacio cuaterniónico $\mathbb{H}^n$

Si queremos obtener en este ámbito los resultados habituales para matrices asociadas a una aplicación lineal, es necesario considerar el espacio cuaterniónico $\mathbb{H}^n$ como un $\mathbb{H}$-espacio vectorial por la derecha.

El producto hermítico que utilizaremos en $\mathbb{H}^n$ es $\langle u, v \rangle = u^* v$.

1.1.3. Matrices cuaterniónicas

Dada una matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$, denotaremos por $\overline{M}$ a su conjugada, M^T a la matriz traspuesta de M y $M^* = (\overline{M})^T$.

Definición 1.1.3. Diremos que $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ es *normal* si $MM^* = M^*M$, *hermítica* si $M^* = M$, *antihermítica* si $M^* = -M$ y *simpléctica o unitaria* si $M^*M = I$; donde I denota la matriz identidad de orden n.

F. Zhang recoge en [92] las principales propiedades de las matrices cuaterniónicas.

1.1.4. Forma compleja de una matriz cuaterniónica

Dada M una matriz cuaterniónica de orden $n \times n$ puede expresarse de forma única como $M = X + \mathbf{j}Y$ con $X, Y \in \mathcal{M}_{n\times n}(\mathbb{C})$.

Definición 1.1.4. Llamamos *forma compleja* de $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ a

$$c(M) = \begin{pmatrix} X & -\overline{Y} \\ Y & \overline{X} \end{pmatrix} \in \mathcal{M}_{2n\times 2n}(\mathbb{C}). \tag{1.1}$$

La base de $\mathbb{C}^{2n} \cong \mathbb{H}^n$ como $\mathbb{C}$-espacio vectorial que aquí se utiliza es

$$\{e_1, \ldots, e_n, \mathbf{j}e_1, \ldots, \mathbf{j}e_n\}$$

con $\{e_1, \ldots, e_n\}$ los vectores de la base canónica de $\mathbb{H}^n$.

Este modo de expresar las matrices de $\mathcal{M}_{n\times n}(\mathbb{H})$ es bastante efectivo pues permite aprovechar muchas de las propiedades de las matrices complejas.

Proposición 1.1.5. *Sean $M, N \in \mathcal{M}_{n\times n}(\mathbb{H})$, entonces:*

1. $c(M+N) = c(M) + c(N)$;
2. $c(tM) = tc(M)$ *para* $t \in \mathbb{R}$;
3. $c(M \cdot N) = c(M) \cdot c(N)$; *en particular,* $c(M)$ *es inversible sii* M *es inversible;*
4. $c(M^*) = c(M)^*$;
5. $c(M)$ *es unitaria, hermítica o normal sii* M *es unitaria, hermítica o normal respectivamente;*
6. $\det c(M) \geq 0$ *es un número real no negativo.*

1.1.5. Autovalores por la derecha

La teoría de autovalores por la derecha de matrices cuaterniónicas está bien establecida, como puede verse, entre otros, en [2, 4, 92]. La existencia de estos autovalores se prueba algebraicamente pero, recientemente, A. Baker [2] ha proporcionado un argumento topológico basado en el teorema del punto fijo de Lefschetz.

Definición 1.1.6. Se dice que un cuaternio $q \in \mathbb{H}$ es un *autovalor por la derecha* de la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si existe un vector $v \in \mathbb{H}^n$, $v \neq 0$, tal que $Mv = vq$.

Proposición 1.1.7. *Sea $q \in \mathbb{H}$ un autovalor por la derecha de $M \in \mathcal{M}_{n\times n}(\mathbb{H})$. Entonces, cualquier cuaternio similar a q es también autovalor de M. En concreto, si v es un q-autovector, dado $x \in \mathbb{H}$ no nulo, vx es un $(x^{-1}qx)$-autovector de M.*

Demostración. $M(vx) = (Mv)x = vqx = (vx)x^{-1}qx$. □

Queda claro pues que, con esta definición, el conjunto $\Sigma(q)$ de autovectores asociados a un autovalor 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.

La siguiente proposición, consecuencia de las propiedades vistas hasta ahora, muestra que el cálculo de los autovalores por la derecha de una matriz cuaterniónica es bastante sencillo.

Proposición 1.1.8. *Los autovalores por la derecha de una matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ son los cuaternios similares a los autovalores complejos de su forma compleja $c(M)$.*

Nótese que para una matriz compleja del tipo $c(M) = \begin{pmatrix} X & -\overline{Y} \\ Y & \overline{X} \end{pmatrix}$ de orden $2n \times 2n$ cada vez que aparece un autovalor aparece también su conjugado, de modo que, como mucho, tendrá n autovalores no conjugados.

Corolario 1.1.9. *M no puede tener más de n autovalores no similares.*

Ejemplo 1.1.10. Sea $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Se comprueba directamente que sus autovalores cuaterniónicos son las soluciones de la ecuación $q^2 = -1$, es decir, todos los elementos de $\mathbb{H}$ similares a $q = \mathbf{i}$, que son los cuaternios de módulo 1 en $\langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$.

Teorema 1.1.11. *Sean $M, N \in \mathcal{M}_{n\times n}(\mathbb{H})$ matrices semejantes, es decir, tales que $N = BMB^{-1}$ con B una matriz cuadrada inversible. Entonces M y N tienen los mismos autovalores por la derecha.*

1.1.6. Diagonalización de matrices normales

Veamos ahora que toda matriz normal puede diagonalizarse. Utilizaremos la siguiente generalización del lema de Schur.

Lema 1.1.12. *Toda matriz cuaterniónica es triangularizable por una matriz simpléctica.*

Este resultado fue probado por J.L. Brenner en [4, p.331].

Teorema 1.1.13. *Una matriz cuaterniónica es diagonalizable si y solo si es normal.*

Demostración. Sea $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ tal que $MM^* = M^*M$. Por el Lema 1.1.12, existen matrices T triangular superior y U unitaria tales que $M = UTU^*$. Despejando T en la igualdad anterior, $T = U^*MU$ y así,

$$\begin{aligned} TT^* &= (U^*MU)(U^*M^*U) = U^*MM^*U = \\ & \quad U^*M^*UU^*MU = (U^*MU)^*(U^*MU) = T^*T. \end{aligned}$$

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

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

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

Se deduce de los resultados que hemos recogido sobre los autovalores por la derecha que las matrices cuaterniónicas normales no sólo son diagonalizables sino que se pueden diagonalizar a una matriz compleja. También pueden extenderse al ámbito cuaterniónico las siguientes descomposiciones clásicas de matrices.

1.1.7. Descomposición polar

Es fácil comprobar que los autovalores de una matriz hermítica son reales. Cuando éstos son no negativos decimos que la matriz es semidefinida positiva.

Teorema 1.1.14. *Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ existen una matriz cuaterniónica hermítica semidefinida positiva H y una matriz cuaterniónica unitaria U tales que*

$$M = HU.$$

Además, si M es de rango máximo, estas matrices son únicas.

Zhang prueba este teorema en [92]. Utiliza que MM^* es una matriz normal y, por tanto, diagonalizable a una matriz cuyas entradas son los cuadrados de los valores singulares de M. Después aplica el teorema de descomposición polar en $\mathcal{M}_{2n\times 2n}(\mathbb{C})$ a la forma compleja de M,

$$c(M) = AB$$

con $A, B \in \mathcal{M}_{2n\times 2n}(\mathbb{C})$ tales que A es hermítica semidefinida positiva y B unitaria. Finalmente, comprueba que $A = c(H)$ y $B = c(U)$ para H y U matrices cuaterniónicas tales que H es hermítica semidefinida positiva y U unitaria.

1.1.8. Descomposición SVD

Para cualquier matriz cuaterniónica M en $\mathcal{M}_{n\times n}(\mathbb{H})$ la matriz M^*M es hermítica y todos sus autovalores son reales no negativos. Además, M^*M y MM^* tienen los mismos autovalores.

Definición 1.1.15. Sea $M \in \mathcal{M}_{n\times n}(\mathbb{H})$. Las raíces cuadradas no negativas de los n autovalores de M^*M son los *valores singulares* de M.

Lema 1.1.16. *Para $M \in \mathcal{M}_{n\times n}(\mathbb{H})$, $t \in \mathbb{R}$ es un valor singular de M sii existen vectores $u, v \neq 0$ tales que $Mv = tu$, $M^*u = tv$.*

De la extensión del teorema de descomposición polar a los cuaternios se deduce la descomposición en valores singulares para matrices cuaterniónicas.

Teorema 1.1.17. *Sea $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ una matriz de rango r. Entonces existen dos matrices $U, V \in \mathcal{M}_{n\times n}(\mathbb{H})$ unitarias tales que*

$$U^*MV = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}$$

donde $D_r = \operatorname{diag}(d_1, \ldots, d_r)$ con d_i los valores singulares positivos de M.

En particular, si M es de rango máximo, podemos tomar

$$D = \operatorname{diag}(d_1, \ldots, d_n),$$

U una matriz unitaria cuyas columnas sean autovectores de MM^* y V una unitaria cuyas columnas sean autovectores de M^*M. Entonces

$$M = UDV^*.$$

1.2. Determinante de Study

No es fácil extender a $\mathbb{H}$ la noción de determinante debido a la no conmutatividad de los cuaternios. El primero en proponer un funcional análogo para las matrices cuaterniónicas fue A. Cayley en el año 1845, pero no logró generalizar las propiedades del determinante usual. Hasta 75 años después no hubo ningún avance significativo. En la segunda edición del *Elements of Quaternions* de W.R. Hamilton, de 1889, el editor añadió un apéndice sobre este tema pero no era más que una reelaboración del artículo de Cayley. Diez años más tarde aparece un artículo de J.M. Pierce que, de nuevo, no pasa de una cuidada estructuración de la teoría del determinante de Cayley. Por fin, en 1920, E. Study propone transformar una matriz cuaterniónica $n \times n$ en una compleja $2n \times 2n$ y a ésta última hacerle el determinante complejo.

Se observa, sin embargo, que no es posible generalizar a los cuaternios la noción de determinante ya que, necesariamente tomará valores en $\mathbb{R}$. Sí se puede generalizar el módulo del determinante, $|\det|$ (es decir, con valores reales), y esto es lo que hizo Study. En [1] y [7] H. Aslaksen y N. Cohen respectivamente resumen la teoría general de los determinantes cuaterniónicos.

1.2.1. Definición y propiedades

Definición 1.2.1. Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ se define el *determinante de Study* de M como

$$\operatorname{Sdet}(M) := (\det c(M))^{\frac{1}{2}},$$

donde $c(M)$ es la forma compleja de M que hemos definido en 1.1.4.

Nótese que salvo el exponente 1/2, este determinante es el mismo que el del Teorema 8.1 de [92]. Aquí normalizamos el exponente para que $\mathrm{Sdet}(M) = |q_1 \ldots q_n|$ cuando consideramos una matriz diagonal $M = \mathrm{diag}(q_1, \ldots q_n)$.

Proposición 1.2.2. *El determinante de Study verifica las siguientes propiedades:*

1. $\mathrm{Sdet}(MN) = \mathrm{Sdet}(M)\,\mathrm{Sdet}(N)$*;*

2. $\mathrm{Sdet}(M) > 0$ *si y sólo si* M *es inversible;*

3. *si la matriz* M *es compleja* $\mathrm{Sdet}(M) = |\det(M)|$*;*

4. *si* M, N *son matrices semejantes, entonces* $\mathrm{Sdet}(M) = \mathrm{Sdet}(N)$*.*

Puede probarse que es el único funcional que verifica estas propiedades (véase [7]).

Además de las que acabamos de ver estas otras facilitan mucho su cálculo.

Proposición 1.2.3. *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 nulas 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. *para las matrices permutación de filas y columnas* P_{ij},

$$\mathrm{Sdet}(P_{ij} M P_{ij}) = \mathrm{Sdet}(M),$$

donde llamamos P_{ij} *a la matriz resultante de intercambiar en la matriz identidad las filas* i, j*;*

5. $\mathrm{Sdet}(M^*) = \mathrm{Sdet}(M)$*; en particular, si* M *es una matriz simpléctica entonces* $\mathrm{Sdet}(M) = 1$*.*

Aunque no las hemos encontrado explícitamente en la literatura, también nos harán falta las siguientes propiedades.

Propiedad 1.2.4. *Sea* $X = \left(\begin{array}{c|ccc} q & x_{12} & \cdots & x_{1n} \\ \hline 0 & & & \\ \vdots & & M & \\ 0 & & & \end{array}\right) \in \mathcal{M}_{n\times n}(\mathbb{H})$ *con* $q \in \mathbb{H}$; *entonces,*

$$\mathrm{Sdet}(X) = |q|\,\mathrm{Sdet}(M).$$

Demostración. Si $q = 0$ es trivial. Si $q \neq 0$, podemos suponer que $q = 1$. En este caso, si $x_{1k} = a_{1k} + \mathbf{j}b_{1k}$ y $M = A + \mathbf{j}B$, al desarrollar el determinante de la forma compleja $c(X)$ por la primera columna obtenemos

$$\begin{aligned}
\det(c(X)) &= \det\left(\begin{array}{ccc|c|ccc} & & & 0 & & & \\ & A & & \vdots & & -\overline{B} & \\ & & & 0 & & & \\ \hline b_{12} & \cdots & b_{1n} & 1 & \overline{a}_{12} & \cdots & \overline{a}_{1n} \\ & & & 0 & & & \\ & B & & \vdots & & \overline{A} & \\ & & & 0 & & & \end{array}\right) \\
&= (-1)^{2n+1}\det\begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix}. \quad \square
\end{aligned}$$

Teorema 1.2.5. *Dada una matriz formada por cajas* M, N *de órdenes* m *y* n *respectivamente, se verifica que*

$$\mathrm{Sdet}\begin{pmatrix} M & 0 \\ * & N \end{pmatrix} = \mathrm{Sdet}(M)\,\mathrm{Sdet}(N).$$

Demostración. Una matriz de orden 2 de este tipo solo puede ser

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$$

y por la Propiedad 1.2.4 sabemos que su Sdet es $|ad|$.

Para orden 3, los dos únicos casos posibles son

$$\left(\begin{array}{c|cc} a & 0 & 0 \\ \hline * & b & c \\ * & d & e \end{array}\right) \quad \text{y} \quad \left(\begin{array}{cc|c} a & b & 0 \\ c & d & 0 \\ \hline * & * & e \end{array}\right).$$

En el caso de la primera matriz basta aplicar la Propiedad 1.2.4 a M^*. Si en la segunda matriz intercambiamos las columnas 1 y 3 y las mismas filas, por 1.2.3 estamos también en las condiciones de 1.2.4.

Veamos qué ocurre para orden 4. Cuando $m = 1, n = 3$ ó $m = 3, n = 1$ se verifica el resultado por la Propiedad 1.2.4. El único caso que queda entonces por estudiar es $n = m = 2$. Podemos suponer que el elemento m_{22} no es nulo, si no se reduciría al caso anterior. Haciendo las transformaciones permitidas para el determinante de Study por filas tenemos:

$$\begin{aligned}
\mathrm{Sdet}\left(\begin{array}{cc|cc} a & b & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ c & d & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$*$}} & p & q \\ & & r & s \end{array}\right) &= \\
\mathrm{Sdet}\left(\begin{array}{cc|cc} a - bd^{-1}c & 0 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ c & d & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$*$}} & p & q \\ & & r & s \end{array}\right) &= \\
|a - bd^{-1}c|\,\mathrm{Sdet}\left(\begin{array}{c|cc} d & \multicolumn{2}{c}{0} \\ \hline \multirow{2}{*}{$*'$} & p & q \\ & r & s \end{array}\right) &= \\
|ad - bd^{-1}cd|\,\mathrm{Sdet}\begin{pmatrix} p & q \\ r & s \end{pmatrix} &= \\
\mathrm{Sdet}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\mathrm{Sdet}\begin{pmatrix} p & q \\ r & s \end{pmatrix}&.
\end{aligned}$$

Análogamente para orden n. Por transformaciones de columnas siempre podemos hacer que los elementos de la primera fila de la matriz M sean todos nulos excepto el m_{11}. Al hacerlo por columnas, estas transformaciones no interfieren en la matriz N y obtenemos así lo que buscábamos. $\square$

Corolario 1.2.6. *Si $T = (t_{ij})$ es una matriz triangular entonces*

$$\mathrm{Sdet}(T) = |t_{11} \cdots t_{nn}|.$$

En particular, por el Lema 1.1.12, si M es una matriz con autovalores por la derecha $q_1 \cdots q_n$, entonces $\mathrm{Sdet}(M) = |q_1 \cdots q_n|$.

1.2.2. Relación con otros determinantes cuaterniónicos

H. Aslaksen resume en [1] diversos determinantes que se han ido proponiendo para las matrices cuaterniónicas.

El primero de ellos fue el de A. Cayley. Se basa en el desarrollo por una fila o columna. Como el resultado depende de la fila o columna que se escoja, fija una de ellas, la primera. Así, su determinante consiste en desarrollar por la primera columna tanto la matriz inicial como todos los menores.

Basándose en las ideas de Cayley, Moore define otro determinante pero solo para las matrices hermíticas. Un poco más adelante (Secc. 1.2.3) lo veremos con detalle pero básicamente se define como sigue.

Consideremos σ una permutación de $I_n = \{1, \dots, n\}$ expresada como producto de ciclos disjuntos de manera que en cada ciclo el elemento menor sea el primero.

$$\sigma = (n_{11} \cdots n_{1l_1})(n_{21} \cdots n_{2l_2}) \cdots (n_{r1} \cdots n_{rl_r}),$$

donde para cada i tenemos que $n_{i1} > n_{ij}$ para todo $j > 1$ y $n_{11} > n_{21} > \cdots > n_{rl}$.

Definición 1.2.7. Dada una matriz hermítica $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ el *determinante de Moore* es

$$\mathrm{Mdet}(M) = \sum_{\sigma \in S_n} |\sigma| m_{n_{11}n_{12}} \cdots m_{n_{1l_1}n_{11}} m_{n_{21}n_{22}} \cdots m_{n_{rl_r}n_{r1}}.$$

Para las matrices reales simétricas y las complejas hermíticas coincide con el determinante usual.

Para cualquier $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ la matriz M^*M es hermítica. Haciendo uso del funcional de Moore se define el *doble determinante,*

$$\mathrm{ddet}(M) = \mathrm{Mdet}(M^*M).$$

Por último, utilizando la forma compleja, se define el *q-determinante*

$$\mathrm{qdet}(M) = \det{}_{\mathbb{C}} c(M).$$

Las diferentes formas de calcular el determinante cuaterniónico guardan mucha relación.

Proposición 1.2.8. *Para $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ una matriz cuaterniónica de orden n, se verifica que:*

1. $\mathrm{qdet}(M) = \mathrm{ddet}(M)$ *y*
2. $\mathrm{Sdet}(M)^2 = \mathrm{Mdet}(MM^*) = \mathrm{ddet}(M)$.

Finalmente, J. Dieudonné propone una definición de determinante para el caso no conmutativo recogida en [10]. Puede probarse que coincide con $\mathrm{Sdet}(M)^4$.

1.2.3. Determinante de una matriz hermítica

Como hemos visto, para una matriz cuaterniónica de orden $n \times n$ no tenemos un determinante que tome valores en $\mathbb{H}$. En el caso particular de una matriz hermítica, los autovalores son todos reales y esto nos permite encontrar un polinomio característico de M.

Inspirándose en el caso complejo, en el ámbito de las matrices cuaterniónicas pueden definirse el desarrollo por una fila o por una columna. La extensión de estos conceptos a $\mathbb{H}$ es reciente (2011) y se debe a I. Kyrchei [51].

Definición 1.2.9. Para una matriz $M = (m_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{H})$ se define el *determinante por la fila i-ésima* como

$$\mathrm{rdet}_i M = \sum_{\sigma} (-1)^{n-r} a_{i i_{k_1}} a_{i_{k_1} i_{k_1+1}} \cdots a_{i_{k_1+l_1} i} \cdots a_{i_{k_r} i_{k_r+1}} \cdots a_{i_{k_r+l_r} i_r}$$

para todo $i = 1, \ldots, n$ donde

$$\sigma = (i\, i_{k_1}\, i_{k_1+1} \cdots i_{k_1+l_1})(i_{k_2}\, i_{k_2+1} \cdots i_{k_2+l_2}) \cdots (i_{k_r}\, i_{k_r+1} \cdots i_{k_r+l_r})$$

es una permutación del conjunto de índices $I_n = \{1, \ldots, n\}$. El índice i abre el primer ciclo y los demás verifican que $i_{k_2} < i_{k_3} < \cdots < i_{k_r}$ y $i_{k_t} < i_{k_{t+s}}$ para todo $t = 2, \ldots, r$ y $s = 1, \ldots, l_t$.

Definición 1.2.10. Para una matriz $M = (m_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{H})$ se define el *determinante por la columna j-ésima* como

$$\mathrm{cdet}_j M = \sum_{\tau} (-1)^{n-r} a_{j_{k_r} j_{k_r+l_r}} \cdots a_{j_{k_r+1} j_{k_r}} \cdots a_{j j_{k_1+l_1}} \cdots a_{j_{k_1+1} j_{k_1}} a_{j_{k_1} j}$$

para todo $j = 1, \ldots, n$ donde

$$\tau = (j_{k_r+l_r} \cdots j_{k_r+1}\, j_{k_r}) \cdots (j_{k_2+l_2} \cdots jk_2+1\, j_{k_2})(j_{k_1+l_1} \cdots j_{k_1+1}\, j_{k_1}\, j)$$

es una permutación del conjunto de índices $I_n = \{1, \ldots, n\}$. El índice j abre el primer ciclo y los demás verifican que $j_{k_2} < j_{k_3} < \cdots < j_{k_r}$ y $j_{k_t} < j_{k_{t+s}}$ para todo $t = 2, \ldots, r$ y $s = 1, \ldots, l_t$.

Estos determinantes verifican propiedades que pueden considerarse análogas a las del determinante complejo. Kyrchei recoge en [51] todas las propiedades que verifican estos determinantes así como algunos lemas que permiten desarrollar $\mathrm{rdet}_i\, M$ por cofactores a lo largo de la fila i-ésima y $\mathrm{cdet}_j\, M$ a lo largo de la columna j-ésima. De todas esas propiedades se deduce el siguiente teorema que permite definir de manera unívoca un determinante para las matrices cuaterniónicas hermíticas que toma valores reales.

Teorema 1.2.11. *Si $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ es tal que $M = M^*$, entonces*

$$\mathrm{rdet}_1 M = \cdots = \mathrm{rdet}_n M = \mathrm{cdet}_1 M = \cdots = \mathrm{cdet}_n M \in \mathbb{R}.$$

Definición 1.2.12. Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ hermítica, se define el *determinante de* M como $\det M = \mathrm{cdet}_i\, M = \mathrm{rdet}_i\, M$ con $i = 1, \ldots, n$.

Este determinante coincide con el determinante de Moore dado en la Sección 1.2.2.

Autovalores de las matrices hermíticas

Es sencillo demostrar la siguiente proposición que recogen varios autores como D. Farenick, I.I. Kyrchei o el propio F. Zhang [19, 51, 92].

Proposición 1.2.13. *Si $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ es hermítica, M tiene exactamente n autovalores por la derecha reales.*

En particular, una matriz hermítica M es normal así que, por los resultados de la Sección 1.1.6, observamos que M se puede diagonalizar a una matriz real.

Kyrchei [51] extiende, en este caso, la definición de polinomio característico real como sigue.

Definición 1.2.14. Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ un matriz cuaterniónica hermítica, llamamos *polinomio característico de* M al polinomio en variable $t \in \mathbb{R}$,

$$p_M(t) = \det(tI - M).$$

Las raíces de este polinomio característico $p_M(t)$ son los autovalores por la derecha de M. Además, se tiene una expresión explícita de los coeficientes [51].

Lema 1.2.15. *Sea $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ hermítica. Su polinomio característico viene dado por*

$$p_M(t) = tn - d_1 t^{n-1} + d_2 t^{n-2} - \cdots + (-1)^n d_n,$$

donde d_r es la suma de los menores principales de M de orden r para $1 \leq r < n$ y $d_r = \det M$.

1.2.4. Quasideterminantes

Los quasideterminantes para matrices con entradas en un anillo de división no conmutativo fueron introducidos por I. Gelfand y V. Retakh en los años 90 y son una herramienta útil en álgebra no conmutativa (ver por ejemplo [27, 89]). Estos autores no se limitan a definir un determinante cuaterniónico sino que desarrollan una teoría mucho más general en la que tratan de dar algún funcional que extienda el determinante usual sobre un anillo de división.

Denotemos por $M^{i,j}$ la submatriz de orden $(n-1)\times(n-1)$ resultante de suprimir en $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ la fila i y la columna j. Entonces, para cada $1 \leq i, j \leq n$ se define inductivamente el quasideterminante $|M|_{ij} \in \mathbb{H}$ como sigue.

Definición 1.2.16. Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$, para cada par (i, j) con $1 \leq i, j \leq n$, se define el (i, j)-*quasideterminante* de M como

$$|M|_{ij} = m_{ij} - \sum m_{iq}(|M^{ij}|_{pq})^{-1}m_{pj},$$

donde la suma se toma sobre los $p, q \in \{1, \ldots, n\}$ tales que $p \neq i, q \neq j$.

Nota–. La definición de quasideterminante que dan I. Gelfand *et al.* en [27] no es del todo correcta. La expresión recogida en este artículo es

$$|M|_{ij} = m_{ij} - \sum m_{ii'}(|M^{ij}|_{j'i'})^{-1}m_{j'j}$$

con $i' \in I - \{i\}, j' \in J - \{j\}$.

La definición correcta es la que habían dado trece años antes, en 1992 [28], que hemos recogido en la Definición 1.2.16.

Ejemplo 1.2.17. Para una matriz de orden 2 hay cuatro qdets. Sea $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$. La expresión para cada uno de sus qdets es:

$$|M|_{11} = m_{11} - m_{12}m_{22}^{-1}m_{21}, \qquad |M|_{12} = m_{12} - m_{11}m_{21}^{-1}m_{22},$$

$$|M|_{21} = m_{21} - m_{22}m_{12}^{-1}m_{11}, \qquad |M|_{22} = m_{22} - m_{21}m_{11}^{-1}m_{12}.$$

Ejemplo 1.2.18. Para una matriz de orden 3 hay nueve qdets. Por ejemplo, el $(2, 1)$-qdet es

$$\begin{aligned} |M|_{21} &= m_{21} - m_{22}(m_{12} - m_{13}m_{33}^{-1}m_{32})^{-1}m_{11} - m_{22}(m_{32} - m_{33}m_{13}^{-1}m_{12})^{-1}m_{31} \\ &- m_{23}(m_{13} - m_{12}m_{32}^{-1}m_{33})^{-1}m_{11} - m_{23}(m_{33} - m_{32}m_{12}^{-1}m_{13})^{-1}m_{31}. \end{aligned}$$

Puesto que esta definición involucra los inversos de los quasideterminantes de orden menor, es necesario dar una definición para cuando alguno de ellos se anula. La siguiente Proposición lo aclara (ver Prop. 1.5 de [28]).

Proposición 1.2.19. *Dada $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ el quasideterminante $|M|_{pq}$ está definido si al menos uno de los qdets de la matriz M^{pq} está definido y no se anula. En este caso,*

$$|M|_{pq} = m_{pq} - \sum m_{pj}|M^{p,q}|_{ij}^{-1} m_{iq},$$

donde la suma la tomamos sobre todos los pares $(i,j), i \neq p, j \neq q$, para los que el determinante $|M^{p,q}|_{ij}$ está definido y no se anula.

Este modo de hacer los quasideterminantes nos proporciona la inversa de la matriz. De hecho, cuando M es inversible, las entradas de la matriz inversa m^{ij} son precisamente los inversos de los qdets traspuestos, $|M|_{ji}^{-1}$. En particular, cuando los elementos de la matriz M conmutan, tenemos que

$$|M|_{ji}^{-1} = (-1)^{i+j} \det M^{j,i} / \det M.$$

Veamos un ejemplo:

Ejemplo 1.2.20. $M = \begin{pmatrix} \mathbf{i} & 0 & 0 \\ \mathbf{k} & \mathbf{j} & 0 \\ -3 & 2\mathbf{k} & \mathbf{k} \end{pmatrix}$

Al calcular los quasideterminantes obtenemos:

$$\begin{array}{ll} |M|_{11} = \mathbf{i} & |M|_{22} = \mathbf{j} \\ |M|_{33} = \mathbf{k} & |M|_{12} = 1 \\ |M|_{13} = -\mathbf{i}(-3\mathbf{i} + 2\mathbf{j})^{-1}\mathbf{k} & |M|_{23} = -\frac{1}{2}\mathbf{j}. \end{array}$$

En cambio, $|M|_{21}, |M|_{31}$ y $|M|_{32}$ no existen ya que en estos casos los $|M^{pq}|_{ij}$ o no están definidos o se anulan.

La matriz obtenida con los inversos de los qdets es, efectivamente, la matriz inversa de M:

$$M^{-1} = \begin{pmatrix} -\mathbf{i} & 0 & 0 \\ 1 & -\mathbf{j} & 0 \\ -2-3\mathbf{k} & 2\mathbf{j} & -\mathbf{k} \end{pmatrix}.$$

La siguiente proposición generaliza el caso complejo y establece la relación entre los quasideterminantes y el determinante de Study.

Proposición 1.2.21. *Para toda matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ se verifica que, siempre que exista $|M|_{pq}$,*

$$\mathrm{Sdet}(M^{p,q})\,||M|_{pq}| = \mathrm{Sdet}(M)$$

donde M^{pq} es la matriz resultante de suprimir en M la fila p y la columna q y $||M|_{pq}|$ es el módulo del (p,q)-qdet.

Esta Proposición es un caso particular de la llamada identidad de Jacobi que comprobaremos a continuación.

1.2.5. Identidad de Jacobi

La llamada identidad de Jacobi del determinante (a pesar de que en 1882 se atribuyó a Kronecker [64]) es una generalización de la fórmula

$$(M^{-1})_{ij} = (-1)^{i+j}\det M^{j,i}/\det M. \tag{1.2}$$

Sea $C \in \mathcal{M}_{m\times m}(\mathbb{C})$ una matriz compleja cuadrada de orden m. Sean $I = \{i_1,\dots,i_p\}$ y $J = \{j_1,\dots,j_p\}$ subconjuntos de $\{1,\dots,m\}$ del mismo tamaño p. Denotamos por $C_{I,J}$ la submatriz formada por las filas con índice en I y las columnas con índice en J. Recíprocamente, denotamos por $C^{I,J}$ la submatriz complementaria obtenida suprimiendo las filas en I y las columnas en J.

Lema 1.2.22. *Supongamos una matriz compleja C inversible. Entonces*

$$\det(C^{-1})_{I,J} = (-1)^{I+J}\det C^{J,I}/\det C,$$

donde $I+J$ significa $i_1+\cdots+i_p+j_1+\cdots+j_p$.

Utilizamos el determinante de Study para establecer un teorema análogo en el caso cuaterniónico.

Teorema 1.2.23. *Supongamos una matriz cuaterniónica M inversible. Entonces*

$$\mathrm{Sdet}\,(M^{-1})_{I,J} = \mathrm{Sdet}\,M^{J,I}/\,\mathrm{Sdet}\,M,$$

para cualesquiera subconjuntos de índices I y J del mismo tamaño.

Demostración. Si $I = \{i_1,\dots,i_p\}$ denotamos

$$I' = I + n = \{i_1+n,\dots,i_p+n\};$$

análogamente para $J' = J + n$. Como

$$c\left((M^{-1})_{I,J}\right) = c(M^{-1})_{I\cup I',J\cup J'} = \left(c(M)^{-1}\right)_{I\cup I',J\cup J'}$$

y

$$c(M^{J,I}) = c(M)^{J\cup J',I\cup I'},$$

el resultado se sigue del caso complejo, Lema 1.2.22. □

Con todo esto encontramos el modo de generalizar la fórmula (1.2) a los cuaternios; basta tomar $p = 1, I = \{i\}$ y $J = \{j\}$.

Corolario 1.2.24. *Sea $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ una matriz cuaterniónica inversible. Sea m^{ij} la entrada (i,j) de la matriz inversa M^{-1}. Entonces su norma viene dada por*

$$|m^{ij}| = \operatorname{Sdet}(M^{j,i})/\operatorname{Sdet}(M).$$

Ejemplo 1.2.25. Sea

$$M = \begin{pmatrix} 2\mathbf{i}-\mathbf{j} & 3\mathbf{j} & \mathbf{i}-3\mathbf{k} \\ \mathbf{i}+2\mathbf{k} & \mathbf{i}-\mathbf{k} & 2\mathbf{i} \\ -\mathbf{j} & 3\mathbf{k} & \mathbf{j} \end{pmatrix}.$$

Calculando los quasideterminantes obtenemos que su matriz inversa es

$$M^{-1} = \frac{1}{77}\begin{pmatrix} -27+12\mathbf{i}+17\mathbf{j}+42\mathbf{k} & 18+36\mathbf{i}+18\mathbf{j}-39\mathbf{k} & -28\mathbf{i}+42\mathbf{j}-42\mathbf{k} \\ -16-24\mathbf{i}-13\mathbf{j} & -15+5\mathbf{i}+27\mathbf{j}+22\mathbf{k} & 28-7\mathbf{i}+189\mathbf{j}-7\mathbf{k} \\ 45-36\mathbf{i}+17\mathbf{j}+3\mathbf{k} & 3-9\mathbf{i}-48\mathbf{j}+42\mathbf{k} & 21+56\mathbf{i}-14\mathbf{j} \end{pmatrix}.$$

Por ejemplo, el quasideterminante $(3,1)$ de M es

$$|M|_{31} = \frac{1}{2}\mathbf{i} - \frac{3}{4}\mathbf{j} + \frac{3}{4}\mathbf{k},$$

y la norma de su inverso es

$$\||M|_{31}^{-1}| = |-\frac{4\mathbf{i}}{11} + \frac{6\mathbf{j}}{11} - \frac{6\mathbf{k}}{11}| = \sqrt{\frac{8}{11}}$$

mientras que

$$\operatorname{Sdet} M^{1,3} = 2|\mathbf{i}+3\mathbf{j}+2\mathbf{k}| = 2\sqrt{14} = \sqrt{56}$$

y

$$\operatorname{Sdet} M = \sqrt{77}.$$

Tomemos ahora, por ejemplo, $I = \{1, 2\}$ y $J = \{1, 3\}$. Entonces $M^{J,I} = (2\mathbf{i})$ y $\operatorname{Sdet} A^{J,I} = 2$, mientras que

$$(M^{-1})_{IJ} = \frac{1}{77}\begin{pmatrix} -27 + 12\mathbf{i} + 17\mathbf{j} + 42\mathbf{k} & -28\mathbf{i} + 42\mathbf{j} - 42\mathbf{k} \\ -16 - 24\mathbf{i} - 13\mathbf{j} & 28 - 7\mathbf{i} + 189\mathbf{j} - 7\mathbf{k} \end{pmatrix}$$

y

$$\operatorname{Sdet}(M^{-1})_{IJ} = \frac{2}{\sqrt{77}}.$$

1.3. Autovalores por la izquierda

La teoría de autovalores por la izquierda de matrices cuaterniónicas está muy poco desarrollada. Un resultado de R. M .W. Wood [90] garantiza su existencia pero, en general, no se sabe cuántos autovalores por la izquierda puede tener una matriz cuaterniónica de orden $n \times n$. L. Huang y W. So probaron que una matriz cuadrada de orden 2 puede tener uno, dos o infinitos autovalores (pertenecientes a diferentes clases de similitud) y caracterizaron este último caso [33].

F. Zhang recoge en [92, 93] sus principales propiedades y algunos ejemplos patológicos; puede verse también [32]. El propio Zhang plantea como cuestión abierta cuántos autovalores por la izquierda puede tener una matriz cuaterniónica cuadrada y sugiere investigar su espectro por la izquierda.

1.3.1. Autovalores de una matriz

Definición 1.3.1. Sea M una matriz cuaterniónica de orden n. Se dice que $\lambda \in \mathbb{H}$ es un *autovalor por la izquierda* de M si existe un $v \in \mathbb{H}^n, v \neq 0$, tal que

$$Mv = \lambda v.$$

Llamamos $\sigma_l(M)$ al espectro por la izquierda de la matriz M.

El interés de esta definición radica en que es equivalente al hecho de que la matriz $M - \lambda I$ sea singular, es decir, para esto no sirven los autovalores por la derecha (en el caso conmutativo no se da esta sutileza). De acuerdo con las propiedades del determinante de Study tenemos

Proposición 1.3.2. *Los autovalores por la izquierda de* $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ *son las raíces de la ecuación* $\operatorname{Sdet}(M - \lambda I) = 0$.

Proposición 1.3.3. *El conjunto de los autovectores asociados a un autovalor por la izquierda λ forma un subespacio vectorial (por la derecha).*

Demostración. Sean λ un autovalor de la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$, v, v' dos λ-autovectores y $q \in \mathbb{H}$; se verifica entonces que

$$M(vq) = (Mv)q = (\lambda v)q = \lambda(vq),$$
$$M(v + v') = Mv + Mv' = \lambda(v + v'). \quad \square$$

Proposición 1.3.4. *El espectro por la izquierda de una matriz cuaterniónica $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ es compacto.*

Demostración. De acuerdo con la Proposición 1.3.2, $\sigma_l(M)$ es cerrado por ser la imagen recíproca de un cerrado por una aplicación continua ya que la aplicación forma compleja, el determinante complejo y la raíz cuadrada del valor absoluto son aplicaciones continuas.

Además, $\sigma_l(M)$ es acotado. Sea $\lambda \in \sigma_l(M)$, entonces

$$|\lambda| = \frac{|\lambda v|}{|v|} \leq \sup_{|w|=1} \frac{|Mw|}{|w|} = \|M\|. \quad \square$$

Recogemos aquí el teorema de existencia de Wood [90].

Teorema 1.3.5 (Wood, 1985)**.** *Toda matriz cuaterniónica de orden $n \times n$ tiene al menos un autovalor por la izquierda.*

Demostración. La prueba es topológica. En primer lugar, si $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ es singular, $0 \in \sigma_l(M)$. Así que tomamos M inversible. Suponemos que $M - \lambda I \in \mathrm{GL}(n, \mathbb{H}), \forall \lambda \neq 0$ y llegamos a una contradicción.

En efecto, si $M - \lambda I$ es inversible para todo $\lambda \in \mathbb{H}$, podemos construir dos homotopías en $\mathrm{GL}(n, \mathbb{H})$:

$$\begin{aligned} f_t(\lambda) &= M - t\lambda I, \\ g_t(\lambda) &= tM - \lambda I. \end{aligned}$$

Como $f_0(\lambda) = M$, $f_1(\lambda) = M - \lambda I = g_1(\lambda)$ y $g_0(\lambda) = -\lambda I$, las aplicaciones f_0 y g_0 son homótopas. Sin embargo, si las consideramos como aplicaciones de la esfera S^3 en $GL(n, \mathbb{H})$, en el tercer grupo de homotopía $\pi_3 \,\mathrm{GL}(n, \mathbb{H}) \cong \mathbb{Z}$ corresponden a los enteros 0 y n [80], luego no pueden ser homótopas. $\square$

Matrices 2×2

En cuanto al número de autovalores solo tenemos el siguiente Teorema de Huang y So para matrices 2×2 probado en 2001 [33].

Teorema 1.3.6. *Una matriz* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{H})$ *tiene uno, dos o infinitos autovalores por la izquierda. Este último caso se da si y sólo si* $a_0, a_1 \in \mathbb{R}$, $a_0 \neq 0$ *y* $\Delta = a_1^2 - 4a_0 < 0$, *donde* $a_0 = -b^{-1}c$ *y* $a_1 = b^{-1}(a-d)$.

La demostración de este teorema se basa en la resolución de polinomios de grado dos que hicieron los mismos autores en [34] y que discutiremos más adelante.

Para estas matrices, podemos calcular explícitamente el espectro por la izquierda. Si la matriz M es triangular (es decir, $bc = 0$), los autovalores de M son los elementos de la diagonal. Para una matriz M no triangular Huang y So [33] obtuvieron el siguiente resultado que nosotros probamos de manera diferente.

Proposición 1.3.7. *Si* $bc \neq 0$, *los autovalores por la izquierda de* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *son de la forma* $\lambda = a + bp$, *donde* p *es cualquier solución del polinomio cuadrático unilateral*

$$p^2 + a_1 p + a_0 = 0, \tag{1.3}$$

con $a_1 = b^{-1}(a-d)$ *y* $a_0 = -b^{-1}c$.

Demostración. Los autovalores de M vienen dados por $\mathrm{Sdet}(M - \lambda I) = 0$. En este caso, si λ es un autovalor por la izquierda de M, $\lambda \neq a, d$. Entonces, utilizando las propiedades de Sdet podemos transformar la matriz $M - \lambda I$ de manera que

$$\mathrm{Sdet}(M - \lambda I) = \mathrm{Sdet}\begin{pmatrix} a-\lambda & b \\ 0 & (d-\lambda) - c(a-\lambda)^{-1}b \end{pmatrix}.$$

Esta matriz será no inversible si y sólo si

$$(d-\lambda) - c(a-\lambda)^{-1}b = 0$$

es decir,

$$(d-\lambda)b^{-1}(a-\lambda) = c.$$

Si ahora hacemos el cambio $p = b^{-1}(\lambda - a)$ nos da el polinomio cuadrático unilateral que buscábamos

$$-bp^2 + (d-a)p - c = 0. \quad \square$$

Nota–. Llamaremos *esférico* al caso en el que hay infinitos autovalores porque, en ese caso, el espectro

$$\sigma_l(A) = \{(1/2)(a + d + bq) \colon q^2 = \Delta\} \tag{1.4}$$

es difeomorfo a la esfera $S^2 \subset \mathbb{H}_0 = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$.

Matrices 3×3

Para orden 3 W. So hizo un estudio caso por caso según las relaciones entre las entradas de la matriz [79]. Obtuvo diferentes polinomios cuaterniónicos de grado menor o igual que 3 tales que sus raíces son los autovalores por la izquierda.

1.4. Polinomios cuaterniónicos

No es sencillo hallar las raíces de un polinomio cuaterniónico. Uno de los primeros problemas que nos encontramos es que un polinomio bilateral puede tener varios términos del mismo grado. Además, ya no se verifica que un polinomio tenga tantas soluciones como indica su grado; por ejemplo, la ecuación $x^2 = -1$, que en $\mathbb{R}$ no tiene solución y en $\mathbb{C}$ tiene dos, tiene infinitas soluciones en el caso cuaterniónico, toda una esfera S^2. Esto nos lleva a plantearnos tanto si todos los polinomios cuaterniónicos tienen raíces como, en el caso de tenerlas, cuántas hay.

