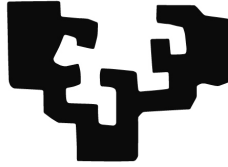


eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Facultad de Ciencia y Tecnología

Departamento de Matemáticas

Pseudoespectros de matrices

Tesis Doctoral

Gorka Armentia Galán

Bilbao, 2015

Memoria realizada por D. Gorka Armentia Galán bajo la dirección del catedrático Dr. D. Juan-Miguel Gracia Melero y del profesor titular Dr. D. Francisco-Enrique Velasco Angulo, ambos del área de Matemática Aplicada, para optar al grado de Doctor en Ciencias, Sección de Matemáticas.

Bilbao, 2015

A la memoria de mi madre y a mi padre.

A Feli.

Índice general

Índice general	I
Introducción	V
Objetivos	V
Hipótesis de partida	V
Resultados alcanzados	VI
Contenidos	VII
Antecedentes y notas históricas	IX
Agradecimientos	XIV
Índice de figuras	XV
Notaciones	XVII
1. Preliminares	1
1.1. Valores propios	1
1.2. Valores singulares, semejanza unitaria	1
1.2.1. Semejanza unitaria	3
1.2.2. Lema de Wielandt	4
1.3. Índice y número de condición de un valor propio	8
1.4. Distancia de Hausdorff	9
2. Pseudoespectros de primer orden	11
2.1. Definiciones	11
2.2. Propiedades de los pseudoespectros de primer orden	12
2.3. Componentes conexas	17
2.4. Conservación de la multiplicidad. Puntos de colisión	18
3. Pseudoespectros de orden superior	27
3.1. Pseudoespectros de orden k	27
3.2. Continuidad del argumento del máximo	29
3.3. Mínimo de la función de Malyshev	32

4. Derivada del área y del diámetro	37
4.1. Elementos de áreas y de diámetros	37
4.1.1. Variación del diámetro y del área de una región plana al someterla a una homotecia.	38
4.2. Ejemplos	40
4.3. Acotación mediante discos cerrados	45
4.4. Acotación mediante diámetros	45
4.5. Derivadas del diámetro	47
4.6. Acotación por medio de áreas	47
4.7. Derivadas del área	50
4.8. Área encerrada por una curva de nivel	52
4.9. Nota final. Radio medio	53
5. Pseudoespectros circulares	55
5.1. Una caracterización de las matrices normales	55
5.2. Bloques de Jordan generalizados	57
5.2.1. Pseudoespectro ordinario de primer orden	57
5.2.2. Pseudoespectro algebraico ordinario de segundo orden	60
5.3. Matrices normales	62
5.3.1. Pseudoespectros algebraicos de segundo orden	62
5.3.2. Pseudoespectros geométricos de cualquier orden	71
6. Pseudoespectros idénticos	73
6.1. Introducción	73
6.2. Órdenes de infinitésimos	73
6.3. Prueba del teorema principal	77
7. Fronteras de los pseudoespectros	83
7.1. Resultados principales	83
7.2. Conjuntos semialgebraicos. Funciones semialgebraicas	84
7.3. Resultados auxiliares	87
7.4. Prueba de los Teoremas principales	90
7.4.1. Prueba del Teorema 7.1.1	90
7.4.2. Prueba del Teorema 7.1.2	92
7.5. Corroboración mediante derivadas	103
7.6. Corroboración mediante curvas algebraicas	110
7.7. Alternativa usando una fórmula de Lippert	111
7.8. Más sobre el caso $(x_0, y_0, \varepsilon_0) = (0, 0, 0)$	112
8. Matrices de orden pequeño	115
8.1. Reducción de matrices 2×2	116
8.2. Reducción de matrices 3×3	124
8.2.1. Caso de un valor propio triple.	129
8.2.2. Caso de un valor propio doble y uno simple	132
8.2.3. Caso de tres valores propios distintos	135

A. Curvas algebraicas reales	139
A.1. Introducción	139
A.2. Caracterización de los puntos de autointersección	140
A.2.1. Aplicaciones afines	140
A.2.2. Número de intersección. Recta tangente	141
A.2.3. Determinación de $m_p(f)$	142
A.2.4. Tangentes en un punto múltiple	143
A.2.5. Clasificación de puntos dobles	144
A.3. Series formales de potencias	150
A.3.1. Sustitución de una serie en otra	152
A.3.2. Series de Laurent y de Puiseux	154
A.4. Ramas de una curva algebraica	158
A.5. Naturaleza de los puntos dobles	159
A.6. Maple	163
A.7. Intersección de curvas	165
A.7.1. Multiplicidad de intersección de una rama con una curva	166
A.7.2. Tangente a una rama y multiplicidad de una rama	167
A.7.2.1. Tacnodo tangente a dos circunferencias	169
B. Conexión	173
B.1. Espacios topológicos	173
B.2. Conexión	175
B.2.1. Componentes conexas de un espacio	176
B.2.2. Componentes conexas de un subconjunto	177
B.3. Conexión por arcos	177
B.3.1. Conexión en \mathbb{R}^n	178
C. Miscelánea algebraica	179
C.1. Resultante de Sylvester	179
C.2. Inversa de Moore-Penrose	180
C.3. Álgebras de conjuntos	181
Epílogo	183
Bibliografía	187
Índice alfabético	193

Introducción

Queremos destacar aquí, de forma breve, los objetivos, las hipótesis de partida y los resultados obtenidos. Más tarde, tanto en el resto de esta Introducción como en el Epílogo (a modo de Problemas Abiertos), daremos más detalles.

Objetivos

Esta tesis se propuso, entre otros, los objetivos siguientes.

1. Prueba de las relaciones de entrelazamiento de los pseudoespectros de órdenes de multiplicidad sucesivos de matrices complejas cuadradas que ocurren en los puntos de coalescencia (o colisión) de las componentes conexas de los pseudoespectros.
2. Dentro del conjunto de matrices que tienen pseudoespectros formados por círculos o unión de círculos, ampliar el estudio de Karow de los pseudoespectros algebraicos ordinarios de primer orden para los bloques generalizados de Jordan (véase [44]) a los pseudoespectros algebraicos ordinarios de primer y segundo orden para matrices bidiagonales con elementos cualesquiera en la superdiagonal primera.
3. Análisis de la naturaleza geométrica de la curva frontera de los pseudoespectros en torno a los puntos de coalescencia como puntos singulares de una curva algebraica real.
4. Caracterización del número de condición de un valor propio en términos de las derivadas del área y diámetro de la componente conexa del ε -pseudoespectro de primer orden en torno al valor propio.
5. Probar que los pseudoespectros algebraicos de segundo orden y los geométricos de cualquier orden son conjuntos semialgebraicos de \mathbb{R}^2 . Averiguar si la función $a : [0, \varepsilon) \rightarrow \mathbb{R}$, dada por el área de una componente conexa del ε -pseudoespectro de primer orden es una función semialgebraica.

Hipótesis de partida

Motivaron nuestro interés por estos problemas algunos trabajos previos. En primer lugar, el artículo de A.N. Malyshev (1999) [54] fue clave. Dados una

matriz $A \in \mathbb{C}^{n \times n}$ y un número complejo z Malyshev demostró una fórmula explícita para el cálculo de la distancia, $h_2(z)$, (respecto de la norma espectral) de A al conjunto de matrices que tenían a z como valor propio de multiplicidad algebraica ≥ 2 . Esta fórmula relacionaba dicha distancia con el máximo respecto de $t \in \mathbb{R}$ del valor singular σ_{2n-1} de una matriz $2n \times 2n$. Esto dió una solución al Problema de Wilkinson: Si A es una matriz simple, hallar la mínima distancia de A al conjunto de matrices que tienen un valor propio múltiple. Aparece así una función $h_2: \mathbb{C} \equiv \mathbb{R}^2 \rightarrow \mathbb{R}$, a la que hemos llamado función de Malyshev de la matriz A . Esta función comparte muchas propiedades con la función $h_1(z) := \sigma_n(zI_n - A)$.

Otra solución al problema de Wilkinson fue dada por Alam y Bora (2005) mediante el primer punto de colisión de los pseudoespectros de A . Su demostración es más bien intuitiva, aunque muchos de sus puntos dificultosos fueron aclarados por Burke, Lewis y Overton.

Resultados alcanzados

1. Respecto del Objetivo 1, se ha establecido el entrelazamiento existente entre los pseudoespectros de órdenes uno y dos en un primer punto de colisión de los pseudoespectros de una matriz: Teorema 3.3.4. También, se ha dado una condición suficiente para dicho entrelazamiento en un segundo punto de colisión: Teorema 3.3.6. No se ha conseguido probarlo para un punto de colisión cualquiera. Tampoco se ha podido demostrar para órdenes de multiplicidad de órdenes k y $k + 1$ con $k \geq 2$.
2. En relación al Objetivo 1, sabemos que $h_1(z) \leq h_2(z)$ para todo $z \in \mathbb{C}$, hemos demostrado que ambas funciones alcanzan el mismo valor en los puntos de colisión y en los puntos aislados de la frontera de los pseudoespectros de primer orden, véanse las Propositiones 3.3.1 y 7.4.9. No ha sido probada la conjetura $h_k(z) \leq h_{k+1}(z)$, para $k \geq 2$.
3. En relación con el Objetivo 2 se ha conseguido caracterizar los pseudoespectros de tales matrices, siendo un círculo en cada caso, como se demuestra en los Teoremas 5.2.1 y 5.2.2.
4. Hemos conseguido probar que la frontera de los pseudoespectros algebraicos estrictos puede diferir de la frontera de los pseudoespectros algebraicos ordinarios en un conjunto finito de puntos, para algunos niveles de ε , como se demuestra en el Teorema 7.1.2. Además, en él, estos puntos se caracterizan por ser máximos locales estrictos para la función h_1 .
5. Se ha indicado una demostración escueta de que los pseudoespectros geométricos de todo orden, y los algebraicos de segundo orden, son conjuntos semialgebraicos de \mathbb{R}^2 . Pero, mientras no haya una demostración completa de la caracterización de los pseudoespectros algebraicos de órdenes tres y superiores como subconjuntos de nivel de funciones de tipo Malyshev (Ikramov-Nazari, Mengi), no se podrá utilizar este abordaje para probar

su semialgebraicidad. No hemos podido demostrar que la función área, a , sea semialgebraica.

6. Con relación al Objetivo 4, se presentan unos teoremas que permiten caracterizar los números de condición en función de la primera derivada respecto de ε , cuando $\varepsilon = 0$, tanto en el caso de los diámetros como de las áreas. Sin embargo, la existencia de la segunda derivada del área a respecto de ε , cuando $\varepsilon = 0$, está sometida a la demostración de que el área a sea una función semialgebraica.

Contenidos

Esta memoria trata de pseudoespectros de matrices complejas cuadradas. Está dividida en ocho capítulos y tres apéndices.

En el Capítulo 1 se introducen algunas definiciones y notaciones. En particular, se incide en la semejanza unitaria, ya que los pseudoespectros son invariantes frente a estas transformaciones; el Lema de Wielandt, de especial interés en el Capítulo 7; índice y número de condición de un valor propio, los cuales se utilizarán en el Capítulo 4; finalmente, se define la distancia de Hausdorff entre dos conjuntos compactos.

En los dos siguientes capítulos se presentan resultados sobre pseudoespectros de primer orden y de orden superior, respectivamente. En el Capítulo 2 se introducen las definiciones de pseudoespectros ordinarios y estrictos de primer orden. En una segunda subsección, mostramos que la geometría de los esos pseudoespectros es invariante bajo transformaciones de semejanza unitaria y afines. Como veremos, los pseudoespectros son conjuntos de puntos del plano real que están formados por componentes conexas. Centramos nuestro interés en los puntos donde coalescen dos de esas componentes conexas, ya que esos puntos dan una solución al problema de Wilkinson sobre la distancia de una matriz al conjunto de matrices que tienen un valor propio múltiple.

En el Capítulo 3, se introducen los pseudoespectros algebraicos y geométricos, tanto ordinarios como estrictos, de orden superior. En particular nos centraremos en los pseudoespectros algebraicos ordinarios de orden 2, los cuales se pueden limitar por los conjuntos de nivel de una función que llamaremos función de Malyshev. Se proporcionan algunas propiedades de esa función concernientes a su continuidad y derivabilidad, así como otras propiedades de los pseudoespectros algebraicos de segundo orden que involucran a la función de Malyshev.

El Capítulo 4 está dedicado al estudio de la relación entre la componente conexa $\mathcal{K}_\lambda(\varepsilon)$ de un ε -pseudoespectro que rodea a un valor propio λ y el número de condición de este valor propio. Dadas las funciones geométricas radio medio, diámetro y área de una de tales componentes conexas, se estudian las derivadas de primer y segundo orden por la derecha en $\varepsilon = 0$. Los resultados reposan sobre un teorema de M. Karow en que halla una especie de “derivada” por la derecha de la función compacto-valorada $\varepsilon \mapsto \mathcal{K}_\lambda(\varepsilon)$ en $\varepsilon = 0$ respecto de la distancia

de Hausdorff entre compactos no vacíos de \mathbb{C} .

El Capítulo 5 tiene su inspiración en algunos de los resultados expuestos por M. Karow en su Tesis. Es sabido que una matriz es normal si y sólo si el pseudoespectro ordinario de primer orden de nivel ε está formado por la unión de círculos de radio ε para todo $\varepsilon \geq 0$. En este capítulo se estudian el pseudoespectro algebraico de segundo orden, que también está formado por la unión de círculos cuyo radio depende de ε , y el pseudoespectro geométrico de cualquier orden, que está formado por la unión de las intersecciones de regiones contenidas en el pseudoespectro ordinario de mismo nivel.

En el Capítulo 6 se analizan las matrices que para cada $\varepsilon > 0$ tienen iguales sus ε -pseudoespectros geométricos de cualquier orden. Estas matrices tienen necesariamente los mismos factores invariantes no triviales. Para ello, se introduce la relación de equivalencia de infinitésimos dada por Hardy.

En el Capítulo 7 se prueba que el pseudoespectro ordinario de primer orden y su frontera son conjuntos semialgebraicos. Asimismo, se extienden esos resultados de conjuntos semialgebraicos a los pseudoespectros geométricos de cualquier orden y sus fronteras, así como a los pseudoespectros algebraicos de segundo orden. Por otro lado, en relación con las fronteras de los pseudoespectros se demuestra que para algunos valores de ε se satisface que $\partial\Lambda_\varepsilon(A) \subsetneq \partial\Lambda'_\varepsilon(A)$, debido a que existen puntos aislados de $\partial\Lambda'_\varepsilon(A)$. Se caracteriza la existencia de esos puntos aislados a través de los máximos relativos de la función $z \mapsto \sigma_n(zI - A)$. Mediante un Lema de Wielandt se tiene que la frontera del ε -pseudoespectro estricto $\Lambda'_\varepsilon(A)$ está contenido en una curva algebraica real $q(x, y, \varepsilon) = 0$, donde $q(x, y, \varepsilon) \in \mathbb{R}[x, y]$ y los coeficientes de $q(x, y, \varepsilon)$ respecto de x, y pertenecen a $\mathbb{R}[\varepsilon]$. Los puntos críticos se determinan mediante la resultante de Sylvester; mientras que la naturaleza de los mismos se comprueba de varias formas: mediante las expresiones de las primeras y segundas derivadas de los valores singulares, que son funciones analíticas reales de x, y (siendo $z = x + yi$); mediante las derivadas parciales segundas de la curva $q(x, y, \varepsilon)$ respecto de x e y para unos valores de ε determinados previamente. Cuando se tiene un valor singular múltiple se utilizan las derivadas direccionales laterales. Véase [39]

En el Capítulo 8 se realiza un estudio de la geometría de los pseudoespectros de matrices 2×2 y 3×3 de una forma más detallada. Esto es posible debido a la existencia de formas canónicas de matrices para la semejanza unitaria, así como la invarianza de las singularidades de la frontera de sus pseudoespectros mediante transformaciones afines. El caso de matrices de orden 2 ha quedado completamente resuelto; mientras que en el caso de matrices de orden 3 sólo hemos resuelto los casos más significativos.

En el Apéndice A encontramos las definiciones y notaciones sobre curvas algebraicas reales utilizadas en los Capítulos 2, 7 y 8. En particular, se tratan con detalle los conceptos de número de intersección de una curva algebraica con una recta en un punto, la multiplicidad de un punto singular y las tangentes a una curva en un punto. Asimismo, se precisa la naturaleza de los puntos dobles: nodos ordinarios, puntos aislados, puntos cuspidales y tacnodos. Mediante el estudio de las series formales de potencias, de Laurent y de Puiseux, se han podido definir las parametrizaciones y ramas locales irreducibles de una curva

algebraica real, las cuales son el marco formal donde contextualizar y entender los resultados dados en esos capítulos. Además, hemos extendido la definición de número de intersección en un punto de una curva con una recta al de una rama con una curva algebraica; y de dos ramas. El paquete `algcurses` del programa de cálculo simbólico `Maple` nos ha permitido utilizando las instrucciones `singularities` y `puiseux` hallar los puntos singulares, y sus invariantes; así como parametrizaciones formales de las ramas que pasan por un punto singular.

El Apéndice B fija las notaciones sobre espacios topológicos y subconjuntos conexos; distinguiendo los conexos por arcos de los conexos. Se fijan definiciones y resultados sobre las componentes conexas de un subconjunto *cualquiera* de un espacio topológico. Como los pseudoespectros son subconjuntos de \mathbb{R}^2 , nos interesan estos conceptos referidos a los espacios euclídeos ordinarios \mathbb{R}^n .

Apéndice C reúne cuatro conceptos de amplio uso en esta Memoria: la resultante de Sylvester de dos polinomios; la matriz inversa en el sentido de Moore-Penrose; un lema de Wielandt sobre valores singulares de matrices que permite ver las fronteras de los pseudoespectros como (parte de) curvas algebraicas reales; y las σ -álgebras de subconjuntos (de Borel, medibles Lebesgue) de \mathbb{R}^2 .

Todos los resultados que aparecen en esta Tesis son originales, excepto los referidos a otros autores.

Antecedentes, notas históricas y temas relacionados

Siguiendo el dictado de Richard E. Bellman, toda tesis debería dar alguna noticia histórica sobre el origen del tema que trata. En cumplimiento de esta obligación, escribimos las siguientes notas, limitadas al alcance de nuestro conocimiento. Nos parece imprescindible contextualizar todo resultado de una investigación.

El origen del tema de los **pseudoespectros de una matriz cuadrada** puede remontarse a diversos autores entre los que se encuentran Landau, Varah, Wilkinson, Mosier, Trefethen-Embree [72], Godunov e Hinrichsen [40] (Sección 5.1.4, páginas 542 y 543). Véase el portal de Internet *Pseudospectra Gateway* dedicado a pseudoespectros [33]. En los años 1990 apareció un artículo de Trefethen en el que llamó a prestar atención a los pseudoespectros de una matriz $A \in \mathbb{C}^{n \times n}$, más que a su espectro, al estudiar cuestiones relativas a la estabilidad de soluciones de los sistemas diferenciales $\dot{x}(t) = Ax(t)$, $t \in [0, \infty)$ y en recurrencias $x(k+1) = Ax(k)$, $k = 0, 1, 2, \dots$. En los años precedentes se había constatado que para ciertos problemas de ciencia e ingeniería, las predicciones basadas en los valores propios no siempre casaban con las observaciones, véase Trefethen (1997) [71]. Se descubrió que en estas aplicaciones las matrices involucradas eran altamente no normales y tenían valores propios volátiles. Lo que sugería que debía buscarse un sustituto del espectro. De ahí surgiría el concepto de los pseudoespectros. Además, la potencia creciente de cálculo de los ordenadores permitiría la visualización de estos conjuntos para matrices de tamaños

medianos.

La cuestión del cálculo de las raíces de un polinomio

$$p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

se planteó con toda su crudeza en los años 1950 con el advenimiento de los primeros ordenadores. Famoso fue el *polinomio pérfido* de Wilkinson,

$$w(\lambda) = (\lambda - 1)(\lambda - 2) \cdots (\lambda - 20),$$

y la experiencia traumática que a éste le produjo ver que el método iterativo de Newton-Raphson fallaba y no convergía a la raíz de mayor valor absoluto de $w(\lambda)$. Véanse Wilkinson [79] y Cohen [21]. En 1986 Ronald G. Mosier [55] estudió las regiones (“root neighborhoods”) donde estaban las raíces de un polinomio $q(\lambda)$ obtenido al perturbar los coeficientes de un polinomio $p(\lambda)$ hasta una cierta magnitud ε . Probó que en las componentes conexas de esas regiones la suma de las multiplicidades de las raíces que $q(\lambda)$ tuviera en ellas era igual a la suma de las multiplicidades de las raíces que $p(\lambda)$ tenía en la misma componente. Y que cada componente conexa tenía al menos una raíz del polinomio original $p(\lambda)$.

Queremos señalar una diferencia entre el estudio de las pseudoraíces de un polinomio y los pseudovalores propios de una matriz cuadrada. Aunque los valores propios de una matriz $M \in \mathbb{C}^{n \times n}$ son las raíces de su polinomio característico $p_M(\lambda)$, un polinomio es esencialmente un objeto *unidimensional*: un vector; mientras que una matriz es un objeto *bidimensional*. De ahí que quepa esperar más riqueza a la hora de expresar las relaciones entre los pseudovalores propios de M y los elementos de M , que la esperable entre las pseudoraíces de $p_M(\lambda)$ y los coeficientes de $p_M(\lambda)$.

El periódico *SIAM News* difundió en 1995 un artículo de B.A. Cipra [20] sobre pseudoespectros de matrices en base al trabajo de Trefethen en los años 1990. En él aparecía la parte de la *derecha* de la Figura 1 con los ε -pseudoespectros de la matriz $A := \begin{pmatrix} -1 & 5 \\ 0 & -2 \end{pmatrix}$ para los niveles de $\varepsilon = 0,05; 0,15; 0,25; 0,35; 0,45; 0,55; 0,65$. Los números de condición de ambos valores propios eran iguales: $\kappa(-1) = \kappa(-2) = 5,0990$. Este hecho apareció probado por vez primera en la Tesis de Karow [44] para cualquier matriz de orden 2 con dos valores propios distintos. Para cierto valor de ε las dos componentes conexas se cortaban en un punto situado a mitad del segmento que une ambos valores propios. La figura guardaba enorme parecido con óvalos de Cassini de focos los valores propios. Incluso las gráficas de los óvalos de Cassini realizadas con `Matlab` (parte izquierda) y la imagen fotocopiada del periódico (parte derecha) coincidían al trasluz al 99%. En todas las figuras de Cipra las dos componentes conexas eran iguales y simétricas, cuando había dos. Mejor dicho, todas las figuras eran simétricas respecto de la recta vertical $x = -1,5$.

Por otra parte, desde años atrás había un enorme interés por el **problema de Wilkinson**: Dada una matriz $A \in \mathbb{C}^{n \times n}$ con n valores propios simples,

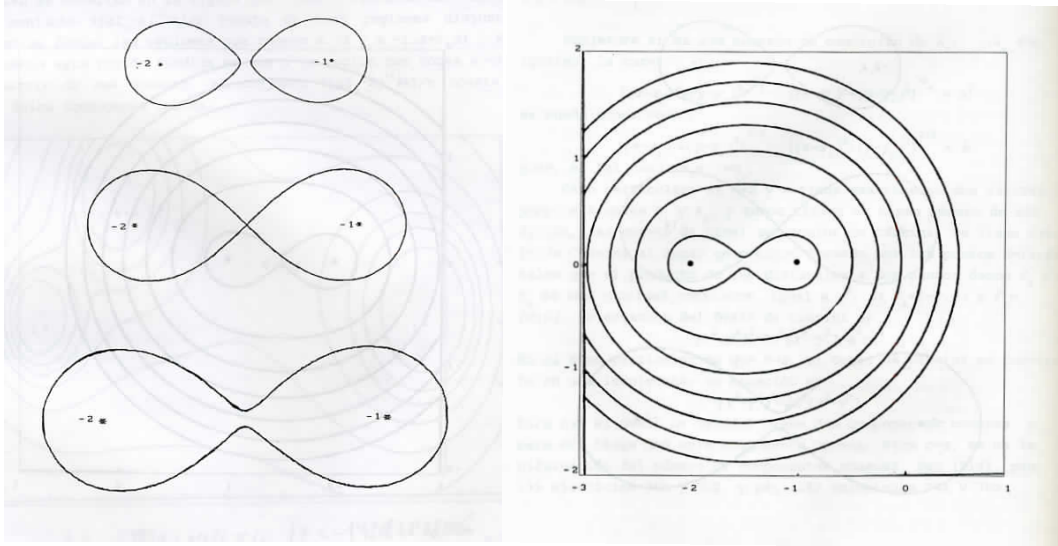


Figura 1: *Izquierda:* Óvalos de Cassini con focos $(-2,0)$ y $(-1,0)$. *Derecha:* Figura original del artículo de Cipra.

hallar la distancia de A al conjunto de matrices que tuviesen algún valor propio múltiple. Wilkinson, Golub y otros habían observado que cuando en una matriz con valores propios simples, el número de condición de uno de ellos fuera muy grande, había una matriz próxima con un valor propio múltiple (por lo general doble), cercano al valor propio mal condicionado. Varias líneas de pensamiento confluyeron en una solución dada por Malyshev a este problema, como vamos a explicar con un poco más de detalle.

En toda esta Memoria, para cualquier espacio de matrices $\mathbb{C}^{m \times n}$ usaremos la **norma espectral** de matrices, también llamada norma ℓ_2 , o norma inducida por las normas euclídeas en $\mathbb{C}^{n \times 1}$ y $\mathbb{C}^{m \times 1}$. Para ello, se asocia a cualquier matriz $A \in \mathbb{C}^{m \times n}$ la aplicación lineal $x \mapsto Ax$ de $\mathbb{C}^{n \times 1}$ en $\mathbb{C}^{m \times 1}$. También se conoce que $\|A\| = \sigma_1(A)$ donde $\sigma_1(A)$ es el mayor valor singular de A .

Es bien conocida la interpretación del k -ésimo valor singular, $\sigma_k(A)$, de una matriz $A \in \mathbb{C}^{m \times n}$ como la distancia de la matriz A al conjunto de matrices $X \in \mathbb{C}^{m \times n}$ cuyo rango es menor que k .

$$\min_{\substack{X \in \mathbb{C}^{m \times n} \\ \text{rg } X < k}} \|X - A\| = \sigma_k(A).$$

¿Qué queda en pie de esta fórmula si sustituimos $X \in \mathbb{C}^{m \times n}$ por $X \in \mathbb{R}^{m \times n}$? Bernhardsson, Rantzer y Qiu probaron que

$$\min_{\substack{X \in \mathbb{R}^{m \times n} \\ \text{rg } X < k}} \|X - A\| = \sup_{\gamma \in (0,1]} \sigma_{2k-1} \begin{pmatrix} \text{Re } A & -\gamma \text{Im } A \\ \gamma^{-1} \text{Im } A & \text{Re } A \end{pmatrix},$$

donde $\operatorname{Re} A := (\operatorname{Re} a_{ij})$, $\operatorname{Im} A := (\operatorname{Im} a_{ij})$, siendo $A = (a_{ij})$. Esta fórmula de aproximación por matrices reales fue comunicada a J.M. Gracia por M. Karow en Barcelona el 21 de julio de 1999 en la “8th ILAS Conference”.

Es bien sabido que una matriz $A \in \mathbb{C}^{n \times n}$ se dice *estable* si tiene todos sus valores propios en el semiplano $\operatorname{Re} z < 0$. Si A es estable, el radio de estabilidad complejo viene dado por

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ X \text{ inestable}}} \|X - A\| = \min_{\omega \in \mathbb{R}} \sigma_n(\omega I_n - A).$$

Tras la publicación de la fórmula

$$\min_{\substack{X \in \mathbb{R}^{n \times n} \\ X \text{ inestable}}} \|X - A\| = \min_{\omega \in \mathbb{R}} \max_{\gamma \in (0,1]} \sigma_{2n-1} \begin{pmatrix} -\operatorname{Re} A & -\gamma \omega I_n \\ \gamma^{-1} \omega I_n & -\operatorname{Re} A \end{pmatrix}$$

para el radio de estabilidad real por Qiu, Bernhardsson, Rantzer, Davison, Young y Doyle [60], Malyshev vió la manera de trasladarla al problema de encontrar la mínima distancia de A al conjunto \mathcal{M} de matrices $X \in \mathbb{C}^{n \times n}$ que tienen un valor propio múltiple.

$$\min_{X \in \mathcal{M}} \|X - A\| = \min_{z \in \mathbb{C}} \max_{\gamma \in \mathbb{R}} \sigma_{2n-1} \begin{pmatrix} zI_n - A & \gamma I_n \\ O & zI_n - A \end{pmatrix}.$$

No parece nada extraño que Malyshev abordara esta suerte de problema. Su director de tesis fue Godunov en Novosibirsk, y Godunov fue uno de los introductores del concepto de pseudoespectros con el nombre de retratos espectrales (“spectral portraits”) o ε -espectros. Wilkinson intuyó la importancia de los pseudoespectros de una matriz a los que llamó “dominios de inclusión fundamentales”; él usó la letra η en lugar del ε . Hizo notar la importancia del dominio $D(\eta)$ del plano complejo formado por los λ tales que

$$\frac{1}{\|(A - \lambda I)^{-1}\|} \leq \eta;$$

pues todos los $\lambda \in D(\eta)$ son valores propios de $A + F$ para perturbaciones F que satisfacen $\|F\| \leq \eta$. Wilkinson definió el concepto de *sucesor* de un valor propio de la siguiente manera: Se dice que un número complejo z es un ε -*sucesor* del valor propio λ de A si existe una matriz E_λ con $\|E_\lambda\| \leq \varepsilon$ tal que z es valor propio de $A + E_\lambda$, $\varphi(0) = \lambda$, $\varphi(1) = z$ siendo $\varphi: [0, 1] \rightarrow \mathbb{C}$ la función continua que describe un valor propio de la matriz $A + tE_\lambda$, cuando t varía de 0 a 1. Puede ocurrir que z sea un ε -sucesor de más de un valor propio de A . El principal interés se centra en aquellos z que son ε -sucesores de más de un valor propio para el mínimo ε . Se define que z es un ε -*sucesor de, digamos, dos valores propios λ y μ de A* si ε es el valor mínimo para el que z es un ε -sucesor de λ y un ε -sucesor de μ . Con todo, estas definiciones no fueron muy afortunadas, y no pudieron clarificarse hasta que se precisó el concepto de ε -pseudoespectros de A . Véanse los artículos de Wilkinson sobre la sensibilidad de los valores propios [77] y [78]. Dejémoslo hablar con sus propias palabras ([78], página 254):

The behaviour of the domain $D(\eta)$ as η increases from zero is of fundamental interest to us. We shall assume that A has s distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_s$ of multiplicities m_1, m_2, \dots, m_s . If all the eigenvalues are simple, then $s = n$ and $m_i = 1$ ($i = 1, \dots, n$). Clearly $D(\eta_2)$ contains $D(\eta_1)$ when $\eta_2 > \eta_1$. Indeed, if λ is any point of $D(\eta_1)$ it is an internal point of $D(\eta_2)$ since the elements of $(A - \lambda I)^{-1}$ are continuous functions of λ apart from poles at the λ_i . $D(0)$ consists of the s isolated points λ_i . When η is sufficiently small, $D(\eta)$ consists of s isolated domains, the i th of which has λ_i as an internal point. For any such η and any specific E (with $\|E\| = 1$), the i th of these domains contains the m_i successors of λ_i . A problem of basic interest to us is the smallest value of η for which one of these domains coalesces with one of the others. We may divide the s values of λ_i into a number of distinct sets and to determine the minimum value of η for which a domain containing the successors of one of these sets coalesces with a domain containing the successors of another of the sets (i.e. ignoring the possible coalescence of individual λ_i in the same set).

(JAMES H. WILKINSON)

Wilkinson también desarrolló trabajos sobre la distancia de A al conjunto de matrices con forma canónica de Jordan menos genérica; véase [76]. La línea de pensamiento de Wilkinson, utilizando ε -sucesores de los valores propios de A , fue seguida por J.W. Demmel en su densa Tesis cuyo sugerente título no defraudó [27]. Véase también su artículo [28]. A Demmel lo siguieron con paso firme R. Alam y S. Bora, en la Tesis de esta última, que dio lugar al artículo [3]; en él apareció otra solución al problema de Wilkinson, dada a través del primer punto de coalescencia de componentes conexas de los pseudoespectros. La primera solución la había dado Malyshev, por un método *minimax* como hemos indicado antes.

Para nosotros, a día de hoy el libro más profundo en sentido matemático que se ha escrito sobre pseudoespectros de matrices lo constituye la Tesis Doctoral de Michael Karow (Bremen, 2003) [44]. Documento al que no se ha prestado la atención suficiente. La presente Tesis tiene una deuda intelectual con la de Karow. El trabajo pionero de este autor sobre pseudoespectros, ha sido guía y fuente de inspiración. A él se deben la semialgebraicidad de los valores singulares como funciones de una matriz y de los pseudoespectros y sus fronteras como subconjuntos de \mathbb{R}^2 , así como las estrechas relaciones entre la componente conexas que rodea a un valor propio λ de un ε -pseudoespectro para valores pequeños de ε y el índice de λ .

El tema de estudio de los pseudoespectros de matrices es objeto de investigación muy activa en los últimos veinticinco años. Y se ha extendido a campos como el análisis de los conjuntos de pseudoceros de sistemas de polinomios de varias variables, habiéndose logrado generalizar los teoremas de Mosier de 1986 sobre conservación de multiplicidades de los pseudoceros en las componentes conexas de los conjuntos de pseudoceros. Véase el trabajo de Hoffman et al. [41] y sus referencias. El tema tiene profundas connotaciones con la geometría algebraica real. En muchos casos de interés natural los conjuntos de pseudoceros son semialgebraicos.

Agradecimientos

En primer lugar, agradezco a mis directores de tesis, Juan-Miguel Gracia y Francisco-Enrique Velasco, por ofrecerme un tema tan interesante, por los conocimientos transmitidos y su calidad humana, y por su plena dedicación a ayudarme en todo lo que he necesitado. En segundo lugar, agradezco a Inmaculada de Hoyos por su inestimable ayuda en la revisión de esta Tesis. Asimismo, quiero dejar constancia de mi gratitud al Grupo de Análisis Matricial y Aplicaciones de la Universidad del País Vasco/Euskal Herriko Unibertsitatea UPV/EHU por la ayuda recibida en sus seminarios, por los comentarios y críticas constructivas, y por la atención que me han dedicado en los viajes a congresos y cursos. Deseo expresar mi reconocimiento a María Jesús de la Puente por su clara explicación de ayuda en el análisis de una curva algebraica real cuártica que surgió en el estudio de la frontera de pseudoespectros de matrices de segundo orden.

A mis padres, por su amor infinito y los esfuerzos que han hecho, por su dedicación y ánimos transmitidos. A Feli, por su apoyo y paciencia.

Finalmente, quiero agradecer a mis compañeros de la Universidad Pública de Navarra por el interés que han mostrado en la realización de este trabajo. En especial a Pedro Pagola.

Índice de figuras

1.	<i>Izquierda:</i> Óvalos de Cassini con focos $(-2, 0)$ y $(-1, 0)$. <i>Derecha:</i> Figura original del artículo de Cipra.	XI
2.1.	Gráfica de una función ρ'	21
2.2.	Gráfica de una función ρ	22
3.1.	Paso angosto para la curva roja.	30
3.2.	Entrelazamiento de pseudoespectros $\Lambda_{\varepsilon_1,1}^a(A)$ y $\Lambda_{\varepsilon_2,2}^a(A)$	34
4.1.	Áreas homotéticas	39
4.2.	Descomposición canónica de un abierto en cuadrados, bajo una homotecia	40
4.3.	ε -pseudoespectros de A_1 para $\varepsilon = 1,00; 1,50; 1,75$	43
5.1.	Pseudoespectro de segundo orden	68
5.2.	Fronteras del ε -pseudoespectro geométrico.	71
7.1.	Pseudoespectros hendidos.	84
7.2.	Superficies $z = \sigma_1(x, y)$, $z = \sigma_2(x, y) = h_1(x, y)$	89
7.3.	$\partial\Lambda'_{\varepsilon_0}(A) \neq \{(x, y) \in \mathbb{R}^2 : q(x, y, \varepsilon_0) = 0\}$	90
7.4.	Punto aislado $(-4/15, 0)$	95
7.5.	ε -pseudoespectros con ε bajo.	97
7.6.	Curva de nivel de $h_1(x, y) = \sigma_3((x + yi)I_3 - A)$ que contiene un punto de ensilladura.	98
7.7.	Máximo relativo de $h_1(x, y) = \sigma_3((x + yi)I_3 - A)$	98
7.8.	Puntos acnodos $(-\frac{4}{19}, \pm\frac{4\sqrt{5}}{95})$	101
7.9.	Máximos relativos de h_1 en los puntos $(-\frac{4}{19}, \pm\frac{4\sqrt{5}}{95})$	102
7.10.	Gráfica de la función f sobre el intervalo $0 \leq t \leq 2\pi$	110
8.1.	Retroceso de primera especie en $(x_2, y_2) = \left(-\frac{\sqrt{3}}{6}, 0\right)$	131
8.2.	Tacnodo en $(x_3, y_3) = (1/2, 0)$ a la curva algebraica $V_{\mathbb{R}}(f_{\varepsilon_3})$	133
8.3.	Conjunto frontera $\partial\Lambda'_{\varepsilon_3}(A)$	133
8.4.	Tacnodo en $(x_4, y_4) = (3/8, 1/8)$ a la curva algebraica $V_{\mathbb{R}}(f_{\varepsilon_4})$	138

8.5. Tacnodo en $(x_5, y_5) = \left(\frac{\sqrt{2}-1}{2}, \frac{\sqrt{2}-1}{2}\right)$ a la curva algebraica $V_{\mathbb{R}}(f_{\varepsilon_5})$.	138
A.1. Curva $y^2 - x^3 - x^2 = 0$; crunodo en $(0, 0)$.	145
A.2. Curva $y^2 - x^3 + x^2 = 0$; acnodo en $(0, 0)$.	145
A.3. Puntos de retroceso: 1ª especie (izquierda), 2ª especie (derecha).	147
A.4. Puntos de retroceso de primera y segunda especies.	147
A.5. Ejemplo de puntos tacnodos.	148
A.6. Cúspide ranfoidea en $(0, 0)$.	162
A.7. Cuatro ramas reales centradas en la vertical $x = (-1 + \sqrt{2})/2$.	166
A.8. Oscnodo de $y^2 - x^6 = 0$.	170
A.9. Tacnodo de $fg = 0$ en el punto $(0, 0)$.	170
C.1. Los seis supervalores propios de la matriz A .	185
C.2. Las funciones h_1, h_2 y h_3 para la matriz A .	185

Notaciones

A^* , matriz traspuesta conjugada de A , página 1

$D(z, \rho)$, disco abierto de centro z y radio ρ , página 1

$\mathbb{C}^{p \times q}$, espacio de las matrices rectangulares complejas de orden $p \times q$, página 1

$\Lambda(A)$, espectro de la matriz A , página 1

$\Lambda_k^g(A)$, espectro geométrico de orden k de la matriz A , página 1

$\Lambda_k(A)$, espectro de orden k de la matriz A , página 1

$\|A\|_F = \sqrt{\operatorname{tr}(A^*A)}$, norma matricial de Frobenius, página 2

$m(\lambda, A)$, multiplicidad algebraica de λ respecto de la matriz A , página 1

$\operatorname{mg}(\lambda, A)$, multiplicidad geométrica de λ respecto de la matriz A , página 1

$\mathcal{D}(z, \rho)$, disco cerrado de centro z y radio ρ , página 1

$\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_k(A)$, valores singulares de la matriz $A \in \mathbb{C}^{p \times q}$, con $k = \min\{p, q\}$, página 2

\approx , semejanza ordinaria, página 3

$\overset{u}{\approx}$, semejanza unitaria, página 3

$f_{\varepsilon, A}(x, y) = f_{\varepsilon}(x, y) = q(x, y, \varepsilon)$ polinomio de Wielandt de la matriz A para el nivel ε , página 6

$\operatorname{cond}_{\omega}(A, \lambda)$, número de condición de orden ω del valor propio λ , página 9

$\nu(\lambda)$, índice del valor propio λ , página 8

$\operatorname{sv}_{(A, \lambda)}(A')$, variación espectral de A' respecto de λ , página 8

d_H , distancia de Hausdorff, página 9

$\Lambda'_{\varepsilon}(A) := \Lambda'_{\varepsilon, 1}(A)$, abreviatura para el pseudoespectro estricto de primer orden, página 11

$\Lambda_\varepsilon(A) := \Lambda_{\varepsilon,1}^a(A)$, abreviatura para el pseudoespectro ordinario de primer orden, página 11

$\Lambda_{\varepsilon,1}^a(A)$, pseudoespectro algebraico estricto de primer orden de nivel ε de la matriz A , página 11

$\Lambda_{\varepsilon,1}^a(A)$, pseudoespectro algebraico ordinario de primer orden de nivel ε de la matriz A , página 11

$h_1(x, y) := \sigma_n((x + yi)I_n - A)$, último valor singular, $(x, y) \in \mathbb{R}^2$, página 12

$h_1(z) := \sigma_n(zI_n - A)$, último valor singular, $z \in \mathbb{C}$, página 12

$\partial S := \partial(S)$, frontera del conjunto $S \subset \mathbb{C}$, página 17

$|X|$, cardinal del conjunto finito X , página 18

$A \dot{\cup} B := A \cup B$, cuando $A \cap B = \emptyset$, página 19

$D(c, b)$, disco abierto de centro $c \in \mathbb{C}$ y que tiene a $b \in \mathbb{C}$ en su frontera, página 23

$\Lambda_{\varepsilon,k}^a(A)$, pseudoespectro algebraico ordinario de orden k de nivel ε de la matriz A , página 27

$\Lambda_{\varepsilon,k}^g(A)$, pseudoespectro geométrico ordinario de orden k de nivel ε de la matriz A , página 27

$F_z(t)$, $\begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix}$, página 28

$h_2(x, y) := \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} (x+yi)I_n - A & tI_n \\ O & (x+yi)I_n - A \end{pmatrix}$, función de Malyshev $h_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$, página 28

$h_2(z) := \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} zI_n - A & tI_n \\ O & zI_n - A \end{pmatrix}$, función de Malyshev $h_2 : \mathbb{C} \rightarrow \mathbb{R}$, página 28

$\mathcal{K}'_\lambda(\varepsilon)$, componente conexa del pseudoespectro estricto $\Lambda'_\varepsilon(A)$ que contiene al valor propio λ , página 35

$\mathcal{K}_\lambda(\varepsilon)$, componente conexa del ε -pseudoespectro que contiene al valor propio λ , página 37

$\partial\mathcal{K}_\lambda(\varepsilon)$, frontera de $\mathcal{K}_\lambda(\varepsilon)$, página 39

$r(\varepsilon)$, radio medio de $\partial\mathcal{K}_\lambda(\varepsilon)$, página 53

$J_n(\lambda, d)$, bloque de Jordan generalizado, página 55

$\text{Eig}(A)$, n -tupla desordenada de valores propios de A , página 62

$B'(z_0, \delta)$, bola perforada de centro $z_0 \in \mathbb{C}$ y radio $\delta > 0$, página 73

- $V'(z_0)$, entorno perforado de centro z_0 , página 73
- $f(z) \asymp g(z)$, ($z \rightarrow z_0$), funciones equivalentes según Hardy cuando $z \rightarrow z_0$, página 73
- $\partial\Lambda'_\varepsilon(A)$, frontera de $\Lambda'_\varepsilon(A)$, página 83
- $\partial\Lambda_\varepsilon(A)$, frontera de $\Lambda_\varepsilon(A)$, página 83
- $\mathbf{A}(x, y) := (x + yi)I_n - A$, página 84
- $\mathbf{A}(z) := zI_n - A$, página 84
- $\text{dist}(A, B)$, distancia entre los subconjuntos A y B de \mathbb{C} , página 100
- K , un cuerpo cualquiera, página 139
- $K[x_1, \dots, x_n]$, anillo de los polinomios en las variables x_1, \dots, x_n con coeficientes en el cuerpo K , página 139
- $V_K(f_1, \dots, f_s)$, conjunto algebraico en el espacio afín K^n , página 139
- $V_{\mathbb{R}}(f)$, curva algebraica real plana, página 139
- $I(p, f, \ell)$, número (o multiplicidad) de intersección de la curva f con la recta ℓ en el punto p , página 141
- $m_p(f)$, multiplicidad del punto p en la curva algebraica f , página 141
- $K[[T]]$, anillos de las series formales de potencias en T , página 151
- $\text{ord}(a)$, orden de una serie formal de potencias a , página 151
- $K((T))$, cuerpo de fracciones del anillo de integridad $K[[T]]$, página 154
- $O(f)$, orden de una serie formal de Laurent f , página 154
- $K\{\{T\}\}$, cuerpo de las series de Puiseux en T con coeficientes en K , página 156
- $O(\bar{a})$, orden de una serie formal de Puiseux \bar{a} , página 156
- $\mathbb{P}(K) := K\{\{T\}\}$, cuerpo de las series formales de Puiseux, página 156
- $I(p, S, f)$, multiplicidad de intersección de una rama S con una curva $V_{\mathbb{C}}(f)$ en el punto p , página 166
- $\text{mult}_p(C, D) := I(p, C, D)$ multiplicidad de la intersección de C y D en el punto p , página 167
- $I(p, R, S)$, multiplicidad en un punto de intersección p de dos ramas R y S , página 169
- $\overset{\circ}{S}$, interior del conjunto $S \subset E$ espacio topológico, página 174

\bar{S} , adherencia o clausura del conjunto $S \subset E$ espacio topológico , página 174

$\partial(S)$, frontera del conjunto $S \subset E$ espacio topológico, página 174

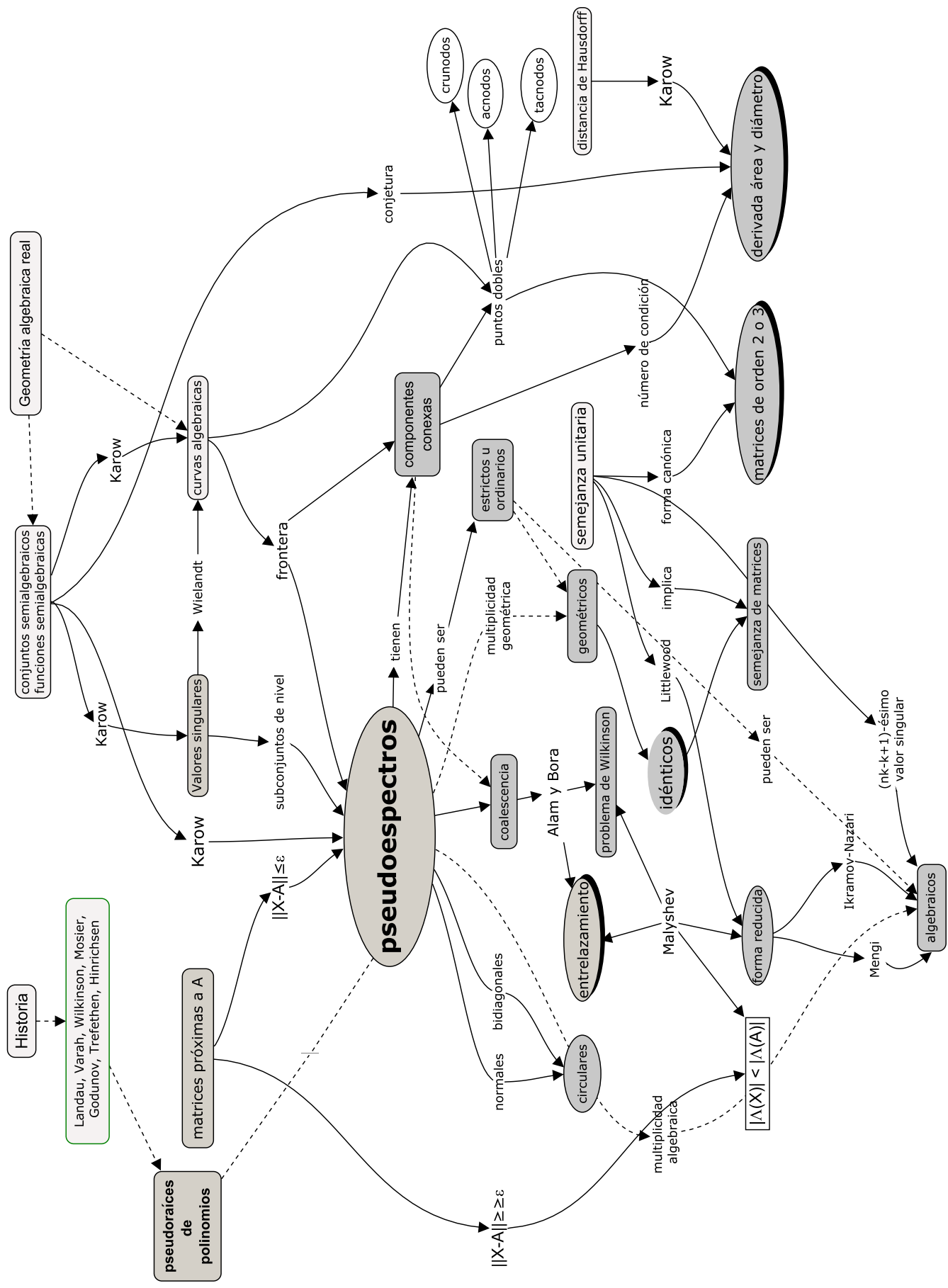
$B(x, r)$, bola abierta de centro x y radio $r > 0$, página 178

A^\dagger , inversa Moore-Penrose de A , página 180

$\text{Card } \mathcal{B}$, cardinal del conjunto infinito \mathcal{B} , página 181

$\sigma(\mathcal{G})$, σ -álgebra generada por \mathcal{G} , página 181

\mathcal{B} , σ -álgebra de los conjuntos de Borel de \mathbb{R}^2 , página 181



Capítulo 1

Preliminares

1.1. Valores propios

En este trabajo denotaremos por $\mathbb{C}^{p \times q}$ al espacio de las matrices rectangulares complejas de orden $p \times q$; así mismo I_n denotará la matriz identidad de $\mathbb{C}^{n \times n}$. Denotaremos por A^* a la *matriz traspuesta conjugada* de A . Así mismo denotaremos por $\mathcal{D}(z, \rho)$ al disco cerrado centrado en $z \in \mathbb{C}$ de radio ρ , y por $D(z, \rho)$ al disco abierto con los mismos centro y radio.

Un número complejo λ se dice que es un **valor propio** de una matriz $A \in \mathbb{C}^{n \times n}$, si $\det(\lambda I_n - A) = 0$. El conjunto de valores propios de una matriz $A \in \mathbb{C}^{n \times n}$ se llama **espectro** de A y lo denotaremos por $\Lambda(A)$. Dada $A \in \mathbb{C}^{n \times n}$ y $\lambda \in \mathbb{C}$, denotemos por $m(\lambda, A) := \dim \ker(\lambda I_n - A)^n$ a la **multiplicidad algebraica** de λ respecto de A . Si λ no es valor propio de A entonces $m(\lambda, A) = 0$, y recíprocamente. De la misma manera, denotemos por $mg(\lambda, A) := \dim \ker(\lambda I_n - A)$ a la **multiplicidad geométrica** del número complejo λ respecto de la matriz A . Igualmente $mg(\lambda, A) = 0$ si y sólo si λ no es valor propio de A . Con estas notaciones, definimos los conjuntos

$$\Lambda_k(A) := \{\lambda \in \mathbb{C} : m(\lambda, A) \geq k\}$$

que llamaremos *espectro de orden k* de A , y

$$\Lambda_k^g(A) := \{\lambda \in \mathbb{C} : mg(\lambda, A) \geq k\}$$

que llamaremos *espectro geométrico de orden k* de A .

1.2. Valores singulares, semejanza unitaria

La norma matricial que usaremos en este trabajo es la norma espectral sobre $\mathbb{C}^{p \times q}$, representada por $\|\cdot\|$, y definida como

$$\|A\| := \max_{\substack{x \in \mathbb{C}^{q \times 1} \\ \|x\|_2 = 1}} \|Ax\|_2.$$

Observación 1.2.1. Utilizaremos también la **norma de Frobenius** sobre $\mathbb{C}^{n \times n}$, representada por $\|\cdot\|_F$, y definida como

$$\|A\|_F := \sqrt{\operatorname{tr}(A^*A)} = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2},$$

siendo $A = (a_{ij})$, con $i = 1, \dots, n$ y $j = 1, \dots, n$.

Definición 1.2.2. Dados $A \in \mathbb{C}^{n \times n}$ y $\delta > 0$, llamamos **bola abierta** de centro A y radio δ , y la denotamos como $B(A, \delta)$, al conjunto

$$B(A, \delta) = \{X \in \mathbb{C}^{n \times n} : \|X - A\| < \delta\}.$$

Análogamente, llamamos **bola cerrada** de centro A y radio δ , y la denotamos como $\overline{B}(A, \delta)$, al conjunto

$$\overline{B}(A, \delta) = \{X \in \mathbb{C}^{n \times n} : \|X - A\| \leq \delta\}.$$

Los *valores singulares* de una matriz $A \in \mathbb{C}^{p \times q}$ se denotarán por

$$\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq \sigma_k(A),$$

donde $k = \min\{p, q\}$. Recordemos que los valores singulares de una matriz A son las raíces cuadradas de los valores propios de la matriz A^*A .

Utilizando las relaciones de Courant-Fisher para los valores singulares, se tienen las siguientes igualdades para una matriz $A \in \mathbb{C}^{n \times n}$.

$$\|A\| = \sigma_1(A), \quad \min_{\substack{x \in \mathbb{C}^{n \times 1} \\ \|x\|_2=1}} \|Ax\|_2 = \sigma_n(A).$$

Pasamos a enunciar algunas propiedades de los valores singulares. La primera es la desigualdad de Weyl.