1.4.1. Teorema fundamental

En 1941 I. Niven prueba [65], basándose en el algoritmo de la división, que toda ecuación del tipo

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0$$

tiene alguna solución. En ese mismo artículo proporciona un método (poco práctico) para obtener las raíces de polinomios de este tipo, discute el número de soluciones y da una condición necesaria y suficiente para que tenga infinitas soluciones.

Tres años más tarde, S. Eilenberg y el propio Niven [14] extendieron el teorema fundamental del álgebra a los cuaternios para polinomios bilaterales pero en el caso particular de que el polinomio solo tenga un término de mayor grado.

Teorema 1.4.1 (Eilenberg y Niven, 1944)**.** *Sea $f(x) = a_0 x a_1 x \dots x a_n + \phi(x)$ donde $a_0, a_1, \dots, a_n \in \mathbb{H}$ no nulos y $\phi(x)$ es la suma de un número finito de monomios $b_0 x b_1 x \dots x b_k$ con $k < n$. Entonces la ecuación $f(x) = 0$ tiene al menos una solución.*

En este caso la demostración es topológica. Extienden f a toda la esfera S^4 haciendo $f(\infty) = \infty$. Esta extensión es continua precisamente porque solo tenemos un monomio de mayor grado. Construyen una homotopía entre $f(x)$ y el monomio $g(x) = x^n$ en S^4. Por último, verifican que la función g es de grado n comprobando que $\mathbf{i}$ es un valor regular (ya que $x^n = \mathbf{i}$ tiene exactamente n soluciones).

Gordon y Motzkin probaron en 1965 que un polinomio mónico estándar (unilateral) de grado n, o tiene infinitas raíces o tiene como mucho n raíces distintas.

El teorema fundamental no puede extenderse a todos los polinomios cuaterniónicos ya que, por ejemplo, el polinomio $x^n a - ax^n - 1$ no tiene ninguna raíz en $\mathbb{H}$.

1.4.2. Raíces de un polinomio cuadrático

L. Huang y W. So [34] proporcionan fórmulas explícitas para hallar las raíces de un polinomio mónico unilateral de grado 2 del tipo

$$x^2 + bx + c.$$

Esto es de utilidad para encontrar autovalores por la izquierda (*cf.* Prop. 1.3.7). El siguiente teorema recoge dichas fórmulas.

Teorema 1.4.2. *Las soluciones de la ecuación cuadrática $x^2 + bx + c = 0$ se pueden obtener mediante las siguientes fórmulas.*

1. *Si $b, c \in \mathbb{R}$ y $b^2 < 4c$ entonces*

$$x = \frac{1}{2}(-b + \beta\mathbf{i} + \gamma\mathbf{j} + \delta\mathbf{k})$$

para todos $\beta, \gamma, \delta \in \mathbb{R}$ tales que $\beta^2 + \gamma^2 + \delta^2 = 4c - b^2$;

2. *si $b, c \in \mathbb{R}$ y $b^2 \geq 4c$ entonces*

$$x = -\frac{1}{2}(-b \pm \sqrt{b^2 - 4c});$$

3. *si $b \in \mathbb{R}$ pero $c \notin \mathbb{R}$ entonces*

$$x = -\frac{b}{2} \pm \frac{\rho}{2} \mp \frac{c_1}{\rho}\mathbf{i} \mp \frac{c_2}{\rho}\mathbf{j} \mp \frac{c_3}{\rho}\mathbf{k},$$

donde $c = c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$ y

$$\rho = \frac{1}{\sqrt{2}}\sqrt{b^2 - 4c_0 + \sqrt{(b^2 - 4c_0)^2 + 16(c_1^2 + c_2^2 + c_3^2)}};$$

4. *por último, si* $b \notin \mathbb{R}$,

$$x = -(\Re(b)/2) - (b' + T)^{-1}(c' - N)$$

donde

$$b' = \Im b \quad y \quad c' = c - (\Re(b)/2)(b - \Re(b)/2)$$

y el par (T, N) *se elige como sigue:*

sean $B = |b'|^2 + 2\Re(c'), E = |c'|^2$ *y* $D = 2\Re(\overline{b'})c'$,

a) $T = 0$ *y* $N = 1/2(B \pm \sqrt{B^2 - 4E})$ *cuando* $D = 0$ *y* $B^2 \geq 4E$.

b) $T = \pm\sqrt{-B + 2\sqrt{E}}$ *y* $N = \sqrt{E}$ *cuando* $D = 0$ *y* $B^2 < 4E$.

c) $T = \pm\sqrt{z}$ *y* $N = \frac{1}{2T}(T^3 + BT + D)$ *cuando* $D \neq 0$ *donde* z *es la única raíz positiva del polinomio real* $z^3 + 2Bz^2 + (B^2 - 4E)z - D^2$.

1.5. Ecuaciones lineales cuaterniónicas

En el mismo volumen del *Bull. Amer. Math. Soc.* de 1944 en que está recogido el artículo de Eilenberg y Niven, se publicó un artículo de R.E. Johnson [48] en el que explica cómo resolver la ecuación $x\alpha = \gamma x + \beta$ sobre un anillo de división. Este autor obtiene condiciones necesarias y suficientes para que una ecuación de este tipo tenga solución y en ese caso da un método sencillo para calcular explícitamente al menos una de las soluciones.

A continuación veremos con un poco más de detalle cómo se resuelve en el caso cuaterniónico ésta ecuación, conocida hoy en día como ecuación de Sylvester, ecuación que usaremos más adelante.

1.5.1. La ecuación de Sylvester $\alpha x + x\beta = \gamma$

Una ecuación del tipo

$$\alpha x + x\beta = \gamma \tag{1.5}$$

tal que la variable y los coeficientes están sobre un anillo no conmutativo, recibe el nombre de ecuación de Sylvester.

El método de Johnson para resolverla puede resumirse como sigue. Dada la ecuación (1.5) multiplicamos toda la ecuación a la izquierda por $\overline{\alpha}$ y a la derecha por $\overline{\beta}$,

$$|\alpha|^2 x\overline{\beta} + \overline{\alpha} x|\beta|^2 = \overline{\alpha}\gamma\overline{\beta}.$$

Volvemos a la ecuación inicial y la multiplicamos por el módulo de α al cuadrado (podría hacerse análogamente con el de β):

$$|\alpha|^2\alpha x + |\alpha|^2 x\beta = |\alpha|^2\gamma.$$

Sumando ambas expresiones obtenemos

$$|\alpha|^2 x\,(2\Re(\beta)) + \left(|\beta|^2\overline{\alpha} + |\alpha|^2\alpha\right)x = \overline{\alpha}\gamma\overline{\beta} + |\alpha|^2\gamma,$$

es decir,

$$\left(2|\alpha|^2\Re(\beta) + |\beta|^2\overline{\alpha} + |\alpha|^2\alpha\right)x = \overline{\alpha}\gamma\overline{\beta} + |\alpha|^2\gamma$$

que se resuelve sencillamente.

1.5.2. Método de Janovská y Opfer

El método anterior permite calcular explícitamente las soluciones de la ecuación (1.5). Para estudiar la existencia y el número de soluciones D. Janovská y G. Opfer [46] la identifican con una matriz real 4×4, puede verse también [29]. Estos artículos son de 2008 y 2009 respectivamente. Hemos adaptado sus resultados.

Dado un polinomio cuaterniónico del tipo $\alpha x + x\beta$ con $\alpha, \beta \in \mathbb{H}$, podemos asociarle un aplicación $\mathbb{R}$-lineal de $\mathbb{R}^4$ en $\mathbb{R}^4$. Basta considerar los cuaternios como vectores de $\mathbb{R}^4$,

$$\begin{array}{rcl}\alpha & = & (\alpha_1, \alpha_2, \alpha_3, \alpha_4)^T \\ \beta & = & (\beta_1, \beta_2, \beta_3, \beta_4)^T.\end{array}$$

Si desarrollamos el producto $\alpha x + x\beta$ con α y β expresados de este modo tenemos que la matriz asociada es

$$M = L_\alpha + R_\beta = \begin{pmatrix} \alpha_1+\beta_1 & -\alpha_2-\beta_2 & -\alpha_3-\beta_3 & -\alpha_4-\beta_4 \\ \alpha_2+\beta_2 & \alpha_1+\beta_1 & -\alpha_4+\beta_4 & \alpha_3-\beta_3 \\ \alpha_3+\beta_3 & \alpha_4-\beta_4 & \alpha_1+\beta_1 & -\alpha_2+\beta_2 \\ \alpha_4+\beta_4 & -\alpha_3+\beta_3 & \alpha_2-\beta_2 & \alpha_1+\beta_1 \end{pmatrix}$$

donde denotamos por L_α la matriz asociada a la traslación izquierda por el cuaternio α y R_β la asociada a la traslación derecha por β.

Proposición 1.5.1. *Dados $\alpha, \beta \in \mathbb{H}$ la matriz $M = L_\alpha + R_\beta$ verifica que:*

1. $M + M^T = 2\Re(\alpha+\beta)I$*;*

2. $\det(M) = 2\Re(\alpha+\beta)^2 \left(|\Im(\alpha)|^2 + |\Im(\beta)|^2\right) + \Re(\alpha+\beta)^4 + \left(|\Im(\alpha)|^2 - |\Im(\beta)|^2\right)^2$;

3. *los cuatro autovalores de M son*

$$\lambda = \Re(\alpha+\beta) \pm \left(|\Im(\alpha)| \pm |\Im(\beta)|\right).$$

Proposición 1.5.2. *Consideremos la ecuación lineal cuaterniónica $\alpha x + x\beta = \gamma$. Sea $M = L_\alpha + R_\beta$ su matriz asociada. Se verifica entonces que:*

1. *el rango de M es par;*

2. $\operatorname{rang}(M) < 4$ *sii* $|\alpha| = |\beta|$ *y* $\Re(\alpha+\beta) = 0$, *es decir, sii α y $-\beta$ son similares;*

3. $\operatorname{rang}(M) = 0$ *sii α es real y $\beta = -\alpha$.*

Demostración. Por la proposición anterior tenemos que los autovalores de M son conjugados dos a dos, luego su rango siempre es par.

De la expresión del determinante dada en la Proposición 1.5.1 se deducen las otras dos afirmaciones. □

1.5.3. Resolución de una ecuación lineal arbitraria

Aunque $\mathbb{H}$ es un anillo no conmutativo la asociatividad nos permite reducir una ecuación lineal del tipo

$$\sum_i p_i x q_i = r, \qquad \text{con } p_i, q_i, r \in \mathbb{H}, \tag{1.6}$$

a un sencillo sistema de ecuaciones escalares.

Proposición 1.5.3. *Sean $p = t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ y $q = s + u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ dos cuaternios. Entonces,*

1. *la matriz asociada a la traslación izquierda por p es $L_p = \Re(p)I + A(p)$ donde*

$$A(p) = \begin{pmatrix} 0 & -x & -y & -z \\ x & 0 & -z & y \\ y & z & 0 & -x \\ z & -y & x & 0 \end{pmatrix};$$

2. *análogamente,* $R_q = \Re(q)I + B(q)$ *es la matriz asociada a la traslación derecha por* q*, donde*
$$B(q) = \begin{pmatrix} 0 & -u & -v & -w \\ u & 0 & w & -v \\ v & -w & 0 & u \\ w & v & -u & 0 \end{pmatrix}.$$

En cuanto a los términos bilaterales, como $p(xq) = (px)q$, las matrices L_p y R_q conmutan. Por tanto, $\Re(p)I + A(p)$ y $\Re(q)I + B(q)$ conmutan, luego $A(p)$ y $B(q)$ conmutan. Así, resolver la ecuación (1.6) es equivalente a resolver el sistema matricial
$$M\hat{x} = \hat{c}$$
donde $\hat{x}$ denota el cuaternio x visto como vector de $\mathbb{R}^4$ y M es la matriz real 4×4
$$M = \sum_i L_{p_i} R_{q_i} = \sum a_i b_i I + \sum_i (a_i B_i + b_i A_i) + \sum_i A_i B_i$$
siendo $a_i = \Re(p_i)$, $b_i = \Re(q_i)$ y $A_i = A(p_i)$, $B_i = B(q_i)$.

En general, los coeficientes de la matriz asociada M ya no guardan ningún tipo de simetría, de modo que el rango ya no tiene por qué ser par.

Ejemplo 1.5.4. La matriz real asociada a la ecuación lineal bilateral
$$\mathbf{k}x + x(2 - \mathbf{i}) - 2\mathbf{j}x\mathbf{j} = 0$$
es
$$M = \begin{pmatrix} 4 & 1 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 0 & 1 & 4 & -1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$
que es de rango 3.

Capítulo 2

Espacios Simétricos

A grandes rasgos, un espacio simétrico es una variedad diferenciable cuyo grupo de simetrías tiene una inversión de simetría para cada punto. Estos espacios podemos estudiarlos o bien a partir de la geometría de Riemann o bien con la teoría de Lie; en este caso, la definición es más general y más algebraica. Los puntos fundamentales de la teoría de espacios simétricos pueden verse en el libro de S. Helgason [31] o en el de S. Kobayashi y K. Nomizu [52].

En geometría de Riemann, las inversiones son simetrías geodésicas que se les pide que sean isometrías; tenemos así un espacio simétrico Riemanniano.

En general, un espacio simétrico es un G-espacio homogéneo G/H para algún grupo de Lie G, tal que el estabilizador de un punto, H, es un subgrupo abierto del conjunto de puntos fijos de una involución de G. Esta definición incluye tanto los espacios globalmente simétricos Riemannianos como los pseudo-Riemannianos y los dotados de una conexión afín.

Recordaremos qué es un espacio simétrico, tanto desde el punto de vista geométrico como desde el algebraico y explicaremos en qué consiste el modelo de Cartan.

2.1. Un punto de vista geométrico

La noción de conexión se define como sigue.

Definición 2.1.1. Sea M una variedad y $\chi(M)$ el álgebra de sus campos de vectores. Una *conexión afín* es una aplicación

$$\nabla : \chi(M) \times \chi(M) \to \chi(M)$$

verificando que

1. $\nabla_{fX+hY}Z = f\nabla_X Z + h\nabla_Y Z$;
2. $\nabla_X(fY) = f\nabla_X Y + (Xf)Y$.

Ejemplo 2.1.2. Para las superficies embebidas en $\mathbb{R}^3$, la métrica es la primera forma fundamental y la conexión ∇ corresponde a los símbolos de Christoffel. Está dada localmente como

$$\nabla_{\partial/\partial xi}\frac{\partial}{\partial x^j} = \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x^k}.$$

Se dice que $\nabla_X Y$ es la *derivada covariante de Y en la dirección de X.*

Definición 2.1.3. La *torsión* de una conexión afín se define como

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$

y la *curvatura* como

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

Definición 2.1.4. Dada una variedad M dotada de una conexión afín, un difeomorfismo ϕ de M es una *transformación afín* si se cumple que

$$\nabla_{\phi_* X}(\phi_* Y) = \phi_*(\nabla_X Y).$$

Es la generalización de la noción de isometría.

Definición 2.1.5. Dada una variedad M con una conexión afín, una curva γ es una *geodésica* si el campo tangente γ' es paralelo a lo largo de γ.

Resolviendo la ecuación $\nabla_{\gamma'}\gamma' = 0$ vemos que dado cualquier $v \in T_pM$ no nulo con norma suficientemente pequeña existe una única geodésica maximal γ_v con $\gamma_v(0) = p$ y $\gamma_v'(0) = v$, *i.e.*, una curva en M que arranca en p con velocidad v.

Consideremos una bola $U_0 = B(0,r) \subset T_pM$ y la llamada aplicación exponencial $\exp_p : v \in T_pM \mapsto \gamma_v(1) \in M$. Para un radio r suficientemente pequeño es un difeomorfismo con su imagen N_p; se dice entonces que tenemos un *entorno normal de p*. De hecho, puede conseguirse que N_p sea entorno normal de todos sus puntos.

Definición 2.1.6. Sea $N_p = \exp(U_0)$ un entorno normal. Dado $q \in N_p$ existe una única geodésica γ tal que $\gamma(0) = p$ y $\gamma(1) = q$. Consideremos el punto simétrico $q' = \gamma(-1)$, la aplicación s_p que lleva q en q' se llama *simetría geodésica respecto de p.*

Claramente su diferencial es $(s_p)_* = -I$ en T_pM.

Definición 2.1.7. La variedad M se dice *localmente simétrica* si todas las simetrías geodésicas $s_p : N_p \to N_p$ son transformaciones afines.

Teorema 2.1.8. *Un espacio (M,∇) es localmente simétrico si y solo si la conexión es sin torsión y la curvatura es paralela.*

2.2. Espacios globalmente simétricos

En un espacio localmente simétrico (M, ∇) las simetrías geodésicas locales son transformaciones afines. Cuando podamos extenderlas a toda la variedad hablaremos de un espacio globalmente simétrico.

Definición 2.2.1. Un espacio (M, ∇), donde M es una variedad y ∇ una conexión afín, es un espacio *globalmente simétrico* si para cada $p \in M$ existe una transformación afín global $s_p : M \to M$ involutiva ($s_p^2 = \mathrm{id}$ pero $s_p \neq \mathrm{id}$) y para la que p es un punto fijo aislado.

Es decir, la simetría global conserva a p y "levanta" las geodésicas pasando por p; si γ es una geodésica tal que $\gamma(0) = p$, entonces $s_p(\gamma(t)) = \gamma(-t)$.

Proposición 2.2.2. *Para todo $p \in M$, M globalmente simétrico, la simetría s_p es la simetría geodésica en un entorno de p.*

Nos referiremos a partir de ahora solo a espacios simétricos Riemannianos, aunque algunos de los resultados son válidos para el caso general de una conexión afín.

El siguiente teorema y su corolario clarifican la relación entre local y global.

Teorema 2.2.3 (Lichnerowicz)**.** *Un espacio localmente simétrico Riemanniano completo y simplemente conexo es globalmente simétrico.*

Corolario 2.2.4. *La cubierta universal de un espacio localmente simétrico Riemanniano completo es globalmente simétrica.*

2.2.1. Pares simétricos

Cuando en un espacio localmente simétrico podemos extender las simetrías geodésicas locales a toda la variedad tendremos un espacio globalmente simétrico. A cada uno de estos espacios le asociamos un *par simétrico* (G, K, σ). Veamos como se construye este par para una variedad de Riemann (M, g) globalmente simétrica.

Definición 2.2.5. Dado un grupo de Lie conexo G, un *par simétrico* es (G, K, σ) donde σ es un automorfismo involutivo de G, $\sigma \neq \mathrm{id}$, y $K \subset G$ es un subgrupo cerrado intermedio entre el grupo G^σ de puntos fijos y su componente conexa de la identidad, $(G^\sigma)_e \subset K \subset G^\sigma$.

La condición sobre K nos dice que es unión de componentes conexas de G^σ.

Sea pues (M, g) una variedad de Riemann globalmente simétrica. El par simétrico asociado se construye de la siguiente forma.

a) Consideramos el grupo de isometrías $\mathrm{Isom}(M, g)$. Es un grupo de Lie con la topología compacto-abierto.

b) Como automorfismo σ tomamos $\sigma : G \to G$ dado por

$$\sigma(\varphi) = s_p \circ \varphi \circ s_p.$$

c) Para escoger el subgrupo K, fijamos un punto $p \in M$ y tomamos su isotropía $G_p = K$, es decir, las isometrías que dejan fijo al punto p. Así, de acuerdo con el teorema fundamental, $M \cong G/K$.

Sea un grupo de Lie G compacto y conexo y σ un automorfismo involutivo. Entonces, si denotamos por $K = G^\sigma$, el espacio homogéneo G/K con la métrica Riemanniana generada por la métrica biinvariante de G es un espacio simétrico.

2.2.2. Descomposición

Como $\sigma^2 = id$, sus autovalores son ± 1. Esto nos permite descomponer el álgebra de Lie $\mathfrak{g}$ en vectores horizontales y verticales, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$,

$$\begin{aligned} \mathfrak{h} &= \{X \in \mathfrak{g} : \sigma_*(X) = X\}, \\ \mathfrak{m} &= \{X \in \mathfrak{g} : \sigma_*(X) = -X\}. \end{aligned}$$

Considerando $M = G/K$, se tiene que, como espacios vectoriales, $T_oM \cong \mathfrak{m}$, donde denotamos por $o = [e]$. Además, utilizando que σ_* es un morfismo de álgebras de Lie, tenemos que $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$, es decir, $\mathfrak{h}$ es subálgebra de Lie; $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}$ y $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{h}$.

2.2.3. Ejemplos

Ejemplo 2.2.6 (Grupos de Lie compactos)**.** Todo grupo de Lie compacto y conexo H es un espacio globalmente simétrico Riemanniano. Consideremos el siguiente par simétrico (G, K, σ):

(a) Como grupo de Lie tomamos el producto directo $G = H \times H$.

(b) Como automorfismo

$$\begin{aligned} \sigma : \quad H \times H &\to H \times H \\ (h, h') &\mapsto (h', h) \end{aligned}$$

(c) El conjunto de puntos fijos $K = G^\sigma$ es el subgrupo diagonal isomorfo a H, imagen de

$$h \in H \mapsto (h,h) \in H \times H$$

que se denota H^*. Tenemos el difeomorfismo

$$\begin{array}{rcl} (H \times H)/H^* & \to & H \\ [(h,h')] & \mapsto & h(h')^{-1} \end{array}$$

correspondiente a la acción transitiva de $H \times H$ sobre H dada por

$$\begin{array}{rcl} (H \times H) \times H & \to & H \\ ((h,h'),h'') & \mapsto & hh''(h')^{-1} \end{array}$$

Las geodésicas son los subgrupos uniparamétricos. Finalmente, la simetría involutiva correspondiente al neutro es la isometría $s_e(g) = g^{-1}$. Para otro punto $p \in G$ es $s_p(g) = pg^{-1}p$.

Los siguientes ejemplos son los tres primeros tipos de espacios simétricos compactos que da É. Cartan en su clasificación.

Ejemplo 2.2.7. (A I) Tomamos como G el grupo especial unitario,

$$G = SU(n) = \{X \in \mathcal{M}_{n\times n}(\mathbb{C})\colon XX^* = \mathrm{I}, \det X = 1\}.$$

El automorfismo es la conjugación, $\sigma(X) = \overline{X}$. Así, el conjunto de puntos fijos es $K = SO(n)$, las matrices ortogonales reales con determinante 1.

La variedad $M = SU(n)/SO(n)$ es difeomorfa a las matrices unitarias especiales simétricas (ver Ejemplo 2.4.2).

Ejemplo 2.2.8. (A II) $G = SU(2n)$ con el automorfismo $\sigma(X) = -J\overline{X}J$, donde J es la forma simpléctica $J = \begin{pmatrix} 0 & \mathrm{I} \\ -\mathrm{I} & 0 \end{pmatrix}$. El conjunto de puntos fijos es exactamente la imagen del morfismo

$$\begin{array}{rcl} Sp(n) & \longrightarrow & SU(2n) \\ A + \mathbf{j}B & \mapsto & \begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix} \end{array}$$

donde $Sp(n)$ es el grupo simpléctico formado por las matrices $n \times n$ cuaterniónicas tales que $XX^* = \mathrm{I}$.

En este caso, la variedad $M = SU(2n)/Sp(n)$ es difeomorfa a las matrices unitarias antisimétricas.

Ejemplo 2.2.9. (A III) Tomemos $G = SU(p+q)$ y el automorfismo $\sigma(X) = \mathrm{I}_{p,q} X \mathrm{I}_{p,q}$ donde

$$I_{p,q} = \begin{pmatrix} -\mathrm{I}_p & 0 \\ 0 & \mathrm{I}_q \end{pmatrix}$$

es una forma simétrica de signatura (p,q). El conjunto de puntos fijos es el subgrupo imagen del morfismo

$$\begin{array}{ccc} U(p) \times U(q) & \longrightarrow & U(p+q) \\ (X,Y) & \mapsto & \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \end{array}$$

que además cumplen $\det X \det Y = 1$. Lo denotamos $S(U_p \times U_q)$.

La variedad $M = SU(p+q)/S(U_p \times U_q)$ es la Grassmanniana compleja formada por los subespacios de dimensión p de $\mathbb{C}^{p+q}$. En efecto, G actúa transitivamente ya que un isomorfismo lineal conserva la dimensión y, si conserva un subespacio, conserva su ortogonal para el producto hermítico.

2.3. Clasificación de los espacios simétricos irreducibles, compactos, clásicos

Para obtener una clasificación de los espacios simétricos riemannianos asociamos a cada uno un par simétrico (G,K,σ) y a éste sus álgebras de Lie y el morfismo inducido $(\mathfrak{g},\mathfrak{h},s)$. Recíprocamente, toda la información queda codificada en el álgebra de Lie $\mathfrak{g}$ y en el automorfismo s pues $\mathfrak{h}$ son los puntos fijos de s y, si tomamos los grupos de Lie simplemente conexos G y K correspondientes a $\mathfrak{g}$ y $\mathfrak{h}$, es posible integrar el morfismo de álgebras de Lie $s\colon \mathfrak{g} \to \mathfrak{g}$ y recuperar el cociente G/K y el automorfismo σ.

Esta descripción algebraica permitió a E. Cartan obtener en 1926 una clasificación completa de todos los espacios simétricos.

Definición 2.3.1. Un par simétrico $(\mathfrak{g},\mathfrak{h},s)$ de álgebras de Lie se dice *irreducible* si $\mathfrak{g}$ es semisimple y $\mathfrak{h}$ no contiene ideales no nulos de $\mathfrak{g}$ y $\mathfrak{h}$ es una subálgebra maximal propia.

Diremos que un espacio simétrico es *irreducible* si lo es el par simétrico asociado.

Se puede ver que cualquier espacio simétrico Riemanniano simplemente conexo es producto de irreducibles. Como la cubierta universal de un espacio simétrico es un espacio simétrico, podemos reducirnos a estudiar, sin pérdida de generalidad, los M simplemente conexos. Por tanto, basta con clasificar los espacios simétricos Riemannianos irreducibles y simplemente conexos.

2.3.1. Clasificación

Con todas estas herramientas pasamos a dar la clasificación de los posibles espacios simétricos Riemannianos irreducibles, compactos y simplemente conexos clásicos. Seguimos la clasificación que dan A.T. Fomenko y Helgason [21, 31].

Tipo	Modelo Cartan	dim	$\sigma(X)$
AI	$SU(n)/SO(n)$	$(n-1)(n+2)/2$	$\overline{X}$
AII	$SU(2n)/Sp(n)$	$(n-1)(2n+1)$	$-J\overline{X}J$
AIII	$SU(p+q)/SU(p)\times SU(q)$	$2pq$	$I_{p,q}X\, I_{p,q}$
BDI	$SO(p+q)/SO(p)\times SO(q)$	pq	$I_{p,q}X\, I_{p,q}$
DIII	$SO(2n)/U(n)\quad [n\geq 4]$	$n(n-1)$	$-J\overline{X}J$
CI	$Sp(n)/U(n)\quad [n\geq 3]$	$n(n+1)$	$-\mathbf{i}X\mathbf{i}$
CII	$Sp(p+q)/Sp(p)\times Sp(q)$	$4pq$	$I_{p,q}X\, I_{p,q}$

M. Mimura añade a esta lista $SO(n+1)/SO(n)$, que no es más que una esfera [61].

2.4. Modelo de Cartan

Una cuestión que surge de manera natural es la de si es posible embeber el espacio simétrico $M\cong G/K$ en G. Si G/H fuese un espacio homogéneo en general no sería posible, pero al ser un espacio simétrico sí.

El modelo de Cartan nos permite afirmar que cualquier espacio simétrico M puede interpretarse como una subvariedad de G. Este modo de ver un espacio simétrico será muy útil para tratar de calcular su categoría LS; es la herramienta que utilizan M. Mimura y K. Sugata en su artículo de 2008 [61].

Proposición 2.4.1. *La aplicación de G/K en G que lleva $[g]$ en $g\sigma(g)^{-1}$ es un difeomorfismo con la imagen.*

Es inmediato ver que es inyectiva. Nótese que mientras que los puntos de K son fijos, M está contenida en $N=\{g\in G:\sigma(g)=g^{-1}\}$. En efecto, si $n=g\sigma(g)^{-1}$, entonces

$$\sigma(n)=\sigma(g)\sigma^2(g^{-1})=n^{-1}.$$

Así pues, si llamamos

$$\begin{aligned} M &= \{h\in G:\ h=g\sigma(g)^{-1}, g\in G\}\\ N &= \{g\in G:\ \sigma(g)=g^{-1}\}, \end{aligned}$$

tenemos que

$$G/K \cong M \subset N.$$

Más aún, M es precisamente la componente conexa de N que contiene al neutro, $M = N_e$ [21, cap. 4].

Describimos ahora el modelo de Cartan de algunos espacios simétricos clásicos para los que calcularemos la categoría LS. Todas las demostraciones las hace S. Ramanujam en [69], también H. Kadzisa y M. Mimura en [49].

Ejemplo 2.4.2. $\mathbf{U(n)/O(n)}$. En este caso el automorfismo es la conjugación, $\sigma(X) = \overline{X}$. El conjunto de puntos fijos es $O(n)$ y las variedades M y N coinciden [49, Teor. 3.1]. Por tanto, podemos identificar el espacio simétrico $U(n)/O(n)$ con las matrices simétricas unitarias,

$$U(n)/O(n) \cong \{Y \in U(n) : Y = Y^T\}.$$

Ejemplo 2.4.3. $\mathbf{SO(2n)/U(n)}$. Para este espacio tomamos $\sigma(X) = -J\overline{X}J$, con $J = \begin{pmatrix} 0 & \mathrm{I} \\ -\mathrm{I} & 0 \end{pmatrix}$. Puesto que las matrices complejas de $U(n)$ son matrices de la forma $A + B\mathbf{i}$ con $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$, podemos considerarlas como las matrices de $SO(2n)$ del tipo $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$. Así, el conjunto de puntos fijos de σ es $U(n)$ y tenemos que

$$\begin{array}{rcl} M & = & \{Y \in SO(2n) \colon Y = -XJX^TJ,\ X \in SO(2n)\} \\ N & = & \{X \in SO(2n) \colon JX = X^TJ\}. \end{array}$$

Con el embebimiento de $U(n)$ en $SO(2n)$ se puede probar que $SO(2n)/U(n) \cong M$ es difeomorfo a las matrices antisimétricas de $SO(2n)$.

Ejemplo 2.4.4. $\mathbf{Sp(n)/U(n)}$. La involución es $\sigma(X) = -\mathbf{i}X\mathbf{i}$ y

$$N = \{X \in Sp(n) : X = -\mathbf{i}X^*\mathbf{i}\}.$$

Mediante un embebimiento de las matrices cuaterniónicas en las complejas se obtiene que este espacio es difeomorfo a

$$Sp(n)/U(n) \cong \{Y \in Sp(n) : Y = Y^T\}.$$

Ejemplo 2.4.5. $\mathbf{U(2n)/Sp(n)}$. Como vimos en el Ejemplo 2.2.8, $\sigma(X) = -J\overline{X}J$ es el automorfismo. Para este espacio simétrico la variedad

$$M = \{Y \in U(2n) : Y = -XJX^TJ, X \in U(2n)\}$$

mientras que

$$N = \{X \in U(2n) : JX^T = XJ\}.$$

Se prueba que

$$U(2n)/Sp(n) \cong \{Y \in U(2n) : Y = -Y^T\}.$$

Capítulo 3

Funciones de Morse

El comienzo del estudio de la categoría LS se sitúa en el intento de acotar el número de puntos críticos de una función diferenciable. De hecho, la categoría LS (más uno) de una variedad compacta es una cota inferior para el número de puntos críticos de cualquier función diferenciable, sea de Morse o no. A grandes rasgos, esto se debe a que el flujo del gradiente determina, para cada punto crítico, un abierto categórico. Nuestro interés en las funciones de Morse está en que para un grupo de Lie este flujo puede pasarse el álgebra de Lie mediante la transformación de Cayley estudiada en el Capítulo 4. Más adelante veremos que este resultado también puede extenderse a los espacios simétricos clásicos.

La diferencia entre la teoría de Morse y la de Lusternik y Schnirelmann radica en que la primera exige que los puntos críticos sean no degenerados. Esto no siempre nos da los resultados más precisos, como ocurre por ejemplo en el toro T^2, para el que podemos obtener funciones diferenciables con sólo tres puntos críticos, mientras que toda función de Morse tiene como mínimo cuatro puntos críticos.

3.1. Funciones de Morse

J.W. Milnor hace un resumen preciso y completo de la teoría de Morse en [60].

Definición 3.1.1. Sea M una variedad diferenciable. Dada una función con valores reales, $f : M \to \mathbb{R}$ diferenciable, se dice que un punto $p \in M$ es un *punto crítico* de f si la diferencial $f_{*p} : T_pM \to T_{f(p)}\mathbb{R}$ 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 p un punto crítico de f, a partir de la matriz formada por las derivadas segundas de f en p, podemos definir una forma bilineal simétrica H_pf en el tangente, el *Hessiano de f en p.*

Definición 3.1.2. Un *punto crítico* p es *no degenerado* si el Hessiano de f en p es no singular.

Definición 3.1.3. Para referirnos al índice de H_pf en el T_pM hablaremos del *índice de f en p.* Es la mayor dimensión de un subespacio de V en el cual H es definida negativa.

El índice 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.

Conviene señalar que los conceptos de punto crítico, índice y ser no degenerado son independientes de la elección de las coordenadas locales que tomemos.

Definición 3.1.4. Una función $f : M \to \mathbb{R}$ es una *función de Morse* si todos sus puntos críticos son no degenerados. En particular, son aislados.

Teorema 3.1.5. *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.

3.2. 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 un toro en horizontal sobre un plano, entonces la función altura respecto de dicho plano tendrá dos subvariedades críticas; la circunferencia inferior sobre la que el toro está apoyado en el plano y la superior.

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

Definición 3.2.1. 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 3.2.2. Una *función de Bott-Morse* es una función diferenciable cuyos puntos críticos forman una subvariedad cerrada y la Hessiana es no degenerada.

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

3.3. Teoría de Morse en los grupos de Lie

Sea $G = O_n(\mathbb{K})$ un grupo de Lie ortogonal embebido en el espacio euclídeo de las matrices $\mathcal{M}_{n\times n}(\mathbb{K})$. La métrica euclídea en este espacio viene dada por el producto escalar parte real de la traza, $\langle A, B\rangle = \Re\operatorname{Tr}(A^*B)$. Clásicamente, las funciones de Morse que se consideran en los grupos de Lie son las llamadas funciones *altura* o *distancia*. Las primeras miden la altura del grupo G con respecto a algún hiperplano; fueron estudiadas, entre otros, por I.A. Dynnikov y A.P. Veselov en [84] y A.N. Kozachko y K.Y. Volchenko en [86]. Las segundas miden la distancia a un punto dado y las estudió H. Duan en [12]. Veremos que ambos tipos de funciones son, salvo una constante, de la forma $h_X(A) = \Re\operatorname{Tr}(X^*A)$ (ver Prop. 3.3.2).

El primero en estudiar este tipo de funciones fue T. Frankel en 1963 [23]. Escoge como X la matriz identidad (el estudio que hace sirve para cualquier matriz $X = tI$ con $t \in \mathbb{R}\setminus\{0\}$). Con esta elección la función altura es invariante por la acción adjunta, de modo que para calcular el conjunto de puntos críticos nos llega con estudiar los puntos críticos de un toro maximal; todos los demás son las órbitas de éstos. Además, como el gradiente $\operatorname{grad} h_I$ es tangente a cada punto de un toro maximal dado T, los puntos críticos de h_I en T son los puntos críticos de la restricción de h_I a T.

En [23] Frankel prueba que el conjunto de puntos críticos $\Sigma(h_I^G)$ está formado por todas las matrices en G tales que $A^2 = I$. Así pues, los puntos críticos ya no son aislados sino que forman subvariedades, de hecho, son Grassmannianas. Es decir, h_I no es estrictamente una función de Morse sino de Bott-Morse.

Después de estudiar detalladamente el comportamiento de esta función en los grupos de Lie clásicos, T. Frankel lo generaliza a variedades de Stiefel [23].

Pocos años después de este estudio, en 1969, S. Ramanujam [70] obtuvo una expresión explícita para las subvariedades críticas de h_I en los grupos de Lie ortogonales. Generalizó también esta función a los espacios simétricos $U(n)/O(n)$, $Sp(n)/U(n)$, $U(2n)/Sp(n)$ y $SO(2n)/U(n)$ obteniendo de nuevo que un punto A de uno de estos espacios simétricos M es un punto crítico de h_I si y solo si $A^2 = I$.

Recientemente, H. Kadzisa y M. Mimura han utilizado la función parte real de la traza para construir la descomposición en conos de variedades de Stiefel y de algunos espacios simétricos Riemannianos y así obtener la categoría LS de estos espacios [49].

3.3.1. Funciones altura y distancia en los grupos ortogonales

La función h_I en un grupo $G = O_n(\mathbb{K})$ no es propiamente una función de Morse pero basta modificarla con un coeficiente para que sí lo sea. K. Volchenko y A. Kozachko [86] estudian en los grupos de Lie clásicos la función altura

$$h_X^G(A) = 2\Re\,\mathrm{Tr}(XA)$$

con

$$X = \mathrm{diag}(x_1, \ldots, x_n),\ x_i \in \mathbb{R} \text{ y } 0 < x_1 < \cdots < x_n.$$

Prueban que los puntos críticos de h_X^G son las matrices diagonales

$$\mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_n), \quad \varepsilon_i = \pm 1,$$

con índice

$$(\mathrm{ind} h_X^G)(A) = \sum_{i=1}^{n} \delta_{\varepsilon_i,1}\left((\dim_{\mathbb{R}} \mathbb{K})i - 1\right).$$

Utilizando las desigualdades de Morse obtienen que h_X^G es una función de Morse perfecta, es decir, con el menor número posible de puntos críticos.

Después generalizan estos resultados a variedades de Stiefel.

Esta misma función la estudian I. Dynnikov y A. Veselov en grupos de Lie y en algunos espacios simétricos embebidos en los grupos de Lie clásicos [84]. Prueban que el conjunto de puntos críticos para h_X^G es

$$\Sigma(h_X^G) = \{A \in G \colon XA = (XA)^*\}.$$

Describen explícitamente el flujo del gradiente y obtienen así una descomposición celular para la variedad.

Todas estas funciones son casos particulares del Teorema 5.1.5, que proporciona una descripción general del conjunto de puntos críticos de una función altura arbitraria h_X^G.

3.3.2. Espacio tangente y ortogonal

En la Sección 4.1.2 recogimos una descripción del espacio tangente de los grupos de Lie clásicos. A partir de ella se obtiene una caracterización de las matrices que están en el espacio perpendicular al tangente de una $X \in G$ dada, $\gamma_X G$, donde G denota al grupo $O_n(\mathbb{K})$.

Lema 3.3.1. *Para $X \in G$,*

1. *El espacio tangente a la variedad en X está formado por:*

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

2. En consecuencia, el espacio normal es:

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

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

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

donde el primer sumando pertenece al $T_I G$ y el segundo a $\gamma_I G$.

Proposición 3.3.2. *Toda función altura en G es de la forma*

$$h_X^G(A) = \Re\operatorname{Tr}(XA).$$

Demostración. Sea $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ 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 solo en una constante), de manera que el hiperplano es simplemente $X^\perp$. Supongamos también, sin pérdida de generalidad, que $\|X\| = 1$.

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

$$A = h_X^G(A)X + V, \text{ con } V \in X^\perp,$$

así

$$\langle X, A\rangle = h_X^G(A)\langle X, X\rangle + 0 = h_X^G(A)$$

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

$$h_X^G(A) = \Re\operatorname{Tr}(X^*A). \quad \square$$

3.3.3. Gradiente y Hessiano de las funciones altura

Lema 3.3.3. *Sea $h_X^G : G \to \mathcal{M}_{n\times n}(\mathbb{K})$ una función altura. El gradiente de h_X^G en un punto $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ viene dado por:*

$$\operatorname{grad} h_X^G(A) = 1/2(X^* - AXA).$$

Demostración. Sea $\widehat{h}_X$ la función altura en $\mathcal{M}_{n\times n}(\mathbb{K})$. Entonces $\operatorname{grad} h_X^G(A)$ en G es la proyección del gradiente de $\widehat{h}_X$ en en el espacio tangente $T_A G$. Por definición, para todo $V \in T_I G$,

$$\langle \operatorname{grad} h_X^G(I), V\rangle = (h_X^G)_{*I}(V)$$

de modo que, tomando W en el ortogonal,

$$\langle \operatorname{grad} h_X^G(I) + W, V\rangle = (h_X^G)_{*I}(V).$$

Pero como $h_X^G(A) = \Re \operatorname{Tr}(XA)$ es lineal,

$$(h_X^G)_{*I}(V) = \Re \operatorname{Tr}(XV).$$

Por otro lado, si llamamos $Y = \operatorname{grad} h_X^G(I) + W$, se verifica que

$$\Re \operatorname{Tr}(Y^*V) = \langle Y, V\rangle = \Re \operatorname{Tr}(XV),$$

luego $Y^* = X$, es decir,

$$\operatorname{grad} h_X^G(I) = X^* - W \qquad \text{con } W \in \gamma_I G.$$

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

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

entonces

$$\operatorname{grad} h_X^G(A) = A\frac{1}{2}[(XA)^* - XA] = \frac{1}{2}(X^* - AXA). \quad \square$$

Corolario 3.3.4. *Un punto A es crítico para la función altura h_X^G si y solo si $X^* = AXA$.*

Lema 3.3.5. *Sea $A \in G$ un punto crítico para la función altura h_X^G y $U \in T_A(G)$, entonces el operador Hessiano viene dado por*

$$(Hh_X^G)_A(U) = -1/2(AXU + UXA).$$

Demostración. Por definición

$$\begin{aligned}
(Hh_X^G)_A(U) &= \lim_{t\to 0} \frac{1}{t}[\operatorname{grad}(A+tU) - \operatorname{grad}(A)] \\
&= \lim_{t\to 0} \frac{1}{2t}[(X^* - (A+tU)X(A+tU)) - (X^* - AXA)] \\
&= \lim_{t\to 0} \frac{1}{2t}(-AXA - tAXU - tUXA - t^2UXU + AXA) \\
&= -\frac{1}{2}(AXU + UXA). \quad \square
\end{aligned}$$

3.3.4. Flujo del gradiente

Recogemos a continuación un resultado de K. Volchenko y A. Kozachko que permite transformar el flujo del gradiente de la función altura en G en un flujo en el álgebra de Lie mediante la transformación de Cayley [86]. Utilizando la expresión de los autores denominamos “linealización” a este proceso. Más adelante generalizaremos este resultado a espacios simétricos y para cualquier punto crítico.

Lema 3.3.6. *Una solución de*

$$X' = \frac{1}{2}(B^* - XBX) \tag{3.1}$$

en un entorno de I es la imagen por la transformación de Cayley de la solución de la ecuación diferencial:

$$c'(X) = -2(Bc(X) + c(X)B). \tag{3.2}$$

Demostración.

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

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

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

Proposición 3.3.7. *La ecuación (3.1) puede resolverse explícitamente para la condición inicial $X(0) = X_0$. La solución es de la forma:*

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

Líneas de flujo de la función de Morse con $B \neq I$

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

$$\psi(+\infty) = \lim_{t\to+\infty} \psi(t) = c(0) = I.$$

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

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

pues B y e^{-tB} conmutan.

Es decir, las líneas de flujo de los puntos de Ω van de I a $-I$.

Parte II

Resultados

Capítulo 4

Transformación de Cayley

4.1. La transformación de Cayley clásica

La transformación de Cayley clásica fue introducida en 1846 [5] como un modo de expresar una transformación ortogonal en coordenadas antisimétricas. Viene dada por $c(X) = (I - X)(I + X)^{-1}$ y está definida en el abierto Ω de todas las matrices que no tengan como autovalor al -1. Es involutiva, $c^2 = \mathrm{id}$, y se puede pensar como una generalización de la proyección estereográfica. Sus principales propiedades se recogen en [67]; puede verse también el libro de Weyl [87].

Como veremos más adelante esta aplicación lleva un grupo de Lie de tipo ortogonal G en su álgebra de Lie $\mathfrak{g}$. Así, la propiedad de ser involutiva resulta especialmente interesante porque nos permite traducir la contracción radial del espacio vectorial $\mathfrak{g}$ en una contracción del abierto Ω.

4.1.1. Definición y propiedades

Para pasar de un grupo de Lie ortogonal a su álgebra de Lie utilizaremos la transformación de Cayley $c\colon \Omega \to \Omega$ tal que

$$c(X) = \frac{I - X}{I + X},$$

donde

$$\Omega = \{X \in \mathcal{M}_{n\times n}(\mathbb{K})\colon \text{existe } (I + X)^{-1}\}.$$

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

$$(I - X)(I + X)^{-1} = (I + X)^{-1}(I - X).$$

Además, como
$$(I + c(X))^{-1} = \frac{1}{2}(I + X),$$
tenemos que la aplicación $c\colon \Omega \to \Omega$ está bien definida.

Recogemos a continuación algunas de la propiedades de esta aplicación que nos serán útiles a la hora de generalizarla.

Proposición 4.1.1. *Sea $A \in \Omega$, se verifica entonces que:*

1. *$A^* \in \Omega$ y $c(A^*) = c(A)^*$;*
2. *si A es inversible, $A^{-1} \in \Omega$ y $c(A^{-1}) = -c(A)$;*
3. *si $-A \in \Omega$, $c(-A) = c(A)^{-1}$;*
4. *si $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ es inversible, $BAB^{-1} \in \Omega$ y $c(BAB^{-1}) = Bc(A)B^{-1}$.*

Demostración. 1. $(I + A^*)^{-1} = \left((I + A)^{-1}\right)^*$;

2. $(I + A^{-1})^{-1} = A(I + A)^{-1}$;
3. utilizando el apartado anterior es inmediato comprobar que $c(A)c(-A) = I$;
4. basta escribir $I + BAB^{-1}$ como $B(I + A)B^{-1}$.

□

Proposición 4.1.2. *La aplicación c es involutiva, es decir, $c^{-1} = c$.*

Demostración. Consideremos $A \in \Omega$, entonces $c(c(A)) = (I - c(A))(I + c(A))^{-1}$. Al comprobar que la transformación de Cayley estaba bien definida obtuvimos que $(I + c(A))^{-1} = \frac{1}{2}(I + A)$. Además, tenemos que
$$\begin{aligned} I - c(A) &= I - (I - A)(I + A)^{-1} \\ &= (I + A)(I + A)^{-1} - (I - A)(I + A)^{-1} \\ &= 2A(I + A)^{-1}. \end{aligned}$$
Por tanto, $c(c(A)) = 0$. □

Nota–. Supongamos que A es una matriz unitaria o simpléctica. Recordemos que en este caso A se puede diagonalizar a una matriz compleja diagonal, $A = UDU^*$, con $D = \mathrm{diag}(z_1, \ldots, z_n)$ (ver Subsecc. 1.1.6). Entonces,
$$c(A) = U\,\mathrm{diag}\left(\pi(z_1), \ldots, \pi(z_n)\right) U^*,$$
donde π es la proyección estereográfica $\pi : S^1 \setminus \{-1\} \to \mathbf{i}\mathbb{R}$, $\pi(z) = \frac{1-z}{1+z}$.

Con la siguiente Proposición establecemos la relación entre los autovalores de una matriz y los de su imagen por la transformación de Cayley.

Proposición 4.1.3. *Si q es un autovalor por la derecha de $A \in \Omega$, entonces $\frac{1-q}{1+q}$ es un autovalor por la derecha de $c(A)$.*

Demostración. Fijemos q un autovalor por la derecha de A. Necesariamente $q \neq -1$, ya que si -1 fuese autovalor $I + A$ no sería inversible. Así, para algún $v \neq 0$ tenemos que

$$(I + A)v = v + Av = v(1 + q),$$

por tanto,

$$c(A)v = \frac{I - A}{I + A}v = \frac{v - Av}{1 + q} = v\frac{1 - q}{1 + q}. \qquad \square$$

4.1.2. Grupos ortogonales clásicos

Consideremos $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ó $\mathbb{H}$. Decimos que una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ es *ortogonal* si $AA^* = I$. Estas matrices se pueden identificar con una aplicación $\mathbb{K}$-lineal (por la derecha) $\mathbb{K}^n \to \mathbb{K}^n$ que conserva el producto hermítico $\langle u, v\rangle = u^*v$. Denotaremos por $O_n(\mathbb{K})$ el grupo de Lie de las matrices ortogonales,

$$O_n(\mathbb{K}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon AA^* = I\}.$$

Este grupo corresponde al grupo ortogonal $O(n)$, al unitario $U(n)$ o al simpléctico $Sp(n)$ dependiendo de que $\mathbb{K}$ sea $\mathbb{R}, \mathbb{C}$ ó $\mathbb{H}$ respectivamente.

Vamos a ver que la transformación de Cayley clásica lleva los grupos de Lie ortogonales clásicos en su álgebra de Lie $\mathfrak{g}$ formada por las matrices antisimétricas (o antihermíticas).

Proposición 4.1.4. $\mathfrak{g} = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon X + X^* = 0\}$.

Proposición 4.1.5. *$c \colon \Omega \cap G \to \Omega \cap \mathfrak{g}$ es un difeomorfismo.*

Demostración. Es un caso particular del Teorema 4.3.2. $\square$

Como espacio vectorial $\mathfrak{g} = T_I G$, de modo que el espacio tangente en cualquier otro punto $A \in G$ viene dado por

$$T_A G = L_A(T_I G) = \{Y \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon A^*Y + Y^*A = 0\}.$$

La Proposición que enunciamos a continuación será clave para poder establecer un difeomorfismo entre el espacio tangente $T_{A^*}G$ y el abierto $\Omega_G(A)$ (ver Subsección 4.3.1) y probar así que este tipo de abiertos son contráctiles.

Proposición 4.1.6. *Sea $X \in \mathfrak{g}$ una matriz antisimétrica (antihermítica). Entonces X no tiene autovalores reales no nulos.*

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

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

Por lo tanto, v^*Xv se anula, *i.e.*, $|v|^2t = 0$, luego $t = 0$. □

Ejemplo 4.1.7. Consideremos $Sp(1) = \{q \in \mathbb{H} : |q| = 1\} = S^3$. Su álgebra de Lie es

$$\mathfrak{sp}(1) = \{w \in \mathbb{H} : w + \overline{w} = 0\} = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle.$$

En este caso el abierto Ω 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 $\mathbb{R}^3$:

$$c\colon S^3 \setminus \{-1\} \to \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle = \mathbb{R}^3$$
$$q \mapsto c(q) = \frac{1-q}{1+q}$$

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 $q = t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ con $t, x, y, z \in \mathbb{R}$,

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

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

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

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

Por otro lado, $SU(2)$ es un subgrupo de $U(2)$ y en el grupo unitario tenemos la aplicación de Cayley $c\colon \Omega \cap U(2) \to \mathfrak{u}(2)$. De hecho, $c(SU(2)) \subseteq \mathfrak{su}(2)$, es decir, c lleva las matrices de determinante 1 en matrices de traza 0. En efecto, evaluemos c sobre un elemento $A \in SU(2)$. Tenemos

$$I + A = \begin{pmatrix} 1+\alpha & -\overline{\beta} \\ \beta & 1+\overline{\alpha} \end{pmatrix},$$

luego

$$\det(I + A) = 2\left(1 + \Re(\alpha)\right).$$

Por tanto, como

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

queda $\operatorname{Tr} c(A) = 0$. Concluimos pues que el isomorfismo $Sp(1) \cong SU(2)$ conmuta con la aplicación de Cayley.

$$\begin{tikzcd} U(2) \arrow[d, "c"] & SU(2) \arrow[l, hook'] \arrow[r, "\cong", "\varphi"'] \arrow[d, "c"] & Sp(1) \arrow[d, "c"] \\ \mathfrak{u}(2) & \mathfrak{su}(2) \arrow[l, hook'] \arrow[r, "\cong", "\varphi_*"'] & \mathfrak{sp}(1) \end{tikzcd}$$

4.2. La transformación de Cayley generalizada

Recordemos que en una variedad compacta la categoría LS (más uno) es una cota inferior para el número de puntos críticos de cualquier función diferenciable, sea de Morse o no. A grandes rasgos, esto se debe a que el flujo del gradiente determina, para cada punto crítico, un abierto categórico. En nuestro contexto, obtenemos que para las funciones altura, tanto en los grupos de Lie como en los espacios simétricos, este flujo viene dado por la contracción asociada a la transformación de Cayley.

La generalización de la transformación de Cayley que hacemos a continuación permite establecer un método que sirva para dar abiertos contráctiles. Modificamos así la definición de c para obtener un recubrimiento por abiertos categóricos de los grupos de Lie ortogonales. A continuación, veremos que los resultados obtenidos para los grupos pueden adaptarse de modo natural a los espacios simétricos clásicos sin más que verificar que todas las construcciones son compatibles con el automorfismo σ.

Además, veremos que para las funciones de Bott-Morse, la contracción de Cayley proporciona un modelo local de las variedades críticas.

4.2.1. Definición y propiedades de c_A

Sea $A \in O_n(\mathbb{K})$ una matriz ortogonal, unitaria o simpléctica según $\mathbb{K}$ sea $\mathbb{R}, \mathbb{C}$ o $\mathbb{H}$. 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 4.2.1. La *transformación de Cayley centrada en* A se define como la aplicación

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

dada por

$$c_A(X) = c(A^*X)A^* = A^*c(XA^*).$$

La transformación de Cayley clásica c corresponde al caso $A = I$. Como veremos en la siguiente Proposición, la aplicación c_A está bien definida, ya que si $X \in \Omega(A)$ entonces $c_A(X) \in \Omega(A^*)$, y es inversible con $c_A^{-1} = c_{A^*}$.

Proposición 4.2.2. *Si* $X \in \Omega(A)$ *entonces,*

1. $c_A(X) = (A + X)^{-1}(I - XA^*) = (I - A^*X)(A + X)^{-1}$;
2. *la inversa de la matriz* $A^* + c_A(X)$ *es* $(1/2)(A + X)$;
3. *si* $X \in \Omega(A)$, *entonces* $c_A(X) \in \Omega(A^*)$;
4. c_A *es un difeomorfismo con* $(c_A)^{-1} = c_{A^*}$.

Demostración. Las dos primeras afirmaciones se deducen de las propiedades de la transformación de Cayley clásica. La tercera afirmación se deduce de la segunda. Haciendo uso de ésta obtenemos que

$$\begin{aligned} c_{A^*}c_A(X) &= c_{A^*}\left(c(A^*X)A^*\right) \\ &= Ac\left[(c(A^*X)A^*)A\right] \\ &= Ac^2(A^*X) \\ &= X. \qquad \square \end{aligned}$$

Necesitaremos también las siguientes propiedades.

Proposición 4.2.3. *Sea $X \in \Omega(A)$. Entonces*

1. $X^* \in \Omega(A^*)$ *y* $c_{A^*}(X^*) = c_A(X)^*$;

2. $UXU^* \in \Omega(UAU^*)$ *para cualquier matriz* $U \in O_n(\mathbb{K})$ *y*

$$c_{UAU^*}(UXU^*) = Uc_A(X)U^*;$$

3. *si la matriz X es inversible, entonces $X^{-1} \in \Omega(A^*)$ ya que*

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

Más aún,

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

4.3. Abiertos categóricos en un grupo de Lie

La transformación de Cayley generalizada c_A establece un difeomorfismo entre $\Omega_G(A)$ y $T_{A^*}G$; esto nos permitirá obtener abiertos contráctiles en el grupo como imagen difeomórfica de un espacio vectorial. A partir de las propiedades de c_A que hemos visto y de la Proposición 4.1.6 obtenemos el siguiente interesante resultado.

4.3.1. Un difeomorfismo entre $T_{A^*}G$ y $\Omega_G(A)$

Proposición 4.3.1. *El espacio vectorial real $T_{A^*}G$ está contenido en $\Omega(A^*)$.*

Demostración. Supongamos que existe algún $Y \in T_{A^*}G$ tal que $A^* + Y$ no es inversible. Entonces, existe algún $v \neq 0$ tal que $A^*v = -Yv$, luego $-v = AYv$. Es decir, -1 sería autovalor de la matriz antisimétrica AY, lo que contradice la Prop. 4.1.6. $\square$

Sea $G = O_n(\mathbb{K})$, denotaremos por $\Omega_G(A)$ al subconjunto abierto $\Omega(A) \cap G \subset G$.

Teorema 4.3.2. *La aplicación c_A lleva difeomórficamente $\Omega_G(A)$ en el espacio vectorial real $T_{A^*}G$ de las matrices X tales que $AX + (AX)^* = 0$. Además, $c_A(A) = 0$. En consecuencia, $\Omega_G(A)$ es un abierto contráctil.*

Demostración. En primer lugar, veamos que c_A envía $\Omega_G(A)$ en $T_{A^*}G \subset \Omega(A^*)$.

Sea $B \in \Omega_G(A)$, entonces $B^{-1} = B^*$ luego, por la Prop. 4.2.3

$$c_A(B)^* = c_{A^*}(B^*) = c_{A^*}(B^{-1}) = -Ac_A(B)A.$$

Llamemos X^* a $c_A(B)$, entonces

$$-AX^*A = X,$$

es decir X^*A es antisimétrica, que es la condición buscada.

Recíprocamente, sea X^* en $T_{A^*}G$, es decir X^*A antisimétrica. Debemos probar que $c_{A^*}(X^*)$ es ortogonal. De nuevo, por la Prop. 4.2.3

$$c_{A^*}(X^*) = c_A(X)^*.$$

Así que basta ver que $c_A(X)^*c_A(X) = I$.

En primer lugar, como $c_A(X) = (A+X)^{-1}(I - XA^*)$ tenemos que $c_A(X)$ es inversible si y solo si $I - XA^*$ lo es. Pero si $I - XA^*$ no fuera inversible, existiría algún $v \neq 0$ tal que $XA^*v = v$ y por tanto, 1 sería un autovalor real de la matriz antisimétrica XA^*, en contradicción con la Prop. 4.1.6.

Ahora bien, como

$$A^*X = -X^*A$$

tenemos que

$$-A^*XA^* = X^*.$$

Pero, por las propiedades de la transformación de Cayley generalizada,

$$\begin{aligned} -A^*XA^* &= -A^*c_A^*\left(c_A(X)\right)A^* = c_A\left(c_A(X)^{-1}\right) \text{ y} \\ X^* &= c_A(c_A^*(X^*)) = c_A(c_A(X)^*). \end{aligned}$$

Como c_A es inyectivo, se sigue que $c_A(X)^* = c_A(X)^{-1}$. $\square$

Nota–. De hecho $\Omega(A) = L_A(\Omega(I)) = R_A(\Omega(I))$, y como

$$T_{A^*}(G) = L_{A^*}(T_IG) = R_{A^*}(T_IG),$$

tenemos $c_A = R_{A^*} \circ c \circ L_{A^*}$.

$$\begin{tikzcd} \Omega(I) \arrow[r, "c"] & \mathfrak{g} = T_IG \arrow[d, "R_A"] \\ \Omega(A) \arrow[u, "L_{A^*}"] \arrow[r, "c_A"] & T_AG \end{tikzcd}$$

4.3.2. Contracción de Cayley generalizada

Sea $A \in G$. La contracción radial que sale de este punto podemos parametrizarla como $c_{A^*}(tX)$, con $t \in [0, \infty)$, para algún X del tangente $T_{A^*}G$. Es decir $AX = -X^*A^*$. Como $AX \in T_IG$ es antisimétrica, es normal y por tanto diagonalizable, $AX = UDU^*$. Entonces

$$\phi(t) = c_{A^*}(tX) = c_{A^*}(tA^*UDU^*) = c(AtA^*UDU^*)A,$$

esto es

$$\phi(t) = c(tUDU^*)A = Uc(tD)U^*A.$$

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

$$\lim_{t\to\infty} c(tD) = \lim_{t\to\infty} \operatorname{diag}((1 - t\lambda_k)/(1 + t\lambda_k)) = -I,$$

por tanto

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

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

4.4. Contracción de Cayley en los espacios simétricos

Análogamente al caso de los grupos de Lie, para buscar abiertos contráctiles en los espacios simétricos G/K usaremos la transformación de Cayley generalizada, c_A, en el modelo de Cartan correspondiente M. Para ello comprobaremos que c_A es compatible con la aplicación σ definida en 2.2.1. Esto nos permitirá tratar de manera simultánea y unificada todos los espacios simétricos. Recordemos que para G/K con automorfismo σ, la variedades M y N son como siguen:

$$\begin{aligned} M &= \{h \in G\colon h = g\sigma(g)^{-1}, g \in G\}, \\ N &= \{g \in G\colon \sigma(g) = g^{-1}\}. \end{aligned}$$

Como vimos en la Sección 2.4, G/K es difeomorfo a M y $M \subset N$. En general, $M \neq N$, de hecho, $M = N_e$ es la componente conexa de N que contiene al neutro. Lo que haremos entonces será contraer mediante c_A en M o, al ser $M = N_e$, basta con que lo hagamos en N.

4.4.1. Compatibilidad

Lema 4.4.1. *Sea* $\sigma\colon \mathcal{M}_{n\times n}(\mathbb{K}) \to \mathcal{M}_{n\times n}(\mathbb{K})$ *un automorfismo involutivo,* $\sigma^2 = \mathrm{id}$*, y de* $\mathbb{R}$*-álgebras. Entonces, para todo* $A \in G$*,*

$$c_{\sigma(A)} \circ \sigma = \sigma \circ c_A,$$

o, equivalentemente,

$$\sigma \circ c_{\sigma(A)} = c_A \circ \sigma.$$

Demostración. Como A es una matriz ortogonal, $\sigma(A)$ también lo es. Entonces, al ser σ de $\mathbb{R}$-álgebras,

$$\sigma(A^*) = \sigma(A^{-1}) = \sigma(A)^{-1} = \sigma(A)^*.$$

Por otro lado, sea $X \in \Omega(A)$, es decir, $A + X$ inversible, entonces, $\sigma(A + X) = \sigma(A) + \sigma(X)$ sigue siendo inversible, luego $\sigma(X)$ está en el abierto que nos interesa, $\sigma(X) \in \Omega(\sigma(A))$.

Podemos entonces hacer la composición sin problema y tenemos que

$$\begin{aligned} c_{\sigma(A)}(\sigma(X)) &= (I - \sigma(A)^*\sigma(X))\,(\sigma(A) + \sigma(X))^{-1} \\ &= (I - \sigma(A^*X))\,\sigma(A + X)^{-1} \\ &= \sigma(I - A^*X)\sigma(A + X)^{-1} \\ &= \sigma(c_A(X)). \quad \square \end{aligned}$$

Nota–. Estas hipótesis se cumplen en todos los casos que hemos visto en los Ejemplos de la Subsección 2.2.3

4.4.2. Abiertos categóricos en espacios simétricos

Las propiedades de la transformación de Cayley generalizada, en particular $c_A^{-1} = c_{A^*}$, nos permitirán pasar al modelo de Cartan la contracción del espacio tangente $T_{A^*}G$.

Teorema 4.4.2. *Fijada* $A \in N$*, para cualquier* $X \in N \cap \Omega$*, el camino*

$$\gamma(t) = c_{A^*}(tc_A(X))$$

permanece en N *para todo* $t \in [0, 1]$*.*

Demostración. 1. En primer lugar, como $A, X \in N$ y $A \in G$, al ser σ de $\mathbb{R}$-álgebras tenemos que $\sigma(A) = A^{-1} = A^*$ y $\sigma(X) = X^{-1}$.

2. Puesto que $c_A(X) \in \Omega_G(A^*)$, se verifica que $\gamma(t) = c_{A^*}(tc_A(X)) \in \Omega_G(A)$, por tanto, $\gamma(t)^{-1} \in \Omega_G(A^*)$.

   Por otro lado, $\gamma(t) \in \Omega_G(A)$ implica que $\sigma(\gamma(t)) \in \Omega_G(\sigma(A)) = \Omega_G(A^*)$.

3. Por último, utilizando las propiedades de la transformación de Cayley generalizada (Prop. 4.2.3) y el Lema 4.4.1 tenemos que

$$\begin{aligned} c_{A^*}(\gamma(t)^{-1}) &= -Ac_A(\gamma(t))A \\ &= -Atc_A(X)A \\ &= tc_{A^*}(X^{-1}) \end{aligned}$$

   y

$$\begin{aligned} c_{A^*}(\sigma(\gamma(t)) &= \sigma(c_{\sigma(A^*)}(\gamma(t))) \\ &= \sigma(c_A(\gamma(t))) \\ &= \sigma(tc_A(X)) \\ &= t\sigma(c_A(X)) \\ &= tc_{\sigma(A)}(\sigma(X)) \\ &= tc_{A^*}(\sigma(X)) \\ &= tc_{A^*}(X^{-1}). \end{aligned}$$

   Se verifica entonces que

$$c_{A^*}(\gamma(t)^{-1}) = tc_{A^*}(X^{-1})$$

   y

$$c_{A^*}(\sigma(\gamma(t)) = tc_{A^*}(X^{-1}).$$

   Por tanto, $\gamma(t) \in N$ para cualesquiera $A, X \in N$.

□

Corolario 4.4.3. *Sea el modelo de Cartan $M \subset G$ y $A \in M$. Denotamos por*

$$\Omega_M(A) = \Omega_G(A) \cap M.$$

Entonces, el abierto $\Omega_M(A)$ es contráctil.

Demostración. Sabemos que $\Omega_G(A)$ es contráctil por la contracción $\gamma(t)$. Basta entonces comprobar que $\gamma(t)$ permanece en M para todo t.

Por el Teorema 4.4.2 tenemos que la contracción permanece en la variedad N, pero al ser $M = N_e$ y $A \in M$, se verifica que la contracción no se sale de la variedad de Cartan M. □

Veremos en el Capítulo 8 como dar, a partir de estos resultados, recubrimientos categóricos explícitos mínimos de los espacios simétricos $U(n)/O(n)$ y $U(2n)/\,Sp(n)$.

Capítulo 5

Subvariedades críticas en grupos de Lie

5.1. Estudio del conjunto de puntos críticos de las funciones altura

Recogemos en esta sección una serie de propiedades que verifica el conjunto de puntos críticos de una función altura en un grupo de Lie, $\Sigma(h_X^G)$. Estas propiedades nos permitirán simplificar la caracterización de las funciones altura en G que son de Morse.

5.1.1. Compatibilidad con la descomposición polar

En primer lugar, veamos que basta estudiar el comportamiento de las funciones del tipo h_D^G con D una matriz diagonal. Los siguientes resultados suponemos que son conocidos, aunque no hemos encontrado ninguna demostración explícita en la literatura.

Lema 5.1.1. *Sea $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz cualquiera y $X = US$, con U ortogonal y S hermítica, su descomposición polar. Se verifica entonces que:*

$$\Sigma(h_X^G) = \Sigma(h_S^G)U^*.$$

Además, el conjunto de puntos críticos no degenerados de la función altura h_X^G, $\Sigma^+(h_X^G)$, es tal que

$$\Sigma^+(h_X^G) = \Sigma^+(h_S^G)U^*.$$

Demostración. Sea $A \in \Sigma(h_X^G)$, es decir,

$$X^* = AXA$$

esto es,

$$SU^* = AUSA$$

pero como U es ortogonal,

$$S = AUSAU$$

y por tanto, $AU \in \Sigma(h_S^G)$. Entonces, al ser A un elemento del grupo, $A \in \Sigma(h_S^G)U^*$.

Análogamente, si $B \in \Sigma(h_S^G)$, $S = BSB$ y por tanto

$$SU^* = BSBU^*$$

pero, de nuevo por ser U ortogonal,

$$SU^* = BU^*USBU^*,$$

es decir, $BU^* \in \Sigma(h_X^G)$.

La afirmación referida a los puntos críticos no degenerados se debe a que, según la expresión obtenida en el Lema 3.3.5, tenemos el siguiente diagrama:

$$\begin{tikzcd}
T_A G \arrow[r, "(Hh_X^G)_A"] \arrow[d, "R_U"] & T_A G \arrow[d, "R_U"] \\
T_{AU} G \arrow[r, "(Hh_S^G)_{AU}"] & T_{AU} G
\end{tikzcd}$$

que como se ve es conmutativo ya que, para $V \in T_{AU}G$:

$$(Hh_S^G)_{AU}(V) = -1/2(AUSV + VSAU)$$

pero si $V = R_U(W)$ para algún $W \in T_A(G)$ entonces

$$\begin{aligned}
(Hh_S^G)_{AU}(V) &= -1/2(AUSWU + WUSAU) \\
&= -1/2(AXW + WXA)U.
\end{aligned}$$

Esto es:

$$R_U \circ (Hh_X^G)_A = (Hh_S^G)_{AU} \circ R_U.$$

Concluimos pues por ser las traslaciones isomorfismos. $\square$

5.1.2. Compatibilidad con la descomposición SVD

La obtención de los valores singulares de la matriz X y la extensión de la descomposición SVD al ámbito cuaterniónico que recogimos en la Subsección 1.1.8 nos permitirán ahora reducir el estudio de cualquier función altura en un grupo de Lie ortogonal al estudio de las funciones altura con respecto a una matriz diagonal real.

Teorema 5.1.2. *Un punto $A \in G$ es un punto crítico de la función altura h_X^G sii V^*AU es un punto crítico de h_D^G, donde $X = UDV^*$ es la descomposición SVD de X. Además*

$$(L_{V^*} \circ R_U) \circ (Hh_X^G)_A = (Hh_D^G)_{V^*AU}.$$

Demostración. La condición de ser crítico establece que

$$X^* = AXA.$$

Entonces

$$D = D^* = V^*AUDV^*AU.$$

El resultado referido a los operadores hessianos se prueba análogamente al del Lema 5.1.1 ya que

$$(L_{V^*} \circ R_U) \circ (Hh_X^G)_A(Z) = -\frac{1}{2}V^*(AUDV^*Z + ZUDV^*A)U$$

y el diagrama correspondiente entre los tangentes es conmutativo. □

5.1.3. Compatibilidad con la diagonalización de matrices

En la línea de los resultados anteriores comprobamos ahora que la diagonalización de matrices en G también respeta el conjunto de puntos críticos de las funciones altura.

Lema 5.1.3. *Sea $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz diagonalizable, $X = UDU^*$. Se verifica que:*

$$\Sigma(h_X^G) = U\Sigma(h_D^G)U^*.$$

Además, el conjunto de puntos críticos no degenerados viene dado por:

$$\Sigma^+(h_X^G) = U\Sigma^+(h_D^G)U^*.$$

Demostración. Sea $A \in \Sigma(h_X^G)$, es decir,

$$X^* = AXA$$

esto es,

$$UD^*U^* = AUDU^*A$$

luego

$$D^* = U^*AUDU^*AU$$

y por tanto, $U^*AU \in \Sigma(h_D^G)$ o, equivalentemente, $A \in U\Sigma(h_S^G)U^*$.

Además, el siguiente diagrama es conmutativo:

$$\begin{tikzcd} T_AG \arrow[r, "(Hh_X^G)_A"] \arrow[d, "L_{U^*}\circ R_U"] & T_AG \arrow[d, "L_{U^*}\circ R_U"] \\ T_{U^*AU}G \arrow[r, "(Hh_D^G)_{U^*AU}"] & T_{U^*AU}G \end{tikzcd}$$

es decir,

$$(L_{U^*} \circ R_U) \circ (Hh_X^G)_A = (Hh_D^G)_{U^*AU} \circ (L_{U^*} \circ R_U). \quad \square$$

El siguiente lema prueba que, una vez diagonalizada la matriz, podemos cambiarla de signo por bloques.

Lema 5.1.4. *Sea la matriz real diagonal por bloques*

$$D = \mathrm{diag}(t_1I, \ldots, t_rI, p_1I, \ldots, p_sI)$$

con $t_i < 0$ *y* $p_j > 0$ *para* $1 \leq i \leq r, 1 \leq j \leq s$ *y* $C = \left(\begin{array}{c|c} -I_r & 0 \\ \hline 0 & I_s \end{array}\right)$. *Entonces* $A \in \Sigma(h_D^G)$ *sii* $CA \in \Sigma(h_{DC}^G)$. *Esto reduce el estudio al caso positivo. Además*

$$L_C \circ (Hh_D^G)_A = (Hh_{CD}^G)_{CA} \circ L_C.$$

Demostración. Supongamos $A \in G$ tal que $D^* = ADA$, es decir,

$$(DC)^* = CD = CADA = CADCCA.$$

La afirmación referida a los hessianos se prueba de la misma manera que en el Lema 5.1.3. $\square$

Concluimos pues que si buscamos caracterizar las funciones altura en un grupo de Lie ortogonal que son de Morse podemos reducirnos a las funciones altura con respecto a una matriz diagonal, real y no negativa, h_D^G.

5.1.4. Descomposición del conjunto de puntos críticos

Consideremos una función altura h_X^G con $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ arbitraria. Para estudiar el conjunto de sus puntos críticos, aplicando los Lemas 5.1.1 y 5.1.3, nos llega con analizar sólo el caso de las matrices diagonales reales con todos sus elementos no negativos.

La ventaja de poder reducirnos a este tipo de matrices es que, como veremos, podemos obtener la estructura de $\Sigma(h_D^G)$ a partir del comportamiento de cada uno de los bloques de la diagonal.

1. Cuando X es la matriz nula, los puntos críticos son todo el grupo $G = O_n(\mathbb{K})$. La función no es de Morse.

2. El caso $X = tI$ con $t \in \mathbb{R}$ es análogo al caso estudiado por Frankel en [23]. $\Sigma(h_X^G)$ está formado por los elementos del grupo que cumplen $A^2 = I$, es decir, las órbitas por la acción adjunta de todos los que tienen ± 1 en la diagonal. Para una matriz de orden n de este tipo denotamos el conjunto de puntos críticos por
$$\Sigma(h_{tI}^G) = \Sigma(n).$$

3. Consideremos ahora X una matriz diagonal cualquiera. Por el Lema 5.1.4 podemos reordenarla de manera que
$$X = \left(\begin{array}{c|c|c} t_1 I & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & t_k I \end{array}\right).$$
Es decir, los elementos de la diagonal van por bloques $t_1 I, \ldots, t_k I$, de dimensiones $n_1, \ldots, n_k$ tal vez completados antes por un bloque nulo de tamaño n_0, con $n_0 + n_1 + \cdots + n_k = n$ de manera que los autovalores estén ordenados y $0 < t_1 \leq \cdots \leq t_k$. El estudio de este caso lo haremos por inducción en el número de bloques.

Teorema 5.1.5. *Sea $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz arbitraria. Sea n_0 la dimensión de su núcleo y $0 < t_1 < \cdots < t_k$ los autovalores reales no nulos de XX^* con multiplicidades $n_1, \ldots, n_k$. Entonces el conjunto de puntos críticos de la función altura h_X^G es difeomorfo a*
$$\Sigma(h_X^G) \cong O_{n_0}(\mathbb{K}) \times \Sigma(n_1) \times \cdots \times \Sigma(n_k)$$
donde $n_0 + n_1 + \cdots + n_k = n$.

Demostración. De acuerdo con lo que acabamos de ver basta estudiar el caso de la matriz X diagonal por bloques

$$X = \left(\begin{array}{c|c|c|c} 0 & & & \\ \hline & t_1 I & & \\ \hline & & \ddots & \\ \hline & & & t_k I \end{array}\right).$$

En primer lugar, supongamos que la matriz no tiene ceros en la diagonal. Si $k = 1$, estamos en el segundo caso, $X = tI$. Sea entonces

$$S = \begin{pmatrix} S_0 & 0 \\ 0 & t_k I \end{pmatrix},$$

Para estudiar los puntos críticos de h_S^G escribimos las matrices del grupo también por bloques:

$$A = \begin{pmatrix} A_0 & U \\ V & A_1 \end{pmatrix}, \quad A_1 \in \mathcal{M}_{n_k \times n_k}(\mathbb{K}).$$

Al imponerle la condición $A \in \Sigma(h_S^G)$

$$\begin{pmatrix} A_0^* & V^* \\ U^* & A_1^* \end{pmatrix} \begin{pmatrix} S_0 & 0 \\ 0 & t_k I \end{pmatrix} = \begin{pmatrix} S_0 & 0 \\ 0 & t_k I \end{pmatrix} \begin{pmatrix} A_0 & U \\ V & A_1 \end{pmatrix}$$

obtenemos

$$\begin{aligned} A_0^* S_0 &= S_0 A_0 && (5.1)\\ t_k V^* &= S_0 U && (5.2)\\ U^* S_0 &= t_k V && (5.3)\\ A_1^* &= A_1. && (5.4) \end{aligned}$$

Por otro lado, como A está en el grupo,

$$\begin{pmatrix} A_0 & U \\ V & A_1 \end{pmatrix} \begin{pmatrix} A_0^* & V^* \\ U^* & A_1 \end{pmatrix} = \begin{pmatrix} A_0^* & V^* \\ U^* & A_1 \end{pmatrix} \begin{pmatrix} A_0 & U \\ V & A_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

de donde, como A_1 es hermítica, $VV^* = U^*U$. Tenemos así que

$$(t_k V^*)^*(t_k V^*) = t_k^2 V V^* = t_k^2 U^* U,$$

luego, por la condición (5.2)

$$U^* S_0^2 U = t_k^2 U^* U.$$

Para la matriz de la primera parte de la igualdad, de orden $n_k \times n_k$, obtenemos que el elemento i-ésimo de la diagonal es de la forma:

$t_1^2|u_{1i}|^2+\ldots+t_1^2|u_{n_1i}|^2+t_2^2|u_{n_1+1i}|^2+\ldots+t_2^2|u_{n_1+n_2i}|^2+\ldots+t_{k-1}^2|u_{(m-n_{k-1})+1i}|^2+\ldots t_{k-1}^2|u_{mi}|^2$

mientras que para la segunda:

$$t_k^2(|u_{1i}|^2+|u_{2i}|^2+\ldots+|u_{ki}|^2).$$

Por tanto, de la igualdad de los dos primeros elementos de la diagonal obtenemos

$$\begin{aligned}(t_k^2-t_1^2)(|u_{11}|^2+\ldots+|u_{n-11}|^2)+(t_k^2-t_2^2)(|u_{n_1+11}|^2+\ldots+|u_{n_1+n_21}|^2)&\\ +\ldots+(t_k^2-t_{k-1}^2)(|u_{m-n_{(k-1)}+11}|^2+\ldots+|u_{m1}|^2) &= 0\end{aligned}$$

y análogamente para todos los demás. Entonces, al ser todos los autovalores positivos y distintos deducimos que U es la matriz nula y por tanto, también $V=0$. Así, los puntos críticos serán de la forma:

$$A=\begin{pmatrix} A_0 & 0\\ 0 & A_1\end{pmatrix}$$

donde A_0 es ortogonal y $A_0 \in \Sigma(h_{S_0})$ (ver condición 5.1) y A_1 es también ortogonal y punto crítico de h_{t_kI}.

Finalmente, si hubiese un bloque de ceros de tamaño n_0 en la diagonal,

$$X=\begin{pmatrix} 0I_{n_0} & 0\\ 0 & S\end{pmatrix},$$

volviendo a considerar las matrices del grupo por bloques,

$$A=\begin{pmatrix} A_0 & M\\ N & A_1\end{pmatrix},\quad A_0\in\mathcal{M}_{n_0\times n_0}(\mathbb{K})$$

la condición de ser punto crítico se reduce a $A_0\in G$, $M=N=0$ y A_1 ortogonal tal que $S^*=A_1SA_1$. Entonces

$$\Sigma(h_X^G)=O_{n_0}(\mathbb{K})\times\Sigma(h_S^G).\quad\square$$

Ejemplo 5.1.6 (Una función altura h_X definida en $U(2)\cong S^3\times S^1$)**.** Si denotamos por G_k^n la Grassmanniana $U(n)/(U(k)\times U(n-k))$, la función h_X puede tener los siguientes conjuntos críticos:

$n_0=2$: $\Sigma=U(2)$, ya que, en este caso, h_X sería la función constante nula;

$n_0=1, n_1=1$: $\Sigma=U(1)\times\Sigma(1)=S^1\sqcup S^1$, unión disjunta de dos circunferencias porque $\Sigma(1)=G_0^1\sqcup G_1^1$ son dos puntos;

$n_0=0, n_1=1, n_2=1$: $\Sigma=\Sigma(1)\times\Sigma(1)$, es decir, cuatro puntos;

$n_0=0, n_1=2$: $\Sigma=\Sigma(2)=G_0^2\sqcup G_1^2\sqcup G_2^2$, dos puntos y una esfera $S^2=S^3/S^1$.

5.2. Caracterización de las funciones de Morse

5.2.1. Prueba mediante la descomposición polar

Del estudio de la estructura del conjunto de puntos críticos y del resultado para la descomposición polar (Lema 5.1.1) se deriva la siguiente caracterización de las funciones altura en un grupo de Lie ortogonal G que son de Morse.

Teorema 5.2.1. *Dada $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ la función altura h_X^G es de Morse si y solo si XX^* es inversible y tiene todos sus autovalores distintos.*

Demostración. Dada $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ arbitraria podemos escribir

$$X = US$$

con $U = U^*$ y S simétrica (ver Teorema 1.1.14). Basándonos en el Lema 5.1.1 podemos afirmar que h_X^G es de Morse sii lo es h_S^G. Pero como S es simétrica, todos sus autovalores son reales, luego

$$S = VDV^*$$

para alguna $D = \operatorname{diag}(t_1, \dots, t_n)$ con $t_i \in \mathbb{R}$. De nuevo, nos reducimos a estudiar sólo el caso de las matrices diagonales ya que, según el Lema 5.1.3, que h_S^G sea de Morse es equivalente a que lo sea h_D^G.

Si XX^* es inversible y tiene todos sus autovalores distintos, entonces los elementos distinguidos de $DD^* = D^2 = \operatorname{diag}(t_1^2, \dots, t_n^2)$ verificarán que $t_i^2 \neq t_j^2$ para $i \neq j$ y $t_i \neq 0$, para $i, j \in \{1, \dots, n\}$. En consecuencia, cuando $i \neq j$, $|t_i| \neq |t_j|$, por tanto, h_D^G es de Morse.

Análogamente, cuando h_D^G sea de Morse, reordenando y cambiando el signo si fuese necesario, tendremos que $0 < t_1 \ldots < t_n$. Se verifica entonces que D es inversible y por tanto, también X. Además, como todos los t_i son positivos y $t_i \neq t_j$ si $i \neq j$, $t_i^2 \neq t_j^2$ para $i \neq j$. Es decir, todos los autovalores de XX^* son distintos. □

Nota–. Este resultado es equivalente a la afirmación que hacen I. Dynnikov y A. Veselov, según la cual una función h_X^G en los grupos de Lie clásicos es de Morse si y solo si los autovalores de la matriz hermítica definida positiva S de la descomposición polar $X = SU$ son diferentes dos a dos [84].

5.2.2. Prueba mediante la descomposición SVD

De manera análoga puede obtenerse este mismo resultado utilizando la descomposición en valores singulares (ver Teorema 5.1.2) en vez de la descomposición polar.

Teorema 5.2.2. *Dada $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ arbitraria, la función h_X^G es de Morse sii los valores singulares de X son distintos y no nulos.*

Demostración. Sea $X = UDV^*$ la descomposición SVD de X. Por el Teorema 5.1.2 tenemos que

$$V^*(\operatorname{grad} h_X^G)(A)U = (\operatorname{grad} h_D^G)(V^*AU)$$

y

$$V^*(Hh_X^G)_A(Y)U = (Hh_D^G)_{V^*AU}(Y).$$

Además, como U y V son matrices ortogonales, se verifica que la $\dim S(A^*)$ para la función h_X^G es la misma que $\dim S(V^*A^*U)$ para h_D^G. □

5.3. Estructura local del conjunto de puntos críticos

El siguiente resultado nos permite dar, a partir de la transformación de Cayley generalizada, una carta local del conjunto de puntos críticos de una función de Bott-Morse en un grupo de Lie. Se prueba con la misma técnica que emplearemos para espacios simétricos (ver Teor. 6.2.1) o bien deduciéndolo como un caso particular de este mismo Teorema.

Sea h_X^G una función altura arbitraria en un grupo de Lie $G = O(n, \mathbb{K})$ y $\Sigma(h_X^G)$ el conjunto de puntos críticos de h_X^G. Si $A \in \Sigma(h_X^G)$ es un punto crítico, denotamos por $S^G(A)$ al espacio vectorial real

$$S^G(A) = \{\beta_0 \in T_{A^*}G\colon\ XA\beta_0 + \beta_0 AX = 0\}.$$

Podemos dar entonces la siguiente carta local para el conjunto de puntos críticos.

Teorema 5.3.1. *La transformación de Cayley generalizada establece un difeomorfismo $c_{A^*}\colon S^G(A) \to \Sigma(h_X^G) \cap \Omega_G(A)$.*

5.3.1. Ejemplos

Ejemplo 5.3.2 (Una función altura en un grupo ortogonal)**.** Supongamos $X = I$ y $\mathbb{K} = \mathbb{C}$. Entonces el conjunto de puntos críticos de la función altura h_I son las matrices unitarias $A \in U(n)$ tales que $A^2 = I$. Una matriz de este tipo, hermítica, se puede diagonalizar a una $D = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_n)$ con $\varepsilon_k = \pm 1$.

Por otro lado, se verifica que $\beta_0 \in T_A G$ si y solo si la matriz $A\beta_0$ es antihermítica, $A\beta_0 = -\beta_0^* A$, mientras que $\beta_0 \in S(A)$ si y solo si $A\beta_0 + \beta_0 A = 0$. De donde se deduce que β_0 también tiene que ser hermítica.