Teorema 1.2.3. Sean las matrices $A, B \in \mathbb{C}^{p \times q}$. Entonces

$$|\sigma_j(A) - \sigma_j(B)| \leq \|A - B\|, \quad j = 1, \dots, \min\{p, q\}.$$

Este resultado también se puede interpretar como la continuidad de los valores singulares $\sigma_j(A)$ respecto de la matriz A .

Recordemos que una matriz $U \in \mathbb{C}^{n \times n}$ se dice unitaria si $U^*U = I_n$.

Proposición 1.2.4. Los valores singulares de una matriz son invariantes bajo transformaciones unitarias. Es decir, dada $A \in \mathbb{C}^{p \times q}$, para todo par de matrices unitarias $U \in \mathbb{C}^{p \times p}$, $V \in \mathbb{C}^{q \times q}$ se tiene

$$\sigma_i(A) = \sigma_i(UAV) \quad i = 1, \dots, \min\{p, q\}.$$

El siguiente resultado es el *Teorema de descomposición de valores singulares*.

Teorema 1.2.5. Sean $A \in \mathbb{C}^{p \times q}$ y $\sigma_1 \geq \dots \geq \sigma_r > 0$, $r = \text{rg}(A)$, sus valores singulares no nulos. Entonces existen matrices unitarias $U \in \mathbb{C}^{p \times p}$ y $V \in \mathbb{C}^{q \times q}$ de manera que

$$A = U\Sigma V^* \quad (1.2.1)$$

donde

$$\Sigma = \begin{matrix} & \begin{matrix} r & q-r \end{matrix} \\ \begin{matrix} r \\ p-r \end{matrix} & \begin{bmatrix} \sigma_1 & & O \\ \cdot & \cdot & \cdot \\ O & \sigma_r & O \\ & & O \end{bmatrix} \end{matrix} \in \mathbb{C}^{p \times q}.$$

Observación 1.2.6. Las columnas de la matriz V (respectivamente, U), vistas como vectores de $\mathbb{C}^{q \times 1}$ (resp. $\mathbb{C}^{p \times 1}$), constituyen una base de vectores propios ortonormal para A^*A (resp. AA^*). Así, para hallar la descomposición en valores singulares de A basta resolver el problema de los valores propios para A^*A y AA^* .

El teorema siguiente nos da la distancia de una matriz cuadrada al conjunto de matrices que tienen a un complejo dado como valor propio cuya multiplicidad geométrica es mayor o igual que k .

Teorema 1.2.7. Sean $z_0 \in \mathbb{C}$ y $A \in \mathbb{C}^{n \times n}$. Entonces,

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda_k^g(X)}} \|X - A\| = \sigma_{n-k+1}(z_0 I_n - A).$$

1.2.1. Semejanza unitaria

Definición 1.2.8. Dos matrices $A, B \in \mathbb{C}^{n \times n}$ se dice que son **semejantes unitariamente** si existe una matriz unitaria, U , tal que

$$B = U^* A U.$$

La notación $B \stackrel{u}{\approx} A$ denotará que las matrices A y B son unitariamente semejantes. Mientras que $B \approx A$ denota la semejanza ordinaria de matrices: $B = P^{-1}AP$ con $P \in \mathbb{C}^{n \times n}$ invertible. Tres resultados sobre la semejanza unitaria son los siguientes.

Lema 1.2.9 (Lema de Schur). Toda matriz $A \in \mathbb{C}^{n \times n}$ es semejante unitariamente a una matriz triangular superior, es decir, existe U unitaria tal que $U^* A U = T$.

Como consecuencia, si A es una matriz normal, es decir $A^* A = A A^*$, del lema anterior se deduce lo siguiente.

Corolario 1.2.10. Una matriz normal $A \in \mathbb{C}^{n \times n}$ es semejante unitariamente a la matriz

$$\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

siendo cada λ_i un valor propio de A , $i = 1, 2, \dots, n$.

Teorema 1.2.11. Sean $A, B \in \mathbb{C}^{n \times n}$ matrices triangulares superiores con los mismos elementos diagonales $a_{ii} = b_{ii}$, para todo $i \in \{1, \dots, n\}$, y tal que $a_{ii} \neq a_{jj}$, $i \neq j$. Supongamos $U \in \mathbb{C}^{n \times n}$ matriz unitaria tal que $U^*AU = B$, entonces U es diagonal.

DEMOSTRACIÓN.

De la relación de semejanza unitaria entre las matrices A y B se tiene:

$$U^*AU = B \Leftrightarrow AU = UB.$$

En la anterior expresión consideramos el elemento que ocupa la posición $(n, 1)$:

$$\sum_{i=1}^n a_{ni}u_{i1} = \sum_{j=1}^n u_{nj}b_{j1} \xRightarrow[\substack{a_{ij}=0=b_{ij}, \\ i>j}]{\implies} a_{nn}u_{n1} = u_{n1}b_{11} \Rightarrow u_{n1}(a_{nn} - b_{11}) = 0.$$

Como $a_{ii} = b_{ii} \neq b_{jj}$, se tiene $u_{n1} = 0$. Análogamente, si calculamos el elemento en la posición $(n, 2)$:

$$\sum_{i=1}^n a_{ni}u_{i2} = \sum_{j=1}^n u_{nj}b_{j2} \xRightarrow[\substack{a_{ij}=0=b_{ij}, \\ i>j}]{\implies} a_{nn}u_{n2} = u_{n2}b_{22} \Rightarrow u_{n2}(a_{nn} - b_{22}) = 0 \Rightarrow u_{n2} = 0.$$

Si siguiendo el razonamiento, tenemos $u_{nj} = 0$, para todo $j \in \{1, \dots, n-1\}$. Si calculamos el valor del elemento $(n-1, 1)$, tenemos:

$$\sum_{i=1}^n a_{n-1,i}u_{i,1} = \sum_{j=1}^n u_{n-1,j}b_{j,1} \Rightarrow a_{n-1,n-1}u_{n-1,1} + a_{n-1,n} \underbrace{u_{n,1}}_{=0} = u_{n-1,1}b_{11} \Rightarrow u_{n-1,1}(a_{n-1,n-1} - b_{11}) = 0,$$

de donde $u_{n-1,1} = 0$.

Si siguiendo este proceso, podemos probar que $u_{ij} = 0$ si $i \neq j$. Por lo tanto, $U = \text{diag}(e^{i\alpha_1}, \dots, e^{i\alpha_n})$, $\alpha_i \in [0, 2\pi]$, para todo $i \in \{1, \dots, n\}$. \square

1.2.2. Lema de Wielandt

Dada una matriz $M \in \mathbb{C}^{n \times n}$, hay una relación entre sus valores singulares y los valores propios de cierta matriz asociada a ella. El resultado es el siguiente.

Lema 1.2.12 (Lema de Wielandt). Sea $M \in \mathbb{C}^{n \times n}$. Sean $s_1 \geq \dots \geq s_n \geq 0$. Entonces

$$s_1 \geq \dots \geq s_n \geq -s_n \geq \dots \geq -s_1,$$

son los valores propios de la matriz hermitica $2n \times 2n$

$$H := \begin{bmatrix} O & M \\ M^* & O \end{bmatrix}$$

ordenados en sentido decreciente si y sólo si $s_1 \geq \dots \geq s_n$ son los valores singulares de M ordenados en sentido decreciente.

La matriz H se conoce como la **matriz de Wielandt** asociada a M .

Así pues el número real $\varepsilon \geq 0$ es un valor singular de M si y sólo si ε es un valor propio de H . De igual modo, por el lema del complemento de Schur (véase Lancaster y Tismenetsky [48, pág. 46, Ejercicio 15],

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D - CA^{-1}B| |A| \text{ si } |A| \neq 0,$$

siendo A y D matrices cuadradas, se sigue que $\varepsilon \geq 0$ es un valor singular de la matriz M si y sólo si ε^2 es un valor propio de la matriz M^*M . Además, los valores propios de las matrices

$$\begin{bmatrix} O & M \\ M^* & O \end{bmatrix} \text{ y } \begin{bmatrix} O & -M \\ -M^* & O \end{bmatrix}$$

son los mismos. Veamos el porqué de esta afirmación última. El número λ_0 es un valor propio de

$$\begin{bmatrix} O & M \\ M^* & O \end{bmatrix}$$

si y sólo si

$$\begin{vmatrix} \lambda_0 I_n & -M \\ -M^* & \lambda_0 I_n \end{vmatrix} = 0.$$

Multiplicando la primera fila de bloques de este determinante por -1 , se tiene que

$$(-1)^n \begin{vmatrix} -\lambda_0 I_n & M \\ -M^* & \lambda_0 I_n \end{vmatrix} = (-1)^{2n} \begin{vmatrix} \lambda_0 I_n & M \\ M^* & \lambda_0 I_n \end{vmatrix} = \begin{vmatrix} \lambda_0 I_n & M \\ M^* & \lambda_0 I_n \end{vmatrix}.$$

Por tanto,

$$\begin{vmatrix} \lambda_0 I_n & -M \\ -M^* & \lambda_0 I_n \end{vmatrix} = 0 \text{ si y sólo si } \begin{vmatrix} \lambda_0 I_n & M \\ M^* & \lambda_0 I_n \end{vmatrix} = 0;$$

es decir, si y sólo si λ_0 es un valor propio de

$$\begin{vmatrix} O & -M \\ -M^* & O \end{vmatrix}.$$

Sea ahora $A \in \mathbb{C}^{n \times n}$, $x, y \in \mathbb{R}$ y $\varepsilon \geq 0$. Por consiguiente, si definimos

$$q(x, y, \varepsilon) := \begin{vmatrix} \varepsilon I_n & (x + yi)I_n - A \\ (x - yi)I_n - A^* & \varepsilon I_n \end{vmatrix} \quad (1.2.2)$$

se tiene que

$$q(x, y, \varepsilon) = \begin{vmatrix} \varepsilon I_n & A - (x + yi)I_n \\ A^* - (x - yi)I_n & \varepsilon I_n \end{vmatrix}.$$

Por tanto, se tiene la proposición siguiente.

Proposición 1.2.13. *Supongamos que $\varepsilon \geq 0$. Entonces se tiene que $q(x, y, \varepsilon) = 0$ si y sólo si ε es un valor singular de $(x + yi)I_n - A$.*

Al polinomio $f_\varepsilon(x, y) = q(x, y, \varepsilon)$ lo llamaremos **polinomio de Wielandt de la matriz A para el nivel ε** y lo denotaremos también por $f_{\varepsilon, A}(x, y)$ cuando sea necesaria la referencia a A . Es de observar que $f_{\varepsilon, A}(x, y)$ es mirado como un polinomio en x, y .

Lema 1.2.14. *Dada $A \in \mathbb{C}^{n \times n}$, el polinomio f_ε dado por*

$$f_\varepsilon(x, y) := q(x, y, \varepsilon) = \sum_{i+j \leq 2n} c_{ij}(\varepsilon) x^i y^j$$

tiene sus coeficientes $c_{ij}(\varepsilon)$ en $\mathbb{R}[\varepsilon]$. Es decir, el polinomio f_ε pertenece al anillo $\mathbb{R}[\varepsilon][x, y]$ de polinomios en x, y con coeficientes en el anillo $\mathbb{R}[\varepsilon]$.

DEMOSTRACIÓN. En principio hay números complejos tanto en las matrices A y A^* como en las expresiones conjugadas $x + yi$ y $x - yi$. Por esto, expresemos $q(x, y, \varepsilon)$ como un elemento de $\mathbb{C}[x, y][\varepsilon]$:

$$q(x, y, \varepsilon) = \varepsilon^{2n} + a_1(x, y)\varepsilon^{2n-1} + \cdots + a_{2n-1}(x, y)\varepsilon + a_{2n}(x, y),$$

visto como polinomio en ε . Observemos que $q(x, y, \varepsilon)$ es el polinomio característico de la matriz hermitica

$$\begin{bmatrix} O & (x + yi)I_n - A \\ (x - yi)I_n - A^* & O \end{bmatrix}.$$

Por tanto sus raíces en ε son reales. Así pues, por las relaciones entre las raíces y los coeficientes de un polinomio, tenemos que los coeficientes de este polinomio en ε son reales. En principio,

$$a_k(x, y) \in \mathbb{C}[x, y].$$

Supongamos que

$$a_k(x, y) = \sum_{i,j} a_{ij}^{(k)} x^i y^j, \text{ con } a_{ij}^{(k)} \in \mathbb{C}, \quad k = 1, \dots, 2n.$$

Para todos $t \in \mathbb{R}$, $z \in \mathbb{C}$, $\operatorname{Re}(tz) = t \operatorname{Re}(z)$, $\operatorname{Im}(tz) = t \operatorname{Im}(z)$. Como para todo $(x, y) \in \mathbb{R}^2$ tenemos que $a_k(x, y) \in \mathbb{R}$, se sigue que $\operatorname{Im}(a_k(x, y)) = 0$. pero

$$\operatorname{Im}(a_k(x, y)) = \sum_{i,j} \operatorname{Im}(a_{ij}^{(k)}) x^i y^j.$$

Por tanto, el polinomio

$$\sum_{i,j} \operatorname{Im}(a_{ij}^{(k)}) x^i y^j \in \mathbb{R}[x, y]$$

se anula en todo punto $(x, y) \in \mathbb{R}^2$. Como el cuerpo \mathbb{R} es infinito,

$$\operatorname{Im}(a_{ij}^{(k)}) = 0, \quad \forall i, j.$$

Por consiguiente,

$$a_{ij}^{(k)} \in \mathbb{R}, \forall i, j,$$

con lo cual, el polinomio $q(x, y, \varepsilon) \in \mathbb{C}[x, y][\varepsilon]$ tiene reales sus coeficientes: dicho de otro modo $q(x, y, \varepsilon) \in \mathbb{R}[x, y][\varepsilon]$. Lo que implica que para todo $\varepsilon \geq 0$, $c_{ij}(\varepsilon)$ es real $\forall i, j$. \square

También se tiene el resultado siguiente.

Lema 1.2.15.

$$q(x, y, \varepsilon) = \det \left(\varepsilon^2 I_n - ((x + yi)I_n - A)^* ((x + yi)I_n - A) \right).$$

DEMOSTRACIÓN. $z = x + yi$, $x, y \in \mathbb{R}$.

$$\begin{aligned} q(x, y, \varepsilon) &= \left| \begin{array}{cc} \varepsilon I_n & zI_n - A \\ \bar{z}I_n - A^* & \varepsilon I_n \end{array} \right| = \left| \varepsilon I_n - (\bar{z}I_n - A^*) \frac{1}{\varepsilon} I_n (zI_n - A) \right| \varepsilon^n \\ &= \left| \frac{1}{\varepsilon} (\varepsilon^2 I_n - (zI_n - A)^* (zI_n - A)) \right| \varepsilon^n = \frac{1}{\varepsilon^n} |\varepsilon^2 I_n - (zI_n - A)^* (zI_n - A)| \varepsilon^{2n}. \end{aligned}$$

\square

Observación 1.2.16. El polinomio $q(x, y, \varepsilon)$ mirado como elemento del anillo $\mathbb{R}[x, \varepsilon][y]$ es un polinomio en y de grado $2n$ donde el coeficiente del monomio y^{2n} vale -1 :

$$q(x, y, \varepsilon) = a_0(x, \varepsilon) + a_1(x, \varepsilon)y + a_2(x, \varepsilon)y^2 + \cdots + a_{2n-1}(x, \varepsilon)y^{2n-1} - y^{2n}.$$

A los efectos prácticos, cambiando de signo, $-q(x, y, \varepsilon)$ es un polinomio mónico en y cualquiera que sea el valor real de ε . Este hecho nos permite usar la versión simplificada del Teorema de Newton (véase el Teorema A.3.5) que aparece en el libro de Abhyankar [1].

Veamos algunas consecuencias del Lema 1.2.12. Dados $(x, y) \in \mathbb{R}^2$, $\varepsilon \geq 0$ y $A \in \mathbb{C}^{n \times n}$, consideramos el polinomio

$$f_{\varepsilon, A}(x, y) := \det \begin{pmatrix} \varepsilon I_n & -(x + yi)I_n + A \\ -(x - yi)I_n + A^* & \varepsilon I_n \end{pmatrix}, \quad (1.2.3)$$

que corresponde al polinomio $q(x, y, \varepsilon)$ definido en (1.2.2). Sea ahora

$$V_{\mathbb{R}}(f_{\varepsilon, A}) := \{(x, y) \in \mathbb{R}^2 : f_{\varepsilon, A}(x, y) = 0\}$$

la curva algebraica real definida por dicho polinomio. Entonces

Lema 1.2.17. *Con las notaciones anteriores se tiene*

- (a) $(x, y) \in V_{\mathbb{R}}(f_{\varepsilon, A})$ si y sólo si ε es un valor singular de $(x + yi)I_n - A$.
- (b) La curva algebraica $V_{\mathbb{R}}(f_{\varepsilon, A})$ está acotada. Concretamente está contenida en el disco cerrado de centro el origen 0 y radio $\varepsilon + \|A\|$.

DEMOSTRACIÓN. El apartado (a) es una consecuencia directa del Lema 1.2.12. Para probar (b) sea $(x_0, y_0) \in V_{\mathbb{R}}(f_{\varepsilon, A})$. Entonces, por (a), existe un vector unitario u tal que $\|((x_0 + y_0i)I_n - A)u\| = \varepsilon$. Luego

$$\begin{aligned} \|(x_0, y_0)\| &= |x_0 + y_0i| = \|(x_0 + y_0i)u\| = \|((x_0 + y_0i)I_n - A)u + Au\| \leq \\ &\leq \|((x_0 + y_0i)I_n - A)u\| + \|Au\| \leq \varepsilon + \|A\|. \end{aligned}$$

□

1.3. Índice y número de condición de un valor propio

Para introducir estos conceptos, precisamos del resultado siguiente.

Teorema 1.3.1 (Teorema de descomposición de Jordan). *Dada $A \in \mathbb{C}^{n \times n}$ sea $\Lambda(A) = \{\lambda_1, \dots, \lambda_r\}$ su espectro. Entonces*

$$A = \sum_{i=1}^r (\lambda_i P_i + N_i),$$

donde

(a) para cada $i \in \{1, \dots, r\}$, P_i es el proyector sobre el subespacio $\ker(\lambda_i I_n - A)^n$, siendo

$$\sum_{i=1}^r P_i = I_n, \text{ y } P_i P_j = 0 \text{ si } i \neq j;$$

(b) $N_i = (A - \lambda_i I_n)P_i$ son matrices nilpotentes, cumpliendo

$$N_i N_j = 0 \text{ si } i \neq j, \text{ y } N_i P_j = P_j N_i = \begin{cases} 0 & \text{si } i \neq j, \\ N_i & \text{si } i = j. \end{cases}$$

El subespacio $\ker(\lambda_i I_n - A)^n$ se conoce como el subespacio radical (o subespacio propio generalizado) de la matriz A asociado al valor propio λ_i , y el proyector P_i se conoce como proyector de Riesz sobre el subespacio radical asociado a λ_i . El orden de nilpotencia de N_i , es decir $\min\{l \in \mathbb{N} : N_i^l = 0\}$, se llama **índice** de λ_i , y se representa por $\nu(\lambda_i)$. Ese número corresponde con el tamaño del bloque de Jordan más grande asociado a λ_i en la forma de Jordan de A . También, dicho número es la multiplicidad de λ_i como raíz del polinomio mínimo de A .

Sea λ un valor propio de $A \in \mathbb{C}^{n \times n}$, de multiplicidad algebraica m . Sea otra matriz dada $A' \in \mathbb{C}^{n \times n}$. La **variación espectral** de A' respecto del valor propio λ de A , denotada por $\text{sv}_{(A, \lambda)}(A')$, está definida como el mínimo de los $\rho \geq 0$ tales que el disco cerrado $\mathcal{D}(\lambda, \rho)$ con centro λ y radio ρ contenga al menos m valores propios de A' (contando multiplicidades). Para fijar las ideas

observemos que si $\lambda \in \Lambda(A')$ y $m(\lambda, A') \geq m$, entonces $\text{sv}_{(A,\lambda)}(A') = 0$; pero si $m(\lambda, A') < m$, entonces

$$\text{sv}_{(A,\lambda)}(A') \geq \min\{|\mu - \lambda| : \mu \in \Lambda(A') \setminus \{\lambda\}\} > 0.$$

El **número de condición de Hölder del valor propio** $\lambda \in \Lambda(A)$ de **orden** $\omega > 0$ está definido por

$$\text{cond}_\omega(A, \lambda) := \lim_{\varepsilon \rightarrow 0^+} \max_{0 < \|X - A\| \leq \varepsilon} \frac{\text{sv}_{(A,\lambda)}(X)}{\|X - A\|^\omega}.$$

Nótese que el número $\text{cond}_\omega(A, \lambda) \in [0, \infty]$ está bien definido para cada $\omega > 0$; y también que hay a lo más un $\omega > 0$ tal que $0 \neq \text{cond}_\omega(A, \lambda) \neq \infty$. Véase Karow [44, Definición 2.6.3, pág. 33]. El límite que define $\text{cond}_\omega(A, \lambda)$ es un número positivo precisamente cuando $\omega = 1/\nu(\lambda)$.

1.4. Distancia de Hausdorff

La distancia de un punto $z \in \mathbb{C}$ a un subconjunto compacto no vacío K de \mathbb{C} viene definida por $\text{dist}(z, K) := \min_{w \in K} |w - z|$. Si K_1, K_2 son dos subconjuntos no vacíos de \mathbb{C} la suma (de Minkowski) viene definida por

$$K_1 + K_2 := \{z_1 + z_2 \mid z_1 \in K_1, z_2 \in K_2\}.$$

Para todo $\alpha \in \mathbb{R}$ y $K \subset \mathbb{C}$, se define $\alpha K := \{\alpha z \mid z \in K\}$. Especial valor tiene para nosotros el espacio métrico, $\mathcal{H}(\mathbb{C})$, formado por los subconjuntos compactos no vacíos de \mathbb{C} y cuya distancia viene definida por

$$d_H(K, L) := \max\left\{\max_{z \in K} \text{dist}(z, L), \max_{w \in L} \text{dist}(w, K)\right\},$$

para todo par (K, L) de elementos de $\mathcal{H}(\mathbb{C})$. Esta es la **distancia de Hausdorff**¹. Resumimos algunas propiedades esenciales de la distancia d_H . Para todo $K \in \mathcal{H}(\mathbb{C})$ y todo $\varepsilon > 0$ denotamos

$$V(K, \varepsilon) := \bigcup_{z \in K} D(z, \varepsilon),$$

siendo $D(z, \varepsilon) := \{w \in \mathbb{C} : |w - z| < \varepsilon\}$ el disco abierto de centro en $z \in \mathbb{C}$ y radio $\varepsilon > 0$.

Proposición 1.4.1. *Para cualesquiera $K, L \in \mathcal{H}(\mathbb{C})$,*

$$\begin{aligned} d_H(K, L) &= \sup_{z \in \mathbb{C}} |\text{dist}(z, K) - \text{dist}(z, L)| \\ &= \inf\{\varepsilon > 0 : K \subset V(L, \varepsilon), L \subset V(K, \varepsilon)\}. \end{aligned}$$

¹también atribuida a Pompeiu

La proposición siguiente nos permitirá asegurar que el espacio métrico $(\mathcal{H}(\mathbb{C}), d_H)$ es localmente compacto.

Proposición 1.4.2. *En el espacio métrico $(\mathcal{H}(\mathbb{C}), d_H)$, toda bola cerrada*

$$B(K, r) := \{K' \in \mathcal{H}(\mathbb{C}) : d_H(K', K) \leq r\}$$

con centro en $K \in \mathcal{H}(\mathbb{C})$ y radio $r \geq 0$ es compacta.

Sea (E, d) un espacio métrico. Se dice que una sucesión $(x_n)_{n=0}^\infty$ de puntos de E es una sucesión de Cauchy si para cada $\varepsilon > 0$ existe un entero positivo n_0 tal que para todos $m, n \geq n_0$ se tiene que $d(x_m, x_n) < \varepsilon$. Es obvio que toda sucesión convergente es sucesión de Cauchy. Un espacio métrico (E, d) se llama **completo** si toda sucesión de Cauchy en él es convergente. La siguiente Proposición es consecuencia de que $\mathbb{C} \equiv \mathbb{R}^2$ es un espacio métrico completo.

Proposición 1.4.3. *El espacio métrico $(\mathcal{H}(\mathbb{C}), d_H)$ es completo.*

De hecho, puede demostrarse que si $(K_n)_{n=0}^\infty$ es una sucesión de Cauchy en $\mathcal{H}(\mathbb{C})$, entonces definiendo

$$L := \{w \in \mathbb{C} : \text{existe una sucesión } (z_n)_{n=0}^\infty, \text{ con } z_n \in K_n, \text{ que converge a } w\}$$

se puede probar que L es un subconjunto compacto no vacío de \mathbb{C} y que

$$\lim_{n \rightarrow \infty} d_H(K_n, L) = 0.$$

Capítulo 2

Pseudoespectros de primer orden

2.1. Definiciones

Definición 2.1.1 ([3],[17]). Sea $A \in \mathbb{C}^{n \times n}$, para todo $\varepsilon \geq 0$ se define el **pseudoespectro algebraico ordinario de primer orden de nivel ε de la matriz A** como el conjunto

$$\Lambda_{\varepsilon,1}^a(A) := \bigcup_{\substack{\|X-A\| \leq \varepsilon \\ X \in \mathbb{C}^{n \times n}}} \Lambda(X).$$

Es decir, el pseudoespectro algebraico ordinario de nivel ε de la matriz A está formado por la unión de los espectros de matrices que disten de A una cantidad menor o igual que ε , respecto de la norma espectral. De igual manera, definimos el **pseudoespectro algebraico estricto de primer orden de nivel ε de la matriz A** como

$$\Lambda'_{\varepsilon,1}(A) := \bigcup_{\substack{\|X-A\| < \varepsilon \\ X \in \mathbb{C}^{n \times n}}} \Lambda(X).$$

Notemos que $\Lambda_{0,1}^a(A) = \Lambda(A)$ y $\Lambda'_{0,1}(A) = \emptyset$.

Para simplificar la exposición, vamos a aliviar de alguna manera las notaciones y conceptos anteriores. Así, cuando no haya lugar a equívocos, denotaremos por

$\Lambda_\varepsilon(A) := \Lambda_{\varepsilon,1}^a(A)$ y lo llamaremos abreviadamente pseudoespectro ordinario,
 $\Lambda'_\varepsilon(A) := \Lambda'_{\varepsilon,1}(A)$ y lo llamaremos abreviadamente pseudoespectro estricto.

Usando el Teorema 1.2.7, podemos caracterizar los pseudoespectros de otra manera. Véase por ejemplo [17]. Previamente vamos a introducir unas notaciones. Dados $A \in \mathbb{C}^{n \times n}$ y $\varepsilon > 0$, para $z \in \mathbb{C}$ definimos

$$h_1(z) := \sigma_n(zI_n - A). \quad (2.1.1)$$

Si identificamos $z = x + yi \in \mathbb{C}$, la notación dada en (2.1.1) se puede reescribir de la forma siguiente: para $(x, y) \in \mathbb{R}^2$ definimos

$$h_1(x, y) := \sigma_n((x + yi)I_n - A). \quad (2.1.2)$$

A lo largo del trabajo usaremos indistintamente $h_1(z)$ y $h_1(x, y)$ según convenga. Pasamos a dar otra caracterización de los pseudoespectros.

Proposición 2.1.2. *Con las notaciones anteriores, para todo $\varepsilon > 0$ se tiene*

$$\begin{aligned} \Lambda_\varepsilon(A) &= \{z \in \mathbb{C} : h_1(z) \leq \varepsilon\}, \\ \Lambda'_\varepsilon(A) &= \{z \in \mathbb{C} : h_1(z) < \varepsilon\}. \end{aligned} \quad (2.1.3)$$

DEMOSTRACIÓN.

Se probará, por doble inclusión, la igualdad correspondiente al conjunto $\Lambda_\varepsilon(A)$. El otro caso se hace de forma similar. Primero supongamos que $z_0 \in \Lambda_\varepsilon(A)$, entonces existe $X_0 \in \mathbb{C}^{n \times n}$ tal que $z_0 \in \Lambda(X_0)$ y $\|X_0 - A\| \leq \varepsilon$. Esto junto con el Teorema 1.2.7 nos conduce a que

$$h_1(z_0) = \min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda(X)}} \|X - A\| \leq \|X_0 - A\| \leq \varepsilon.$$

Por lo tanto $h_1(z_0) \leq \varepsilon$. Con lo cual, $\Lambda_\varepsilon(A) \subset \{z \in \mathbb{C} : h_1(z) \leq \varepsilon\}$.

Recíprocamente, sea $z_0 \in \mathbb{C}$ tal que $h_1(z_0) \leq \varepsilon$. Entonces, por el Teorema 1.2.7, resulta

$$h_1(z_0) = \sigma_n(z_0 I_n - A) = \min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda(X)}} \|X - A\|.$$

Esto implica que existe una $X_0 \in \mathbb{C}^{n \times n}$ tal que $h_1(z_0) = \|X_0 - A\| \leq \varepsilon$ y $z_0 \in \Lambda(X_0)$. Por lo tanto, $z_0 \in \Lambda_\varepsilon(A)$. En consecuencia, $\{z \in \mathbb{C} : h_1(z) \leq \varepsilon\} \subset \Lambda_\varepsilon(A)$. \square

2.2. Propiedades de los pseudoespectros de primer orden

Vamos a ver algunas propiedades de los pseudoespectros. Aunque son hechos bien conocidos, incluiremos algunas demostraciones, por razones de completitud. Véanse, por ejemplo, el portal de Internet *Pseudospectra Gateway*: [33]; o la Sección 5.1, páginas 68 y 69, de la Tesis de Karow [44].

Lema 2.2.1 (Invarianza por semejanza unitaria). Sean $A \in \mathbb{C}^{n \times n}$, $\varepsilon \geq 0$ y $U \in \mathbb{C}^{n \times n}$ una matriz unitaria. Entonces

$$\Lambda_\varepsilon(U^*AU) = \Lambda_\varepsilon(A), \quad \Lambda'_\varepsilon(U^*AU) = \Lambda'_\varepsilon(A). \quad (2.2.4)$$

Lema 2.2.2 (Invarianza por rotaciones, homotecias y traslaciones). Sean $A \in \mathbb{C}^{n \times n}$, $\alpha, \beta \in \mathbb{C}$ y $\varepsilon \geq 0$. Entonces

$$\begin{aligned} \Lambda_{\varepsilon|\beta|}(\alpha I_n + \beta A) &= \alpha + \beta \Lambda_\varepsilon(A), \\ \Lambda'_{\varepsilon|\beta|}(\alpha I_n + \beta A) &= \alpha + \beta \Lambda'_\varepsilon(A). \end{aligned} \quad (2.2.5)$$

DEMOSTRACIÓN. Probamos la igualdad correspondiente al conjunto $\Lambda_\varepsilon(A)$, demostrándose el otro caso de forma similar.

Si $\beta = 0$, la igualdad es obvia, pues ambos conjuntos se reducen a $\{\alpha\}$. Supongamos que $\beta \neq 0$. Si $z \in \Lambda_\varepsilon(A)$, entonces existe X tal que $z \in \Lambda(X)$ y $\|X - A\| \leq \varepsilon$; lo que implica que $\alpha + \beta z \in \Lambda(\alpha I_n + \beta X)$, con

$$\|\alpha I_n + \beta X - \alpha I_n - \beta A\| = |\beta| \|X - A\| \leq |\beta| \varepsilon;$$

de donde $\alpha + \beta z \in \Lambda_{\varepsilon|\beta|}(\alpha I_n + \beta A)$; por tanto

$$\alpha + \beta \Lambda_\varepsilon(A) \subset \Lambda_{\varepsilon|\beta|}(\alpha I_n + \beta A).$$

Recíprocamente, sea $w \in \Lambda_{\varepsilon|\beta|}(\alpha I_n + \beta A)$; esto implica que existe Y tal que $w \in \Lambda(Y)$ y

$$\|Y - \alpha I_n - \beta A\| \leq \varepsilon |\beta|; \quad (2.2.6)$$

lo que implica que $w - \alpha \in \Lambda(Y - \alpha I_n)$; de donde se obtiene que

$$\frac{1}{\beta}(w - \alpha) \in \Lambda\left(\frac{1}{\beta}Y - \frac{\alpha}{\beta}I_n\right).$$

Por lo tanto, teniendo en cuenta (2.2.6),

$$\frac{1}{|\beta|} \|Y - \alpha I_n - \beta A\| \leq \varepsilon, \text{ implica que } \frac{1}{\beta}(w - \alpha) \in \Lambda_\varepsilon(A);$$

por consiguiente existe $z \in \Lambda_\varepsilon(A)$ tal que $\frac{1}{\beta}(w - \alpha) = z$, de donde se obtiene que $w = \alpha + \beta z$ y $w \in \alpha + \beta \Lambda_\varepsilon(A)$; en consecuencia

$$\Lambda_{\varepsilon|\beta|}(\alpha I_n + \beta A) \subset \alpha + \beta \Lambda_\varepsilon(A).$$

□

Observación 2.2.3. Lo que nos dicen las identidades (2.2.4) y (2.2.5) es que a la hora de estudiar la geometría de los pseudoespectros *no se pierde generalidad* si hacemos alguna o varias de las transformaciones siguientes: (1) cambiar la matriz A por una *matriz unitariamente semejante* a ella $\Lambda_\varepsilon(U^*AU) = \Lambda_\varepsilon(A)$, con U matriz unitaria; (2) *rotar* el pseudoespectro y someterlo a una *homotecia*; (3) *trasladar* el pseudoespectro. De hecho, como se probará más adelante que la frontera de un pseudoespectro está contenida en una curva algebraica real, la naturaleza de sus puntos singulares no cambia al someter los pseudoespectros a *transformaciones afines*.

Vamos a englobar algunas propiedades de los pseudoespectros en la proposición siguiente.

Proposición 2.2.4. Sean $A \in \mathbb{C}^{n \times n}$ y $\varepsilon > 0$.

(a) $\Lambda_\varepsilon(A)$ es un compacto no vacío contenido en el disco cerrado de centro el origen 0 y radio $\varepsilon + \|A\|$.

(b) $\Lambda'_\varepsilon(A)$ es un abierto no vacío contenido en $\Lambda_\varepsilon(A)$.

(c) La función h_1 no es constante en ningún abierto de \mathbb{C} .

(d) La función

$$z \mapsto \|(zI_n - A)^{-1}\|$$

no es constante en ningún subconjunto abierto de $\mathbb{C} \setminus \Lambda(A)$.

(e) Los únicos mínimos locales de la función h_1 son los valores propios de la matriz A .

(f) $\overline{\Lambda'_\varepsilon(A)} = \Lambda_\varepsilon(A)$.

(g) Si $B \in \mathbb{C}^{m \times m}$ entonces

$$\Lambda_\varepsilon(\text{diag}(A, B)) = \Lambda_\varepsilon(A) \cup \Lambda_\varepsilon(B), \quad \Lambda'_\varepsilon(\text{diag}(A, B)) = \Lambda'_\varepsilon(A) \cup \Lambda'_\varepsilon(B).$$

(h)

$$\bigcup_{\lambda \in \Lambda(A)} \mathcal{D}(\lambda, \varepsilon) \subset \Lambda_\varepsilon(A).$$

DEMOSTRACIÓN. Observemos en primer lugar que el conjunto $\Lambda_\varepsilon(A)$ es no vacío puesto que contiene al espectro de A . Además, si $z_0 \in \Lambda_\varepsilon(A)$, entonces por (2.1.3) resulta que $h_1(z_0) \leq \varepsilon$. Ahora, por el Lema 1.2.17 (b) se tiene que $\Lambda_\varepsilon(A)$ está contenido en el disco cerrado de centro el origen 0 y radio $\varepsilon + \|A\|$. Por lo tanto es un conjunto acotado. De otro lado, a partir de (2.1.3), se tiene que $\Lambda_\varepsilon(A) = h_1^{-1}[0, \varepsilon]$ y $\Lambda'_\varepsilon(A) = h_1^{-1}(-1, \varepsilon)$. Luego, como por el Teorema 1.2.3 la función $h_1(z)$ es continua, resulta que la contraimagen de un cerrado (abierto) es un cerrado (abierto). Con esto probamos (a) y (b).

Vamos a probar el apartado (c), por reducción al absurdo. Supongamos que existen un número real c y un conjunto abierto $U \subset \mathbb{C}$ tales que $h_1(z) = c$, para todo $z \in U$. Si $c = 0$, para cada $z \in U$ se tiene $h_1(z) = \sigma_n(zI_n - A) = 0$; es decir, para todo $z \in U$, $z \in \Lambda(A)$, lo cual es imposible. Si $c \neq 0$, tomamos $w \in U$ y se tiene $h_1(w) = c > 0$. Por una parte, existe $X \in \mathbb{C}^{n \times n}$ tal que $\|X - A\| = c$ y $w \in \Lambda(X)$. Por otro lado, para cada entero $k \geq 1$, definimos la sucesión de matrices siguiente

$$X_k := \frac{1}{k}A + \left(1 - \frac{1}{k}\right)X.$$

Observemos que

$$\|X_k - A\| = \left\| \left(1 - \frac{1}{k}\right) (X - A) \right\| = \left(1 - \frac{1}{k}\right) \|X - A\| = \left(1 - \frac{1}{k}\right) c. \quad (2.2.7)$$

Como $X_k \rightarrow X$ cuando $k \rightarrow \infty$, por la continuidad del espectro se tiene que

$$\lim_{k \rightarrow \infty} \Lambda(X_k) = \Lambda(X),$$

respecto de la métrica Hausdorff. En consecuencia, existen una sucesión $z_k \rightarrow w$ y $k_0 \geq 1$ tal que, $\forall k \geq k_0$, se cumple que

$$z_k \in \Lambda(X_k) \cap U.$$

Como $\sigma_n(z_k I_n - X_k) = 0$, resulta que

$$\begin{aligned} h_1(z_k) &= \sigma_n(z_k I_n - A) = \sigma_n(z_k I_n - X_k + X_k - A) \leq \sigma_n(z_k I_n - X_k) + \|X_k - A\| \\ &= \|X_k - A\|. \end{aligned}$$

Si utilizamos (2.2.7), deducimos que $h_1(z_k) \leq \left(1 - \frac{1}{k}\right) c < c$ y que $z_k \in U$. Lo que es una contradicción.

Otra prueba del apartado (c) es la siguiente. Supongamos nuevamente que existe $z_0 \in \mathbb{C}$ y $\delta > 0$ de manera que para todo $|z - z_0| < \delta$ se cumple $h_1(z) = h_1(z_0)$. Si $h_1(z_0) = 0$ entonces $h_1(z) = 0$, para todo $|z - z_0| < \delta$. En consecuencia $\sigma_n(z I_n - A) = 0$. Luego $\text{rg}(z I_n - A) < n$; de donde se sigue que $z \in \Lambda(A)$ para todo $z \in D(z_0, \delta)$. Lo que es absurdo.

Analicemos el caso $h_1(z_0) = c > 0$. Identificando \mathbb{C} con \mathbb{R}^2 , existe $\delta > 0$ de manera que si

$$\|(x - x_0, y - y_0)\| < \delta \Rightarrow h_1(x, y) = c > 0,$$

por lo tanto c es un valor singular de $(x + yi)I_n - A$, para todo (x, y) tal que $\|(x - x_0, y - y_0)\| < \delta$. Sea $f_{c,A}(x, y)$ el polinomio de Wielandt de la matriz A para el nivel c . Sea $\alpha > 0$ un número real suficientemente pequeño para que el cuadrado $C := [x_0 - \alpha, x_0 + \alpha] \times [y_0 - \alpha, y_0 + \alpha]$ esté contenido en el disco $D((x_0, y_0), \delta)$ de \mathbb{R}^2 de centro (x_0, y_0) y radio δ .

Ahora, por el Lema 1.2.17 (a) y teniendo en cuenta que

$$f_{c,A}(x, y) = 0$$

para todo $(x, y) \in C$, se sigue¹ que $f_{c,A} = 0$ (polinomio cero). Lo que contradice el aserto de la Observación 1.2.16 de que $-q(x, y, c) := -f_{c,A}(x, y)$ es un polinomio mónico² de grado $2n$ en y .

El apartado (d) es una consecuencia de (c). Se puede ver una prueba directa de (d) en [2, Proposición 9]. Como corolario se probaría (c).

¹Sea K un cuerpo, y sean S_1, \dots, S_n subconjuntos infinitos de K . Sea $f(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$. Si $f(a_1, \dots, a_n) = 0$ para todos $a_i \in S_i (i = 1, \dots, n)$, entonces $f = 0$. Véase el Corolario 1.6 de la página 176 del libro[49] de Lang.

²Se dice que un polinomio $f(y) \in K[y]$ es **mónico** si el coeficiente del término de mayor grado es 1.

Para probar (e), observemos en primer lugar que, como $h_1(z) \geq 0$, para todo $z \in \mathbb{C}$ y además $h_1(z) = 0$ si y sólo si z es un valor propio de A , entonces $h_1(z)$ tiene un mínimo relativo en cada valor propio de A . Luego, para probar este apartado, basta probar que $\|(zI_n - A)^{-1}\|$ no alcanza ningún máximo relativo en el conjunto $\mathbb{C} \setminus \Lambda(A)$. Supongamos lo contrario, es decir supongamos que existe un $z_0 \notin \Lambda(A)$ tal que $\|(zI_n - A)^{-1}\|$ alcanza un máximo relativo en z_0 . Ahora, como existen vectores unitarios $u, v \in \mathbb{C}^n$ tales que

$$\|(z_0I_n - A)^{-1}\| = |u^*(z_0I_n - A)^{-1}v|,$$

para todo $z \simeq z_0$ tenemos las desigualdades

$$|u^*(zI_n - A)^{-1}v| \leq \|(zI_n - A)^{-1}\| \leq \|(z_0I_n - A)^{-1}\| = |u^*(z_0I_n - A)^{-1}v|.$$

Por lo tanto, el módulo de la función analítica $z \mapsto u^*(zI_n - A)v$ tendría un máximo relativo en z_0 . Con lo cual, por el principio del módulo máximo, es constante en un entorno de z_0 . Lo que contradice el apartado (d).

Veamos la prueba del apartado (f). Por los apartados (a) y (b) se tiene inmediatamente que $\overline{\Lambda'_\varepsilon(A)} \subset \Lambda_\varepsilon(A)$. Para probar la otra inclusión, supongamos que existe $z_0 \in \Lambda_\varepsilon(A) \setminus \overline{\Lambda'_\varepsilon(A)}$. Entonces

$$\left. \begin{array}{l} z_0 \notin \overline{\Lambda'_\varepsilon(A)} \Rightarrow z_0 \notin \Lambda'_\varepsilon(A) \Rightarrow h_1(z_0) \geq \varepsilon, \\ z_0 \in \Lambda_\varepsilon(A) \Rightarrow h_1(z_0) \leq \varepsilon \end{array} \right\} \Rightarrow h_1(z_0) = \varepsilon.$$

Como $z_0 \notin \overline{\Lambda'_\varepsilon(A)}$, existe un $\delta > 0$ tal que el $D(z_0, \delta) \cap \Lambda'_\varepsilon(A) = \emptyset$. Por lo tanto, $\forall z \in D(z_0, \delta)$ resulta que $h_1(z) \geq \varepsilon$, lo que implica que h_1 tiene un mínimo local en z_0 . Luego por (e), resulta que $z_0 \in \Lambda(A)$ y, por lo tanto, $h_1(z_0) = 0$, lo cual es absurdo.

(g) se deduce directamente de la propiedad $\Lambda(\text{diag}(A, B)) = \Lambda(A) \cup \Lambda(B)$. Finalmente, para probar (h), sea $z_0 \in \mathcal{D}(\lambda_0, \varepsilon)$; es decir $|z_0 - \lambda_0| \leq \varepsilon$. Sea $u \in \mathbb{C}^n$ un vector unitario tal que $Au = \lambda_0 u$. Entonces

$$\|(z_0I_n - A)u\| = \|(z_0 - \lambda_0)u\| = |z_0 - \lambda_0| \leq \varepsilon.$$

Por lo tanto $h_1(z_0) \leq \varepsilon$, lo que implica que $z_0 \in \Lambda_\varepsilon(A)$ por (2.1.3). \square

Observación 2.2.5. Para matrices $A \in \mathbb{C}^{n \times n}$ se tiene que $\overline{\Lambda'_\varepsilon(A)} = \Lambda_\varepsilon(A)$. Pero si T es un operador lineal acotado actuando sobre un espacio de Banach de dimensión infinita X , el contenido

$$\overline{\{z \in \mathbb{C} : \|(zI - T)^{-1}\| > 1/\varepsilon\}} \subset \{z \in \mathbb{C} : \|(zI - T)^{-1}\| \geq 1/\varepsilon\}$$

puede ser *estricto*. De hecho, en el artículo de Shargorodsky [65], pueden verse ejemplos de operadores T tales que

$$\bigcup_{\|\Delta\| \leq \varepsilon} \Lambda(T + \Delta) \not\subseteq \{z \in \mathbb{C} : \|(zI - T)^{-1}\| \geq 1/\varepsilon\},$$

e incluso de operadores S tales que

$$\overline{\bigcup_{\|\Delta\| \leq \varepsilon} \Lambda(S + \Delta)} \not\subseteq \{z \in \mathbb{C} : \|(zI - S)^{-1}\| \geq 1/\varepsilon\}.$$

Estos ejemplos significan que hay un salto significativo al pasar de los pseudo-espectros de operadores en espacios de dimensión finita a infinita, como podía esperarse.

2.3. Componentes conexas

Dado un subconjunto cualquiera S de \mathbb{C} , recordemos que un punto $z_0 \in \mathbb{C}$ se llama **punto frontera** de S si en todo entorno $D(z_0, r)$, $r > 0$, de z_0 hay puntos de S y puntos que no están en S . Se llama la **frontera** de S al conjunto de sus puntos frontera; la denotaremos por $\partial(S)$ o ∂S si no hubiera lugar a confusión.

Es conocido que los pseudoespectros de primer orden están compuestos por la unión de componentes conexas, abiertas para el estricto y cerradas para el ordinario. Veamos algunas propiedades.

Teorema 2.3.1 (Pseudoespectros estrictos). *Sea $\varepsilon > 0$, para cada $X \in \mathbb{C}^{n \times n}$ tal que $\|X - A\| < \varepsilon$, se tiene que toda componente conexa T de $\Lambda'_\varepsilon(A)$ contiene al menos un valor propio de X .*

DEMOSTRACIÓN. Sea z_0 un elemento cualquiera de T ; por la definición de pseudoespectro estricto, existe una matriz X_0 tal que $\|X_0 - A\| < \varepsilon$ y z_0 es valor propio de X_0 . Sea $Z(t) := X_0 + t(X - X_0)$, $t \in [0, 1]$. Es claro que para todo $t \in [0, 1]$, $Z(t) \in B(A, \varepsilon)$, pues toda bola en un espacio normado es convexo [48, pág. 356, Ejercicio 2]; por tanto,

$$\Lambda(Z(t)) \subset \Lambda'_\varepsilon(A).$$

Por Bhatia [12], Corolario VI.1.6, página 155, y Kato [45], Teorema 5.2, página 126, existen funciones continuas

$$\lambda_1, \dots, \lambda_n: [0, 1] \rightarrow \mathbb{C}$$

tales que para cada $t \in [0, 1]$,

$$\Lambda(Z(t)) = \{\lambda_1(t), \lambda_2(t), \dots, \lambda_n(t)\} \quad (\text{puede haber repeticiones}).$$

Entonces,

$$\Lambda(X_0) = \{\lambda_1(0), \lambda_2(0), \dots, \lambda_n(0)\};$$

por consiguiente, existe un índice $i \in \{1, \dots, n\}$ tal que $z_0 = \lambda_i(0)$. Como

$$\Lambda(X) = \{\lambda_1(1), \lambda_2(1), \dots, \lambda_n(1)\},$$

se sigue que $\lambda_i(1) \in \Lambda(X)$, y dado que existe un camino continuo $\lambda_i(t)$ que conecta $z_0 = \lambda_i(0)$ con $\lambda_i(1)$ dentro de $\Lambda'_\varepsilon(A)$, se tiene que $\lambda_i(1) \in T$. Así pues, T contiene un valor propio de X . \square

Observación 2.3.2. En particular, el teorema que acabamos de demostrar implica que *toda componente conexa del pseudoespectro estricto, $\Lambda'_\varepsilon(A)$, contiene al menos un valor propio de A* . Otros autores demuestran primero esta aserción recurriendo al principio del módulo máximo, [17], [32].

Proposición 2.3.3. *Para todo $\varepsilon > 0$ las componentes conexas de $\Lambda'_\varepsilon(A)$ son conjuntos abiertos conexos por arcos y cada componente conexa tiene al menos un valor propio de A .*

DEMOSTRACIÓN. Como $\Lambda'_\varepsilon(A)$ es un conjunto abierto, la proposición se sigue de los Teoremas B.3.3 y B.3.5. \square

Proposición 2.3.4 (Pseudoespectros ordinarios. Karow [44], Proposición 2.2.3 (i), página 17). *Para cada $\varepsilon \geq 0$, las componentes conexas de $\Lambda_\varepsilon A$ son conjuntos cerrados conexos por arcos y cada componente conexa contiene al menos un valor propio de A .*

Proposición 2.3.5 (Pseudoespectros ordinarios. Karow [44], Proposición 2.2.5 (ii), página 18). *Sean $\varepsilon > 0$ y z un punto de la frontera de $\Lambda_\varepsilon(A)$. Entonces*

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ z \in \Lambda(X)}} \|X - A\| = \varepsilon.$$

2.4. Conservación de la multiplicidad. Puntos de colisión

Según vimos en la Observación 2.3.2 del Teorema 2.3.1 y la Proposición 2.3.4 cada componente conexa de los pseudoespectros (ordinario y estricto de una matriz A), tiene al menos un valor propio de A . Por lo tanto el número de componentes conexas es finito. Para cada nivel $\varepsilon \geq 0$, el número de ellas viene dado por las siguientes funciones.

Definición 2.4.1. Para cada $\varepsilon \geq 0$ llamaremos $\rho(\varepsilon)$ al número de componentes conexas de $\Lambda_\varepsilon(A)$. De igual manera, para cada $\varepsilon \geq 0$, $\rho'(\varepsilon)$ denotará el número de componentes conexas de $\Lambda'_\varepsilon(A)$. Quedan así definidas dos funciones $\rho: [0, \infty) \rightarrow \mathbb{R}$ y $\rho': [0, \infty) \rightarrow \mathbb{R}$. *Nótese que $\rho'(\varepsilon)$ no denota la derivada de $\rho(\varepsilon)$, sino que la prima ' se añade por coherencia con la notación $\Lambda'_\varepsilon(A)$ para el pseudoespectro estricto de A de nivel ε . Véanse las Figuras 2.1 y 2.2.*

Lema 2.4.2 (Propiedades de las funciones ρ y ρ' de $[0, \infty)$ en \mathbb{R}). *Las funciones ρ y ρ' gozan de las propiedades siguientes.*

- (I) *Las funciones $\rho: [0, \infty) \rightarrow \mathbb{R}$ y $\rho': (0, \infty) \rightarrow \mathbb{R}$ son decrecientes en los intervalos indicados, i.e. $\alpha < \beta$ implica que $\rho(\alpha) \geq \rho(\beta)$, $\rho'(\alpha) \geq \rho'(\beta)$.*
- (II) *$\rho(0) = |\Lambda(A)|$ (cardinal del espectro de A). Por el principio del palomar de Dirichlet, la función ρ sólo tiene un número finito de discontinuidades; al menos una si $|\Lambda(A)| \geq 2$; ninguna si $|\Lambda(A)| = 1$. Por otro lado, $\rho'(0) = 0$ pues $\Lambda'_0(A) = \emptyset$. La función ρ' es discontinua en 0; si $|\Lambda(A)| = 1$, ésta es su única discontinuidad.*

- (III) Sean $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q$ los puntos de discontinuidad de ρ , donde $0 < \varepsilon_1 < \varepsilon_2 < \dots < \varepsilon_q$. Si $|\Lambda(A)| \geq 2$, las funciones ρ y ρ' son discontinuas en los mismos puntos $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q$ de $(0, \infty)$. En los puntos $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q$ la función ρ' es continua por la izquierda y la función ρ es continua por la derecha.
- (IV) Para todo $\varepsilon \in [0, \infty) \setminus \{0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_q\}$, se tiene que $\rho(\varepsilon) = \rho'(\varepsilon)$. Sin embargo, $\rho(0) > \rho'(0)$ y para cada $i = 1, \dots, q$ se tiene que $\rho'(\varepsilon_i) > \rho(\varepsilon_i)$.
- (V) Existe un número real $E > 0$ tal que para todo $\varepsilon \geq E$

$$\rho'(\varepsilon) = 1 = \rho(\varepsilon).$$

Es decir, las funciones ρ' y ρ son finalmente constantes. Si fuera $|\Lambda(A)| = 1$ la función ρ sería constante en $[0, \infty)$ e igual a 1.

Para cada $\varepsilon \geq 0$, toda componente conexa \mathcal{C} de $\Lambda_\varepsilon(A)$ es un conjunto cerrado. Esto es debido a que la clausura de cualquier conjunto conexo es conexa. Véase el Teorema B.2.4. Por tanto, ya que $\mathcal{C} \subset \overline{\mathcal{C}}$, el conjunto $\Lambda_\varepsilon(A)$ es cerrado y $\mathcal{C} \subset \Lambda_\varepsilon(A)$, se tiene que $\overline{\mathcal{C}} \subset \Lambda_\varepsilon(A)$. Y como las componentes conexas de $\Lambda_\varepsilon(A)$ son los subconjuntos conexos maximales, se sigue que $\mathcal{C} = \overline{\mathcal{C}}$; es decir, \mathcal{C} es cerrado.