Así, por ejemplo, la I y su opuesta, $-I$ son puntos críticos y son aislados ya que $S(\pm I) = 0$. Por otro lado, si tomamos $A = \operatorname{diag}(I_p, -I_q)$, entonces para que β_0 esté en el tangente $T_A G$ debe ser de la forma

$$\beta_0 = \begin{pmatrix} 0 & V^* \\ V & 0 \end{pmatrix},$$

lo que implica que $\dim S(A) = 2pq$. De hecho, ésta es la dimensión de la órbita (crítica) de A que es difeomorfa a la Grassmaniana compleja $U(p+q)/(U(p) \times U(q))$.

Ejemplo 5.3.3 (Una matriz diagonal de rango máximo)**.** Consideremos ahora $X = \operatorname{diag}(q_1, \ldots, q_n)$ una matriz diagonal con $q_k \neq 0$ y supongamos que $|q_1| < \ldots < |q_n|$. En este caso, la condición del gradiente $A^* X^* = XA$ implica que los puntos críticos sean de la forma

$$A = \operatorname{diag}(\pm |q_1|/q_1, \ldots, \pm |q_n|/q_n).$$

Ahora bien, como

$$XA = AX = \operatorname{diag}(\varepsilon_1 |q_1|, \ldots, \varepsilon_n |q_n|), \quad \text{con} \quad \varepsilon_k = \pm 1,$$

tenemos que $S(A) = 0$, es decir, todos los puntos críticos son aislados.

Capítulo 6

Funciones de Bott-Morse en los espacios simétricos

En el Capítulo 3 recogimos cómo calcular los puntos críticos de funciones altura en un grupo de Lie G de tipo ortogonal, unitario o simpléctico; ahora tratamos de analizar qué puntos críticos tienen estas funciones cuando las estudiamos en un espacio simétrico M. La dificultad está en que estos puntos críticos no son simplemente los puntos críticos de la función en el grupo que están en el modelo de Cartan del cociente.

Como vimos en 3.3, T. Frankel aplicó la teoría de Bott-Morse a los grupos de Lie clásicos y a las variedades de Stiefel tomando la representación matricial de estos grupos y usando la aplicación traza como función de Morse [23]. S. Ramanujam obtuvo en [69] una descomposición de Morse de ciertos espacios simétricos G/K usando una técnica similar a la de Frankel. Un poco más tarde, el propio Ramanujam [70] propuso un método alternativo basándose justamente en el modelo de Cartan de modo que identifica cada espacio simétrico con las matrices que verifican determinadas condiciones. Así, muestra que las subvariedades críticas del cociente G/K son, en su caso concreto, la intersección del espacio G/K con las subvariedades críticas de G. Obtiene a partir de éstas últimas los índices de las subvariedades críticas del cociente.

El estudio que proponemos aquí incluye los resultados de Ramanujam y Frankel como casos particulares, pero es mucho más amplio pues se establece para una función altura arbitraria.

Los resultados que veremos en este Capítulo los aplicaremos en el Capítulo 8 para calcular rápidamente la categoría LS del grupo unitario $U(n)$ y de los espacios simétricos compactos de tipo AI $- SU(n)/SO(n)-$ y AII $- SU(2n)/Sp(n)-$, que han sido estudiados recientemente por M. Mimura y K. Sugata en [61]. En el caso del grupo simpléctico $Sp(2)$ proporcionaremos a partir de la teoría de Morse un recubrimiento explícito por cuatro abiertos contráctiles.

6.1. Funciones altura en el modelo de Cartan

Necesitamos ver quién es el tangente al espacio simétrico en cada punto, calcular el gradiente de estas funciones en el cociente y ver cuándo se anula la expresión obtenida. Asimismo, para comprobar qué puntos críticos son no degenerados, es preciso obtener también el operador Hessiano.

6.1.1. Espacio tangente

Recordemos que en un espacio simétrico $G/K \cong M$ el grupo G actúa transitivamente sobre $M \subset N$ mediante la acción $g \cdot m = g\sigma(g)^{-1}$, siendo

$$M = \{h \in G \colon h = g\sigma(g)^{-1}\}$$

la componente conexa que contiene al neutro de

$$N = \{g \in G \colon g = \sigma(g)^{-1}\}.$$

Buscamos una expresión que caracterice las matrices del tangente en un punto arbitrario.

Lema 6.1.1. *Dado M el modelo de Cartan de un espacio simétrico,*

1. *el tangente en el neutro a la variedad M viene dado por*

$$T_I M = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon X + X^* = 0,\ \sigma(X) = -X\};$$

2. *luego, para una matriz cualquiera $A \in M$,*

$$T_A M = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon A^*X + X^*A = 0,\ \sigma(X) = -A^*XA^*\}.$$

Demostración. 1. Como $M = N_e$, $T_I M = T_I N$ y al ser σ lineal, derivando en la construcción de $N = \{X \in G \colon X^{-1} = \sigma(X)\}$, se tiene que

$$T_I N = \{X \in T_I G \colon -X = \sigma(X)\}$$

es decir,

$$T_I N = \{X \in \mathcal{M}_{n\times n}(\mathbb{K}) \colon X + X^* = 0,\ \sigma(X) = -X\}.$$

2. Sea $A \in M$, i.e, $AA^* = I$ y tal que $A = Y\sigma(Y)^{-1}$ para algún $Y \in G$; como G actúa transitivamente en M mediante $l_Y^M(m) = Ym\sigma(Y)^{-1}$ entonces
$$T_A M = (l_Y^M)_{*I}(T_I M).$$
Luego $X \in T_A M$ sii $(l_Y^M)_{*I}^{-1}(X) \in T_I M$. Pero la acción $l_Y^M(m)$ es lineal en m de modo que
$$X \in T_A M \text{ sii } (l_Y^M)^{-1}(X) \in T_I M,$$
es decir,
$$Y^* X \sigma(Y) \in T_I M.$$
De la expresión obtenida para el tangente en el neutro y utilizando que el automorfismo σ es involutivo se deduce la condición que buscábamos.

□

La métrica en el espacio de las matrices $\mathcal{M}_{n\times n}(\mathbb{K})$ nos da un producto escalar en el espacio tangente al grupo $T_g G$. Además, para $U \in G$ las traslaciones en el espacio simétrico inducido son de la forma $l_U^M(m) = Um\sigma(U)^{-1}$.

Lema 6.1.2. *Para un grupo ortogonal $G \subset \mathcal{M}_{n\times n}(\mathbb{K})$ se verifica que*

1. *la métrica inducida es bi-invariante para las traslaciones de G;*
2. *para $U \in G$ las traslaciones l_U^M son isometrías.*

Demostración. 1. Para toda U del grupo
$$\langle UA, UB\rangle = \Re \operatorname{Tr}(A^* U^* U B) = \Re \operatorname{Tr}(A^* B) = \langle A, B\rangle$$
y
$$\langle AU, BU\rangle = \Re \operatorname{Tr}(U^* A^* B U) = \Re \operatorname{Tr}(A^* B U U^*) = \langle A, B\rangle.$$

2. Como $T_m M \subset T_m G$,
$$\begin{aligned}\langle l_U^M(A), l_U^M(B)\rangle &= \langle UA\sigma(U)^{-1}, UB\sigma(U)^{-1}\rangle \\ &= \Re \operatorname{Tr}\left(\sigma(U) A^* U^* U B \sigma(U)^*\right) = \langle A, B\rangle. \quad \square\end{aligned}$$

Lema 6.1.3. *La proyección canónica $p \colon G \mapsto M \subset G$ tal que $p(g) = g\sigma(g)^{-1}$ es una submersión riemanniana.*

6.1.2. Gradiente

Dada la función altura $\widehat{h}_X(A) = \Re \operatorname{Tr}(XA)$ en el espacio vectorial de las matrices $\mathcal{M} = \mathcal{M}_{n\times n}(\mathbb{K})$, su gradiente viene dado por $(\operatorname{grad} \widehat{h}_X)(A) = X^*$. Para obtener el gradiente en un espacio simétrico M basta con proyectarlo al $T_A M$.

Lema 6.1.4. *Dado $Z \in T_I\mathcal{M}$, la proyección de Z en $T_I G$ es $\frac{1}{2}(Z-Z^*)$. La proyección de éste en $\mathfrak{m} = T_I M$ corresponde a la parte de Z antiinvariante por σ,*

$$\frac{1}{4}\left((Z - Z^*) - \sigma(Z - Z^*)\right).$$

Demostración. Podemos descomponer el tangente a $\mathcal{M}$ en el neutro como

$$T_I\mathcal{M} = T_I G \oplus \mathfrak{g}^\perp = \mathfrak{m} \oplus \mathfrak{h} \oplus \mathfrak{g}^\perp.$$

Así, dado $Z \in T_I\mathcal{M}$, su descomposición en parte antihermítica y parte hermítica es

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

Llamemos $X = \frac{1}{2}(Z - Z^*)$ al sumando antihermítico. Entonces, $X = A + B$ donde $A = -A^*, \sigma(A) = -A$ y $B = -B^*, \sigma(B) = B$. Utilizando que σ es de álgebras se obtiene

$$X = \frac{1}{2}\left[(X - \sigma(X)) + (X + \sigma(X)\right],$$

de modo que

$$Z = \frac{1}{2}\left[\frac{1}{2}(Z - Z^*) - \frac{1}{2}\sigma(Z - Z^*)\right] + \frac{1}{2}\left[\frac{1}{2}(Z - Z^*) + \frac{1}{2}\sigma(Z - Z^*)\right] + \frac{1}{2}(Z+Z^*). \quad \square$$

Proposición 6.1.5. *Dada una función altura en el modelo de Cartan M de un espacio simétrico, $h_X \colon M \to \mathbb{R}$, el gradiente en un punto $A \in M$ viene dado por*

$$(\operatorname{grad} h_X^M)(A) = \frac{1}{4}\left(\widehat{X} - A\sigma(\widehat{X})A\right),$$

donde $\widehat{X} = X^* + \sigma(X)$.

Demostración. Por el Lema 6.1.4, tenemos el gradiente en el neutro

$$(\operatorname{grad} h_X^M)(I) = \frac{1}{4}\left[(X^* - X) - \sigma(X^* - X)\right].$$

Es inmediato comprobar que $((\operatorname{grad} h_X^M)(I))^* = -(\operatorname{grad} h_X^M)(I) = \sigma\left((\operatorname{grad} h_X^M)(I)\right)$.

Como las traslaciones son isometrías (véase Lema 6.1.2), para calcular el gradiente en $A = Y\sigma(Y)^{-1}$, basta hacer la traslación por l_Y^M de la parte tangente de

$$l_Y^M\left((\operatorname{grad}\widehat{h}_X)(A)\right) = l_Y^M(X^*),$$

es decir,

$$\begin{aligned} Y\left[\frac{1}{4}(Y^{-1}X^*\sigma(Y) - \sigma(Y)^*XY - \sigma(Y^{-1}X^*\sigma(Y) - \sigma(Y)^*XY))\right]\sigma(Y)^* &= \\ \frac{1}{4}\big[(X^* + \sigma(X)) - A(X + \sigma(X)^*)A\big]. & \qquad \square \end{aligned}$$

Corolario 6.1.6. *$A \in M$ es un punto crítico de h_X^M sii $\widehat{X} = A\sigma(\widehat{X})A$.*

Nótese que $\sigma(\widehat{X})^* = \widehat{X}$, de modo que de la condición del gradiente en grupos de Lie (Lema 3.3.3) se deriva el siguiente teorema.

Teorema 6.1.7. *Dado un espacio simétrico $M \cong G/K$ se verifica que $A \in M$ es un punto crítico de h_X^M sii es un punto crítico de $h^G_{\sigma(\widehat{X})}$.*

6.1.3. Hessiano

A partir de la expresión que hemos obtenido para el gradiente podemos obtener una expresión explícita del Hessiano de una función altura en un espacio simétrico.

Proposición 6.1.8. *Dada una función altura h_X^M su Hessiano $(Hh_X^M)_A \colon T_AM \to T_AM$ en un punto $A \in M$, viene dado por*

$$H(h_X^M)_A(U) = -\frac{1}{4}\left(A\sigma(\widehat{X})U + U\sigma(\widehat{X})A\right),$$

donde $\widehat{X} = X^* + \sigma(X)$.

Demostración.

$$\begin{aligned} \lim_{t\to 0}\frac{1}{t}\left[(\operatorname{grad} h_X^M)(A+tU) - (\operatorname{grad} h_X M)(A)\right] &= \\ \lim_{t\to 0}\frac{1}{4t}\left[-A\sigma(\widehat{X})A - A\sigma(\widehat{X})tU - tU\sigma(\widehat{X})A - tU\sigma(\widehat{X})tU + A\sigma(\widehat{X})A\right] &= \\ \lim_{t\to 0}\frac{1}{4}\left[-A\sigma(\widehat{X})U - U\sigma(\widehat{X})A - tU\sigma(\widehat{X})U\right]. \end{aligned}$$

Comprobemos que la expresión obtenida efectivamente está en el T_AM:

$$\begin{aligned}\left[A\sigma(\widehat{X})U + U\sigma(\widehat{X})A\right] A^* + A\left[U^*\widehat{X}A^* + A^*\widehat{X}U^*\right] &= \\ A\sigma(\widehat{X})UA^* + U\sigma(\widehat{X}) + AU^*\widehat{X}A^* + \widehat{X}U^* &= 0\end{aligned}$$

ya que, como $U \in T_AM$ y A es un punto crítico,

$$AU^*\widehat{X}A^* = -U\sigma(\widehat{X}) \quad \text{y} \quad A\sigma(\widehat{X})UA^* = -\widehat{X}U^*.$$

Además, utilizando que $\sigma(A) = A^*$ y que A es un punto crítico es inmediato comprobar que

$$\sigma(H(h_X^M)_A(U)) = -A^*(H(h_X^M)_A(U))A^*. \quad \square$$

6.1.4. Linealización de la ecuación diferencial del gradiente

Análogamente al caso de los grupos de Lie, la transformación de Cayley generalizada nos permite "linealizar" el flujo del gradiente de las funciones altura en espacios simétricos. Este resultado generaliza el de Volchenko y Kozachko que recogimos en la Subsección 3.3.4, véase también [84].

Teorema 6.1.9. *Sea h_X^M una función altura arbitraria en un espacio simétrico M y A un punto crítico. La solución de la ecuación del gradiente*

$$4\alpha' = \widehat{X} - \alpha\sigma(\widehat{X})\alpha,$$

con condición inicial $\alpha(0) \in \Omega_M(A)$, es la imagen por c_A^* de la curva

$$\beta(t) = \exp(-\frac{t}{4}A^*\widehat{X})\beta_0 \exp(-\frac{t}{4}\widehat{X}A^*),$$

donde $\widehat{X} = X^* + \sigma(X)$ y $\beta_0 = c_A(\alpha(0)) \in T_{A^*}M$.

Demostración. En primer lugar, veamos que $\beta(t) \in T_{A^*}M$. Nótese que por ser A una matriz normal y un punto crítico, las matrices $A^*\widehat{X} = \widehat{X}A^*$ son hermíticas.

Utilizando que para una matriz A unitaria o simpléctica se verifica que $A\exp(XA) = \exp(AX)A$ tenemos que

$$\begin{aligned}\beta(t)A + A^*\beta(t) &= \\ \left(\exp(-\frac{t}{4}A^*\widehat{X})\beta_0 \exp(-\frac{t}{4}\widehat{X}A^*)\right) A + \left(\exp(-\frac{t}{4}A^*\widehat{X})A^*\beta_0^* \exp(-\frac{t}{4}A^*\widehat{X})\right)\end{aligned}$$

pero como β_0 está en $T_{A^*}M$, esto es igual a

$$\left(\exp(-\frac{t}{4}A^*\widehat{X})\beta_0\exp(-\frac{t}{4}\widehat{X}A^*)\right)A+\left(\exp(-\frac{t}{4}A^*\widehat{X})(-\beta_0 A)\exp(-\frac{t}{4}A^*\widehat{X})\right) = \exp(-\frac{t}{4}A^*\widehat{X})\beta_0\left[\exp(-\frac{t}{4}\widehat{X}A^*)A-A\exp(-\frac{t}{4}A^*\widehat{X})\right]$$

que se anula utilizando de nuevo que $A\exp(XA)=\exp(AX)A$.

Además, como σ es de grupos, $\sigma(\beta(t))$ es

$$\sigma\left(\exp(-\frac{t}{4}A^*\widehat{X})\right)\sigma(\beta_0)\sigma\left(\exp(-\frac{t}{4}\widehat{X}A^*)\right) \tag{6.1}$$

y utilizando de nuevo la linealidad de σ y que β_0 está en $T_{A^*}M$ y por tanto, $\sigma(\beta_0)=-A\beta_0 A$, tenemos que (6.1) es igual a

$$\exp\left(-\frac{t}{4}\sigma(A^*\widehat{X})\right)(-A\beta_0 A)\exp\left(-\frac{t}{4}\sigma(\widehat{X}A^*)\right) = -\exp\left(-\frac{t}{4}\sigma(A^*)\sigma(\widehat{X})\right)A\beta_0 A\exp\left(-\frac{t}{4}\sigma(\widehat{X})\sigma(A^*)\right).$$

Pero como $A\in M\subset N$, $\sigma(A)=A^{-1}=A^*$ es igual a

$$-\exp\left(-\frac{t}{4}A(\widehat{X})^*\right)A\beta_0 A\exp\left(-\frac{t}{4}(\widehat{X})^*A\right)$$

lo que, aplicando que $A\exp(XA)=\exp(AX)A$ y que $(\widehat{X})^*A$ y $A(\widehat{X})^*$ son matrices hermíticas, equivale a

$$\begin{aligned}-A\exp\left(-\frac{t}{4}(\widehat{X})^*A\right)\beta_0\exp\left(-\frac{t}{4}A(\widehat{X})^*\right)A &= \\ -A\left[\exp\left(-\frac{t}{4}A^*\widehat{X}\right)\beta_0\exp\left(-\frac{t}{4}\widehat{X}A^*\right)\right] &= \\ -A\beta(t)A.\end{aligned}$$

Ahora tiene sentido comprobar que $c_{A^*}(\beta)$ es solución de la ecuación diferencial. Derivando en la definición de beta tenemos que

$$\beta'=(-1/4)\left(\beta\widehat{X}A^*+A^*\widehat{X}\beta\right).$$

Sea

$$\alpha = c_{A^*} \circ \beta = (I - A\beta)(A^* + \beta)^{-1},$$

entonces

$$\alpha(A^* + \beta) = I - A\beta.$$

Derivando

$$\alpha'(A^* + \beta) + \alpha\beta' = -A\beta',$$

es decir,

$$\alpha'(A^* + \beta) = -(A + \alpha)\beta'.$$

Por la Proposición 4.2.2, la inversa de $A^* + \beta$ es $\frac{1}{2}(A + c_{A^*}(\beta)) = \frac{1}{2}(A + \alpha)$, entonces

$$\begin{aligned} \alpha' &= \\ -\frac{1}{2}(A+\alpha)\beta'(A+\alpha) &= \\ \frac{1}{8}(A+\alpha)(\beta\widehat{X}A^* + A^*\widehat{X}\beta)(A+\alpha). \end{aligned} \tag{6.2}$$

Ahora bien,

$$\beta = c_A(\alpha) = (I - A^*\alpha)(A + \alpha)^{-1} = (A + \alpha)^{-1}(I - \alpha A^*)$$

implica que

$$\beta(A + \alpha) = I - A^*\alpha$$

y

$$(A + \alpha)\beta = I - \alpha A^*.$$

Así que por la Ecuación (6.2) tenemos que

$$\begin{aligned} 8\alpha' &= \\ (A+\alpha)\beta\widehat{X}A^*(A+\alpha) + (A+\alpha)A^*\widehat{X}\beta(A+\alpha) &= \\ (I-\alpha A^*)\widehat{X}A^*(A+\alpha) + (A+\alpha)A^*\widehat{X}(I-A^*\alpha) &= \\ 2(\widehat{X} - \alpha A^*\widehat{X}A^*\alpha) &= \\ 2(\widehat{X} - \alpha\sigma(\widehat{X})\alpha). \quad \square \end{aligned}$$

Nota–. De hecho, cuando $X = A = I$, $\beta(t) = \exp(-t)\beta_0$ es la contracción radial a β_0.

6.2. Estructura local del conjunto de puntos críticos

Otra de las aplicaciones de la transformación de Cayley generalizada es que nos permite dar una carta local para el conjunto de puntos críticos de una función de Bott-Morse.

Sea $h_X^M(A) = \Re\,\mathrm{Tr}(XA)$ una función altura arbitraria en el espacio simétrico M. Dado $A \in \Sigma(h_X^M)$ un punto crítico denotamos por $S^M(A)$ al siguiente espacio vectorial

$$S^M(A) = \{\beta_0 \in T_{A^*}M \colon A^*\widehat{X}\beta_0 + \beta_0\widehat{X}A^* = 0\}.$$

Teorema 6.2.1. *La transformación de Cayley generalizada establece un difeomorfismo, $c_{A^*}\colon S^M(A) \to \Sigma(h_X^M) \cap \Omega_M(A)$.*

Demostración. Veamos que la curva $\beta(t)$ del Teorema 6.1.9 es constante si y solo si $\beta_0 \in S^M(A)$. En efecto, derivando en la definición de β,

$$\beta' = -\frac{1}{4}(A^*\widehat{X}\beta + \beta\widehat{X}A^*).$$

Entonces, de nuevo por el Teorema 6.1.9, $c_{A^*}(\beta_0)$ es un punto crítico si y solo si $\beta'(t) = 0$ para todo t, o equivalentemente,

$$A^*\widehat{X}\beta_0 + \beta_0\widehat{X}A^* = 0. \qquad \square$$

Nótese que el espacio vectorial $S^M(A)$ es justamente el núcleo del operador Hessiano $H(h_X^M)_A$. Es decir, una función h_X^M será de Morse si todos sus puntos críticos A verifican que $\dim S^M(A) = 0$.

6.2.1. Ejemplos

Con las técnicas desarrolladas en este Capítulo es sencillo obtener los resultados de Ramanujam relativos a los puntos críticos de la función altura h_I^M. Recogemos dos de los casos que estudia este autor como ejemplo.

Ejemplo 6.2.2. La función altura h_I^M en el espacio simétrico $SU(n)/SO(n)$. Su modelo de Cartan es (difeomorfo a) $M = \{X \in SU(n)\colon X = X^T\}$. En primer lugar, identificamos los tangentes,

$$\begin{aligned} T_IM &= \{Y \in \mathcal{M}_{n\times n}(\mathbb{C})\colon Y + Y^* = 0, \sigma(Y) = -Y\}, \\ T_AM &= \{Y \in \mathcal{M}_{n\times n}(\mathbb{C})\colon A^*Y + Y^*A = 0, \sigma(Y) = -A * YA^*\}. \end{aligned}$$

Como en este caso el automorfismo σ es la conjugación, el T_IM está formado por las matrices complejas simétricas con todos los elementos complejos puros y T_AM por las Y simétricas complejas tales que $\overline{Y} = -A^*YA^*$.

Como

$$(\operatorname{grad} h_I^M)(A) = \frac{1}{2}(I - A^2),$$

el conjunto de puntos críticos está formado por las matrices involutivas de M, es decir, por las matrices ortogonales simétricas. Además, dado $A \in \Sigma(h_I^M)$, para $U \in T_AM$,

$$(Hh_I^M)_A(U) = -\frac{1}{2}(AU + UA),$$

de modo que todos los puntos críticos son degenerados ya que el núcleo del operador hessiano está formado por todas las matrices reales simétricas del T_AM.

Ejemplo 6.2.3. El modelo de Cartan del espacio simétrico $U(2n)/\,Sp(n)$ es difeomorfo al conjunto de las matrices antisimétricas de $U(2n)$. El espacio tangente en el neutro viene dado por

$$T_IM = \{Y \in \mathcal{M}_{2n\times 2n}(\mathbb{C})\colon Y + Y^* = 0, J\overline{Y} = -YJ\}$$

y T_AM es el trasladado correspondiente.

Consideramos de nuevo la función altura h_I^M y tenemos que

$$\begin{aligned}(\operatorname{grad} h_I^M)(A) &= \frac{1}{2}(I - A^2),\\ (Hh_I^M)_A(U) &= -\frac{1}{2}(AU + UA).\end{aligned}$$

De modo que los puntos críticos, todos ellos degenerados, son las matrices unitarias, antisimétricas con todos los elementos complejos puros. Todas las matrices hermíticas están en el núcleo del hessiano

Ejemplo. Sea la esfera $U(2)/(U(1)\times U(1)) \cong S^2 \subset U(2) \cong S^3 \times S^1$. Su modelo de Cartan está formado por los puntos $(s,z) \in \mathbb{R}\times\mathbb{C}$ con $s^2 + |z|^2 = 1$.

Si tomamos la función altura h_X con $X = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ entonces $\hat{X} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ y los puntos críticos son los dos polos. Lo mismo ocurre cuando $X = I$ es la matriz identidad o cuando $X = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

6.2.2. Relación entre los flujos de los gradientes en el espacio simétrico y en el grupo

A partir del caso $X = I$ podría parecer que los puntos críticos de una función altura h_X^M en el modelo de Cartan M de un espacio simétrico no son más que los puntos críticos de esa misma función en el grupo que estén en M. Sin embargo, los teoremas que acabamos de probar nos muestran que no es así. El siguiente ejemplo lo ilustra.

Ejemplo 6.2.4. Consideramos la función altura h_X con $X = (\mathbf{i}+\mathbf{j}+\mathbf{k})I$ en el espacio simétrico $Sp(1)/U(1)$. El automorfismo es $\sigma(X) = -\mathbf{i}X\mathbf{i}$, de donde se obtiene que el modelo de Cartan M está formado por los cuaternios unitarios q' tales que $q'\mathbf{i} = q\mathbf{i}\overline{q}$ para algún $q \in \mathbb{H}$, $|q| = 1$. Se puede ver que los elementos de M siempre verifican que $\Re(q'\mathbf{i}) = 0$

1. Si estudiamos la función altura en el grupo, por el Lema 3.3.3 tenemos que los puntos críticos de h_X^G son los elementos $q \in Sp(1)$ tales que
$$(\mathbf{i}+\mathbf{j}+\mathbf{k})^* = q(\mathbf{i}+\mathbf{j}+\mathbf{k})q$$
o, equivalentemente
$$-(\mathbf{i}+\mathbf{j}+\mathbf{k})\overline{q} = q(\mathbf{i}+\mathbf{j}+\mathbf{k}).$$
De esta ecuación se deduce que $q = t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ es crítico para h_X^G si y solo si
$$\begin{aligned} t &= 0 \\ x &= y = z. \end{aligned}$$
Pero como q es unitario se tiene que los únicos puntos críticos en $G = Sp(1)$ son
$$\Sigma(h_X^G) = \{\pm\frac{1}{\sqrt{3}}(\mathbf{i}+\mathbf{j}+\mathbf{k})\}.$$
Estos dos puntos no están en M porque $\Re\left(\pm\frac{1}{\sqrt{3}}(\mathbf{i}+\mathbf{j}+\mathbf{k})\mathbf{i}\right) = \mp\frac{1}{\sqrt{3}}$. Sin embargo, veamos que sí hay puntos de M que son críticos para la función altura restringida al modelo de Cartan.

2. Pasamos entonces a calcular los puntos críticos de esta misma función altura en el modelo de Cartan de $Sp(1)/U(1)$. Esta vez, según la Proposición 6.1.5, la condición para que un punto q' en M sea crítico para h_X^M es que $\widehat{X} = q\sigma(\widehat{X})q'$ donde $\widehat{X} = X^* + \sigma(X)$. Para esta elección de X, tenemos que $\widehat{X} = -2(\mathbf{j}+\mathbf{k})$ y

$\sigma(\widehat{X}) = 2(\mathbf{j}+\mathbf{k})$. Así, utilizando que los puntos q' de M verifican que $\Re(q'\mathbf{i}) = 0$, de la condición

$$-2\overline{q'}(\mathbf{j}+\mathbf{k}) = 2(\mathbf{j}+\mathbf{k})q'$$

se obtiene que

$$\Sigma(h_X^M) = \{\pm\frac{1}{\sqrt{2}}(\mathbf{j}+\mathbf{k})\}.$$

Nota–. Según Ramanujam [69], el flujo del gradiente de la función altura h_X^M que pasa por un punto $A \in M$ se mantiene dentro de M. Esto es cierto porque este autor estudia únicamente el caso de Frankel, la función altura con respecto a la identidad h_I. En [84] Dynnikov y Veselov afirman que los espacios simétricos $U(2n)/\,Sp(n)$, $Sp(n)/U(n)$ y $U(n)/O(n)$ permanecen invariantes por el flujo del gradiente de las funciones altura en los grupos de Lie correspondientes.

Sin embargo, por el Teorema 6.1.7, esto únicamente podemos asegurarlo cuando el automorfismo σ es tal que $X = \sigma(\widehat{X})$, caso en el que podría reformularse 6.1.7 como sigue:

Teorema 6.2.5. *Dado un espacio simétrico $M \cong G/K$ y una $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que $\sigma(\widehat{X}) = X$ entonces, los puntos críticos de la función altura h_X^M restringida a M verifican que*

$$\Sigma(h_X^M) = \Sigma(h_X^G) \cap M.$$

Para las funciones altura que escogen Dynnikov y Veselov la matrix X es siempre una matriz diagonal real, pero en general el resultado no es cierto.

Para los espacios simétricos tales que el automorfismo σ verifica que $\sigma(D) = D$ para cualquier D diagonal real, tenemos el siguiente resultado.

Lema 6.2.6. *Un punto $A \in M$ es un punto crítico de h_X^M si y solo si U^*AV es un punto crítico de h_D^M, donde D es la matriz diagonal de la descomposición SVD de $\widehat{X}$, $\widehat{X} = UDV^*$. Además,*

$$U^*(\operatorname{grad} h_X^M)(A)V = \frac{1}{2}((\operatorname{grad} h_D^M)(U^*AV))$$

Demostración. Por la Proposición 6.1.5 tenemos que

$$(\operatorname{grad} h_X^M)(A) = \frac{1}{4}\left(\widehat{X} - A\sigma(\widehat{X})A\right),$$

donde $\widehat{X} = X^* + \sigma(X)$. Luego, como $\sigma(\widehat{X}) = \widehat{X}^*$,

$$(\operatorname{grad} h_X^M)(A) = \frac{1}{4}(UDV^* - AVDU^*A),$$

es decir,

$$U^*(\operatorname{grad} h_X^M)(A)V = \frac{1}{4}(D - U^*AVDU^*AV).$$

Por otro lado, como nos hemos restringido a los casos en los que $\sigma(D) = D$, tenemos que $\widehat{D} = 2D$. Así el gradiente para la función altura con respecto a la matriz diagonal D es precisamente $(\operatorname{grad} h_D^M)(B) = \frac{1}{2}(D - BDB)$. $\square$

Esto se verifica para todos los espacios simétricos de tipo clásico recogidos en la Tabla de la Subsección 2.3.1 excepto para $U(2n)/\,Sp(n)$ y $SO(2n)/U(n)$.

A la vista del Lema 6.2.6 podría pensarse que en orden a obtener una caracterización análoga a la del Teorema 5.2.2 bastaría ahora estudiar la dimensión del subespacio vectorial $S^M(A)$ para la función altura h_D^M . Sin embargo, en el ámbito de los espacios simétricos esto carece de sentido pues en general, el T_AM y el T_{U^*AV} no guardan relación (únicamente podrían relacionarse si las matrices U y V de la descomposición SVD de $\widehat{X}$ y el automorfismo σ verificasen que $V = \widehat{U}$). Por tanto, en principio, el caracter degenerado o no de los puntos críticos de h_X^M no tiene que ver con el de los de la función h_D^M.

Si buscamos una caracterización análoga a la que obtuvimos en el ámbito de los grupos de Lie (ver Teorema 5.2.1), la dificultad mayor radica en que la descomposición polar no es compatible con la estructura del modelo de Cartan. De nuevo, esto nos impide reducirnos a la funciones altura del tipo h_D^M con D diagonal. Pensamos entonces que sería interesante investigar si es posible obtener una descomposición polar generalizada del tipo de la que proponen autores como H.Z. Munthe-Kaas *et al.* [63] o J.D. Lawson [56] en el ámbito de los espacios simétricos. Si pudiésemos construir tal descomposición, podría permitirnos reducir una matriz X de M a una diagonal sin salirnos del modelo de Cartan. De hecho, la descomposición polar generalizada en un grupo de Lie guarda relación estrecha con el automorfismo involutivo y la descomposición que éste induce en el grupo. En el ámbito de los grupos de Lie, esta descomposición permite expresar un elemento del grupo como producto de un factor en un espacio simétrico por otro en un subgrupo de Lie del grupo dado. En el artículo de Lawson citado este autor caracteriza los grupos de Lie en los que se tiene la existencia y unicidad de dicha descomposición. Los trabajos de Munthe-Kaas *et al.* proponen una prueba alternativa para la existencia (y unicidad) local de la descomposición polar y obtiene explícitamente series en el álgebra de Lie que determinan de manera única los dos factores de la descomposición para matrices cercanas a la identidad.

Capítulo 7

Autovalores cuaterniónicos por la izquierda

En este capítulo desarrollamos los resultados que hemos obtenido en el ámbito de los autovalores por la izquierda de matrices cuaterniónicas. En particular, proporcionamos una nueva demostración del resultado de Huang y So (Teor. 1.3.6) así como una caracterización de las matrices 2×2 simplécticas cuyo espectro por la izquierda es infinito. Este resultado nos permitirá discutir al final de la Memoria, en la Sección 8.8, la posibilidad de recubrir el grupo $Sp(2)$ por abiertos asociados a autovalores.

Los estudios conocidos hasta el momento para autovalores por la izquierda son de naturaleza algebraica y difíciles de generalizar para matrices de orden $n > 3$. Proponemos aquí una noción de función característica μ para las matrices cuaterniónicas tal que las raíces de la ecuación $\mu(\lambda) = 0$ son los autovalores por la izquierda de la matriz. En general, esta μ será una función racional y nos permitirá realizar un estudio topológico del espectro por la izquierda de las matrices de órdenes dos y tres calculando su diferencial y utilizando la teoría del grado. Aparentemente, este nuevo método es más fácil de generalizar para las matrices de orden $n > 3$.

Un tema que ha surgido de manera natural al estudiar esta función característica, aunque no tiene relación directa con la categoría LS, es la posibilidad de extender el teorema de Cayley-Hamilton. Comprobamos que para matrices de órdenes dos y tres esta función característica lo verifica, es decir la extensión natural de μ a las matrices cumple que $\mu_M(M) = 0$. Es de esperar que este resultado sea cierto para cualquier orden mayor que tres pero no hemos podido comprobarlo ya que ninguna de las demostraciones de este teorema en el ámbito conmutativo parecen poder aplicarse en el caso cuaterniónico.

7.1. Autovalores por la izquierda

Los autovalores por la izquierda de una matriz cuaterniónica están muy poco estudiados. Es obvio que no pueden extenderse de manera sencilla los métodos propios del ámbito conmutativo y hasta hoy no hay un método sistemático que permita acometer su estudio. Como afirmaba T. Suzuki en el año 2008 [82],

> "para una matriz con entradas no conmutativas, no están definidos el determinante o la función característica por el problema del orden. Por lo tanto, los "autovalores" usados en la descomposición espectral no están definidos y, hasta ahora, no tenemos un método sistemático para calcular una función de una matriz con entradas no conmutativas."

De la definición de autovalor por la izquierda (Def. 1.3.1) se deriva que $\lambda \in \sigma_l(M)$ si y sólo si $M - \lambda I$ no es inversible. La idea que aquí proponemos es utilizar el determinante de Study (ver Sección 1.2) para estudiar el espectro por la izquierda.

En principio, la ecuación $\mathrm{Sdet}(M - \lambda I) = 0$ no tiene por qué ser un polinomio y en general no se sabe resolver. En [7, 33] pueden verse dos métodos diferentes para resolverla en el caso 2×2. No está claro que todos los autovalores de una matriz cuaterniónica de orden $n \times n$ puedan hallarse resolviendo polinomios cuaterniónicos de grado menor o igual que n. Por ahora, sólo se ha podido comprobar que hasta orden 3 sí es posible pues en 2005 W. So obtuvo fórmulas explícitas para orden 3 recogidas en [79]. Como ya hemos visto, al margen de estos cálculos específicos, en el año 1985, R.M.W. Wood [90] probó, usando métodos homotópicos, que cualquier matriz cuaterniónica tiene al menos un autovalor por la izquierda.

7.1.1. Autovalores de las matrices 2×2

Para $n = 2$, L. Huang y W. So dieron en [33] una caracterización de las matrices con infinitos autovalores por la izquierda. Su resultado (ver Teorema 1.3.6) está basado en una serie de fórmulas explícitas para resolver algunas ecuaciones cuadráticas obtenidas previamente por los mismos autores (ver Teorema 1.4.2) [34]. Más tarde, De Leo *et al.* propusieron en [54] un método alternativo de resolver ecuaciones polinómicas unilaterales que reduce el problema a encontrar los autovalores por la derecha de una matriz asociada a la ecuación (matriz compañera). Proponemos aquí un modo nuevo de probar los resultados de Huang y So basándonos en el método de De Leo *et al.*

A lo largo de toda la sección A denota una matriz cuaterniónica 2×2,

$$A = \begin{pmatrix} a & b \\ c & b \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{H})$$

y a_1, a_0 los coeficientes $a_1 = b^{-1}(a-d)$ y $a_0 = -b^{-1}c$.

Nueva demostración del teorema de Huang y So

El Teorema 1.3.6 de Huang y So caracteriza completamente las matrices de orden 2×2 con infinitos autovalores. La prueba original consiste en ir estudiando qué ocurre en cada caso utilizando las fórmulas que habían obtenido en [33] para resolver ecuaciones del tipo (7.1). En particular probaron que, en el caso de que haya infinitos autovalores, éstos vienen dados por la expresión $(a+d+b\xi)/2$, donde ξ es un cuaternio sin parte real, $\xi \in \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$, tal que el cuadrado del módulo es $|\xi|^2 = |\Delta|$ con $\Delta = a_1^2 - 4a_0$.

Es fácil ver que las condiciones del Teorema 1.3.6 son suficientes. De hecho, si $a_0 = s$, $a_1 = t$, $s, t \in \mathbb{R}$, entonces, la definición de autovalor nos da la expresión $\lambda^2 + t\lambda + s = 0$, que, después de hacer el cambio $p = \lambda + t/2$, nos da la ecuación $p^2 = \Delta/4 < 0$ que tiene infinitas soluciones del tipo

$$p = (\sqrt{-\Delta}/2)\omega, \quad \omega \in \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle, |\omega| = 1.$$

En cuanto a la necesidad de las condiciones daremos una prueba alternativa. La construimos basándonos en un método elegante para resolver ecuaciones, propuesto por De Leo *et al.* en [54], como mejora de un algoritmo previo de Serôdio *et al.* [75]. Veamos explícitamente en qué consiste este algoritmo.

Algoritmo de De Leo

Sea

$$M = \begin{pmatrix} -a_1 & -a_0 \\ 1 & 0 \end{pmatrix}$$

la *matriz compañera* de la ecuación

$$p^2 + a_1 p + a_0 = 0, \tag{7.1}$$

$$M \begin{pmatrix} p \\ 1 \end{pmatrix} = \begin{pmatrix} -a_1 p - a_0 \\ p \end{pmatrix} = \begin{pmatrix} p^2 \\ p \end{pmatrix} = \begin{pmatrix} p \\ 1 \end{pmatrix} p,$$

lo cual nos muestra que, si queremos encontrar soluciones de (7.1) debemos buscar autovalores *por la derecha* p de M correspondientes a autovectores de la forma $(p, 1)$. Siguiendo a [54] llamaremos a p autovalor por la derecha *especial*.

Luego, para resolver la ecuación (7.1) necesitamos, en primer lugar, encontrar los autovalores por la derecha de la matriz compañera M. Según la Proposición 1.1.8, éstos se corresponden con la clase de similitud de los autovalores de su forma compleja 4×4, $c(M)$, y se pueden calcular resolviendo la ecuación característica $\det(c(M) - qI) = 0$. Como ya habíamos visto, debido a la forma de $c(M)$ sus autovalores aparecen en pares $q_1, \overline{q_1}, q_2, \overline{q_2}$.

Autovectores por la derecha.

Para calcular los autovectores, consideremos el $\mathbb{C}$-isomorfismo

$$(z', z) \in \mathbb{C}^2 \mapsto z' + \mathbf{j}z \in \mathbb{H}. \tag{7.2}$$

Proposición 7.1.1. *$(u', u, v', v) \in \mathbb{C}^4$ es un z-autovector de la forma compleja $c(M)$ si y solo si $(u' + \mathbf{j}v', u + \mathbf{j}v)$ es z-autovector por la derecha de M.*

Soluciones de la ecuación.

Pasemos entonces a calcular las soluciones de (7.1). Una vez que hayamos encontrado un autovalor complejo q de la matriz compañera M y algún q-autovector (u', u) observamos que, debido a la forma específica de M, $u' = uq$. Por tanto, por la Proposición 1.1.7 el vector

$$\begin{pmatrix} u' \\ u \end{pmatrix} u^{-1} = \begin{pmatrix} uqu^{-1} \\ 1 \end{pmatrix}$$

es un uqu^{-1}-autovector, esto es $p = uqu^{-1}$ es un autovalor especial de la clase de similitud de $[q]$ y por tanto, es la solución deseada.

Nótese que, por la Proposición 1.1.7, dos autovectores $\mathbb{H}$-linealmente dependientes dan lugar al mismo autovalor especial

Ejemplo 7.1.2. Tomamos $A = \begin{pmatrix} \mathbf{i}+\mathbf{j} & \mathbf{j} \\ -1-\mathbf{j} & \mathbf{i}+2\mathbf{j} \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{H})$. La matriz compañera del polinomio $p^2 + a_1 p + a_0$ es $M = \begin{pmatrix} 1 & -1+\mathbf{j} \\ 1 & 0 \end{pmatrix}$ cuya forma compleja viene dada por

$$c(M) = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Los autovalores de $c(M)$ son $\{1+\mathbf{i}, 1-\mathbf{i}, \mathbf{i}, -\mathbf{i}\}$ con autovectores asociados los subespacios de $\mathbb{C}^4$ generados por $(-1+\mathbf{i}, \mathbf{i}, 1+\mathbf{i}, 1), (-1-\mathbf{i}, -\mathbf{i}, 1-\mathbf{i}, 1), (1, -\mathbf{i}, \mathbf{i}, 1)$ y $(1, \mathbf{i}, -\mathbf{i}, 1)$ respectivamente. Entonces, los autovalores por la derecha de M son todos los cuaternios de la clase de similitud de $q_1 = 1+\mathbf{i}$ y de $q_2 = \mathbf{i}$.

Tenemos pues que $q_1 = 1+\mathbf{i}$ es un autovalor por la derecha de M con autovector asociado $(-1+\mathbf{i}+\mathbf{j}-\mathbf{k}, \mathbf{i}+\mathbf{j})$ (Prop. 7.1.1), luego $p_1 = 1+\mathbf{j}$ es un autovalor especial de la matriz compañera M. Análogamente, $(1-\mathbf{k}, -\mathbf{i}+\mathbf{j})$ es un $\mathbf{i}$-autovector por la derecha, a partir del cual obtenemos el autovalor especial $p_2 = -\mathbf{j}$.

Es decir, $p_1 = 1+\mathbf{j}$ y $p_2 = -\mathbf{j}$ son las soluciones de la ecuación $p^2 - p + (1+\mathbf{j}) = 0$ y por tanto, $\lambda_1 = -1+\mathbf{i}+2\mathbf{j}$ y $\lambda_2 = 1+\mathbf{i}+\mathbf{j}$ son los autovalores por la izquierda de la matriz A.

Número de autovalores de las matrices 2×2.

Estamos ahora en condiciones de discutir el número de soluciones de la ecuación (7.1). Esto es lo que nos dará una nueva prueba del Teorema 1.3.6. Daremos una forma topológica de ver este resultado en la Subsección 7.3.3.

Cada uno de los autovalores q de la forma compleja $c(M) \in \mathcal{M}_{4\times 4}(\mathbb{C})$ tiene asociado un espacio vectorial complejo de autovalores al que denotaremos por $V(q) \subset \mathbb{C}^4$. Sabiendo que sus dimensiones complejas suman 4 podemos ver cuáles son las posibles dimensiones de $V(q_k)$ y $V(\overline{q_k})$, $1 \leq k \leq 2$.

1. Si los cuatro autovalores $q_1, \overline{q_1}, q_2, \overline{q_2}$ son diferentes, entonces cada $V(q_k)$ tiene dimensión 1. Es el caso en el que nos dan dos complejos diferentes y sus conjugados. Los espacios de autovectores $V(q_1), V(q_2)$ son rectas y nos dan dos autovalores especiales p_1 y p_2. Las rectas de sus conjugados nos darán los mismos autovalores especiales, por tanto, M tiene exactamente dos autovalores especiales, es decir, hay exactamente dos soluciones.

2. Si algún autovalor es real, por ejemplo $q_1 \in \mathbb{R}$, entonces los únicos cuaternios similares vq_1v^{-1} son él mismo y $p_1 = q_1$, independientemente de la dimensión de $V(q_1)$. Entonces habrá una o dos soluciones dependiendo de que $q_1 = q_2$ o no.

3. El único caso en el que pueden aparecer infinitos autovalores especiales es cuando los autovalores son iguales y complejos no reales, $q_1 = q_2 \notin \mathbb{R}$, lo que lleva consigo que $\dim_{\mathbb{C}} V(q_1) = 2 = \dim_{\mathbb{C}} V(\overline{q_1})$.

Nos fijamos entonces en el tercer caso, el único en el que M puede tener infinitos autovalores por la izquierda, cuando $c(M)$ sólo tiene como autovalores un complejo y su conjugado $q_1, \overline{q_1}$. La siguiente proposición prueba que en este caso tenemos realmente un número infinito de soluciones (nótese que $q_1 \notin \mathbb{R}$).

Proposición 7.1.3. *En el tercer caso todos los cuaternios similares a q_1 son autovalores especiales por la derecha de M.*

Demostración. Tomemos una $\mathbb{C}$-base $\tilde{u}, \tilde{v}$ del espacio de autovectores $V(q_1) \subset \mathbb{C}^4$ y los correspondientes vectores v, w en $\mathbb{H}^2$ por el ismorfismo (7.2). Como éstos últimos son de la forma (uq, q), se tiene que las segundas coordenadas v_2, w_2 de v y w son $\mathbb{C}$-independientes en $\mathbb{H}$ y por tanto, una $\mathbb{C}$-base. Esto significa que los autovalores especiales $p = uq_1u^{-1}$, donde u es una $\mathbb{C}$-combinación lineal de v_2 y w_2, recorren todos los posibles cuaternios de la clase de similitud de q_1. $\square$

Las condiciones de Huang y So.

Queda ahora comprobar que en el caso 3 se verifican las condiciones del Teorema 1.3.6 de Huang y So.

Sea $q_1 = x+\mathbf{i}y$, $y \neq 0$, uno de los dos autovalores complejos de $c(M)$ y sea $p \in [q_1]$ cualquier autovalor especial de M. Como $\Re(p) = \Re(q_1)$ y $|p| = |q_1|$, podemos escribir $p = x + |y|\omega$, donde $\omega \in \mathbb{H}_0 = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$ con $|\omega| = 1$.

Pongamos $a_1 = t + \xi_1$, con $t \in \mathbb{R}$ y $\xi_1 \in \mathbb{H}_0$. Si escribimos la ecuación (7.1) de la forma $a_0 = -(p + a_1)p$ podemos deducir que

$$\Re(a_0) = xt + x^2 - |y|^2 + |y|\Re(\xi_1\omega).$$

Por tanto, $|y|\langle \xi_1, \omega \rangle$ no depende de ω. Como $y \neq 0$, el siguiente Lema 7.1.4 afirma que $\xi_1 = 0$, *i.e.* $a_1 \in \mathbb{R}$.

Lema 7.1.4. *Sea $\xi \in \mathbb{H}_0$ verificando que para cualquier par de vectores $\omega, \omega' \in \Omega$, se anula el producto $\langle \xi, \omega - \omega' \rangle = 0$. Entonces $\xi = 0$.*

Demostración. Sea $\xi = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \neq 0$. Podemos suponer que $x \neq 0$ (los otros casos son análogos). Tomemos cualquier $\omega \in \Omega$ ortogonal a ξ y $\omega' = \mathbf{i}$. Entonces $\langle \xi, \omega - \omega' \rangle = x \neq 0$. $\square$

Una consecuencia de la Proposición 7.1.3 es que $p = q_1$ es un autovalor especial de M, de manera que, por la ecuación (7.1), a_0 es un número complejo. Como $\overline{q_1}$ es también una solución, deducimos que $a_0 = \overline{a_0}$ es un número real. Finalmente, de que $q_1^2 + a_1q_1 + a_0 = 0$ se sigue que $a_1^2 - 4a_0 < 0$ porque $q_1 \notin \mathbb{R}$.

Esto completa la verificación de las condiciones de Huang y So dadas en el Teorema 1.3.6.

7.1.2. Autovalores por la izquierda de las matrices simplécticas

Proposición 7.1.5. *Los autovalores por la izquierda de una matriz simpléctica tienen módulo 1.*

Demostración. Sea $A \in Sp(n)$, es decir, $A^*A = I$. Si $Av = \lambda v$ entonces

$$0 < |v|^2 = \langle Av, Av \rangle = \langle \lambda v, \lambda v \rangle = v^* \overline{\lambda} \lambda v = v^* |\lambda|^2 v = |\lambda|^2 |v|^2. \quad \square$$

Usamos ahora la clasificación general de las matrices de orden 2 según el número de autovalores para caracterizar las matrices de $Sp(2)$ cuyo espectro por la izquierda es infinito. Asimismo, damos la forma explícita de los autovalores de estas matrices. Con esta caracterización probaremos que dados cuatro cuaternios arbitrarios de norma 1 siempre hay alguna matriz de $Sp(2)$ para la cual estos cuaternios son autovalores por la izquierda. Este resultado no es trivial y nos será de utilidad para estudiar la categoría LS de $Sp(2)$, como veremos en el Capítulo 8.

Forma general de las matrices de $Sp(2)$

Consideremos el grupo de Lie 10-dimensional $Sp(2)$ formado por las matrices simplécticas 2×2, es decir, matrices cuaterniónicas A tales que $A^*A = I$. Desde el punto de vista geométrico, estas matrices corresponden a los endomorfismos $\mathbb{H}$-lineales por la derecha de $\mathbb{H}^2$ que conservan el producto hermítico $\langle u, v \rangle = u^* v$. Así, una matriz es simpléctica si y solo si sus columnas forman una base ortonormal para este producto hermítico.

Buscamos una expresión general de una matriz simpléctica cualquiera.

Proposición 7.1.6. *Una matriz simpléctica $A \in Sp(2)$ o bien es diagonal o bien es de la forma*

$$A = \begin{pmatrix} \alpha & -\overline{\beta}\gamma \\ \beta & \beta\overline{\alpha}\overline{\beta}\gamma/|\beta|^2 \end{pmatrix}, \ \textit{con } \alpha, \beta \in \mathbb{H}, \quad \beta \neq 0, \quad |\alpha|^2 + |\beta|^2 = 1, \quad |\gamma| = 1.$$

Demostración. Por definición, las dos columnas A_1, A_2 de A forman una base ortonormal de $\mathbb{H}^2$ para el producto hermítico. Sea la primera

$$A_1 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \alpha, \beta \in \mathbb{H}, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Si $\beta = 0$, A es una matriz diagonal $\operatorname{diag}(\alpha, \delta)$ con $|\alpha| = 1 = |\delta|$.

Si $\beta \neq 0$, consideremos la aplicación $\mathbb{H}$-lineal por la derecha $\langle A_1, - \rangle \colon \mathbb{H}^2 \to \mathbb{H}$, que es sobre porque $|A_1| = 1$. Entonces su núcleo $K = (A_1)^\perp$ tiene dimensión $\dim_{\mathbb{H}} K = 1$. Claramente, el vector

$$u = \begin{pmatrix} -\overline{\beta} \\ \beta\overline{\alpha}\overline{\beta}/|\beta|^2 \end{pmatrix} \neq 0,$$

con $|u|=1$, es ortogonal a A_1, de manera que cualquier otro vector en K debe ser un múltiplo cuaterniónico de u. Además, como A_2 tiene norma 1, tenemos

$$A_2 = u\gamma = \begin{pmatrix} -\overline{\beta}\gamma \\ \beta\overline{\alpha}\overline{\beta}\gamma/|\beta|^2 \end{pmatrix}, \quad \gamma \in \mathbb{H}, \quad |\gamma| = 1. \quad \square$$

Matrices simplécticas con infinitos autovalores

Pasamos entonces a aplicar el Teorema 1.3.6 a las matrices simplécticas.

Teorema 7.1.7. *Las únicas matrices simplécticas $A \in Sp(2)$ con infinitos autovalores por la izquierda son de la forma*

$$A = \begin{pmatrix} q\cos\theta & -q\operatorname{sen}\theta \\ q\operatorname{sen}\theta & q\cos\theta \end{pmatrix}, \quad q \in \mathbb{H}, |q| = 1, \quad \theta \in \mathbb{R}, \operatorname{sen}\theta \neq 0.$$

Es decir, A es la matriz correspondiente a la composición $L_q \circ R_\theta$ de una rotación real de ángulo θ, $R_\theta \neq I$, con una traslación por la izquierda L_q por un cuaternio unitario, $|q| = 1$.

Demostración. Claramente, este tipo de matrices cumplen las condiciones de Huang y So recogidas en el Teorema 1.3.6. Recíprocamente, teniendo en cuenta la forma de estas matrices (Prop. 7.1.6), tenemos que comprobar estas condiciones para

$$\begin{aligned} a &= \alpha, & b &= -\overline{\beta}\gamma, \\ c &= \beta, & d &= \beta\overline{\alpha}\overline{\beta}\gamma/|\beta|^2, \end{aligned}$$

con $\beta \neq 0$. Como

$$a_0 = -b^{-1}c = \overline{\gamma}\beta^2/|\beta^2| = s \in \mathbb{R},$$

tenemos que $s = |\gamma| = 1$ (nótese que la condición $a_1 - 4a_0 < 0$ implica $a_0 > 0$). Entonces

$$\gamma = \left(\frac{\beta}{|\beta|}\right)^2.$$

Sustituyendo γ obtenemos $b = -\beta$ y $d = \beta\overline{\alpha}\beta/|\beta|^2$. Por tanto,

$$a_0 = -b^{-1}c = -\beta^{-1}\beta = 1.$$

Si ahora calculamos a_1 queda

$$a_1 = b^{-1}(a-d) = -\beta^{-1}(\alpha - \beta\overline{\alpha}\beta/|\beta|^2) = \frac{-1}{|\beta|^2}(\overline{\beta}\alpha - \overline{\alpha}\beta).$$

De donde $\Re(\alpha_1) = 0$, de modo que la condición $a_1 \in \mathbb{R}$ implica $a_1 = 0$. Luego $\overline{\beta}\alpha$ es igual a su conjugado, $\overline{\alpha}\beta$, *i.e.* es un número real, $\overline{\beta}\alpha = r \in \mathbb{R}$.

Por otro lado, como $|\alpha|^2 + |\beta|^2 = 1$ y $\beta \neq 0$, se verifica que $0 < |\beta| \leq 1$. Tomamos un ángulo cualquiera θ tal que $|\beta| = \operatorname{sen}\theta, \operatorname{sen}\theta \neq 0$ y definimos $q = \beta/|\beta|$ de manera que $\beta = q \operatorname{sen}\theta$ con $|q| = 1$. Por otro lado, las relaciones $|\alpha| = |\cos\theta|$ y $|r| = |\overline{\beta}||\alpha|$ implican que $r = \pm \operatorname{sen}\theta\cos\theta$. Podemos asumir, cambiando el ángulo si fuera necesario pero sin cambiar $\operatorname{sen}\theta$, que $r = \operatorname{sen}\theta$. Entonces,

$$\alpha = r(\overline{\beta})^{-1} = r\beta/|\beta|^2 = q\cos\theta.$$

Finalmente, como $d = \beta\overline{\alpha}\beta/|\beta|^2 = q\cos\theta$, queda completa la prueba. □

Proposición 7.1.8. *Dada $A = L_q \circ R_\theta \in Sp(2)$ sus autovalores por la izquierda son de la forma*

$$\lambda = q(\cos\theta + \operatorname{sen}\theta\omega)$$

con $\omega \in \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle = \mathbb{H}_0, |\omega| = 1$ y $|q| = 1$.

Demostración. El polinomio compañero de A es $r^2 + 1 = 0$ de manera que sus autovalores λ son de la forma $\lambda = q\cos\theta - q\operatorname{sen}\theta\omega$, donde $\omega \in \mathbb{H}$ es raíz de $r^2 + 1 = 0$. □

Corolario 7.1.9. *Un cuaternio $\lambda \in \mathbb{H}$ es autovalor por la izquierda de $A = L_q \circ R_\theta$ si y sólo si $|\lambda| = 1$ y $\Re(\overline{q}\lambda) = \cos\theta$.*

Demostración. Los autovalores por la izquierda de una matriz simpléctica tienen módulo 1 (ver Proposición 7.1.5).

Si λ es un autovalor de A, por la Proposición 7.1.8 se verifica que

$$\begin{aligned} \Re(\overline{q}\lambda) &= \overline{q}q(\cos\theta + \operatorname{sen}\theta\omega) \\ &= |q|^2(\cos\theta + \operatorname{sen}\theta\omega), \end{aligned}$$

luego $\Re(\overline{q}\lambda) = \cos\theta$.

Recíprocamente, si tenemos que $\Re(\overline{q}\lambda) = \cos\theta$ entonces

$$\overline{q}\lambda = \cos\theta + v, \text{ con } v \in \mathbb{H}_0.$$

Así

$$\lambda = q\overline{q}\lambda = q(\cos\theta + v), \text{ con } v \in \mathbb{H}_0.$$

Como $|\lambda| = 1$ y $|q| = 1$, entonces $|\cos\theta + v| = 1$ donde el primer sumando es real y el segundo está en $\mathbb{H}_0$, de manera que

$$1 = |\cos\theta + v|^2 = |\cos\theta|^2 + |v|^2 = \cos^2\theta + |v|^2,$$

luego $|v|^2 = \operatorname{sen}^2\theta$ y por tanto, $v = \operatorname{sen}\theta\omega$, con $\omega \in \mathbb{H}_0, |\omega| = 1$. □

Corolario 7.1.10. *Los autovalores por la izquierda de $A = L_q \circ R_\theta$ son todos los λ tales que $\overline{q}\lambda$ es similar a $e^{\mathbf{i}\theta}$.*

Matrices simplécticas cuyo espectro contiene al menos cuatro cuaternios independientes

Recogemos a continuación dos lemas que necesitaremos para probar que dados cuatro cuaternios unitarios arbitrarios siempre podemos encontrar una matriz simpléctica para la cual dichos cuaternios sean autovalores por la izquierda

Lema 7.1.11. *Si α_i son las filas de una matriz $\alpha \in \mathcal{M}_{n\times n}(\mathbb{R})$, para todo $w \in \mathbb{R}^n$ se verifica que*

$$|\alpha(w)| = [\sum_{i=1}^{n} \langle w, \alpha_i^T \rangle^2]^{1/2},$$

donde $|\cdot|$ denota la norma euclidiana.

Demostración. Si $w = \sum_{j=1}^{n} w_j e_j$ donde e_j denotan los vectores de la base canónica y $\alpha = (\alpha_{ij})$ tenemos que, con la norma euclídea,

$$\begin{aligned} |\alpha w| &= |\alpha(\sum_{j=1}^{n} w_j e_j)| \\ &= |\sum_{j=1}^{n} w_j \alpha(e_j)| = |\sum_{j=1}^{n} w_j (\sum_{i=1}^{n} \alpha_{ij} e_i)| \\ &= |\sum_{j=1}^{n} (\sum_{i=1}^{n} w_j \alpha_{ij}) e_i| = [\sum_{i=1}^{n} (\sum_{j=1}^{n} w_j \alpha_{ij})^2]^{1/2} \\ &= [\sum_{i=1}^{n} \langle w, \alpha_i \rangle^2]^{1/2}. \quad \square \end{aligned}$$

Corolario 7.1.12. *Si $\alpha \in \mathcal{M}_{n\times n}(\mathbb{R})$ tiene rango máximo y sus filas son unitarias, para todo $w \in \mathbb{R}^n, w \neq 0$, se tiene*

$$|\alpha w| < \sqrt{n}|w|.$$

Demostración. En este caso

$$\begin{aligned} \langle w, \alpha_i \rangle &= |w||\alpha_i| \cos \sphericalangle(w, \alpha_i) \\ &= |w| \cos \sphericalangle(w, \alpha_i) \leqslant |w|. \end{aligned}$$

La igualdad sólo se da si $w \in \langle \alpha_i \rangle$. Como los α_i son independientes, $w \neq 0$ no puede estar en la dirección de los cuatro α_i al mismo tiempo. Entonces hay algún $1 \leqslant i \leqslant 4$

tal que $w \notin \langle \alpha_i \rangle$ por tanto,

$$|\alpha w| = [\sum_{i=1}^{n} \langle w, \alpha_i \rangle^2]^{1/2} < (n|w|^2)^{1/2} = \sqrt{n}|w|. \quad \square$$

Teorema 7.1.13. *Dados cuatro cuaternios $\lambda_0, \ldots, \lambda_3 \in \mathbb{H}$ con $|\lambda_s| = 1$ siempre podemos encontrar una matriz $A \in Sp(2)$ tal que $\lambda_0, \ldots, \lambda_3 \in \sigma_l(A)$.*

Demostración. Considerando $A = L_q \circ R_\theta \in Sp(2)$, fijado $\cos\theta$ y escribiendo

$$\begin{aligned} q &= t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \\ \lambda_s &= t_s + x_s\mathbf{i} + y_s\mathbf{j} + z_s\mathbf{k} \end{aligned}$$

podemos expresar la condición del corolario 7.1.9 mediante el siguiente sistema

$$\begin{pmatrix} t_0 & x_0 & y_0 & z_0 \\ \vdots & \vdots & \vdots & \vdots \\ t_3 & x_3 & y_3 & z_3 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = (\cos\theta) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Si denotamos por α la matriz cuyas filas son las coordenadas de los λ_s y $u = (1, 1, 1, 1)^T$, el sistema se escribe

$$\alpha q = \cos\theta u.$$

(a) Si $0 < \text{rang}\alpha < 4$, podemos tomar $\cos\theta = 0$ de manera que tenemos un sistema lineal homogéneo. Este sistema es compatible y el conjunto de soluciones forma un espacio vectorial, en el que tenemos siempre alguna solución q de módulo 1. En este caso, una matriz tal que tiene a $\lambda_0, \ldots, \lambda_3$ como autovalores es, por ejemplo, $A = L_q \circ R_{\pi/2}$.

(b) Cuando $\lambda_0, \ldots, \lambda_3$ son independientes α es de rango máximo y, por tanto, inversible, luego

$$q = (\cos\theta)\alpha^{-1}u.$$

Por el lema 7.1.12 tenemos que $2 = |u| = |\alpha\alpha^{-1}u| < 2|\alpha^{-1}u|$ luego $|\alpha^{-1}u| > 1$ y así podemos escoger θ tal que

$$|\cos\theta| = \frac{1}{|\alpha^{-1}u|} < 1$$

de manera que

$$|q| = |\cos\theta||\alpha^{-1}u| = 1.$$

Hemos obtenido entonces que $\lambda_0, \dots, \lambda_3$ son autovalores por la izquierda de A. $\square$

Ejemplo 7.1.14. Escogemos el conjunto formado por $\{-1, \mathbf{i}, -\mathbf{i}, -\mathbf{k}\}$. Fijamos entonces $\cos\theta = 0$ el sistema es

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

con solución $q \in \langle \mathbf{j} \rangle$. Es decir, las matrices

$$A = \pm \begin{pmatrix} 0 & -\mathbf{j} \\ \mathbf{j} & 0 \end{pmatrix} = L_{\pm\mathbf{j}} \circ R_{\pi/2}$$

tienen a $1, \mathbf{i}, -\mathbf{i}, \mathbf{k}$ por autovalores.

Autovalores por la derecha

Un problema interesante planteado por Zhang recientemente (en 2007) [93] consiste en comparar el espectro por la izquierda y por la derecha de una matriz cuaterniónica. Los siguientes lemas nos dan los autovalores por la derecha de las matrices $L_q \circ R_\theta$.

Lema 7.1.15. *Los autovalores por la derecha de $L_q \circ R_\theta$ no dependen más que de la clase de similitud de q.*

Lema 7.1.16. *Los cuatro autovalores por la derecha de la forma compleja de la matriz $A = L_q \circ R_\theta$ con $q = a + \mathbf{j}b$ quedan completamente determinados por las ecuaciones*

$$\begin{cases} \Re(p_1) + \Re(p_2) = 2\Re(a)c \\ \\ \Re(p_1)\Re(p_2) = \Re(a) - s^2 \end{cases}$$

donde $s = \operatorname{sen}\theta, c = \cos\theta$ y $p_1 \neq \overline{p_2}$ son dos de los autovalores por la derecha de A.

Recordemos que los autovalores de $c(A)$ son conjugados dos a dos y unitarios, de manera que con saber la parte real de dos no conjugados ya quedan completamente determinados todos.

Demostración. Por el Lema 7.1.15 podemos suponer

$$c(A) = \left(\begin{array}{cc|cc} ac & -as & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ as & ac & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & \overline{a}c & -\overline{a}s \\ & & \overline{a}s & \overline{a}c \end{array}\right)$$

con $a \in \mathbb{C}, |a| = 1$ de manera que

$$\begin{array}{rcl} \det(c(A) - pI) & = & \\ (p^2 - 2acp + a^2)(p^2 - 2\overline{a}cp + \overline{a}^2) & = & \\ p^4 - 2\overline{a}cp^3 + \overline{a}^2p^2 - 2acp^3 + 4|a|^2c^2p^2 - 2|a|^2\overline{a}cp + a^2p^2 - 2|a|^2acp + |a|^4 & = & \\ p^4 - 4\Re(a)cp^3 + (\overline{a}^2 + 4|a|^2c^2 + a^2)p^2 - (2|a|^2\overline{a}c + 2|a|^2ac)p + |a|^4 & = & \\ p^4 - 4\Re(a)cp^3 + (4\Re(a) + 2c^2 - 2s^2)p^2 - 4\Re(a)cp + 1. & & \end{array}$$

Por otro lado, por la forma que tiene en este caso $c(A)$, sus cuatro autovalores son conjugados dos a dos luego,

$$\begin{array}{rcl} \det(c(A) - pI) & = & \\ (p - p_1)(p - \overline{p}_1)(p - p_2)(p - \overline{p}_2) & = & \\ p^4 - 2(\Re(p_1) + \Re(p_2)p^3 + (2 + 4\Re(p_1)\Re(p_2))p^2 - 2(\Re(p_1) + \Re(p_2))p + 1. & & \end{array}$$

de donde

$$\begin{array}{l} -4\Re(a)c = -2(\Re(p_1) + \Re(p_2)) \\ 4\Re(a) + 2c^2 - 2s^2 = 2 + 4\Re(p_1)\Re(p_2) \quad \square \end{array}$$

Ejemplo 7.1.17. Matrices de la forma $\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$ con $|\omega| = 1, \omega \in \mathbb{H}_0$.

En este caso tenemos $\cos\theta = 0, \operatorname{sen}\theta = -1$ y $\omega = a + \mathbf{j}b$ con $a, b \in \mathbb{C}$, a imaginario puro. Sean p_1, p_2 dos autovalores por la derecha, entonces:

$$\begin{array}{l} \Re(p_1) + \Re(p_2) = 0, \\ \Re(p_1)\Re(p_2) = -1, \end{array}$$

luego $1 = \Re(p_1) = -\Re(p_2)$, de manera que estas matrices sólo tienen dos autovalores por la derecha y son reales, $p_1 = 1, p_2 = -1$.

7.2. Funciones características de las matrices cuaterniónicas

Al final del artículo [90] al que nos referíamos al introducir los autovalores por la izquierda, Wood hace notar que

> "en el caso 2×2 de la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ hay un determinante parcialmente definido $b - ac^{-1}d$ y una función característica parcialmente definida
>
> $$\lambda c^{-1}\lambda - \lambda c^{-1}d - ac^{-1}\lambda - b + [ac^{-1}d] = 0 \tag{7.3}$$
>
> que reduce el problema de los autovalores al teorema fundamental [del álgebra]. Las dificultades empiezan con las matrices 3×3".

Wood dice también que

> "desafortunadamente, parece no haber una función determinante adecuada en el caso cuaterniónico para reducir el problema de los autovalores al teorema fundamental del álgebra."

7.2.1. Antecedentes

El problema es encontrar una función característica para matrices cuaterniónicas de orden n que proporcione los autovalores por la izquierda. En el caso particular de las matrices hermíticas, como todos sus autovalores por la derecha son reales (ver Prop. 1.2.13), los espectros por la izquierda y por la derecha coinciden, de modo que puede construirse un polinomio característico con variable real

$$p(t) = \det(M - tI), \quad \text{con } t \in \mathbb{R},$$

donde det es el determinante de las matrices hermíticas definido en la Subsección 1.2.3. Las raíces de este polinomio son los autovalores (por la izquierda y por la derecha) de M. Este polinomio característico coincide con la nueva función característica que propondremos en la Subsección 7.2.2.

Zhang propone utilizar el polinomio característico complejo para obtener una nueva función en $\mathbb{C} \times \mathbb{C}$ a partir de la cual se pueden obtener los autovalores por la izquierda. En [92] utiliza que un cuaternio λ es un autovalor por la izquierda de la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si existe algún $v \in \mathbb{C}^{2n}, v \neq 0$, tal que

$$c(M - \lambda I)v = 0,$$

ya que la forma compleja de la matriz cuaterniónica $M - \lambda I$ es singular. Si expresamos $M = X + \mathbf{j}Y$, con $X, Y \in \mathcal{M}_{n\times n}(\mathbb{C})$ y $\lambda = x + \mathbf{j}y$, con $x, y \in \mathbb{C}$, λ es un autovalor por la izquierda de M si y solo si es raíz de la función $\sigma\colon \mathbb{C}\times\mathbb{C} \to \mathbb{R}$ dada por

$$\sigma(x,y) = \det \begin{pmatrix} X - xI & -\overline{Y} + \overline{y}I \\ Y - yI & \overline{X} - \overline{x}I \end{pmatrix}. \tag{7.4}$$

Otra aproximación a la noción de función característica cuaterniónica es la de I. Gelfand *et alii*. Estos autores utilizan los quasideterminantes para construir funciones características de matrices cuaterniónicas [27]. Para cada matriz cuadrada de orden n definen n^2 funciones características como sigue.

Definición 7.2.1. Sea $A \in \mathcal{M}_{n\times n}(\mathbb{H})$, se definen las *GR-funciones características* de A como

$$f_{ij}(\lambda) = |\lambda I - A|_{ij} \text{ para } 1 \leqslant i, j \leqslant n,$$

donde $|\cdot|_{ij}$ es el $(i,j)-$quasideterminante.

A partir del teorema que generaliza la identidad de Jacobi al caso cuaterniónico (Teor. 1.2.23) podemos afirmar que estas GR-funciones características verifican que todas sus raíces son autovalores por la izquierda. Sin embargo, ninguna de ellas nos da todos los autovalores por la izquierda. Veamos algunos ejemplos.

Ejemplo 7.2.2. Tomamos la matriz triangular $B = \begin{pmatrix} \mathbf{i} & 0 & 0 \\ \mathbf{k} & \mathbf{j} & 0 \\ -3\mathbf{i} & 2\mathbf{k} & \mathbf{k} \end{pmatrix}$. Su espectro por la izquierda está formado por los elementos de la diagonal, es decir, $\sigma_l(B) = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Sin utilizar la Proposición 1.2.19, la única GR-función característica que podemos calcular es

$$f_{11}(\lambda) = \mathbf{i} - \lambda.$$

Utilizando 1.2.19 obtenemos

- $f_{12}(\lambda) = (\mathbf{i} - \lambda)\mathbf{k}(\mathbf{j} - \lambda)$,
- $f_{13}(\lambda) = (\mathbf{i} - \lambda)\left(3\mathbf{i} + 2\mathbf{k}(\mathbf{j} - \lambda)^{-1}\mathbf{k}\right)^{-1}(\mathbf{k} - \lambda)$,
- $f_{22}(\lambda) = \mathbf{j} - \lambda$,
- $f_{23}(\lambda) = \frac{1}{2}(\mathbf{j} - \lambda)\mathbf{k}(\mathbf{k} - \lambda)$,
- $f_{33}(\lambda) = \mathbf{k} - \lambda$

y las funciones $f_{21}(\lambda)$, $f_{31}(\lambda)$ y $f_{32}(\lambda)$ no existen.

Se ve claramente que ninguna de las f_{ij} tiene a todos los elementos del espectro por la izquierda como raíz.

Ejemplo 7.2.3. Las GR-funciones características no proporcionan los autovalores por la derecha. Sea

$$D = \begin{pmatrix} \mathbf{i} & -1 & 0 \\ 0 & \mathbf{k} & \mathbf{i} \\ 0 & \mathbf{k} & \mathbf{j} \end{pmatrix}$$

La única GR-función que existe para esta matriz es :

$$f_{11}(\lambda) = \mathbf{i} - \lambda$$

Sin embargo, el polinomio característico de la complexificada de D es $2+2x+4x^2+2x^3+3x^4+x^6$. Puede comprobarse que ninguna de sus raíces es similar a $\mathbf{i}$.

7.2.2. Definición

Con estos antecedentes pasamos a introducir la noción de función característica para una matriz cuaterniónica, que generaliza el polinomio característico usual de los casos real y complejo. En particular, sus raíces son los autovalores por la izquierda. Como veremos, esta definición concuerda de manera natural con la ecuación para orden 2 que da R.M.W. Wood (7.3), así como con el método propuesto por W. So en [79] para calcular los autovalores de matrices de orden 3. Además muestra que, en contra de lo que dice Wood en [90], sí hay una función que permita acometer el estudio de los autovalores.

En el caso 2×2 esta función característica nos permitirá probar de un modo geométrico el resultado de Huang y So (Teorema 1.3.6).

Definición 7.2.4. Una aplicación $\mu : \mathbb{H} \to \mathbb{H}$ es una *función característica* de la matriz $M \in \mathcal{M}_{n\times n}(\mathbb{H})$ si, salvo una constante, su módulo $|\mu(\lambda)| = \mathrm{Sdet}(M - \lambda I)$ para todo $\lambda \in \mathbb{H}$.

Nótese que λ es un autovalor por la izquierda de M si y solo si $\mu(\lambda) = 0$.

Nota–. Como ya dijimos, es un hecho conocido que el espectro por la izquierda no es invariante por semejanza. Sin embargo, por la Proposición 1.2.2, si P es una matriz real inversible,

$$\mathrm{Sdet}(M - \lambda I) = \mathrm{Sdet}(PMP^{-1} - \lambda I).$$

Es decir, M y PMP^{-1} tienen las mismas funciones características.

Ejemplo 7.2.5. Matrices diagonales y triangulares. Si

$$D = \operatorname{diag}(q_1, \dots, q_n)$$

entonces

$$\mu(\lambda) = (q_1 - \lambda) \cdots (q_n - \lambda)$$

es una función característica para D. Análogamente para una matriz triangular.

7.2.3. Una función característica para orden dos

El determinante de Study nos permite dar, para orden 2, una función característica polinómica de grado 2. El comentario de Wood citado en 7.2 puede reformularse así:

Teorema 7.2.6. *Toda matriz* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{H})$ *tiene una función característica polinómica* $\mu(\lambda)$.

- *Si* $b = 0$ *viene dada por*

$$\mu(\lambda) = (d - \lambda)(a - \lambda).$$

- *Si* $b \neq 0$,

$$\mu(\lambda) = c - (d - \lambda)b^{-1}(a - \lambda).$$

Demostración. Si $b = 0$, $A - \lambda I$ es una matriz triangular,

$$\operatorname{Sdet}\begin{pmatrix} a - \lambda & 0 \\ c & d - \lambda \end{pmatrix} = |(d - \lambda)(a - \lambda)|.$$

Si $b \neq 0$, por las propiedades de Sdet podemos hacer la siguiente transformación

$$\operatorname{Sdet}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \operatorname{Sdet}\begin{pmatrix} 0 & b \\ c - db^{-1}a & d \end{pmatrix} = |b||c - db^{-1}a|.$$

luego $\operatorname{Sdet}(A - \lambda I) = 0$ equivale a que

$$c - (d - \lambda)b^{-1}(a - \lambda) = 0. \quad \square$$

Nota–. L. Huang [32] propone otra ecuación característica para las matrices de orden 2. Separa el caso $c = 0$ y para $c \neq 0$ dice que $\lambda \in \sigma_l(A)$ si y solo si:

$$(\lambda - a)c^{-1}(\lambda - d) - b = 0.$$

Este polinomio se obtiene también de manera inmediata sumándole $(\lambda - a)c^{-1}f_2$ a la fila f_1 en la matriz $\lambda I - A$. Esta expresión es equivalente a

$$b - (a - \lambda)c^{-1}(d - \lambda) = 0,$$

que es la expresión que da Wood al final de [90] (hay un error tipográfico en el original).

7.2.4. Función característica para matrices 3×3

Recordemos que lo único que se sabe para las matrices de tamaño 3×3 se lo debemos a W. So [79]. Lo que So ofrece es un estudio caso por caso, según las relaciones entre las entradas de las matrices, en el que da una familia de polinomios cuaterniónicos de grado menor o igual que tres cuyas raíces son los autovalores por la izquierda. Las ecuaciones resultantes no se saben resolver.

Desarrollamos ahora un método similar al que hemos visto para orden 2. De nuevo nos basamos en que los autovalores por la izquierda de una matriz $A \in \mathcal{M}_{3\times 3}(\mathbb{H})$ son las raíces de $\mathrm{Sdet}(A - \lambda I) = 0$. En general, la función característica que obtenemos es una función racional. Sin embargo, en el caso de que la matriz tenga alguna entrada nula fuera de la diagonal, es posible obtener una función característica polinómica. Podría emplearse un algoritmo similar al que veremos a continuación para calcular el $\mathrm{Sdet}(A)$ para cualquier orden $n > 3$.

Consideramos a lo largo de todo la sección la matriz cuaterniónica

$$A = \begin{pmatrix} a & b & c \\ f & g & h \\ p & q & r \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{H}).$$

Caso $c = 0$

Comenzamos estudiando el caso más sencillo, cuando algún elemento no diagonal de la matriz se anula. Las propiedades de Sdet y en particular la Prop. 1.2.2 nos permiten ir haciendo transformaciones de filas y columnas hasta situar el cero en el puesto a_{13}. La ventaja de este caso es que sus autovalores son las raíces de un polinomio cuaterniónico de grado tres.

Partimos de

$$\mathrm{Sdet}(A-\lambda I) = \mathrm{Sdet}\begin{pmatrix} a-\lambda & b & 0 \\ f & g-\lambda & h \\ p & q & r-\lambda \end{pmatrix}.$$

1. En primer lugar, si $b, h = 0$, tenemos una matriz triangular así que podemos tomar

$$\mu(\lambda) = (r-\lambda)(g-\lambda)(a-\lambda). \tag{7.5}$$

2. Si $b = 0$ pero $h \neq 0$, el Lema 1.2.5 permite reducirnos al caso 2×2 y obtenemos

$$\mu(\lambda) = \left(q - (r-\lambda)h^{-1}(g-\lambda)\right)(a-\lambda). \tag{7.6}$$

3. Finalmente, si $b \neq 0$, procedemos como sigue. Hacemos un cero en la primera fila restando a la primera columna C_1 la $C_2 b^{-1}(a-\lambda)$,

$$\mathrm{Sdet}(A) = \mathrm{Sdet}\begin{pmatrix} 0 & b & 0 \\ f-(g-\lambda)b^{-1}(a-\lambda) & g-\lambda & h \\ p-qb^{-1}(a-\lambda) & q & r-\lambda \end{pmatrix}$$

y permutamos la segunda y la tercera columna para reducirla al caso 2×2. Tomamos entonces como función característica:

$$\mu(\lambda) = p - qb^{-1}(a-\lambda) - (r-\lambda)h^{-1}\left(f-(g-\lambda)b^{-1}(a-\lambda)\right).$$

Nota–. Alternativamente, podríamos simplemente permutar en la matriz inicial las filas y las columnas segunda y tercera para obtener una matriz PAP^{-1} con la misma función característica. Esta nueva matriz verifica que $c \neq 0$ y podríamos obtener su función como veremos a continuación (Prop. 7.2.7). Sin embargo, nótese que con este último método, en lugar de obtener un polinomio, obtenemos una función racional.

Caso $c \neq 0$

En la situación más general, cuando $c \neq 0$, utilizando las propiedades de Sdet (Prop. 1.2.3) podemos hacer ceros en la primera fila,

$$\begin{pmatrix} 0 & 0 & c \\ f-hc^{-1}a & g-hc^{-1}b & h \\ p-rc^{-1}a & q-rc^{-1}b & r \end{pmatrix}.$$

Del Lema 1.2.5 y los resultados de orden 2×2 se sigue:

Proposición 7.2.7. *Si $c \neq 0$, entonces* $\mathrm{Sdet}(A)$ *viene dado:*

1. *cuando $g - hc^{-1}b \neq 0$, por*

$$|c| \cdot |g - hc^{-1}b| \cdot |p - rc^{-1}a - (q - rc^{-1}b)(g - hc^{-1}b)^{-1}(f - hc^{-1}a)|;$$

2. *cuando $g - hc^{-1}b = 0$, por*

$$|c| \cdot |q - rc^{-1}b| \cdot |f - hc^{-1}a|.$$

Definición 7.2.8. Dada la matriz $A \in \mathcal{M}_{3\times 3}(\mathbb{H})$ llamamos *polo de A* al punto $\lambda_0 = g - hc^{-1}b$.

Aplicando la Proposición 7.2.7 a la matriz $A - \lambda I$ obtenemos la siguiente función característica de A.

Proposición 7.2.9. *Sea A una matriz de orden 3×3 tal que $c \neq 0$. Se puede definir una función característica para A como sigue:*

1. *si $\lambda_0 = g - hc^{-1}b$ es el polo de A,*

$$\mu(\lambda_0) = \left(q - (r - \lambda_0)c^{-1}b\right)\left(f - hc^{-1}(a - \lambda_0)\right);$$

2. *en otro caso*

$$\begin{aligned}\mu(\lambda) \quad = \quad & (\lambda_0 - \lambda)\left[\left(p - (r - \lambda)c^{-1}(a - \lambda)\right) - \right. \\ & \left. \left(q - (r - \lambda)c^{-1}b\right)(\lambda_0 - \lambda)^{-1}\left(f - hc^{-1}(a - \lambda)\right)\right].\end{aligned}$$

Esta función está bien definida y cumple que:

1. $\mathrm{Sdet}(A - \lambda I) = |\mu(\lambda)|$ para todo $\lambda \in \mathbb{H}$;

2. $\mu(\lambda) = 0$ si y solo si λ es autovalor por la izquierda de A.

Nota–. Como recordamos antes, en [79], W. So probó que los autovalores por la izquierda de estas matrices son las raíces de ciertos polinomios cuaterniónicos de grado ≤ 3. A pesar de que nuestro método es totalmente diferente del suyo, obtenemos que la función de la Proposición 7.2.9 es exactamente la fórmula que da So en [79, p. 563]. Este es el motivo por el que hemos elegido hacer los cálculos de los determinantes comenzando por la esquina superior derecha.

Nota–. El polo λ_0 puede tanto estar como no estar en el espectro de la matriz. Los siguientes ejemplos nos lo muestran.

Ejemplo 7.2.10. Una matriz para la que el polo es autovalor.

$$A = \begin{pmatrix} 0 & -\mathbf{j} & \mathbf{i} \\ -1+\mathbf{j} & \mathbf{j} & \mathbf{k} \\ p & q & r \end{pmatrix},$$

con $p, q, r \in \mathbb{H}$ cualesquiera. Para esta matriz el polo toma el valor $\lambda_0 = 1 + \mathbf{j}$. Su imagen por la función característica de la Prop. 7.2.9 es

$$\mu(\lambda_0) = (q - (r - 1 - \mathbf{j})\mathbf{k})\,(-1 + \mathbf{j} + \mathbf{j}(-1 - \mathbf{j})) = 0.$$

Ejemplo 7.2.11. Matriz para la que el polo no es autovalor.

$$A = \begin{pmatrix} 0 & \mathbf{i} & 1 \\ 3\mathbf{i} - \mathbf{k} & 0 & 1 \\ \mathbf{k} & -1 + \mathbf{j} + \mathbf{k} & 0 \end{pmatrix}.$$

En este caso, $\lambda_0 = -\mathbf{i}$ y

$$\mu(\lambda_0) = (-1 + \mathbf{j} + \mathbf{k} + 1)(3\mathbf{i} - \mathbf{k} - \mathbf{i}) = 1 - \mathbf{i} + 2\mathbf{j} - 2\mathbf{k}.$$

Continuidad en el polo

El módulo de la que hemos definido como función característica es exactamente el determinante de Study de la matriz $A - \lambda I$. Como Sdet es una función continua podemos afirmar que

$$|\mu(\lambda_0)| = \lim_{\lambda \to \lambda_0} |\mu(\lambda)|,$$

es decir, el módulo del límite existe y es finito.