Proposición 2.4.3. Sea $A \in \mathbb{C}^{n \times n}$ y sea $\varepsilon > 0$ tal que $\rho(\varepsilon) = \rho'(\varepsilon)$. Sean $\mathcal{C}(\varepsilon)$ y $\mathcal{C}'(\varepsilon)$ componentes conexas de $\Lambda_\varepsilon(A)$ y de $\Lambda'_\varepsilon(A)$, respectivamente, que contienen a un mismo valor propio μ de A . Entonces

$$\overline{\mathcal{C}'(\varepsilon)} = \mathcal{C}(\varepsilon). \quad (2.4.8)$$

DEMOSTRACIÓN. En primer lugar veamos que $\mathcal{C}'(\varepsilon) \subset \mathcal{C}(\varepsilon)$. En efecto, $\mathcal{C}'(\varepsilon)$ es el conjunto de puntos z de $\Lambda'_\varepsilon(A)$ que se pueden conectar a μ mediante un arco contenido en $\Lambda'_\varepsilon(A)$. Como $\Lambda'_\varepsilon(A) \subset \Lambda_\varepsilon(A)$, estos puntos z son también puntos de $\Lambda_\varepsilon(A)$ que se pueden conectar a μ mediante un arco contenido en $\Lambda_\varepsilon(A)$.

Denotemos las componentes conexas del pseudoespectro ordinario de nivel ε de A por $\mathcal{C}_i(\varepsilon)$, $i = 1, \dots, \rho(\varepsilon)$, y por $\mathcal{C}'_i(\varepsilon)$, $i = 1, \dots, \rho(\varepsilon)$, las del estricto, de manera que $\mathcal{C}'_i(\varepsilon)$ y $\mathcal{C}_i(\varepsilon)$ tengan un valor propio común λ_i de A para $i = 1, 2, \dots, \rho(\varepsilon)$; donde el conjunto de $\rho(\varepsilon)$ elementos

$$\{\lambda_1, \lambda_2, \dots, \lambda_{\rho(\varepsilon)}\} \subset \Lambda(A).$$

Teniendo en cuenta que

$$\Lambda_\varepsilon(A) = \mathcal{C}_1(\varepsilon) \dot{\cup} \mathcal{C}_2(\varepsilon) \dot{\cup} \dots \dot{\cup} \mathcal{C}_{\rho(\varepsilon)}(\varepsilon), \quad (2.4.9)$$

$$\Lambda'_\varepsilon(A) = \mathcal{C}'_1(\varepsilon) \dot{\cup} \mathcal{C}'_2(\varepsilon) \dot{\cup} \dots \dot{\cup} \mathcal{C}'_{\rho(\varepsilon)}(\varepsilon), \quad (2.4.10)$$

donde $\dot{\cup}$ denota la unión disjunta, que $\overline{\Lambda'_\varepsilon(A)} = \Lambda_\varepsilon(A)$ y que la clausura de una unión finita de subconjuntos S_1, S_2, \dots, S_r de \mathbb{C} es igual a la unión de sus clausuras, obtenemos lo siguiente:

$$\Lambda_\varepsilon(A) = \overline{\mathcal{C}'_1(\varepsilon)} \cup \overline{\mathcal{C}'_2(\varepsilon)} \cup \dots \cup \overline{\mathcal{C}'_{\rho(\varepsilon)}(\varepsilon)}. \quad (2.4.11)$$

Dado que para todo $i = 1, 2, \dots, \rho(\varepsilon)$ se tiene que $\overline{\mathcal{C}'_i(\varepsilon)} \subset \mathcal{C}_i(\varepsilon)$, comparando las uniones (2.4.9) y (2.4.11), se tiene que

$$\begin{aligned} \overline{\mathcal{C}'_1(\varepsilon)} &= \mathcal{C}_1(\varepsilon) \\ &\vdots \\ \overline{\mathcal{C}'_{\rho(\varepsilon)}(\varepsilon)} &= \mathcal{C}_{\rho(\varepsilon)}(\varepsilon) \end{aligned}$$

□

Hemos visto algunas propiedades de $\rho(\varepsilon)$ y $\rho'(\varepsilon)$. Ambas funciones están relacionadas con los llamados puntos de colisión de componentes conexas. Damos su definición.

Definición 2.4.4. Sean $A \in \mathbb{C}^{n \times n}$, $z_0 \in \mathbb{C} \setminus \Lambda(A)$. Decimos que z_0 es un **punto de colisión (o de coalescencia) de los pseudoespectros** de A si z_0 es un punto frontera de dos componentes conexas del pseudoespectro estricto $\Lambda'_{\sigma_n(z_0 I - A)}(A)$.

Observación 2.4.5. Sean $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q$ los puntos de discontinuidad de las funciones ρ y ρ' en $(0, \infty)$, entonces los puntos de colisión de los pseudoespectros de A están en las fronteras de los pseudoespectros $\Lambda'_{\varepsilon_1}(A), \Lambda'_{\varepsilon_2}(A), \dots, \Lambda'_{\varepsilon_q}(A)$.

Definición 2.4.6. Se llama **primer punto de colisión de los pseudoespectros** de A a cualquier punto de colisión que esté en la frontera de $\Lambda'_{\varepsilon_1}(A)$.

Proposición 2.4.7. Si $\varepsilon_k \in \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q\}$ y $\mathcal{C}'_{i_1}(\varepsilon_k), \dots, \mathcal{C}'_{i_k}(\varepsilon_k)$ son las componentes conexas de $\Lambda'_{\varepsilon_k}(A)$ cuyas fronteras comparten puntos de colisión de los pseudoespectros de A correspondientes al nivel $\varepsilon = \varepsilon_k$, entonces

$$\overline{\mathcal{C}'_{i_1}(\varepsilon_k)} \cup \dots \cup \overline{\mathcal{C}'_{i_k}(\varepsilon_k)} = \mathcal{C}_j(\varepsilon_k),$$

donde $\mathcal{C}_j(\varepsilon_k)$ es una de las $\rho(\varepsilon_k)$ componentes conexas de $\Lambda_{\varepsilon_k}(A)$.

Observación 2.4.8. Si el número $\varepsilon > 0$ es suficientemente grande, en la Proposición 2.4.3 las componentes $\mathcal{C}(\varepsilon)$ y $\mathcal{C}'(\varepsilon)$ pueden tener en común más de un valor propio de A .

Las componentes conexas cumplen la siguiente propiedad de *conservación de las multiplicidades*, según probaron J. M. Gracia [38], Teorema 16 y M. Karow [44], Proposición 2.2.3 (iii).

Proposición 2.4.9. Supongamos que $\Lambda_\varepsilon(A) = \mathcal{C}_1(\varepsilon) \dot{\cup} \mathcal{C}_2(\varepsilon) \dot{\cup} \dots \dot{\cup} \mathcal{C}_{\rho(\varepsilon)}(\varepsilon)$. Entonces, para $k = 1, 2, \dots, \rho(\varepsilon)$ y para toda matriz X tal que $\|X - A\| \leq \varepsilon$ se cumple

$$\sum_{\xi \in \Lambda(X) \cap \mathcal{C}_k(\varepsilon)} m(\xi, X) = \sum_{\alpha \in \Lambda(A) \cap \mathcal{C}_k(\varepsilon)} m(\alpha, A).$$

Análogamente para el pseudoespectro estricto $\Lambda'_\varepsilon(A)$, se tiene la siguiente proposición.

Proposición 2.4.10. *Supongamos que $\Lambda'_\varepsilon(A) = \mathcal{C}'_1(\varepsilon) \dot{\cup} \mathcal{C}'_2(\varepsilon) \dot{\cup} \dots \dot{\cup} \mathcal{C}'_{\rho'(\varepsilon)}(\varepsilon)$. Entonces, para $k = 1, 2, \dots, \rho'(\varepsilon)$ y para toda matriz X tal que $\|X - A\| < \varepsilon$ se cumple*

$$\sum_{\xi \in \Lambda(X) \cap \mathcal{C}'_k(\varepsilon)} m(\xi, X) = \sum_{\alpha \in \Lambda(A) \cap \mathcal{C}'_k(\varepsilon)} m(\alpha, A).$$

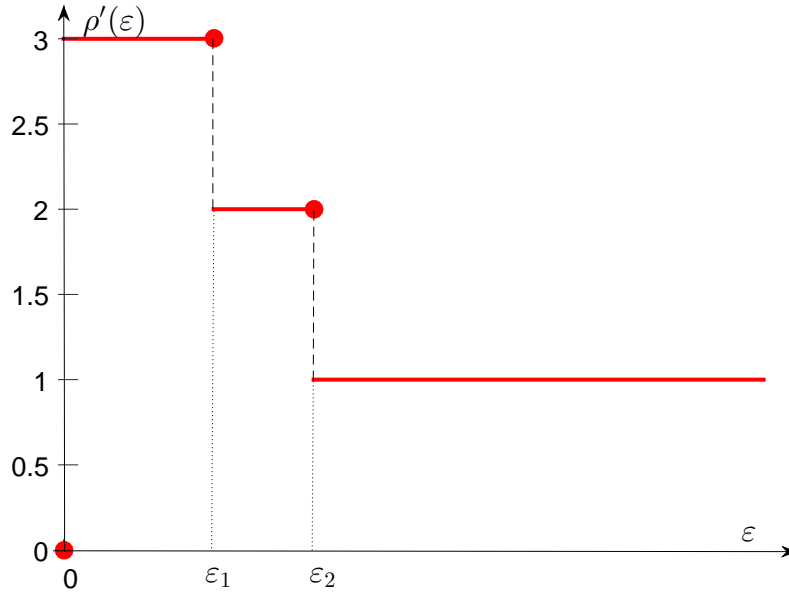
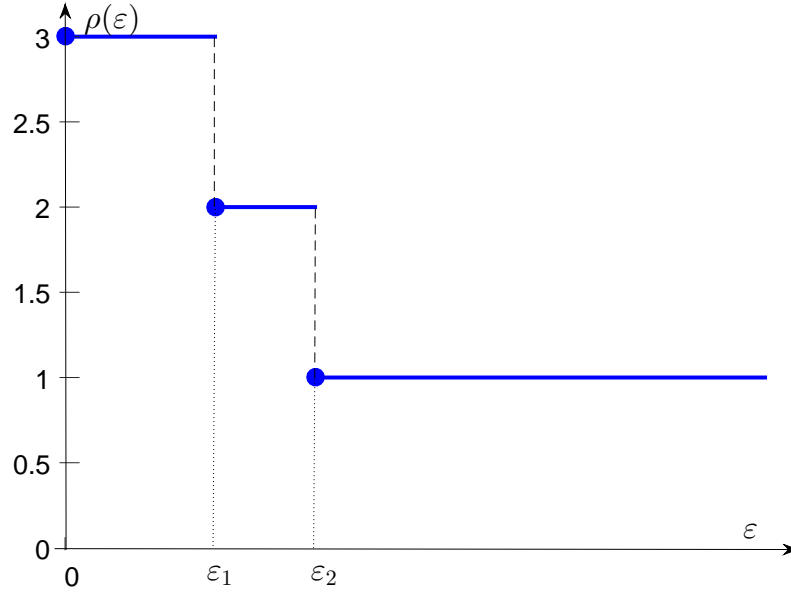


Figura 2.1: Gráfica de una función ρ' .

Observemos que $\Lambda_0(A) = \Lambda(A)$ tiene como componentes conexas los valores propios de A y $\Lambda'_0(A) = \emptyset$. Según crece ε , el pseudoespectro $\Lambda'_\varepsilon(A)$ tiene por componentes conexas unos conjuntos abiertos conexos que contienen algún valor propio de A . De manera intuitiva podemos decir que cuando las fronteras de dos componentes conexas se juntan, el punto común a ambas fronteras se llama punto de colisión; es decir, un punto de colisión es un punto común a la frontera de dos componentes conexas de $\Lambda'_\varepsilon(A)$. En el resultado siguiente se dan algunas propiedades sobre el número de componentes conexas de $\Lambda'_\varepsilon(A)$ y el número de puntos de colisión.

Teorema 2.4.11. *Dada $A \in \mathbb{C}^{n \times n}$, y con las notaciones de la Proposición 2.4.10 se tienen las siguientes propiedades.*

- (a) *El conjunto $\Lambda_{\|A\|}(A)$ tiene sólo una componente conexa.*
- (b) *El conjunto de puntos de colisión de pseudoespectros de A es finito.*

Figura 2.2: Gráfica de una función ρ .

DEMOSTRACIÓN.

Comencemos con el apartado (a). Observemos que si $\lambda \in \Lambda(A)$, entonces $|\lambda| \leq \|A\|$. En efecto, si $\lambda \in \Lambda(A)$, existe un vector unitario x tal que $Ax = \lambda x$. Por lo tanto,

$$|\lambda| = \|\lambda x\| = \|Ax\| \leq \|A\| \|x\| = \|A\|.$$

Por la Proposición 2.2.4 (h) resulta,

$$\Lambda_{\|A\|}(A) \supseteq \bigcup_{\lambda \in \Lambda(A)} \mathcal{D}(\lambda, \|A\|).$$

Observamos que el término derecho de la anterior expresión es un conexo, ya que $\Lambda(A)$ está contenido en el disco cerrado de centro 0 y radio $\|A\|$. Por lo tanto, si $\Lambda_{\|A\|}(A)$ tiene al menos dos componentes conexas, una de ellas no tiene en su interior ningún valor propio de A , lo que es una contradicción.

Para probar (b), usamos el siguiente resultado, véase [3, Proposición 4.2] : *el número de puntos de colisión de dos componentes conexas de $\Lambda'_\varepsilon(A)$ es finito*. Esencialmente el argumento de Alam y Bora radica en la observación de que: (1) la frontera $\partial\Lambda'_{\varepsilon_k}(A)$ está contenida en la curva algebraica real $q(x, y, \varepsilon_k) = 0$, donde $f(x, y) := q(x, y, \varepsilon_k) \in \mathbb{R}[x, y]$ es el polinomio de Wielandt de la matriz A para $\varepsilon = \varepsilon_k$; (2) todo punto de coalescencia de pseudoespectros de A es un punto singular de una de las curvas $q(x, y, \varepsilon_k) = 0$, $k = 1, \dots, q$; (3) la curva $V_{\mathbb{C}}(f)$ en el plano afín \mathbb{C}^2 admite sólo un número finito de punto singulares, por

tanto la curva $V_{\mathbb{R}}(f)$ admite un número finito de puntos singulares en el plano afín \mathbb{R}^2 . \square

Definición 2.4.12. Dada una matriz $A \in \mathbb{C}^{n \times n}$, decimos que es una **matriz simple si todos sus valores propios son simples**.

Los puntos de colisión tienen unas propiedades a destacar. La primera es que dan una solución al *Problema de Wilkinson* sobre la distancia de una matriz al conjunto de matrices que tienen un valor propio múltiple. El resultado, que enunciamos a continuación, puede verse en [3].

Teorema 2.4.13 (Alam-Bora). *Sean $A \in \mathbb{C}^{n \times n}$ una matriz simple, $z_0 \notin \Lambda(A)$ y $\sigma_0 := h_1(z_0) = \sigma_n(z_0 I_n - A)$.*

1. *Supongamos que z_0 es un punto de colisión de los pseudoespectros de A . Entonces si σ_0 es un valor singular simple de $z_0 I_n - A$ se tiene que:*

- 1-1 *existen un par de vectores singulares ortogonales u, v para σ_0 . Además, para la matriz*

$$A' = A - \sigma_0 uv^*$$

se tiene que:

- 1-2 $\|A' - A\| = \sigma_0$.

- 1-3 z_0 es un valor propio no derogatorio y defectuoso de A' , es decir, tiene una cadena de Jordan de longitud al menos 2.

2. *Supongamos que z_0 es un punto de colisión de los pseudoespectros de A y que σ_0 es un valor singular múltiple de $z_0 I_n - A$ con multiplicidad m . Sean $U, V \in \mathbb{C}^{n \times m}$ matrices cuyas columnas son vectores singulares para σ_0 y tales que $U^*U = I_m$ y $V^*V = I_m$. Entonces, para la matriz*

$$A' = A - \sigma_0 UV^*$$

se tiene:

- 2-1 $\|A' - A\| = \sigma_0$.

- 2-2 z_0 es un valor propio múltiple de A' , cuya multiplicidad geométrica es m .

Sean $b, c \in \mathbb{C}$, al disco abierto de centro c y que tiene a b en su frontera lo denotaremos por $D(c, b)$. Si $c = b$, el disco es vacío, en caso contrario su radio es $|b - c|$.

Otra propiedad de los puntos de colisión z_0 es que tienen un par de vectores singulares ortogonales asociados al valor singular $h_1(z_0)$. Antes de demostrarla, introducimos un teorema auxiliar, cuya demostración se puede encontrar en [4].

Teorema 2.4.14 (Alam-Bora-Byers-Overton). *Sean $A \in \mathbb{C}^{n \times n}$ una matriz simple y z_0 un primer punto de coalescencia de los pseudoespectros de A . Sean Ω y $\hat{\Omega}$ dos componentes conexas del pseudoespectro $\Lambda'_{\varepsilon_1}(A)$, con $\varepsilon_1 := h_1(z_0)$, en cuyas*

fronteras está z_0 . Supongamos que no existe ninguna sucesión $\{z_k\} \subset \Lambda'_{\varepsilon_1}(A)$ tal que $z_k \rightarrow z_0$ y $\nabla h_1(z_k) \rightarrow 0$ cuando $k \rightarrow \infty$, siendo $\nabla h_1(z) = (h'_{1x}(z), h'_{1y}(z))$, con $z = x + yi$, donde $x, y \in \mathbb{R}$. Entonces:

1. Sean $z_k \in \Omega$ y $\hat{z}_k \in \hat{\Omega}$ sucesiones que convergen a z_0 . Sean (u_k, v_k) y (\hat{u}_k, \hat{v}_k) pares de vectores singulares para $h_1(z_k)$ y $h_1(\hat{z}_k)$ respectivamente. Supongamos, tomando subsucesiones si es preciso, que

$$(u_k, v_k) \rightarrow (u, v) \quad \text{y} \quad (\hat{u}_k, \hat{v}_k) \rightarrow (\hat{u}, \hat{v}).$$

Entonces (u, v) y (\hat{u}, \hat{v}) son pares de vectores singulares de $h_1(z_0)$ y, además los discos abiertos

$$D(v^*Av, z_0), \quad D(\hat{v}^*A\hat{v}, z_0)$$

son no vacíos y están contenidos en Ω y $\hat{\Omega}$, respectivamente.

2. z_0 no pertenece a la frontera de una tercera componente conexa de $\Lambda'_{\varepsilon_1}(A)$.
3. Los vectores v, \hat{v} son linealmente independientes. Además, $h_1(z_0) = \varepsilon_1$ es un valor singular de multiplicidad 2 de la matriz $z_0 I_n - A$.
4. Sean $g := -v^*u$ y $\hat{g} := -\hat{v}^*\hat{u}$. Entonces, existen los límites

$$\lim_{k \rightarrow \infty} \nabla h_1(z_k) \quad \text{y} \quad \lim_{k \rightarrow \infty} \nabla h_1(\hat{z}_k),$$

y son iguales a g y \hat{g} , respectivamente. Además existe $\mu \in (0, 1)$ tal que

$$\mu g + (1 - \mu)\hat{g} = 0.$$

Observación 2.4.15. Las clausuras de los discos $D(v^*Av, z_0)$ y $D(\hat{v}^*A\hat{v}, z_0)$ son tangentes exteriormente en z_0 .

Como consecuencia de este teorema tenemos el siguiente resultado.

Corolario 2.4.16. Bajo las condiciones del Teorema 2.4.14, en z_0 no pueden colisionar más de dos componentes conexas de $\Lambda'_{\varepsilon_1}(A)$.

Teorema 2.4.17. Sean $A \in \mathbb{C}^{n \times n}$ una matriz simple y z_0 un primer punto de colisión de los pseudoespectros de A . Entonces existen un par de vectores singulares ortogonales u, v para el valor singular $\sigma_0 := h_1(z_0)$ de la matriz $z_0 I_n - A$.

DEMOSTRACIÓN.

Si existe una sucesión $\{z_k\} \subset \Lambda'_{h_1(z_0)}(A)$ tal que $z_k \rightarrow z_0$ y $\nabla h_1(z_k) \rightarrow 0$ cuando $k \rightarrow \infty$, entonces por la compacidad de la bola unidad cerrada de $\mathbb{C}^{n \times 1}$ no hay pérdida de generalidad en suponer que $(u_k, v_k) \rightarrow (u, v)$. En este caso, resulta que

$$v^*u = \lim_{k \rightarrow \infty} v_k^*u_k = - \lim_{k \rightarrow \infty} \nabla h_1(z_k) = 0.$$

Supongamos, por el contrario, que se satisfacen las hipótesis del Teorema 2.4.14. Por la Observación 2.2.3, aplicando una traslación y un giro, no se pierde generalidad si suponemos que $z_0 = 0$ y que la tangente común a las clausuras de los discos $D(v^*Av, z_0)$ y $D(\hat{v}^*A\hat{v}, z_0)$ es la recta $x = 0$. Además, existen v, \hat{v} , vectores singulares de A para σ_0 , tales que

$$\begin{aligned} a &:= v^*Av < 0, \\ b &:= \hat{v}^*A\hat{v} > 0. \end{aligned} \quad (2.4.12)$$

Observemos que

$$Av = \sigma_0 u, A^*u = \sigma_0 v, A\hat{v} = \sigma_0 \hat{u}, A^*\hat{u} = \sigma_0 \hat{v}. \quad (2.4.13)$$

Si sustituimos los resultados de (2.4.13) en (2.4.12), tenemos

$$\begin{aligned} c &:= v^*u = \frac{a}{\sigma_0}, \\ d &:= \hat{v}^*\hat{u} = \frac{b}{\sigma_0}, \\ \text{rg}(u, \hat{u}) &= \text{rg}(v, \hat{v}) = 2 \end{aligned} \quad (2.4.14)$$

Para concluir la demostración tenemos que probar que existen $\alpha, \beta \in \mathbb{C}$ tales que $|\alpha|^2 + |\beta|^2 \neq 0$ y, además,

- (1) $\alpha u + \beta \hat{u}$, $\alpha v + \beta \hat{v}$ es un par de vectores singulares no necesariamente unitarios de σ_0 para A ,
- (2) $(\alpha u + \beta \hat{u})^*(\alpha v + \beta \hat{v}) = 0$.

Para probar (1), de (2.4.13) se sigue inmediatamente

$$A(\alpha v + \beta \hat{v}) = \sigma_0(\alpha u + \beta \hat{u}), \text{ y } A^*(\alpha u + \beta \hat{u}) = \sigma_0(\alpha v + \beta \hat{v}).$$

Veamos que ambos vectores tienen la misma norma, es decir, $\|\alpha u + \beta \hat{u}\| = \|\alpha v + \beta \hat{v}\|$. En primer lugar,

$$\begin{aligned} \|\alpha u + \beta \hat{u}\|^2 &= |\alpha|^2 + |\beta|^2 + 2 \text{Re}(\bar{\alpha}\beta u^*\hat{u}), \text{ y} \\ \|\alpha v + \beta \hat{v}\|^2 &= |\alpha|^2 + |\beta|^2 + 2 \text{Re}(\bar{\alpha}\beta v^*\hat{v}). \end{aligned}$$

Por otro lado, de (2.4.13) se tiene $u = Av/\sigma_0$ y $\hat{u} = A\hat{v}/\sigma_0$, por lo tanto,

$$u^*\hat{u} = \frac{v^*A^*A\hat{v}}{\sigma_0^2} = v^*\hat{v},$$

donde se ha utilizado que $A^*A\hat{v} = \sigma_0^2\hat{v}$.

Notemos que por ser $\text{rg}(u, \hat{u}) = \text{rg}(v, \hat{v}) = 2$, los vectores $\alpha u + \beta \hat{u}$ y $\alpha v + \beta \hat{v}$ son no nulos.

(2) Para probar este apartado tenemos que encontrar $\alpha, \beta \in \mathbb{C}$ tales que $|\alpha|^2 + |\beta|^2 \neq 0$ y $(\alpha u + \beta \hat{u})^*(\alpha v + \beta \hat{v}) = 0$, es decir, por (2.4.14)

$$|\alpha|^2 c + |\beta|^2 d + \bar{\alpha}\beta u^*\hat{v} + \alpha\bar{\beta}\hat{u}^*v = 0. \quad (2.4.15)$$

Tomemos $\alpha \in \mathbb{R}$ no nulo, y sean

$$\beta := \beta_1 + i\beta_2, \quad \hat{u}^* v := r_1 + ir_2, \quad u^* \hat{v} := s_1 + is_2,$$

donde $\beta_1, \beta_2, r_1, r_2, s_1, s_2 \in \mathbb{R}$.

De (2.4.15), separando las partes real e imaginaria, obtenemos

$$\begin{aligned} c\alpha^2 + d(\beta_1^2 + \beta_2^2) + \alpha\beta_1(s_1 + r_1) + \alpha\beta_2(r_2 - s_2) &= 0, \\ \beta_1(s_2 + r_2) + \beta_2(s_1 - r_1) &= 0, \end{aligned} \quad (2.4.16)$$

Si $r_1 = s_1$, tomando $\beta_1 = 0, \beta_2 = 1$, se satisface la segunda ecuación en (2.4.16). Para la primera se tiene que

$$p(\alpha) := c\alpha^2 + d + \alpha(r_2 - s_2) = 0.$$

Como $p(0) = d > 0$ y $\lim_{\alpha \rightarrow \infty} p(\alpha) = -\infty < 0$, existe un $\alpha_0 \in (0, \infty)$ tal que $p(\alpha_0) = 0$. Si, por el contrario, $r_1 \neq s_1$, tomando $\alpha = 1$ y despejando β_2 en la segunda ecuación de (2.4.16), se tiene que

$$q(\beta_1) := c + d(1 + w^2)\beta_1^2 + \beta_1(s_1 + r_1 + (r_2 - s_2)w) = 0,$$

donde $w = (s_2 + r_2)/(r_1 - s_1)$. Como $q(0) = c < 0$ y $\lim_{\beta_1 \rightarrow \infty} q(\beta_1) = \infty > 0$, existe un $\tilde{\beta}_1 \in (0, \infty)$ tal que $q(\tilde{\beta}_1) = 0$. □

Otra demostración de este teorema puede encontrarse en [4, Teorema 5].

Capítulo 3

Pseudoespectros de orden superior

3.1. Pseudoespectros de orden k

Con los conjuntos $\Lambda_k(A)$ y $\Lambda_k^g(A)$, definidos en los preliminares, vamos a generalizar el concepto de pseudoespectro ordinario.

Definición 3.1.1. Sea $A \in \mathbb{C}^{n \times n}$, para todo $\varepsilon \geq 0$ definimos el **pseudoespectro algebraico ordinario de orden k de nivel ε de la matriz A** como el conjunto

$$\Lambda_{\varepsilon,k}^a(A) := \bigcup_{\substack{\|X-A\| \leq \varepsilon \\ X \in \mathbb{C}^{n \times n}}} \Lambda_k(X).$$

Así mismo, el **pseudoespectro geométrico ordinario de orden k de nivel ε de la matriz A** será el conjunto

$$\Lambda_{\varepsilon,k}^g(A) := \bigcup_{\substack{\|X-A\| \leq \varepsilon \\ X \in \mathbb{C}^{n \times n}}} \Lambda_k^g(X).$$

De forma similar se pueden definir los pseudoespectros estrictos correspondientes, $\Lambda'_{\varepsilon,k}^a(A)$ y $\Lambda'_{\varepsilon,k}^g(A)$.

Notemos que $\Lambda_{\varepsilon,1}^a(A) = \Lambda_{\varepsilon,1}^g(A)$ y $\Lambda'_{\varepsilon,1}^a(A) = \Lambda'_{\varepsilon,1}^g(A)$. Además $\Lambda_{0,k}^a(A) = \Lambda_k(A)$, $\Lambda_{0,k}^g(A) = \Lambda_k^g(A)$ y $\Lambda_{\varepsilon,k}^g(A) \subset \Lambda_{\varepsilon,k}^a(A)$.

Vamos a relacionar los pseudoespectros geométricos de orden k con los valores singulares. Para ello, partiendo del Teorema 1.2.7, de forma similar a la Proposición 2.1.2, se demuestra el siguiente resultado.

Proposición 3.1.2. Dada $A \in \mathbb{C}^{n \times n}$ y $\varepsilon \geq 0$. Para todo k se tiene

$$\begin{aligned} \Lambda_{\varepsilon,k}^g(A) &= \{z \in \mathbb{C} : \sigma_{n-k+1}(zI_n - A) \leq \varepsilon\}, \\ \Lambda'_{\varepsilon,k}^g(A) &= \{z \in \mathbb{C} : \sigma_{n-k+1}(zI_n - A) < \varepsilon\}. \end{aligned} \tag{3.1.1}$$

Para el caso de pseudoespectros algebraicos de nivel k , la generalización de la Proposición 2.1.2 pasa por determinar la distancia de la matriz A al conjunto de matrices que tienen a z_0 como valor propio de multiplicidad algebraica $\geq k$. Es decir, calcular

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda_k(X)}} \|X - A\|.$$

Este problema sólo está resuelto, de forma completa, para el caso $k = 2$. Véase Malyshev [54] y Hinrichsen y Pritchard [40], Proposición 4.3.21, página 443. Previamente a su enunciado, vamos a introducir una definición.

Definición 3.1.3. Sea $A \in \mathbb{C}^{n \times n}$. Llamaremos **función de Malyshev** a la función $h_2 : \mathbb{C} \rightarrow \mathbb{R}$ dada por

$$h_2(z) := \max_{t \in \mathbb{R}} \sigma_{2n-1} \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix}, \quad z \in \mathbb{C}.$$

Como la función

$$f(t) := \sigma_{2n-1} \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix} \quad (3.1.2)$$

es par, i.e. $f(-t) = f(t)$ para todo t real, podemos restringir el dominio en el que se maximiza f para expresar la función h_2 :

$$h_2(z) := \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix} \quad (3.1.3)$$

Aunque consideramos a \mathbb{C} el dominio de definición de h_2 , no utilizamos el hecho de que \mathbb{C} es un cuerpo. Por eso, muchas veces cometeremos el abuso de lenguaje de identificar \mathbb{C} con \mathbb{R}^2 y de considerar a h_2 como una función de \mathbb{R}^2 en \mathbb{R} , escribiendo $h_2(x, y) := h_2(x + yi)$ para todo $(x, y) \in \mathbb{R}^2$. Hemos adoptado la notación $h_2(z), z \in \mathbb{C}$, pues por definición cada uno de los puntos de un pseudoespectro de la matriz $A \in \mathbb{C}^{n \times n}$ es un valor propio de una determinada matriz $X \in \mathbb{C}^{n \times n}$.

Más adelante, para abreviar, utilizaremos la siguiente notación

$$F_z(t) = \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix} \quad (3.1.4)$$

Con estas consideraciones tenemos el resultado siguiente.

Teorema 3.1.4 (Malyshev). Sean $z_0 \in \mathbb{C}$ y $A \in \mathbb{C}^{n \times n}$. Entonces,

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda_2(X)}} \|X - A\| = h_2(z_0).$$

Estamos ya en condiciones de generalizar la Proposición 2.1.2 para $\Lambda_{\varepsilon, 2}^a(A)$ y $\Lambda'_{\varepsilon, 2}(A)$.

Proposición 3.1.5. *Con las notaciones precedentes, para todo $\varepsilon \geq 0$ se tiene*

$$\begin{aligned}\Lambda_{\varepsilon,2}^a(A) &= \{z \in \mathbb{C} : h_2(z) \leq \varepsilon\}, \\ \Lambda'_{\varepsilon,2}(A) &= \{z \in \mathbb{C} : h_2(z) < \varepsilon\}.\end{aligned}\tag{3.1.5}$$

DEMOSTRACIÓN.

Demostremos la igualdad correspondiente a $\Lambda_{\varepsilon,2}^a(A)$. La otra se hace igual.

Primero supongamos que $z_0 \in \Lambda_{\varepsilon,2}^a(A)$, entonces existe $X_0 \in \mathbb{C}^{n \times n}$ tal que $z_0 \in \Lambda_2(X_0)$ y $\|X_0 - A\| \leq \varepsilon$. Esto junto con el Teorema 3.1.4 nos conduce a que

$$h_2(z_0) = \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} z_0 I_n - A & t I_n \\ O & z_0 I_n - A \end{pmatrix} = \min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda_2(X)}} \|X - A\| \leq \|X_0 - A\| \leq \varepsilon.$$

Por lo tanto,

$$\Lambda_{\varepsilon,2}^a(A) \subset \{z \in \mathbb{C} : h_2(z) \leq \varepsilon\}.$$

Recíprocamente, sea $z_0 \in \mathbb{C}$ tal que $h_2(z_0) \leq \varepsilon$. Aplicando el Teorema 3.1.4, resulta

$$\min_{\substack{X \in \mathbb{C}^{n \times n} \\ z_0 \in \Lambda_2(X)}} \|X - A\| \leq \varepsilon.$$

Esto implica que existe una matriz $X_0 \in \mathbb{C}^{n \times n}$ que cumple

$$z_0 \in \Lambda_2(X_0) \text{ y además } \|X_0 - A\| \leq \varepsilon.$$

Luego $z_0 \in \Lambda_{\varepsilon,2}^a(A)$. □

3.2. Continuidad del argumento del máximo

Para cada $(x, y, t) \in \mathbb{R} \times \mathbb{R} \times [0, \infty)$ definamos

$$s_{2n-1}(x, y, t) := \sigma_{2n-1} \begin{pmatrix} (x + yi)I_n - A & tI_n \\ 0 & (x + yi)I_n - A \end{pmatrix}.$$

Malyshev demostró que para cada $(x_0, y_0) \in \mathbb{R}^2$ la función $t \mapsto s_{2n-1}(t)$ tiende a 0 cuando $t \rightarrow \infty$; si no es una función idénticamente nula, alcanza su máximo absoluto en $[0, \infty)$ en un único punto t_0 ; además, no tiene otros extremos relativos. Por consiguiente, la correspondencia

$$(x_0, y_0) \mapsto t_0$$

define una función de \mathbb{R}^2 en $[0, \infty)$. Llamemos

$$\begin{aligned}\varphi : \quad \mathbb{R}^2 &\rightarrow [0, \infty) \\ (x_0, y_0) &\mapsto t_0\end{aligned}$$

a esta función. En la literatura también se la ha llamado argmax

$$\varphi(x_0, y_0) := t_0 = \operatorname{argmax}_{t \geq 0} s_{2n-1}(x_0, y_0, t).$$

Vamos a probar que dicha función es continua en \mathbb{R}^2 .

Lema 3.2.1. La función $\varphi : \mathbb{R}^2 \rightarrow [0, \infty)$ es continua en cada punto de \mathbb{R}^2 .

DEMOSTRACIÓN.

Sea $(x_0, y_0) \in \mathbb{R}^2$. Pongamos nuestra atención en la curva representativa de la función

$$t \mapsto s_{2n-1}(x_0, y_0, t).$$

Sea M_0 el valor máximo de esta función; así pues, existe un $t_0 \geq 0$ tal que $M_0 = s_{2n-1}(x_0, y_0, t_0)$. Supondremos en primer lugar que $t_0 > 0$. Tomemos un $\delta \in (0, M_0)$ y consideremos las curvas de las funciones

$$t \mapsto s_{2n-1}(x_0, y_0, t) + \delta$$

y

$$t \mapsto s_{2n-1}(x_0, y_0, t) - \delta.$$

Consideremos los puntos A y B que son las intersecciones de la curva

$$s = s_{2n-1}(x_0, y_0, t) + \delta$$

con la recta horizontal

$$s = M_0 - \delta$$

en el plano t, s . Véase la Figura 3.1.

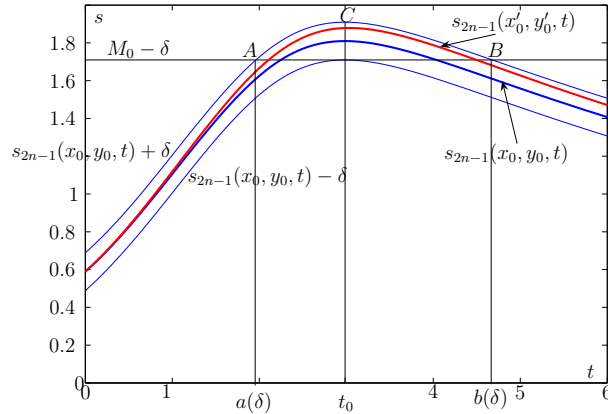


Figura 3.1: Paso angosto para la curva roja.

Sean $a(\delta)$ y $b(\delta)$ las abscisas de los puntos A y B , respectivamente. Es claro que

$$t_0 \in [a(\delta), b(\delta)]. \quad (3.2.6)$$

Por ser las curvas

$$s = s_{2n-1}(x_0, y_0, t) \pm \delta$$

traslaciones de la curva

$$s = s_{2n-1}(x_0, y_0, t)$$

en la dirección paralela al eje vertical, estas curvas conservan la pendiente y la curvatura para cada $t \in [0, \infty)$. Es fácil ver que $a(\delta)$ es una función estrictamente decreciente de δ en $(0, M_0)$, i.e., si $0 < \delta' < \delta$ entonces $a(\delta) < a(\delta')$. Asimismo puede verse que $b(\delta)$ es una función estrictamente creciente en $(0, M_0)$. Por tanto, existen los límites laterales

$$\lim_{\delta \rightarrow 0^+} a(\delta) \quad \text{y} \quad \lim_{\delta \rightarrow 0^+} b(\delta)$$

y, por (3.2.6), estos límites valen t_0 .

Dado ahora un $\varepsilon > 0$, buscamos un $\delta_0 > 0$ tal que para todo punto $(x'_0, y'_0) \in \mathbb{R}^2$ que satisfaga $\|(x'_0, y'_0) - (x_0, y_0)\| < \delta_0$ se tenga $|\varphi(x'_0, y'_0) - \varphi(x_0, y_0)| < \varepsilon$; lo que equivaldría a que φ fuera continua en el punto (x_0, y_0) .

Como $\lim_{\delta \rightarrow 0^+} a(\delta) = t_0$, existe un $\delta_0^a > 0$ tal que para todo $\delta : 0 < \delta \leq \delta_0^a$ se tiene que

$$|t_0 - a(\delta)| < \frac{\varepsilon}{2}; \quad (3.2.7)$$

de igual manera, como $\lim_{\delta \rightarrow 0^+} b(\delta) = t_0$, existe un $\delta_0^b > 0$ tal que para todo $\delta : 0 < \delta \leq \delta_0^b$ se tiene que

$$|t_0 - b(\delta)| < \frac{\varepsilon}{2}. \quad (3.2.8)$$

Tomemos $\delta_0 := \min(\delta_0^a, \delta_0^b)$. Si ahora (x'_0, y'_0) es un punto que satisface la desigualdad

$$\|(x'_0, y'_0) - (x_0, y_0)\| < \delta_0, \quad (3.2.9)$$

entonces por el Teorema 1.2.3 de continuidad de los valores singulares, se sigue que

$$\forall t \geq 0, |s_{2n-1}(x'_0, y'_0, t) - s_{2n-1}(x_0, y_0, t)| \leq \|(x'_0, y'_0) - (x_0, y_0)\| < \delta_0. \quad (3.2.10)$$

Sea $t'_0 \geq 0$ tal que

$$s_{2n-1}(x'_0, y'_0, t'_0) = \max_{t \geq 0} s_{2n-1}(x'_0, y'_0, t).$$

En este momento surgen dos casos posibles

- (1) O bien $s_{2n-1}(x'_0, y'_0, t'_0) \geq s_{2n-1}(x_0, y_0, t'_0) - \delta_0$.
- (2) O bien $s_{2n-1}(x'_0, y'_0, t'_0) < s_{2n-1}(x_0, y_0, t'_0) - \delta_0$.

Pero (2) no puede ocurrir por la desigualdad (3.2.10). En consecuencia, $t'_0 \in [a(\delta_0), b(\delta_0)]$. Por la elección de δ_0 se tiene por (3.2.7) y (3.2.8) que

$$|b(\delta_0) - a(\delta_0)| \leq |b(\delta_0) - t_0| + |t_0 - a(\delta_0)| < \varepsilon.$$

Dado que $t'_0, t_0 \in [a(\delta_0), b(\delta_0)]$,

$$|t'_0 - t_0| \leq |b(\delta_0) - a(\delta_0)| < \varepsilon.$$

Es decir, hemos probado que

$$|\varphi(x'_0, y'_0) - \varphi(x_0, y_0)| < \varepsilon.$$

En el caso en que $\varphi(x_0, y_0) = 0$, es decir, en el caso en que $s_{2n-1}(x_0, y_0, t)$ alcanza su máximo valor en $t_0 = 0$, la demostración se realiza de manera análoga, pero en vez del intervalo $[a(\delta), b(\delta)]$ debemos considerar el intervalo $[0, b(\delta)]$. \square

Observación 3.2.2. Reparemos en la región “triangular” cerrada Δ , de vértices A , B y C en la Figura 3.1. Está limitada por el segmento rectilíneo \overline{AB} y el arco de curva $s = s_{2n-1}(x_0, y_0, t) + \delta$ situado encima de él. Consideremos las coordenadas del punto

$$C' := (t'_0, s_{2n-1}(x'_0, y'_0, t'_0))$$

que es el punto más alto de la curva roja; es decir, el máximo de

$$t \mapsto s_{2n-1}(x'_0, y'_0, t).$$

Por la propiedad unimodal de esta función, el punto C' debe estar en la región Δ . Esta es una manera geométrica de probar que en la demostración anterior, la posibilidad (2) no puede darse.

3.3. Mínimo de la función de Malyshev

En esta sección queremos probar que si z_1 es un *primer punto de colisión* de dos componentes conexas del pseudoespectro ordinario de una matriz $A \in \mathbb{C}^{n \times n}$ con todos sus valores propios simples, entonces la función de Malyshev, h_2 , alcanza su mínimo absoluto en z_1 . Otras propiedades de la función h_2 se dan en las proposiciones que siguen. Véase la Definición 2.4.6 de primer punto de colisión de pseudoespectros.

Proposición 3.3.1. *Sea z_0 un punto de colisión de dos componentes conexas de $\Lambda'_{\sigma_0, 1}(A)$. Entonces*

$$h_2(z_0) = h_1(z_0).$$

DEMOSTRACIÓN. Sabemos que $\sigma_0 = h_1(z_0) := \sigma_n(z_0 I_n - A)$. Si σ_0 es un valor singular múltiple de $z_0 I_n - A$, entonces $\sigma_{n-1}(z_0 I_n - A) = \sigma_0$ es la distancia de A al conjunto de matrices que tienen a z_0 como valor propio de multiplicidad geométrica mayor o igual a 2. Por el Teorema 3.1.4 resulta que

$$\begin{aligned} \sigma_0 &= \sigma_{n-1}(z_0 I_n - A) = \min_{\text{mg}(z_0, X) \geq 2} \|X - A\| \geq \min_{m(z_0, X) \geq 2} \|X - A\| \\ &= \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} z_0 I_n - A & t I_n \\ 0 & z_0 I_n - A \end{pmatrix} \geq \sigma_{2n-1} \begin{pmatrix} z_0 I_n - A & 0 \\ 0 & z_0 I_n - A \end{pmatrix} \\ &= \sigma_n(z_0 I_n - A) = \sigma_0. \end{aligned}$$

Supongamos, por el contrario, que σ_0 es un valor singular simple de $z_0 I_n - A$. Entonces por el Teorema 2.4.13, existe una matriz X_0 que tiene a z_0 como valor propio múltiple y además $\|X_0 - A\| = \sigma_0$. Por lo tanto, por el Teorema 3.1.4 resulta que

$$\begin{aligned} \sigma_0 &\geq \min_{m(z_0, X) \geq 2} \|X - A\| \\ &= \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} z_0 I_n - A & t I_n \\ 0 & z_0 I_n - A \end{pmatrix} \geq \sigma_{2n-1} \begin{pmatrix} z_0 I_n - A & 0 \\ 0 & z_0 I_n - A \end{pmatrix} \\ &= \sigma_n(z_0 I_n - A) = \sigma_0. \end{aligned}$$

□

Observación 3.3.2. Obsérvese que para todo $z \in \mathbb{C}$, se tiene que

$$h_2(z) \geq h_1(z),$$

pues

$$\begin{aligned} h_2(z) &= \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} z I_n - A & t I_n \\ 0 & z I_n - A \end{pmatrix} \geq \sigma_{2n-1} \begin{pmatrix} z I_n - A & 0 \\ 0 & z I_n - A \end{pmatrix} \\ &= \sigma_n(z I_n - A) = h_1(z). \end{aligned}$$

Proposición 3.3.3. *Sea z_0 un punto de colisión de dos componentes conexas de $\Lambda'_{\sigma_0, 1}(A)$. Entonces, para todo $z \in \mathbb{C}$ que no pertenece a $\Lambda'_{\sigma_0, 1}(A)$ se cumple que*

$$h_2(z) \geq h_2(z_0).$$

DEMOSTRACIÓN. Como $h_2(z) \geq h_1(z)$ y utilizando la Proposición 3.3.1, resulta $h_2(z) \geq h_1(z) \geq h_1(z_0) = h_2(z_0)$. □

Teorema 3.3.4. *Sea $A \in \mathbb{C}^{n \times n}$ una matriz cuyos valores propios son simples. Sea z_1 un primer punto de colisión de dos componentes conexas de $\Lambda'_{\sigma_1, 1}(A)$. Entonces, la función h_2 alcanza el mínimo absoluto en z_1 .*

DEMOSTRACIÓN. Recordemos que $\sigma_1 = h_1(z_1)$. Como z_1 es un punto de colisión de dos componentes conexas de $\Lambda'_{\sigma_1, 1}(A)$, por la Proposición 3.3.1 se tiene que

$$h_2(z_1) = \sigma_1.$$

Supongamos que existiese un punto $w \in \mathbb{C}$ tal que

$$h_2(w) < h_2(z_1).$$

Por la Proposición 3.3.3, w tendría que pertenecer a $\Lambda'_{\sigma_1, 1}(A)$. Por tanto, se cumpliría que $h_1(w) < \sigma_1$. Sea Ω la componente conexas de $\Lambda'_{\sigma_1, 1}(A)$ que contendría a w . Por ser z_1 un primer punto de colisión, Ω sólo contendría un valor propio, α , de A . El Teorema 3.1.4 nos aseguraría que existiría una matriz X_w

que tendría a w como valor propio múltiple y tal que $\|X_w - A\| = h_2(w)$. Por la Proposición 2.4.10 resultaría que

$$2 \leq \sum_{\xi \in \Lambda(X_w) \cap \Omega} m(\xi, X_w) = \sum_{\alpha \in \Lambda(A) \cap \Omega} m(\alpha, A) = 1;$$

lo cual es absurdo. Por lo tanto, para todo $z \in \mathbb{C}$, se tiene que $h_2(z) \geq h_2(z_1)$. Lo que implica que

$$\min_{z \in \mathbb{C}} h_2(z) = h_2(z_1).$$

□

Observación 3.3.5. En el caso en que $A \in \mathbb{C}^{n \times n}$ sea una matriz con todos sus valores propios simples, el Teorema 3.3.4 sustancia la *Conjetura del entrelazamiento de pseudoespectros* para los pseudoespectros algebraicos de órdenes uno y dos en un *primer* punto de colisión z_1 : Sea $\sigma_1 := h_1(z_1)$, entonces $\Lambda_{\varepsilon,2}^a(A) = \emptyset$ si $0 \leq \varepsilon < \sigma_1$, $\Lambda_{\sigma_1,2}^a(A) = \{z_1\}$, y cuando $\varepsilon \geq \sigma_1$ crece desde σ_1 , el pseudoespectro $\Lambda_{\varepsilon,2}^a(A)$ es no vacío y crece alrededor de z_1 . Véase la Figura 3.2 referida a la matriz $A \in \mathbb{C}^{3 \times 3}$ que viene a continuación y donde a σ_1 se lo llamó ε_1 .

$$A = \begin{pmatrix} -1 & 5 & 6 \\ 0 & -2 & 0 \\ 0 & 0 & -1,8 + 0,2i \end{pmatrix}, \quad \varepsilon_1 < \varepsilon_2 < \varepsilon_3 < \varepsilon_4.$$

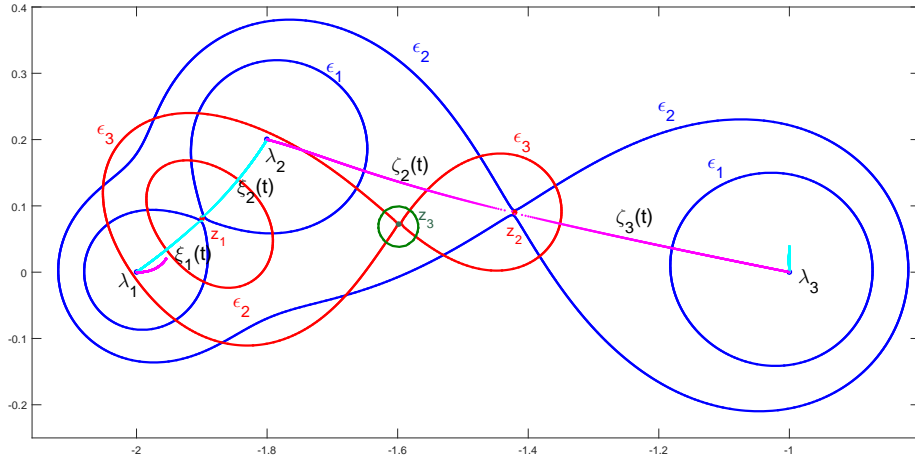


Figura 3.2: Entrelazamiento de pseudoespectros $\Lambda_{\varepsilon_1,1}^a(A)$ y $\Lambda_{\varepsilon_2,2}^a(A)$.

Diremos que z_0 es un **segundo punto de colisión de los pseudoespectros de una matriz** $A \in \mathbb{C}^{n \times n}$ si está en la frontera de $\Lambda'_{\varepsilon_2}(A)$, siendo

ε_2 el segundo punto de discontinuidad de ρ en el semieje positivo de ε . Sea $P \in \mathbb{C}^{n \times n}$ una matriz invertible; llamaremos **número de condición** de P al valor $\text{cond}(P) := \|P\| \|P^{-1}\|$. Sean λ un valor propio de A y $\varepsilon > 0$, entonces $\mathcal{K}'_\lambda(\varepsilon)$ denotará la componente conexa del pseudoespectro estricto $\Lambda'_\varepsilon(A)$ que contiene a λ .

Utilizando el Teorema de Bauer-Fike (véase el Teorema 2, página 388, de Lancaster-Tismenetsky [48]), las Proposiciones 3.3.1, 3.3.3 y el Teorema 3.3.4 se prueba el resultado que sigue.

Teorema 3.3.6 (Una condición suficiente de entrelazamiento). *Sea $A \in \mathbb{C}^{3 \times 3}$ con valores propios distintos $\lambda_1, \lambda_2, \lambda_3$. Sea $P \in \mathbb{C}^{3 \times 3}$ una matriz invertible que transforma A en la matriz $D = P^{-1}AP$, donde $D := \text{diag}(\lambda_1, \lambda_2, \lambda_3)$. Sean z_1 y z_2 el primer y el segundo puntos de colisión de pseudoespectros de A . Llamemos $\varepsilon_j := h_1(z_j), j = 1, 2$.*

Supongamos que z_1 es punto frontera de $\mathcal{K}'_{\lambda_1}(\varepsilon_1)$ y $\mathcal{K}'_{\lambda_2}(\varepsilon_1)$ y punto exterior a $\mathcal{K}'_{\lambda_3}(\varepsilon_1)$. Sea $A_1 \in \mathbb{C}^{3 \times 3}$ una matriz a la distancia $\sigma_3(z_1 I_3 - A)$ de A que tiene a z_1 como valor propio. Supongamos que los valores propios $\xi_1(t), \xi_2(t)$ de $X(t) := (1-t)A + tA_1$ describen caminos desde λ_1, λ_2 a z_1 al variar t de 0 a 1.

Supongamos además que z_2 es un punto frontera de $\mathcal{K}'_{\lambda_1}(\varepsilon_2)$ y de $\mathcal{K}'_{\lambda_3}(\varepsilon_2)$. Sea $A_2 \in \mathbb{C}^{3 \times 3}$ una matriz a la distancia $\sigma_3(z_2 I_3 - A)$ de A que tiene a z_2 como valor propio. Supongamos que los valores propios $\zeta_2(t), \zeta_3(t)$ de $Z(t) := (1-t)A + tA_2$ describen caminos desde λ_2, λ_3 a z_2 al variar t de 0 a 1.

Si

$$\text{cond}(P)\varepsilon_2 < |z_2 - \lambda_1|,$$

entonces la función h_2 tiene un mínimo local en z_2 .

Capítulo 4

Derivada del área y del diámetro

Dado $\varepsilon \geq 0$ y $A \in \mathbb{C}^{n \times n}$, denotemos por $\mathcal{K}_\lambda(\varepsilon)$ a la componente conexa del pseudoespectro ordinario $\Lambda_\varepsilon(A)$, que contiene al valor propio λ de A de índice ν . Denotemos también por $\delta(\varepsilon)$, $r(\varepsilon)$ y $a(\varepsilon)$ el diámetro, el radio medio y el área de dicha componente conexa. El objeto de este capítulo es establecer la relación que existe entre el número de condición de λ de orden $1/\nu$ y las derivadas por la derecha primera y segunda de δ , r y a en $\varepsilon = 0$. Todo ello es consecuencia de que la “derivada” de $\mathcal{K}_\lambda(\varepsilon)$ respecto de ε en $\varepsilon = 0$, en el espacio métrico formado por los compactos no vacíos de \mathbb{C} con la métrica de Hausdorff¹, es un disco cerrado de \mathbb{C} centrado en 0 y de radio r el número de condición:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\mathcal{K}_\lambda(\varepsilon) - \mathcal{K}_\lambda(0)}{\varepsilon} = \mathcal{D}(0, r), \text{ para el caso en que } \nu = 1,$$

según probó Karow en su tesis [44, Teorema 2.6.6 (ii), pág. 36]. Téngase en cuenta que $\mathcal{K}_\lambda(0) = \{\lambda\}$, por lo que el conjunto $\mathcal{K}_\lambda(\varepsilon) - \mathcal{K}_\lambda(0) = \mathcal{K}_\lambda(\varepsilon) - \lambda$ es el resultado de someter $\mathcal{K}_\lambda(\varepsilon)$ a la traslación $z \mapsto z - \lambda$. Después actúa la homotecia $z \mapsto \frac{z}{\varepsilon}$ de razón $\frac{1}{\varepsilon}$. Véase también [7].