Sin embargo, el mismo ejemplo 7.2.11 nos muestra que a pesar de que la norma $|\mu|$ de la función característica de la Proposición 7.2.9 es continua, μ puede no serlo en el polo λ_0.

Sea

$$\begin{pmatrix} 0 & \mathbf{i} & 1 \\ 3\mathbf{i} - \mathbf{k} & 0 & 1 \\ \mathbf{k} & -1 + \mathbf{j} + \mathbf{k} & 0 \end{pmatrix}.$$

Como hemos visto, su polo es $\lambda_0 = -\mathbf{i}$ y $\mu(\lambda_0) = (\mathbf{j} + \mathbf{k})(2\mathbf{i} - \mathbf{k}) = 1 - \mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$.

Pero tomando límites restringidos se comprueba fácilmente que el valor del límite que buscamos depende de por dónde nos acerquemos al polo. De hecho, el límite

$$\lim_{\varepsilon\to 0} \mu(-\mathbf{i}+\varepsilon q) = -q(\mathbf{j}+\mathbf{k})q^{-1}(2\mathbf{i}-\mathbf{k})$$

depende de $q \in \mathbb{H}$, luego no existe $\lim\limits_{\lambda\to\lambda_0} \mu(\lambda)$.

Si es o no posible encontrar siempre, aparte del caso polinómico, una función característica que sea continua es una cuestión abierta.

7.3. Estudio mediante el grado de una función característica

En [33] Huang y So probaron que una matriz 2×2 puede tener uno, dos o infinitos autovalores. En la sección 7.1.1 propusimos una nueva prueba de ese mismo resultado. Ambos métodos, aparentemente, son difíciles de generalizar para orden $n > 2$.

Frente a estos estudios, ambos de naturaleza algebraica, proponemos aquí una visión más topológica. La idea fundamental consiste en estudiar el grado topológico de las funciones características 7.2.6 y 7.2.9 y el rango de sus linealizadas. El caso de la matrices 2×2 queda completamente cerrado ya que podemos discutir las ecuaciones cuaterniónicas que obtenemos al linealizar. Para orden 3×3 no obtenemos tanta información pero sí podemos decir cuáles son las ecuaciones que habría que discutir.

7.3.1. Teoría del grado

Veamos a continuación algunos resultados topológicos que necesitamos para nuestro estudio.

El grado topológico (también llamado grado de Brouwer) de una aplicación continua puede definirse utilizando técnicas o bien de topología algebraica [11] o bien del análisis funcional [9, 68]. Nuestra intención es aplicar los siguientes resultados conocidos (ver, por ejemplo, [58, pag 101]).

Definición 7.3.1. Dada una aplicación diferenciable ente variedades, $\rho : M \to N$ con $\dim M = \dim N$, diremos que $r \in N$ es un *valor regular* de ρ si la diferencial $\rho_{*q} : T_qM \to T_rN$ es de rango máximo para todo $q \in \rho^{-1}(r)$.

Lema 7.3.2 (de Sard)**.** *El conjunto de valores regulares es denso en N.*

Teorema 7.3.3. *Sea M una variedad orientable conexa y cerrada y $\nu\colon M \to M$ un aplicación diferenciable de grado k. Sea $m \in M$ un valor regular tal que la diferencial $\nu_{*\lambda}$ conserva la orientación para cualquier λ de la fibra $\nu^{-1}(m)$. Entonces, la imagen recíproca $\nu^{-1}(m)$ es un conjunto finito con k elementos.*

En el libro de Eilenberg-Steenrod se puede encontrar una prueba rigurosa del siguiente resultado [15, pags 304–310].

Proposición 7.3.4. *Dos aplicaciones continuas en la esfera $S^n \to S^n$ son homótopas si y solo si tienen el mismo grado.*

7.3.2. Derivación en espacios de Banach

La existencia de una norma multiplicativa $|q| = (q\overline{q})^{1/2}$ en $\mathbb{H}$ garantiza que la prueba usual de la regla de Leibniz sigue pudiendo aplicarse en este contexto. Veamos con detalle cómo obtener las reglas de derivación en el ámbito no conmutativo. Observemos que, siempre que respetemos el orden, en el caso cuaterniónico se conservan las reglas de derivación del producto y del cociente:

Lema 7.3.5. *Sean $A, B\colon \mathbb{R} \to \mathcal{M}_{n\times n}(\mathbb{H})$ diferenciables. Entonces*

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

Demostración.

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

□

A partir de esta propiedad podemos establecer una fórmula para la diferencial del producto de aplicaciones cuaterniónicas.

Lema 7.3.6. *Sean $f, g\colon \mathbb{H} \to \mathbb{H}$ dos aplicaciones diferenciables. Entonces, la diferencial del producto viene dada por*

$$(fg)_{*\lambda}(X) = f_{*\lambda}(X)g(\lambda) + f(\lambda)g_{*\lambda}(X).$$

Lema 7.3.7. *Sea* $B\colon \mathbb{R} \to GL(n,\mathbb{H}) \subset \mathcal{M}_{n\times n}(\mathbb{H})$, *entonces*

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

Demostración.

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

□

Linealización

Consideramos el espacio vectorial $\mathbb{H}$ como una variedad diferenciable difeomorfa a $\mathbb{R}^4$. Entonces, para una aplicación $\rho\colon \mathbb{H} \to \mathbb{H}$ diferenciable, su diferencial en cualquier punto $\lambda \in \mathbb{H}$, $\rho_{*\lambda}\colon \mathbb{H} \to \mathbb{H}$, puede calcularse con la siguiente fórmula

$$\rho_{*\lambda}(X) = \frac{d}{dt}_{|t=0}\rho(\lambda + tX) = \lim_{t\to 0}\frac{1}{t}\left(\rho(\lambda + tX) - \rho(\lambda)\right).$$

7.3.3. Estudio topológico del caso 2×2

A continuación desarrollamos con detalle el caso 2×2, interpretando los resultados conocidos desde un punto de vista geométrico. Sea la matriz cuaterniónica de orden dos

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

El caso $b = 0$ está completamente cerrado, A es una matriz triangular y sus autovalores son $\lambda_1 = a$ y $\lambda_2 = c$. Así que, a partir de ahora podemos suponer, sin pérdida de generalidad, $b \neq 0$ ya que para cualquier matriz real inversible P, $\sigma_l(PAP^{-1}) = \sigma_l(A)$, lo que permite permutar filas y columnas.

Como habíamos visto en la Sección 7.2.3, calcular el espectro es equivalente a encontrar las raíces de la función característica

$$\mu(\lambda) = c - (d - \lambda)b^{-1}(a - \lambda). \tag{7.7}$$

Nótese que, como veremos a continuación, esta función característica polinómica $\mu\colon \mathbb{H} \to \mathbb{H}$ se puede extender de manera continua (e incluso diferenciable) a una aplicación $\mu\colon S^4 \to S^4$ en la esfera $S^4 = \mathbb{H} \cup \{\infty\}$. Esto se debe a que $\lim |\mu(\lambda)| = \infty$ cuando $|\lambda| \to \infty$.

Para este caso puede reformularse la Proposición 7.3.4 como sigue.

Proposición 7.3.8. *Una aplicación polinómica como μ y la aplicación λ^2 son homótopas en la esfera S^4, por tanto, tienen el mismo grado topológico,* 2.

Demostración. Probamos que podemos extender la función característica a

$$\mu : \mathbb{H} \cup \{\infty\} \mapsto \mathbb{H} \cup \{\infty\}.$$

En efecto,

$$|\mu(\lambda)| \geq |(d - \lambda)b^{-1}(a - \lambda)| - |c|$$

pero

$$|(d - \lambda)b^{-1}(a - \lambda)| = |(d - \lambda)| \cdot |b^{-1}| \cdot |(a - \lambda)|$$

y como a, d son constantes,

$$\lim_{|\lambda| \to \infty} |d - \lambda| = \infty \quad \text{y} \quad \lim_{|\lambda| \to \infty} |a - \lambda| = \infty$$

por tanto,

$$\lim_{|\lambda| \to \infty} |\mu(\lambda)| = \infty.$$

Veamos ahora que μ es homótopa a $-\lambda^2$. Podemos tomar en $\mathbb{H}$ los caminos $a(t) = at$, $c(t) = ct$ y $d(t) = dt$ que llevan a, c y d en 0 respectivamente. Así,

$$\mu(\lambda) \simeq ct - (dt - \lambda)b^{-1}(at - \lambda),$$

de modo que nos queda solo el término de mayor grado, $-\lambda b^{-1}\lambda$. Ahora bien, como $\mathbb{H} - \{0\}$ es conexo por caminos, podemos encontrar un camino $\gamma(t)$ de b^{-1} a 1 sin pasar por 0, por tanto, $\mu(\lambda) \simeq -\lambda^2$ y la homotopía conserva el punto del infinito. $\square$

Como consecuencia, teniendo en cuenta que el determinante de la matriz asociada a la aplicación diferencial es siempre mayor o igual que cero (ver Prop. 1.5.1), obtenemos el resultado que queríamos.

Corolario 7.3.9. *Si* 0 *es un valor regular de* μ*, entonces*

$$\mu^{-1}(0) = \sigma_l(A)$$

está formado por dos puntos.

Linealización

Dada una función característica μ clasificaremos las matrices en función de que 0 sea o no un valor regular para μ. Hacemos la diferencial de μ para calcular su rango.

Proposición 7.3.10. *Para cualquier autovalor izquierda* λ *de* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $b \neq 0$, *la diferencial de su función característica* $\mu : \mathbb{H} \longrightarrow \mathbb{H}$ *viene dada por*

$$\mu_{*\lambda}(X) = Xb^{-1}(a - \lambda) + (d - \lambda)b^{-1}X. \tag{7.8}$$

La prueba es una aplicación directa de las reglas de derivación del producto y del inverso en el caso cuaterniónico vistas en la Subsección 7.3.2

La expresión obtenida es un polinomio cuaterniónico lineal tipo ecuación de Sylvester. Utilizando algunos resultados conocidos recogidos en la Sección 1.5.1 para calcular el rango de esta aplicación podemos clasificar los posibles espectros de una matriz 2×2 según los valores que tome dicho rango. El método es completamente nuevo; la clasificación que damos coincide con la de L. Huang y W. So [33].

Clasificación

Sea λ un autovalor de A, es decir $\mu(\lambda) = 0$ para la aplicación μ en (7.7). Siguiendo la notación que hemos usado para la ecuación de Sylvester en la Sección 1.5.1 llamaremos $\alpha = (d - \lambda)b^{-1}$ y $\beta = b^{-1}(a - \lambda)$.

Proposición 7.3.11. *Si* $\operatorname{rang} \mu_{*\lambda} = 0$*, entonces* a_0, a_1 *son números reales y* $\Delta = 0$*. Más aún,* $\lambda = (a + d)/2$ *y la matriz tiene un único autovalor*

Demostración. Sea $\lambda \in \sigma_l(A)$ y supongamos $\operatorname{rang} \mu_{*\lambda}$ es nulo. Por la Proposición 1.5.2, esto se da cuando y solo cuando

$$\begin{aligned} &(d - \lambda)b^{-1} = s \in \mathbb{R} \\ &b^{-1}(a - \lambda) = -s. \end{aligned} \tag{7.9}$$

Como λ es autovalor, $c = (d-\lambda)b^{-1}(a-\lambda) = -bs^2$. Además, de (7.9) se deduce que $-2sb = a - d$ y por tanto, $\lambda = \frac{1}{2}(a-d)$ es el único autovalor. □

Lema 7.3.12. *Sean q_1, q_2 dos cuaternios similares que no conmutan. Entonces, la ecuación $\lambda^2 - (q_1+q_2)\lambda + q_1q_2 = 0$ tiene una única solución $\lambda = q_2$.*

Demostración. Si $\lambda \neq q_2$ es una solución, de la igualdad $(\lambda - q_2)\lambda = q_1(\lambda - q_2)$ se sigue que λ y q_1 son similares, entonces, $\Re(\lambda) = \Re(q_1) = \Re(q_2)$ y $|\lambda| = |q_1| = |q_2|$. Sustituyendo en la ecuación, se van las partes reales y las normas, así que podemos suponer que q_1, q_2, λ son imaginarios puros unitarios. Esto es, $q_1, q_2, \lambda \in S^2 \subset \mathbb{H}_0$. Por tanto, $\lambda^2 = -1 = q_2^2$, de manera que la ecuación se reduce a $(q_1+q_2)\lambda = (q_1+q_2)q_2$ es decir, $\lambda = q_2$, pero esto es una contradicción. □

Proposición 7.3.13. *Si* $\operatorname{rang}\mu_{*\lambda} = 2$ *pueden ocurrir dos cosas:*

1. *El espectro es esférico tipo (1.4) y todos los autovalores de A son de rango 2;*
2. *o la matriz A solo tiene un autovalor.*

Demostración. Usando el difeomorfismo $a + b\sigma_l(A') = \sigma_l(A)$ podemos sustituir A por su "matriz compañera"

$$\begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}.$$

Como el rango es 2, por el Teorema 1.5.2 tenemos que

$$\begin{aligned} \alpha &= t + \omega_1, \\ \beta &= -t + \omega_2 \end{aligned}$$

con $\omega_1, \omega_2 \in \mathbb{H}_0 = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$, $|\omega_1| = |\omega_2| \neq 0$. Entonces

$$a_1 = -2t + \omega_2 - \omega_1.$$

La primera posibilidad es que $\omega_2 = \omega_1$, entonces $a_1 = -2t$. Como los autovalores anulan la función característica, $\mu(\lambda) = 0$,

$$\begin{aligned} a_0 &= t^2 + |\omega_2|^2 \neq 0, \\ \Delta &= -4|\omega_2|^2 < 0. \end{aligned}$$

Entonces, tenemos el caso esférico (1.4). En particular,

$$2\lambda = -a_1 + q$$

con $q = -2\omega_2$. Los otros autovalores son de la forma

$$(-a_1 + q)/2$$

con $q^2 = -4|\omega_2|^2$, entonces la diferencial de μ verifica que

$$\begin{aligned} \alpha &= t - q^{-1}/2, \\ \beta &= -t - q/2 \end{aligned}$$

de modo que todos los demás autovalores también tienen rango 2.

La segunda posibilidad es que $\omega_2 \neq \omega_1$. Entonces

$$\begin{aligned} a_1 &= -2t + \omega_2 - \omega_1, \\ a_0 &= (t + \omega_1)(t - \omega_2) \end{aligned}$$

y el Lema 7.3.12 nos muestra que el único autovalor es $\lambda = t - \omega_2$. □

Proposición 7.3.14. *El rango de la diferencial es máximo,* $\operatorname{rang} \mu_{*\lambda} = 4$*, si y solo si la matriz tiene exactamente dos autovalores.*

Demostración. Como para un autovalor λ la diferencial es de rango máximo, la matriz A no puede estar en las condiciones de las Proposiciones 7.3.11 o 7.3.13, por lo tanto, todos sus autovalores son de rango máximo.

Entonces, por el teorema de la función inversa, la fibra $\mu^{-1}(0)$ es discreta (de hecho, compacta) y por el Teorema 7.3.3, su cardinal es igual al grado de la aplicación μ , que es 2 (Prop. 7.3.4). □

Obsérvese que en el caso particular en el que la matriz sea triangular, es decir, $bc = 0$, $\sigma_l(A) = \{a, d\}$. Como

$$\mu_{*\lambda}(x) = x(a - \lambda) + (d - \lambda)x,$$

El rango de la diferencial solo puede ser 0 ó 4. $\operatorname{rang}(\mu_{*a}) = 4$ sii $a \neq d$. Luego $\operatorname{rang}(\mu_{*a}) = 0$ sii $\sigma_l(A) = \{a\}$.

Nota–. En [47], Janovská y Opfer consideran polinomios cuaterniónicos y muestran que hay varios tipos de ceros según el rango de ciertas matrices reales de orden 4×4, pero su procedimiento no parece tener ningún significado geométrico evidente.

Ejemplo 7.3.15. Una matriz con un solo autovalor para la que la diferencial de la función característica es de rango 2.

Sea $M = \begin{pmatrix} \mathbf{j} & 1 \\ \mathbf{k} & \mathbf{i} \end{pmatrix}$. Tomamos como función característica

$$\mu(\lambda) = -\lambda^2 + \lambda\mathbf{j} + \mathbf{i}\lambda.$$

La ecuación $\mu(\lambda) = 0$ tiene como única raíz $\lambda = 0$, es decir, $\sigma_l(M) = \{0\}$. La diferencial es $\mu_{*0}(x) = x\mathbf{j} + \mathbf{i}x$ con matriz asociada

$$\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Ejemplo 7.3.16. Una matriz con infinitos autovalores.

Sea $M = \begin{pmatrix} \mathbf{i} & 1 \\ -1 & \mathbf{i} \end{pmatrix}$. Entonces,

$$\begin{aligned} \mu(\lambda) &= \mathbf{i}\lambda + \lambda\mathbf{i} - \lambda^2, \\ \mu_{*\lambda}(x) &= x(\mathbf{i} - \lambda) + (\mathbf{i} - \lambda)x. \end{aligned}$$

En este caso $b^{-1}c = -1$, $b^{-1}(a-d) = 0 \in \mathbb{R}$ y $\Delta = -4 < 0$. El conjunto de sus autovalores por la izquierda está formado por una esfera centrada en $\mathbf{i}$ de radio 2.

$$\sigma_l(M) = \{\mathbf{i} + \frac{q}{2} : \Re(q) = 0, |q|^2 = 4\}.$$

Si fijamos $\lambda \in \sigma_l(M)$, $\lambda = \mathbf{i} + \frac{q}{2}$ con $q = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, la matriz real asociada a $\mu_{*\lambda}(x)$ es

$$\begin{pmatrix} 0 & -2q_1 & -2q_2 & -2q_3 \\ 2q_1 & 0 & 0 & 0 \\ 2q_2 & 0 & 0 & 0 \\ 2q_3 & 0 & 0 & 0 \end{pmatrix},$$

que es siempre de rango 2, ya que $q_1^2 + q_2^2 + q_3^2 = 4$.

Ejemplo 7.3.17. Una matriz cuya diferencial es de rango 4 y, por tanto, con dos autovalores.

Sea $M = \begin{pmatrix} \mathbf{j} & \mathbf{i} \\ \mathbf{i} & \mathbf{j} \end{pmatrix}$. Esta vez

$$\begin{aligned} \mu(\lambda) &= \mathbf{i} + (\mathbf{j} - \lambda)\mathbf{i}(\mathbf{j} - \lambda), \\ \mu_{*\lambda}(x) &= -x - x = -2x. \end{aligned}$$

Tenemos que el espectro por la izquierda es $\sigma_l(M) = \{\mathbf{i} + \mathbf{j}, -\mathbf{i} + \mathbf{j}\}$ y es inmediato que la diferencial es de rango 4.

7.3.4. Estudio topológico del caso 3×3

Para matrices de orden 3, al hacer la diferencial de la función característica 7.2.9, obtenemos de nuevo una ecuación lineal que podemos discutir con el método visto en 1.5.3. Por ahora no podemos establecer una clasificación completa pero el método nos permite hallar el rango de la diferencial $\mu_{*\lambda}$ para cada uno de los autovalores por la izquierda de una matriz cuaterniónica de orden tres arbitaria.

Una vez más estudiamos separadamente el caso en el que la matriz tenga alguna entrada nula fuera de la diagonal. Recordemos que en este caso la función característica es continua de modo que el comportamiento es análogo al de las matrices 2×2.

Caso polinómico

Comenzamos tomando matrices con $c = 0$, de modo que su función característica μ sea un polinomio de grado 3. Para este tipo de matrices podemos extender μ de manera continua a la esfera S^4.

Proposición 7.3.18. *La función μ definida en la Subsección 7.2.4 verifica que* $\lim\limits_{|\lambda|\to\infty} |\mu(\lambda)| = \infty$.

Demostración.

1. Si $b = h = 0$, $\mu(\lambda) = (r-\lambda)(g-\lambda)(a-\lambda)$ pero
$$\lim_{|\lambda|\to\infty} |r-\lambda| = \lim_{|\lambda|\to\infty} |g-\lambda| = \lim_{|\lambda|\to\infty} |a-\lambda| = \infty.$$

2. Si $b = 0, h \neq 0$, la función característica $\mu(\lambda) = \left(q - (r-\lambda)h^{-1}(g-\lambda)\right)(a-\lambda)$ se extiende de manera continua a la esfera por el resultado obtenido para el caso 2×2 (Prop. 7.3.8) y porque $\lim\limits_{|\lambda|\to\infty} |a-\lambda| = \infty$.

3. Por último, si $b \neq 0$, la expresión de μ es
$$\mu(\lambda) = p - qb^{-1}(a-\lambda) - (r-\lambda)h^{-1}f + (r-\lambda)h^{-1}(g-\lambda)b^{-1}(a-\lambda)$$
de modo que
$$\frac{|\mu(\lambda)|}{|\lambda|^2} \geq \frac{|(r-\lambda)h^{-1}(g-\lambda)b^{-1}(a-\lambda)|}{|\lambda|^2} - \frac{|-p+qb^{-1}(a-\lambda)+(r-\lambda)h^{-1}f|}{|\lambda|^2}.$$
Pero
$$\begin{aligned}\lim_{|\lambda|\to\infty} \frac{|(r-\lambda)h^{-1}(g-\lambda)b^{-1}(a-\lambda)|}{|\lambda|^2} &= \\ \lim_{|\lambda|\to\infty} \frac{|r-\lambda|}{|\lambda|}\frac{|g-\lambda|}{|\lambda|}|h^{-1}b^{-1}||a-\lambda| &= \infty\end{aligned}$$

mientras que el límite del sustraendo es cero. Por tanto, $\lim_{|\lambda|\to\infty} \frac{|\mu(\lambda)|}{|\lambda|^2} = \infty$ y así también $\lim_{|\lambda|\to\infty} |\mu(\lambda)| = \infty$.

$\square$

En este caso tendremos siempre grado topológico 3, por ser μ un polinomio con un único término de mayor grado, de modo que $\mu \sim \lambda^3$.

Linealización

De nuevo, calculamos la diferencial de la función característica para calcular su rango y poder estudiar si el 0 es o no un valor regular.

Proposición 7.3.19. *Para cualquier autovalor izquierda λ de una matriz del tipo* $A = \begin{pmatrix} a & b & 0 \\ f & g & h \\ p & q & r \end{pmatrix}$, *la diferencial de su función característica $\mu : \mathbb{H} \longrightarrow \mathbb{H}$ viene dada por*

1. *si $b, h = 0$,*

$$\mu_{*\lambda}(X) = -X(g-\lambda)(a-\lambda) - (r-\lambda)(g-\lambda)X - (r-\lambda)X(a-\lambda);$$

2. *si $b = 0, h \neq 0$,*

$$\mu_{*\lambda}(X) = Xh^{-1}(g-\lambda)(a-\lambda) - \left(q - (r-\lambda)h^{-1}(g-\lambda)\right) X + (r-\lambda)h^{-1}X(a-\lambda);$$

3. *en otro caso,*

$$\begin{aligned} \mu_{*\lambda}(X) &= \left(qb^{-1} - (r-\lambda)h^{-1}(g-\lambda)b^{-1}\right) X \\ &+ Xh^{-1}\left(f - (g-\lambda)b^{-1}(a-\lambda)\right) - (r-\lambda)h^{-1}Xb^{-1}(a-\lambda). \end{aligned}$$

La prueba es una aplicación directa de las reglas de derivación del producto y del inverso en el caso cuaterniónico vistas en la sección 7.3.2

La expresión obtenida es una ecuación $Px + xQ + RxS = 0$, cuyo rango podemos calcular según el método visto en la Subsección 1.5.3.

Ejemplo 7.3.20. Matrices triangulares $A = \begin{pmatrix} a & 0 & 0 \\ f & g & 0 \\ p & q & r \end{pmatrix}$.

Su espectro por la izquierda está formado por $\lambda_1 = a, \lambda_2 = g$ y $\lambda_3 = r$. La función característica es

$$\mu(\lambda) = (r-\lambda)(g-\lambda)(a-\lambda),$$

con diferencial

$$\mu_{*\lambda}(X) = -X(g-\lambda)(a-\lambda) - (r-\lambda)(g-\lambda)X - (r-\lambda)X(a-\lambda).$$

La diferencial en cada uno de los autovalores es o bien una traslación izquierda, o bien una traslación derecha, o bien una composición de ambas:

$$\begin{array}{rcl} \mu_{*a}(X) & = & L_{(a-r)(g-a)}(X), \\ \mu_{*g}(X) & = & L_{(g-r)} \circ R_{(a-g)}(X), \\ \mu_{*r}(X) & = & L_{(r-g)(a-r)}(X). \end{array}$$

Luego, en general, una matriz diagonal tiene tres autovalores distintos, todos ellos con diferencial de rango máximo. Si hay dos autovalores iguales y uno distinto, para el que se repite la diferencial es idénticamente nula y para el otro de rango máximo. Por último, si todos los elementos de la diagonal fuesen iguales, la diferencial sería de rango nulo para este único autovalor.

Nótese que, como se observa en este caso, a diferencia de lo que ocurre para matrices 2×2, no todos lo autovalores tienen por qué ser del mismo rango.

Nota–. Los siguientes ejemplos están calculados usando Mathematica.

Ejemplo 7.3.21. Una matriz con tres autovalores distintos, todos ellos de rango máximo. Sea $A = \begin{pmatrix} \mathbf{k} & 0 & 0 \\ 3\mathbf{i}-\mathbf{j} & -\mathbf{i} & \mathbf{i} \\ 1-2\mathbf{k} & \mathbf{j} & -\mathbf{j} \end{pmatrix}$. Su función característica es

$$\mu(\lambda) = (-1-\mathbf{k}+\lambda\mathbf{i})\lambda(\mathbf{k}-\lambda)$$

y por tanto $\sigma_l(A) = \{\mathbf{k}, 0, -\mathbf{i}-\mathbf{j}\}$. La diferencial de μ en cada uno de los autovalores es siempre de rango máximo,

$$\begin{array}{rcl} \mu_{*\mathbf{k}}(X) & = & (-1-\mathbf{i}+\mathbf{k})X; \\ \mu_{*0}(X) & = & -(1+\mathbf{k})X\mathbf{k}; \\ \mu_{*(-\mathbf{i}-\mathbf{j})}(X) & = & -X(-1-\mathbf{i}+\mathbf{j}) + X(\mathbf{i}+\mathbf{j}+\mathbf{k}). \end{array}$$

Ejemplo 7.3.22. Una matriz con dos autovalores, uno de ellos de rango nulo y otro de rango máximo. Sea $A = \begin{pmatrix} -\mathbf{i}-\mathbf{j} & 0 & 0 \\ \mathbf{k} & -\mathbf{i} & \mathbf{i} \\ 1-\mathbf{i} & \mathbf{j} & -\mathbf{j} \end{pmatrix}$. Esta vez,

$$\mu(\lambda) = (1 + k - \lambda\mathbf{i})\lambda(\mathbf{i} + \mathbf{j} + \lambda)$$

de modo que $\sigma_l(A) = \{0, -\mathbf{i}-\mathbf{j}\}$. En cuanto a la diferencial, para el autovalor nulo es de rango máximo y para $\lambda = -\mathbf{i}-\mathbf{j}$ se anula.

$$\begin{aligned} \mu_{*0}(X) &= (1+\mathbf{k})X(\mathbf{i}+\mathbf{j}); \\ \mu_{*(-\mathbf{i}-\mathbf{j})}(X) &= 0. \end{aligned}$$

Ejemplo 7.3.23. La matriz $A = \begin{pmatrix} -1 & 1 & 0 \\ \mathbf{j} & 0 & -1 \\ 1-\mathbf{j} & -1 & 1 \end{pmatrix}$. Para esta matriz, $\mu(\lambda) = -\lambda(2+\mathbf{j}) + \lambda^3$ y sus autovalores son exactamente tres, $\lambda_1 = 0$ y

$$\begin{aligned} \lambda_2 &= -\sqrt{-1+\sqrt{2}} - \sqrt{2(-1+\sqrt{2})} - (\sqrt{-1+\sqrt{2}})\mathbf{j} \\ \lambda_3 &= \sqrt{-1+\sqrt{2}} + \sqrt{2(-1+\sqrt{2})} + (\sqrt{-1+\sqrt{2}})\mathbf{j} \end{aligned}$$

La diferencial es

$$\begin{aligned} \mu_{*\lambda}(X) = X(-2-\mathbf{j}+\lambda^2) + \lambda^2 X + \lambda X \lambda &= \\ L_{\lambda^2} + R_{-2-\mathbf{j}+\lambda^2} + L_\lambda R_\lambda, & \end{aligned}$$

que en λ_1 es simplemente una traslación derecha, $\mu_{*0}(X) = -X(2+\mathbf{j})$ y para λ_2 y λ_3 tiene como matriz asociada

$$M = \begin{pmatrix} 5 & 0 & -5 & 0 \\ 0 & 7 & 0 & 1 \\ 5 & 0 & 5 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

que también es de rango máximo.

Ejemplo 7.3.24. Sea $A = \begin{pmatrix} \mathbf{j} & 1 & 0 \\ 2\mathbf{i} & -\mathbf{k} & 1 \\ 2-\mathbf{i}-2\mathbf{j} & -1-\mathbf{j}+\mathbf{k} & -\mathbf{i}-\mathbf{k} \end{pmatrix}$. La matriz A es del tipo $bc \neq 0$ y su función carcterística es

$$\mu(\lambda) = 2 - \mathbf{i} - 2\mathbf{j} + (1+\mathbf{j}-\mathbf{k})(\mathbf{j}-\lambda) + (\mathbf{i}+\mathbf{k}-\lambda)\left(2\mathbf{i} + (\mathbf{k}+\lambda)(\mathbf{j}-\lambda)\right).$$

Se comprueba que $\lambda = 0$ es un autovalor por la izquierda de A. En este punto, la expresión de la diferencial es $\mu_{*0}(X) = \mathbf{k}X + X\mathbf{i} + (\mathbf{i}+\mathbf{k})X\mathbf{j}$ con matriz real asociada de rango 3,

$$M = \begin{pmatrix} 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & -2 & 0 \end{pmatrix}.$$

Caso no polinómico

En el caso en el que ninguno de los elementos no diagonales se anule, la función característica de la Prop. 7.2.9 es racional con un punto de discontinuidad, el polo λ_0, de modo que no puede extenderse de manera continua a la esfera. Así, para poder hacer un estudio topológico análogo al del caso polinómico es necesario emplear la teoría del grado local (véase, por ejemplo el libro de K. Deimling [9]). En este caso nos limitamos simplemente a dar la expresión de la diferencial.

Proposición 7.3.25. *Para una matriz 3×3 con $c \neq 0$ la diferencial de la la función característica μ de la Prop. 7.2.9 es*

$$\begin{aligned} \mu_{*\lambda}(X) &= \\ X\left[-p + (r-\lambda)c^{-1}(a-\lambda) + (q-(r-\lambda)c^{-1}b)(\lambda_0-\lambda)^{-1}(f-hc^{-1}(a-\lambda))\right] & \\ +(\lambda_0-\lambda)Xc^{-1}(a-\lambda) - (\lambda_0-\lambda)Xc^{-1}b(\lambda_0-\lambda)^{-1}(f-hc^{-1}(a-\lambda)) & \\ -(\lambda_0-\lambda)(q-(r-\lambda)c^{-1}b)(\lambda_0-\lambda)^{-1}X(\lambda_0-\lambda)^{-1}(f-hc^{-1}(a-\lambda)) & \\ +\left[(\lambda_0-\lambda)(r-\lambda)c^{-1} - (\lambda_0-\lambda)(q-(r-\lambda)c^{-1}b)(\lambda_0-\lambda)^{-1}hc^{-1}\right]X. & \end{aligned}$$

7.4. Teorema de Cayley-Hamilton

Discutimos a continuación la posible extensión del teorema de Cayley-Hamilton para funciones características de autovalores por la izquierda de matrices cuaterniónicas. Para orden dos es sencillo pues se trata de un polinomio. Para $n = 3$, en general, μ será una función racional del tipo $P(\lambda) - Q(\lambda)(\lambda_0-\lambda)^{-1}F(\lambda)$ donde P, Q, F son polinomios, definida salvo en un punto de discontinuidad $\lambda = \lambda_0$. Comprobaremos que el teorema de Cayley-Hamilton se verifica en todos los casos.

Dos son los principales problemas con los que nos encontramos. El primero, es que en el caso conmutativo, tenemos polinomios característicos en variable compleja que se extienden de manera natural a un polinomio con una matriz como variable. Sin embargo, en el ámbito cuaterniónico, no hay una extensión obvia para una función característica cualquiera. Afortunadamente, hemos encontrado funciones características que son polinomios o funciones racionales, que se extienden de manera natural a una aplicación $\mu\colon \mathcal{M}_{n\times n}(\mathbb{H}) \to \mathcal{M}_{n\times n}(\mathbb{H})$.

El segundo problema, es que ninguna de las demostraciones conocidas para el caso conmutativo parece poder adaptarse a nuestro ámbito, de modo que hemos tenido que probarlo directamente.

Es interesante tratar de extender este teorema a $\mathcal{M}_{n\times n}(\mathbb{H})$ pues su múltiples aplicaciones facilitan mucho los cálculos de las funciones exponencial y logarítmica, tan útiles en el ámbito de la robótica [25, 85]. Por ejemplo, dada una matriz A de orden n sobre un cuerpo, el teorema de Cayley-Hamilton permite reducir un polinomio en A de cualquier grado a un polinomio de grado menor o igual que $n-1$; en general, cualquier función analítica puede expresarse en términos de polinomios de grado menor o igual que $n-1$, en particular, esto se verifica para la exponencial. En principio, estas aplicaciones no se extienden al ámbito cuaterniónico, ya que ni siquiera tenemos una "buena expresión" de una serie de potencias, pero podría estudiarse cómo extender la serie de la exponencial $\exp(At)$ ya que, en este caso, los coeficientes pueden obtenerse a partir de los autovalores de la matriz A [72].

Nota–. Para matrices 4×4 puede construirse de modo análogo una función característica racional. En el caso de que la matriz tenga dos entradas nulas fuera de la diagonal puede probarse que esta función verifica el teorema de Cayley-Hamilton. Para todos los ejemplos (arbitrarios) de matrices sin entradas nulas que hemos probado, hemos obtenido el mismo resultado pero no tenemos ninguna prueba genérica.

7.4.1. Caso $n = 2$

Hemos visto que para orden dos podemos encontrar una función característica polinómica $\mu(\lambda)$ para la cual es fácil probar que $\mu(A) = 0$.

Teorema 7.4.1. *Sea* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{H})$ *y* $\mu(\lambda) = c - (d-\lambda)b^{-1}(a-\lambda)$ *la función característica definida en el Teorema 7.2.6. Entonces* $\mu(A) = 0$.

Demostración. Tenemos que

$$\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} - \begin{pmatrix} d-a & -b \\ -c & 0 \end{pmatrix} \begin{pmatrix} b^{-1} & 0 \\ 0 & b^{-1} \end{pmatrix} \begin{pmatrix} 0 & -b \\ -c & a-d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \square$$

Nota–. Se sigue que

$$Ac^{-1}A = Ac^{-1}d + ac^{-1}A + (b - ac^{-1})I,$$

que es una generalización de la fórmula $A^2 = \mathrm{Tr}(A)A - (\det A)I$ en el ámbito conmutativo.

7.4.2. Orden tres. Caso polinomial

Como vimos anteriormente, para $n = 3$, si la matriz tiene algún cero fuera de la diagonal, también podemos escoger la función característica μ de modo que sea un polinomio cuaterniónico (bilateral) de grado 3. En este caso, un cálculo directo nos muestra que se verifica el teorema de Cayley-Hamilton.

Teorema 7.4.2. *Sea A una matriz cuaterniónica 3×3 con alguna entrada nula fuera de la diagonal. Entonces existe una función característica polinómica μ tal que $\mu(A) = 0$.*

Lo probaremos primero para las matrices con $c = 0$.

Proposición 7.4.3. *Sean $A = \begin{pmatrix} a & b & 0 \\ f & g & h \\ p & q & r \end{pmatrix}$ y $\mu(\lambda)$ la función característica definida en la Subsección 7.2.4. Entonces $\mu(A) = 0$.*

Demostración. Si $b, h = 0$, tomamos la fórmula (7.5), de modo que $\mu(A)$ es

$$\begin{pmatrix} r-a & 0 & 0 \\ -f & r-g & 0 \\ -p & -q & 0 \end{pmatrix} \begin{pmatrix} g-a & 0 & 0 \\ -f & 0 & 0 \\ -p & -q & g-r \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ -f & a-g & 0 \\ -p & -q & a-r \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Si $b = 0$, $h \neq 0$, tomamos la fórmula (7.6), y comprobamos que

$$\begin{pmatrix} r-a & 0 & 0 \\ -f & r-g & -h \\ -p & -q & 0 \end{pmatrix} h^{-1} \begin{pmatrix} g-a & 0 & 0 \\ -f & 0 & -h \\ -p & -q & (g-r) \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ -f & a-g & -h \\ -p & -q & a-r \end{pmatrix} = q \begin{pmatrix} 0 & 0 & 0 \\ -f & a-g & -h \\ -p & -q & a-r \end{pmatrix},$$

es decir,

$$(rI - A)h^{-1}(gI - A)(aI - A) = q(aI - A),$$

por tanto $\mu(A) = 0$.

Si $b \neq 0$, tenemos que

$$\operatorname{Sdet}(A - \lambda I) = \operatorname{Sdet} \begin{pmatrix} 0 & 0 & b \\ f - (g-\lambda)b^{-1}(a-\lambda) & h & g-\lambda \\ p - qb^{-1}(a-\lambda) & r-\lambda & q \end{pmatrix},$$

de manera que, por el Lema 1.2.4, estamos en el caso 2×2.

En primer lugar, asumamos que $h = 0$ y tomemos

$$\mu(\lambda) = (r-\lambda)\left(f - (g-\lambda)b^{-1}(a-\lambda)\right).$$

Comprobamos que

$$\begin{pmatrix} r-a & -b & 0 \\ -f & r-g & 0 \\ -p & -q & 0 \end{pmatrix} \begin{pmatrix} g-a & -b & 0 \\ -f & 0 & 0 \\ -p & -q & g-r \end{pmatrix} b^{-1} \begin{pmatrix} 0 & -b & 0 \\ -f & a-g & 0 \\ -p & -q & a-r \end{pmatrix} = \begin{pmatrix} r-a & -b & 0 \\ -f & r-g & 0 \\ -p & -q & 0 \end{pmatrix} f,$$

esto es,

$$(rI-A)(gI-A)b^{-1}(aI-A) = (rI-A)f,$$

por tanto, $\mu(A) = 0$.

Por otro lado, si $h \neq 0$ tomamos

$$\mu(\lambda) = p - qb^{-1}(a-\lambda) - (r-\lambda)h^{-1}\left(f - (g-\lambda)b^{-1}(a-\lambda)\right), \tag{7.10}$$

calculamos

$$pI - qb^{-1}(aI-A) - (rI-A)h^{-1}f =$$

$$\begin{pmatrix} p-(r-a)h^{-1}f & q+bh^{-1}f & 0 \\ qb^{1}f+fh^{-1}f & p-qb^{-1}(a-g)-(r-g)h^{-1}f & qb^{-1}h+f \\ qb^{-1}p+ph^{-1}f & qb^{-1}q+qh^{-1}f & p-qb^{-1}(a-r) \end{pmatrix}$$

y comprobamos que es igual a

$$-\begin{pmatrix} r-a & -b & 0 \\ -f & r-g & -h \\ -p & -q & 0 \end{pmatrix} h^{-1} \begin{pmatrix} g-a & -b & 0 \\ -f & 0 & -h \\ -p & -q & g-r \end{pmatrix} b^{-1} \begin{pmatrix} 0 & -b & 0 \\ -f & a-g & -h \\ -p & -q & a-r \end{pmatrix} = -(rI-A)h^{-1}(gI-A)b^{-1}(aI-A),$$

luego $\mu(A) = 0$. □

Lema 7.4.4. *Sea A una matriz cuaterniónica tal que $\mu(A) = 0$ para algún polinomio cuaterniónico $\mu(\lambda)$. Sea $B = PAP^{-1}$, con P una matriz real. Entonces $\mu(B) = 0$.*

Demostración. Basta observar que si $\nu(\lambda) = q_1\lambda q_2\lambda\cdots q_k\lambda q_{k+1}$ es un monomio, entonces $\nu(B) = P\nu(A)P^{-1}$. □

Nótese que se verifica el mismo resultado cuando $\mu(\lambda)$ es una función racional.

Permutando consecutivamente filas y columnas, se deduce el Teorema 7.4.2.

Ejemplo 7.4.5. Consideremos la matriz $A = \begin{pmatrix} 1 & \mathbf{i} & \mathbf{i} \\ \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & -1 & \mathbf{j} \end{pmatrix}$. Es similar vía una matriz real a $\begin{pmatrix} \mathbf{j} & -1 & 0 \\ \mathbf{k} & \mathbf{j} & \mathbf{i} \\ \mathbf{i} & \mathbf{i} & 1 \end{pmatrix}$, cuya función característica viene dada por la función (7.10), es decir,

$$\mu(\lambda) = \mathbf{i} + \mathbf{i}(\mathbf{j} - \lambda) + (1 - \lambda)\mathbf{i}\left(\mathbf{k} + (\mathbf{j} - \lambda)^2\right).$$

Entonces se verifica la siguiente ecuación:

$$A\mathbf{i}A^2 = A\mathbf{i}A\mathbf{j} + A\mathbf{k}A + \mathbf{i}A^2 - \mathbf{i}A\mathbf{j} + A(\mathbf{i} + \mathbf{j}) - (\mathbf{i} + \mathbf{k})A + (\mathbf{k} - \mathbf{j})I.$$

7.4.3. Orden tres. Caso no polinomial

Cuando todas las entradas fuera de la diagonal son no nulas, la función característica es una función racional con un polo (ver Proposición 7.2.9). Podemos extenderla de manera natural al espacio de matrices como sigue.