4.1. Elementos de áreas y de diámetros

Por la Proposición 2.2.4 (a), para todo $\varepsilon \geq 0$, los pseudoespectros ordinarios $\Lambda_\varepsilon(A)$ pertenecen a $\mathcal{H}(\mathbb{R}^2) \equiv \mathcal{H}(\mathbb{C})$. Sea $K \in \mathcal{H}(\mathbb{C})$, sabemos que su **diámetro** se define por

$$\delta(K) := \sup_{z, w \in K} |z - w|$$

¹Este concepto encajaría en el incipiente *Análisis Mutacional* que trata de ecuaciones diferenciales cuyas incógnitas son funciones con valores en un espacio métrico que no es espacio vectorial; véanse los libros de Aubin [11] y de Lorenz [52]

y su **área**, $a(K)$ por la *medida Lebesgue* $m(K)$, pues K es un *conjunto de Borel*² por ser cerrado. Como todo conjunto de Borel es *medible Lebesgue* se sigue que la medida o área de K está bien definida. Recordemos que un subconjunto A de un espacio topológico E se dice *denso en E* si la clausura de A es igual a E . Véase B.1.3.

Proposición 4.1.1. *Sea $(\mathcal{H}(\mathbb{C}), d_H)$ el espacio métrico formado por todos los subconjuntos compactos no vacíos de \mathbb{C} , donde d_H es la distancia de Hausdorff. Entonces,*

- *El conjunto, $\mathcal{F} \subset \mathcal{H}(\mathbb{C})$, formado por los subconjuntos finitos no vacíos de \mathbb{C} es denso en $\mathcal{H}(\mathbb{C})$.*
- *Las funciones diámetro y área $\delta, a: \mathcal{H}(\mathbb{C}) \rightarrow \mathbb{R}$ son*
 - *$K \mapsto \delta(K)$ continua en todo K .*
 - *$K \mapsto a(K)$ continua en K_0 si y sólo si $a(K_0) = 0$.*
 - *monótonas crecientes respecto de las relaciones de contenido \subset en \mathbb{C} y de orden habitual \leq en \mathbb{R} : $K_1 \subset K_2$ implica que $\delta(K_1) \leq \delta(K_2)$ y $a(K_1) \leq a(K_2)$.*

4.1.1. Variación del diámetro y del área de una región plana al someterla a una homotecia.

Para hablar del *área* de una componente conexa como el pseudoespectro estricto, $\Lambda'_\varepsilon(A)$, es un conjunto abierto acotado, todo conjunto abierto acotado Ω de \mathbb{R}^2 es igual a la unión de una sucesión de cuadrados de lados paralelos a los ejes de coordenadas $\{Q_i\}_{i=1}^\infty$ no rampantes, i.e. $\overset{\circ}{Q}_i \cap \overset{\circ}{Q}_j = \emptyset$ si $i \neq j$, con clausura contenida en Ω . Véase el Teorema C.3.2. Así mismo, al ser $\Lambda'_\varepsilon(A)$ abierto, será un conjunto de Borel y por tanto *medible Lebesgue*; y como la medida de Lebesgue es σ -aditiva, se tiene que

$$a(\Lambda'_{\varepsilon,1}(A)) = \sum_{i=1}^{\infty} a(Q_i).$$

Lo mismo puede decirse del conjunto $\mathcal{K}'_\lambda(\varepsilon)$, componente conexa de $\Lambda'_\varepsilon(A)$ que contiene al valor propio λ . Es decir, que es medible Lebesgue y la unión de una sucesión de cuadrados abiertos no rampantes. De otro lado, como por la Proposición 2.2.4(c), la función $h_1(x, y) = \sigma_n((x + yi)I_n - A) : \mathbb{R}^2 \rightarrow \mathbb{R}$ no es constante en ninguna bola abierta de radio positivo, se tiene que la frontera de $\Lambda'_\varepsilon(A)$ es un conjunto de medida nula; por tanto,

$$a(\Lambda_\varepsilon(A)) = a(\Lambda'_\varepsilon(A)).$$

²Se dice que $B \subset \mathbb{R}^2$ es un **conjunto de Borel** si B puede obtenerse por un número de operaciones numerable, partiendo de los conjuntos abiertos, consistiendo cada operación en tomar uniones, intersecciones o complementarios.

Por todo ello, para establecer como afecta una homotecia con centro el origen y razón r a $\mathcal{K}_\lambda(\varepsilon)$, basta con estudiar el comportamiento del área de un cuadrado de lados paralelos a los ejes al someterlo a dicha homotecia. La cuestión análoga sobre los diámetros es mucho más sencilla. Pues bien, si $Q \subset \mathbb{R}^2$ es un cuadrado de lados paralelos a los ejes, y $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es la homotecia o dilatación de centro $(0, 0)$ y razón $r > 0$, entonces

$$a(T(Q)) = r^2 a(Q).$$

Véanse las Figuras 4.1 y 4.2. En consecuencia,

$$a(T(\mathcal{K}_\lambda(\varepsilon))) = r^2 a(\mathcal{K}_\lambda(\varepsilon)).$$

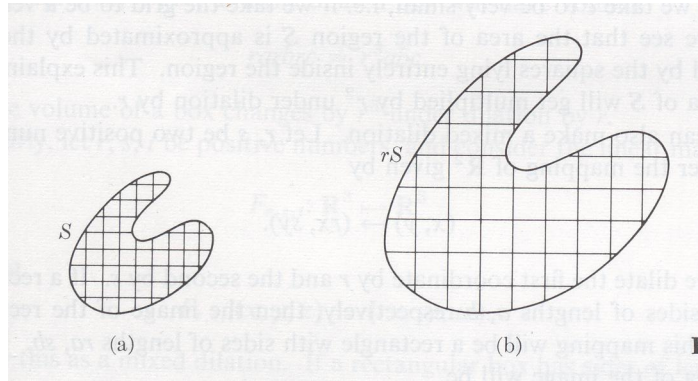


Figura 4.1: Áreas homotéticas

Dado $\varepsilon > 0$, denotemos por $\partial\mathcal{K}_\lambda(\varepsilon)$ la frontera de $\mathcal{K}_\lambda(\varepsilon)$. De Karow [44, Proposición 2.6.5] deducimos

$$\text{cond}_\omega(A, \lambda) = \lim_{\varepsilon \rightarrow 0^+} \frac{\max_{z \in \partial\mathcal{K}_\lambda(\varepsilon)} |z - \lambda|}{\varepsilon^\omega}. \quad (4.1.1)$$

Observemos que

$$\rho(\varepsilon) := \max_{z \in \partial\mathcal{K}_\lambda(\varepsilon)} |z - \lambda|$$

es el máximo radio de $\mathcal{K}_\lambda(\varepsilon)$ respecto de λ . Así pues,

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\rho(\varepsilon)}{\varepsilon^\omega} = \text{cond}_\omega(A, \lambda).$$

Para cada $\varepsilon \geq 0$ el diámetro de $\mathcal{K}_\lambda(\varepsilon)$ está definido por

$$\delta(\varepsilon) := \sup_{z_1, z_2 \in \mathcal{K}_\lambda(\varepsilon)} |z_1 - z_2|.$$

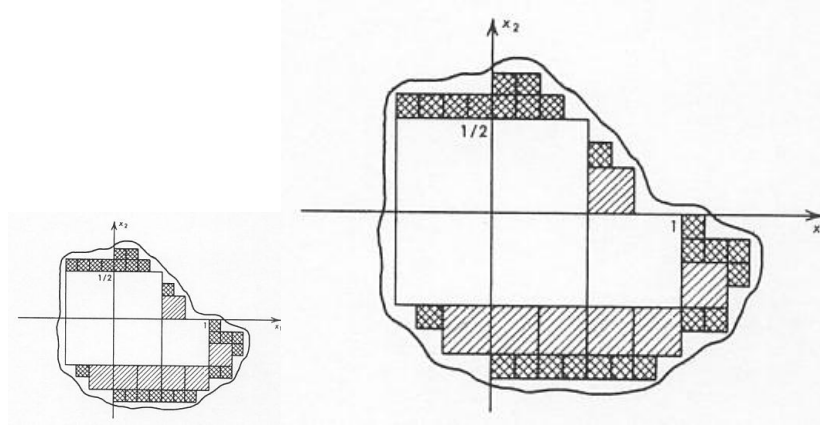


Figura 4.2: Descomposición canónica de un abierto en cuadrados, bajo una homotecia

4.2. Ejemplos

Antes de entrar en una discusión teórica vamos a considerar cuatro ejemplos.

Ejemplo primero. Sea $A \in \mathbb{C}^{n \times n}$ una matriz normal. Entonces $\Lambda_\varepsilon(A)$ es la unión de discos cerrados de radio ε centrados en los valores propios de A . Así pues, para $\varepsilon \geq 0$ suficientemente pequeño tenemos que

$$\mathcal{K}_\lambda(\varepsilon) = \mathcal{D}(\lambda, \varepsilon).$$

Por lo tanto, $\rho(\varepsilon) = \varepsilon$, y ya que los valores propios de una matriz normal son semisimples (i.e. de índice 1), resulta que

$$\text{cond}_1(A, \lambda) = \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\varepsilon} = 1.$$

El diámetro de $\mathcal{D}(\lambda, \varepsilon)$ es 2ε . Por consiguiente $\delta'(\varepsilon) = 2$ y $\delta'_+(0) = 2$. Si denotamos por $a(\varepsilon)$ el área de este disco, $a(\varepsilon) = \pi\varepsilon^2$; de aquí que $a'(\varepsilon) = 2\pi\varepsilon$, $a''(\varepsilon) = 2\pi$. En consecuencia, $a'_+(0) = 0$, $a''_+(0) = 2\pi$.

Ejemplo segundo. Sea la matriz

$$J_2(\lambda, d) := \begin{bmatrix} \lambda & d \\ 0 & \lambda \end{bmatrix},$$

análoga a un bloque de Jordan, donde λ, d son números complejos y $d \neq 0$. En [44, Teorema 5.4.1, p. 74] Karow demostró que para cada $\varepsilon \geq 0$, el pseudospectro $\Lambda_\varepsilon(J_2(\lambda, d))$ es un disco cerrado centrado en λ con radio $r_2(\varepsilon) =$

máx\{r > 0 | \sigma_2(J_2(r, |d|)) \le \varepsilon\}. Como

$$\sigma_2(J_2(r, |d|)) = \sqrt{r^2 + \frac{|d|^2}{2}} - \sqrt{r^2|d|^2 + \frac{|d|^4}{4}},$$

resolviendo la ecuación que sigue en la incógnita r

$$r^2 + \frac{|d|^2}{2} - \sqrt{r^2|d|^2 + \frac{|d|^4}{4}} = \varepsilon^2$$

hallamos que r_2(\varepsilon) = \sqrt{\varepsilon^2 + |d|\varepsilon}. Por tanto, \delta(\varepsilon) = 2\sqrt{\varepsilon^2 + |d|\varepsilon}. Consecuentemente,

$$\delta'(\varepsilon) = \frac{2\varepsilon + |d|}{\sqrt{\varepsilon^2 + |d|\varepsilon}}, \quad \Rightarrow \quad \delta'_+(0) = \infty.$$

Es obvio que \nu(\lambda) = 2. Observemos que

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon^{1/2}} = 2\sqrt{|d|}. \quad (4.2.2)$$

Ahora, vamos a ver que c := \text{cond}_{1/2}(J_2(\lambda, d), \lambda) = \sqrt{|d|}. Necesitamos hacer algunas consideraciones previas. Para una matriz cualquiera A \in \mathbb{C}^{n \times n} sea \Lambda(A) = \{\lambda_1, \dots, \lambda_r\} su espectro y

$$A = \sum_{j=1}^r (\lambda_j P_j + N_j)$$

la descomposición de Jordan de A, ver 1.3.1. Por Karow [44, Teorema 5.4.4 (viii), p. 78], si \nu_j := \nu(\lambda_j), entonces

$$\text{cond}_{1/\nu_j}(A, \lambda_j) = \|N_j^{\nu_j-1}\|^{1/\nu_j}. \quad (4.2.3)$$

En nuestro ejemplo presente A = J_2(\lambda, d) = \lambda I_2 + N, siendo

$$N = \begin{bmatrix} 0 & d \\ 0 & 0 \end{bmatrix},$$

así pues, c = \|N\|^{1/2} = \sqrt{|d|}. Por (4.2.2), deducimos que

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon^{1/2}} = 2c.$$

El área a(\varepsilon) está dada por \pi(\varepsilon^2 + |d|\varepsilon). Por lo tanto, a'(\varepsilon) = \pi(2\varepsilon + |d|); y de aquí, a'_+(0) = \pi|d| = \pi c^2.

Ejemplo tercero. Este ejemplo es una pequeña generalización del segundo. Sea A cualquier matriz compleja n \times n que tenga un único valor propio: \lambda. Además, supongamos que \nu(\lambda) = 2. Si N := A - \lambda I_n, entonces \Lambda(N) = \{0\}.

Así pues, N es una matriz nilpotente. Por consiguiente, $A = \lambda I_n + N$ es la descomposición de Jordan de A . De aquí que, por (4.2.3), $c := \text{cond}_{1/2}(A, \lambda) = \|N\|^{1/2} = \sqrt{\sigma_1(\lambda I_n - A)}$. En [8, Corolario 20] fue probado que para cada $\varepsilon \geq 0$,

$$\Lambda_\varepsilon(A) = \mathcal{D}(\lambda, \sqrt{\varepsilon^2 + \varepsilon\sigma_1(\lambda I_n - A)}).$$

Por lo que $\delta(\varepsilon) = 2\sqrt{\varepsilon^2 + \varepsilon\sigma_1(\lambda I_n - A)}$, $a(\varepsilon) = \pi(\varepsilon^2 + \varepsilon\sigma_1(\lambda I_n - A))$. En consecuencia,

$$\delta'(\varepsilon) = \frac{2\varepsilon + \sigma_1(\lambda I_n - A)}{\sqrt{\varepsilon^2 + \varepsilon\sigma_1(\lambda I_n - A)}} \Rightarrow \delta'_+(0) = \infty,$$

y

$$a'(\varepsilon) = \pi(2\varepsilon + \sigma_1(\lambda I_n - A)) \Rightarrow a'_+(0) = \pi\sigma_1(\lambda I_n - A) = \pi c^2.$$

Ejemplo cuarto. Sean λ_1, λ_2 dos números complejos distintos. Sea $A \in \mathbb{C}^{2 \times 2}$ una matriz cuyos valores propios son λ_1 y λ_2 . Definamos

$$d(A) := \sqrt{\|A\|_F^2 - |\lambda_1|^2 - |\lambda_2|^2}.$$

El número $d(A)$ es la desviación de la normalidad de A respecto de la norma de Frobenius. En [44, Proposición 5.5.3, p. 80] Karow probó que para cada $\varepsilon > 0$,

$$\Lambda_\varepsilon(A) = \mathcal{D}(\lambda_1, \varepsilon) \cup \mathcal{D}(\lambda_2, \varepsilon) \cup \mathcal{M}_{\lambda_1, \lambda_2}(d(A), \varepsilon) \quad (4.2.4)$$

siendo

$$\mathcal{M}_{\lambda_1, \lambda_2}(d, \varepsilon) := \{z \in \mathbb{C} : (|z - \lambda_1|^2 - \varepsilon^2)(|z - \lambda_2|^2 - \varepsilon^2) \leq \varepsilon^2 d^2\}, \quad d \geq 0. \quad (4.2.5)$$

Consideremos la Figura 4.3 que muestra el pseudoespectro ordinario de nivel ε , para los valores de $\varepsilon = 1,00; 1,50; 1,75$, de la matriz

$$A_1 := \begin{bmatrix} 1 + 2i & 3 \\ 0 & -1 - 3i \end{bmatrix}.$$

Así pues, $\lambda_1 = -1 - 3i$, $\lambda_2 = 1 + 2i$, $\|A_1\|_F = \sqrt{|1 + 2i|^2 + 3^2 + |-1 - 3i|^2} = \sqrt{5 + 9 + 10} = \sqrt{24}$. Por lo tanto, $d(A_1) = \sqrt{24 - 5 - 10} = \sqrt{9} = 3$.

Vemos que para valores de ε suficientemente pequeños $\Lambda_\varepsilon(A_1)$ tiene dos componentes conexas. Fijemos nuestra atención en $\varepsilon = 1,50$. La Figura 4.3 muestra una línea recta que pasa por λ_1 y λ_2 . Esta recta corta a la frontera $\partial\Lambda_\varepsilon(A_1)$ en los puntos z_1, z_2, z_3, z_4 . Si consideramos el valor propio λ_2 , el diámetro $\delta(\varepsilon)$ de la componente conexa $\mathcal{K}_{\lambda_2}(\varepsilon)$ es igual a la distancia entre z_3 y z_4 . De (4.2.4) y (4.2.5), podemos deducir que

$$\delta(\varepsilon) = \frac{1}{\sqrt{116}} \left(\sqrt{116\varepsilon^2 + 116\sqrt{38}\varepsilon + 841} - \sqrt{116\varepsilon^2 - 116\sqrt{38}\varepsilon + 841} \right).$$

Por consiguiente, $\delta'(0) = 2\sqrt{\frac{38}{29}}$. Como demostraremos más adelante este valor es igual a dos veces el número de condición de orden 1 de λ_2 . En vez de hacer un razonamiento particular para la matriz A_1 , vamos a encontrar una expresión general para el diámetro $\delta(\varepsilon)$ de $\mathcal{K}_{\lambda_2}(\varepsilon)$ para *cualquier* matriz $A \in \mathbb{C}^{2 \times 2}$ que tenga por valores propios a λ_1 y λ_2 . Por (4.2.5), la frontera $\partial\Lambda_\varepsilon(A)$ son cuasi-óvalos de Cassini con focos en los puntos λ_1 y λ_2 . De hecho, el conjunto $\Lambda_\varepsilon(A)$ es simétrico respecto de la recta que pasa por los puntos λ_1 y λ_2 . Esto es una consecuencia de (4.2.4) y (4.2.5). Además, $\delta(\varepsilon) = |z_4 - z_3| = |z_2 - z_1|$. Más tarde necesitaremos los números de condición de λ_1 y λ_2 de orden 1. Estos números son iguales. De hecho Karow [44, Proposición 5.5.8, p. 83] dio la expresión siguiente

$$\text{cond}_1(A, \lambda_k) = \sqrt{1 + \left(\frac{d(A)}{|\lambda_2 - \lambda_1|}\right)^2}, \quad k = 1, 2. \quad (4.2.6)$$

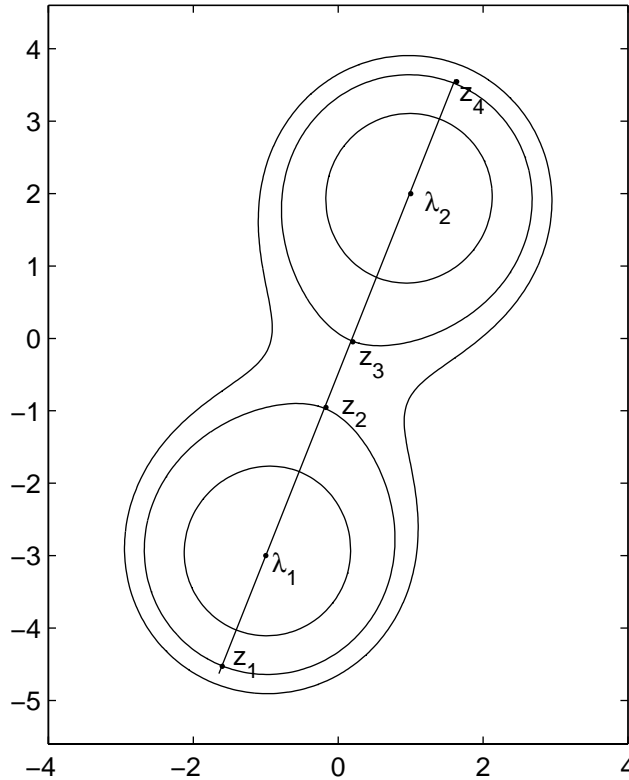


Figura 4.3: ε -pseudoespectros de A_1 para $\varepsilon = 1,00; 1,50; 1,75$.

Así pues, sea $\varepsilon \geq 0$ tal que $\Lambda_\varepsilon(A)$ tenga dos componentes conexas. A continuación calculamos los puntos de intersección z_1, z_2, z_3, z_4 de la recta $z(t) := (1-t)\lambda_1 + t\lambda_2, t \in \mathbb{R}$, con la curva $\partial\Lambda_\varepsilon(A)$. Por (4.2.5), esta curva está dada por la ecuación

$$(|z - \lambda_1|^2 - \varepsilon^2)(|z - \lambda_2|^2 - \varepsilon^2) - \varepsilon^2 d(A)^2 = 0. \quad (4.2.7)$$

Para determinar los valores del parámetro t que corresponden a los puntos z_1, z_2, z_3 y z_4 sustituimos $z(t)$ en (4.2.7),

$$\begin{aligned} & |\lambda_1 - \lambda_2|^4 t^4 - 2|\lambda_1 - \lambda_2|^4 t^3 + (|\lambda_1 - \lambda_2|^4 - 2\varepsilon^2|\lambda_1 - \lambda_2|^2) t^2 \\ & + 2\varepsilon^2|\lambda_1 - \lambda_2|^2 t - \varepsilon^2|\lambda_1 - \lambda_2|^2 + \varepsilon^4 - \varepsilon^2 d(A)^2 = 0. \end{aligned}$$

Para simplificar abreviamos $\theta := |\lambda_1 - \lambda_2|$,

$$\theta^4 t^4 - 2\theta^4 t^3 + (\theta^4 - 2\varepsilon^2\theta^2) t^2 + 2\varepsilon^2\theta^2 t - \varepsilon^2\theta^2 + \varepsilon^4 - \varepsilon^2 d(A)^2 = 0.$$

Usando el comando `solve` de MAPLE 13 resolvemos esta ecuación en t obteniendo

$$\begin{aligned} t_1 &= \frac{1}{2} - \frac{1}{2\theta} \sqrt{\theta^2 + 4\varepsilon^2 + 4\varepsilon \sqrt{\theta^2 + d(A)^2}} \\ t_2 &= \frac{1}{2} - \frac{1}{2\theta} \sqrt{\theta^2 + 4\varepsilon^2 - 4\varepsilon \sqrt{\theta^2 + d(A)^2}} \\ t_3 &= \frac{1}{2} + \frac{1}{2\theta} \sqrt{\theta^2 + 4\varepsilon^2 - 4\varepsilon \sqrt{\theta^2 + d(A)^2}} \\ t_4 &= \frac{1}{2} + \frac{1}{2\theta} \sqrt{\theta^2 + 4\varepsilon^2 + 4\varepsilon \sqrt{\theta^2 + d(A)^2}} \end{aligned}$$

donde $t_1 < 0 < t_2 < \frac{1}{2} < t_3 < 1 < t_4$. De aquí que, $z_3 = z(t_3)$ y $z_4 = z(t_4)$. Por tanto,

$$\delta(\varepsilon) = |z_4 - z_3| = \frac{1}{2} \sqrt{\theta^2 + 4\varepsilon^2 + 4\varepsilon \sqrt{\theta^2 + d(A)^2}} - \frac{1}{2} \sqrt{\theta^2 + 4\varepsilon^2 - 4\varepsilon \sqrt{\theta^2 + d(A)^2}}$$

De donde se sigue,

$$\begin{aligned} \delta'(\varepsilon) &= \frac{2\varepsilon + \sqrt{\theta^2 + d(A)^2}}{\sqrt{\theta^2 + 4\varepsilon^2 + 4\varepsilon \sqrt{\theta^2 + d(A)^2}}} - \frac{2\varepsilon - \sqrt{\theta^2 + d(A)^2}}{\sqrt{\theta^2 + 4\varepsilon^2 - 4\varepsilon \sqrt{\theta^2 + d(A)^2}}} \\ \delta'(0) &= \frac{2\sqrt{\theta^2 + d(A)^2}}{\sqrt{\theta^2}} = 2\sqrt{1 + \left(\frac{d(A)}{\theta}\right)^2}. \end{aligned}$$

Por (4.2.6), tenemos $\delta'(0) = 2 \operatorname{cond}_1(A, \lambda_2)$.

Veremos que los resultados de estos ejemplos no son casuales. De hecho, el número de condición de λ de orden $1/\nu$ está relacionado con las funciones δ y a .

4.3. Acotación mediante discos cerrados

De los teoremas, probados por Karow, [44, Teoremas 2.6.6 and 5.4.4] deducimos el teorema siguiente.

Teorema 4.3.1. *Sea λ un valor propio de $A \in \mathbb{C}^{n \times n}$ de índice ν . Para cada $\eta > 0$ existe un $\varepsilon_\eta > 0$ tal que para cualquier $\varepsilon \in (0, \varepsilon_\eta]$,*

$$\mathcal{D}(\lambda, ((1 - \eta)\varepsilon)^{1/\nu} c) \subset \mathcal{K}_\lambda(\varepsilon) \subset \mathcal{D}(\lambda, ((1 + \eta)\varepsilon)^{1/\nu} c) \quad (4.3.8)$$

siendo c el número de condición de λ de orden $1/\nu$.

De (4.3.8) deducimos que

$$\mathcal{D}(0, ((1 - \eta)\varepsilon)^{1/\nu} c) \subset \mathcal{K}_\lambda(\varepsilon) - \lambda \subset \mathcal{D}(0, ((1 + \eta)\varepsilon)^{1/\nu} c)$$

Considerando la homotecia

$$z \mapsto \frac{z}{\varepsilon^{1/\nu}}, \quad z \in \mathbb{C},$$

vemos que

$$\mathcal{D}(0, (1 - \eta)^{1/\nu} c) \subset \frac{\mathcal{K}_\lambda(\varepsilon) - \lambda}{\varepsilon^{1/\nu}} \subset \mathcal{D}(0, (1 + \eta)^{1/\nu} c). \quad (4.3.9)$$

4.4. Acotación mediante diámetros

Ya que el diámetro es una función monótona creciente respecto de \mathbb{C} , por (4.3.9),

$$2c(1 - \eta)^{1/\nu} \leq \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} \leq 2c(1 + \eta)^{1/\nu}. \quad (4.4.10)$$

De este modo, llegamos al teorema siguiente.

Teorema 4.4.1. *Sea ν el índice de λ y sea c el número de condición de orden $1/\nu$ de λ . Entonces*

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} = 2c. \quad (4.4.11)$$

DEMOSTRACIÓN. Definamos las funciones

$$f_1(\eta) := 1 - (1 - \eta)^{1/\nu}, \quad (4.4.12)$$

$$f_2(\eta) := (1 + \eta)^{1/\nu} - 1, \quad (4.4.13)$$

donde $\eta \in [0, 1]$. Si fuera $\nu = 1$, entonces

$$f_1(\eta) := 1 - 1 + \eta = \eta,$$

$$f_2(\eta) := 1 + \eta - 1 = \eta;$$

de donde tenemos que, $f_1(\eta) = f_2(\eta)$ sobre $[0, 1]$.

Si fuera $\nu = 2$, entonces $f_2(\eta) \leq f_1(\eta)$, porque

$$(\sqrt{1+\eta} - 1)^2 \leq (1 - \sqrt{1-\eta})^2$$

es equivalente a $\eta^4 \geq 0$.

En el caso de que $\nu \geq 3$, deduciremos que $f_2(\eta) \leq f_1(\eta)$. Esta desigualdad última es equivalente a

$$(1+\eta)^{1/\nu} - 1 \leq 1 - (1-\eta)^{1/\nu}, \quad \eta \in [0, 1],$$

$$(1+\eta)^{1/\nu} + (1-\eta)^{1/\nu} \leq 2, \quad \eta \in [0, 1].$$

Definamos

$$g(\eta) := (1+\eta)^{1/\nu} + (1-\eta)^{1/\nu}, \quad \eta \in [0, 1].$$

Entonces

$$g'(\eta) = \frac{1}{\nu} [(1+\eta)^{1/\nu-1} - (1-\eta)^{1/\nu-1}].$$

Ya que $x^\alpha := e^{\alpha \ln x}$, cuando $\alpha < 0$ la función $x \mapsto x^\alpha$ es decreciente en $[0, \infty)$; por lo tanto $x \mapsto x^{1/\nu-1}$ es decreciente en $[0, \infty)$. De donde, si $0 < \eta < 1$, inferimos que $1 - \eta < 1 + \eta$; lo que implica que

$$(1-\eta)^{1/\nu-1} > (1+\eta)^{1/\nu-1}.$$

Así pues, $g'(\eta) < 0$. Por ende, g es decreciente sobre $[0, 1]$. Por consiguiente $g(\eta) < g(0) = 2$. Es decir, $f_2(\eta) \leq f_1(\eta)$.

Dado que

$$\begin{aligned} 2c - 2cf_1(\eta) &= 2c(1-\eta)^{1/\nu}, \\ 2c + 2cf_2(\eta) &= 2c(1+\eta)^{1/\nu}, \end{aligned}$$

por (4.4.10), tenemos que

$$2c - 2cf_1(\eta) \leq \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} \leq 2c + 2cf_2(\eta).$$

Para cada $\nu \geq 1$, $f_2(\eta) \leq f_1(\eta)$ para $\eta \in [0, 1]$. Por tanto,

$$2c - 2cf_1(\eta) \leq \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} \leq 2c + 2cf_1(\eta),$$

$$-2cf_1(\eta) \leq \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} - 2c \leq 2cf_1(\eta),$$

o

$$\left| \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} - 2c \right| \leq 2cf_1(\eta).$$

Pero $\lim_{\eta \rightarrow 0^+} f_1(\eta) = 0$ y $f_1(\eta) > 0$ para $\eta > 0$. Así pues, fijemos un $\eta_0 > 0$, entonces existe un $\eta_1 > 0$ tal que $2cf_1(\eta_1) < \eta_0$. Para este η_1 hay un $\varepsilon_{\eta_1} > 0$ tal que para $\varepsilon \in (0, \varepsilon_{\eta_1}]$,

$$\left| \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} - 2c \right| \leq 2cf_1(\eta_1) < \eta_0.$$

Por consiguiente, existe el límite

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}}$$

y es igual a $2c$. \square

4.5. Derivadas del diámetro

En esta sección relacionaremos la derivada por la derecha del diámetro δ en 0 con el número de condición del valor propio λ , cuando $\nu = 1$. En primer lugar, tenemos el lema siguiente.

Lema 4.5.1. *Sean $A \in \mathbb{C}^{n \times n}$ y λ un valor propio de A de índice ν . Sea $\delta(\varepsilon)$ el diámetro de la componente conexa de $\Lambda_\varepsilon(A)$ que contiene a λ . Sea c el número de condición de λ de orden $1/\nu$. Entonces existe el límite*

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}}$$

y es igual a

$$\begin{cases} 2c & \text{si } \nu = 1, \\ \infty & \text{si } \nu \geq 2. \end{cases}$$

DEMOSTRACIÓN. Si fuera $\nu = 1$, el Teorema 4.4.1 implica que

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\delta(\varepsilon)}{\varepsilon} = 2c.$$

Si fuera $\nu \geq 2$, del mismo Teorema se sigue que

$$\frac{\delta(\varepsilon)}{\varepsilon} = \frac{\delta(\varepsilon)}{\varepsilon^{1/\nu}} \frac{1}{\varepsilon^{1-1/\nu}} \rightarrow 2c \cdot \infty,$$

cuando $\varepsilon \rightarrow 0^+$. \square

A partir de este lema el próximo teorema se sigue inmediatamente.

Teorema 4.5.2. *Sean $A \in \mathbb{C}^{n \times n}$ y λ un valor propio de A de índice ν . Sea $\delta(\varepsilon)$ el diámetro de la componente conexa de $\Lambda_\varepsilon(A)$ que contiene a λ . Sea c el número de condición de λ de orden $1/\nu$. Entonces, existe la derivada de δ en 0 por la derecha, $\delta'_+(0)$, y*

$$\delta'_+(0) = \begin{cases} 2c & \text{si } \nu = 1, \\ \infty & \text{si } \nu \geq 2. \end{cases}$$

4.6. Acotación por medio de áreas

Como la función del área es monótona creciente con respecto a \subset , por (4.3.9),

$$\pi c^2 (1 - \eta)^{2/\nu} \leq \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \leq \pi c^2 (1 + \eta)^{2/\nu}, \quad (4.6.14)$$

donde $a(\varepsilon) :=$ área o medida de Lebesgue de $\mathcal{K}_\lambda(\varepsilon)$. Así, llegamos al siguiente teorema.

Teorema 4.6.1. *Sea ν el índice de λ , y sea c el número de condición de λ de orden $1/\nu$. Entonces*

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} = \pi c^2. \quad (4.6.15)$$

DEMOSTRACIÓN. Definamos las funciones

$$\varphi_1(\eta) := 1 - (1 - \eta)^{2/\nu}, \quad (4.6.16)$$

$$\varphi_2(\eta) := (1 + \eta)^{2/\nu} - 1. \quad (4.6.17)$$

Si $\nu = 1$, entonces

$$\begin{aligned} \varphi_1(\eta) &= 1 - (1 - \eta)^2 = 2\eta - \eta^2, \\ \varphi_2(\eta) &= (1 + \eta)^2 - 1 = 2\eta + \eta^2; \end{aligned}$$

es obvio que $\varphi_1(\eta) \leq \varphi_2(\eta)$ cuando $\eta \geq 0$.

Si $\nu = 2$, entonces

$$\varphi_1(\eta) = 1 - (1 - \eta) = \eta = (1 + \eta) - 1 = \varphi_2(\eta).$$

Si $\nu \geq 3$, veremos que

$$\varphi_2(\eta) \leq \varphi_1(\eta), \quad \eta \in [0, 1],$$

o, equivalentemente,

$$(1 + \eta)^{2/\nu} - 1 \leq 1 - (1 - \eta)^{2/\nu}, \quad \eta \in [0, 1],$$

$$(1 + \eta)^{2/\nu} + (1 - \eta)^{2/\nu} \leq 2, \quad \eta \in [0, 1].$$

Definamos

$$\psi(\eta) := (1 + \eta)^{2/\nu} + (1 - \eta)^{2/\nu}, \quad \eta \in [0, 1].$$

Entonces

$$\psi'(\eta) = \frac{2}{\nu} \left[(1 + \eta)^{2/\nu-1} - (1 - \eta)^{2/\nu-1} \right].$$

Como la función $x \mapsto x^{2/\nu-1}$ es decreciente en $[0, \infty)$, si $0 < \eta < 1$, entonces $1 - \eta < 1 + \eta$ implica que $(1 - \eta)^{2/\nu-1} > (1 + \eta)^{2/\nu-1}$ y $\psi'(\eta) < 0$. Por tanto, ψ es decreciente en $[0, 1]$. Así, para $\eta \in (0, 1]$, $\psi(\eta) < \psi(0) = 2$. En consecuencia,

$$\varphi_2(\eta) \leq \varphi_1(\eta), \quad \eta \in [0, 1].$$

Por (4.6.16) y (4.6.17),

$$\begin{aligned} (1 - \eta)^{2/\nu} &= 1 - \varphi_1(\eta), \\ (1 + \eta)^{2/\nu} &= 1 + \varphi_2(\eta). \end{aligned}$$

De las desigualdades (4.6.14) se deduce que

$$\pi c^2 - \pi c^2 \varphi_1(\eta) \leq \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \leq \pi c^2 + \pi c^2 \varphi_2(\eta). \quad (4.6.18)$$

Caso $\nu \leq 2$. Como $\varphi_1(\eta) \leq \varphi_2(\eta)$ en $[0, 1]$,

$$-\pi c^2 \varphi_2(\eta) \leq -\pi c^2 \varphi_1(\eta).$$

Por tanto, por (4.6.18),

$$\pi c^2 - \pi c^2 \varphi_2(\eta) \leq \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \leq \pi c^2 + \pi c^2 \varphi_2(\eta); \quad (4.6.19)$$

lo que es equivalente a

$$\left| \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} - \pi c^2 \right| \leq \pi c^2 \varphi_2(\eta).$$

Pero

$$\lim_{\eta \rightarrow 0^+} \varphi_2(\eta) = 0, \quad \varphi_2(\eta) > 0, \forall \eta > 0;$$

así,

$$\lim_{\eta \rightarrow 0^+} \pi c^2 \varphi_2(\eta) = 0.$$

Fijemos $\eta_0 > 0$, entonces existe un $\eta_1 > 0$ tal que $\pi c^2 \varphi_2(\eta_1) < \eta_0$. Para este η_1 existe un $\varepsilon_{\eta_1} > 0$ tal que para $\varepsilon \in (0, \varepsilon_{\eta_1}]$,

$$\left| \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} - \pi c^2 \right| \leq \pi c^2 \varphi_2(\eta_1) < \eta_0.$$

Esto prueba que existe el límite

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^{2/\nu}}$$

y es igual a πc^2 .

Caso $\nu \geq 3$. Como $\varphi_2(\eta) \leq \varphi_1(\eta)$ en $[0, 1]$, de (4.6.18) deducimos que

$$\pi c^2 - \pi c^2 \varphi_1(\eta) \leq \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \leq \pi c^2 + \pi c^2 \varphi_1(\eta)$$

y, como $\varphi_1(\eta) > 0$ si $\eta > 0$ y

$$\lim_{\eta \rightarrow 0^+} \varphi_1(\eta) = 0,$$

por un razonamiento análogo al anterior, deducimos que existe el límite

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} = \pi c^2. \square$$

4.7. Derivadas del área

En esta sección establecemos la relación entre la primera y segundas derivadas de a en 0 por la derecha y el número de condición del valor propio λ de orden $1/\nu$, cuando $\nu = 1$ o 2. Primero, probamos el siguiente lema.

Lema 4.7.1. *Sean $A \in \mathbb{C}^{n \times n}$ y λ un valor propio de A de índice ν . Sea $a(\varepsilon)$ el área de la componente conexa del pseudoespectro $\Lambda_\varepsilon(A)$ que contiene a λ . Sea c el número de condición de λ de orden $1/\nu$. Entonces, existe el límite*

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon}$$

y es igual a

$$\begin{cases} 0 & \text{si } \nu = 1, \\ \pi c^2 & \text{si } \nu = 2, \\ \infty & \text{si } \nu \geq 3. \end{cases}$$

DEMOSTRACIÓN. Primero, supongamos que $\nu = 1$. Entonces el Teorema 4.6.1 implica que

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^2} = \pi c^2.$$

Por tanto,

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \varepsilon \frac{a(\varepsilon)}{\varepsilon^2} = \left(\lim_{\varepsilon \rightarrow 0^+} \varepsilon \right) \left(\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^2} \right) = 0 \cdot \pi c^2 = 0.$$

Segundo, asumamos que $\nu = 2$. Del Teorema 4.6.1 se deduce que

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon} = \pi c^2.$$

Finalmente, cuando $\nu \geq 3$, existen m, M, ε_0 positivos tales que

$$m < \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} < M$$

en $(0, \varepsilon_0)$; y como la función

$$\frac{1}{\varepsilon^{(\nu-2)/\nu}}$$

diverge a ∞ cuando $\varepsilon \rightarrow 0^+$, el producto

$$\frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \cdot \frac{1}{\varepsilon^{(\nu-2)/\nu}}$$

diverge a ∞ cuando $\varepsilon \rightarrow 0^+$. Pero

$$\frac{a(\varepsilon)}{\varepsilon} = \frac{a(\varepsilon)}{\varepsilon^{2/\nu}} \cdot \frac{1}{\varepsilon^{(\nu-2)/\nu}}.$$

Luego,

$$\lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon} = \infty. \square$$

Teorema 4.7.2. Sean $A \in \mathbb{C}^{n \times n}$ y λ un valor propio de A de índice ν . Sea $a(\varepsilon)$ el área de la componente conexa del pseudoespectro $\Lambda_\varepsilon(A)$ que contiene a λ . Entonces, existe la derivada de a en 0 por la derecha, $a'_+(0)$, y

$$a'_+(0) = \begin{cases} 0 & \text{si } \nu = 1, \\ \pi c^2 & \text{si } \nu = 2, \\ \infty & \text{si } \nu \geq 3. \end{cases}$$

DEMOSTRACIÓN. Por el Lema 4.7.1 deducimos que

si $\nu = 1$,

$$a'_+(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon) - a(0)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon} = 0;$$

si $\nu = 2$,

$$a'_+(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon) - a(0)}{\varepsilon} = \pi c^2;$$

si $\nu \geq 3$,

$$a'_+(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon) - a(0)}{\varepsilon} = \infty. \square$$

Teorema 4.7.3. Sean $A \in \mathbb{C}^{n \times n}$ y λ un valor propio semisimple. Sea $a(\varepsilon)$ el área de la componente conexa del pseudoespectro $\Lambda_\varepsilon(A)$ que contiene a λ . Asumamos que existe $a''_+(0)$. Entonces

$$a''_+(0) = 2\pi c^2.$$

DEMOSTRACIÓN. Definamos la función

$$A(\varepsilon) := \begin{cases} a(\varepsilon), & \text{si } \varepsilon \geq 0, \\ a(-\varepsilon), & \text{si } \varepsilon < 0. \end{cases}$$

Por el Teorema 4.7.2, como $\nu = 1$, existe $A'(0)$, y $A'(0) = 0$. Si suponemos que existe la derivada $A''(0)$, entonces

$$A''(0) = \lim_{\varepsilon \rightarrow 0} \frac{A(\varepsilon) - 2A(0) + A(-\varepsilon)}{\varepsilon^2}.$$

Pero, la existencia de este límite no implica la existencia de $A''(0)$. Véase [6, Ejercicio 5–20].

Cuando $\varepsilon > 0$, $A(\varepsilon) = a(\varepsilon)$, $A(-\varepsilon) = a(-(-\varepsilon)) = a(\varepsilon)$; por lo tanto, $A(\varepsilon) + A(-\varepsilon) = 2a(\varepsilon)$.

Cuando $\varepsilon < 0$, $A(\varepsilon) = a(-\varepsilon)$, $A(-\varepsilon) = a(-\varepsilon)$; por consiguiente $A(\varepsilon) + A(-\varepsilon) = 2a(-\varepsilon)$.

Entonces, por el Teorema 4.6.1, se tiene que

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{A(\varepsilon) - 2A(0) + A(-\varepsilon)}{\varepsilon^2} &= \lim_{\varepsilon \rightarrow 0^+} \frac{2a(\varepsilon)}{\varepsilon^2} = 2\pi c^2, \\ \lim_{\varepsilon \rightarrow 0^-} \frac{A(\varepsilon) - 2A(0) + A(-\varepsilon)}{\varepsilon^2} &= \lim_{\varepsilon \rightarrow 0^-} \frac{2a(-\varepsilon)}{(-\varepsilon)^2} = \lim_{\beta \rightarrow 0^+} \frac{2a(\beta)}{\beta^2} = 2\pi c^2. \end{aligned}$$

Consecuentemente, existe el límite

$$\lim_{\varepsilon \rightarrow 0} \frac{A(\varepsilon) - 2A(0) + A(-\varepsilon)}{\varepsilon^2}$$

y es igual a $2\pi c^2$. Así, $A''(0) = 2\pi c^2$, y $a_+''(0) = 2\pi c^2$. \square

Finalmente, remitimos a la Conjetura 7.4.4, en la página 91, sobre la existencia de la derivada $a_+''(0)$, que puede leerse una vez definidas las funciones semialgebraicas.

4.8. Área encerrada por una curva de nivel

Planteamos en esta sección el comienzo de un método alternativo que podría servir para abordar los problemas relativos a las derivadas de la función área $a(\varepsilon)$ en $\varepsilon = 0$.

Veamos una fórmula dada por Koutis [47, Teorema 14 y Corolario 3, pág. 10] que expresa el área encerrada por una curva de nivel de una función de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de clase C^∞ , por medio de una integral en la que aparece su gradiente. Primero se cita un teorema general de Konrod.

Teorema 4.8.1. *Sean $\Omega \subset \mathbb{R}^n$ un dominio (abierto y conexo) y $f : \Omega \rightarrow \mathbb{R}$ una función de clase C^∞ acotada sobre Ω . Sea*

$$E_t := \{x \in \Omega \mid f(x) = t\}$$

el conjunto de nivel t de la función f y ds el elemento $(n-1)$ -dimensional de superficie sobre E_t . Entonces la medida Lebesgue de Ω viene dada por

$$m(\Omega) = \int_{\inf f(\Omega)}^{\sup f(\Omega)} \left(\int_{E_t} \frac{ds}{\|\nabla f\|} \right) dt.$$

Sean $A \in \mathbb{C}^{n \times n}$, $z \in \mathbb{C}$. Denotemos por $u_n(z), v_n(z)$ los vectores singulares por la izquierda y derecha correspondientes al valor singular $\sigma_n(zI - A)$ de la matriz $zI - A$.

Corolario 4.8.2. *Para cada $t \geq 0$, sea $\partial\mathcal{K}_\lambda(t)$ la frontera de $\mathcal{K}_\lambda(t)$ y ds la diferencial de la longitud de arco sobre $\partial\mathcal{K}_\lambda(t)$. Entonces*

$$a(\varepsilon) = \int_0^\varepsilon \left(\int_{\partial\mathcal{K}_\lambda(t)} |v_n(s)^* u_n(s)|^{-1} ds \right) dt.$$

El Teorema Fundamental del Cálculo nos asegura que si $f : [a, b] \rightarrow \mathbb{R}$ es una función continua, entonces la función

$$F(x) := \int_a^x f(t) dt$$

es derivable y su derivada satisface la igualdad $F'(x) = f(x)$ para todo $x \in [a, b]$. Por lo tanto, como la función

$$t \mapsto \int_{\partial\mathcal{K}_\lambda(t)} |v_n(s)^* u_n(s)|^{-1} ds$$

es continua, se sigue que existe $a'(\varepsilon)$ y

$$a'(\varepsilon) = \int_{\partial\mathcal{K}_\lambda(\varepsilon)} |v_n(s)^* u_n(s)|^{-1} ds. \quad (4.8.20)$$

Observación 4.8.3. Sea C una curva plana de Jordan orientada positivamente y sea D la componente conexa interior a C de $\mathbb{R}^2 \setminus C$; definamos $R = D \cup C$. Sean P y Q funciones reales continuamente diferenciables definidas en \mathbb{R}^2 . Por el Teorema de Green sabemos que

$$\int_C P(x, y) dx + Q(x, y) dy = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy,$$

donde el miembro de la izquierda es una integral curvilínea y el miembro de la derecha es una integral doble. Si ahora elegimos P y Q de manera que $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \equiv 1$ sobre R , se tiene que

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \text{área}(R).$$

Por ejemplo, eso ocurre si $P(x, y) = 1$ y $Q(x, y) = x$ para todo $(x, y) \in \mathbb{R}^2$. En este caso, por el Teorema de Green,

$$\int_C 1 dx + x dy = \iint_R (1 - 0) dx dy = \text{área}(R).$$

También podemos deducir de manera parecida que

$$\text{área}(R) = \frac{1}{2} \int_C x dy - y dx.$$

Para esta cuestión puede consultarse el Courant-John [22].

4.9. Nota final. Radio medio

El **radio medio** de $\partial\mathcal{K}_\lambda(\varepsilon)$ alrededor de λ se define como el valor medio de $r(\theta, \varepsilon)$ al variar θ de 0 a 2π , es decir:

$$r(\varepsilon) := \frac{1}{2\pi} \int_0^{2\pi} r(\theta, \varepsilon) d\theta.$$

La explicación del significado de la función $r(\theta, \varepsilon)$ se da aquí: Sea E el semieje paralelo al semieje x positivo que arranca de λ . Consideremos la semirrecta L_t de origen λ que forma un ángulo de θ radianes con el semieje E . Hallemos el punto de corte $P_{\theta, \varepsilon}$ de esta semirrecta con la curva $\{z \in \mathbb{C} : \sigma_n(zI_n - A) = \varepsilon\}$. Entonces $r(\theta, \varepsilon)$ es igual a la distancia desde $P_{\theta, \varepsilon}$ a λ . Así pues $r(\theta, \varepsilon)$ es la distancia de λ al punto variable $P_{\theta, \varepsilon} = \lambda + r(\theta, \varepsilon)e^{i\theta} \in \mathcal{K}_\lambda(\varepsilon)$. Por tanto, el radio medio de la componente $\mathcal{K}_\lambda(\varepsilon)$ es el valor medio de la función $f_\varepsilon : \theta \mapsto r(\theta, \varepsilon)$ definida de $[0, 2\pi]$ en \mathbb{R} :

$$r(\varepsilon) := \overline{f_\varepsilon} := \frac{1}{2\pi} \int_0^{2\pi} r(\theta, \varepsilon) d\theta$$

Bajo condiciones suficientes podemos derivar $r(\varepsilon)$ respecto de ε mediante la regla de Leibniz de derivación bajo el signo integral, así se tiene que

$$r'(\varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \varepsilon} r(\theta, \varepsilon) d\theta.$$

De donde tenemos que

$$r'(0) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \varepsilon} r(\theta, 0) d\theta.$$

Aunque hemos centrado nuestra atención en las funciones diámetro y área, también podríamos haber hecho un estudio análogo para la función radio medio. E igualmente para los radios máximo y mínimo de $\partial\mathcal{K}_\lambda(\varepsilon)$ en función de ε .

Capítulo 5

Pseudoespectros circulares

Es un hecho conocido que el pseudoespectro ordinario de primer orden de una matriz normal está formado por la unión de discos, pero el recíproco no es cierto en general. Concretamente, M. Karow en [44, Sección 5.4, páginas 74–79] estudia los pseudoespectros ordinarios de primer orden de **bloques de Jordan generalizados** de la forma

$$J_n(\lambda, d) = \begin{bmatrix} \lambda & d & & & \\ & \lambda & d & & \\ & & \ddots & \ddots & \\ & & & \lambda & d \\ & & & & \lambda \end{bmatrix}, \quad (5.0.1)$$

donde λ es el único valor propio y $d \in \mathbb{C}$. Mediante transformaciones de semejanza unitaria, transforma la matriz $J_n(z - \lambda, d)$ en $J_n(|z - \lambda|, |d|)$. De esta forma prueba que la única componente del pseudoespectro posee una simetría rotacional alrededor de λ , es decir, es un círculo centrado en λ . En este capítulo probaremos que dicha propiedad se mantiene si los elementos de la superdiagonal no son todos iguales. Así mismo, generalizaremos el resultado de Karow para el pseudoespectro algebraico de orden dos. También, estudiaremos los pseudoespectros ordinarios algebraicos de segundo orden y geométrico de cualquier orden, para matrices normales.

Comenzamos el capítulo con una caracterización de las matrices normales, vía pseudoespectros.

5.1. Una caracterización de las matrices normales

El objetivo de esta sección es probar la siguiente caracterización de las matrices normales.

Teorema 5.1.1. Sean $\varepsilon \geq 0$ y $A \in \mathbb{C}^{n \times n}$ tal que $\Lambda(A) = \{\lambda_1, \dots, \lambda_n\}$. Entonces, la matriz A es normal si y sólo si

$$\Lambda_{\varepsilon,1}^a(A) = \bigcup_{j=1}^n \mathcal{D}(\lambda_j, \varepsilon). \quad (5.1.2)$$

Es decir, una matriz es normal si y sólo si el pseudoespectro ordinario algebraico de primer orden de nivel ε es la unión de discos con centro en cada valor propio λ_i y radio ε , $i = 1, \dots, n$.

Para probar el Teorema 5.1.1 es preciso dar unos resultados previos que comprenden la descomposición de Jordan para una matriz, dada en el Teorema 1.3.1.

Lema 5.1.2. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_v\}$, $v \leq n$. Entonces, A es normal si y sólo si

$$A = \sum_{i=1}^v \lambda_i P_i,$$

con $P_i P_j = \delta_{ij} P_j$ y $P_j^* = P_j$.

Una prueba de este lema se puede ver en [80, Teorema 8.1 (10), p. 241].

A partir de la definición del número de Hölder, dada en la página 29, y de [44, Teorema 5.4.4 (iv),(viii), p. 77-78], se tiene el siguiente lema

Lema 5.1.3.

$$c_{\frac{1}{\nu_j}}(\lambda_j) = \begin{cases} \|P_j\|, & \nu_j = 1, \\ \|N_j^{\nu_j-1}\|, \|\frac{1}{\nu_j}\|, & \nu_j \geq 2. \end{cases}$$

El último resultado previo precisa de la siguiente notación: O y O_t denotan la matriz nula y la matriz nula de tamaño $t \times t$, respectivamente. Por otro lado, recordamos que un proyector $P \in \mathbb{C}^{n \times n}$ es una matriz tal que $P^2 = P$. En el siguiente lema se utiliza un resultado de valores singulares de un proyector.

Lema 5.1.4. Sea $P \in \mathbb{C}^{n \times n}$, $P \neq O$, un proyector tal que $\|P\| = 1$. Entonces, $P^* = P$.

DEMOSTRACIÓN.

A partir de [42, p. 235], deducimos que si σ es un valor singular de P , entonces $\sigma = 0$ o $\sigma = 1$. Véase [5]. Mediante un resultado de Djoković, [35, Teorema 4], se puede afirmar que existe una matriz unitaria $U \in \mathbb{C}^{n \times n}$ tal que

$$U^* P U = \begin{pmatrix} 1 & s_1 \\ 0 & 0 \end{pmatrix} \oplus \dots \oplus \begin{pmatrix} 1 & s_k \\ 0 & 0 \end{pmatrix} \oplus I_m \oplus O_t,$$

si P tiene valores singulares mayores que la unidad, o

$$U^* P U = I_m \oplus O_t,$$

en los demás casos. Por lo tanto, si $\|P\| = 1$, entonces los valores singulares de P son 0 y 1. Así, $P = U(I_m \oplus O_t)U^*$. Como consecuencia $P^* = P$.