Sea $\lambda_0 = g - hc^{-1}b$ el polo de A. Sean

$$f_0 = f - hc^{-1}(a - \lambda_0),$$

$$q_0 = q - (r - \lambda_0)c^{-1}b.$$

Lema 7.4.6. *La matriz $\lambda_0 I - A$ es inversible si y solo si $f_0, q_0 \neq 0$.*

Demostración. Por la Proposición 7.2.7, $\mathrm{Sdet}(\lambda_0 I - A) = |c| \cdot |q_0 f_0|$. □

Definición 7.4.7. De acuerdo con la Proposición 7.2.9, definimos $\mu\colon \mathcal{M}_{n\times n}(\mathbb{H}) \to \mathcal{M}_{n\times n}(\mathbb{H})$ como sigue:

1. si $\lambda_0 I - B$ no es inversible, entonces $\mu(B) = q_0 f_0 I$;

2. en otro caso,

$$\begin{aligned} \mu(B) \;=\; & (\lambda_0 I - B)\left[\left(pI - (rI - B)c^{-1}(aI - B)\right) - \right. \\ & \left. \left(qI - (rI - B)c^{-1}b\right)(\lambda_0 - B)^{-1}\left(fI - hc^{-1}(aI - B)\right)\right]. \end{aligned}$$

Proposición 7.4.8. *La aplicación μ de la Definición 7.4.7 satisface el teorema de Cayley-Hamilton, es decir, $\mu(A) = 0$.*

Demostración. Si $\lambda_0 I - A$ no es inversible, entonces, $\mu(A) = q_0 f_0 I = 0$ por el Lema 7.4.6. En otro caso basta probar que

$$pI - (rI - A)c^{-1}(aI - A) \tag{7.11}$$

es igual a

$$\big(qI - (rI - A)c^{-1}b\big)\,(\lambda_0 I - A)^{-1}\left(fI - hc^{-1}(aI - A)\right). \tag{7.12}$$

Lema 7.4.9. *Un cálculo directo muestra que el primer término (7.11) es igual a*

$$\begin{bmatrix} -bc^{-1}f & -q + (r-a)c^{-1}b + bc^{-1}(a-g) & \vdots \\ (r-g)c^{-1}f - hc^{-1}p & p - fc^{-1}b - hc^{-1}q - (r-g)c^{-1}(a-g) & \vdots \\ -qc^{-1}f & -pc^{-1}b + qc^{-1}(a-g) & \vdots \end{bmatrix}$$

$$\begin{matrix} \vdots & -bc^{-1}h \\ \vdots & -f + (r-g)c^{-1}h + hc^{-1}(a-r) \\ \vdots & -qc^{-1}h \end{matrix}\Bigg] .$$

Calculemos ahora el término (7.12).

Comenzamos calculando $(\lambda_0 I - A)^{-1}$ por eliminación Gaussiana. Sea

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c^{-1}(\lambda_0 - a) & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c^{-1}b & 1 \end{pmatrix}.$$

Entonces

$$(\lambda_0 I - A)P_1 P_2 = \begin{pmatrix} 0 & 0 & -c \\ -f_0 & 0 & -h \\ -p_0 & -q_0 & \lambda_0 - r \end{pmatrix}, \tag{7.13}$$

donde

$$p_0 = p - (\lambda_0 - r)c^{-1}(\lambda_0 - a).$$

La inversa de la matriz $(\lambda_0 I - A)P_1 P_2$ en (7.13) es

$$B = \begin{pmatrix} f_0^{-1}hc^{-1} & -f_0^{-1} & 0 \\ -q_0^{-1}\left(p_0 f_0^{-1}hc^{-1} + (\lambda_0 - r)c^{-1}\right) & q_0^{-1}p_0 f_0^{-1} & -q_0^{-1} \\ -c^{-1} & 0 & 0 \end{pmatrix}.$$

Se sigue que $(\lambda_0 I - A)^{-1}$ es igual a

$$P_1P_2B = \left[\begin{array}{l} f_0^{-1}hc^{-1} \\ -q_0^{-1}\left(p_0f_0^{-1}hc^{-1} + (\lambda_0 - r)c^{-1}\right) \\ -c^{-1} + c^{-1}(\lambda_0 - a)f_0^{-1}hc^{-1} + c^{-1}bq_0^{-1}\left(p_0f_0^{-1}hc^{-1} + (\lambda_0 - r)c^{-1}\right) \end{array}\right. \begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array}$$

$$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \left.\begin{array}{rr} -f_0^{-1} & 0 \\ q_0^{-1}p_0f_0^{-1} & -q_0^{-1} \\ -c^{-1}(\lambda_0 - a)f_0^{-1} - c^{-1}bq_0^{-1}p_0f_0^{-1} & c^{-1}bq_0^{-1} \end{array}\right].$$

Más aún,

$$F = fI - hc^{-1}(aI - A) \;=\; \begin{pmatrix} f & hc^{-1}b & h \\ hc^{-1}f & f - hc^{-1}(a-g) & hc^{-1}h \\ hc^{-1}p & hc^{-1}q & f - hc^{-1}(a-r) \end{pmatrix},$$

mientras que

$$Q = qI - (rI - A)c^{-1}b \;=\; \begin{pmatrix} q - (r-a)c^{-1}b & bc^{-1}b & b \\ fc^{-1}b & q - (r-g)c^{-1}b & hc^{-1}b \\ pc^{-1}b & qc^{-1}b & q \end{pmatrix}.$$

Si ahora calculamos $(P_1P_2B)F$, tenemos que es igual a

$$\left[\begin{array}{l} 0 \\ -q_0^{-1}\left((\lambda_0 - r)c^{-1}f + hc^{-1}p\right) \\ c^{-1}\left(bq_0^{-1}\left((\lambda_0 - r)c^{-1}f + hc^{-1}p\right) - f\right) \end{array}\right. \begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array}$$

$$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \begin{array}{r} -1 \\ q_0^{-1}\left(p_0 - (\lambda_0 - r)c^{-1}hc^{-1}b - hc^{-1}q\right) \\ -c^{-1}bq_0^{-1}\left(p - (\lambda_0 - r)c^{-1}(g-a) - hc^{-1}q\right) + c^{-1}(a-g) \end{array} \begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array}$$

$$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \left.\begin{array}{r} 0 \\ -q_0^{-1}\left((\lambda_0 - r)c^{-1}h + f - hc^{-1}(a-r)\right) \\ -c^{-1}h + c^{-1}bq_0^{-1}\left((\lambda_0 - r)c^{-1}h + f - hc^{-1}(a-r)\right) \end{array}\right].$$

Finalmente, el (7.12) es $Q(P_1P_2BF)$, obtenemos así el resultado que queríamos. $\square$

Ejemplo 7.4.10. Sea $A = \begin{pmatrix} 1 & \mathbf{i} & -\mathbf{j} \\ \mathbf{i} & -1 & \mathbf{k} \\ 1 & -1 & \mathbf{j} \end{pmatrix}$. El polo es $\lambda_0 = -2$ y su imagen $\mu(\lambda_0) = -4 - 4\mathbf{i} + 8\mathbf{j}$. Para $\lambda \neq -2$, la función característica es

$$\mu(\lambda) = -(2+\lambda)\left(2 + \lambda(-1+\mathbf{j}) - \lambda\mathbf{j}\lambda + (-1+\mathbf{i}-\lambda\mathbf{k})(2+\lambda)^{-1}\mathbf{i}(2-\lambda)\right).$$

Con la notaciones de la demostración de la Proposición 7.4.8, tenemos que

$$(\lambda_0 I - A)^{-1} = (1/12)\begin{pmatrix} -3 & 3\mathbf{i} & 0 \\ 2\mathbf{i}-\mathbf{j}-\mathbf{k} & -8+2\mathbf{i}+\mathbf{j}+3\mathbf{k} & 2+2\mathbf{i}+4\mathbf{k} \\ 1+\mathbf{i}-\mathbf{j} & -3-\mathbf{i}+2\mathbf{j}+\mathbf{k} & -4+2\mathbf{j}+2\mathbf{k} \end{pmatrix},$$

$$P = \begin{pmatrix} -\mathbf{j} & 1-\mathbf{i}+3\mathbf{k} & 1 \\ -\mathbf{k} & 3-\mathbf{i}-3\mathbf{j} & -\mathbf{i} \\ -\mathbf{k} & -2\mathbf{j}+\mathbf{k} & \mathbf{i} \end{pmatrix},$$

$$Q = \begin{pmatrix} -1+\mathbf{i}-\mathbf{k} & \mathbf{j} & \mathbf{i} \\ \mathbf{j} & -1+\mathbf{i}+\mathbf{k} & 1 \\ -\mathbf{k} & \mathbf{k} & -1 \end{pmatrix}$$

y

$$F = \begin{pmatrix} \mathbf{i} & 1 & \mathbf{k} \\ 1 & 3\mathbf{i} & \mathbf{j} \\ -\mathbf{i} & \mathbf{i} & 2\mathbf{i}-\mathbf{k} \end{pmatrix}.$$

Se verifica que $P - Q(\lambda_0 I - A)^{-1}F = 0$.

Ejemplo 7.4.11. Si en la matriz del ejemplo polinomial 7.4.5 no consideramos ninguna transformación de filas y columnas, tenemos que la matriz $A = \begin{pmatrix} 1 & \mathbf{i} & \mathbf{i} \\ \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & -1 & \mathbf{j} \end{pmatrix}$ también verifica la ecuación

$$A^2\mathbf{i}A = -(1+\mathbf{i})I - \mathbf{i}A + A(1+\mathbf{k}) + (\mathbf{k}-\mathbf{j})A\mathbf{i} + A^2\mathbf{i} + (\mathbf{j}-\mathbf{k})A\mathbf{i}A - A\mathbf{k}A.$$

Además, considerando que los autovalores por la derecha de A son las clases de similitud de los autovalores de la forma compleja $c(A)$ y que ésta anula su propio polinomio característico también se da la siguiente igualdad.

$$-A^6 = (2-4\mathbf{i})I + (4-8\mathbf{i})A + (10-6\mathbf{i})A^2 - 8A^3 + (5+2\mathbf{i})A^4 - 2A^5.$$

7.4.4. Una nota final

Planteamos en esta última sección qué sentido tiene hablar del teorema de Cayley-Hamilton si escogemos un enfoque distinto de la definición de funciones características para autovalores por la izquierda.

En cuanto a la función (7.4) de la pag. 97, podría pensarse en enunciar el teorema de Cayley-Hamilton como $\sigma(X,Y)=0$, suponiendo que esto tenga algún sentido. Sin embargo, tenemos el siguiente contraejemplo para una matriz de orden 2.

Ejemplo 7.4.12. Sea $A=\begin{pmatrix}0 & \mathbf{i}\\ \mathbf{j} & 0\end{pmatrix}$. Sean $x=x_1+\mathbf{i}x_2$, $y=y_1+\mathbf{i}y_2$. Entonces

$$\sigma(x,y)=1+(x_1^2+x_2^2+y_1^2+y_2^2)^2-4x_2y_1.$$

Sin embargo, tenemos que $X=\begin{pmatrix}0 & \mathbf{i}\\ 0 & 0\end{pmatrix}$ e $Y=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$, de modo que $X_1=0$, $X_2=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$, $Y_1=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$ e $Y_2=0$. Entonces $\sigma(X,Y)=\begin{pmatrix}-3 & 0\\ 0 & 1\end{pmatrix}\neq 0$.

Entonces, aunque $\sigma(x,y)$ es igual al cuadrado de la norma de la función característica $\mu(\lambda)$ que satisface el teorema de Cayley-Hamilton, la extensión de σ a las matrices no está relacionada con la extensión de μ.

Según Gelfand *et alii*, las funciones características obtenidas a partir de los quasideterminantes también verifican el teorema de Cayley-Hamilton [27].

Teorema 7.4.13. *$\tilde{f}_{ij}(A)=0$ para todo $i,j=1,\ldots,n$, donde $\tilde{f}_{ij}$ se construye sustituyendo cada coeficiente a_{ij} por $\tilde{a}_{ij}=a_{ij}I$.*

Sin embargo, este teorema no es cierto para matrices que tengan alguna entrada nula, tanto para las f_{ij} que se pueden obtener directamente como para las que es necesario utilizar la Proposición 1.2.19.

Ejemplo 7.4.14. Se ve claramente con la matriz $B=\begin{pmatrix}\mathbf{i} & 0 & 0\\ \mathbf{k} & \mathbf{j} & 0\\ -3\mathbf{i} & 2\mathbf{k} & \mathbf{k}\end{pmatrix}$. El $(1,1)$-qdet, que involucra b_{23}^{-1}, es $|B-\lambda I|_{11}=\mathbf{i}-\lambda$ de modo que

$$f_{11}(B)=\begin{pmatrix}0 & 0 & 0\\ -\mathbf{k} & \mathbf{i}-\mathbf{j} & 0\\ 3\mathbf{i} & -2\mathbf{k} & \mathbf{i}-\mathbf{k}\end{pmatrix}\neq 0.$$

El $(1,2)$-qdet es $|B-\lambda I|_{12} = \mathbf{k} + (\mathbf{i}-\lambda)\mathbf{k}(\mathbf{j}-\lambda)$ de modo que

$$f_{12}(B) = \begin{pmatrix} \mathbf{k} & 0 & 0 \\ 0 & \mathbf{k} & 0 \\ 3+3\mathbf{i}-2\mathbf{k} & 2\mathbf{i}-2\mathbf{k} & 1+\mathbf{i}+\mathbf{j} \end{pmatrix} \neq 0.$$

Nota–. Aunque un polinomio con variable en $\mathbb{K}$ se extiende de modo natural a $\mathcal{M}_{n\times n}(\mathbb{K})$, hay que hacer la extensión con cuidado ya que una elección incorrecta puede dar lugar a demostraciones falsas. Por ejemplo, en el artículo [30] se obtiene una generalización del teorema de Cayley-Hamilton que podría servir para obtener Cayley-Hamilton en el caso cuaterniónico, pero esta generalización es falsa.

Se afirma que si A y B son dos matrices que conmutan se verifica la fórmula

$$(adj(B-A))\,(B-A) = \chi_A(B) = (B-A)adj(B-A), \tag{7.14}$$

donde $\chi_A(t)$ es el polinomio característico de A y $adj(A)$ la matriz adjunta de A (es decir, la matriz tal que en la entrada $a_{i,j}$ es el determinante de la matriz resultante de suprimir en A la fila i y la columna j).

A partir de esta fórmula se tiene que $Q = adj(tI-A)$ es un polinomio mónico de grado $n-1$ cuyos coeficientes son polinomios en la matriz A. De este modo, si B es una matriz que conmuta con A, entonces B conmuta con todos los coeficientes de Q y se obtiene el siguiente “teorema”.

> *“Teorema” Sobre un anillo conmutativo arbitrario, si las matrices A y B conmutan, entonces*
>
> $$\chi_A(B) \equiv 0 \quad \mathrm{mod} B-A$$
>
> *donde $\chi_A(t)$ es el polinomio característico de A.*

Sin embargo, los coeficientes del polinomio Q son coeficientes en los adjuntos de la matriz A y no en A y la fórmula (7.14) no es cierta; basta comprobarlo para las matrices $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ donde $\chi_A(t) = t^2 - 1$.

En cuanto a los autovalores por la derecha, recordemos que no son más que los cuaternios de la clase de similitud de los autovalores de la forma compleja $c(A) \in \mathcal{M}_{2n\times 2n}(\mathbb{C})$. A partir de la versión compleja del teorema de Cayley-Hamilton podemos enunciar el siguiente resultado.

Teorema 7.4.15 ([92])**.** *Sea $p(z) = \det(c(A) - zI) = \sum_{k=0}^{2n} c_k z^k$, $c_k \in \mathbb{R}$, el polinomio característico de $c(A)$. Entonces $p(A) = \sum_{k=0}^{2n} c_k A^k = 0$.*

Capítulo 8

Categoría de Lusternik-Schnirelmann

Un problema clásico en el estudio cualitativo de las ecuaciones diferenciales es relacionar la forma de las posibles soluciones con la complejidad topológica de la variedad subyacente. En este aspecto, uno de los primeros pasos fue estimar el número de puntos estacionarios para el caso particular del flujo de un gradiente, *i.e,* el menor número de puntos críticos de una función en la variedad. A finales de los años 20, Morse desarrolla estas ideas para un tipo concreto de funciones, aquellas cuyos puntos críticos son no degenerados, hoy conocidas como funciones de Morse. Es bien conocido que los números de Betti proporcionan una cota inferior para el número de puntos críticos de este tipo de funciones. En efecto, si f es una función de Morse en una variedad M, entonces el número de puntos críticos de índice r es siempre menor que β_r, donde $\beta_r = \dim H_r(M; A)$ y A es $\mathbb{R}, \mathbb{Z}$ ó $\mathbb{Z}_p$ con $p \neq 2$ primo [21].

Sin embargo, cuando permitimos que f sea un función diferenciable cualquiera, pudiendo tener puntos críticos degenerados, la situación es mucho más complicada. Este es el marco en el que L.A. Lusternik y L.G. Schnirelmann desarrollaron su invariante homotópico, que relaciona análisis, álgebra, geometría y topología [57]. De hecho, Lusternik y Schnirelmann utilizaron sus resultados para probar la existencia de tres geodésicas cerradas en la esfera. Es precisamente la categoría de Lusternik-Schnirelmann o categoría LS la que nos da una cota inferior para el número de puntos críticos de una función diferenciable arbitraria. Este resultado fue generalizado por M. Reeken [71] de la siguiente forma: sea M una variedad y f una función diferenciable en M, entonces $\operatorname{cat} M \leq \operatorname{cat} \Sigma(f)$, donde $\Sigma(f)$ es el conjunto de puntos críticos de f.

En su artículo de 1978, I.M. James recoge las principales propiedades de la categoría LS así como la definición de categoría de una aplicación [43]. Más tarde, en 1995, vuelve a dar las ideas más importantes sobre este invariante y su relación con otras variantes como la categoría seccional, la categoría punteada o la categoría polar [44]. La principal monografía sobre este tema escrita hasta el momento es la de O. Cornea, G. Lupton, J. Oprea y D. Tanré [8].

En 1941 R.H. Fox [22] redefinió este invariante homotópico para un espacio topológico X como el menor entero n tal que existen $n+1$ abiertos contráctiles en X que lo recubren. La categoría LS aporta mucha información topológica sobre la variedad pero es difícil de calcular. Un método usado con frecuencia para hallar la categoría de un espacio topológico dado es tratar de acotarla. Por ejemplo, como veremos en la Sección 8.2, la longitud del cup producto en cohomología nos da una cota inferior, mientras que la longitud de conos o la dimensión son cotas superiores. Hasta este momento el cálculo de la categoría sigue siendo, en general, un problema abierto que se aborda mediante el uso de invariantes homotópicos, acotaciones cohomológicas y métodos de geometría diferencial. La importancia del problema radica no solo en sí mismo, sino en que ha revitalizado temas clásicos y ha estimulado el desarrollo de herramientas cohomológicas muy potentes.

Asimismo, los avances en el campo de la categoría LS pueden aplicarse a la robótica. Para estudiar uno de los problemas centrales de la robótica, el cálculo de la complejidad topológica de un espacio de configuraciones X de un sistema dado, $\mathrm{TC}(X)$, se utilizan técnicas análogas a las de la teoría de Lusternik y Schnirelmann [16, 17].

Buscar explícitamente abiertos contráctiles que recubran suele ser una tarea ardua y en pocos casos se logra encontrar dichos abiertos directamente. Nuestra idea inicial para calcular la categoría LS de los grupos simplécticos consiste en considerar su representación matricial y aplicar el método clásico que desarrolló W. Singhof para $U(n)$. Lo que proponemos aquí es utilizar este método pero probando que los abiertos son contráctiles con la transformación de Cayley. Esto nos ha llevado a establecer una nueva técnica para el cálculo de la categoría LS. Como veremos a continuación, esta técnica nos permitirá calcular rápidamente la categoría LS del grupo unitario $U(n)$ y de los espacios simétricos compactos de tipo $\mathrm{AI}-SU(n)/SO(n)-$ y $\mathrm{AII}-SU(2n)/\,Sp(n)-$, proporcionando una demostración más sencilla que la dada recientemente por Mimura y Sugata [61].

Los abiertos categóricos en los que está definida la transformación de Cayley del grupo unitario $U(n)$ están asociados a autovalores en el sentido de que, dado un $z \in \mathbb{C}$ de módulo $|z| = 1$, una matriz $X \in \mathcal{M}_{n\times n}(\mathbb{C})$ está en $\Omega(z) = \Omega(z \cdot I)$ si y solo si $X - zI$ es inversible, es decir, z no es un autovalor de X. La dificultad en el caso cuaterniónico radica en que la condición para que $A \in \mathcal{M}_{n\times n}(\mathbb{H})$ esté en $\Omega(\lambda)$, $A-\lambda I$ inversible, es equivalente a que λ no esté en el espectro por la izquierda de A, pero no guarda ninguna relación con los autovalores por la derecha. Sin embargo, conforme al estudio establecido en la Sección 1.3, el único caso en el que los autovalores por la izquierda de una matriz cuaterniónica están completamente estudiados es para matrices 2×2. De ahí nuestro interés en caracterizar las matrices simplécticas 2×2 con infinitos autovalores (7.1.2). Este resultado nos permitirá probar que, aunque $\operatorname{cat} Sp(2) = 3$, hacen falta al menos cinco abiertos asociados a autovalores para recubrir el grupo.

Finalmente, nuestro estudio de las funciones de Bott-Morse nos permitirá acotar la categoría de $Sp(n)$ mediante las categorías relativas de las componentes conexas del conjunto de puntos críticos de cualquier función altura, generalizando así un resultado de M. Hunziker y M.R. Sepanski de 2009 [35].

8.1. Definición

Siguiendo a R.H. Fox [22] definimos la categoría utilizando conjuntos abiertos, aunque inicialmente se estableció esta teoría con subconjuntos cerrados.

Definición 8.1.1. Dado un espacio topológico 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 8.1.2. 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 $\operatorname{cat}(X)$.

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

Definición 8.1.3. Análogamente, se define la *categoría relativa de un subespacio* $A \subset X$, $\operatorname{cat}_X(A)$, como el menor entero n tal que existe $n+1$ abiertos contráctiles en X que cubren A.

Podemos afirmar que en un cierto sentido la categoría es subaditiva.

Proposición 8.1.4. *Si $X = Y \cup Z$ donde Y y Z son subespacios abiertos de X, entonces*

$$\operatorname{cat} X \leq \operatorname{cat} Y + \operatorname{cat} Z + 1.$$

Análogamente, si $X = \bigsqcup_\alpha X_\alpha$ donde X_α son las componentes conexas del espacio X localmente conexo por caminos, entonces

$$\operatorname{cat} X + 1 \leq \sum_\alpha (\operatorname{cat} X_\alpha + 1).$$

También se verifica que la categoría de un producto es menor o igual que la suma de las categorías [43].

Proposición 8.1.5. *Sean X e Y espacios topológicos conexos y paracompactos, entonces*

$$\operatorname{cat}(X \times Y) \leq \operatorname{cat} X + \operatorname{cat} Y.$$

Hay algunos ejemplos en los que la desigualdad es estricta. Ganea conjeturó la igualdad en esta relación si uno de los espacios era un esfera, es decir, $\operatorname{cat}(X \times S^n) = \operatorname{cat} X + 1$ [26]. Sin embargo, N. Iwase obtuvo un ejemplo de un espacio de categoría dos tal que al multiplicarlo por la esfera sigue teniendo categoría dos [36].

8.2. Acotaciones

Como comentamos al principio del presente capítulo, buscar explícitamente abiertos contráctiles que recubran una variedad no es tarea fácil. Las siguientes acotaciones son útiles para delimitar la categoría.

Proposición 8.2.1. *Para una variedad diferenciable M conexa,*

$$\operatorname{cat}(M) \leq \dim M.$$

Si trabajamos sobre grupos de Lie tenemos una cota superior en función de sus toros maximales.

Proposición 8.2.2. *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*

$$\operatorname{cat}(G) \leq (\operatorname{cat}(j) + 1)[\frac{1}{2}(\dim G - \text{rango } G) + 1] - 1,$$

donde, como j es la inclusión, $\operatorname{cat}(j) = \operatorname{cat}_G(T)$.

8.2.1. Una cota inferior

Dado un espacio X, para un anillo conmutativo R, la cohomología reducida es una R-álgebra graduada con el cup producto

$$H^*(X; R) \otimes H^*(X; R) \longrightarrow H^*(X; R).$$

Definición 8.2.3. La longitud del producto no trivial más largo en $H^*(X; \mathbb{K})$ es la *longitud del cup producto* (l.c.p.).

A partir de la cohomología podemos obtener información sobre la categoría; de hecho, la longitud del cup producto proporciona una cota inferior clásica para la categoría de un espacio.

Proposición 8.2.4. *La $\mathbb{K}$-longitud del cup producto de un espacio topológico X es menor o igual que la categoría de todo el espacio:*

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

Usando operaciones en cohomología de orden superior, P. Schweitzer y W. Singhof [74, 78] proporcionaron nuevos métodos con los que obtener todavía más información de las clases de cohomología. Un estudio más reciente es el que presenta J. Strom en [81].

En el caso de los grupos ortogonales puede verse, por ejemplo en [62, 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:

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

8.3. 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 8.3.1. *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,*

$$cat(M) + 1 \leq \#\Sigma(f).$$

El resultado ya citado de Reeken [71] relaciona la categoría de la variedad ambiente con la del conjunto crítico. Este resultado ha sido refinado por Y.B. Rudyak en [73] de la siguiente forma.

Proposición 8.3.2. *Sea f una función diferenciable en una variedad cerrada M, sea $\{a_0, \ldots, a_n\}$ el conjunto (finito) de niveles críticos de f y sea Σ el conjunto de puntos críticos de f. Entonces*

$$\operatorname{cat} M + 1 \leq \sum_{k=0}^{n} (\operatorname{cat}_M(\Sigma \cap f^{-1}(a_k)) + 1).$$

Nótese que la definición de categoría que usa Rudyak es la no reducida, por lo que hemos adaptado el enunciado.

Ejemplo 8.3.3. De acuerdo con el Ejemplo 5.1.6, podemos tomar en $G = U(2)$ la función altura h_X con $X = I$, de modo que

$$\operatorname{cat} G + 1 \leq 2(\operatorname{cat}_G(*) + 1) + (\operatorname{cat}_G(G_1^2) + 1),$$

es decir, $\operatorname{cat} U(2) \leq 2$ ya que $\operatorname{cat}_G G_1^2 = 0$.

En efecto, $G_1^2 \cong S^2$ es la órbita por la acción adjunta de $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, que está contenida en $\Omega_G(\mathbf{i})$.

8.4. Grupos de Lie y espacios homogéneos

El primero en la lista de problemas de la teoría de invariantes homotópicos numéricos de T. Ganea es “calcular la categoría de variedades familiares: variedades de Stiefel, grupos de Lie, etc.” [26]. La pregunta de Ganea es de 1970 y, sin embargo, aún no se ha podido responder completamente pues en este ámbito los avances son lentos y difíciles. La dificultad del cálculo directo ha tratado de paliarse introduciendo aproximaciones algebraicas como la categoría LS racional en homotopía racional.

En 1976, I. Berstein [3] estudió la longitud del cup producto de las Grassmannianas reales $O(n)/O(n-k) \times O(k)$. Para los grupos ortogonales $SO(n)$ solo se conoce la categoría de los de baja dimensión. I.M. James y W. Singhof [45] calcularon la categoría de $SO(5)$ y en [42] se llega hasta $SO(9)$. La de $SO(10)$ está calculada en [38] mientras que la categoría de $Spin(9)$ se ha obtenido en [39].

En 1975, W. Singhof [76] calculó la categoría de $SU(n)$, dando un recubrimiento explícito por n abiertos categóricos; se tiene entonces que $\operatorname{cat} U(n) = n$, ya que $U(n)$

es homeomorfo a $S^1 \times SU(n)$. Para obtener el recubrimiento considera abiertos asociados a autovalores del siguiente modo: dado $z \in \mathbb{C}$ un complejo unitario toma el abierto $\Omega(z)$ formados por la matrices complejas $n \times n$ tales que $A - zI$ es inversible. Utiliza la aplicación exponencial para probar que estos conjuntos son contráctiles. Además, dado que una matriz unitaria $n \times n$ no puede tener más de n autovalores distintos, obtiene que los abiertos $\Omega(z_1), \ldots, \Omega(z_{n+1})$ con todos los z_i distintos recubren el grupo. Utilizando la cota inferior cohomológica queda determinada la categoría de todos los grupos unitarios.

En el caso simpléctico, el problema se presenta aún más difícil. La categoría de $Sp(2)$ fue calculada por P. Schweitzer en 1965 [74]. Hubo que esperar al año 2002 hasta que L. Fernández-Suárez, A. Gómez-Tato, J. Strom y D. Tanré obtuvieron la categoría de $Sp(3)$ [20]. De manera independiente, N. Iwase y M. Mimura [40] llegaron al mismo resultado con técnicas distintas y probaron que para $n \geq 3$, $\operatorname{cat} Sp(n) \geq n + 2$ (recuérdese que Singhof había probado en [77] que para $n \geq 2$, $\operatorname{cat} Sp(n) \geq n+1$). El caso de $Sp(4)$ ha sido estudiado por J. Strom. Recientemente, M. Hunziker y M. Sepanski [35] han conjeturado que $\operatorname{cat} Sp(n) \leq \lfloor \frac{(n+2)^2}{4} \rfloor - 1$ puede ser una cota superior.

En cuanto a la categoría LS de los espacios homogéneos, el primero en dar algún resultado fue Singhof; en [77] establece la categoría de las variedades de Stiefel complejas, $\operatorname{cat} U(n)/U(n-m) = m+1$. Entre los resultados más recientes podemos citar los de J. Korbaš [53] para algunas Grasmannianas $O(n)/O(k) \times O(n-k)$, el trabajo de T. Fukaya [24] sobre $SO(n+3)/SO(n) \times SO(3)$ y el de T. Nishimoto [66] sobre la categoría LS de algunas variedades de Stiefel.

Para los espacios simétricos compactos es de destacar el trabajo de M. Mimura y K. Sugata [61], en el que prueban por el método de Singhof que $\operatorname{cat}(U(n)/O(n)) = n$ y $\operatorname{cat}(U(2n)/Sp(n)) = n$. Otros casos se han resuelto usando la longitud del cup producto, por ser variedades kahlerianas. La estructura celular de estos espacios fue estudiada en [49] por H. Kadzisa y M. Mimura. Recientemente, estos mismos autores han usado el gradiente de funciones de Bott-Morse para determinar la longitud en conos de algunos casos ya conocidos [50]. M. Mimura y K. Sugata recogen la categoría LS de todos los espacios simétricos Riemannianos irreducibles, compactos y simplemente conexos clásicos (excepto la de algunas grassmannianas y otros casos particulares) [61].

Tipo	Modelo Cartan	dim	LS cat
AI	$SU(n)/SO(n)$	$(n-1)(n+2)/2$	$n-1$
AII	$SU(2n)/Sp(n)$	$(n-1)(2n+1)$	$n-1$
AIII	$SU(p+q)/SU(p)\times SU(q)$	$2pq$	pq
BDI	$SO(p+q)/SO(p)\times SO(q)$	pq	?
DIII	$SO(2n)/U(n)\quad [n\geq 4]$	$n(n-1)$	$n(n-1)/2$
CI	$Sp(n)/U(n)\quad [n\geq 3]$	$n(n+1)$	$n(n+1)/2$
CII	$Sp(p+q)/Sp(p)\times Sp(q)$	$4pq$	pq

Para probar su resultado utilizan el siguiente teorema de Ganea.

Teorema 8.4.1. *Sea X un CW complejo $(r-1)$-conexo $(r\geq 1)$, entonces*

$$\operatorname{cat} X \leq \dim X/r.$$

Este teorema unido a la Proposición 8.4.2 permite a M. Mimura y K. Sugata obtener la categoría LS de los espacios simétricos que tengan una métrica Kähler.

Teorema 8.4.2. *Sea V una d-variedad compleja simplemente conexa tal que admite una métrica Kähler, entonces*

$$\operatorname{cat} V = d.$$

S. Helgason [31] probó que la métrica hermítica de un espacio simétrico hemítico es kähleriana. De modo que, como un espacio simétrico hermítico puede verse como una variedad compleja, obtienen la categoría de los espacios simétricos de tipo AIII, BDI (para $q\neq 2$), BDII (con $n\neq 2$), DIII y CI.

Ahora, únicamente falta calcular la categoría de la Grassmanniana simpléctica $Sp(p+q)/Sp(p)\times Sp(q)$. En este caso, la cota superior proporcionada por el Teorema 8.4.1 y la cota inferior cohomológica coinciden.

8.5. Categoría de $U(n)$: nueva demostración

En [76] W. Singhof utiliza la aplicación exponencial para probar que los abiertos $\Omega(z)$ formados por la matrices complejas $n\times n$ tales que $A-zI$ es inversible son contráctiles. Pasa del grupo al álgebra de Lie mediante la exponencial y una vez en el espacio vectorial escoge una contracción. Este método tiene el inconveniente de que es necesario elegir una rama de logaritmo para poder volver al grupo, lo que lleva consigo complicaciones técnicas que dificultan su extensión a otros grupos de Lie.

Con la transformación de Cayley tenemos en cada punto $A \in G = U(n)$ un abierto contráctil $\Omega_G(A)$ difeomorfo al álgebra de Lie $\mathfrak{u}(n)$ (ver Teorema 4.3.2). Es entonces sencillo encontrar un recubrimiento categórico explícito del grupo formado por $n+1$ abiertos.

Teorema 8.5.1. $\operatorname{cat} U(n) = n$.

Demostración. Sea $X \in U(n)$ una matriz unitaria y sea $z \cdot I \in U(n)$ la matriz diagonal $\operatorname{diag}(z, \dots, z)$, donde $z \in \mathbb{C}$ es un complejo arbitrario con $|z| = 1$. Recordemos que $X \in \Omega(z)$ si y solo si la matriz $zI + X$ es inversible.

Tomemos pues $n+1$ complejos diferentes $z_0, \dots, z_n$, con $|z_k| = 1$. Como cualquier matriz $X \in U(n)$ tiene a lo sumo n autovalores diferentes, entonces siempre podemos encontrar algún $-z_k$ que no sea autovalor de X, es decir, $X \in \Omega(z_k)$. Esto prueba que

$$U(n) = \bigcup_{k=0}^{n} \Omega_G(z_k),$$

por tanto

$$\operatorname{cat} U(n) \leq n.$$

Por otro lado, la cota cohomológica de [8, p. 273], nos dice que $n = \text{l.c.p.} \leq \operatorname{cat} U(n)$. □

8.6. Categoría de $U(n)/O(n)$

En el año 2008 M. Mimura y K. Sugata [61] probaron utilizando una adaptación del método de W. Singhof para $U(n)$ que

$$\operatorname{cat}(SU(n)/SO(n)) = n - 1.$$

8.6.1. Nueva demostración

Proponemos aquí un nuevo método para calcular la categoría de este espacios simétricos utilizando el modelo de Cartan y la extensión de la transformación de Cayley generalizada a este tipo de espacios que hicimos en la Sección 4.4.

Teorema 8.6.1. $\operatorname{cat} U(n)/O(n) = n$.

Demostración. Como vimos en la Sección 2.4, el modelo de Cartan de este espacio simétrico está formado por las matrices unitarias simétricas,

$$U(n)/O(n) \cong \{X \in U(n) \colon X = X^T\}.$$

Al considerar el modelo de Cartan M podemos considerar $U(n)/O(n)$ como una subvariedad de $U(n)$ de modo que los $\Omega(z)$ son abiertos. Sabemos que recubren M porque una matriz de $U(n)$ no puede tener más de n autovalores distintos y además, en el Corolario 4.4.3 obtuvimos que los abiertos del tipo $\Omega_M(A)$ son contráctiles.

Tenemos entonces que, dados $n+1$ complejos unitarios distintos $z_0, \ldots, z_n$, los abiertos $\Omega(z_0 \cdot I), \ldots, \Omega(z_n \cdot I)$ forman un recubrimiento categórico de $U(n)/O(n)$.

Recordemos que, en este caso, la cota cohomológica de la pag. 131 nos dice que

$$n \leq \operatorname{cat} M. \quad \square$$

8.7. Categoría de $U(2n)/\,Sp(n)$

De manera análoga al caso anterior, Mimura y Sugata probaron en [61] que

$$\operatorname{cat}(SU(2n)/\,Sp(n)) = n - 1.$$

Nota–. En el artículo que acabamos de citar, [61, Lema 1.2, pag 324], para poder construir abiertos asociados a autovalores prueban el siguiente Lema.

Lema 8.7.1. *$SU(2n)/Sp(n) = \{X \in SU(2n) : X^T = -JXJ^T\}$, donde $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$.*

Sin embargo, al construir el modelo de Cartan del espacio simétrico $SU(2n)/Sp(n)$ vemos que esto no es del todo cierto. Esto es importante porque no queda claro si, a pesar de ello, la contracción que construyen queda dentro del espacio simétrico.

En efecto, sea $G = SU(2n)$ y el automorfismo $\sigma\colon G \to G$ tal que, para cada $X \in G$ $\sigma(X) = -J\overline{X}J$. El conjunto de puntos fijos de σ es $Sp(n)$ de modo que

$$\begin{aligned} SU(2n)/\,Sp(n) &\cong \{Y \in SU(2n)\colon Y = X(-J\overline{X}J)^{-1}, X \in SU(2n)\} \\ &= \{Y \in SU(2n)\colon Y = -XJX^TJ, X \in SU(2n)\}, \end{aligned}$$

ahora bien, como MJ es difeomorfa a M tenemos que

$$SU(2n)/\,Sp(n) \cong \{Y \in SU(2n)\colon Y = XJX^T, X \in SU(2n)\}.$$

Por otro lado, la variedad N es

$$\begin{aligned} N &= \{X \in SU(2n)\colon -J\overline{X}J = X^{-1}\} \\ &= \{X \in SU(2n)\colon JX^T = XJ\}, \end{aligned}$$

que no es conexa, mientras que la variedad M sí. Basta verlo, por ejemplo, para $n = 1$; en este caso, las únicas matrices J-antisimétricas son $N = \{\pm I\}$ mientras que la variedad $M = N_e = \{I\}$. El problema que esto lleva consigo es que, en general, el conjunto de las matrices J-antisimétricas de $SU(2n)$ no tiene por qué ser conexo, luego, en principio, aún cuando cada componente conexa de los abiertos $\Omega_G(-z_r)$ sea contráctiles, no podemos afirmar que el propio $\Omega_G(-z_r)$ también lo sea.

Para ver que, en general, $SU(2n)/Sp(n) = \{X \in SU(2n) : X^T = -JXJ^T\}$ no tiene porqué ser conexo basta tomar, por ejemplo, $n = 1$. En ese caso se ve fácilmente que las únicas matrices J-antisimétricas son $N = \{\pm I\}$ mientras que la variedad $M = N_e = I$.

8.7.1. Nueva demostración

Teorema 8.7.2. $\operatorname{cat} U(2n)/\,Sp(n) = n.$

Demostración. Si utilizamos la forma compleja de una matriz cuaterniónica (Subsecc. 1.1.4) para identificar el grupo simpléctico $Sp(n)$ con el subgrupo de $U \in U(2n)$ formado por las matrices tales que $UJU^T = J$, donde $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, el modelo de Cartan nos dice que el espacio simétrico $U(2n)/Sp(n)$ es difeomorfo a

$$M \cong \{X \in U(2n) \colon X + X^T = 0\},$$

las matrices unitarias antisimétricas (ver Subsecc. 2.4).

Consideremos ahora la variedad

$$M' = JM = \{Y \in U(2n) \colon Y^T = -JYJ\}$$

que es difeomorfa a M porque $J^2 = I$.

Sea $Y \in M'$, y $z\lambda \in \mathbb{C}$ con $|\lambda| = 1$. Entonces $Y \in \Omega(z \cdot I)$ si y solo si $-z$ no es autovalor de Y. Pero, por el Lema 8.7.3, $Yv = -vz$ implica que $Y(J\bar{v}) = -(J\bar{v})z$, luego Y se puede diagonalizar como

$$Y = U \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} U^*, \qquad \text{donde } D = \operatorname{diag}(z_1, \ldots, z_n),$$

entonces

$$Y + zI = U \begin{pmatrix} D + zI & 0 \\ 0 & D + zI \end{pmatrix} U^*,$$

lo que nos muestra que el máximo número de autovalores diferentes de Y es n. Esto implica que si tomamos $n+1$ complejos diferentes $z_0, \ldots, z_n$, con $|z_i| = 1$, los abiertos $\Omega(z_k \cdot I)$ cubrirán M'.

Además, de nuevo por 4.4.3 tenemos que estos abiertos son contráctiles. Por tanto,

$$\operatorname{cat} U(2n)/Sp(n) \leq n.$$

Por otro lado [62, p. 149],

$$H(U(2n)/Sp(n)) = \Lambda(x_1, x_5, x_9, \ldots, x_{4n-3})$$

luego $n = \text{l.c.p.} \leq \operatorname{cat} U(2n)/Sp(n)$. □

Hemos usado la siguiente relación entre los autovectores de las matrices de M' que probaron Mimura y Sugata [61]:

Lema 8.7.3. *Sea $Y \in M'$. Si v es un z-autovector entonces $J\bar{z}$ también es un z-autovector, linealmente independiente del anterior.*

Demostración. $Yv = vz$ implica $v = Y^*vz$ implica $\bar{v} = Y^T\bar{v}\bar{z} = -JYJ\bar{v}\bar{z}$ por tanto $J\bar{v} = YJ\bar{v}\bar{z}$ y como $|z| = 1$ se sigue $J\bar{v}z = YJ\bar{v}$. □

8.8. El grupo simpléctico $Sp(2)$

La categoría LS de $Sp(2)$ fue calculada por primera vez por P. Schweitzer en [74]; obtuvo que $\operatorname{cat} Sp(2) = 3$ pero no da un recubrimiento explícito.

8.8.1. Un recubrimiento explícito mínimo

Utilizaremos ahora la teoría de Morse establecida en el Capítulo 5 para obtener de modo sencillo un recubrimiento explícito por cuatro abiertos categóricos del grupo simpléctico de orden 2. Este resultado completa la demostración abstracta dada por Schweitzer.

Consideremos los cuatro puntos críticos de la función altura h_X^G, con $G = Sp(2)$ y $X = \operatorname{diag}(1, 2)$. Por el Lema 3.3.3 tenemos que $\Sigma(h_I^G)$ está formado por la identidad $\pm I$ y las matrices $P = \operatorname{diag}(-1, 1)$ y $-P$.

Teorema 8.8.1. $\{\pm\Omega_G(I), \pm\Omega_G(P)\}$ *es un recubrimiento categórico de $G = Sp(2)$.*

Demostración. Recordemos que la condición de pertenencia $X \in \Omega_G(A)$ es que la matriz $A + X$ sea inversible.

En primer lugar, tenemos que la unión de los abiertos $\Omega_G(I)$ y $\Omega_G(-I)$ recubren todo el grupo excepto la órbita por la acción adjunta UPU^* de la matriz P. En efecto, al diagonalizar tenemos que esta órbita está formada por las matrices que tienen al 1 y al -1 como autovalores y es difeomorfa a las esfera S^4, de modo que puede cubrirse mediante la proyección estereográfica.