□

DEMOSTRACIÓN DEL TEOREMA 5.1.1

Si A es normal, entonces existe una matriz unitaria $U \in \mathbb{C}^{n \times n}$ tal que $U^*AU = \text{diag}(\lambda_1, \dots, \lambda_n)$. Luego, como $\Lambda_{\varepsilon,1}^a(A) = \Lambda_{\varepsilon,1}^a(U^*AU)$, se tiene

$$\Lambda_{\varepsilon,1}^a(A) = \Lambda_{\varepsilon,1}^a(\text{diag}(\lambda_1, \dots, \lambda_n)) = \bigcup_{i=1}^n \mathcal{D}(\lambda_i, \varepsilon).$$

Recíprocamente, si para cada $\varepsilon > 0$

$$\Lambda_{\varepsilon,1}^a(A) = \bigcup_{i=1}^n \mathcal{D}(\lambda_i, \varepsilon),$$

entonces para cada valor propio λ_j de A y para cada ε no negativo suficientemente pequeño tenemos $\mathcal{K}_j(\varepsilon) = \mathcal{D}(\lambda_j, \varepsilon)$. Así, el diámetro de $\mathcal{K}_j(\varepsilon)$ es igual a $\delta_j(\varepsilon) = 2\varepsilon$. Por lo tanto, $\delta'_j(\varepsilon) = 2$ y $\delta'_j(0^+) = 2$. Entonces, por el Teorema 4.5.2, resulta que el índice correspondiente al valor propio λ_j es igual a $\nu_j = 1$. Ahora por (4.1.1), resulta inmediatamente que el número de condición $c_1(\lambda_j) = 1$. Lo que lleva, por el Lema 5.1.3, a que $\|P_j\| = 1$. Finalmente, por el Lema 5.1.4, se deduce que $P_j^* = P_j$. Con lo cual, Lema 5.1.2, la matriz A es normal. □

5.2. Bloques de Jordan generalizados

Pasamos al análisis de los pseudoespectros de un bloque generalizado de Jordan.

5.2.1. Pseudoespectro ordinario de primer orden

En esta subsección vamos a probar el resultado siguiente.

Teorema 5.2.1. *Sea la matriz bidiagonal de $\mathbb{C}^{n \times n}$*

$$B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n}) := \begin{bmatrix} \lambda & a_{12} & & & \mathbf{0} \\ & \lambda & a_{23} & & \\ & & \ddots & \ddots & \\ \mathbf{0} & & & \lambda & a_{n-1,n} \\ & & & & \lambda \end{bmatrix}. \quad (5.2.3)$$

El pseudoespectro ordinario algebraico de primer orden,

$$\Lambda_{\varepsilon,1}^a(B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n})),$$

es un disco cerrado centrado en λ y cuyo radio viene dado por la expresión

$$r(\varepsilon) := \text{máx}\{r > 0 \mid \sigma_n(B(r, -|a_{12}|, \dots, -|a_{n-1,n}|)) \leq \varepsilon\}. \quad (5.2.4)$$

La prueba del resultado se fundamenta en que los valores singulares de la matriz

$$zI_n - B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n})$$

permanecen constantes al rotar z alrededor de λ ; es decir, dado $z = \lambda + re^{i\theta}$, para $r > 0$ fijo y $\theta \in [0, 2\pi]$, se tiene que el pseudoespectro ordinario algebraico de primer orden de la matriz

$$zI_n - B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n}) = (\lambda + re^{i\theta})I_n - B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n}),$$

es invariante para cualquier valor de θ .

Antes de comenzar la demostración, introduciremos la siguiente notación: si $\lambda = 0$, expresaremos la matriz (5.2.3) como

$$B_0 := B(0, a_{12}, \dots, a_{n-1,n}).$$

Por otro lado, los elementos en la superdiagonal de esa misma matriz se pueden expresar en forma exponencial como

$$a_{i,i+1} = |a_{i,i+1}|e^{i\theta_{i,i+1}}, \quad i = 1, \dots, n-1,$$

siendo $\theta_{i,i+1} = \arg(a_{i,i+1})$; finalmente, dado $z \in \mathbb{C}$, como en el párrafo anterior, definimos el siguiente número

$$w := \frac{z - \lambda}{|z - \lambda|} = e^{i\alpha}.$$

DEMOSTRACIÓN.

Definimos la matriz unitaria $D_n(w) := \text{diag}(w, w^2, \dots, w^n)$. Como es una matriz diagonal, para los w definidos previamente se satisface que

$$D_n(w)^* = D_n(w)^{-1} = \text{diag}(w^{-1}, \dots, w^{-n}). \quad (5.2.5)$$

Es fácil comprobar que las transformaciones de semejanza unitarias mediante una matriz unitaria diagonal aplicadas a una matriz nilpotente triangular dan como resultado otra matriz nilpotente triangular. Por lo tanto, considerando la definición de $D_n(w)$ y (5.2.5), podemos expresar los elementos de la matriz $D_n(w)^* B_0 D_n(w)$ como

$$(D_n(w)^* B_0 D_n(w))_{j,k} = \begin{cases} w^{-j} a_{j,j+1} w^{j+1} = e^{i\alpha} a_{j,j+1}, & \text{si } k = j+1, \\ 0, & \text{en los demás casos,} \end{cases}$$

con $j = 1, \dots, n$ y $k = 1, \dots, n$. Por lo tanto, tenemos la siguiente relación de semejanza unitaria,

$$D_n(w)^* B_0 D_n(w) = e^{i\alpha} B_0. \quad (5.2.6)$$

Sea ahora

$$U := \text{diag}(e^{-i\theta_{1,2}}, e^{-i(\theta_{1,2} + \theta_{2,3})}, \dots, e^{-i(\theta_{1,2} + \theta_{2,3} + \dots + \theta_{n-1,n})}).$$

A partir de la Ecuación (5.2.6) se tiene que los elementos de la matriz $U^*(D_n(w)^*B_0D_n(w))U$ vienen dados por

$$(U^*(D_n(w)^*B_0D_n(w))U)_{j,k} = \begin{cases} e^{i\alpha} e^{i(\sum_{\ell=1}^j \theta_{\ell, \ell+1})} a_{j, j+1} e^{-i(\sum_{\ell=1}^{j+1} \theta_{\ell, \ell+1})} = e^{i\alpha} |a_{j, j+1}|, & \text{si } k = j + 1, \\ 0, & \text{en los demás casos,} \end{cases}$$

con $j = 1, \dots, n$ y $k = 1, \dots, n$. Por lo tanto, (5.2.6) se transformará en

$$U^* e^{i\alpha} \begin{bmatrix} 0 & a_{12} & & & \mathbf{0} \\ & 0 & a_{23} & & \\ & & \ddots & \ddots & \\ & & & 0 & a_{n-1, n} \\ \mathbf{0} & & & & 0 \end{bmatrix} U = e^{i\alpha} \begin{bmatrix} 0 & |a_{12}| & & & \mathbf{0} \\ & 0 & |a_{23}| & & \\ & & \ddots & \ddots & \\ & & & 0 & |a_{n-1, n}| \\ \mathbf{0} & & & & 0 \end{bmatrix} \quad (5.2.7)$$

Si aplicamos sucesivamente las dos transformaciones de semejanza unitaria a la matriz característica de $B(\lambda, a_{1,2}, \dots, a_{n-1,n})$, llegamos al siguiente resultado

$$\begin{aligned} & (D_n(w)U)^* (zI_n - B(\lambda, a_{12}, a_{23}, \dots, a_{n-1,n})) D_n(w)U = \\ & (D_n(w)U)^* ((z - \lambda)I_n - B_0) D_n(w)U = (z - \lambda)I_n - (D_n(w)U)^* B_0 D_n(w)U = \\ & e^{i\alpha} \left\{ |z - \lambda| I_n - \begin{bmatrix} 0 & |a_{12}| & & & \mathbf{0} \\ & 0 & |a_{23}| & & \\ & & \ddots & \ddots & \\ & & & 0 & |a_{n-1, n}| \\ \mathbf{0} & & & & 0 \end{bmatrix} \right\} = \\ & e^{i\alpha} \begin{bmatrix} |z - \lambda| & |a_{12}| & & & \mathbf{0} \\ & |z - \lambda| & |a_{23}| & & \\ & & \ddots & \ddots & \\ & & & |z - \lambda| & |a_{n-1, n}| \\ \mathbf{0} & & & & |z - \lambda| \end{bmatrix}. \end{aligned} \quad (5.2.8)$$

Ahora, como

$$\sigma_k(\alpha A) = |\alpha| \sigma_k(A), \quad k = 1, \dots, n; \quad (5.2.9)$$

aplicando este resultado en (5.2.8), tenemos que

$$\begin{aligned}
& |e^{i\alpha}| \sigma_n \left(\begin{bmatrix} |z - \lambda| & |a_{12}| & & & 0 \\ & |z - \lambda| & |a_{23}| & & \\ & & \ddots & \ddots & \\ & & & |z - \lambda| & |a_{n-1,n}| \\ 0 & & & & |z - \lambda| \end{bmatrix} \right) = \\
& \sigma_n \left(\begin{bmatrix} |z - \lambda| & |a_{12}| & & & 0 \\ & |z - \lambda| & |a_{23}| & & \\ & & \ddots & \ddots & \\ & & & |z - \lambda| & |a_{n-1,n}| \\ 0 & & & & |z - \lambda| \end{bmatrix} \right). \quad (5.2.10)
\end{aligned}$$

Por lo tanto, sustituyendo en (5.2.8) $z = \lambda + re^{i\theta}$, siendo $r > 0$ fijo y θ cualquier ángulo en el intervalo $[0, 2\pi]$, se obtiene una matriz independiente del valor de θ . Con lo cual, sus valores singulares, en particular σ_n , no dependen de θ . Esto, junto con la igualdad en (5.2.10), nos permite concluir que el pseudoespectro es un círculo centrado en λ y cuyo radio viene dado por (5.2.4). \square

5.2.2. Pseudoespectro algebraico ordinario de segundo orden

En esta subsección consideraremos los bloques de Jordan generalizados de la forma (5.0.1), y probaremos que los pseudoespectros ordinarios algebraicos de segundo orden son círculos centrados en su único valor propio. Las técnicas que utilizamos para su demostración son las mismas que se han dado en la prueba del Teorema 5.2.1.

Teorema 5.2.2. *Sean $\lambda, d \in \mathbb{C}$ y $n \in \mathbb{N}^*$. El pseudoespectro $\Lambda_{\varepsilon,2}^a(J_n(\lambda, d))$ es el disco cerrado con centro en λ y radio*

$$r(\varepsilon) := \max\{r > 0 \mid \max_{t \geq 0} \sigma_{2n-1} \left(\begin{bmatrix} J_n(r, -|d|) & tI_n \\ O & J_n(r, -|d|) \end{bmatrix} \right) \leq \varepsilon\}. \quad (5.2.11)$$

DEMOSTRACIÓN.

Sea $z \in \mathbb{C}$, como queremos determinar el pseudoespectro ordinario algebraico de segundo orden de la matriz $J_n(\lambda, d)$, consideramos la siguiente matriz

$$\begin{bmatrix} zI_n - J_n(\lambda, d) & tI_n \\ O & zI_n - J_n(\lambda, d) \end{bmatrix} = \begin{bmatrix} J_n(z - \lambda, -d) & tI_n \\ O & J_n(z - \lambda, -d) \end{bmatrix}, \quad (5.2.12)$$

Asimismo, definimos las siguientes matrices unitarias de $\mathbb{C}^{2n \times 2n}$

$$D_{2n}(w) := \text{diag}(w, \dots, w^{2n}), \quad \text{donde } w = \frac{(z - \lambda)\bar{d}}{|(z - \lambda)d|}, \quad (5.2.13)$$

y

$$U := \text{diag}(\underbrace{1, \dots, 1}_{n \text{ veces}}, \underbrace{e^{-i((n-1)\alpha+n\beta)}, \dots, e^{-i((n-1)\alpha+n\beta)}}_{n \text{ veces}}), \quad (5.2.14)$$

$$\text{donde } \alpha = \arg\left(\frac{(z - \lambda)}{|z - \lambda|}\right) \text{ y } \beta = \arg\left(\frac{\bar{d}}{d}\right).$$

Aplicando la transformación de semejanza unitaria a la matriz (5.2.12) mediante la matriz dada en (5.2.13), obtenemos la siguiente expresión

$$\begin{aligned} D_{2n}(w)^* \begin{bmatrix} J_n(z - \lambda, -d) & {}^t I_n \\ O & J_n(z - \lambda, -d) \end{bmatrix} D_{2n}(w) = \\ e^{i\alpha} \begin{bmatrix} J_n(|z - \lambda|, -|d|) & te^{i((n-1)\alpha+n\beta)} I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix}. \end{aligned} \quad (5.2.15)$$

Si aplicamos la transformación de semejanza unitaria a (5.2.15) mediante la matriz dada en (5.2.14), obtenemos

$$\begin{aligned} U^* e^{i\alpha} \begin{bmatrix} J_n(|z - \lambda|, -|d|) & te^{i((n-1)\alpha+n\beta)} I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix} U = \\ e^{i\alpha} \begin{bmatrix} J_n(|z - \lambda|, -|d|) & {}^t I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix} \end{aligned} \quad (5.2.16)$$

Tomando la matriz en (5.2.16) y utilizando (5.2.9), se tiene

$$\begin{aligned} \sigma_{2n-1} \left(e^{i\alpha} \begin{bmatrix} J_n(|z - \lambda|, -|d|) & {}^t I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix} \right) = \\ \sigma_{2n-1} \left(\begin{bmatrix} J_n(|z - \lambda|, -|d|) & {}^t I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix} \right). \end{aligned} \quad (5.2.17)$$

Utilizando (5.2.17), podemos caracterizar el pseudoespectro ordinario algebraico de segundo orden, $\Lambda_{\varepsilon,2}^a(J_n(\lambda, d))$, como

$$\{z \in \mathbb{C} \mid \max_{t \geq 0} \sigma_{2n-1} \left(\begin{bmatrix} J_n(|z - \lambda|, -|d|) & {}^t I_n \\ O & J_n(|z - \lambda|, -|d|) \end{bmatrix} \right) \leq \varepsilon\}.$$

Por lo tanto, $\Lambda_{\varepsilon,2}^a(J_n(\lambda, d))$ está formado por una única componente circular centrada en λ , ya que para todos los puntos de la forma $z = \lambda + re^{i\theta}$, con $r > 0$ fijo y θ variando en el intervalo $[0, 2\pi]$, el valor de $|z - \lambda|$ es igual a r independientemente del valor de θ . El radio de ese círculo viene dado por (5.2.9). \square

Observación 5.2.3. El valor del parámetro t para el que se alcanza el máximo de σ_{2n-1} en el punto $z = \lambda$, no tiene por qué ser igual a 0.

5.3. Matrices normales

En esta sección, vamos a estudiar los pseudoespectros ordinarios algebraicos de segundo orden y geométrico de cualquier orden para matrices normales. Para los algebraicos de segundo orden, probaremos que el pseudoespectro es la unión de discos cerrados centrados en el punto medio que une dos valores propios.

Previamente, vamos a introducir una notación correspondiente a las n -tuplas desordenadas de números complejos que las denotaremos por $\text{Eig}(A)$. Decimos que dos n -tuplas $(\lambda_1, \lambda_2, \dots, \lambda_n)$, $(\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{C}^n$ son equivalentes si existe una permutación π del conjunto $\{1, 2, \dots, n\}$ tal que

$$(\mu_1, \mu_2, \dots, \mu_n) \sim (\lambda_{\pi(1)}, \lambda_{\pi(2)}, \dots, \lambda_{\pi(n)}).$$

Esto define una relación de equivalencia \sim en \mathbb{C}^n . Denotamos por \mathbb{C}_{sim}^n el conjunto cociente \mathbb{C}^n / \sim . Llamamos n -tupla desordenada de números complejos a cada elemento $[(\lambda_1, \lambda_2, \dots, \lambda_n)] \in \mathbb{C}_{sim}^n$, que denotaremos en forma abreviada como $(\lambda_1, \lambda_2, \dots, \lambda_n)$ sin riesgo de confusión. Supongamos ahora que $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores propios de una matriz $A \in \mathbb{C}^{n \times n}$ donde los valores múltiples aparecen repetidos, denotamos por $\text{Eig}(A)$ la n -tupla desordenada $(\lambda_1, \lambda_2, \dots, \lambda_n)$. Así, $\text{Eig}(A) \in \mathbb{C}_{sim}^n$. Por lo tanto, hacemos notar que los conjuntos $\Lambda(A)$ y $\text{Eig}(A)$ pueden ser distintos, de hecho, el espectro $\Lambda(A)$ es un subconjunto de $\text{Eig}(A)$ y puede tener menos de n elementos. Véase [12, p. 153].

5.3.1. Pseudoespectros algebraicos de segundo orden

El resultado que vamos a demostrar en esta subsección es el siguiente.

Teorema 5.3.1. *Sea una matriz $A \in \mathbb{C}^{n \times n}$ normal tal que*

$$\text{Eig}(A) = (\lambda_1, \dots, \lambda_n),$$

y sean ε un número real no negativo y (j, k) un par de números enteros tales que $1 \leq j < k \leq n$. Si $\varepsilon \geq \frac{|\lambda_j - \lambda_k|}{2}$, definimos

$$\rho_{jk}(\varepsilon) := \sqrt{\varepsilon^2 - \frac{|\lambda_j - \lambda_k|^2}{4}}, \quad (5.3.18)$$

$$\mathcal{D}_{jk}(\varepsilon) := \mathcal{D}\left(\frac{\lambda_j + \lambda_k}{2}, \rho_{jk}(\varepsilon)\right). \quad (5.3.19)$$

Si $\varepsilon < \frac{|\lambda_j - \lambda_k|}{2}$, definimos $\mathcal{D}_{jk}(\varepsilon) := \emptyset$.

Entonces,

$$\Lambda_{\varepsilon,2}^a(A) = \bigcup_{1 \leq j < k \leq n} \mathcal{D}_{jk}(\varepsilon).$$

Para su demostración, necesitamos unos resultados previos. Comencemos con una notación. Dados $z \in \mathbb{C}$, $t \geq 0$ y $A \in \mathbb{C}^{n \times n}$, recordemos de (3.1.4) la función $F_z(t)$ definida por

$$F_z(t) := \begin{pmatrix} zI_n - A & tI_n \\ 0 & zI_n - A \end{pmatrix}.$$

Como pretendemos calcular sus valores singulares, entonces si A es una matriz normal, no se pierde generalidad si $A = \text{diag}(\lambda_1, \dots, \lambda_n)$. Para esa función, los valores singulares vienen dados por el siguiente resultado.

Proposición 5.3.2. *Dado $z \in \mathbb{C}$, los valores singulares de $F_z(t)$, no necesariamente ordenados, vienen dados por*

$$f_{z,1}^-(t), f_{z,1}^+(t), f_{z,2}^-(t), f_{z,2}^+(t), \dots, f_{z,n}^-(t), f_{z,n}^+(t),$$

donde, para cada $i = 1, 2, \dots, n$

$$\begin{cases} f_{z,i}^-(t) := \sqrt{\frac{t^2 + 2|z - \lambda_i| - \sqrt{t^4 + 4t^2|z - \lambda_i|^2}}{2}}, \\ f_{z,i}^+(t) := \sqrt{\frac{t^2 + 2|z - \lambda_i| + \sqrt{t^4 + 4t^2|z - \lambda_i|^2}}{2}}. \end{cases} \quad (5.3.20)$$

DEMOSTRACIÓN.

Es suficiente observar que, permutando filas y columnas adecuadamente, el sistema de valores singulares de $F_z(t)$ se corresponde al de la matriz

$$\bigoplus_{i=1}^n \begin{pmatrix} z - \lambda_i & t \\ 0 & z - \lambda_i \end{pmatrix}.$$

□

Pasamos a enunciar y demostrar propiedades de la funciones $f_{z,i}^-(t)$ y $f_{z,i}^+(t)$ definidas para $t \geq 0$.

Proposición 5.3.3. *Para cualquier $i = 1, 2, \dots, n$, se tiene:*

(a) $f_{z,i}^-(0) = f_{z,i}^+(0) = |z - \lambda_i|.$

(b) Si $z = \lambda_j$, entonces $f_{z,j}^-(t) = 0$ para $t \geq 0$.

(c)

$$\lim_{t \rightarrow \infty} f_{z,i}^+(t) = \infty \text{ y } \lim_{t \rightarrow \infty} f_{z,i}^-(t) = 0.$$

(d) $f_{z,i}^+(t)$ es creciente.

(e) $f_{z,i}^-(t)$ es decreciente.

DEMOSTRACIÓN.

Los asertos (a), (b), (c) y (d) son inmediatos por (5.3.20). Probar (e) es equivalente a probar que la función

$$2f_{z,i}^-(t)^2 = t^2 + 2|z - \lambda_i| - \sqrt{t^4 + 4t^2|z - \lambda_i|^2}$$

es decreciente, o equivalentemente, que la función

$$g(t) := t^2 - \sqrt{t^4 + 4t^2|z - \lambda_i|^2}$$

es decreciente para $t \geq 0$. Si derivamos respecto de t , con $t > 0$, tenemos

$$g'(t) \leq 0 \Leftrightarrow 2t - \frac{2t^3 + 4t^2|z - \lambda_i|^2}{\sqrt{t^4 + 4t^2|z - \lambda_i|^2}} \leq 0 \Leftrightarrow 1 \leq \frac{t^2 + 2t|z - \lambda_i|^2}{\sqrt{t^4 + 4t^2|z - \lambda_i|^2}} \Leftrightarrow$$

$$t^4 + 4t^2|z - \lambda_i|^2 \leq (t^2 + 2|z - \lambda_i|^2)^2,$$

lo cual es cierto siempre. \square

Proposición 5.3.4. *Si $|z - \lambda_j| \leq |z - \lambda_k|$, para $t \geq 0$ se tiene:*

$$(a) f_{z,j}^+(t) \leq f_{z,k}^+(t).$$

$$(b) f_{z,j}^-(t) \leq f_{z,k}^-(t).$$

DEMOSTRACIÓN. El apartado (a) se puede deducir inmediatamente de (5.3.20). Probar (b) es equivalente a probar $2f_{z,j}^-(t)^2 \leq 2f_{z,k}^-(t)^2$; utilizando la expresión en (5.3.20), es suficiente probar que la función

$$g(x) := 2x^2 - \sqrt{t^4 + 4t^2x^2}$$

es creciente para $x \geq 0$. Si derivamos respecto de x , con $x > 0$, tenemos

$$g'(x) \geq 0 \Leftrightarrow x - \frac{t^2x}{\sqrt{t^4 + 4t^2x^2}} \geq 0 \Leftrightarrow 1 \geq \frac{t^2}{\sqrt{t^4 + 4t^2x^2}},$$

lo cual es cierto. \square

Proposición 5.3.5. *Supongamos que $|z - \lambda_j| \leq |z - \lambda_k|$. Entonces, $f_{z,j}^+(t) = f_{z,k}^-(t)$ si y sólo si $t = t_{jk}^z$, donde*

$$t_{jk}^z = \begin{cases} 0, & \text{si } \lambda_j = \lambda_k, \\ \frac{|z - \lambda_j|^2 - |z - \lambda_k|^2}{\sqrt{2(|z - \lambda_j|^2 + |z - \lambda_k|^2)}}, & \text{si } \lambda_j \neq \lambda_k. \end{cases} \quad (5.3.21)$$

DEMOSTRACIÓN.

El caso $|z - \lambda_j| = |z - \lambda_k|$ se deduce inmediatamente de (5.3.20). Supongamos que $|z - \lambda_j| < |z - \lambda_k|$. Probar que $f_{z,j}^+(t) = f_{z,k}^-(t)$ es equivalente a probar que

$2f_{z,j}^+(t)^2 = 2f_{z,k}^-(t)^2$; a partir de la expresión (5.3.20), se tiene que $f_{z,j}^+(t) = f_{z,k}^-(t)$ si y sólo si

$$\begin{aligned} 2|z - \lambda_j| - \sqrt{t^4 + 4t^2|z - \lambda_j|^2} &= 2|z - \lambda_k| - \sqrt{t^4 + 4t^2|z - \lambda_k|^2} \Leftrightarrow \\ \sqrt{t^4 + 4t^2|z - \lambda_j|^2} &= 2(|z - \lambda_k|^2 - |z - \lambda_j|^2) - \sqrt{t^4 + 4t^2|z - \lambda_k|^2}. \end{aligned}$$

Si definimos el número positivo $\alpha := 2(|z - \lambda_k|^2 - |z - \lambda_j|^2)$, y elevamos al cuadrado la anterior expresión, deducimos que $f_{z,j}^+(t) = f_{z,k}^-(t)$ si y sólo si

$$\begin{aligned} t^4 + 4t^2|z - \lambda_j|^2 &= \alpha^2 + t^4 + 4t^2 - 2\alpha\sqrt{t^4 + 4t^2|z - \lambda_k|^2} \Leftrightarrow \\ 2\alpha\sqrt{t^4 + 4t^2|z - \lambda_k|^2} &= \alpha^2 + 2\alpha t^2 \Leftrightarrow 2\alpha\sqrt{t^4 + 4t^2|z - \lambda_k|^2} = \alpha + 2t^2 \Leftrightarrow \\ 4t^4 + 16t^2|z - \lambda_k|^2 &= \alpha^2 + 4t^4 + 4\alpha t^2. \end{aligned}$$

De donde deducimos que

$$t^2 = \frac{(|z - \lambda_j|^2 - |z - \lambda_k|^2)^2}{2(|z - \lambda_j|^2 + |z - \lambda_k|^2)}.$$

□

La siguiente proposición describe el valor singular $\sigma_{2n-1}(F_z(t))$

Proposición 5.3.6. *Supongamos que*

$$|z - \lambda_j| \leq |z - \lambda_k| \leq |z - \lambda_i|, \quad i = 1, 2, \dots, n, \quad i \neq j, k.$$

Entonces,

$$\sigma_{2n-1}(F_z(t)) = \min\{f_{z,j}^+(t), f_{z,k}^-(t)\}.$$

Aún más, si $\lambda_j = \lambda_k$,

$$\sigma_{2n-1}(F_z(t)) = f_{z,j}^+(t).$$

DEMOSTRACIÓN.

Como $|z - \lambda_j| \leq |z - \lambda_i|$, para $i \neq j$, por la Proposición 5.3.4, se tiene para $t \geq 0$

$$f_{z,j}^-(t) \leq f_{z,i}^-(t), \quad \text{y} \quad f_{z,j}^-(t) \leq f_{z,j}^+(t) \leq f_{z,i}^+(t),$$

para todo $i \neq j$. Así,

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t). \quad (5.3.22)$$

Por otro lado, sea t_{jk}^z el punto donde $f_{z,j}^+(t) = f_{z,k}^-(t)$. La Proposición 5.3.3 establece que $f_{z,j}^+(t)$ es creciente y $f_{z,k}^-(t)$ decreciente, además se tiene que $f_{z,j}^+(t_{jk}^z) = f_{z,k}^-(t_{jk}^z)$; con lo cual, podemos concluir que $f_{z,j}^+(t) \leq f_{z,k}^-(t)$ en el intervalo $[0, t_{jk}^z]$. Si tenemos en cuenta este último resultado, la Ecuación (5.3.22) y la Proposición 5.3.4, podemos establecer que

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t) \leq f_{z,j}^+(t) \leq f_{z,k}^-(t) \leq \begin{cases} f_{z,i}^-(t), & \text{si } i \neq j, k, \\ f_{z,i}^+(t), & \text{si } i \neq j. \end{cases}$$

En consecuencia, $\sigma_{2n-1}(F_z(t)) = f_{z,j}^+(t)$, para $t \in [0, t_{jk}^z]$.

Ahora nos fijamos en $t \in [t_{jk}^z, \infty)$. Recordamos nuevamente que la Proposición 5.3.3 establece que $f_{z,j}^+(t)$ es creciente y $f_{z,k}^-(t)$ decreciente, además se tiene que $f_{z,j}^+(t_{jk}^z) = f_{z,k}^-(t_{jk}^z)$; con lo cual, podemos concluir que $f_{z,j}^+(t) \geq f_{z,k}^-(t)$ en el intervalo $[t_{jk}^z, \infty)$. Si tenemos en cuenta este último resultado, la Ecuación (5.3.22) y la Proposición 5.3.4, podemos establecer que

$$\sigma_{2n}(F_z(t)) = f_{z,j}^-(t) \leq f_{z,k}^-(t) \leq \begin{cases} f_{z,j}^+(t) \leq f_{z,i}^+(t), & \text{si } i \neq j, k, \\ f_{z,i}^-(t), & \text{si } i \neq j, k. \end{cases}$$

Por lo tanto, $\sigma_{2n-1}(F_z(t)) = f_{z,j}^-(t)$, para $t \in [t_{jk}^z, \infty)$. Observamos que si $\lambda_j = \lambda_k$, se tiene $t_{jk}^z = 0$ y $\sigma_{2n-1}(F_z(t)) = f_{z,k}^-(t) = f_{z,j}^-(t)$. □

Para la siguiente propiedad necesitamos introducir alguna notación. Sean $z \in \mathbb{C}$ y λ_j, λ_k dos valores propios de A , supongamos que $|z - \lambda_j| \leq |z - \lambda_k|$. Con estas consideraciones, definimos

$$g_{jk}(z, t) := \min\{f_{z,j}^+(t), f_{z,k}^-(t)\}, \quad (5.3.23)$$

$$g_{jk}^z := \max_{t \geq 0} g_{jk}(z, t). \quad (5.3.24)$$

Con estas notaciones, presentamos el siguiente resultado

Proposición 5.3.7. *Si t_{jk}^z es el valor definido en (5.3.21), entonces $g_{jk}^z = g_{jk}(z, t_{jk}^z)$.*

DEMOSTRACIÓN.

Por la Proposición 5.3.5, se satisface que $f_{z,j}^+(t_{jk}^z) = f_{z,j}^+(t_{jk}^z)$. Al igual que en la Proposición 5.3.6, podemos deducir que $f_{z,j}^+(t) \leq f_{z,k}^-(t)$ en el intervalo $[0, t_{jk}^z]$, y $f_{z,j}^+(t) \geq f_{z,k}^-(t)$ en el intervalo $[t_{jk}^z, \infty)$. Así, podemos concluir la prueba utilizando el resultado de la Proposición 5.3.3 que establece que $f_{z,j}^+(t)$ es creciente y $f_{z,k}^-(t)$ decreciente. □

Proposición 5.3.8. *Supongamos que $|z - \lambda_j| \leq |z - \lambda_k| \leq |z - \lambda_m|$. Entonces, se tiene que*

(a) *Para toda $t \geq 0$, $g_{jk}(z, t) \leq g_{km}(z, t)$ y $g_{jk}(z, t) \leq g_{jm}(z, t)$.*

(b) *$g_{jk}^z \leq g_{km}^z$ y $g_{jk}^z \leq g_{jm}^z$.*

DEMOSTRACIÓN.

Como $|z - \lambda_j| \leq |z - \lambda_k| \leq |z - \lambda_m|$, según la Proposición 5.3.4, para $t \geq 0$ se tiene que

$$f_{z,j}^-(t) \leq f_{z,k}^-(t) \leq f_{z,m}^-(t), \text{ y } f_{z,j}^+(t) \leq f_{z,k}^+(t) \leq f_{z,m}^+(t).$$

Así,

$$\min\{f_{z,j}^+(t), f_{z,k}^-(t)\} \leq \min\{f_{z,k}^+(t), f_{z,m}^-(t)\},$$

y

$$\min\{f_{z,j}^+(t), f_{z,k}^-(t)\} \leq \min\{f_{z,j}^+(t), f_{z,m}^-(t)\}.$$

Por lo tanto, (a) se sigue de la expresión dada para $g_{pq}(t)$ en (5.3.24). (b) se sigue inmediatamente de (a). \square

Para concluir esta sección, veamos una relación entre la función h_2 y g_{jk}^z , definidas en (3.1.3) y (5.3.24), respectivamente.

Proposición 5.3.9. *Sean $z \in \mathbb{C}$ y λ_j, λ_k dos valores propios de A . Supongamos que $|z - \lambda_j| < |z - \lambda_k|$. Entonces,*

(a) *Si $|z - \lambda_k| < |z - \lambda_i|$, $i \neq j, k$, se tiene $h_2(z) = g_{jk}^z$.*

(b) $h_2(z) \leq g_{jk}^z$.

DEMOSTRACIÓN.

Observemos que si λ_p, λ_q son valores propios de A tales que

$$|z - \lambda_p| \leq |z - \lambda_q| \leq |z - \lambda_i|, \quad i \neq p, q,$$

entonces, por la Proposición 5.3.4 y (5.3.23), se tiene

$$\sigma_{2n-1}(F_z(t)) = \min\{f_{z,p}^+(t), f_{z,q}^-(t)\} = g_{pq}(z, t).$$

Por lo tanto, $h_2(z) = g_{jk}^z$.

Para probar (b), consideramos varias situaciones: (1) $j = p, k = q$; (2) $j = p, k \neq q$; (3) $j = q$; (4) $j \neq p, q$. El caso (1) es inmediato por (a). Para el caso (2), como $|z - \lambda_j| = |z - \lambda_p| \leq |z - \lambda_q| \leq |z - \lambda_k|$, utilizando la Proposición 5.3.8 y el aserto (a), se tiene $h_2(z) = g_{jq}^z \leq g_{jk}^z$. Los casos (3) y (4) se resuelve de forma análoga. \square

Teorema 5.3.10. *Sean λ_j, λ_k términos de $\text{Eig}(A)$. Supongamos que $\varepsilon \geq \frac{|\lambda_j - \lambda_k|}{2}$. Definimos el punto medio del segmento de extremos λ_j y λ_k ,*

$$z_0 := \frac{\lambda_j + \lambda_k}{2}.$$

Entonces, para cada $z \in \mathbb{C}$ tal que $|z - z_0| = \rho_{jk}(\varepsilon)$, donde $\rho_{jk}(\varepsilon)$ se define en (5.3.18), si $|z - \lambda_j| \leq |z - \lambda_k|$ o $|z - \lambda_k| \leq |z - \lambda_j|$, se tiene

$$g_{jk}^z = \varepsilon \text{ o } g_{kj}^z = \varepsilon, \text{ respectivamente.} \quad (5.3.25)$$

DEMOSTRACIÓN.

Primero, asumamos que $\lambda_j \neq \lambda_k$; de lo contrario, la prueba es fácil. Para abreviar la notación, escribimos ρ en lugar de $\rho_{jk}(\varepsilon)$ y

$$\begin{cases} a := \text{Re}\left(\frac{\lambda_j - \lambda_k}{2}\right), b := \text{Im}\left(\frac{\lambda_j - \lambda_k}{2}\right). \\ \theta_1 := \frac{\pi}{2} + \arctan\left(\frac{b}{a}\right), \text{ si } a \neq 0; \theta_1 := \pi, \text{ si } a = 0. \\ \alpha(\theta) := a \cos \theta + b \sin \theta, \theta_1 \leq \theta \leq \theta_1 + \pi. \end{cases} \quad (5.3.26)$$

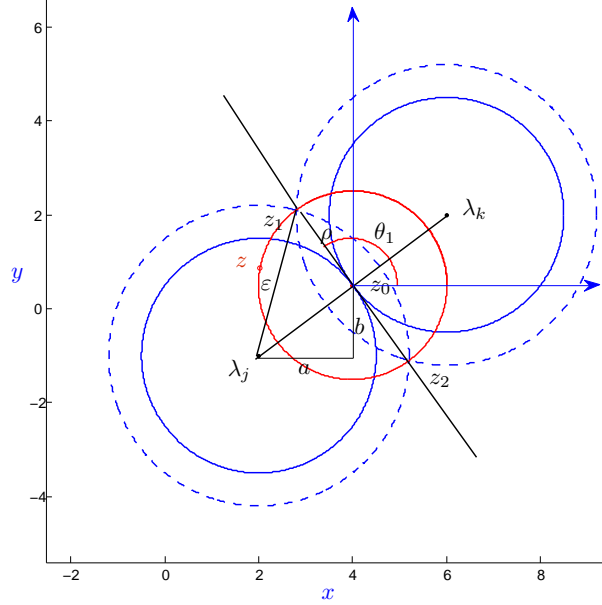


Figura 5.1: En el entorno de z_0 , punto medio entre dos valores propios próximos λ_j y λ_k . Además, $z_1 = z_0 + \rho e^{i\theta_1}$ y $z_2 = z_0 + \rho e^{i(\theta_1+\pi)}$ son los puntos de intersección de los círculos con centro en λ_j y λ_k , y de radio ε .

Es fácil demostrar que para cada $\theta \in (\theta_1, \theta_1 + \pi)$, $\alpha(\theta) < 0$; así mismo, $\alpha(\theta_1) = 0 = \alpha(\theta_1 + \pi)$. Hacemos notar que, a partir de la expresión (5.3.18), se puede inferir que

$$\varepsilon^2 = a^2 + b^2 + \rho^2. \quad (5.3.27)$$

Aún más, de (5.3.26) se puede probar que

$$\varepsilon^2 \geq \rho |\alpha(\theta)|. \quad (5.3.28)$$

Definamos $z := z_0 + \rho e^{i\theta}$. A partir de (5.3.26), se puede concluir que

$$z - \lambda_j = z_0 + \rho e^{i\theta} - \lambda_j = \frac{\lambda_j - \lambda_k}{2} + \rho e^{i\theta} = a + \rho \cos \theta + i(b + \rho \sin \theta).$$

De forma análoga, $z - \lambda_k = -a + \rho \cos \theta + i(-b + \rho \sin \theta)$. Así, por las Ecuaciones (5.3.26) y (5.3.27), tenemos que

$$\begin{cases} |z - \lambda_j|^2 = a^2 + b^2 + \rho^2 + 2\rho(a \cos \theta + b \sin \theta), \\ |z - \lambda_k|^2 = a^2 + b^2 + \rho^2 - 2\rho(a \cos \theta + b \sin \theta), \end{cases}$$

esto es,

$$\begin{cases} |z - \lambda_j|^2 = \varepsilon^2 + 2\alpha(\theta)\rho, \\ |z - \lambda_k|^2 = \varepsilon^2 - 2\alpha(\theta)\rho. \end{cases} \quad (5.3.29)$$

Supongamos que $|z - \lambda_j| \leq |z - \lambda_k|$, por la Proposición 5.3.7 tenemos $g_{jk}^z = g_{jk}(z, t_{jk}^z)$, donde t_{jk}^z está definido en (5.3.21). Utilizando (5.3.29), tenemos

$$t_{jk}^z = \frac{-2\alpha(\theta)\rho}{\varepsilon}. \quad (5.3.30)$$

Por otro lado, mediante las expresiones (5.3.23) y (5.3.24), y considerando el resultado de la Proposición 5.3.7, $g_{jk}^z = f_{z,j}^+(t_{jk}^z)$, es suficiente ver que $g_{jk}^z = f_{z,j}^+(t_{jk}^z) = \varepsilon$ para probar el lema. Si utilizamos la expresión para $f_{z,j}^+(t_{jk}^z)$ dada en (5.3.20), deducimos que

$$2(g_{jk}^z)^2 = (t_{jk}^z)^2 + 2|z - \lambda_j|^2 + t_{jk}^z \sqrt{(t_{jk}^z)^2 + 4|z - \lambda_j|^2}.$$

Utilizando las expresiones para $|z - \lambda_j|^2$ y t_{jk}^z derivadas de (5.3.29) y (5.3.30), respectivamente, concluimos que

$$\begin{aligned} (g_{jk}^z)^2 &= \frac{2\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho - \frac{2\alpha(\theta)\rho}{\varepsilon} \sqrt{\frac{\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho} = \\ &= \frac{2\alpha(\theta)^2\rho^2}{\varepsilon^2} + \varepsilon^2 + 2\alpha(\theta)\rho - \frac{2\alpha(\theta)\rho}{\varepsilon} \left| \frac{\alpha(\theta)\rho}{\varepsilon} + \varepsilon \right|. \end{aligned}$$

Según (5.3.28), $\frac{\alpha(\theta)\rho}{\varepsilon} + \varepsilon \geq 0$, de donde se sigue inmediatamente que $(g_{jk}^z)^2 = \varepsilon^2$.

El caso $|z - \lambda_k| \leq |z - \lambda_j|$ se puede razonar de forma similar. □

Como consecuencia de este teorema y de la Proposición 5.3.9, podemos deducir el siguiente resultado

Corolario 5.3.11. *Con las notaciones del Teorema 5.3.10,*

(a) *sea $z \in \mathbb{C}$ tal que $|z - z_0| = \rho_{jk}(\varepsilon)$ y*

$$\max\{|z - \lambda_j|, |z - \lambda_k|\} \leq |z - \lambda_i|, \quad i \neq j, k.$$

Entonces, $h_2(z) = \varepsilon$.

(b) *para todo $z \in \mathcal{D}_{jk}(\varepsilon)$, donde $\mathcal{D}_{jk}(\varepsilon)$ está definido en (5.3.19), se satisface que $h_2(z) \leq \varepsilon$.*

Ahora podemos probar el Teorema 5.3.1

DEMOSTRACIÓN. (Teorema 5.3.1)

Supongamos que $z \in \mathcal{D}_{jk}(\varepsilon)$. Entonces, por el Corolario 5.3.11(b), se tiene $h_2(z) \leq \varepsilon$. Como $\Lambda_{\varepsilon,2}^a(A) = \{z \in \mathbb{C} \mid h_2(z) \leq \varepsilon\}$, se tiene $z \in \Lambda_{\varepsilon,2}^a(A)$.

Recíprocamente, sea $w \in \Lambda_{\varepsilon,2}^a(A)$, equivalentemente, $h_2(w) \leq \varepsilon$. Elijamos λ_j, λ_k tales que

$$|w - \lambda_j| \leq |w - \lambda_k| \leq |w - \lambda_i|, \quad i \neq j, k.$$

Definiendo $z_0 := \frac{\lambda_j + \lambda_k}{2}$, podemos escribir $w = z_0 + \rho_0 e^{i\theta}$, para cierto $\rho_0 \geq 0$ y $\theta \in [0, 2\pi)$. Observemos que por el Corolario 5.3.11(a) se tiene $h_2(w) = \varepsilon_0$, donde

$$\varepsilon_0 := \sqrt{\rho_0^2 + \left(\frac{|\lambda_j - \lambda_k|}{2}\right)^2}.$$

Así, como $h_2(w) \leq \varepsilon$, podemos concluir que

$$\sqrt{\rho_{jk}(\varepsilon)^2 + \left(\frac{|\lambda_j - \lambda_k|}{2}\right)^2} = \varepsilon \geq \varepsilon_0;$$

es decir, $\rho_0 \leq \rho_{jk}(\varepsilon)$. □

Como consecuencia del Teorema 5.3.1, tenemos el siguiente resultado que describe el conjunto $\Lambda_{\varepsilon,2}^a(A)$ cuando todos los valores propios de A son múltiples.

Corolario 5.3.12. *Sea $A \in \mathbb{C}^{n \times n}$ una matriz normal, y supongamos que $\Lambda(A) = \Lambda_2^a(A) = \{\lambda_1, \dots, \lambda_u\}$. Entonces, para todo $\varepsilon \geq 0$, se tiene que*

$$\Lambda_{\varepsilon,2}^a(A) = \bigcup_{i=1}^u \mathcal{D}(\lambda_i, \varepsilon).$$

DEMOSTRACIÓN.

Observemos que, por el Teorema 5.3.1, es suficiente probar que para cualquier par de valores propios $\lambda_j \neq \lambda_k$ de A se ha de satisfacer

$$\mathcal{D}_{jk}(\varepsilon) \subset \mathcal{D}(\lambda_j, \varepsilon) \cup \mathcal{D}(\lambda_k, \varepsilon),$$

para cada $\varepsilon \geq 0$. Observemos que si $\varepsilon < \frac{|\lambda_j - \lambda_k|}{2}$, entonces $\mathcal{D}_{jk}(\varepsilon) = \emptyset$; es decir, es un caso trivial. Supongamos que $\varepsilon \geq \frac{|\lambda_j - \lambda_k|}{2}$. Si $z \in \text{scr} \mathcal{D}_{jk}(\varepsilon)$, con las notaciones en (5.3.26), de la prueba del Teorema 5.3.10 se sigue que

$$\begin{cases} |z - \lambda_j|^2 = \varepsilon^2 + 2\alpha(\theta)\rho, \\ |z - \lambda_k|^2 = \varepsilon^2 - 2\alpha(\theta)\rho. \end{cases}$$

Así, si $|z - \lambda_j|z \leq |z - \lambda_k|$, entonces $\alpha(\theta) \leq 0$ y $|z - \lambda_j| \leq \varepsilon$; por lo tanto, $z \in \mathcal{D}(\lambda_j, \varepsilon)$. Por otro lado, si $|z - \lambda_k|z \leq |z - \lambda_j|$, entonces $\alpha(\theta) \geq 0$ y $|z - \lambda_k| \leq \varepsilon$; por lo tanto, $z \in \mathcal{D}(\lambda_k, \varepsilon)$. □

Para finalizar esta subsección, consideramos apropiado enunciar la siguiente conjetura

Conjetura 5.3.13. *Si $A \in \mathbb{C}^{n \times n}$ es una matriz normal, entonces para $\varepsilon \geq 0$ y $k = 1, 2, \dots, n$ se tiene que el pseudoespectro algebraico ordinario de orden k de nivel ε , $\Lambda_{\varepsilon,k}^a(A)$, viene dado por una unión de discos cerrados, o el conjunto vacío.*

5.3.2. Pseudoespectros geométricos de cualquier orden

Vamos a analizar los pseudoespectros geométricos de cualquier orden, para matrices normales. Comenzamos con un ejemplo ilustrativo.

Ejemplo 5.3.1. Sea $A_0 \in \mathbb{C}^{4 \times 4}$ la matriz

$$\text{diag} \left(2, -1 + \sqrt{3}i, -1 - \sqrt{3}i, \frac{5}{2} + \frac{5}{2}\sqrt{3}i \right).$$

En la figura 5.2 se muestran las fronteras de $\Lambda_{\varepsilon,1}^g(A_0)$ en azul, $\Lambda_{\varepsilon,2}^g(A_0)$ en rojo y $\Lambda_{\varepsilon,3}^g(A_0)$ en verde, para $\varepsilon = 2, 5$.

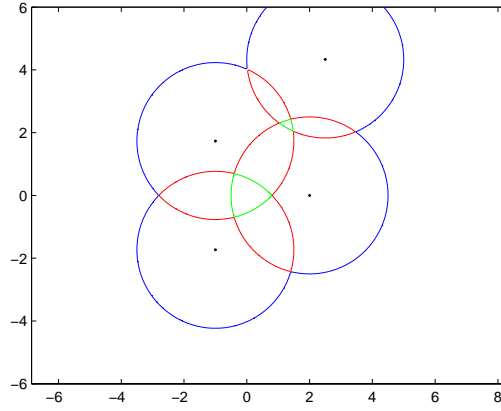


Figura 5.2: Fronteras del ε -pseudoespectro geométrico.

El resultado a demostrar, en esta subsección, es el siguiente.

Teorema 5.3.14. Sea $A \in \mathbb{C}^{n \times n}$ matriz normal. Supongamos que $\text{Eig}(A) = (\lambda_1, \lambda_2, \dots, \lambda_n)$. Dado $\varepsilon \geq 0$, para $k = 1, 2, \dots, n$ se tiene

$$\Lambda_{\varepsilon,k}^g(A) = \bigcup_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \bigcap_{j=1}^k \mathcal{D}(\lambda_{i_j}, \varepsilon).$$

En particular, si $k = 2$,

$$\Lambda_{\varepsilon,2}^g(A) = \bigcup_{1 \leq i < j \leq n} (\mathcal{D}(\lambda_i, \varepsilon) \cap \mathcal{D}(\lambda_j, \varepsilon)).$$

DEMOSTRACIÓN.

Como A es normal, podemos asumir sin pérdida de generalidad que $A = \text{diag}(\lambda_1, \dots, \lambda_n)$. Por lo tanto, dado $z \in \mathbb{C}$, el sistema de valores singulares de $zI_n - A$ es

$$|z - \lambda_1|, |z - \lambda_2|, \dots, |z - \lambda_n|,$$

no necesariamente ordenados. Recordamos la caracterización de los pseudoespectros ordinarios geométricos de orden k y nivel ε , dada en (3.1.1)

$$\Lambda_{\varepsilon, k}^g(A) = \{z \in \mathbb{C} : \sigma_{n-k+1}(zI_n - A) \leq \varepsilon\}.$$

Si $z \in \Lambda_{\varepsilon, k}^g(A)$, se tiene que $\sigma_{n-k+1}(zI_n - A) \leq \varepsilon$, es decir, existen k términos de la n -tupla desordenada $\text{Eig}(A)$, denotados por λ_{i_j} , $j = 1, 2, \dots, k$ tales que

$$|z - \lambda_{i_j}| \leq \varepsilon.$$

Con lo cual, $z \in \mathcal{D}(\lambda_{i_j}, \varepsilon)$ para $j = 1, 2, \dots, k$. La implicación contraria se puede probar siguiendo un razonamiento inverso. \square

Capítulo 6

Pseudoespectros idénticos

6.1. Introducción

Si dos matrices cualesquiera $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$ satisfacen que para todo $\varepsilon > 0$

$$\Lambda_\varepsilon(A) = \Lambda_\varepsilon(B),$$

entonces los polinomios mínimos de A y B son iguales. Véase [31]. En este capítulo vamos a establecer que los pseudoespectros geométricos para valores pequeños de ε determinan la forma canónica de Jordan de A o, equivalentemente, determinan sus factores invariantes. Concretamente, vamos a probar el resultado siguiente.

Teorema 6.1.1. *Sean $A, B \in \mathbb{C}^{n \times n}$. Supongamos que para todo $z \in \mathbb{C}$ se satisface que $\sigma_j(zI_n - A) = \sigma_j(zI_n - B)$, $j \in \{1, \dots, n\}$. Entonces $A \approx B$.*

6.2. Órdenes de los valores singulares de una matriz característica como infinitésimos

Dado un $z_0 \in \mathbb{C}$, denotamos por $V'(z_0)$ un entorno perforado de centro z_0 ; por otro lado, denotamos por $B'(z_0, \delta)$ la bola perforada de centro z_0 y radio $\delta > 0$. Vamos a considerar el siguiente conjunto de funciones $\mathcal{F} := \{f \mid f : V'(z_0) \rightarrow \mathbb{R}\}$.

Definición 6.2.1. Sean $f, g \in \mathcal{F}$. Diremos que f y g son **equivalentes según Hardy** cuando $z \rightarrow z_0$, y representamos $f(z) \asymp g(z)$, ($z \rightarrow z_0$), si existen constantes δ, Δ, d positivas, tales que para cualquier $z \in B'(z_0, \delta)$ se satisface que

$$f(z) > 0, g(z) > 0 \text{ y } \delta \leq \frac{f(z)}{g(z)} \leq \Delta.$$

Diremos que $f \in \mathcal{F}$ es un **infinitésimo** cuando $z \rightarrow z_0$ si $\lim_{z \rightarrow z_0} f(z) = 0$. Si $f(z) \asymp |z - z_0|^k$, con $k \geq 1$ entero, diremos que f es un infinitésimo de orden k cuando $z \rightarrow z_0$. La relación \asymp es una relación de equivalencia, damos la demostración a continuación.

DEMOSTRACIÓN.

Sean $f, g, h \in \mathcal{F}$ tales que $f(z) > 0, g(z) > 0, h(z) > 0$.

Reflexiva: Sea $d > 0$ tal que para toda $z \in B'(z_0, d)$ se tiene que existen $0 < \delta \leq 1, \Delta \geq 1$ tales que

$$\delta \leq \frac{f(z)}{f(z)} = 1 \leq \Delta.$$

Entonces, $f(z) \asymp f(z)$ ($z \rightarrow z_0$).

Simétrica: Supongamos que existen δ, Δ, d positivas tales que

$$\delta \leq \frac{f(z)}{g(z)} \leq \Delta, \quad \forall z \in B'(z_0, d), \quad (6.2.1)$$

es decir, $f(z) \asymp g(z)$ ($z \rightarrow z_0$). Entonces, las siguientes desigualdades también son ciertas

$$\frac{1}{\delta} \geq \frac{g(z)}{f(z)} \geq \frac{1}{\Delta}, \quad \forall z \in B'(z_0, d). \quad (6.2.2)$$

Si definimos las siguientes cantidades $\delta' := \frac{1}{\Delta}$ y $\Delta' := \frac{1}{\delta}$, podemos reescribir (6.2.2) como

$$\delta' \leq \frac{g(z)}{f(z)} \leq \Delta', \quad \forall z \in B'(z_0, d).$$

Entonces, $g(z) \asymp f(z)$ ($z \rightarrow z_0$).

Transitiva: Si existen $\delta_1, \delta_2, \Delta_1, \Delta_2, d_1, d_2$ positivas tales que

$$\begin{aligned} \delta_1 \leq \frac{f(z)}{g(z)} \leq \Delta_1, \quad \forall z \in B'(z_0, d_1), \\ \delta_2 \leq \frac{g(z)}{h(z)} \leq \Delta_2, \quad \forall z \in B'(z_0, d_2), \end{aligned} \quad (6.2.3)$$

y definimos $d := \min\{d_1, d_2\}$, las desigualdades de (6.2.3) se satisfacen para todo $z \in B'(z_0, d)$. Por lo tanto,

$$\delta_1 \delta_2 \leq \frac{f(z)}{g(z)} \frac{g(z)}{h(z)} \leq \Delta_1 \Delta_2.$$

Definiendo $\delta := \delta_1 \delta_2$ y $\Delta := \Delta_1 \Delta_2$, podemos escribir las anteriores desigualdades como

$$\delta \leq \frac{f(z)}{h(z)} \leq \Delta, \quad \forall z \in B'(z_0, d).$$

De donde se deduce que $f(z) \asymp h(z)$ ($z \rightarrow z_0$). \square

Podemos establecer dos consecuencias de la equivalencia en sentido Hardy en las siguientes observaciones.

Observación 6.2.2. Sean $j, k \in \mathbb{Z}^+ \cup \{0\}$ tales que $|z - z_0|^j \asymp |z - z_0|^k$ ($z \rightarrow z_0$), entonces $j = k$.

DEMOSTRACIÓN.

Como $|z - z_0|^j \asymp |z - z_0|^k$ ($z \rightarrow z_0$), sabemos que existen δ, Δ, d positivas tales que

$$\delta \leq \frac{|z - z_0|^j}{|z - z_0|^k} \leq \Delta,$$

de donde tenemos que

$$\delta \leq |z - z_0|^{j-k} \leq \Delta. \quad (6.2.4)$$

La implicación $j = k$ se deduce inmediatamente de la acotación superior e inferior de $|z - z_0|^{j-k}$ en $B'(z_0, d)$. Supongamos que $j > k$, entonces para todo $z \in \{z_0 + re^{i\theta}, 0 < r < \delta^{\frac{1}{j-k}}\}$ se tendría que $|z - z_0|^{j-k} < \delta$, lo cual es absurdo por (6.2.4). Se puede hacer un razonamiento análogo para $j < k$, tomando para todo $z \in \{z_0 + Re^{i\theta}, R > \Delta^{\frac{1}{j-k}}\}$ y utilizando la otra desigualdad en (6.2.4). \square

Observación 6.2.3. Si tomamos $f, g \in \mathcal{F}$ tales que $f(z), g(z) > 0$, la relación de equivalencia $f(z) \sim g(z)$, cuando $z \rightarrow z_0$, significa que

$$\lim_{z \rightarrow z_0} \frac{f(z)}{g(z)} = 1.$$

Es evidente que $f(z) \sim g(z)$, cuando $z \rightarrow z_0$, implica que $f(z) \asymp g(z)$, cuando $z \rightarrow z_0$.

DEMOSTRACIÓN.

$$\lim_{z \rightarrow z_0} \frac{f(z)}{g(z)} = 1 \iff \forall \varepsilon > 0, \exists \delta(\varepsilon) > 0 : \forall z \in B'(z_0, \delta), \left| \frac{f(z)}{g(z)} - 1 \right| \leq \varepsilon,$$

donde la última desigualdad es equivalente a

$$1 - \varepsilon \leq \frac{f(z)}{g(z)} \leq 1 + \varepsilon.$$

Entonces, si definimos $\delta_\varepsilon := 1 - \varepsilon$, $\Delta_\varepsilon := 1 + \varepsilon$ y $d_\varepsilon := d(\varepsilon)$, se satisface que

$$\forall z \in B'(z_0, \delta), \quad \delta_\varepsilon \leq \frac{f(z)}{g(z)} \leq \Delta_\varepsilon.$$

\square

Para concluir esta sección, vamos a probar el siguiente teorema.

Teorema 6.2.4. Sea $J_k(z_0) \in \mathbb{C}^{k \times k}$ el bloque de Jordan asociado al valor $z_0 \in \mathbb{C}$. Entonces, cuando $z \rightarrow z_0$,

$$\sigma_j(zI_k - J_k(z_0)) \sim \begin{cases} 1, & j = 1, \dots, k-1, \\ |z - z_0|^k, & j = k. \end{cases}$$

DEMOSTRACIÓN.

Definamos la siguiente matriz

$$A(z) := zI_k - J_k(z_0) = \begin{pmatrix} z - z_0 & -1 & & \\ & z - z_0 & \ddots & \\ & & \ddots & -1 \\ & & & z - z_0 \end{pmatrix}. \quad (6.2.5)$$

De donde se tiene que,

$$\det(zI_k - J_k(z_0)) = (z - z_0)^k, \quad \Lambda(zI_k - J_k(z_0)) = \{z - z_0\}, \quad m(z - z_0, A(z)) = k.$$

Por otro lado,

$$A(z)A^*(z) = \begin{pmatrix} |z - z_0|^2 + 1 & \bar{z} - \bar{z}_0 & & \\ z - z_0 & |z - z_0|^2 + 1 & \bar{z} - \bar{z}_0 & \\ & \ddots & \ddots & \ddots \\ & & z - z_0 & |z - z_0|^2 + 1 & \bar{z} - \bar{z}_0 \\ & & & z - z_0 & |z - z_0|^2 \end{pmatrix}. \quad (6.2.6)$$

Los valores propios de (6.2.6) son los cuadrados de los valores singulares de (6.2.5), es decir,

$$\lambda_j(A(z)A^*(z)) = \sigma_j^2(zI_k - J_k(z_0)), \quad j = 1, \dots, k. \quad (6.2.7)$$

Tanto las matrices en (6.2.5) y (6.2.6), como las funciones en (6.2.7) son continuas respecto de z . Tomando el límite cuando $z \rightarrow z_0$ se tiene

$$A(z_0) = \begin{pmatrix} 0 & -1 & & \\ & 0 & \ddots & \\ & & \ddots & -1 \\ & & & 0 \end{pmatrix}.$$

Por lo tanto,

$$A(z_0)A^*(z_0) = \text{diag}(1, \dots, 1, 0) \in \mathbb{C}^{k \times k}.$$

Como $\Lambda(A(z_0)A^*(z_0)) = \{0, 1\}$, siendo las multiplicidades algebraicas

$$m(0, A(z_0)A^*(z_0)) = 1 \text{ y } m(1, A(z_0)A^*(z_0)) = k - 1,$$

se tiene que

$$\forall j \in \{1, \dots, k - 1\}, \quad \sigma_j(zI_k - J_k(z_0)) \sim 1, \text{ cuando } z \rightarrow z_0, \quad (6.2.8)$$

mientras que si $j = k$,

$$\sigma_k(zI_k - J_k(z_0)) \sim |z - z_0|^k, \text{ cuando } z \rightarrow z_0. \quad (6.2.9)$$

Vamos a comprobar (6.2.9), para ello tenemos en cuenta que

$$\prod_{j=1}^k \sigma_j^2(zI_k - J_k(z_0)) = \det((zI_k - J_k(z_0))(zI_k - J_k(z_0))^*) = |z - z_0|^{2k}. \quad (6.2.10)$$

Por otro lado,

$$\prod_{j=1}^k \sigma_j^2(zI_k - J_k(z_0)) = \sigma_1^2(zI_k - J_k(z_0)) \cdots \sigma_{k-1}^2(zI_k - J_k(z_0)) \sigma_k^2(zI_k - J_k(z_0)). \quad (6.2.11)$$

Igualando (6.2.10) y (6.2.11), y dividiendo en ambos miembros por $|z - z_0|^{2k}$, tenemos que

$$\frac{\sigma_1^2(zI_k - J_k(z_0)) \cdots \sigma_{k-1}^2(zI_k - J_k(z_0)) \sigma_k^2(zI_k - J_k(z_0))}{|z - z_0|^{2k}} = 1,$$

que podemos escribir como

$$\frac{\sigma_k^2(zI_k - J_k(z_0))}{|z - z_0|^{2k}} = \frac{1}{\sigma_1^2(zI_k - J_k(z_0)) \cdots \sigma_{k-1}^2(zI_k - J_k(z_0))}.$$

Si calculamos el límite cuando $z \rightarrow z_0$, tenemos

$$\lim_{z \rightarrow z_0} \frac{\sigma_k^2(zI_k - J_k(z_0))}{|z - z_0|^{2k}} = \lim_{z \rightarrow z_0} \frac{1}{\sigma_1^2(zI_k - J_k(z_0)) \cdots \sigma_{k-1}^2(zI_k - J_k(z_0))} = 1,$$

de donde se deduce que $\sigma_k^2(zI_k - J_k(z_0)) \sim |z - z_0|^{2k}$ o, equivalentemente, $\sigma_k(zI_k - J_k(z_0)) \sim |z - z_0|^k$. \square

6.3. Prueba del teorema principal

En esta sección vamos a presentar un resultado que permite determinar el número de bloques de Jordan de un tamaño dado correspondientes a un valor propio z_0 , a partir de la equivalencia en sentido Hardy entre los valores singulares de la matriz $zI_k - J_k(z_0)$ y las potencias $|z - z_0|^k$. Para ello, precisamos establecer unos lemas previos. El primero es una aplicación del Teorema 1, pág. 44 de [61]

Lema 6.3.1. Sean $M_1, M_2, M_3 \in \mathbb{C}^{n \times n}$. Se satisface, para $k = 1, \dots, n$, que

$$\sigma_n(M_1) \sigma_k(M_2) \sigma_n(M_3) \leq \sigma_k(M_1 M_2 M_3) \leq \sigma_1(M_1) \sigma_k(M_2) \sigma_1(M_3) \quad (6.3.12)$$

Lema 6.3.2. Sean $M \in \mathbb{C}^{n \times n}$, $P \in \text{GL}_n(\mathbb{C})$ y $z_0 \in \Lambda(M)$. Entonces, para $j = 1, \dots, n$,

$$\sigma_j(zI_n - P^{-1}MP) \asymp \sigma_j(zI_n - M), \quad z \rightarrow z_0. \quad (6.3.13)$$

DEMOSTRACIÓN.

Aplicando el Lema 6.3.1 a las matrices $M_1 := P^{-1}$, $M_2 := zI_n - M$ y $M_3 := P$, tenemos las siguientes desigualdades

$$\sigma_n(P^{-1})\sigma_j(zI_n - M)\sigma_n(P) \leq \sigma_j(P^{-1}(zI_n - M)P) \leq \sigma_1(P^{-1})\sigma_j(zI_n - M)\sigma_1(P). \quad (6.3.14)$$

Observamos que $\sigma_j(P^{-1}(zI_n - M)P) = \sigma_j(zI_n - P^{-1}MP)$ y, si $z \notin \Lambda(M) = \Lambda(P^{-1}MP)$, $\sigma_j(zI_n - P^{-1}MP) > 0$ y $\sigma_j(zI_n - M) > 0$, para todo $j \in \{1, \dots, n\}$. Finalmente, podemos definir las siguientes constantes no negativas $\delta := \sigma_n(P^{-1})\sigma_n(P)$ y $\Delta := \sigma_1(P^{-1})\sigma_1(P)$. Por lo tanto, si dividimos por $\sigma_j(zI_n - M)$ en (6.3.14) y tomamos en cuenta lo dicho en este párrafo, podemos asegurar que

$$\exists d > 0, \text{ t.q. } \delta \leq \frac{\sigma_j(P^{-1}(zI_n - M)P)}{\sigma_j(zI_n - M)} \leq \Delta, \quad \forall z \in B'(z_0, d).$$

Podemos expresar lo anterior mediante la notación de Hardy, de la siguiente forma

$$\sigma_j(zI_n - P^{-1}MP) \asymp \sigma_j(zI_n - M), \quad z \rightarrow z_0.$$

□

Lema 6.3.3. Sean $L \in \mathbb{C}^{q \times q}$ y $z_0 \in \mathbb{C}$ tal que $z_0 \notin \Lambda(L)$. Entonces, para $j = 1, \dots, q$,

$$\sigma_j(zI_q - L) \asymp 1, \quad z \rightarrow z_0.$$

DEMOSTRACIÓN.

Calculamos el límite de los valores singulares $\sigma_j(zI_q - L)$, $j \in \{1, \dots, q\}$, cuando $z \rightarrow z_0$; por ser funciones continuas en la variable z , tenemos

$$\lim_{z \rightarrow z_0} \sigma_j(zI_q - L) = \sigma_j(z_0I_q - L) \in \mathbb{R}^+.$$

Tomando en cuenta lo dicho en la Observación 6.2.3, podemos asegurar que

$$\sigma_j(zI_q - L) \asymp \sigma_j(z_0I_q - L), \quad z \rightarrow z_0.$$

Como $z_0 \notin \Lambda(L)$, podemos considerar que tiene asociado un bloque de Jordan de tamaño 0. Esto implica que, en la notación de Hardy,

$$\forall j \in \{1, \dots, q\}, \quad \sigma_j(zI_q - L) \asymp 1, \quad z \rightarrow z_0.$$

□

Una vez enunciados estos lemas, podemos probar el resultado siguiente.

Lema 6.3.4. Sean J la forma de Jordan de una matriz $M \in \mathbb{C}^{n \times n}$, $z_0 \in \mathbb{C}$ y $k \in \{1, \dots, n\}$. Entonces, el número de bloques de Jordan en J con tamaño $k \times k$ correspondientes al autovalor z_0 , contando repeticiones, es igual al número de subíndices $j \in \{1, \dots, n\}$ tales que $\sigma_j(zI_n - M) \asymp |z - z_0|^k$, cuando $z \rightarrow z_0$.

DEMOSTRACIÓN.

Si $z_0 \notin \Lambda(M)$, utilizando el Lema 6.3.3, $\sigma_j(zI_n - M) \asymp 1$, cuando $z \rightarrow z_0$. Con lo cual, no existe $j \in \{1, \dots, n\}$ tal que $\sigma_j(zI_n - M) \asymp |z - z_0|^k$, cuando $z \rightarrow z_0$.

Por el contrario, si $z_0 \in \Lambda(M)$, según el Lema 6.3.2 se tiene que para $j \in \{1, \dots, n\}$,

$$\sigma_j(zI_n - M) \asymp \sigma_j(zI_n - J), \quad \text{cuando } z \rightarrow z_0.$$

Sea $J = J_0 \oplus J_1$ la descomposición en suma directa de J , donde $J_0 \in \mathbb{C}^{n_0 \times n_0}$ es la suma directa de los t bloques de Jordan asociados a z_0 , con $z_0 \notin \Lambda(J_1)$. Cuando z está *suficientemente cerca* de z_0 , los t últimos valores singulares de $zI_n - J$ son justamente los valores singulares infinitesimales de $zI_{n_0} - J_0$. Así,

$$\forall j \in \{n - t + 1, \dots, n\}, \quad \lim_{z \rightarrow z_0} \sigma_j(zI_n - J) = 0.$$

El valor de $j \in \{n - t + 1, \dots, n\}$ tal que el orden del infinitésimo correspondiente a $\sigma_j(zI_n - J)$, cuando $z \rightarrow z_0$, es k es igual al número de bloques de Jordan en J de orden $k \times k$ asociados a z_0 . Para $j \in \{1, \dots, n - t\}$, se tiene $\sigma_j(zI_n - J) \asymp 1$, cuando $z \rightarrow z_0$. \square

Observación 6.3.5. En J_0 , para cada bloque de Jordan de orden $k \times k$ asociado al valor propio z_0 , se satisface

$$\forall j \in \{1, \dots, k - 1\}, \quad \sigma_j(zI_{n_0} - J_0) \asymp 1,$$

y $\sigma_k(zI_{n_0} - J_0) \asymp |z - z_0|^k$. Por otro lado, en J_1 , para cada bloque de Jordan de orden $k \times k$ no asociado a z_0 , se tiene

$$\forall j \in \{1, \dots, k\}, \quad \sigma_j(zI_{n_1} - J_1) \asymp 1.$$

A continuación probaremos el Teorema 6.1.1.

DEMOSTRACIÓN.

Primero observaremos que ambas matrices tienen el mismo espectro. En efecto, una caracterización de los valores propios de una matriz $M \in \mathbb{C}^{n \times n}$ a través de sus valores singulares es la siguiente:

$$z_0 \in \Lambda(M) \iff \sigma_n(z_0 I_n - M) = 0. \quad (6.3.15)$$

Como por hipótesis se satisface que $\sigma_n(zI_n - A) = \sigma_n(zI_n - B)$, $\forall z \in \mathbb{C}$, se tiene que $\Lambda(A) = \Lambda(B) = \{\lambda_1, \dots, \lambda_p\}$.

Por otro lado, como $\sigma_j(zI_n - A) = \sigma_j(zI_n - B)$ para cualquier $z \in \mathbb{C}$ y $j \in \{1, \dots, n\}$, se tiene que para cada $k \in \{1, \dots, n\}$ y $\lambda_i \in \Lambda(A)$ el valor de $j \in \{1, \dots, n\}$ tal que $\sigma_j(zI_n - A) \asymp |z - z_0|^k$, cuando $z \rightarrow \lambda_i$, es igual al valor de $j \in \{1, \dots, n\}$ tal que $\sigma_j(zI_n - B) \asymp |z - z_0|^k$, cuando $z \rightarrow \lambda_i$. Así, por el Lema (6.3.4), el número de bloques de Jordan de tamaño $k \times k$ asociados a λ_i en las formas de Jordan de A y B , respectivamente, es el mismo. Como eso se verifica para todo $\lambda_i \in \Lambda(A) = \Lambda(B)$, deducimos que $A \approx B$. \square

En la prueba del teorema principal de este capítulo utilizaremos los factores invariantes, los divisores elementales y las particiones asociadas a cada valor propio $\lambda_i, i = 1, \dots, p$, que conforman la característica de Segre de una matriz $M \in \mathbb{C}^{n \times n}$. Vamos a recordar la relación existente entre ellos. Sean los divisores elementales asociados a un valor propio λ_i

$$(\lambda - \lambda_i)^{n_{i1}}, (\lambda - \lambda_i)^{n_{i2}}, \dots, (\lambda - \lambda_i)^{n_{it_i}},$$

siendo $n_{i1} \geq n_{i2} \geq \dots \geq n_{it_i} > 0$. A partir de esa sucesión de exponentes podemos construir la partición de Segre correspondiente a cada valor propio λ_i ,

$$s(\lambda_i, M) = (n_{i1}, n_{i2}, \dots, n_{it_i}, 0, \dots).$$

Utilizaremos la siguiente notación para facilitar la escritura de las expresiones.

$$s(\lambda_i, M) = (s_1(\lambda_i, M), s_2(\lambda_i, M), \dots, s_{t_i}(\lambda_i, M), \dots, s_n(\lambda_i, M)).$$

Los factores invariantes, contruidos a partir de los divisores elementales, se pueden expresar del modo siguiente:

$$\begin{aligned} f_n(\lambda) &= (\lambda - \lambda_1)^{s_1(\lambda_1, M)} \dots (\lambda - \lambda_h)^{s_1(\lambda_h, M)} \dots (\lambda - \lambda_p)^{s_1(\lambda_p, M)} \\ f_{n-1}(\lambda) &= (\lambda - \lambda_1)^{s_2(\lambda_1, M)} \dots (\lambda - \lambda_h)^{s_2(\lambda_h, M)} \dots (\lambda - \lambda_p)^{s_2(\lambda_p, M)} \\ &\vdots \\ f_j(\lambda) &= (\lambda - \lambda_1)^{s_{n-j+1}(\lambda_1, M)} \dots (\lambda - \lambda_h)^{s_{n-j+1}(\lambda_h, M)} \dots (\lambda - \lambda_p)^{s_{n-j+1}(\lambda_p, M)} \\ &\vdots \\ f_1(\lambda) &= (\lambda - \lambda_1)^{s_n(\lambda_1, M)} \dots (\lambda - \lambda_h)^{s_n(\lambda_h, M)} \dots (\lambda - \lambda_p)^{s_n(\lambda_p, M)}. \end{aligned}$$

Teorema 6.3.6. Sean $A \in \mathbb{C}^{n \times n}$ y $B \in \mathbb{C}^{m \times m}$. Supongamos que $n \geq m$ y que

$$g_i(\lambda) \mid g_{i+1}(\lambda) \mid \dots \mid g_{m-1}(\lambda) \mid g_m(\lambda)$$

sean los factores invariantes no triviales de B . Supongamos que para todo $z \in \mathbb{C}$ y $k = 1, \dots, m - i + 1$ se satisface que

$$\sigma_{n-k+1}(zI_n - A) = \sigma_{m-k+1}(zI_n - B). \quad (6.3.16)$$

Entonces los últimos $m - i + 1$ factores invariantes de A ,

$$f_{n-m+i}(\lambda) \mid f_{n-m+i+1}(\lambda) \mid \dots \mid f_{n-1}(\lambda) \mid f_n(\lambda) \quad (6.3.17)$$

son no triviales y satisfacen que

$$f_n(\lambda) = g_m(\lambda), \quad f_{n-1}(\lambda) = g_{m-1}(\lambda), \dots, f_{n-m+i}(\lambda) = g_i(\lambda). \quad (6.3.18)$$

DEMOSTRACIÓN.

Primero veremos que ambas matrices tienen el mismo espectro. En efecto, por hipótesis se tiene que

$$\Lambda_{\varepsilon,1}^g(A) = \{z \in \mathbb{C} \mid \sigma_n(zI_n - A) \leq \varepsilon\} = \{z \in \mathbb{C} \mid \sigma_m(zI_n - B) \leq \varepsilon\} = \Lambda_{\varepsilon,1}^g(B)$$

y, por otro lado, tenemos que

$$\lim_{\varepsilon \rightarrow 0^+} \Lambda_{\varepsilon,1}^g(A) = \Lambda(A), \quad \lim_{\varepsilon \rightarrow 0^+} \Lambda_{\varepsilon,1}^g(B) = \Lambda(B),$$

respecto de la distancia Hausdorff. Por lo tanto, podemos concluir que $\Lambda(A) = \Lambda(B) = \{\lambda_1, \dots, \lambda_p\}$.

A continuación, tenemos que probar las igualdades entre los factores invariantes de las matrices A y B , expresadas en (6.3.18), lo cual es equivalente a probar que las características de Segre correspondientes a cada valor propio en las matrices A y B son las mismas para $h = 1, \dots, p$ y $k = 1, \dots, m - i + 1$, es decir,

$$s_k(\lambda_h, A) = s_k(\lambda_h, B). \quad (6.3.19)$$

Por el Lema 6.3.4, si $z \rightarrow \lambda_h$, $h = 1 \dots, p$, se tiene que

$$\sigma_{n-k+1}(zI_n - A) \asymp |z - \lambda_h|^{s_k(\lambda_h, A)}, \quad \sigma_{m-k+1}(zI_m - B) \asymp |z - \lambda_h|^{s_k(\lambda_h, B)}, \quad (6.3.20)$$

para $k = 1, \dots, m - i + 1$. De donde se puede deducir que, si $z \rightarrow \lambda_h$,

$$|z - \lambda_h|^{s_k(\lambda_h, A)} \asymp |z - \lambda_h|^{s_k(\lambda_h, B)},$$

pero por la Observación 6.2.2 se tiene que $s_k(\lambda_h, A) = s_k(\lambda_h, B)$, $k = 1 \dots, m - i + 1$. Con lo cual, se tiene (6.3.18). □

Capítulo 7

Fronteras de los pseudoespectros

7.1. Resultados principales

La *frontera* del pseudoespectro ordinario de una matriz $A \in \mathbb{C}^{n \times n}$ de nivel ε está contenida en la frontera del pseudoespectro estricto, Véase [3, p. 280, Observación 3.2]. En este capítulo demostraremos que, en general, no son iguales. Un problema equivalente es determinar si la función $h_1(z) = \sigma_n(zI_n - A)$ puede tener máximos locales. La respuesta es afirmativa, de hecho probaremos que $z_0 \in \partial\Lambda'_\varepsilon(A) \setminus \partial\Lambda_\varepsilon(A)$ si y sólo si h_1 tiene un máximo local en z_0 .

Una pregunta que aparece de forma natural es ¿Cómo puede ser el conjunto $\partial\Lambda'_\varepsilon(A) \setminus \partial\Lambda_\varepsilon(A)$? Para tratar de responder a esta cuestión, introduciremos la noción de conjunto semialgebraico. Citemos que, tanto el pseudoespectro ordinario de una matriz A de nivel ε como su frontera son conjuntos semialgebraicos, véase [44]. Probaremos que dicha propiedad también se cumple para el pseudoespectro estricto. La Figura 7.1 que representa pseudoespectros hendidos de una matriz 3×3 que sólo tiene un valor propio prelude la aparición de puntos aislados y puntos crunodos en la frontera de algunos pseudoespectros.

Recapitulando, los teoremas principales de este capítulo son los siguientes.

Teorema 7.1.1. *Para todo $\varepsilon > 0$, los conjuntos $\Lambda'_\varepsilon(A)$ y $\partial\Lambda'_\varepsilon(A)$ son semialgebraicos, de \mathbb{C} o de \mathbb{R}^2 .*

Teorema 7.1.2. *Dado un conjunto finito de puntos \mathcal{P} de \mathbb{C} , se satisface que:*

1. *existen un $\varepsilon_0 > 0$ y una matriz $A_0 \in \mathbb{C}^{n \times n}$ tales que*

$$\mathcal{P} \subset \partial\Lambda'_{\varepsilon_0}(A_0) \setminus \partial\Lambda_{\varepsilon_0}(A_0),$$

siendo cada $z_0 \in \mathcal{P}$ un punto aislado de $\partial\Lambda'_{\varepsilon_0}(A_0)$.

2. *Para cada $z_0 \in \mathcal{P}$, la función $z \mapsto \sigma_n(zI_n - A_0)$ tiene un máximo local estricto en z_0 .*

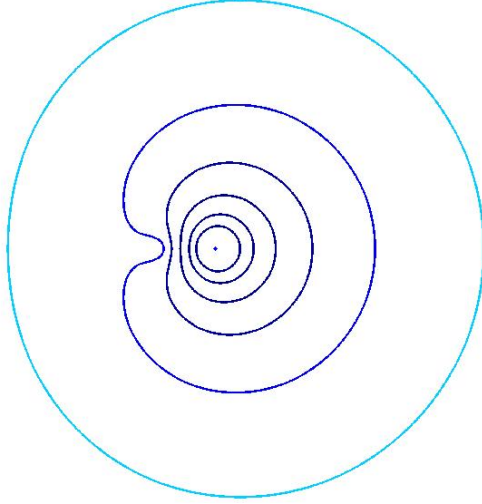


Figura 7.1: Pseudoespectros hendidos.

Por simplificación, vamos a introducir unas notaciones. Dada $A \in \mathbb{C}^{n \times n}$ y $z \in \mathbb{C}$, denotaremos por

$$\mathbf{A}(z) := zI_n - A.$$

Si identificamos $z = x + yi$, con $(x, y) \in \mathbb{R}^2$, la notación anterior puede reescribirse de la forma

$$\mathbf{A}(x, y) := (x + yi)I_n - A.$$

Observemos que con estas notaciones, se tiene $h_1(z) = \sigma_n(\mathbf{A}(z))$ y $h_1(x, y) := \sigma_n(\mathbf{A}(x, y))$.

7.2. Conjuntos semialgebraicos. Funciones semialgebraicas

En esta sección vamos a presentar los conceptos de conjunto semialgebraico y función semialgebraica, (véanse [13] y [44]).

Sea E un conjunto. Denotemos por $\mathcal{P}(E)$ el conjunto de las partes o subconjuntos de E . Se llama **álgebra de Boole de subconjuntos** de E a todo conjunto no vacío $\mathcal{B} \subset \mathcal{P}(E)$ tal que

- Si $A, B \in \mathcal{B}$, entonces $A \cup B \in \mathcal{B}$;
- Si $A \in \mathcal{B}$, entonces $A^c \in \mathcal{B}$. Recordemos que A^c denota el conjunto *complementario* de A respecto de E .

Dado un conjunto $\mathcal{G} \subset \mathcal{P}(E)$, se llama **álgebra de Boole generada por** \mathcal{G} a la intersección de todas las álgebras de Boole que contienen a \mathcal{G} .

Definición 7.2.1. Un subconjunto semialgebraico de \mathbb{R}^n es un elemento del álgebra de Boole de subconjuntos de \mathbb{R}^n que está engendrada por los conjuntos

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) > 0\}, \quad p \in \mathbb{R}[X_1, \dots, X_n],$$

donde $\mathbb{R}[X_1, \dots, X_n]$ denota el anillo de polinomios reales en las variables X_1, \dots, X_n .

Se sigue inmediatamente de la definición que los conjuntos

$$\begin{aligned} &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) < 0\}, \\ &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) > 0\}, \\ &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) \geq 0\}, \\ &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) \leq 0\}, \\ &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) = 0\}, \\ &\{(x_1, \dots, x_n) \in \mathbb{R}^n : p(x_1, \dots, x_n) \neq 0\}, \end{aligned}$$

con $p \in \mathbb{R}[X_1, \dots, X_n]$ son semialgebraicos. Se ve fácilmente que un subconjunto \mathcal{S} de \mathbb{R}^n es semialgebraico si y sólo si \mathcal{S} es una unión finita de conjuntos de la forma

$$\{x \in \mathbb{R}^n : p_1(x) = 0, \dots, p_r(x) = 0, q_1(x) > 0, \dots, q_s(x) > 0\}, \quad (7.2.1)$$

con $p_1, \dots, p_r, q_1, \dots, q_s \in \mathbb{R}[X_1, \dots, X_n]$.

Observación 7.2.2. El conjunto (7.2.1) es unión finita de conjuntos de la forma

$$\{x \in \mathbb{R}^n : p(x) = 0, q_1(x) > 0, \dots, q_k(x) > 0\}$$

donde $p \in \mathbb{R}[X_1, \dots, X_n]$ es irreducible. Si $\mathcal{S} \subset \mathbb{R}^n$ y $\mathcal{M} \subset \mathbb{R}^m$ son semialgebraicos, entonces $\mathcal{S} \times \mathcal{M} \subset \mathbb{R}^n \times \mathbb{R}^m$ es semialgebraico.

Observación 7.2.3. La Definición 7.2.1 se puede extender a subconjuntos \mathbb{C}^n , identificando \mathbb{C}^n con \mathbb{R}^{2n} . Así, dado un conjunto $S \subset \mathbb{C}^n$, identificamos cada elemento $s \in S$ por el par (w, ℓ) , con $w, \ell \in \mathbb{R}^n$, siendo $s = w + \ell i$. Por lo tanto, podemos suponer que S es un subconjunto de \mathbb{R}^{2n} . Así, por la Definición 7.2.1, diremos que S es semialgebraico, de \mathbb{C}^n o \mathbb{R}^{2n} indistintamente, si existen polinomios $p_j, q_l^j \in \mathbb{R}[W_1, \dots, W_n, L_1, \dots, L_n]$, tales que

$$S = \bigcup_{j=1}^k \{(w, \ell) \in \mathbb{R}^n \times \mathbb{R}^n : p_j(w, \ell) = 0, q_l^j(w, \ell) > 0, l = 1, \dots, r_j\}.$$

Definición 7.2.4. Dados dos conjuntos semialgebraicos no vacíos $\mathcal{S}_i \subset \mathbb{R}^{n_i}$, $i = 1, 2$, una función $f : \mathcal{S}_1 \rightarrow \mathcal{S}_2$ se dice **semialgebraica** si su grafo es un conjunto semialgebraico de $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \equiv \mathbb{R}^{n_1+n_2}$.

Definición 7.2.5. Sea $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Sea \mathcal{S} un conjunto semialgebraico de \mathbb{R}^n . Se dice que una función $f : \mathcal{S} \rightarrow \overline{\mathbb{R}}$ es **semialgebraica** si los conjuntos $f^{-1}(-\infty)$ y $f^{-1}(\infty)$ son semialgebraicos y la restricción $f|_{\mathcal{S} \setminus f^{-1}(\{-\infty, \infty\})}$ es una función semialgebraica.

Definición 7.2.6. Un conjunto $\mathcal{M} \subset \mathbb{R}^2$ es una **subvariedad real analítica unidimensional** si para todo $p \in \mathcal{M}$ existen un abierto $U \subset \mathbb{R}$ y una función real analítica $f : U \rightarrow \mathbb{R}^2$ que transforma abiertos de U en abiertos (con la topología relativa) de \mathcal{M} y además

$$p \in f(U) \quad \text{y} \quad \text{rg}[Df(u)] = 1 \quad \forall u \in U,$$

donde $Df(u)$ denota la diferencial de f en u .

Por tanto, de la definición se sigue que \mathcal{M} es localmente parametrizable. Una parametrización viene dada por las dos componentes de la función f . En el lema siguiente se presentan algunas propiedades de los conjuntos semialgebraicos y de las funciones semialgebraicas.

Lema 7.2.7. [Bochnak et al., Karow]

- (a) *La composición de funciones semialgebraicas es semialgebraica.*
- (b) *Sea $f : \mathcal{S}_1 \rightarrow \mathcal{S}_2$ semialgebraica. Supongamos que $\mathcal{A} \subset \mathcal{S}_1$ y $\mathcal{B} \subset \mathcal{S}_2$ son conjuntos semialgebraicos. Entonces $f(\mathcal{A})$ y $f^{-1}(\mathcal{B})$ son conjuntos semialgebraicos.*
- (c) *Para $k = 1, 2, \dots$ la función $\sigma_k : \mathbb{C}^{n \times n} \rightarrow [0, \infty)$ es semialgebraica. Asimismo, el conjunto $\Lambda_\varepsilon(A)$ es semialgebraico.*
- (d) *La frontera de un conjunto semialgebraico es un conjunto semialgebraico.*
- (e) *Si $\mathcal{S} \subset \mathbb{R}^2$ es un conjunto semialgebraico no vacío, su frontera $\partial\mathcal{S}$ puede descomponerse en la unión disjunta*

$$\partial\mathcal{S} = \mathcal{M} \dot{\cup} \mathcal{Q},$$

donde \mathcal{M} es el conjunto vacío o una subvariedad analítica de \mathbb{R}^2 , unidimensional, formada por un número finito de componentes conexas y \mathcal{Q} es un conjunto finito o el conjunto vacío.

- (f) *Sea $f : \mathcal{S} \times \mathcal{M} \rightarrow \mathbb{R}$ una función semialgebraica. Entonces las funciones*

$$f_1, f_2 : \mathcal{S} \rightarrow \overline{\mathbb{R}}, \quad f_1(x) := \inf_{y \in \mathcal{M}} f(x, y), \quad f_2(x) := \sup_{y \in \mathcal{M}} f(x, y)$$

son semialgebraicas.

DEMOSTRACIÓN.

Las partes (a) y (b) están en Proposición 2.2.6 y 2.2.7 de [13]. Los apartados (c), (d) y (e) están en Corolario 3.1.22, Proposición 3.1.8, Proposición 3.1.9 y Corolario 3.1.10 de [44].

□

7.3. Resultados auxiliares

Comenzamos esta sección introduciendo algunas propiedades de las fronteras de los pseudoespectros.

Proposición 7.3.1. *Sea $A \in \mathbb{C}^{n \times n}$ y $\varepsilon > 0$. Entonces*

$$(a) \partial\Lambda'_\varepsilon(A) = \{z \in \mathbb{C} : h_1(z) = \varepsilon\}.$$

$$(b) \partial\Lambda_\varepsilon(A) \subset \partial\Lambda'_\varepsilon(A).$$

DEMOSTRACIÓN.

(a) Sea $z_0 \in \partial\Lambda'_\varepsilon(A)$, entonces aplicando (2.1.3) y la Proposición 2.2.4 (f), resulta

$$\left. \begin{array}{l} z_0 \in \overline{\Lambda'_\varepsilon(A)} \subset \Lambda_\varepsilon(A) \Rightarrow h_1(z_0) \leq \varepsilon, \\ z_0 \in \overline{(\Lambda'_\varepsilon(A))^c} \Rightarrow h_1(z_0) \geq \varepsilon \end{array} \right\} \Rightarrow h_1(z_0) = \varepsilon.$$

Recíprocamente, sea z_0 tal que $h_1(z_0) = \varepsilon$. Si $z_0 \notin \partial\Lambda'_\varepsilon(A)$, entonces $z_0 \in (\Lambda'_\varepsilon(A))^c$. Luego existe un $\delta > 0$ tal que el disco $\mathcal{D} = \mathcal{D}(z_0, \delta)$ satisface $\mathcal{D} \cap \Lambda'_\varepsilon(A) = \emptyset$ y, por lo tanto, $\mathcal{D} \subset (\Lambda'_\varepsilon(A))^c$. En estas condiciones, $\forall z \in \mathcal{D}$ se tiene que $h_1(z) \geq \varepsilon = h_1(z_0)$. Luego, por la Proposición 2.2.4 (e), z_0 es un mínimo local de h_1 , es decir, $z_0 \in \Lambda(A)$ y $h_1(z_0) = 0$, lo que es absurdo.

(b) Observemos en primer lugar que para todo $X \subset \mathbb{C}$ se cumple $\overline{X} = \overset{\circ}{X} \cup \partial X$. Ahora, por la Proposición 2.2.4 (a)(f), como $\Lambda_\varepsilon(A)$ es cerrado y $\overline{\Lambda'_\varepsilon(A)} = \Lambda_\varepsilon(A)$, resulta que

$$\overline{\Lambda_\varepsilon(A)} \cup \partial\Lambda_\varepsilon(A) = \overline{\Lambda_\varepsilon(A)} = \Lambda_\varepsilon(A),$$

$$\overline{\Lambda'_\varepsilon(A)} \cup \partial\Lambda'_\varepsilon(A) = \overline{\Lambda'_\varepsilon(A)} = \Lambda_\varepsilon(A).$$

Por lo tanto,

$$\overline{\Lambda_\varepsilon(A)} \cup \partial\Lambda_\varepsilon(A) = \overline{\Lambda'_\varepsilon(A)} \cup \partial\Lambda'_\varepsilon(A).$$

Luego, por la Proposición 2.2.4 (b), tenemos que $\Lambda'_\varepsilon(A) \subset \Lambda_\varepsilon(A)$, entonces

$$\overline{\Lambda'_\varepsilon(A)} \subset \overline{\Lambda_\varepsilon(A)} \Rightarrow \partial\Lambda_\varepsilon(A) \subset \partial\Lambda'_\varepsilon(A).$$

□

Como consecuencia, tenemos el corolario siguiente:

Corolario 7.3.2. *Son equivalentes:*

1. $\partial\Lambda_\varepsilon(A) \subsetneq \partial\Lambda'_\varepsilon(A)$.
2. $\Lambda'_\varepsilon(A) \subsetneq \overline{\Lambda'_\varepsilon(A)}$.

El siguiente resultado es clave para demostrar el Teorema 7.1.2.

Teorema 7.3.3. *Dado ε_0 , sea $z_0 \in \partial\Lambda'_{\varepsilon_0}(A)$. Entonces*

1. z_0 es un punto aislado de $\partial\Lambda'_{\varepsilon_0}(A)$ si y sólo si la función h_1 alcanza un máximo local estricto en z_0 .
2. Si z_0 es un punto aislado de $\partial\Lambda'_{\varepsilon_0}(A)$, entonces $z_0 \notin \partial\Lambda_{\varepsilon_0}(A)$. Por lo tanto, si $\partial\Lambda'_{\varepsilon_0}(A)$ tiene puntos aislados, se cumple que $\partial\Lambda_{\varepsilon_0}(A) \subsetneq \partial\Lambda'_{\varepsilon_0}(A)$.
3. $z_0 \in \partial\Lambda'_{\varepsilon_0}(A) \setminus \partial\Lambda_{\varepsilon_0}(A)$ si y sólo si la función h_1 alcanza un máximo local en z_0 .

DEMOSTRACIÓN.

(1) Supongamos que z_0 es un punto aislado de $\partial\Lambda'_{\varepsilon_0}(A)$. Entonces existe un disco abierto $D := D(z_0, \delta)$, con $\delta > 0$, de modo que $\forall z \in \mathcal{B}' = D \setminus \{z_0\}$ se cumple $h_1(z) \neq \varepsilon_0$. Veremos que $h_1(z) > \varepsilon_0, \forall z \in \mathcal{B}'$, o $h_1(z) < \varepsilon_0, \forall z \in \mathcal{B}'$.

Supongamos lo contrario. Sean $z_1, z_2 \in \mathcal{B}'$ tal que $h_1(z_1) < \varepsilon_0 < h_1(z_2)$. Sea $\gamma(t)$ una curva continua en $[0, 1]$, con $\gamma(0) = z_1, \gamma(1) = z_2$, tal que $\gamma(t) \in \mathcal{B}', \forall t \in [0, 1]$. Como $h_1(\gamma(t))$ es continua en $[0, 1]$, por el Teorema de Bolzano existe un $z_3 \in \mathcal{B}'$ tal que $h_1(z_3) = \varepsilon_0$. Lo cual es absurdo.

Si fuera $h_1(z) < \varepsilon_0, \forall z \in \mathcal{B}'$, tendríamos que z_0 sería un mínimo local de h_1 . Por lo tanto, por la Proposición 2.2.4 (e), z_0 sería un valor propio de A , es decir, $\varepsilon_0 = 0$, lo cual es absurdo. Esto implica que $h_1(z) < \varepsilon_0, \forall z \in \mathcal{B}'$, es decir, h_1 tiene un máximo estricto local en z_0 . El recíproco es inmediato.

(2) Como z_0 es aislado en $\partial\Lambda'_{\varepsilon_0}(A)$, utilizando el apartado (1), se tiene que existe un $\delta > 0$ de manera que $\forall z \in \mathcal{D}'$ se cumple $h_1(z) \leq \varepsilon_0$. Luego $D \cap (\Lambda_{\varepsilon_0}(A))^c = \emptyset$. Por lo tanto, $z_0 \notin \partial\Lambda_{\varepsilon_0}(A)$.

(3) Tomamos la notación introducida en la demostración de (1). Para probar este apartado es suficiente notar que

$$z_0 \in \partial\Lambda'_{\varepsilon_0}(A) \setminus \partial\Lambda_{\varepsilon_0}(A) \text{ si y sólo si } \begin{cases} h_1(z_0) = \varepsilon_0, \\ \forall z \in D, h_1(z) \leq \varepsilon_0. \end{cases}$$

□

Lema 7.3.4. *Dada $A \in \mathbb{C}^{n \times n}$, para $(x, y, \varepsilon) \in \mathbb{R}^3$, con $\varepsilon \geq 0$, consideramos*

$$q(x, y, \varepsilon) := \det \begin{pmatrix} \varepsilon I_n & \mathbf{A}(x, y) \\ \mathbf{A}^*(x, y) & \varepsilon I_n \end{pmatrix}$$

con $\mathbf{A}(x, y)$ definida en (2.1.2). Sean $\varepsilon_0 > 0$ y (x_0, y_0) un punto aislado de la curva $q(x, y, \varepsilon_0) = 0$. Entonces, ε_0 es un valor singular de $\mathbf{A}(x_0, y_0)$. Además, si $\varepsilon_0 = h_1(x_0, y_0)$, entonces (x_0, y_0) es un punto aislado de $\partial\Lambda'_{\varepsilon_0}(A)$.

DEMOSTRACIÓN.

Por el Lema 1.2.17 se tiene que ε_0 es un valor singular de $\mathbf{A}(x_0, y_0)$ si y sólo si $q(x_0, y_0, \varepsilon_0) = 0$. Supongamos que $h_1(x_0, y_0) = \varepsilon_0$. Entonces, por la Proposición 7.3.1 (a), se tiene que

$$(x_0, y_0) \in \partial\Lambda'_{\varepsilon_0}(A) \subset \{(x, y) \in \mathbb{R}^2 : q(x, y, \varepsilon_0) = 0\}.$$

Por lo tanto, si (x_0, y_0) es un punto aislado de la curva $q(x, y, \varepsilon_0) = 0$, también lo es de $\partial\Lambda'_{\varepsilon_0}(A)$. \square

En la Figura 7.3 podemos ver un ejemplo en el que $\partial\Lambda'_{\varepsilon_0}(A) \neq \{(x, y) \in \mathbb{R}^2 : q(x, y, \varepsilon_0) = 0\}$, mediante la matriz

$$A := \begin{pmatrix} 0 & 0,1 \\ 0 & 1 \end{pmatrix}, \text{ con } \varepsilon_0 := 0,6.$$

Estos hechos son ilustrados por la Figura 7.2 que representa las superficies $z = \sigma_1(x, y)$ y $z = \sigma_2(x, y)$, siendo $\sigma_1(x, y) := \sigma_1((x + yi)I_2 - A)$ y $\sigma_2(x, y) := \sigma_2((x + yi)I_2 - A)$. La Figura 7.3 muestra las curvas de nivel producidas por la intersección de estas dos superficies por el plano horizontal $z = 0,6$.

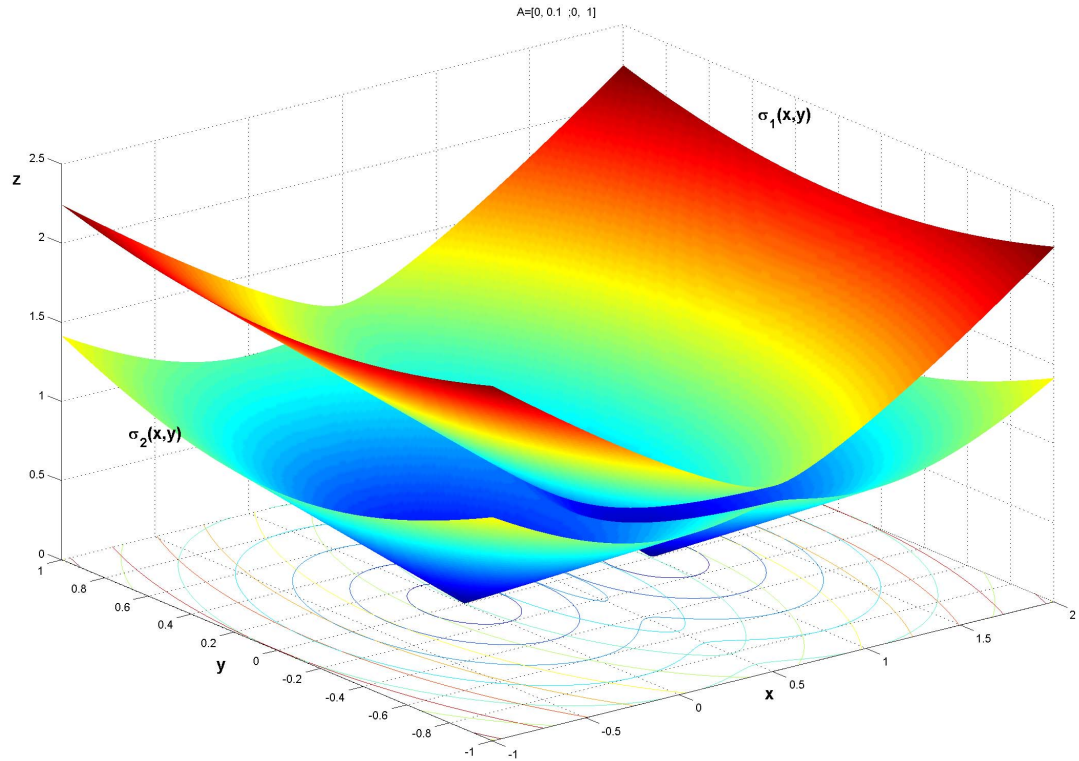


Figura 7.2: Superficies $z = \sigma_1(x, y), z = \sigma_2(x, y) = h_1(x, y)$.

Recapitulando, llamamos a este fenómeno *lo que el lema de Wielandt oculta*, pues no siempre que $q(x, y, \varepsilon) = 0$ significa que el punto (x, y) esté necesariamente contenido en la frontera del ε -pseudoespectro estricto $\Lambda'_\varepsilon(A)$ de la matriz

A. En la prueba del Teorema 7.1.2, utilizaremos las técnicas de Geometría Algebraica descritas en el Apéndice A. En concreto, el apartado II del Teorema A.5.1.

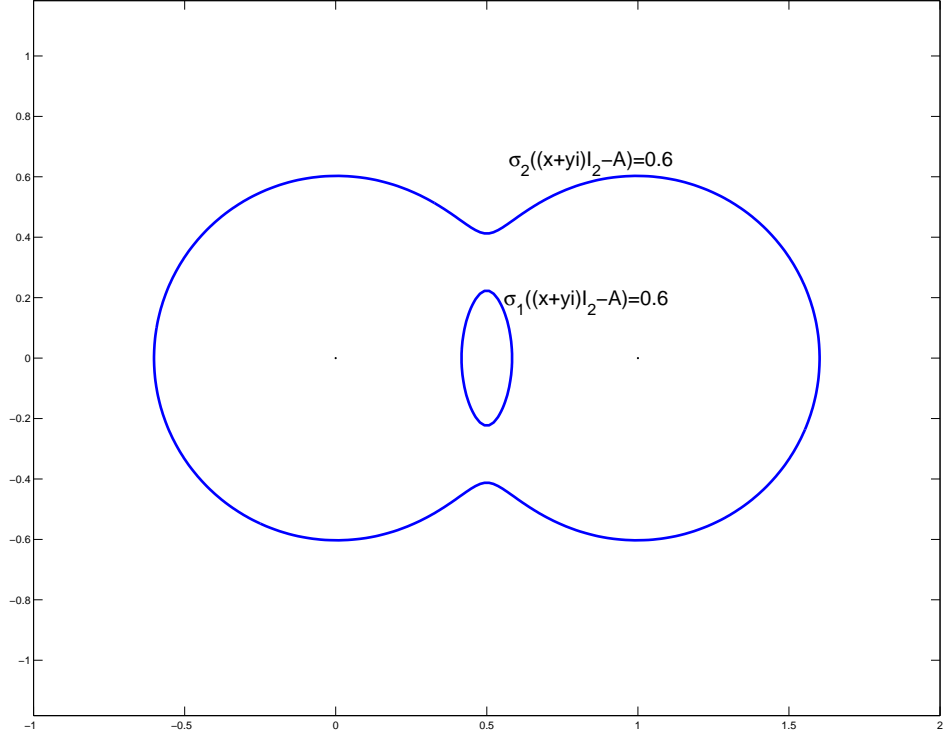


Figura 7.3: $\partial\Lambda'_{\varepsilon_0}(A) \neq \{(x, y) \in \mathbb{R}^2 : q(x, y, \varepsilon_0) = 0\}$.

7.4. Prueba de los Teoremas principales

Vamos a dar la prueba de los Teoremas 7.1.1 y 7.1.2.

7.4.1. Prueba del Teorema 7.1.1

Consideremos el diagrama siguiente

$$\begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{g} & \mathbb{C}^{n \times n} & \xrightarrow{\sigma_n} & [0, \infty) \\ (x, y) & \mapsto & (x + yi)I_n - A & \mapsto & \sigma_n((x + yi)I_n - A). \end{array}$$

Denotando por

$$A := (a_{ij} + b_{ij}i)_{i,j=1}^n, \quad a_{ij}, b_{ij} \in \mathbb{R},$$

obtenemos que el grafo de g es

$$\begin{aligned} \text{grafo}(g) = \{ & (x, y, x_{ij}, y_{ij}) : x_{ii} = x - a_{ii}, y_{ii} = y - b_{ii}, \\ & x_{ij} = -a_{ij}, y_{ij} = -b_{ij}, i, j = 1, \dots, n, j \neq i \}. \end{aligned}$$

Si consideramos el polinomio

$$P(X, Y, X_{ij}, Y_{ij}) := \sum_{i=1}^n ((X_{ii} - X + a_{ii})^2 + (Y_{ii} - Y + b_{ii})^2) + \sum_{\substack{i,j=1 \\ j \neq i}}^n ((X_{ij} + a_{ij})^2 + (Y_{ij} + b_{ij})^2),$$

se tiene

$$\text{grafo}(g) = \{(x, y, x_{ij}, y_{ij}) : P(x, y, x_{ij}, y_{ij}) = 0\}.$$

Por lo tanto, de la Definición 7.2.1 se deduce que $\text{grafo}(g)$ es un conjunto semialgebraico, de hecho es algebraico, y por la Definición 7.2.4, la función g es semialgebraica. Si aplicamos los apartados (c) y (a) del Lema 7.2.7, concluimos que

$$\begin{aligned} h_1 : \mathbb{R}^2 & \rightarrow [0, \infty) \\ (x, y) & \rightarrow h_1(x, y) := \sigma_n((x + yi)I_n - A) \end{aligned}$$

es semialgebraica. Por otro lado, como el conjunto $[0, \varepsilon)$ es semialgebraico, por el apartado (c) del Lema 7.2.7 deducimos que $h_1^{-1}([0, \varepsilon))$ es semialgebraico. Pero

$$h_1^{-1}([0, \varepsilon)) = \{(x, y) \in \mathbb{R}^2 : h_1(x, y) < \varepsilon\}.$$

Recordando la identificación de \mathbb{C} con \mathbb{R}^2 , por la Proposición 2.1.2 concluimos que $\Lambda'_\varepsilon(A)$ es un conjunto semialgebraico. Finalmente, por el apartado (d) del Lema 7.2.7, concluimos que $\partial\Lambda'_\varepsilon(A)$ es un conjunto semialgebraico. \square

Observación 7.4.1. Todos los resultados obtenidos sobre la semialgebraicidad de los pseudoespectros ordinario y estricto se pueden extender a los geométricos de cualquier orden. La demostración es similar, partiendo de la caracterización de los pseudoespectros geométricos dada en la Proposición 3.1.2, y usando que la función

$$\sigma_{n-k+1} : \mathbb{C}^{n \times n} \rightarrow [0, \infty)$$

es una función semialgebraica. Véase el apartado (c) del Lema 7.2.7.

Observación 7.4.2. Igualmente, partiendo de la caracterización dada en (3.1.5) y utilizando el Lema 7.2.7 (f), se demuestra que los pseudoespectros algebraicos de segundo orden son conjuntos semialgebraicos.

Observación 7.4.3. Sea λ un valor propio semisimple de la matriz $A \in \mathbb{C}^{n \times n}$. Consideremos las funciones $a(\varepsilon)$ y $\delta(\varepsilon)$, área y diámetro de la componente conexa de $\Lambda_\varepsilon(A)$ que contiene al valor propio λ , introducidas en el Capítulo 4. Una conjetura que nos planteamos es la siguiente.