Veamos ahora que dada $X = (x_{ij}) = UPU^*$ en la órbita de P, o bien $P + X$ es inversible o bien lo es $-P + X$.

Como $P^2 = X^2 = I$ tenemos que

$$(P + X)^2 = 2I + PX + XP = 2\,\mathrm{diag}(1 - x_{11}, 1 + x_{22}).$$

Entonces $P + X$ es inversible si y solo si $x_{11} \neq 1$ y $x_{22} \neq -1$. Supongamos que $x_{11} = 1$. La condición $X^*X = I$ significa que las columnas de X forman una base ortonormal del espacio vectorial $\mathbb{H}^2$ para el producto hermítico $\langle v, w\rangle = v^*w$. Entonces, $x_{11} = 1$ implica que $x_{12} = x_{21} = 0$. Pero como X está en la órbita de P, entonces $x_{22} = -1$. Por tanto, $X = -P$. Se obtiene la misma conclusión para $x_{22} = -1$. Luego $\Omega_G(P)$ cubre toda la órbita de P, excepto $-P$. Pero como $-P \in \Omega_G(-P)$ obtenemos el recubrimiento que queríamos. □

Nota–. La cohomología del grupo simpléctico es [62, p. 119]

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

de manera que el producto no nulo más largo es $x_3 \wedge x_7 \wedge \cdots \wedge x_{4n-1}$ y por tanto l.c.p. $Sp(n) = n$. Sin embargo, W. Singhof [77] demostró que

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

Además, N. Iwase y M. Mimura [41] probaron que

$$\mathrm{cat}\, Sp(n) \geq n + 2 \text{ para } n \geq 3.$$

8.8.2. Extensión del método de Singhof

Para estudiar la categoría del grupo complejo $U(n)$ [76] y de los espacios simétricos $SU(2n)/Sp(n)$ y $SU(n)/SO(n)$ [61] se ha aplicado la técnica estándar que vimos para el caso de $U(n)$ iniciada por Singhof [76]. Consiste en considerar abiertos del tipo $\Omega(-zI)$ formados por las matrices A tales que $A - zI$ es inversible, para $z \in \mathbb{C}$ un complejo unimodular. Como hemos visto, se prueba que estos abiertos son contráctiles y recubren. Este recubrimiento junto con la cota cohomológica nos permite dar el valor exacto de la categoría LS de estos espacios.

En el contexto cuaterniónico, si queremos trabajar con este tipo de abiertos debemos considerar autovalores por la izquierda ya que la condición $A - \lambda I$ inversible, para $\lambda \in \mathbb{H}$, es equivalente a que λ no sea un autovalor por la izquierda de la matriz A. Como vimos en la Sección 4.4.2, los abiertos del tipo

$$\Omega(-\lambda) = \{A \in Sp(2) \colon A - \lambda I \text{ es inversible}\}$$

se pueden contraer mediante la transformación de Cayley generalizada.

Para los grupos simplécticos el método de W. Singhof ya no es tan efectivo ya que, por ejemplo, como veremos en este apartado, aunque la categoría de $Sp(2)$ es tres, cuatro abiertos de este tipo nunca recubren todo el grupo. De todas formas, sí podemos encontrar un recubrimiento categórico formado por cinco abiertos asociados a autovalores. La dificultad a la hora de extender este método al caso simpléctico radica en el peculiar comportamiento de los autovalores por la izquierda de matrices cuaterniónicas (ver Capítulo 7); de hecho, ya no se verifica que una matriz de orden n pueda tener a lo sumo n autovalores distintos (Teorema 1.3.6).

8.8.3. Cuatro abiertos asociados a autovalores no recubren

Dado $\lambda \in \mathbb{H}$ denotamos por $\Omega(\lambda) \subset Sp(2)$ al abierto $\Omega(\lambda I)$ formado por las matrices simplécticas que no tienen a $-\lambda$ como autovalor por la izquierda. Entonces, el Teorema 7.1.13 puede reformularse así.

Teorema 8.8.2. *Dados cuatro cuaternios $\lambda_0, \ldots, \lambda_3 \in \mathbb{H}$ con $|\lambda_s| = 1$, los cuatro abiertos $\Omega(\lambda_s)$ no forman nunca un recubrimiento de $Sp(2)$.*

Recordemos que las únicas matrices de $Sp(2)$ con infinitos autovalores por la izquierda son composición de una rotación con una traslación izquierda, $A = L_q \circ R_\theta$ con q un cuaternio unitario y $\theta \in \mathbb{R}$ tal que $\operatorname{sen}\theta \neq 0$ (ver Teorema 7.1.7). Para este tipo de matrices, según el Corolario 7.1.9, $\lambda \in \mathbb{H}$ es un autovalor por la izquierda de A si y solo si $|\lambda| = 1$, $\Re(\bar{q}\lambda) = \cos\theta$.

En la Sección 7.1.2 habíamos visto un ejemplo de dos matriz cuyo espectro por la izquierda contiene a $\{-1, \mathbf{i}, -\mathbf{i}, -\mathbf{k}\}$. Recogemos ahora el caso de dos matrices para la que $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ son autovalores por la izquierda.

Ejemplo 8.8.3. Tomamos $-\lambda_0 = 1, -\lambda_1 = \mathbf{i}, -\lambda_2 = \mathbf{j}, -\lambda_3 = \mathbf{k}$. Estos cuatro cuaternios independientes dan lugar al sistema

$$Iq = \cos\theta u,$$

luego

$$t = x = y = z = \cos\theta,$$

es decir,

$$q = \cos\theta(1 + \mathbf{i} + \mathbf{j} + \mathbf{k}).$$

Como $|q| = 1, \cos\theta = \pm 1/2$ luego $\theta = \pm\pi/3$ ó $\theta = \pm 2\pi/3$. Si denotamos por $\xi = 1 + \mathbf{i} + \mathbf{j} + \mathbf{k}$, las dos únicas matrices que quedan sin recubrir son entonces

$$A_1 = L_q \circ R_{\pi/3} = \frac{1}{4}\begin{pmatrix} \xi & -\sqrt{3}\xi \\ \sqrt{3}\xi & \xi \end{pmatrix}$$

y

$$A_2 = L_q \circ R_{2\pi/3} = \frac{1}{4}\begin{pmatrix} \xi & \sqrt{3}\xi \\ -\sqrt{3}\xi & \xi \end{pmatrix}.$$

Nota–. Obsérvese que tampoco podemos obtener un recubrimiento por cuatro abiertos categóricos de este tipo obteniendo alguno como unión disjunta de dos abiertos $\Omega(\lambda_i)$ distintos. En efecto, para cualesquiera $\lambda_1, \lambda_2 \in \mathbb{H}$, $|\lambda_1| = |\lambda_2| = 1$, $\Omega(\lambda_1) \cap \Omega(\lambda_2) \neq \emptyset$. Esto se debe a que siempre podemos escoger un tercer $\lambda_3 \in \mathbb{H}$ unitario no conjugado a $-\lambda_1$ ni a $-\lambda_2$ de manera que la matriz

$$A = \begin{pmatrix} \lambda_3 & 0 \\ 0 & \overline{\lambda_3} \end{pmatrix} \in Sp(2)$$

es tal que $A \in \Omega(\lambda_1) \cap \Omega(\lambda_2)$.

8.8.4. Un recubrimiento por cinco abiertos asociados a autovalores

En el apartado anterior queda claro que el método de W. Singhof no proporciona una cota superior mínima para la categoría LS de los grupos simplécticos. Sin embargo, la demostración de que cuatro abiertos no recubren nos ha llevado a encontrar cinco abiertos contráctiles en sí mismos que forman un recubrimiento categórico de $Sp(2)$. Este método resulta interesante pues, en la medida en que se vaya avanzando en el estudio de los autovalores cuaterniónicos por la izquierda para matrices de órdenes superiores, puede permitir encontrar recubrimientos explícitos de los grupos simplécticos.

Por la demostración del Teorema 7.1.13, tenemos que la unión de los abiertos $\Omega(-1), \Omega(-\mathbf{i}), \Omega(-\mathbf{j}), \Omega(-\mathbf{k})$ sólo deja fuera dos matrices que pueden ser recubiertas por $\Omega(-\lambda_4)$ donde λ_4 es el cuaternio unitario $\lambda_4 = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j})$.

En efecto, el ejemplo 8.8.3 vimos que las únicas matrices que quedan fuera son

$$\begin{aligned} A_1 &= L_q \circ R_{\pi/3}, \text{ con } q = (1/2)\xi \\ A_2 &= L_q \circ R_{2\pi/3}, \text{ con } q = -(1/2)\xi. \end{aligned}$$

Por el corolario 7.1.10 sabemos que los autovalores de estas matrices son todos los $\lambda \in \mathbb{H}$ de la forma $\lambda = q(\cos\theta + \operatorname{sen}\theta\omega)$ tales que $\overline{q}\lambda$ es conjugado a

$$\begin{aligned} \xi_1 &= \frac{1}{2} + \frac{\sqrt{3}}{2}\mathbf{i} \quad \text{para } A_1 \\ \xi_2 &= -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathbf{i} \quad \text{para } A_2. \end{aligned}$$

Tomamos pues $\lambda_4 = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j})$, de manera que $|\lambda_4| = 1$ y $\overline{q}\lambda_4$ no es conjugado ni a ξ_1 ni a ξ_2 pues $\Re(\overline{q}p) = \frac{1}{\sqrt{2}} \neq \pm\frac{1}{2}$. Hemos obtenido entonces que

$$Sp(2) = \bigcup_{i=0}^{4} \Omega(-\lambda_i),$$

donde cada uno de los cinco abiertos es contráctil.

8.9. Sobre la categoría de $Sp(n)$

8.9.1. Categoría LS y órbitas distinguidas

Recientemente, M. Hunziker y M.R. Sepankski [35] han probado que la categoría LS de un grupo de Lie simple, compacto, conexo y simplemente conexo G está acotada superiormente por la suma de las categorías relativas de ciertas clases de conjugación en G (determinadas por la acción del grupo de Weyl en un toro maximal). Para $G = SU(n)$ se obtiene el resultado de Singhof. Para el grupo simpléctico la desigualdad es

Teorema 8.9.1. *Sea $G = Sp(n)$. Entonces*

$$\operatorname{cat} G + 1 \leq \sum_{k=0}^{n} (\operatorname{cat}_G(\mathcal{O}_k) + 1)$$

donde $\mathcal{O}_k$, $0 \leq k \leq n$, son las órbitas por la acción adjunta de las matrices

$$E_k = \begin{pmatrix} -I_k & 0 \\ 0 & +I_{n-k} \end{pmatrix}.$$

Estas subvariedades son Grassmannianas

$$\frac{Sp(n)}{Sp(k) \times Sp(n-k)}.$$

Veamos que es posible dar una demostración más sencilla de este resultado. En primer lugar consideremos la función altura h_I en G. Como sabemos, sus puntos críticos son las matrices $A \in G$ que cumplen $A^2 = I$. Al diagonalizarlas observamos que sus autovalores deben ser ± 1, por lo que el conjunto de puntos críticos es exactamente

$$\Sigma(h_I) = \mathcal{O}_0 \sqcup \mathcal{O}_1 \sqcup \cdots \cup \mathcal{O}_n.$$

Entonces el Teorema 8.9.1 es un corolario inmediato del resultado de Rudyak citado en la Sección 8.3

Nota–. En una versión previa del artículo [35] se probaba que $\operatorname{cat}_G(\mathcal{O}_k) = \text{mín}\{k, n-k\}$. Sin embargo, la demostración contenía un error consistente en confundir los autovalores derecha con los izquierda, por lo que ahora esa igualdad es solamente una conjetura, sin demasiado apoyo, en opinión de N. Iwase (MR2546946 (2010j:55002)).

8.9.2. Categoría LS y componentes del conjunto crítico

Aplicando el resultado de Rudyak y nuestra descripción del conjunto crítico de las funciones de Morse (Teorema 5.1.5), tenemos que el teorema de Hunziker y Sepankski es un caso muy particular del siguiente

Teorema 8.9.2. *Sea $G = Sp(n)$ y $\Sigma = Sp(n_0) \times \Sigma(n_1) \times \cdots \times \Sigma(n_k)$ el conjunto crítico de la función altura h_X. Sean sus componentes conexas*

$$\Sigma[i_1, \ldots, i_k] = Sp(n_0) \times G^{n_k}_{i_k} \times \cdots \times G^{n_1}_{i_1}, \quad 0 \le i_j \le n_j,$$

donde G^p_q denota la grassmanniana $Sp(p)/(Sp(q) \times Sp(p-q))$. Entonces

$$\operatorname{cat} Sp(n) + 1 \le \sum (\operatorname{cat}_G \Sigma[i_1, \ldots, i_k] + 1).$$

Ejemplo 8.9.3. Tomando $n_0 = n - 1$ se obtiene

$$\operatorname{cat} Sp(n) \le 2 \operatorname{cat}_{Sp(n)} Sp(n-1) + 1.$$

Contando el número de *niveles* críticos de una función de Morse se obtiene

Teorema 8.9.4.

$$\operatorname{cat} Sp(n) \le \binom{n+1}{2}.$$

8.10. Estudio de algunos espacios simétricos kählerianos

8.10.1. Recubrimiento explícito de $Sp(2)/U(2)$

Las propiedades del determinante de Study (Secc. 1.2), la función característica obtenida para las matrices cuaterniónicas de orden dos (Secc. 7.2.3) y la expresión explícita de la matrices de $Sp(2)$ (Secc. 7.1.2) nos permiten dar un recubrimiento explícito mínimo del espacio $Sp(2)/U(2)$. Su categoría es cat $Sp(2)/U(2) =$ cat $(SO(5)/(SO(2) \times SO(3))) = 3$ y podemos encontrar cuatro abiertos contráctiles que lo recubren.

Sea M el modelo de Cartan del espacio $Sp(2)/U(2)$. Tomamos en M los abiertos $\Omega_M(\pm\mathbf{i}I)$ y $\Omega_M(\pm E)$, donde $E = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{pmatrix}$.

Proposición 8.10.1. $\Omega(\pm\mathbf{i}I), \Omega(\pm E)$ *recubren* $Sp(2)/U(2)$

Demostración. Dada una matriz $A \in M$, la condición $A \in \Omega_M(X)$ es equivalente a que $\mathrm{Sdet}(A + X) \neq 0$. Veamos si puede haber alguna matriz de $Sp(2)/U(2)$ que no esté en $\Omega(\pm\mathbf{i}I) \cup \Omega(\pm E)$

(a) Supongamos, en primer lugar, que A es una matriz diagonal, $A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$. Entonces,

$$\begin{aligned}
\mathrm{Sdet}^2(A + \mathbf{i}I) &= |(\alpha + \mathbf{i})(\beta + \mathbf{i})| \\
\mathrm{Sdet}^2(A - \mathbf{i}I) &= |(\alpha - \mathbf{i})(\beta - \mathbf{i})| \\
\mathrm{Sdet}^2(A + E) &= |(\alpha + \mathbf{i})(\beta - \mathbf{i})| \\
\mathrm{Sdet}^2(A - E) &= |(\alpha - \mathbf{i})(\beta + \mathbf{i})|
\end{aligned}$$

De modo que la única diagonal que deja sin cubrir $\Omega_M(\mathbf{i}I)$ es $-\mathbf{i}I$ y viceversa y la única diagonal que deja sin cubrir $\Omega_M(E)$ es $-E$ y al revés. Por tanto, quedan cubiertas todas las diagonales.

(b) Sea entonces A una matriz no diagonal, $A = \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \beta\bar{\alpha}\beta^{-1} \end{pmatrix}$. Esta vez,

$$\begin{array}{rcl} \mathrm{Sdet}^2(A+\mathbf{i}I) & = & |-\overline{\beta}| \cdot \left|\beta - (\beta\bar{\alpha}\beta^{-1} + \mathbf{i})(-\overline{\beta})^{-1}(\alpha+\mathbf{i})\right| \\ \mathrm{Sdet}^2(A-\mathbf{i}I) & = & |-\overline{\beta}| \cdot \left|\beta - (\beta\bar{\alpha}\beta^{-1} - \mathbf{i})(-\overline{\beta})^{-1}(\alpha-\mathbf{i})\right| \\ \mathrm{Sdet}^2(A+E) & = & |-\overline{\beta}| \cdot \left|\beta - (\beta\bar{\alpha}\beta^{-1} - \mathbf{i})(-\overline{\beta})^{-1}(\alpha+\mathbf{i})\right| \\ \mathrm{Sdet}^2(A-E) & = & |-\overline{\beta}| \cdot \left|\beta - (\beta\bar{\alpha}\beta^{-1} + \mathbf{i})(-\overline{\beta})^{-1}(\alpha-\mathbf{i})\right| \end{array}$$

Como todos tienen que anularse, prescindiendo del coeficiente no nulo $-\beta|\beta|^2$, tiene que verificarse

$$\begin{array}{rl} \beta|\alpha|^2 - \beta\overline{\alpha}\mathbf{i} + \mathbf{i}\beta\alpha - \mathbf{i}\beta\mathbf{i} & \\ \beta|\alpha|^2 + \beta\overline{\alpha}\mathbf{i} - \mathbf{i}\beta\alpha - \mathbf{i}\beta\mathbf{i} & = \\ \beta|\alpha|^2 - \beta\overline{\alpha}\mathbf{i} - \mathbf{i}\beta\alpha + \mathbf{i}\beta\mathbf{i} & = \\ \beta|\alpha|^2 + \beta\overline{\alpha}\mathbf{i} + \mathbf{i}\beta\alpha + \mathbf{i}\beta\mathbf{i} & = \end{array}$$

Pero esto es incompatible; en efecto, de la igualdad entre las dos primeras ecuaciones por un lado y las dos últimas por otro tenemos

$$\begin{array}{rcl} -\beta\bar{\alpha}\mathbf{i} + \mathbf{i}\beta\alpha & = & 0 \\ -\beta\bar{\alpha}\mathbf{i} - \mathbf{i}\beta\alpha & = & 0 \end{array}$$

luego $\mathbf{i}\beta\alpha = 0$, $\beta\bar{\alpha}\mathbf{i} = 0$, y como $\beta \neq 0$, $\alpha = 0$ y por tanto, sustituyendo en el Sdet tenemos $-\mathbf{i}\beta\mathbf{i} = \mathbf{i}\beta\mathbf{i}$, es decir, $\beta = 0$.

Además, la contracción de Cayley nos permite probar que estos abiertos son contráctiles (Corol. 4.4.3), luego

$$\mathrm{cat}(Sp(2)/U(2)) \leq 3. \quad \square$$

8.10.2. Recubrimiento de $Sp(3)/U(3)$

Análogamente al caso $Sp(2)/U(2)$ podemos dar un recubrimiento explícito del cociente $Sp(3)/U(3)$, aunque en este caso no es mínimo, ya que $\mathrm{cat}(Sp(3)/U(3)) = 6$ pero necesitamos ocho abiertos $\Omega_M(A)$ para cubrir todo $Sp(3)/U(3)$.

Teorema 8.10.2. *Los ocho abiertos contráctiles*

$$\Omega\left(\mathrm{diag}(\pm\mathbf{i}, \pm\mathbf{i}, \pm\mathbf{i})\right)$$

forman un recubrimiento de $Sp(3)/U(3)$.

Demostración. Utilizando el modelo de Cartan de este espacio simétrico podemos identificar $Sp(3)/U(3)$ con las matrices simplécticas de orden 3 antisimétricas. De esta condición, para

$$A = \begin{pmatrix} a & b & c \\ f & g & h \\ p & q & r \end{pmatrix} \in Sp(3)/U(3),$$

se deduce que la partes reales de los elementos de la diagonal se anulan,

$$\Re(a) = \Re(g) = \Re(r) = 0$$

y las relaciones

$$b = -\overline{f}, \quad c = -\overline{p}, \quad h = -\overline{q}$$

y

$$\begin{aligned} -ab + b\overline{g} + c\overline{h} &= 0, \\ -ac + bh + c\overline{r} &= 0, \\ -fc + gh + h\overline{r} &= 0. \end{aligned}$$

Además, al ser una matriz simpléctica, sabemos que las columnas (y las filas) son unitarias.

Es sencillo comprobar que cada $\Omega(E_r)$ cubre todas las matrices triangulares excepto las triangulares cuya diagonal sea como la de las Es restantes, con $s \neq r$, donde $E_r = \operatorname{diag}(\varepsilon_{r1}, \varepsilon_{r2}, \varepsilon_{r3})$ con $\varepsilon_{ri} = \pm 1$

En cuanto a las no triangulares, estudiamos separadamente cada caso.

(a) Cuando $c = 0$ y $b = 0$, según la expresión (7.6) de la Subsección 7.2.4, las ecuaciones que se obtienen para que una matriz $A \in Sp(3)/U(3)$ del tipo

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & g & h \\ 0 & -\overline{h} & r \end{pmatrix},$$

con $\Re(a) = \Re(g) = \Re(r) = 0$ no esté en ninguno de los abiertos $\Omega_M(E_r)$ son

$$\begin{array}{rcl}
\left(\overline{h} + (r - \mathbf{i})h^{-1}(g - \mathbf{i})\right)(a - \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r + \mathbf{i})h^{-1}(g + \mathbf{i})\right)(a + \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r - \mathbf{i})h^{-1}(g + \mathbf{i})\right)(a + \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r + \mathbf{i})h^{-1}(g - \mathbf{i})\right)(a + \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r + \mathbf{i})h^{-1}(g + \mathbf{i})\right)(a - \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r - \mathbf{i})h^{-1}(g - \mathbf{i})\right)(a + \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r - \mathbf{i})h^{-1}(g + \mathbf{i})\right)(a - \mathbf{i}) & = & 0 \\
\left(\overline{h} + (r + \mathbf{i})h^{-1}(g - \mathbf{i})\right)(a - \mathbf{i}) & = & 0
\end{array}$$

Estudiando estas ecuaciones se comprueba que, si $a = \mathbf{i}$ ó $a = -\mathbf{i}$ necesitamos seis abiertos $\Omega_M(E_r)$ para cubrir. Si $a \neq \pm\mathbf{i}$ pero $g = \pm\mathbf{i}$ ó $r = \pm\mathbf{i}$, entonces nos llegaría con cuatro y, por último, en el caso en el que ninguno de los elelmentos de la diagonal fuese $\pm\mathbf{i}$, solo dos abiertos bastarían para cubrir. Puede verse que para cubrir todas las matrices con $c = 0$ y $b = 0$ necesitamos los ocho $\Omega_M(E_r)$.

(b) Análogamente se estudia el caso $c = 0, b \neq 0$ (Subsecc. 7.2.4).

(c) En el caso genérico, $c \neq 0$, lo primero que vemos es qué ocurre cuando el polo $\lambda_0 = \pm\mathbf{i}$ y es autovalor. Es sencillo comprobar que bastan tres abiertos para cubrir este tipo de matrices. Cuando el polo $\lambda_0 \neq \pm\mathbf{i}$, siguiendo la expresión obtenida para la función característica en la Proposición 7.2.9, se obtienen ecuaciones análogas a las del caso (a) que, de nuevo, muestran que son necesarios los ocho abiertos. □

El siguiente ejemplo nos muestra que podemos mejorar la desigualdad anterior, aunque no de forma óptima.

Ejemplo 8.10.3. Tomamos una función altura h_X^G en $G = Sp(3)$ con X de manera que $n_0 = 1$, $n_1 = 2$. Se tiene

$$\operatorname{cat} Sp(3) + 1 \leq 2(\operatorname{cat} Sp(1) + 1) + (\operatorname{cat} Sp(1) \times G_1^2 + 1).$$

Como, haciendo el modelo de Cartan, $G_1^2 \cong S^4$ y $\operatorname{cat} S^3 = 1$, $\operatorname{cat} S^3 \times S^4 = 2$, se sigue que $\operatorname{cat} Sp(3) \leq 6$.

8.11. Otras aplicaciones

Ademas del cálculo de la categoría LS las construcciones que hemos desarrollado en la memoria pueden tener otras aplicaciones.

8.11.1. Cálculo de autovalores de matrices simplécticas mediante la transformación de Cayley

Recordemos que, según lo visto en la demostración del teorema 8.8.1, $Sp(2)$ puede ser recubierto por los cuatro abiertos contráctiles $\Omega(\pm I), \Omega(\pm P)$ donde $P = \mathrm{diag}(-1, 1)$. Esto nos permite calcular los posibles autovalores por la izquierda de las matrices de $Sp(2)$ de un modo más sencillo. Dada $A \in Sp(2)$, A estará en alguno de estos cuatro abiertos. Supongamos, por ejemplo, $A \in \Omega(I)$, entonces $A = c(X)$ para alguna $X \in \Omega(I)$ tal que $X + X^* = 0$.

Teorema 8.11.1. *Un cuaternio $\lambda \in \mathbb{H}$ es un autovalor por la izquierda de A con autovector asociado v sii $c(\lambda)$ es un autovalor por la izquierda de $X = c(A)$ con autovector asociado $(1 + \lambda)v$.*

Demostración. Como $A = c(X)$ la condición $Av = \lambda v$ es equivalente a

$$(I + X)^{-1}(I - X)v = \lambda v,$$

es decir,

$$(I - X)v = (I + X)\lambda v$$

luego

$$(1 - \lambda)v = X((1 + \lambda)v)$$

pero como $A \in \Omega(I)$, $1 + \lambda \neq 0$, luego

$$\begin{aligned} \frac{1-\lambda}{1+\lambda}(1 + \lambda)v &= X((1 + \lambda)v) \\ c(\lambda)((1 + \lambda)v) &= X((1 + \lambda)v). \quad \square \end{aligned}$$

Así, estudiar los autovalores de las matrices $A \in Sp(2) \cap \Omega(I)$ es equivalente a hacerlo para las $c(A) \in \mathfrak{sp}(2) \cap \Omega(I) = \mathfrak{sp}(2)$.

Autovalores de las matrices antihermíticas

Sea $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{H})$ antihermítica. De la condición $X + X^* = 0$ se tiene que $X = \begin{pmatrix} a & b \\ -\bar{b} & d \end{pmatrix}$ con $\Re(a) = \Re(d) = 0$.

Proposición 8.11.2. *Las matrices cuaterniónicas antihermíticas 2×2 con infinitos autovalores son de la forma*

$$X = \begin{pmatrix} \sigma & t \\ -t & \sigma \end{pmatrix} \qquad \text{con } t \in \mathbb{R},\ t \neq 0 \text{ y } \Re(\sigma) = 0.$$

Demostración. Para que una matriz de este tipo tenga infinitos autovalores no puede ser triangular y, según el Teorema 1.3.6 debe verificarse:

$$a_1 = b^{-1}(a-d) \in \mathbb{R} \tag{8.1}$$

$$a_0 = -b^{-1}(-\bar{b}) \in \mathbb{R} \tag{8.2}$$

$$\Delta = a_1^2 - 4a_0 < 0 \tag{8.3}$$

De la segunda condición se obtiene $b \in \mathbb{R}$ y $a_0 = 1$ o $\bar{b} = -b$ y $a_0 = -1$, pero este último caso no puede darse porque sino el discriminante sería $\Delta = a_1^2 + 4$, positivo siempre. Por (8.1) tenemos que $a = d$ y automáticamente se cumple (8.3). □

Además, la transformación de Cayley respeta las matrices con infinitos autovalores.

Proposición 8.11.3. *Sea $A \in Sp(2)$ de la forma $A = L_q R_\theta$, entonces la matriz antihermítica $X = c(A)$ también tiene infinitos autovalores.*

8.11.2. Descomposición polar generalizada

Es bien conocido el siguiente resultado (veáse por ejemplo el artículo de Dynnikov y Veselov [84]).

Teorema 8.11.4. *El flujo gradiente de la función altura h_X en el grupo G resuelve el problema de la descomposición polar de la matriz X.*

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

$$\Re Tr(HY) \leq \Re \operatorname{Tr}(H)$$

y entonces

$$\begin{aligned} h_X(A) = \Re \operatorname{Tr}(XA) = \Re \operatorname{Tr}(HUA) \leq \\ \Re \operatorname{Tr}(XU^*) = h_X(U^*) = \Re Tr(H), \end{aligned}$$

luego U^* es el máximo de la función h_X en G.

Nuestro interés en obtener una descomposición de este tipo en el ámbito de los espacios simétricos se debe a que, como vimos al final del Capítulo 6, pensamos que puede permitirnos obtener una caracterización de las funciones altura en el modelo de Cartan de un espacio simétrico que son de Morse.

La generalización de la descomposición a los espacios simétricos está íntimamente relacionada con la descomposición en subespacios que induce el automorfismo involutivo σ (ver Subsec. 2.2.2). La existencia y unicidad de esta descomposición polar generalizada en un entorno de la identidad depende del automorfismo elegido. Para determinados grupos de Lie la prueba Lawson en [56]; en [63] Munthe-Kaas *et al.* proponen una prueba alternativa.

Uno de los principales problemas que se plantea al trabajar con esta descomposición es cómo obtener los factores. Munthe-Kaas *et al.* dan una expresión explícita a partir de una serie de potencias para algunos casos [63].

Bibliografía

[1] Aslaksen, H. Quaternionic determinants. *Math. Intell.* **18** No.3, 57-65 (1996).

[2] Baker, A. Right eigenvalues for quaternionic matrices: A topological approach. *Linear Algebra Appl.* **286** No.1-3, 303-309 (1999).

[3] Berstein, I. On the Lusternik-Schnirelmann category of Grassmannians, *Math. Proc. Cambridge Philos. Soc.* **79** 129-134 (1976).

[4] Brenner, J. L. Matrices of quaternions. *Pac. J. Appl. Math.* **1** No. 3, 329-335 (1951).

[5] Cayley, A. Sur quelques propriétés des déterminants gauches. *J. Reine Angew. Math.* **32** 119-123 (1846).

[6] Chevalley, C. *Theory of Lie groups I.* Princeton Math. Ser., Vol. 8. Princeton University Press, 1946.

[7] Cohen, N.; De Leo, S. The quaternionic determinant. *Electron. J. Linear Algebra* **7** 100-111 (2000).

[8] Cornea, O.; Lupton, G.; Oprea, J.; Tanré, D. *Lusternik-Schnirelmann category.* Math. Surveys Monogr. **103**. Amer. Math. Soc., 2003.

[9] Deimling, K. *Nonlinear functional analysis.* Springer, Berlin, 1985.

[10] Dieudonné, J. Les déterminants sur un corps non-commutatif. *Bull. Soc. Math. France* **71** 27-45 (1943).

[11] Dieudonné, J. *A history of algebraic and differential topology 1900-1960.* Modern Birkhäuser Classics, Boston, 2009.

[12] Duan, H. Morse functions and Cohomology of Homogeneous Spaces, en Pragacz, Piotr (ed.), *Topics in cohomological studies of algebraic varieties.* Impanga lecture notes. Trends in Mathematics. Birkhäuser, 105-133 (2005).

[13] Duan, H. Multiplicative rule of Schubert classes. *Invent. Math.* **159** 407-436 (2005).

[14] Eilenberg, S.; Niven, I. The "fundamental theorem of algebra"for quaternions. *Bull. Amer. Math. Soc.* **50** No. 4, 246-248 (1944).

[15] Eilenberg, S.; Steenrod, N. *Foundations of algebraic topology.* Princeton Math. Ser., No.15, Princeton: University Press, XIV, 1952.

[16] Farber, M. Topological Complexity of motion planning *Discrete Comput. Geom.* **29** 211-221 (2003).

[17] Farber, M.; Grant, M. Robot motion planning, weights of cohomology classes, and cohomology operations. *Procc of the AMS* **136** No. 9, 3339-3349 (2008).

[18] Farber, M.; Grant, M Topological Complexity of configuration spaces. *Procc of the AMS* **137** No. 5, 1841-1847 (2009).

[19] Farenick, D. R.; Pidkowich, B. A. F. The spectral theorem in quaternions. *Linear Algebra Appl.* **371** 75-102 (2003).

[20] Fernández-Suárez, L.; Gómez-Tato, A.; Strom, J.; Tanré, D. The Lyusternik-Schnirelmann category of $Sp(3)$. *Proc. Amer. Math. Soc.* **132** No. 2, 587-595 (2004).

[21] Fomenko, A.T. *Differential Geometry and Topology.* Contemporary Soviet Mathematics, Consultants Bureau, 1987.

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

[23] Frankel, T. Critical submanifolds of the classical groups and Stiefel manifolds. *Differ. and Combinat. Topology, Sympos. Marston Morse*, Princeton, 37-53 (1963).

[24] Fukaya, T. Gröbner bases of oriented Grassmann manifolds *Homology Homotopy Appl.* **10** No. 2, 195-209 (2008).

[25] Gallier, J.; Xu, D. Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices. *Int. J. Robotics and Automotion* **17** No. 4 (2002).

[26] Ganea, T. Some Problems on Numerical Homotopy Invariants. *Lecture Notes in Math.* **249** 13-22 (1971).

[27] Gelfand, I.; Gelfand, S.; Retakh, V.; Lee Wilson, R. Quasideterminants. *Adv. Math.* **193** 56-141 (2005).

[28] Gelfand, I.M.; Retakh, V.S. A Theory of Noncommutative Determinants and Chracterisitic Functions of Graphs. *Funct. Anal. Appl* **26** No. 4, 1-20 (1992).

[29] Georgiev, G.; Ivanov, I.; Mihaylova, M.; Dinkova, T. An algorithm for solving a Sylvester quaternion equation. *Proc. Annual Conf. Rousse Univ. and Sc. Union* **48** No. 6.1, 35-39 (2009).

[30] Greenberg, M.J. Note on the Cayley-Hamilton theorem. *Amer. Math. Monthly* **91** No. 3, 193-195 (1984).

[31] Helgason, S. *Differential geometry, Lie groups and symmetric spaces.* Academic Press, 1978.

[32] Huang, L. On two questions about quaternion matrices. *Linear Algebra Appl.* **318** No.1-3, 79-86 (2000).

[33] Huang, L.; So, W. On left eigenvalues of a quaternionic matrix. *Linear Algebra Appl.* **323** No.1-3, 105-116 (2001).

[34] Huang, L.; So, W. Quadratic formulas for Quaternions *Appl. Math. Lett* **15** 533-540 (2002).

[35] Hunziker, M.; Sepanski, M.R. Distinguished orbits and the L-S category of simply connected compact Lie groups. *Topology Appl.* **156** 2443-2451 (2009).

[36] Iwase, N. Ganea's conjecture on Lusternik-Schnirelmann category. *Bull. London Math. Soc.* **30** No. 6, 623-634 (1998).

[37] Iwase, N. The Ganea conjecture and recent developments on Lusternik-Schnirelmann category [translation of Sugaku 56 (2004), 281-296], *Sugaku Expositions* **20** 43-63 (2007).

[38] Iwase,N.; Kikuchi, K.; Miyauchi, M. On the Lusternik-Schnirelmann category of SO(10), *arXiv:0712.3637v1.*

[39] Iwase, N.; Kono, A. Lusternik-Schnirelmann category of $Spin(9)$ *Trans. Amer. Math. Soc.* **359** 1517-1526 (2007).

[40] Iwase, N.; Mimura, M.; Nishimoto, T, On the cellular decomposition and the Lusternik-Schnirelmann category of $Spin(7)$ *Topology Appl.* **133** 1-14 (2003).

[41] Iwase, N.; Mimura, M. L-S categories of simply-connected compact simple Lie groups of low rank *Progress Math.* **215** 199-212 (2004).

[42] Iwase, N.; Mimura, M.; Nishimoto, T. Lusternik-Schnirelmann categories of non-simply connected compact simple Lie groups. *Topology Appl.* **150** No. 1-3, 111-123 (2005).

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

[44] James, I.M. Lusternik-Schnirelmann category, James, I. M. (ed.), *Handbook of algebraic topology.* North-Holland, 1293-1310 (1995).

[45] James, I.M.; Singhof, W. On the category of fibre bundles, Lie groups, and Frobenius maps, *Contemp. Math.* **227** 177-189 (1999)

[46] Janovská, D.; Opfer, G. Linear equations in quaternionic variables. *Mitt. Math. Ges. Hamburg.* **27** 223-234 (2008).

[47] Janovská, D.; Opfer, G. The classification and the computation of the zeros of quaternionic, two-sided polynomials. *Numer. Math.* **115** No. 1, 81-100 (2010).

[48] Johnson, R.E. On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring. *Bull. Amer. Math. Soc.* **50** No. 4, 202-207 (1944).

[49] Kadzisa, H.; Mimura, M. Cartan models and cellular decompositions of symmetric Riemannian spaces. *Topology Appl.* **156** 348-364 (2008).

[50] Kadzisa, H.; Mimura, M. Morse-Bott functions and the Lusternik-Schnirelmann category. *J. Fixed Point Theory Appl.* **10** No. 1, 63-85 (2011).

[51] Kyrchei, I. I. Determinantal representations of the Moore-Penrose inverse over the quaternion skew field and corresponding Cramer's rules. *Linear Multilinear Algebra* **59** No. 4 , 413-431 (2011).

[52] Kobayashi, S.; Nomizu, K. *Foundations of Differential Geometry,* vol. II. Wiley Classics Library Edition, 1996.

[53] Korbaš, J. Bounds for the cup-length of Poincaré spaces and their applications. *Topology Appl.* **153** 2976-2986 (2006).

[54] de Leo, S.; Ducati, G.; Leonardi, V. Zeros of unilateral quaternionic polynomials. *Electron. J. Linear Algebra* **15** 297-313 (2006).

[55] Liebeck, H. A proof of the equality of column and row rank of a matrix. *Amer. Math. Monthly* **73** No. 10,11-14 (1966).

[56] Lawson, J. D. Polar and Ol'shanskii decompositions. *J. Reine Angew. Math.* **448** 191-219 (1994).

[57] Lusternik, L. A.; Schnirelmann, L. G. *Méthodes Topologiques dans les Problèmes Variationnels.* Hermann (1934).

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

[59] Mare, A.L. On two results of Singhof. *Comment. Math. Univ. Carolin.* **38** No. 2, 379-383 (1997).

[60] Milnor, J.W. *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.

[61] Mimura, M.; Sugata, K. On the Lusternik-Schnirelmann category of symmetric spaces of classical type. Iwase, N. (ed.) et al., *Proceedings of the conference on groups, homotopy and configuration spaces*, University of Tokyo, Japan, July 5-11, 2005 in honor of the 60th birthday of Fred Cohen. Coventry: Geometry & Topology Publications. Geometry and Topology Monographs **13** 323-334 (2008).

[62] Mimura, M.; Toda, H. *Topology of Lie groups, I and II.* Translations of Mathematical Monographs **91**. American Mathematical Society, 1991.

[63] Munthe-Kaas, H. Z.; Quispel, G. R. W; Zanna, A. Generalized polar decompositions on Lie Groups with involutive automorphisms. *Found. Comput. Math.* **1** 297-324 (2001).

[64] Nanson, J. On certain determinant theorems. *J. Reine Angew. Math* **122** 179-185 (1900).

[65] Niven, I. Equations in quaternions. *Amer. Math.Monthly* **48** No. 10, 654-661 (1941).

[66] Nishimoto, T. On the LS category of Stiefel manifolds. *Topology Appl.* **154** No. 9, 1956-1960 (2007).

[67] Postnikov, M. *Lie Groups and Lie Algebras.* Lectures in Geometry, vol. 5. Nauka, Moscow, 1994.

[68] Pumplün, S.; Walcher, S. On the zeros of polynomials over quaternions. *Comm. Algebra*, **30** No. 8, 4007-4018 (2002).

[69] Ramanujan, S. An application of Morse theory to some homogeneous spaces. *Tohoku Math. J.* **21** No. 3, 344-353 (1969).

[70] Ramanujan, S. Morse theory of certain symmetric spaces. *J. Differential Geom.* **3** 213-229 (1969).

[71] Reeken, M. Stability of critical points under small perturbations. Part I: Topological theory. *Manuscripta Math.* **7** 387-411 (1972).

[72] Rowell, D. Computing the matrix exponential, the Cayley-Hamilton method. Departament of Mechanical Engineering, MIT (2004).

[73] Rudyak, Y. B. A remark on fixed point sets of gradient-like flows. *Arxiv:math/9908017v1* (1999).

[74] Schweitzer, P. Secondary cohomology operations induced by the diagonal mapping. *Topology* **3** 337-355 (1965).

[75] Serôdio, R.; Pereira, R.; Vitória, J. Computing the zeros of quaternion polynomials. *Comput. Appl. Math.* **42** (8-9) 1229-1237 (2001).

[76] Singhof, W. On the Lusternik-Schnirelmann category of Lie Groups I. *Math Z.* **145** 111-116 (1975).

[77] Singhof, W. On the Lusternik-Schnirelmann Category of Lie Groups II. *Math. Z.* **151** 143-148 (1976).

[78] Singhof, W. Generalized higher order cohomology operations induced by the diagonal mapping. *Math. Z.* **162** 7-26 (1978).

[79] So, W. Quaternionic left eigenvalue problem. *Southeast Asian Bull. Math.* **29** No. 3, 555-565 (2005).

[80] Steenrod, N. *The topology of fibre bundles.* Princeton University Press, 1951.

[81] Strom, J. Decomposition of the diagonal mapping. *Topology* **42** 349-364 (2003).

[82] Suzuki, T. Noncommutative spectral decomposition with quasideterminant. *Adv. Math.* **217** 2141-2158 (2008).

[83] Takeuchi, M. Cell decompositions and Morse inequalities on certain symmetric spaces. *J. Fac. Sci. Univ. Tokyo Sect. I* **12** 81-192 (1965).

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

[85] Visser, M.; Stramigioli, S.; Heemskerk, C, Cayley-Hamilton for roboticists. *Proceedings of the 2006 IEEE/RSJ* (2006).

[86] Volchenko, K. Y.; Kozachko, A. N. Integrable gradient flows on classic manifolds and Morse theory. *Moscow Univ. Math. Bull.* **52** No.3, 9-15 (1997); traducción de *Vestnik. Moskow. Univ. Ser. Mat. I,* **3** 9-15 (1997).

[87] Weyl, H. *The classical groups, their invariants and representations.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton (1997).

[88] Wiegmann, N.A. Some theorems on matrices with real quaternion elements, *Canad. J. Math.* **7** 191-201(1955).

[89] Wilson, R. L.; Retakh, V. Advanced course on Quasideterminants and Universal Localization *Quaderns* **41** Centre de Recerca Matemàtica, Barcelona (2007).

[90] Wood, R.M.W. Quaternionic eigenvalues. *Bull. Lond. Math. Soc.* **17** 137-138 (1985).

[91] Yokota, I. On the cellular decompositions of unitary groups. *Journal of the Institute of Polytechnics.* Osaka City University **7** No.1-2 Series A, 39-49 (1956).

[92] Zhang, F. Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251** 21-57 (1997).

[93] Zhang, F. Geršgorin type theorems for quaternionic matrices. *Linear Algebra Appl.* **424** No. 1, 139-153 (2007).