Conjetura 7.4.4. *Las funciones $\delta, a : [0, \infty) \rightarrow \mathbb{R}$ son semialgebraicas.*

El siguiente teorema se prueba en [14, Lema 3.1 (ii)]

Teorema 7.4.5 (Regla de de l'Hôpital inversa). *Si $f, g: [0, \infty) \rightarrow \mathbb{R}$ son funciones semialgebraicas, $f(0) = g(0) = 0$, y existe un $\varepsilon_0 > 0$ tal que $g'(\varepsilon) > 0$ para $\varepsilon \in (0, \varepsilon_0)$, entonces*

$$\lim_{\varepsilon \rightarrow 0^+} \frac{f(\varepsilon)}{g(\varepsilon)} = \ell \in \mathbb{R} \implies \lim_{\varepsilon \rightarrow 0^+} \frac{f'(\varepsilon)}{g'(\varepsilon)} = \ell.$$

Si la Conjetura 7.4.4 fuera cierta, entonces del Teorema 4.7.2 deduciríamos que

$$\pi c^2 = \lim_{\varepsilon \rightarrow 0^+} \frac{a(\varepsilon)}{\varepsilon^2} = \lim_{\varepsilon \rightarrow 0^+} \frac{a'(\varepsilon)}{2\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{a''(\varepsilon)}{2}.$$

Así, la derivada $a''_+(0)$ existiría y sería igual a $2\pi c^2$. Ya que $a''_+(0) = \lim_{\varepsilon \rightarrow 0^+} a''(\varepsilon)$. Esto nos permitiría eliminar la hipótesis de la existencia de $a''_+(0)$, en el Teorema 4.7.3.

Como el área de una región puede ser expresada por medio de una integral de línea, y tomando en cuenta que una integral paramétrica es derivable con respecto al parámetro cuando el integrando lo es, las derivadas $a'(\varepsilon)$ y $a''(\varepsilon)$ existen para $\varepsilon > 0$ suficientemente pequeño.

7.4.2. Prueba del Teorema 7.1.2

Para la demostración de este Teorema precisamos de un resultado previo.

Proposición 7.4.6. *Dada la matriz*

$$A := \begin{pmatrix} 0 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (7.4.2)$$

se cumple que $z_0 = -\frac{4}{15} + 0i$ es un punto aislado de $\partial\Lambda'_{\frac{1}{15}}(A)$.

DEMOSTRACIÓN. Dado $(x, y) \in \mathbb{R}^2$, sigamos las notaciones del Lema 7.3.4; es decir, sea $\mathbf{A}(x, y) := (x + yi)I_3 - A$ y llamemos

$$q(x, y, \varepsilon) := \det \begin{pmatrix} \varepsilon I_3 & \mathbf{A}(x, y) \\ \mathbf{A}^*(x, y) & \varepsilon I_3 \end{pmatrix}.$$

En primer lugar, pretendemos encontrar los puntos singulares de las curvas de la familia, de curvas, $\{q(x, y, \varepsilon) = 0\}_{\varepsilon \in (0, \infty)}$. Es decir, tratamos de encontrar $(x, y, \varepsilon) \in \mathbb{R}^3$ que cumplan el sistema

$$q(x, y, \varepsilon) = 0, \quad q'_x(x, y, \varepsilon) = 0, \quad q'_y(x, y, \varepsilon) = 0.$$

Operando, las expresiones del polinomio q y sus derivadas parciales son:

$$\begin{aligned} q(x, y, \varepsilon) &= \varepsilon^6 - 3(x^2 + y^2 + 6)\varepsilon^4 + [3(x^2 + y^2)^2 + 18(x^2 + y^2) + 8x + 1]\varepsilon^2 - (x^2 + y^2)^3, \\ q'_x(x, y, \varepsilon) &= -6x^5 - 12x^3y^2 + 12x^3\varepsilon^2 - 6xy^4 + 12xy^2\varepsilon^2 - 6x\varepsilon^4 + 36x\varepsilon^2 + 8\varepsilon^2, \\ q'_y(x, y, \varepsilon) &= -6y[(x^2 + y^2 - \varepsilon^2)^2 + \varepsilon^4]. \end{aligned}$$

Si observamos la derivada parcial de q respecto de y , se tiene que:

$$q'_y(x, y, \varepsilon) = 0 \Rightarrow \begin{cases} (x^2 + y^2 - \varepsilon^2)^2 + \varepsilon^4 = 0 & \Leftrightarrow (x, y, \varepsilon) = (0, 0, 0) \\ \text{ó} \\ y = 0. \end{cases}$$

Pero, como $\varepsilon > 0$, resulta que $y = 0$. Luego, al ser $q'_y(x, 0, \varepsilon) \equiv 0$, para buscar los puntos singulares, basta con considerar el sistema

$$\begin{cases} q(x, 0, \varepsilon) = 0, \\ q'_x(x, 0, \varepsilon) = 0. \end{cases} \quad (7.4.3)$$

Buscamos las soluciones $(x, \varepsilon) \in \mathbb{R}^2$ de este sistema. Para ello consideramos $q(x, 0, \varepsilon)$ y $q'_x(x, 0, \varepsilon)$ como polinomios en ε con coeficientes en $\mathbb{R}[x]$.

$$\begin{cases} q(x, 0, \varepsilon) = \varepsilon^6 + (-3x^2 - 18)\varepsilon^4 + (3x^4 + 18x^2 + 8x + 1)\varepsilon^2 - x^6 \\ \quad = \varepsilon^6 + p_4(x)\varepsilon^4 + p_2(x)\varepsilon^2 + p_0(x), \\ q'_x(x, 0, \varepsilon) = -6x\varepsilon^4 + (12x^3 + 36x + 8)\varepsilon^2 - 6x^5 \\ \quad = \tilde{p}_4(x)\varepsilon^4 + \tilde{p}_2(x)\varepsilon^2 + \tilde{p}_0(x). \end{cases}$$

Sus grados, respecto de ε , son 6 y 4, respectivamente.

Mediante la resultante de Sylvester, hallaremos los valores de x para los que estos polinomios tienen raíces comunes en ε . Véase (C.1.2). Procedamos, la matriz resultante de Sylvester de $q(x, 0, \varepsilon)$ y $q'_x(x, 0, \varepsilon)$ respecto de ε es la matriz siguiente de orden $10 = 6 + 4$:

$$\begin{pmatrix} 1 & 0 & p_4(x) & 0 & p_2(x) & 0 & p_0(x) & 0 & 0 & 0 \\ 0 & 1 & 0 & p_4(x) & 0 & p_2(x) & 0 & p_0(x) & 0 & 0 \\ 0 & 0 & 1 & 0 & p_4(x) & 0 & p_2(x) & 0 & p_0(x) & 0 \\ 0 & 0 & 0 & 1 & 0 & p_4(x) & 0 & p_2(x) & 0 & p_0(x) \\ \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) & 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) & 0 \\ 0 & 0 & 0 & 0 & 0 & \tilde{p}_4(x) & 0 & \tilde{p}_2(x) & 0 & \tilde{p}_0(x) \end{pmatrix}$$

El determinante de esta matriz, $R(q(x, 0, \varepsilon), q'_x(x, 0, \varepsilon), \varepsilon)$, nos da la resultante de Sylvester de los dos polinomios:

$$\begin{aligned} R(q(x, 0, \varepsilon), q'_x(x, 0, \varepsilon), \varepsilon) = \\ 64x^{10} (21600x^4 + 15120x^3 + 2781x^2 - 104x - 48)^2 = 0. \end{aligned}$$

Simplificando, tenemos

$$64x = 0, \quad 21600x^4 + 15120x^3 + 2781x^2 - 104x - 48 = 0$$

cuyas raíces en x son

$$0, \frac{-2 - \sqrt{22}}{24}, \frac{-2 + \sqrt{22}}{24}, -\frac{4}{15}. \quad (7.4.4)$$

Sustituimos $x = -\frac{4}{15}$ en (7.4.3) para hallar las raíces comunes de esos polinomios en la variable ε

$$\begin{cases} q(-4/15, 0, \varepsilon) = \varepsilon^6 - \frac{1366}{75} \varepsilon^4 + \frac{2731}{16875} \varepsilon^2 - \frac{4096}{11390625} = 0 \\ q'_x(-4/15, 0, \varepsilon) = \frac{8}{5} \varepsilon^4 - \frac{2056}{1125} \varepsilon^2 + \frac{2048}{253125} = 0. \end{cases}$$

Las respectivas raíces son

$$\frac{1}{15}, \frac{-1}{15}, \frac{1}{15}, \frac{-1}{15}, \frac{64}{15}, \frac{-64}{15}, \frac{1}{15}, \frac{16}{15}, \frac{-16}{15}, \frac{-1}{15}.$$

Como se puede observar la raíz $\frac{1}{15}$, es doble para el polinomio q y simple para q'_x ; es decir, el punto $(-\frac{4}{15}, 0)$ es un punto doble de la curva $q(x, y, \frac{1}{15}) = 0$ y, además,

$$q''_{xy}\left(-\frac{4}{15}, 0, \frac{1}{15}\right)^2 - q''_{xx}\left(-\frac{4}{15}, 0, \frac{1}{15}\right)q''_{yy}\left(-\frac{4}{15}, 0, \frac{1}{15}\right) < 0;$$

por lo tanto, en virtud del Apartado II del Teorema A.5.1, $(-\frac{4}{15}, 0)$ es un punto aislado de la curva $q(x, y, 1/15) = 0$. Para finalizar la prueba de la Proposición debemos probar que $\frac{1}{15}$ es el menor valor singular de la matriz $-\frac{4}{15}I_3 - A$, como indica el Lema 7.3.4. Los valores singulares de $-\frac{4}{15}I_3 - A$ son las raíces cuadradas positivas de los ceros del polinomio

$$p(\lambda) = \det[\lambda I_3 - (-\frac{4}{15}I_3 - A)^*(-\frac{4}{15}I_3 - A)] = \lambda^3 - \frac{1366}{75}\lambda^2 + \frac{2731}{16875}\lambda - \frac{4096}{11390625},$$

que son

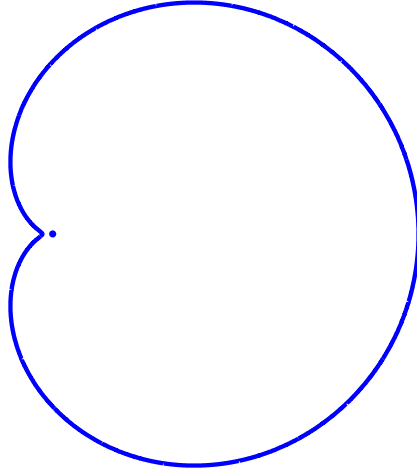
$$\left(\frac{64}{15}\right)^2, \left(\frac{1}{15}\right)^2, \left(\frac{1}{15}\right)^2.$$

Por lo tanto, $h_1(-4/15) = \sigma_3(-\frac{4}{15}I_3 - A) = 1/15$. Volviendo a aplicar el Lema 7.3.4, deducimos que $-4/15 + 0i$ es un punto aislado de $\partial\Lambda'_{\frac{1}{15}}(A)$. \square

La Figura 7.4 representa la frontera del pseudoespectro estricto de la matriz A de nivel $\varepsilon = 1/15$. Obsérvese que el punto $(-4/15, 0)$ es un punto aislado.

Analizamos a continuación las otras raíces en (7.4.4) para el polinomio q .

1. Punto $(-\frac{2+\sqrt{22}}{24}, 0)$.

Figura 7.4: Punto aislado $(-4/15, 0)$.

Sustituimos $x = -\frac{2+\sqrt{22}}{24}$ en (7.4.3) para hallar las raíces comunes de esos polinomios en la variable ε

$$\begin{cases} q\left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon\right) &= \varepsilon^6 - \frac{1741+2\sqrt{22}}{96}\varepsilon^4 + \frac{31937-5708\sqrt{22}}{6912}\varepsilon^2 - \frac{5629+1190\sqrt{22}}{23887872} \\ &= 0 \\ q'_x\left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon\right) &= \frac{2+\sqrt{22}}{4}\varepsilon^4 + \frac{2810-881\sqrt{22}}{576}\varepsilon^2 + \frac{1658+361\sqrt{22}}{331776} = 0. \end{cases}$$

Las raíces de $q\left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon\right)$ son

$$\begin{aligned} &\frac{4 + \sqrt{22}}{4} - \frac{4 - \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ &\frac{4 + \sqrt{22}}{4} + \frac{4 - \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ &-\frac{4 + \sqrt{22}}{4} + \frac{4 - \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ &-\frac{4 + \sqrt{22}}{4} - \frac{4 - \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ &-\frac{25}{12} + \frac{11}{24}\sqrt{22}, \quad \frac{25}{12} - \frac{11}{24}\sqrt{22}. \end{aligned}$$

Las raíces de $q'_x \left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon \right)$ son

$$\begin{aligned} & \frac{7}{12} + \frac{1}{8}\sqrt{22}, \quad -\frac{7}{12} - \frac{1}{8}\sqrt{22}, \\ & -\frac{25}{12} + \frac{11}{24}\sqrt{22}, \quad \frac{25}{12} - \frac{11}{24}\sqrt{22}. \end{aligned}$$

Como se puede observar, la raíz $-\frac{25}{12} + \frac{11}{24}\sqrt{22}$ es simple para los polinomios q y q'_x ; es decir, el punto $\left(-\frac{2+\sqrt{22}}{24}, 0 \right)$ es un punto doble de la curva $q \left(x, y, -\frac{25}{12} + \frac{11}{24}\sqrt{22} \right) = 0$.

2. Punto $\left(-\frac{2+\sqrt{22}}{24}, 0 \right)$.

Sustituimos $x = -\frac{2+\sqrt{22}}{24}$ en (7.4.3) para hallar las raíces comunes de esos polinomios en la variable ε

$$\begin{cases} q \left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon \right) &= \varepsilon^6 - \frac{1741-2\sqrt{22}}{96}\varepsilon^4 + \frac{31937+5708\sqrt{22}}{27648}\varepsilon^2 - \frac{5629-1190\sqrt{22}}{23887872} \\ &= 0 \\ q'_x \left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon \right) &= \frac{2+\sqrt{22}}{4}\varepsilon^4 + \frac{2810-881\sqrt{22}}{576}\varepsilon^2 + \frac{1658+361\sqrt{22}}{331776} = 0. \end{cases}$$

Las raíces de $q \left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon \right)$ son

$$\begin{aligned} & \frac{4 - \sqrt{22}}{4} - \frac{4 + \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ & \frac{4 - \sqrt{22}}{4} + \frac{4 + \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ & -\frac{4 - \sqrt{22}}{4} + \frac{4 + \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ & -\frac{4 - \sqrt{22}}{4} - \frac{4 + \sqrt{22}}{24} \sqrt{2655 + 566\sqrt{22}}, \\ & \frac{25}{12} + \frac{11}{24}\sqrt{22}, \quad -\frac{25}{12} - \frac{11}{24}\sqrt{22}. \end{aligned}$$

Las raíces de $q'_x \left(-\frac{2+\sqrt{22}}{24}, 0, \varepsilon \right)$ son

$$\begin{aligned} & -\frac{7}{12} + \frac{1}{8}\sqrt{22}, \quad \frac{7}{12} - \frac{1}{8}\sqrt{22}, \\ & \frac{25}{12} + \frac{11}{24}\sqrt{22}, \quad -\frac{25}{12} - \frac{11}{24}\sqrt{22}. \end{aligned}$$

Como se puede observar, la raíz $\frac{25}{12} + \frac{11}{24}\sqrt{22}$ es simple para los polinomios q y q'_x ; es decir, el punto $\left(-\frac{2+\sqrt{22}}{24}, 0 \right)$ es un punto doble de la curva

$q\left(x, y, \frac{25}{12} + \frac{11}{24}\sqrt{22}\right) = 0$. Es preciso observar que ese punto no pertenece a la $\partial\Lambda_{\frac{25}{12} + \frac{11}{24}\sqrt{22}}(A)$. En efecto, se tiene que $\sigma_1\left(\frac{-2+\sqrt{22}}{24}I_3 - A\right) = \frac{25}{12} + \frac{11}{24}\sqrt{22}$.

3. Punto $(0, 0)$.

Sustituimos $x = 0$ en (7.4.3) para hallar las raíces comunes de esos polinomios en la variable ε

$$\begin{cases} q(0, 0, \varepsilon) &= \varepsilon^6 - 18\varepsilon^4 + \varepsilon^2 = 0 \\ q'_x(0, 0, \varepsilon) &= 8\varepsilon^2 = 0. \end{cases}$$

Las raíces de $q(0, 0, \varepsilon)$ son

$$0, 0, -2 + \sqrt{5}, -2 - \sqrt{5}, 2 + \sqrt{5}, 2 - \sqrt{5}.$$

Las raíces de $q'_x(0, 0, \varepsilon)$ son

$$0, 0.$$

Como se puede observar, la raíz 0 es doble para los polinomios q y q'_x ; es decir, el punto $(0, 0)$ es un punto singular (pero no doble) de la curva $q(x, y, 0) = -(x^2 + y^2)^3 = 0$. Claro que esto simplemente confirma que $z = 0$ es el valor propio de A ; por lo que $\sigma_3(0I_3 - A) = \sigma_3(A) = 0$. Cosa también sabida pues $\text{rg } A = 2$.

A continuación veremos en las Figuras 7.5, 7.6 y 7.7 una ilustración de las ideas que venimos desarrollando.

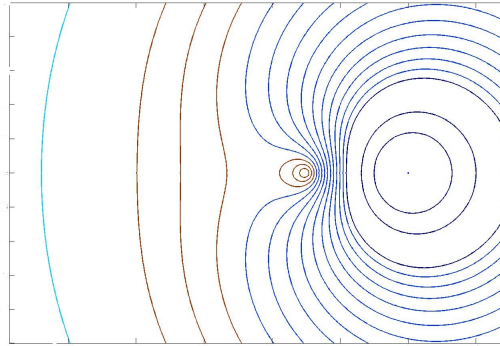


Figura 7.5: ε -pseudoespectros con ε bajo.

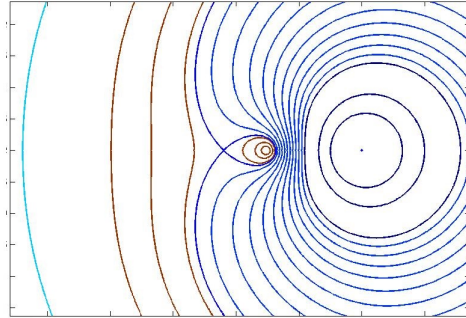


Figura 7.6: Curva de nivel de $h_1(x, y) = \sigma_3((x + yi)I_3 - A)$ que contiene un punto de ensilladura.

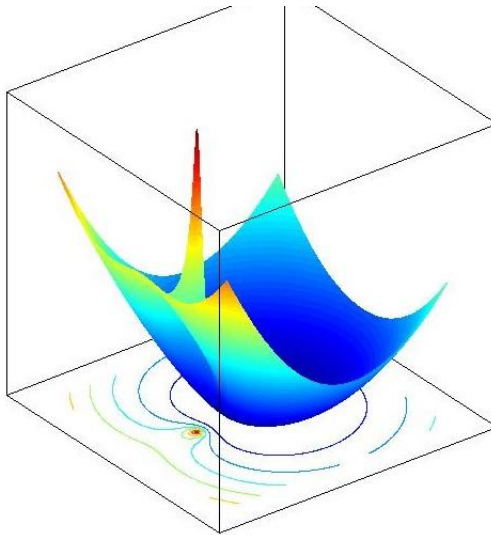


Figura 7.7: Máximo relativo de $h_1(x, y) = \sigma_3((x + yi)I_3 - A)$.

Una vez visto este resultado, vamos a demostrar a continuación el Teorema 7.1.2

DEMOSTRACIÓN.

Para simplificar la notación, vamos a definir

$$w_0 := -\frac{4}{15}, \quad \varepsilon_1 := \frac{1}{15}. \quad (7.4.5)$$

Sea $\mathcal{P} := \{z_1, \dots, z_k\} \subset \mathbb{C}$. Tomando la matriz dada en (7.4.2), denotamos por

$$\delta := \frac{1}{4} \min_{1 \leq p \leq q \leq k} |z_p - z_q|, \quad \gamma := \frac{\delta}{\varepsilon_1 + \|A\|}, \quad \varepsilon_0 = \gamma \varepsilon_1 = \frac{\gamma}{15}. \quad (7.4.6)$$

Aplicando el Lema 2.2.2, para $p = 1, 2, \dots, k$ resulta

$$\begin{aligned} \Lambda'_{\varepsilon_0}(\gamma A + (z_p - \gamma w_0)I_3) &= z_p - \gamma w_0 + \gamma \Lambda'_{\varepsilon_0/|\gamma|}(A) \\ &= z_p - \gamma w_0 + \gamma \Lambda'_{\varepsilon_1}(A). \end{aligned} \quad (7.4.7)$$

Por la Proposición 7.4.6 deducimos que z_p es un punto aislado de $\partial \Lambda'_{\varepsilon_0}(\gamma A + (z_p - \gamma w_0)I_3)$.

De otro lado, por la Proposición 2.2.4 (a), de (7.4.7) resulta

$$\begin{aligned} \Lambda'_{\varepsilon_0}(\gamma A + (z_p - \gamma w_0)I_3) &= z_p - \gamma w_0 + \gamma \Lambda'_{\varepsilon_1}(A) \\ &\subset z_p - \gamma w_0 + \gamma \mathcal{D}(0, \varepsilon_1 + \|A\|) = \mathcal{D}(z_p - \gamma w_0, \delta). \end{aligned} \quad (7.4.8)$$

Si definimos

$$A_0 := \text{diag}(\gamma A + (z_p - \gamma w_0)I_3)_{p=1}^k,$$

y utilizamos el apartado (g) de la Proposición 2.2.4 y la expresión (7.4.8), resulta

$$\Lambda'_{\varepsilon_0}(A_0) = \bigcup_{p=1}^k \Lambda'_{\varepsilon_0}(\gamma A + (z_p - \gamma w_0)I_3) \subset \bigcup_{p=1}^k \mathcal{D}(z_p - \gamma w_0, \delta).$$

Para finalizar la prueba del apartado (a), observemos que $z_p \in \mathcal{D}(z_p - \gamma w_0, \delta)$. Además, esos discos son disjuntos dos a dos, por lo tanto, z_p es un punto aislado de $\partial \Lambda'_{\varepsilon_0}(A_0)$. El apartado (b) se deduce del apartado (a) del Teorema 7.3.3. \square

Observación 7.4.7. En la demostración del Teorema 7.1.2 hemos probado que cada punto del conjunto \mathcal{P} está en la frontera de una componente conexa *distinta* del pseudoespectro estricto $\Lambda'_{\varepsilon_0}(A_0)$. La cuestión que surge es: dado un conjunto finito de números complejos \mathcal{P} , ¿existe una matriz A de modo que para algún ε_0 se cumple que cada $z \in \mathcal{P}$ es un punto aislado de $\partial \Lambda'_{\varepsilon_0}(A_0)$, y además el conjunto \mathcal{P} está contenido en la frontera de una misma componente conexa de $\Lambda'_{\varepsilon_0}(A_0)$? La respuesta es afirmativa como demostramos a continuación.

Sea la matriz A_0 y consideremos las notaciones de la demostración del Teorema 7.1.2. Además, denotamos por $\text{dist}(A, B)$ la distancia entre los subconjuntos A y B de \mathbb{C} , que se define como

$$\text{dist}(A, B) := \inf_{(z, w) \in A \times B} |z - w|.$$

En el caso de que los conjuntos A y B sean compactos (cerrados y acotados), el ínfimo se puede sustituir por el mínimo. Dados $z_p, z_q \in \mathcal{P}$, consideremos un camino formado por un número finito de arcos Γ_{pq} que una los puntos z_p y z_q , y tales que

$$\left(\bigcup_{z \in \Gamma_{pq}} \mathcal{D}(z, \varepsilon_0) \right) \cap \mathcal{D}(z_r - \gamma w_0, \delta) = \emptyset$$

para $r \neq p, q$, de modo que podamos conectar cualquier par de puntos de dos discos distintos. Esto es posible ya que, como $\|A\| = 2 + \sqrt{5}$, se tiene

$$\varepsilon_0 = \frac{15\sqrt{5} - 31}{164} \delta,$$

y además, para $p \neq q$ se cumple

$$\text{dist}(\mathcal{D}(z_p - \gamma w_0, \delta), \mathcal{D}(z_q - \gamma w_0, \delta)) = |z_p - z_q| - 2\delta \geq 2\delta.$$

Sea el arco Γ_{pq} que une los puntos z_p y z_q . Como ambos puntos son aislados para $\partial\Lambda'_{\varepsilon_0}(A_0)$, existen dos discos $\mathcal{D}(z_p, \eta)$ y $\mathcal{D}(z_q, \eta)$, de modo que están contenidos en $\Lambda_{\varepsilon_0}(A_0)$. Elegimos un conjunto finito de puntos del arco γ_{pq} :

$$\{w_1, w_2, \dots, w_m\},$$

tales que

$$\begin{aligned} |w_r - w_s| &< 2\varepsilon_0, \\ |z_p - w_1| &< |z_p - w_r|, \quad r > 1, \\ |z_q - w_m| &< |z_q - w_r|, \quad r < m, \\ z_p &\notin \mathcal{D}(w_1, \varepsilon_0), \quad z_q \notin \mathcal{D}(w_m, \varepsilon_0), \\ \mathcal{D}(z_p, \eta) \cap \mathcal{D}(w_1, \varepsilon_0) &\neq \emptyset, \quad \mathcal{D}(z_q, \eta) \cap \mathcal{D}(w_m, \varepsilon_0) \neq \emptyset \end{aligned}$$

Es inmediato comprobar que los puntos del conjunto \mathcal{P} son aislados para la frontera del pseudoespectro estricto de nivel ε_0 de la matriz

$$A_1 = \text{diag}(A_0, w_1, w_2, \dots, w_m).$$

Además, z_p, z_q están en la frontera de una componente conexa de $\Lambda'_{\varepsilon_0}(A_1)$. Repitiendo el razonamiento, podemos construir en un número finito de pasos una matriz A , tal que el conjunto \mathcal{P} esté contenido en la frontera de alguna componente conexa de $\Lambda'_{\varepsilon_0}(A)$.

Hemos probado que una matriz puede tener un número finito de puntos aislados para la frontera del pseudoespectro estricto, y todos esos puntos están en

una misma componente conexas. Sin embargo, podemos preguntarnos si puede haber dos puntos aislados en la frontera de una componente conexas del pseudo-espectro estricto, que contiene sólo un valor propio. La respuesta es afirmativa. En efecto, si tomamos la siguiente matriz

$$\begin{pmatrix} 0 & 1 & 4 & 20 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

repetiendo la demostración de la Proposición 7.4.6, se comprueba que los puntos

$$-\frac{4}{19} \pm \frac{4\sqrt{5}}{95}i$$

son aislados del pseudoespectro estricto de nivel $\varepsilon = \frac{\sqrt{5}}{95}$. Sea $q(x, y, \varepsilon)$ el polinomio $\det(\varepsilon^2 I_4 - ((x - yi)I_4 - A^*)((x + yi)I_4 - A))$. En la Figura 7.8 está representada la curva algebraica real $q(x, y, \sqrt{5}/95) = 0$ en color azul; en color verde aparecen los puntos dobles aislados $(-\frac{4}{19}, \pm \frac{4\sqrt{5}}{95})$, son acnodos, y forman parte de esta curva. En la Figura 7.9 podemos ver la gráfica de la superficie $z = h_1(x, y) = \sigma_4((x + yi)I_4 - A)$, donde aparecen dos máximos relativos de esta función en los puntos $(-\frac{4}{19}, \pm \frac{4\sqrt{5}}{95})$.

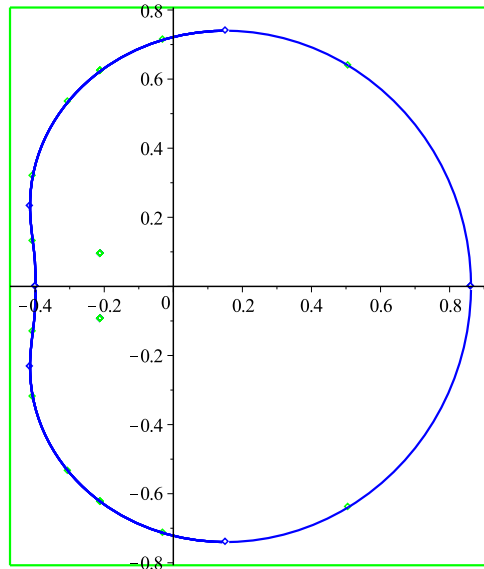


Figura 7.8: Puntos acnodos $(-\frac{4}{19}, \pm \frac{4\sqrt{5}}{95})$.

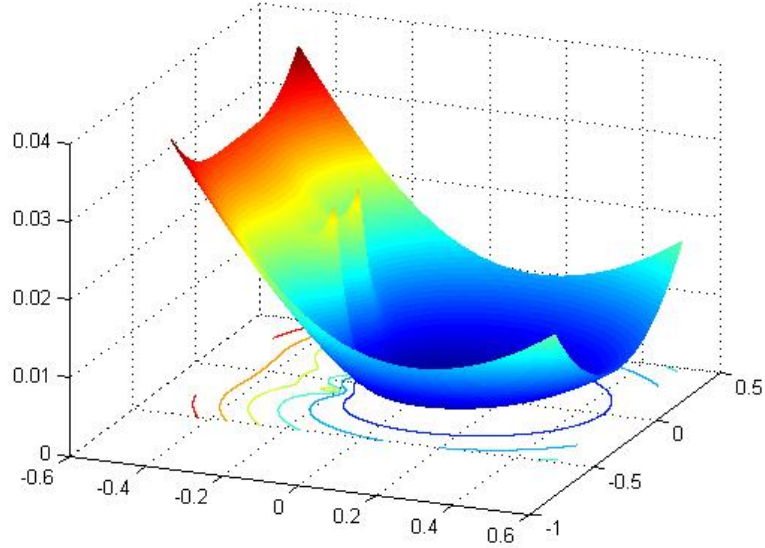


Figura 7.9: Máximos relativos de h_1 en los puntos $(-\frac{4}{19}, \pm \frac{4\sqrt{5}}{95})$.

Observación 7.4.8. Hemos probado que $\partial\Lambda_\varepsilon(A) \subset \partial\Lambda'_\varepsilon(A)$, y además pueden diferir en un conjunto finito de puntos, en algún caso. La pregunta que surge es: ¿pueden diferir ambos conjuntos en un número infinito de puntos?

Supongamos que $\partial\Lambda'_\varepsilon(A) \setminus \partial\Lambda_\varepsilon(A)$ tiene infinitos puntos. Entonces una de las componentes conexas de $\partial\Lambda'_\varepsilon(A)$, que denotamos por \mathcal{C} , tiene infinitos puntos que no están en $\partial\Lambda_\varepsilon(A)$. Denotemos a ese conjunto por $\mathcal{X} := \mathcal{C} \cap (\partial\Lambda_\varepsilon(A))^c$. Si \mathcal{X} tiene infinitos puntos aislados (respecto al mismo conjunto) entonces $\mathcal{C} \cap \partial\Lambda_\varepsilon(A)$ tiene infinitas componentes conexas, lo que es absurdo. Luego, \mathcal{X} no tiene infinitos puntos aislados, es decir es a lo sumo una unión finita de subvariedades reales analíticas 1–dimensionales de \mathbb{R}^2 .

Luego, los conjuntos $\partial\Lambda_\varepsilon(A)$ y $\partial\Lambda'_\varepsilon(A)$ difieren a lo sumo en un conjunto finito de puntos y en una subvariedad real analítica 1–dimensional de \mathbb{R}^2 , con un número finito de componentes conexas.

Concluimos esta sección probando que las funciones h_1 y h_2 coinciden en los máximos locales de h_1 .

Proposición 7.4.9. *Sea $A \in \mathbb{C}^{n \times n}$. Supongamos que la función h_1 tiene un máximo local en z_0 . Entonces,*

$$h_2(z_0) = h_1(z_0).$$

DEMOSTRACIÓN. Sin pérdida de generalidad podemos considerar $z_0 = 0$. Denotemos por $\sigma_0 := h_1(0) > 0$.

Vamos a suponer en primer lugar que el valor singular σ_0 de $\mathbf{A}(0)$ es simple. Sea (u, v) un par de vectores singulares de $\mathbf{A}(0)$ asociado a σ_0 . Identificando z con $(x, y) \in \mathbb{R}^2$, existe un $\delta > 0$ tal que h_1 es una función real analítica en $D((0, 0), \delta)$, véase [70]. Además

$$h_1(x, y) = \sigma_0 + \operatorname{Re}(-u^*v)x + \operatorname{Im}(-u^*v)y + O(x^2 + y^2)$$

cuando $(x, y) \rightarrow (0, 0)$. Por ser h_1 una función real analítica y tener un máximo local en $(0, 0)$, sus derivadas parciales de primer orden se anulan en ese punto, por lo tanto, $u^*v = 0$. Según [54], se tiene que la matriz $\mathbf{A}(0) - \sigma uv^*$, que dista de $\mathbf{A}(0)$ exactamente σ_0 , tiene a cero como valor propio múltiple. Con lo cual, por el Teorema 3.1.4, resulta

$$\sigma_0 \geq h_2(0) = \max_{t \geq 0} \sigma_{2n-1} \begin{pmatrix} \mathbf{A}(0) & tI_n \\ O & \mathbf{A}(0) \end{pmatrix} \geq \sigma_{2n-1} \begin{pmatrix} \mathbf{A}(0) & O \\ O & \mathbf{A}(0) \end{pmatrix} = \sigma_0.$$

Supongamos ahora que σ_0 sea un valor singular múltiple de $\mathbf{A}(0)$. Para este caso, resulta $\sigma_{n-1}(\mathbf{A}(0)) = \sigma_0$. Pero como

$$\sigma_{n-1}(\mathbf{A}(0)) \geq \sigma_{2n-1} \begin{pmatrix} \mathbf{A}(0) & X \\ 0 & \mathbf{A}(0) \end{pmatrix},$$

para cualquier matriz $X \in \mathbb{C}^{n \times n}$, entonces

$$\sigma_0 = \sigma_{n-1}(\mathbf{A}(0)) \geq h_2(0) \geq \sigma_0.$$

□

Observación 7.4.10. Por el Teorema 3.1.4, si h_1 tiene un máximo local en z_0 , entonces $h_1(z_0)$ es la distancia de A al conjunto de matrices que tienen a z_0 como valor propio múltiple.

Observación 7.4.11. Notemos que la misma propiedad de la Proposición 7.4.9 la cumplen los puntos de colisión de dos componentes conexas, como se ha probado en la Proposición 3.3.1. De hecho, las funciones h_1 y h_2 alcanzan el mismo valor en los puntos de $\partial\Lambda'_\varepsilon(A)$ que son puntos singulares de la curva algebraica real $q(x, y, \varepsilon) = 0$, donde $q(x, y, \varepsilon)$ es el polinomio de Wielandt de la matriz A correspondiente al valor $\varepsilon > 0$.

7.5. Corroboración mediante derivadas

En la Subsección 7.4.2 hemos obtenido cuatro pares (x, y) con sus correspondientes ε , soluciones del sistema (7.4.3). Uno de ellos es un punto singular aislado de la curva algebraica real $q(x, y, \frac{1}{15}) = 0$ (que contiene a $\partial\Lambda'_{1/15}(A)$), y que ha sido útil en la demostración del Teorema 7.1.2. En esta sección, el objetivo es analizar los otros puntos, para ello utilizaremos el resultado siguiente, que es una adaptación de [53, Teorema 8, p. 161] y [53, Teorema 11, p. 166] sobre las derivadas parciales de un valor singular.

Teorema 7.5.1. Sean una matriz $A \in \mathbb{C}^{n \times n}$ y σ_0 un valor singular simple positivo de A . Sean (p, q) un par de vectores singulares unitarios de A para σ_0 . Sean

$$Y := \begin{pmatrix} \sigma_0 I_n & -A \\ -A^* & \sigma_0 I_n \end{pmatrix}, \quad w = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{C}^{2n \times 1}.$$

Dado un $(x, y) \in \mathbb{R}^2$, consideramos las funciones

$$\mathbf{A}(x, y) = (x + yi)I_n - A, \quad B(x, y) = \begin{pmatrix} O & \mathbf{A}(x, y) \\ \mathbf{A}(x, y)^* & O \end{pmatrix}.$$

Entonces, existen un $\delta > 0$, un entorno del origen $D = D((0, 0), \delta)$ y una función real analítica f en D tal que

1. $f(0, 0) = \sigma_0$,
2. para todo $(x, y) \in D$, $f(x, y)$ es un valor singular de $\mathbf{A}(x, y)$, y
- 3.

$$\begin{aligned} f'_x(0, 0) &= \frac{1}{2} w^* B'_x(0, 0) w, & f'_y(0, 0) &= \frac{1}{2} w^* B'_y(0, 0) w, \\ f''_{xx}(0, 0) &= w^* B''_{xx}(0, 0) Y^\dagger B'_x(0, 0) w, & f''_{yy}(0, 0) &= w^* B''_{yy}(0, 0) Y^\dagger B'_y(0, 0) w, \\ f''_{xy}(0, 0) &= \frac{1}{2} (w^* B''_{xy}(0, 0) Y^\dagger B'_y(0, 0) w + w^* B''_{yx}(0, 0) Y^\dagger B'_x(0, 0) w). \end{aligned}$$

Caso: $x_2 = -\frac{2+\sqrt{22}}{24}$.

La matriz correspondiente es

$$\mathbf{A}(x_1, 0) = \begin{pmatrix} -\frac{2+\sqrt{22}}{24} & -1 & -4 \\ 0 & -\frac{2+\sqrt{22}}{24} & -1 \\ 0 & 0 & -\frac{2+\sqrt{22}}{24} \end{pmatrix}.$$

El número $\varepsilon_2 = \frac{-50+11\sqrt{22}}{24}$ es el menor valor singular de $\mathbf{A}(x_1, 0)$. Si consideramos la descomposición en valores singulares de la matriz $\mathbf{A}(x_1, 0)$ y aplicamos el Teorema 7.5.1, tenemos

$$w = \left(\frac{-7+\sqrt{22}}{12}, \frac{2+\sqrt{22}}{12}, \frac{5+\sqrt{22}}{12}, \frac{5+\sqrt{22}}{12}, \frac{2+\sqrt{22}}{12}, \frac{-7+\sqrt{22}}{12} \right)^T.$$

Por otro lado, la matriz Y^\dagger se puede escribir como

$$Y^\dagger = \begin{pmatrix} M & N \\ N^T & Q \end{pmatrix},$$

donde

$$M = \begin{pmatrix} -\frac{24415\sqrt{22}}{228096} - \frac{739}{1296} & \frac{33593\sqrt{22}}{57024} + \frac{29053}{10368} & -\frac{121339\sqrt{22}}{228096} - \frac{6479}{2592} \\ \frac{33593\sqrt{22}}{57024} + \frac{29053}{10368} & -\frac{341621\sqrt{22}}{114048} - \frac{18505}{1296} & \frac{142751\sqrt{22}}{57024} + \frac{121729}{10368} \\ -\frac{121339\sqrt{22}}{228096} - \frac{6479}{2592} & \frac{142751\sqrt{22}}{57024} + \frac{121729}{10368} & -\frac{325927\sqrt{22}}{228096} - \frac{562}{81} \end{pmatrix},$$

$$N = \begin{pmatrix} \frac{137861\sqrt{22}}{228096} + \frac{7345}{2592} & -\frac{38983\sqrt{22}}{57024} - \frac{33155}{10368} & \frac{34337\sqrt{22}}{228096} + \frac{989}{1296} \\ -\frac{162529\sqrt{22}}{57024} - \frac{138623}{10368} & \frac{397963\sqrt{22}}{114048} + \frac{21239}{1296} & -\frac{38983\sqrt{22}}{57024} - \frac{33155}{10368} \\ \frac{390233\sqrt{22}}{228096} + \frac{650}{81} & -\frac{162529\sqrt{22}}{57024} - \frac{138623}{10368} & \frac{137861\sqrt{22}}{228096} + \frac{7345}{2592} \end{pmatrix},$$

$$Q = \begin{pmatrix} -\frac{325927\sqrt{22}}{228096} - \frac{562}{81} & \frac{142751\sqrt{22}}{57024} + \frac{121729}{10368} & -\frac{121339\sqrt{22}}{228096} - \frac{6479}{2592} \\ \frac{142751\sqrt{22}}{57024} + \frac{121729}{10368} & -\frac{341621\sqrt{22}}{114048} - \frac{18505}{1296} & \frac{33593\sqrt{22}}{57024} + \frac{29053}{10368} \\ -\frac{121339\sqrt{22}}{228096} - \frac{6479}{2592} & \frac{142751\sqrt{22}}{57024} + \frac{121729}{10368} & -\frac{24415\sqrt{22}}{228096} - \frac{739}{1296} \end{pmatrix}.$$

Utilizando el apartado (3) del Teorema 7.5.1, se tiene que

$$f'_x(0,0) = f'_y(0,0) = f''_{xy}(0,0) = 0,$$

$$f''_{xx}(0,0) = \frac{44+13\sqrt{22}}{36}, \quad f''_{yy}(0,0) = -\frac{616+128\sqrt{22}}{33}.$$

Lo cual prueba que la función h_1 tiene un punto de ensilladura en $(x_2, 0)$.

Caso: $x_3 = \frac{-2+\sqrt{22}}{24}$.

La matriz correspondiente es

$$\mathbf{A}(x_3, 0) = \begin{pmatrix} \frac{-2+\sqrt{22}}{24} & -1 & -4 \\ 0 & \frac{-2+\sqrt{22}}{24} & -1 \\ 0 & 0 & \frac{-2+\sqrt{22}}{24} \end{pmatrix}.$$

El número $\varepsilon_3 = \frac{50+11\sqrt{22}}{24}$ es un valor singular simple de $\mathbf{A}(x_3, 0)$. Si consideramos la descomposición en valores singulares de la matriz $\mathbf{A}(x_3, 0)$ y aplicamos el Teorema 7.5.1, tenemos

$$w = \left(\frac{-7+\sqrt{22}}{12}, \frac{2-\sqrt{22}}{12}, \frac{5-\sqrt{22}}{12}, \frac{-5+\sqrt{22}}{12}, \frac{-2+\sqrt{22}}{12}, \frac{-7-\sqrt{22}}{12} \right)^T.$$

Por otro lado, la matriz Y^\dagger se puede escribir como

$$Y^\dagger = \begin{pmatrix} M & N \\ N^T & Q \end{pmatrix},$$

donde

$$M = \begin{pmatrix} -\frac{24415\sqrt{22}}{228096} + \frac{739}{1296} & \frac{33593\sqrt{22}}{57024} - \frac{29053}{10368} & -\frac{121339\sqrt{22}}{228096} + \frac{6479}{2592} \\ \frac{33593\sqrt{22}}{57024} - \frac{29053}{10368} & -\frac{341621\sqrt{22}}{114048} + \frac{18505}{1296} & \frac{142751\sqrt{22}}{57024} - \frac{121729}{10368} \\ -\frac{121339\sqrt{22}}{228096} + \frac{6479}{2592} & \frac{142751\sqrt{22}}{57024} - \frac{121729}{10368} & -\frac{325927\sqrt{22}}{228096} + \frac{562}{81} \end{pmatrix},$$

$$N = \begin{pmatrix} -\frac{137861\sqrt{22}}{228096} + \frac{7345}{2592} & \frac{38983\sqrt{22}}{57024} + \frac{33155}{10368} & -\frac{34337\sqrt{22}}{228096} + \frac{989}{1296} \\ \frac{162529\sqrt{22}}{57024} - \frac{138623}{10368} & -\frac{397963\sqrt{22}}{114048} + \frac{21239}{1296} & \frac{38983\sqrt{22}}{57024} + \frac{33155}{10368} \\ -\frac{390233\sqrt{22}}{228096} + \frac{650}{81} & \frac{162529\sqrt{22}}{57024} - \frac{138623}{10368} & -\frac{137861\sqrt{22}}{228096} + \frac{7345}{2592} \end{pmatrix},$$

$$Q = \begin{pmatrix} -\frac{325927\sqrt{22}}{228096} + \frac{562}{81} & \frac{142751\sqrt{22}}{57024} - \frac{121729}{10368} & -\frac{121339\sqrt{22}}{228096} + \frac{6479}{2592} \\ \frac{142751\sqrt{22}}{57024} - \frac{121729}{10368} & -\frac{341621\sqrt{22}}{114048} + \frac{18505}{1296} & \frac{33593\sqrt{22}}{57024} - \frac{29053}{10368} \\ -\frac{121339\sqrt{22}}{228096} + \frac{6479}{2592} & \frac{33593\sqrt{22}}{57024} - \frac{29053}{10368} & -\frac{24415\sqrt{22}}{228096} + \frac{739}{1296} \end{pmatrix}.$$

Utilizando el apartado (3) del Teorema 7.5.1, se tiene que

$$f'_x(0, 0) = f'_y(0, 0) = f''_{xy}(0, 0) = 0,$$

$$f''_{xx}(0, 0) = \frac{-44+13\sqrt{22}}{36}, \quad f''_{yy}(0, 0) = \frac{616-128\sqrt{22}}{33}.$$

Lo cual prueba que la función $\|\mathbf{A}(x, y)\| = \sigma_1(\mathbf{A}(x, y))$ tiene un mínimo local en $(x_3, 0)$.

Aunque ya hemos caracterizado el punto singular doble $(-\frac{4}{15}, 0)$ mediante la naturaleza de los puntos singulares de curvas algebraicas reales planas, vamos a comprobar de modo alternativo que la función $h_1(x, y) := \sigma_3(\mathbf{A}(x, y)) = \sigma_3((x + yi)I_3 - A)$ tiene un máximo relativo estricto en $(-\frac{4}{15}, 0)$.

Antes vamos a recapitular brevemente los resultados recogidos en el artículo de revisión [39] sobre derivadas direccionales por la derecha de funciones definidas mediante los valores singulares de matrices que dependen de varias variables reales. Sea $\varphi : \Omega \rightarrow \mathbb{R}$ una función definida en un subconjunto abierto Ω de \mathbb{R}^p con valores reales. Sean el punto $\mathbf{x}_0 \in \Omega$ y $\mathbf{d} \in \mathbb{R}^p$ un vector unitario en la norma euclídea. Llamaremos **derivada direccional por la derecha** de φ en \mathbf{x}_0 respecto de \mathbf{d} al límite

$$\varphi'_+(\mathbf{x}_0, \mathbf{d}) := \lim_{h \rightarrow 0^+} \frac{\varphi(\mathbf{x}_0 + h\mathbf{d}) - \varphi(\mathbf{x}_0)}{h}$$

si este límite existe. Al límite más restrictivo

$$\varphi'(\mathbf{x}_0, \mathbf{d}) := \lim_{h \rightarrow 0} \frac{\varphi(\mathbf{x}_0 + h\mathbf{d}) - \varphi(\mathbf{x}_0)}{h}$$

lo llamaremos **derivada direccional** de φ en \mathbf{x}_0 respecto de \mathbf{d} . Sea $\mathbf{A} : \Omega \rightarrow \mathbb{C}^{n \times n}$ una función matricial de clase C^1 definida sobre el abierto $\Omega \subset \mathbb{R}^p$. Para cada $\mathbf{x} \in \Omega$, denotemos por $s_k(\mathbf{x})$, $k = 1, \dots, n$ los valores singulares de la matriz $\mathbf{A}(\mathbf{x})$ ordenados en sentido decreciente. Obtenemos así n funciones $s_k : \Omega \rightarrow \mathbb{R}$,

$k = 1, \dots, n$. Hiriart-Urruty y Ye (también J.G. Sun) probaron que para todo $\mathbf{x}_0 \in \Omega$ y para cada vector unitario $\mathbf{d} \in \mathbb{R}^p$ existe la derivada direccional por la derecha de s_k en \mathbf{x}_0 respecto de \mathbf{d}

$$s'_{k+}(\mathbf{x}_0, \mathbf{d})$$

y probaron que era un determinado valor propio de una matriz hermítica $F(\mathbf{x}_0, \mathbf{d})$, que depende de \mathbf{x}_0, \mathbf{d} y de las derivadas parciales primeras de la función \mathbf{A} en \mathbf{x}_0 . Véanse los Teoremas 5 y 6 de Gracia [39]. Ahora diremos sucintamente cómo hallar $F(\mathbf{x}_0, \mathbf{d})$ y cuál de sus valores propios se trata. Por el Lema de Wielandt sabemos que los valores propios de la matriz hermítica

$$M(\mathbf{x}) := \begin{pmatrix} \mathbf{O} & \mathbf{A}(\mathbf{x}) \\ \mathbf{A}(\mathbf{x})^* & \mathbf{O} \end{pmatrix}$$

son

$$s_1(\mathbf{x}) \geq \dots \geq s_n(\mathbf{x}) \geq -s_n(\mathbf{x}) \geq \dots - s_1(\mathbf{x}).$$

Escribamos esta lista con más detalle donde se pone la atención sobre el valor singular k -ésimo de $\mathbf{A}(\mathbf{x}_0)$. Ya sabemos que para todo vector unitario $\mathbf{d} = (d_1, \dots, d_p) \in \mathbb{R}^p$ la derivada direccional por la derecha de la función $s_k : \Omega \rightarrow \mathbb{R}$ en \mathbf{x}_0 respecto de \mathbf{d} , $s'_{k+}(\mathbf{x}_0, \mathbf{d})$, existe. Para determinar su valor hay que precisar cuántos valores singulares de $\mathbf{A}(\mathbf{x}_0)$ iguales a $s_k(\mathbf{A}(\mathbf{x}_0))$ existen delante y cuántos detrás de su posición. Supongamos que

$$\begin{aligned} s_1(\mathbf{x}_0) &\geq \dots \geq s_{k-i_k}(\mathbf{x}_0) > s_{k-i_k+1}(\mathbf{x}_0) = \dots = s_k(\mathbf{x}_0) \\ &= s_{k+1}(\mathbf{x}_0) = \dots = s_{k+j_k}(\mathbf{x}_0) > s_{k+j_k+1}(\mathbf{x}_0) \geq \dots \geq s_n(\mathbf{x}_0). \end{aligned} \quad (7.5.9)$$

Esto quiere decir que $s_k(\mathbf{x}_0)$ es un valor singular de multiplicidad $r_k := i_k + j_k$, siendo i_k el número de valores singulares iguales a $s_k(\mathbf{x}_0)$ situados delante de la posición $k+1$, y j_k el número de valores singulares iguales a $s_k(\mathbf{x}_0)$ situados detrás de la posición k . Obsérvese que $i_k \geq 1$ y que j_k podría ser cero. Cuando $k = 1$, tenemos que $i_1 = 1, j_1 = r_1 - 1$. En el caso de que $s_k(\mathbf{x}_0)$ sea simple, $i_k = 1, j_k = 0$. Cuando $k = n$, que es el caso que va a ocurrir en el Ejemplo que sigue, tenemos que $i_n = r_n$ y $j_n = 0$. Aunque la notación no lo indica los números i_k, j_k y r_k dependen de \mathbf{x}_0 .

Sea

$$\mathbf{H} := M(\mathbf{x}_0).$$

Por el Lema de Wielandt, $s_k(\mathbf{x}_0)$ es un valor propio de multiplicidad r_k de la matriz hermítica \mathbf{H} . Como \mathbf{H} es diagonalizable, tenemos que r_k es igual a la dimensión del subespacio propio asociado a este valor propio:

$$r_k = \dim \text{Ker} (s_k(\mathbf{x}_0)\mathbf{I}_{2n} - \mathbf{H}).$$

Sea $\mathbf{v}_1, \dots, \mathbf{v}_{r_k} \in \mathbb{C}^{2n \times 1}$ una base de este subespacio. Mediante el método de ortogonalización de Gram-Schmidt obtenemos una base ortonormal $\mathbf{u}_1, \dots, \mathbf{u}_{r_k} \in \mathbb{C}^{2n \times 1}$ del mismo. Formemos la matriz W_2 que los tiene por columnas

$$W_2 := [\mathbf{u}_1, \dots, \mathbf{u}_{r_k}] \in \mathbb{C}^{2n \times r_k};$$

partamos esta matriz en la forma

$$\begin{bmatrix} U_2 \\ V_2 \end{bmatrix}$$

donde $U_2, V_2 \in \mathbb{C}^{n \times r_k}$.

Del Corolario 7 de [39] se sigue que: *Para cada vector unitario $\mathbf{d} = (d_1, \dots, d_p) \in \mathbb{R}^p$ se tiene que*

$$s'_{k+}(\mathbf{x}_0, \mathbf{d}) = \mu_{i_k},$$

siendo μ_{i_k} el valor propio de la matriz hermitica de orden r_k

$$G(\mathbf{d}) := U_1^* \left(\sum_{j=i}^p d_j \frac{\partial A}{\partial x_j}(\mathbf{x}_0) \right) V_2 + \left(U_1^* \left(\sum_{j=i}^p d_j \frac{\partial A}{\partial x_j}(\mathbf{x}_0) \right) V_2 \right)^*$$

que ocupa la posición i_k -ésima en la ordenación decreciente

$$\mu_1 \geq \dots \geq \mu_{r_k}$$

de los valores propios de esta matriz.

Considerando la matriz B definida en (7.4.2) como $\mathbf{A}(x, y) := (x + yi)I_3 - A$, entonces

$$\frac{\partial \mathbf{A}}{\partial x}(x, y) = I_3, \quad \frac{\partial \mathbf{A}}{\partial y}(x, y) = iI_3,$$

Sea $\mathbf{A}_0 := \mathbf{A}(-4/15, 0)$, llamemos

$$\mathbf{H} := \begin{pmatrix} \mathbf{O} & \mathbf{A}_0 \\ \mathbf{A}_0^* & \mathbf{O} \end{pmatrix} \in \mathbb{R}^{6 \times 6}.$$

El número $1/15$ es valor propio doble de la matriz \mathbf{H} y un par de vectores propios asociados a él linealmente independientes son

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -4 \\ 0 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 4 \\ -15 \\ -4 \\ -16 \\ 0 \\ 1 \end{pmatrix}$$

Así pues, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base del subespacio propio $\text{Ker}((1/15)I_6 - \mathbf{H})$. Mediante el método de Gram-Schmidt obtenemos una base ortogonal del mismo:

$$\mathbf{w}_1 := \mathbf{v}_1, \quad \mathbf{w}_2 := \mathbf{v}_2 - \frac{\mathbf{w}_1 \cdot \mathbf{v}_2}{\mathbf{w}_1 \cdot \mathbf{w}_1} \mathbf{w}_1.$$

Finalmente, obtenemos una base ortonormal de este subespacio propio:

$$\mathbf{u}_1 := \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \quad \mathbf{u}_2 := \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}.$$

Lo que nos da:

$$\mathbf{u}_1 = \frac{1}{\sqrt{34}} \begin{pmatrix} 1 \\ -4 \\ 0 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{9282}} \begin{pmatrix} 4 \\ 1 \\ -68 \\ -16 \\ -64 \\ 17 \end{pmatrix}.$$

Con las Notaciones que preceden a este Ejemplo tenemos que

$$W_2 := [\mathbf{u}_1, \mathbf{u}_2];$$

es decir, W_2 es la matriz 6×2 cuyas columnas son los vectores \mathbf{u}_1 y \mathbf{u}_2 . Particionando W_2 en dos submatrices $U_2, V_2 \in \mathbb{R}^{3 \times 2}$

$$W_2 = \begin{pmatrix} U_2 \\ V_2 \end{pmatrix},$$

obtenemos

$$U_2 = \begin{pmatrix} \frac{1}{\sqrt{34}} & \frac{4}{\sqrt{9282}} \\ \frac{-4}{\sqrt{34}} & \frac{1}{\sqrt{9282}} \\ 0 & \frac{-68}{\sqrt{9282}} \end{pmatrix}, \quad V_2 = \begin{pmatrix} \frac{-4}{\sqrt{34}} & \frac{-16}{\sqrt{9282}} \\ \frac{1}{\sqrt{34}} & \frac{-64}{\sqrt{9282}} \\ 0 & \frac{17}{\sqrt{9282}} \end{pmatrix}.$$

Por lo dicho más arriba se tiene que la derivada direccional por la derecha

$$h'_{1+}((-4/15, 0), (\cos t, \sin t)) := \lim_{\delta \rightarrow 0^+} \frac{h_1(-4/15 + \delta \cos t, \delta \sin t) - h_1(-4/15, 0)}{\delta}$$

viene dada por el menor valor propio de la matriz 2×2 (aquí $i_3 = 2, r_3 = 2$).

$$G(\cos t, \sin t) := e^{it} U_2^* V_2 + e^{-it} V_2^* U_2.$$

De donde, con ayuda de **Maple 17**, deducimos que

$$h'_{1+}((-4/15, 0), (\cos t, \sin t)) = -\frac{34}{91} \cos t - \frac{5}{182} \sqrt{249 - 24 \cos 2t}.$$

Llamando

$$f(t) := -\frac{34}{91} \cos t - \frac{5}{182} \sqrt{249 - 24 \cos 2t},$$

se tiene que el máximo valor de f en el intervalo $0 \leq t \leq 2\pi$ se alcanza en π y vale $f(\pi) = -1/26$. Véase la Figura 7.10. Esto se deduce viendo que $f'(\pi) = 0$ y $f''(\pi) < 0$. Por tanto, para todo $t \in [0, 2\pi)$

$$h'_{1+}((-4/15, 0), (\cos t, \sin t)) < 0.$$

Como todas las derivadas direccionales por la derecha de h_1 en $(-4/15, 0)$ son negativas, la función h_1 tiene un máximo relativo estricto en $(-4/15, 0)$.

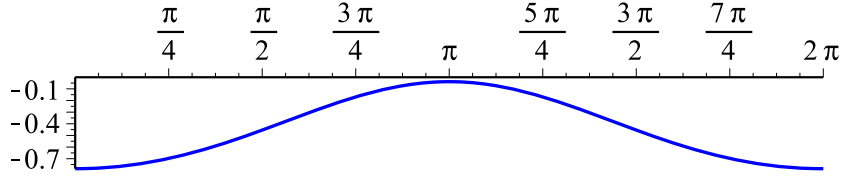


Figura 7.10: Gráfica de la función f sobre el intervalo $0 \leq t \leq 2\pi$.

7.6. Corroboración mediante curvas algebraicas

Otra forma de determinar la naturaleza de los puntos que son solución del sistema (7.4.3) es mediante el Teorema A.5.1. En efecto, cada uno de los pares (x, y) obtenido es un punto singular de la curva $q(x, y, \varepsilon) = 0$, siendo ε la tercera componente en cada una de las ternas solución de (7.4.3). Para cada uno de esos ε , calcularemos las derivadas parciales segundas del polinomio $q(x, y, \varepsilon)$ y, según el signo que tome Δ , clasificaremos el correspondiente punto singular (x, y) .

Las derivadas parciales q''_{xx} , q''_{yy} y q''_{xy} son las siguientes

$$\begin{aligned} q''_{xx}(x, y, \varepsilon) &= -36x^2y^2 - 6y^4 + 36\varepsilon^2 + 36\varepsilon^2x^2 - 30x^4 + 12\varepsilon^2y^2 - 6\varepsilon^4 \\ q''_{yy}(x, y, \varepsilon) &= 36\varepsilon^2 - 6\varepsilon^4 + 36\varepsilon^2y^2 - 6x^4 - 36x^2y^2 - 30y^4 + 12\varepsilon^2x^2 \\ q''_{xy}(x, y, \varepsilon) &= -24x^3y - 24xy^3 + 24\varepsilon^2xy \end{aligned} \tag{7.6.10}$$

Caso: $(x_0, y_0, \varepsilon_0) = (0, 0, 0)$. Este es un caso trivial a la luz de la correspondiente curva algebraica real $q(x, y, 0) = 0$. En efecto,

$$C_0 := \{(x, y) \in \mathbb{R}^2 : q(x, y, 0) = 0\} = \{(x, y) \in \mathbb{R}^2 : -(x^2 + y^2)^3 = 0\} = \{(0, 0)\}.$$

Por lo tanto $(0, 0)$ es un punto aislado de C_0 , pues el único punto de dicha curva. Por tanto, $(0, 0)$ es un punto singular de C_0 . De hecho, sabemos que 0 es el único valor propio de A .

Caso: $(x_1, y_1, \varepsilon_1) = \left(-\frac{4}{15}, 0, \frac{1}{15}\right)$.

Sustituyendo los valores en (7.6.10), tenemos

$$\begin{aligned} A &:= q''_{xx}\left(-\frac{4}{15}, 0, \frac{1}{15}\right) = \frac{22}{1125}, \\ C &:= q''_{yy}\left(-\frac{4}{15}, 0, \frac{1}{15}\right) = \frac{2}{15}, \\ B &:= q''_{xy}\left(-\frac{4}{15}, 0, \frac{1}{15}\right) = 0. \end{aligned}$$

Con lo cual,

$$\Delta := B^2 - AC = -\frac{44}{16875} < 0.$$

Aplicando el apartado II del Teorema A.5.1, se tiene que $\left(-\frac{4}{15}, 0\right)$ es un punto aislado de $q\left(x, y, \frac{1}{15}\right) = 0$.

Caso: $(x_2, y_2, \varepsilon_2) = \left(\frac{-2-\sqrt{22}}{24}, 0, \frac{11\sqrt{22}-50}{24}\right)$.

Sustituyendo los valores en (7.6.10), tenemos

$$A := q''_{xx} \left(\frac{-2-\sqrt{22}}{24}, 0, \frac{11\sqrt{22}-50}{24}\right) = \frac{-91685+19547\sqrt{22}}{144},$$

$$C := q''_{yy} \left(\frac{-2-\sqrt{22}}{24}, 0, \frac{11\sqrt{22}-50}{24}\right) = \frac{-1918+409\sqrt{22}}{3},$$

$$B := q''_{xy} \left(\frac{-2-\sqrt{22}}{24}, 0, \frac{11\sqrt{22}-50}{24}\right) = 0.$$

Con lo cual,

$$\Delta := B^2 - AC = -\frac{11(-8335+1777\sqrt{22})(-1918+409\sqrt{22})}{432} > 0.$$

Aplicando el apartado I del Teorema A.5.1, se tiene que $\left(\frac{-2-\sqrt{22}}{24}, 0\right)$ es un punto crunodo o nodo ordinario de $q\left(x, y, \frac{11\sqrt{22}-50}{24}\right) = 0$.

Caso: $(x_3, y_3, \varepsilon_3) = \left(\frac{-2+\sqrt{22}}{24}, 0, \frac{11\sqrt{22}+50}{24}\right)$.

Sustituyendo los valores en (7.6.10), tenemos

$$A := q''_{xx} \left(\frac{-2+\sqrt{22}}{24}, 0, \frac{11\sqrt{22}+50}{24}\right) = \frac{-91685-19547\sqrt{22}}{144},$$

$$C := q''_{yy} \left(\frac{-2+\sqrt{22}}{24}, 0, \frac{11\sqrt{22}+50}{24}\right) = \frac{-1918-409\sqrt{22}}{3},$$

$$B := q''_{xy} \left(\frac{-2+\sqrt{22}}{24}, 0, \frac{11\sqrt{22}+50}{24}\right) = 0.$$

Con lo cual,

$$\Delta := B^2 - AC = -\frac{11(8335+1777\sqrt{22})(1918+409\sqrt{22})}{432} < 0.$$

Aplicando el apartado II del Teorema A.5.1, se tiene que $\left(\frac{-2+\sqrt{22}}{24}, 0\right)$ es un punto aislado de $q\left(x, y, \frac{11\sqrt{22}+50}{24}\right) = 0$.

7.7. Alternativa usando una fórmula de Lippert

Una forma alternativa de probar que la función $h_1(x, y) = \sigma_n((x+yi)I_n - A)$ tiene un máximo relativo estricto en $(-4/45, 0)$ puede ser basada en el lema siguiente de Lippert [51], Lema A.5. Para la notaciones sobre i_k, j_k seguimos lo indicado en (7.5.9).

Lema 7.7.1 (Lippert, Hiriart-Urruty et al.). *Dada una matriz $A \in \mathbb{C}^{n \times n}$, sea $\sigma_k(A) > 0$ un valor singular de multiplicidad $r_k = i_k + j_k$; es decir*

$$\begin{aligned} \sigma_1(A) &\geq \sigma_2(A) \geq \dots \geq \sigma_{k-i_k}(A) > \sigma_{k-i_k+1} = \dots = \sigma_k(A) \\ &= \sigma_{k+1}(A) = \dots = \sigma_{k+j_k} > \sigma_{k+j_k+1}(A) \geq \dots \geq \sigma_n(A). \end{aligned}$$

Sean $U, V \in \mathbb{C}^{n \times r_k}$ matrices ortogonales por columnas, las cuales están formadas por r_k vectores singulares a izquierda y derecha, respectivamente, asociados

a $\sigma_k(A)$. Sea $(x, y) \in \mathbb{R}^2$ tal que $x^2 + y^2 = 1$. Entonces, cuando $t \rightarrow 0^+$, se cumple

$$\sigma_k(t(x + yi)I_n + A) = \sigma_k(A) + \mu_{i_k}(x, y)t + O(t^2),$$

siendo $\mu_{i_k}(x, y)$ el valor propio que ocupa la posición i_k -ésima en la ordenación decreciente de los r_k valores propios de la matriz hermitica

$$\frac{1}{2}((x + yi)U^*V + (x - yi)V^*U).$$

Vamos a aplicar este resultado a la matriz

$$A = \begin{pmatrix} -\frac{4}{15} & 1 & 4 \\ 0 & -\frac{4}{15} & 1 \\ 0 & 0 & -\frac{4}{15} \end{pmatrix},$$

cuyos valores singulares son: $\sigma_1(A) = \frac{64}{15} > \sigma_2(A) = \sigma_3(A) = \frac{1}{15}$. Como $\sigma_2(A) = \sigma_3(A)$, podemos elegir dos vectores singulares unitarios a izquierda y otros dos a derecha, asociados al valor propio $\frac{1}{15}$, de modo que las matrices U y V , respectivamente, sean ortogonales por columnas.

$$U = \begin{pmatrix} \frac{\sqrt{17}}{17} & \frac{4\sqrt{4641}}{4641} \\ -\frac{4\sqrt{17}}{17} & \frac{\sqrt{4641}}{4641} \\ 0 & -\frac{4\sqrt{4641}}{273} \end{pmatrix}, \quad V = \begin{pmatrix} -\frac{4\sqrt{17}}{17} & -\frac{16\sqrt{4641}}{4641} \\ \frac{\sqrt{17}}{17} & -\frac{64\sqrt{4641}}{4641} \\ 0 & \frac{4\sqrt{4641}}{273} \end{pmatrix}.$$

Aquí $i_3 = 2$ Calculamos los valores propios de la matriz $\frac{1}{2}((x + yi)U^*V + (x - yi)V^*U)$.

$$\begin{aligned} \det\left(\lambda I_2 - \frac{1}{2}((x + yi)U^*V + (x - yi)V^*U)\right) &= \lambda^2 + \frac{68x}{91}\lambda - \frac{11x^2 + 75y^2}{364} \\ &= \lambda^2 + \frac{68x}{91}\lambda - \frac{64x^2 - 75}{364}, \end{aligned}$$

cuyas soluciones son

$$\mu_1(x) = \frac{-68x + 5\sqrt{3(91 - x^2)}}{182}, \quad \mu_2(x) = \frac{-68x - 5\sqrt{3(91 - x^2)}}{182},$$

donde se ha tenido en cuenta que $y^2 = 1 - x^2$. El menor valor propio es $\mu_2(x)$, cuyo máximo en el intervalo $[-1, 1]$ se alcanza en $x = -1$ y vale $-\frac{1}{26}$. Según el Lema 7.7.1, se tiene que $(-4/15, 0)$ es un máximo relativo estricto de h_1 .

7.8. Más sobre el caso $(x_0, y_0, \varepsilon_0) = (0, 0, 0)$

Queremos hacer unos comentarios finales para completar el análisis del caso en que $(x_0, y_0, \varepsilon_0) = (0, 0, 0)$, para el que ya vimos que $(x_0, y_0) = (0, 0)$ es

un punto singular aislado de la curva algebraica $q(x, y, 0) = 0$. Por otra parte sabemos que $0 = 0 + 0i$ es el único valor propio de la matriz

$$A := \begin{pmatrix} 0 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix};$$

esto es, $\sigma_3((x + yi)I_3 - A) = 0$ si y sólo si $(x, y) = (0, 0)$. Por tanto, los pseudoespectros ordinarios y estrictos de nivel 0 de A son $\Lambda_0(A) = \{0\}$, $\Lambda'_0(A) = \emptyset$ y, por ende, $\partial\Lambda_0(A) = \{0\}$, $\partial\Lambda'_0(A) = \emptyset$; que nos da un ejemplo trivial de la relación $\overline{\Lambda'_0(A)} \subsetneq \Lambda_0(A)$. Por otro lado, tenemos el teorema siguiente

Teorema 7.8.1 ([17], Teorema 4.2, página 87). *Sea $B \in \mathbb{C}^{n \times n}$. Sea $h_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función dada por*

$$h_1(x, y) := \sigma_n((x + yi)I_n - B), \quad (x, y) \in \mathbb{R}^2.$$

Entonces $\tilde{x} + \tilde{y}i$ es un valor propio de B si y sólo si la función h_1 tiene un mínimo relativo en (\tilde{x}, \tilde{y}) .

Véase también la Proposición 2.2.4 (b). Ya que para todo $(x, y) \in \mathbb{R}^2$, $h_1(x, y) \geq 0$ y $\tilde{x} + \tilde{y}i$ es un valor propio de A si y sólo si $\sigma_n((x + yi)I_n - A) = 0$, todos los mínimos relativos de h_1 son mínimos absolutos y valen 0. Es decir, en nuestro caso 0 es el único valor propio de A y $h_1(0, 0) = \sigma_3(-A) = 0$. Por consiguiente la curva de nivel cero de la función h_1 ,

$$\{(x, y) \in \mathbb{R}^2 : \sigma_3((x + yi)I_3 - A) = 0\}$$

sólo contiene el punto $(0, 0)$.

Conclusión.- $\Lambda_0(A) = \{(x, y) : h_1(x, y) = 0\} = \{(0, 0)\}$. *Por tanto, $(0, 0)$ es el único punto aislado de $\partial\Lambda_0(A)$.*

Vamos a demostrar que todas las derivadas direccionales por la derecha de h_1 en $(0, 0)$ son iguales a cero. Por lo que todas las derivadas direccionales (en particular las derivadas parciales respecto de x y de y) en $(0, 0)$ valdrán 0. *Sabemos que esta condición es necesaria para un mínimo relativo “suave” en $(0, 0)$; pero no es una condición suficiente.*

Los valores singulares de la matriz $(0 + 0i)I_3 - A$ son: $\sqrt{5} + 2 > \sqrt{5} - 2 > 0$. Por tanto, 0 es un valor singular simple de $-A$, pero ¡ajo! 0 es un valor propio doble de la matriz

$$H := \begin{pmatrix} O & -A \\ -A^* & O \end{pmatrix}$$

Los 6 valores propios de esta matriz H ordenados en sentido decreciente son:

$$\sqrt{5} + 2 > \sqrt{5} - 2 > 0 = 0 > -(\sqrt{5} - 2) > -(\sqrt{5} + 2).$$

Busquemos dos vectores propios de H linealmente independientes asociados a 0 con ayuda de la función `Eigenvectors(H, output='list')` de Maple 17:

$$\mathbf{u}_1 := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Así pues, hemos obtenido una base ortonormal del subespacio propio $\text{Ker}(0I_6 - H)$. Sean

$$\begin{aligned} s_1(x, y) &:= \sigma_1((x + yi)I_3 - A) \geq s_2(x, y) := \sigma_2((x + yi)I_3 - A) \geq \\ s_3(x, y) &:= \sigma_3((x + yi)I_3 - A) \geq -s_3(x, y) := -\sigma_3((x + yi)I_3 - A) \geq \\ &-s_2(x, y) := -\sigma_2((x + yi)I_3 - A) \geq -s_1(x, y) := -\sigma_1((x + yi)I_3 - A) \end{aligned}$$

los 6 valores propios de la matriz

$$\mathbf{H}(x, y) := \begin{pmatrix} \mathbf{O} & (x + yi)I_3 - A \\ (x - yi)I_3 - A & \mathbf{O} \end{pmatrix}$$

ordenados en sentido decreciente. Entonces,

$$s'_{3+}((0, 0), (\cos t, \sin t))$$

es igual al mayor valor propio de la matriz hermitica 2×2 dada por

$$G(\cos t, \sin t) := e^{it}U_2^*V_2 + e^{-it}V_2^*U_2,$$

siendo

$$U_2 := \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad V_2 := \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Pero, la posición de los ceros de estas matrices implica que $G(\cos t, \sin t)$ es la matriz cero O . En consecuencia,

$$s'_{3+}((0, 0), (\cos t, \sin t)) = 0; \quad \implies \quad s'_{3-}((0, 0), (\cos t, \sin t)) = 0$$

es decir, que $h'_1((0, 0), (\cos t, \sin t)) = 0$ para todo $t \in [0, 2\pi)$: En particular, existen las derivadas parciales

$$\frac{\partial h_1}{\partial x}(0, 0), \quad \frac{\partial h_1}{\partial y}(0, 0)$$

y son iguales a cero. Aunque esto no implica que la función h_1 sea diferenciable en $(0, 0)$.

Capítulo 8

Matrices de orden pequeño

Debido a la existencia nítida de formas canónicas de matrices de órdenes dos y tres para la semejanza unitaria, así como la invarianza de las singularidades de la frontera de sus pseudoespectros mediante transformaciones afines que también relacionan los pseudoespectros de una matriz A al sumarle una matriz escalar o multiplicarla por un número complejo, hemos podido realizar un estudio más pormenorizado de la geometría de los pseudoespectros de estas matrices. El caso de matrices de $\mathbb{C}^{2 \times 2}$ ha quedado completamente resuelto en toda su generalidad. Para el caso de matrices de $\mathbb{C}^{3 \times 3}$ hemos resuelto los casos más significativos, no habiendo podido abordarlos en toda su generalidad ni con ayuda del programa de cálculo simbólico `Maple`.

Recordemos que dos matrices $A, B \in \mathbb{C}^{n \times n}$ se dicen unitariamente semejantes si existe una matriz unitaria $U \in \mathbb{C}^{n \times n}$ tal que $B = U^*AU$. Usamos la notación $B \stackrel{u}{\approx} A$ para esta relación de equivalencia en $\mathbb{C}^{n \times n}$. Llamaremos **forma canónica** para $\stackrel{u}{\approx}$ a una función $c : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$ que satisface:

- (I) Para toda $A \in \mathbb{C}^{n \times n}$, $c(A) \stackrel{u}{\approx} A$.
- (II) Si $A \stackrel{u}{\approx} B$, entonces $c(A) = c(B)$.

Si existe una matriz $P \in \mathbb{C}^{n \times n}$ invertible tal que $B = P^{-1}AP$ decimos que A es *semejante* a B y lo denotamos por $A \approx B$. Así como se conocen sistemas completos de invariantes y formas canónicas para \approx , el problema para $\stackrel{u}{\approx}$ tardó más tiempo en resolverse: un sistema completo de invariantes fue obtenido mediante las trazas de las matrices $w(A, A^*)$ donde $w(x, y)$ recorre el conjunto de **todas las palabras** en las variables no conmutativas x e y (Criterio de Specht (1940)). Pero este sistema de infinitos invariantes era redundante. Bastaba con un conjunto finito de ellos. Percy (1962) probó que eran suficientes las trazas de $w(A, A^*)$ para todas las palabras $w(x, y)$ de grado menor o igual que $2n^2$. Ya que el número de palabras de grado $\leq 2n^2$ es 4^{n^2} , esto da un sistema completo de invariantes que contiene 4^{n^2} elementos. La cota superior del número de palabras necesarias fue rebajada a $\lceil \frac{(2n^2+4)}{3} \rceil$ por Laffey (1986) y por Pappacena (1997) a

$n \cdot \sqrt{\frac{2n^2}{n-1} + \frac{1}{4}} + \frac{n}{2} - 2$. Como veremos en la Sección 1, para las matrices 2×2 el sistema $(\text{tr}(A), \text{tr}(A^2), \text{tr}(A^*A))$ es un sistema completo de invariantes minimal. Para las matrices 3×3 , Sibirskii [67], [68] dio el sistema

$$(\text{tr}(A), \text{tr}(A^2), \text{tr}(A^3), \text{tr}(AA^*), \text{tr}(A^2A^*), \text{tr}(A^2(A^*)^2), \text{tr}(A^2(A^*)^2AA^*)), \quad (8.0.1)$$

que resultó ser un sistema completo de invariantes minimal. Es decir, en la colección de palabras

$$x, x^2, x^3, xy, x^2y, x^2y^2, x^2y^2xy$$

no hay ninguna redundante. Véase también Đoković *et al.* [57]. Teniendo en cuenta que para dos matrices $M, N \in \mathbb{C}^{n \times n}$, $\text{tr}(MN) = \text{tr}(NM)$, la presentación por diversos autores de los invariantes de Sibirskii puede ser aparentemente diferente. Para la matrices 4×4 bastan 20 palabras, pero no se sabe si hay redundancias.

Littlewood (1953) construyó un algoritmo que transforma cualquier matriz $A \in \mathbb{C}^{n \times n}$ en una matriz A_{can} que es unitariamente semejante a A y es tal que $A \stackrel{u}{\approx} B$ si y sólo si $A_{\text{can}} = B_{\text{can}}$. Otros autores también dieron *definiciones inductivas de formas canónicas* para la relación de semejanza unitaria, *pero es difícil ver la forma canónica final*. Como veremos más adelante, Percy (1962) [59] dio una forma canónica para las matrices 3×3 .

8.1. Reducción de matrices 2×2

Cualquier matriz $A \in \mathbb{C}^{2 \times 2}$ se puede representar como una matriz triangular mediante la aplicación del Lema de Schur, es decir, existe $U \in \mathbb{C}^{2 \times 2}$ unitaria t.q.

$$U^*AU = \begin{pmatrix} \lambda_1 & z \\ 0 & \lambda_2 \end{pmatrix}, \text{ donde } \lambda_1, \lambda_2 \in \Lambda(A) \text{ y } z = re^{i\alpha}, \text{ con } r \geq 0, \alpha \in [0, 2\pi].$$

Además, se puede convertir a z en un número real no negativo. Para probar esto, es suficiente elegir la matriz unitaria

$$V = \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}, \text{ donde } \alpha = \arg(z),$$

y aplicar el Teorema 1.2.11. En efecto,

$$V^* \begin{pmatrix} \lambda_1 & re^{i\alpha} \\ 0 & \lambda_2 \end{pmatrix} V = \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix}.$$

Es más, podemos expresar el número r en función de los valores propios de la matriz A y de la traza de A^*A , para ello realizamos el siguiente cálculo

$$A^*A = \begin{pmatrix} \bar{\lambda}_1 & 0 \\ r & \bar{\lambda}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & \bar{\lambda}_1 r \\ \lambda_1 r & |\lambda_2|^2 + r^2 \end{pmatrix}.$$

La traza de esa matriz es igual a $\text{tr}(A^*A) = |\lambda_1|^2 + |\lambda_2|^2 + r^2$, de donde se tiene que

$$r = \sqrt{\text{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}.$$

La traza de A^*A está relacionada con la norma de Frobenius. Por tanto, la expresión para el número r se puede escribir como

$$r = \sqrt{\|A\|_F^2 - |\lambda_1|^2 - |\lambda_2|^2}.$$

Ese número r no cambia bajo transformaciones de semejanza unitaria, por lo tanto, es un invariante de la matriz A . De hecho, r junto con los valores propios forman un sistema completo de invariantes de la matriz A para la semejanza unitaria, y determinan unívocamente su forma canónica, excepto permutación de sus valores propios. Alternativamente, podemos dar como sistema completo de invariantes los valores propios y los valores singulares de la matriz. Un sistema completo de invariantes para la semejanza unitaria en $\mathbb{C}^{2 \times 2}$ está formado por las trazas de A , A^2 y A^*A : $(\text{tr}(A), \text{tr}(A^2), \text{tr}(A^*A))$. Llamando λ_1, λ_2 a los dos valores propios de A , iguales o no, $\text{tr}(A) = \lambda_1 + \lambda_2$, $\text{tr}(A^2) = \lambda_1^2 + \lambda_2^2$. Por lo tanto, en virtud de la relación de las sumas de Newton con las funciones simétricas elementales los valores de $\lambda_1 + \lambda_2$ y $\lambda_1^2 + \lambda_2^2$ determinan unívocamente $\lambda_1 + \lambda_2$ y $\lambda_1\lambda_2$; y por tanto λ_1 y λ_2 . Así pues $\text{tr}(A)$ y $\text{tr}(A^2)$ determinan λ_1 y λ_2 . Finalmente, las trazas $\text{tr}(A)$, $\text{tr}(A^2)$, $\text{tr}(A^*A)$ determinan r .

El resultado siguiente caracteriza los puntos singulares de las curvas algebraicas reales

$$V_{\mathbb{R}}(f_{\varepsilon,A}) := \{(x, y) \in \mathbb{R}^2 : f_{\varepsilon,A}(x, y) = 0\},$$

siendo $f_{\varepsilon,A}(x, y)$ el polinomio definido en (1.2.2)

$$f_{\varepsilon,A}(x, y) := \det \begin{pmatrix} \varepsilon I_2 & -(x + yi)I_2 + A \\ -(x - yi)I_2 + A^* & \varepsilon I_2 \end{pmatrix}, \quad (8.1.2)$$

donde $\varepsilon \geq 0$ y $A \in \mathbb{C}^{2 \times 2}$ tiene dos valores propios distintos.

Teorema 8.1.1. *Sea $A \in \mathbb{C}^{2 \times 2}$ con dos valores propios distintos, $\lambda_1 = x_1 + y_1i$ y $\lambda_2 = x_2 + y_2i$, donde $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, 2$, y $x_1 \geq x_2$. Sean*

$$\begin{aligned} \delta(A) &:= \sqrt{\text{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}, \\ \varepsilon_0(A) &:= \frac{-\delta(A) + \sqrt{4|\lambda_2 - \lambda_1|^2 + \delta(A)^2}}{2}, \\ \varepsilon_1(A) &:= \frac{\delta(A) + \sqrt{4|\lambda_2 - \lambda_1|^2 + \delta(A)^2}}{2}. \end{aligned} \quad (8.1.3)$$

Sea (x_0, y_0) el punto medio del segmento que une los puntos (x_1, y_1) y (x_2, y_2) . Entonces

$$1. \ \varepsilon_0(A) = \sigma_2((x_0 + y_0i)I_2 - A) \ \text{y} \ \varepsilon_1(A) = \sigma_1((x_0 + y_0i)I_2 - A).$$

2. La curva $V_{\mathbb{R}}(f_{0,A})$ consta de dos puntos (x_1, y_1) y (x_2, y_2) , que son aislados.
3. Supongamos que la matriz A es normal. Entonces, la curva $V_{\mathbb{R}}(f_{\varepsilon,A})$, con $\varepsilon > 0$, no tiene puntos singulares si $\varepsilon < \varepsilon_0(A)$. La curva $V_{\mathbb{R}}(f_{\varepsilon_0(A),A})$ tiene un tacnodo en el punto (x_0, y_0) , su único punto singular. Para todo $\varepsilon > \varepsilon_1(A)$, los únicos puntos singulares de la curva $V_{\mathbb{R}}(f_{\varepsilon,A})$ vienen dados por

$$(x_0, y_0) \pm \frac{\sqrt{4\varepsilon^2 - |\lambda_2 - \lambda_1|^2}}{2|\lambda_2 - \lambda_1|} (y_1 - y_2, x_2 - x_1) \quad (8.1.4)$$

que son nodos con dos tangentes reales.

4. Si A es no normal, la curva $V_{\mathbb{R}}(f_{\varepsilon,A})$, con $\varepsilon > 0$, no tiene puntos singulares, excepto para $\varepsilon_0(A)$ y $\varepsilon_1(A)$. Además (x_0, y_0) es un nodo con dos tangentes reales de la curva $V_{\mathbb{R}}(f_{\varepsilon_0(A),A})$; y (x_0, y_0) es un punto aislado de $V_{\mathbb{R}}(f_{\varepsilon_1(A),A})$.
5. Para cualquier ε tal que $0 \leq \varepsilon < \varepsilon_0(A)$, $\Lambda_{\varepsilon}(A)$ tiene dos componentes conexas y $\Lambda_{\varepsilon_0(A)}(A)$ tiene sólo una.

DEMOSTRACIÓN.

Para simplificar la demostración vamos a realizar unas transformaciones sobre la matriz A . En primer lugar denotemos por $\lambda_0 := x_0 + y_0i$, el punto medio de λ_1 y λ_2 . Como $\lambda_1 \neq \lambda_2$, sea

$$\frac{\lambda_1 - \lambda_2}{2} = \mu e^{\theta i}, \quad \mu > 0.$$

Ahora, sea $A_1 := e^{-\theta i}(A - \lambda_0 I_2)$. Notemos que los valores propios de A_1 son $\{-\mu, \mu\}$, cuyo punto medio es el origen. Además, a partir de (8.1.3) se deduce que $\delta(A_1) = \delta(A)$, $\varepsilon_0(A_1) = \varepsilon_0(A)$ y $\varepsilon_1(A_1) = \varepsilon_1(A)$. Ahora, existe una matriz unitaria $U \in \mathbb{C}^{2 \times 2}$ tal que

$$B := U^* A_1 U = \begin{pmatrix} \mu & r \\ 0 & -\mu \end{pmatrix}, \quad \mu > 0, r = \delta(A) \geq 0. \quad (8.1.5)$$

Luego $A = U(e^{\theta i} B + \lambda_0 I_2)U^*$, con $\delta(B) = \delta(A)$, $\varepsilon_0(B) = \varepsilon_0(A)$ y $\varepsilon_1(B) = \varepsilon_1(A)$. Veamos la relación entre las curvas $V_{\mathbb{R}}(f_{\varepsilon,A})$ y $V_{\mathbb{R}}(f_{\varepsilon,B})$. Para ello, por simplificación, vamos a identificar \mathbb{R}^2 con \mathbb{C} . Así entonces, si $z = x + yi$, diremos que $z \in V_{\mathbb{R}}(f_{\varepsilon,A})$ si $(x, y) \in V_{\mathbb{R}}(f_{\varepsilon,A})$. De igual manera $f_{\varepsilon,A}(z) := f_{\varepsilon,A}(x, y)$. Ahora, operando en (8.1.2) obtenemos

$$\begin{aligned} f_{\varepsilon,A}(z) &= \det(\varepsilon^2 I_2 - (A - z I_2)(A - z I_2)^*) \\ &= \det(\varepsilon^2 I_2 - (B - e^{-\theta i}(z - \lambda_0) I_2)(B - e^{-\theta i}(z - \lambda_0) I_2)^*) \\ &= f_{\varepsilon,B}(e^{-\theta i}(z - \lambda_0)). \end{aligned}$$

Luego $V_{\mathbb{R}}(f_{\varepsilon,A}) = \lambda_0 + e^{\theta i} V_{\mathbb{R}}(f_{\varepsilon,B})$. Es decir, la curva $V_{\mathbb{R}}(f_{\varepsilon,A})$ se puede obtener a partir de $V_{\mathbb{R}}(f_{\varepsilon,B})$ mediante un giro y una traslación. Así, para probar el teorema

no se pierde generalidad si suponemos que la matriz A es de la forma (8.1.5). Pasamos a demostrar el teorema.

Notemos que de (8.1.3) y (8.1.5), resulta $\delta(A) = r$ y

$$\begin{aligned} \varepsilon_0(A) &= \sigma_2((0 + 0i)I_2 - A) = \frac{-r + \sqrt{4\mu^2 + r^2}}{2}, \\ \varepsilon_1(A) &= \sigma_2((0 + 0i)I_2 - A) = \frac{r + \sqrt{4\mu^2 + r^2}}{2}. \end{aligned} \quad (8.1.6)$$

Pasamos a probar los apartados 3 y 4, es decir dada la familia de curvas algebraicas reales

$$V_{\mathbb{R}}(f_{\varepsilon}, A) := \{(x, y) \in \mathbb{R}^2 \mid f_{\varepsilon, A}(x, y) = 0\},$$

queremos encontrar los valores de ε para los cuales hay puntos singulares, es decir, puntos $(x, y) \in V_{\mathbb{R}}(f_{\varepsilon, A})$ que satisfacen el sistema

$$f_{\varepsilon, A}(x, y) = 0, \quad \frac{\partial f_{\varepsilon, A}}{\partial x}(x, y) = 0, \quad \frac{\partial f_{\varepsilon, A}}{\partial y}(x, y) = 0. \quad (8.1.7)$$

Para ello, observemos en primer lugar que de la Ecuación (8.1.2) se tiene que

$$f_{\varepsilon, A}(x, y) = (x^2 + y^2)^2 - 2\varepsilon^2(x^2 + y^2) + 2\mu^2(y^2 - x^2) - r^2\varepsilon^2 + (\varepsilon^2 - \mu^2)^2. \quad (8.1.8)$$

Derivando

$$\begin{cases} \frac{\partial f_{\varepsilon, A}}{\partial x}(x, y) = 4x(x^2 + y^2 - \varepsilon^2 - \mu^2) = 0, \\ \frac{\partial f_{\varepsilon, A}}{\partial y}(x, y) = 4y(x^2 + y^2 - \varepsilon^2 + \mu^2) = 0. \end{cases}$$

Luego, utilizando (8.1.8), el sistema anterior y simplificando, el sistema (8.1.7) es equivalente a

$$\begin{cases} (x^2 + y^2)^2 - 2\varepsilon^2(x^2 + y^2) + 2\mu^2(y^2 - x^2) - r^2\varepsilon^2 + (\varepsilon^2 - \mu^2)^2 = 0, \\ x(x^2 + y^2 - \varepsilon^2 - \mu^2) = 0, \\ y(x^2 + y^2 - \varepsilon^2 + \mu^2) = 0. \end{cases} \quad (8.1.9)$$

De la segunda ecuación, se deduce que si $x \neq 0$ entonces $x^2 + y^2 = \varepsilon^2 + \mu^2$. Es decir

$$y = \pm \sqrt{\varepsilon^2 + \mu^2 - x^2}.$$

Sustituyendo, el sistema (8.1.9) se reduce a

$$\begin{cases} -4\mu^2 x^2 - r^2 \varepsilon^2 + 4\mu^4 = 0, \\ \pm \mu^2 \sqrt{\varepsilon^2 + \mu^2 - x^2} = 0. \end{cases}$$

De la primera ecuación, deducimos

$$x^2 = \frac{4\mu^4 - r^2\varepsilon^2}{4\mu^2},$$

que puesta en la segunda, resulta $\pm\varepsilon\mu\sqrt{r^2 + 4\mu^2} = 0$, lo que es absurdo. Luego se ha de cumplir $x = 0$. Así, el sistema (8.1.9) queda reducido al sistema

$$\begin{cases} \varepsilon^4 - \varepsilon^2(r^2 + 2y^2 + 2\mu^2) + (y^2 + \mu^2)^2 = 0, \\ -y\varepsilon^2 + y(y^2 + \mu^2) = 0. \end{cases} \quad (8.1.10)$$

Para resolver este sistema, es decir las raíces comunes a ambos polinomios en las variables ε, y , utilizamos la resultante de Sylvester de esos polinomios en la variable ε , con coeficientes en $\mathbb{R}[y]$ que viene dado por :

$$\mathcal{S} := \begin{vmatrix} 1 & 0 & p_1(y) & 0 & p_2(y) & 0 \\ 0 & 1 & 0 & p_1(y) & 0 & p_2(y) \\ -y & 0 & q(y) & 0 & 0 & 0 \\ 0 & -y & 0 & q(y) & 0 & 0 \\ 0 & 0 & -y & 0 & q(y) & 0 \\ 0 & 0 & 0 & -y & 0 & q(y) \end{vmatrix},$$

donde

$$p_1(y) = -(r^2 + 2y^2 + 2\mu^2), \quad p_2(y) = (y^2 + \mu^2)^2, \quad q(y) = y(y^2 + \mu^2).$$

Las soluciones de (8.1.10) vienen dadas por los valores de y que anulan dicho determinante. Operando, resulta $\mathcal{S} = r^4 y^4 (y^2 + \mu^2)^2$. Luego tenemos dos posibilidades: si $y = 0$, resolviendo en (8.1.10) obtenemos $\varepsilon = (\pm r \sqrt{4\mu^2 + r^2})/2$. Si $y \neq 0$, entonces $r = 0$, es decir A es una matriz normal. Para este caso deducimos de (8.1.10) que $y = \pm \sqrt{\varepsilon^2 - \mu^2}$. Por lo tanto el sistema tiene solución si $\varepsilon \geq \mu$. En la siguiente tabla resumimos todos los puntos singulares (x, y) y los correspondientes valores de ε de la familia de curvas $V_{\mathbb{R}}(f_{\varepsilon, A})$, dependiendo de si $r > 0$ ó $r = 0$, matriz normal.

$$\begin{array}{lll} \text{Si } r > 0, & x = 0, & y = 0, & \varepsilon = \frac{-r + \sqrt{4\mu^2 + r^2}}{2}. \\ \text{Si } r > 0, & x = 0, & y = 0, & \varepsilon = \frac{r + \sqrt{4\mu^2 + r^2}}{2}. \\ \text{Si } r = 0, & x = 0, & y = -\sqrt{\varepsilon^2 - \mu^2}, & \varepsilon \geq \mu. \\ \text{Si } r = 0, & x = 0, & y = \sqrt{\varepsilon^2 - \mu^2}, & \varepsilon \geq \mu. \end{array}$$

Ahora, como de (8.1.5) se tiene que que $\varepsilon_0(A) = \frac{-r + \sqrt{4\mu^2 + r^2}}{2}$, deducimos que el único punto singular de la curva $V_{\mathbb{R}}(f_{\varepsilon_0, A})$ es $(0, 0)$.

Una vez determinados los puntos singulares, los clasificamos utilizando el

Teorema A.5.1. Para ello calculamos las derivadas parciales segundas de $f_{\varepsilon,A}(x, y)$.

$$\begin{aligned}\frac{\partial^2 f_{\varepsilon,A}}{\partial x^2}(x, y) &= -4\varepsilon^2 + 4(3x^2 + y^2 - \mu^2), \\ \frac{\partial^2 f_{\varepsilon,A}}{\partial y^2}(x, y) &= -4\varepsilon^2 + 4(x^2 + 3y^2 + \mu^2), \\ \frac{\partial^2 f_{\varepsilon,A}}{\partial x \partial y}(x, y) &= 8xy.\end{aligned}\tag{8.1.11}$$

Denotemos por

$$\Delta_\varepsilon(x, y) := \left(\frac{\partial^2 f_{\varepsilon,A}}{\partial x \partial y}(x, y) \right)^2 - \frac{\partial^2 f_{\varepsilon,A}}{\partial x^2}(x, y) \frac{\partial^2 f_{\varepsilon,A}}{\partial y^2}(x, y).\tag{8.1.12}$$

Ahora, a partir de (8.1.11) y (8.1.12), si $r > 0$, resulta

$$\Delta_\varepsilon(0, 0) = 16(\mu^2 + \varepsilon^2)(\mu^2 - \varepsilon^2).$$

Por (8.1.6), como $\mu^2 - \varepsilon_0(A)^2 = r(\sqrt{4\mu^2 + r^2} - r)/2 > 0$, y por el Teorema A.5.1, deducimos que $(0, 0)$ es un nodo con dos rectas tangentes de la curva $V_{\mathbb{R}}(f_{\varepsilon_0(A),A})$. De igual manera, al ser $\mu^2 - \varepsilon_1(A)^2 = -r(\sqrt{4\mu^2 + r^2} + r)/2 < 0$, por el teorema A.5.1 se tiene que $(0, 0)$ es punto aislado de la curva $V_{\mathbb{R}}(f_{\varepsilon_1(A),A})$.

Cuando $r = 0$, como en este caso $\varepsilon_0 = \mu$, entonces $\Delta_\varepsilon(0, 0) = 0$. En consecuencia, del Teorema A.5.6 deducimos que $(0, 0)$ es un punto singular con tangente única para $V_{\mathbb{R}}(f_{\varepsilon_0(A),A})$. Falta por probar que es un tacnodo, para ello calculamos el número de intersección en $(0, 0)$ del polinomio (8.1.8) con la recta

$$\ell = \begin{cases} x = 0 + \alpha t, \\ y = 0 + \beta t. \end{cases}$$

Así, tenemos lo siguiente:

$$f_{\varepsilon_0(A),A}(\alpha t, \beta t) = t^2(t(\alpha^2 + \beta^2) + 2\alpha\mu)(t(\alpha^2 + \beta^2) - 2\alpha\mu)$$

Como se puede comprobar $I((0, 0), f_{\varepsilon_0(A),A}, \ell)$ es igual a 2 siempre que $\alpha \neq 0$; sin embargo, para $\alpha = 0$, $I((0, 0), f_\varepsilon, T) = 4$, siendo T la recta tangente que pasa por $(0, 0)$. Con lo cual, se puede afirmar que $(0, 0)$ es un punto tacnodo. Para concluir con la demostración del apartado 3 notemos que de (8.1.10) y (8.1.12) resulta $\Delta_\varepsilon(0, \pm\sqrt{\varepsilon^2 - \mu^2}) = 64\mu^2(\varepsilon^2 - \mu^2) > 0$, para $\varepsilon > \mu$. Luego los puntos $(0, \sqrt{\varepsilon^2 - \mu^2})$ y $(0, -\sqrt{\varepsilon^2 - \mu^2})$ son nodos con dos rectas tangentes de la curva $V_{\mathbb{R}}(f_{\varepsilon,A})$.

Hay dos componentes conexas de $\Lambda_\varepsilon(A)$ para valores suficientemente pequeños de ε . Véase [44, Prop. 5.5.5]. Cuando $\varepsilon = \varepsilon_0(A)$ el pseudoespectro $\Lambda_{\varepsilon_0(A)}(A)$ es conexo y contiene a $(0, 0)$. \square

Observación 8.1.2. Observemos que cuando la matriz A es normal, es decir $\delta(A) = 0$, entonces existe una matriz unitaria $U \in \mathbb{C}^{2 \times 2}$ tal que $U^*AU = \text{diag}(\lambda_1, \lambda_2)$. A partir de la Ecuación (8.1.2) y con las notaciones del Teorema 8.1.1 resulta que

$$f_{\varepsilon, A}(x, y) = ((x - x_1)^2 + (y - y_1)^2 - \varepsilon^2)((x - x_2)^2 + (y - y_2)^2 - \varepsilon^2).$$

Por lo tanto la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ es igual a la unión de dos circunferencias de radio ε con centros en (x_1, y_1) y (x_2, y_2) . En consecuencia, los puntos singulares son los puntos intersección de dichas circunferencias. En concreto, si $\varepsilon := \frac{1}{2}\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$, por lo dicho en la Subsección A.7.2.1, la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ tiene un punto doble *tacnodo* en $(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$. Para valores de ε mayores que éste la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ tiene dos *crunodos* situados en la mediatriz del segmento que une los centros de las circunferencias.

En el siguiente teorema se analiza la frontera del pseudoespectro para una matriz de orden dos con valores propios distintos, en particular si es una curva algebraica o parte de una curva algebraica. Agradecemos la ayuda recibida de M.J. de la Puente [26] y también hemos usado su conocido libro [25].

Teorema 8.1.3. *Bajo las hipótesis del Teorema 8.1.1,*

1. *si la matriz A es no normal, el polinomio $f_{\varepsilon, A} \in \mathbb{R}[x, y]$ es irreducible para todo $\varepsilon > 0$;*
2. *sea $\varepsilon_1 := \sigma_1((x_0 + y_0i)I_2 - A)$, entonces $\partial\Lambda'_\varepsilon(A) \subsetneq V_{\mathbb{R}}(f_{\varepsilon, A})$ para todo $\varepsilon \geq \varepsilon_1$.*

DEMOSTRACIÓN.

(1) Observemos que no se pierde generalidad si suponemos que la matriz A está en la forma (8.1.5). Por lo tanto, a partir de (8.1.8), se tiene que

$$f_{\varepsilon, A}(x, y) = (x^2 + y^2)^2 - 2\varepsilon^2(x^2 + y^2) + 2\mu^2(y^2 - x^2) - r^2\varepsilon^2 + (\varepsilon^2 - \mu^2)^2.$$

Si $f_{\varepsilon, A}$ es reducible, entonces existen dos polinomios $p, q \in \mathbb{R}[x, y]$, de grado ≥ 1 , tales que $f_{\varepsilon, A}(x, y) = p(x, y)q(x, y)$. Ahora, como la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ es acotada, como $\deg f_{\varepsilon, A} = 4$, entonces $\deg p = \deg q = 2$. Luego

$$\begin{aligned} p(x, y) &= a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{10}x + a_{01}y + a_{00}, \\ q(x, y) &= b_{20}x^2 + b_{11}xy + b_{02}y^2 + b_{10}x + b_{01}y + b_{00}. \end{aligned}$$

Ahora, como las curvas $V_{\mathbb{R}}(p)$ y $V_{\mathbb{R}}(q)$ son acotadas, entonces los signos de a_{20} y a_{02} coinciden y ambos coeficientes son no nulos; lo mismo ocurre para b_{20} y b_{02} . Observando la expresión de $f_{\varepsilon, A}$, como $f_{\varepsilon, A}(x, y) = p(x, y)q(x, y)$, no se pierde generalidad si suponemos que $a_{20} = b_{20} = 1$. Notemos que $b_{02} = 1/a_{02}$ Así, podemos suponer que

$$p(x, y) = x^2 + axy + by^2 + cx + dy + e, \quad q(x, y) = x^2 + fxy + y^2/b + gx + hy + k,$$

con b positivo. Ahora operando en $f_{\varepsilon,A}(x, y) = p(x, y)q(x, y)$, obtenemos inmediatamente que $f = -a$, $g = -c$, $d = -hb^2$, $k = c^2 - e - 2(\varepsilon^2 + \mu^2)$. Además, los coeficientes que acompañan a los monomios x^2y^2 y xy^3 son $a^2b - (b-1)^2$ y $a(b^2 - 1)/b$. Como ambos coeficientes han de ser cero, resolviendo el sistema

$$\begin{cases} a^2b - (b-1)^2 = 0, \\ a(b^2 - 1) = 0, \end{cases}$$

obtenemos que la única solución posible es $x = 0, b = 1$. Luego, con todos los datos obtenidos, de $f_{\varepsilon,A}(x, y) = p(x, y)q(x, y)$ se obtiene el sistema

$$\begin{cases} ch = 0, \\ c(c^2 - 2e - 2(\varepsilon^2 + \mu^2)) = 0, \\ -c^2 + h^2 + 4\mu^2 = 0, \\ h\left(\left(b^2(c^2 - e - 2(\varepsilon^2 + \mu^2)) - e\right) - e\right) = 0, \\ -c^2e + e^2 + 2e(\varepsilon^2 + \mu^2) - r^2\varepsilon^2 + (\varepsilon^2 - \mu^2)^2 = 0. \end{cases}$$

De la primera ecuación $c = 0$ o $h = 0$. Pero si $c = 0$, en la tercera resulta $h^2 + 4\mu^2 = 0$, absurdo puesto que $\mu > 0$. Luego $h = 0$ y $c \neq 0$. Así, el sistema anterior se reduce a

$$\begin{cases} c^2 - 2e - 2(\varepsilon^2 + \mu^2) = 0, \\ -c^2 + 4\mu^2 = 0, \\ -c^2e + e^2 + 2e(\varepsilon^2 + \mu^2) - r^2\varepsilon^2 + (\varepsilon^2 - \mu^2)^2 = 0. \end{cases}$$

Resolviendo, obtenemos $c^2 = 4\mu^2$ y $e = \mu^2 - \varepsilon^2$, y la igualdad $r\varepsilon = 0$. En consecuencia, $r = 0$ lo que contradice que A sea no normal.

Otra prueba, dada por M.J. de la Puente, es la siguiente: observemos que el polinomio

$$f_{\varepsilon,A}(x, y) = y^4 + (2x^2 - 2\varepsilon^2 + 2\mu^2)y^2 - r^2\varepsilon^2 + x^4 - 2x^2(\varepsilon^2 + \mu^2) + (\varepsilon^2 - \mu^2)^2,$$

es bicuadrático en y . Luego, resolviendo y^2 , el discriminante es $4\mu^2x^2 + r^2\varepsilon^2$ que no es un cuadrado, salvo si $r = 0$. Esto impide escribir $f_{\varepsilon,A}(x, y)$ como producto de dos polinomios de grado 2.

(2) Cuando la matriz A es normal, la demostración del segundo apartado es inmediata. Supongamos que A es no normal. Como

$$\begin{aligned} \Lambda'_\varepsilon(A) &= \{(x, y) | \sigma_2((x + yi)I_2 - A) = \varepsilon\}, \\ V_{\mathbb{R}}(f_{\varepsilon,A}) &= \{(x, y) | \sigma_1((x + yi)I_2 - A) = \varepsilon \text{ o } \sigma_2((x + yi)I_2 - A) = \varepsilon\}, \end{aligned}$$

resulta

$$V_{\mathbb{R}}(f_{\varepsilon,A}) \setminus \Lambda'_\varepsilon(A) = \{(x, y) | \sigma_1((x + yi)I_2 - A) = \varepsilon > \sigma_2((x + yi)I_2 - A)\}.$$

Sea ahora ε tal que $\partial\Lambda'_\varepsilon(A) \subsetneq V_{\mathbb{R}}(f_{\varepsilon,A})$. Probemos en primer lugar que $\varepsilon > \varepsilon_0$. En caso contrario, deducimos que existe (x_1, y_1) tal que

$$\varepsilon_0 \geq \sigma_1((x_1 + y_1i)I_2 - A) = \varepsilon > \sigma_2((x_1 + y_1i)I_2 - A).$$

Luego, por un lado, existe una matriz $X \in \mathbb{C}^{2 \times 2}$ tal que $\|X - A\| = \varepsilon \leq \varepsilon_0$, de manera que X tiene a $x_1 + y_1i$ como valor propio doble, además de multiplicidad geométrica = 2. Pero como $\sigma_2((x_1 + y_1i)I_2 - A) < \varepsilon \leq \varepsilon_0$, resulta que $x_1 + y_1i \in \Lambda'_\varepsilon(A)$ que tiene dos componentes conexas. Luego, por la conservación de las multiplicidades, resulta que no puede existir una matriz X que diste $\leq \varepsilon$ de A de manera que $x_1 + y_1i$ sea un valor propio múltiple de X . Luego $\varepsilon > \varepsilon_0$.

Para probar que $\varepsilon \geq \varepsilon_1$, observemos que no se pierde generalidad si suponemos que la matriz A está en la forma (8.1.5), con $r > 0$. Calculando el mínimo de la norma de $(x + yi)I_2 - A$, resulta que dicho mínimo se alcanza en $(x, y) = (0, 0)$, y vale ε_1 . Luego, para todo (x, y) resulta que $\sigma_1((x + yi)I_2 - A) \geq \varepsilon_1$. En consecuencia $\varepsilon \geq \varepsilon_1$. Pero para el caso $\varepsilon = \varepsilon_1$, por la Ecuación (8.1.6), resulta que

$$\sigma_1((0 + 0i)I_2 - A) = \varepsilon_1 > \sigma_2((0 + 0i)I_2 - A) = \varepsilon_0,$$

es decir $\partial\Lambda'_{\varepsilon_1}(A) \subsetneq V_{\mathbb{R}}(f_{\varepsilon_1, A})$.

□

8.2. Reducción de matrices 3×3

En el caso $\mathbb{C}^{3 \times 3}$, el estudio de la naturaleza de los puntos dobles de la frontera del pseudoespectro estricto no se ha podido hacer de una forma genérica; sin embargo, podemos presentar ejemplos que ilustran distintos casos. La existencia de una forma canónica para las matrices $\mathbb{C}^{3 \times 3}$ para la relación de semejanza unitaria $\overset{u}{\approx}$ facilita la reducción de los casos. Además, permite destacar el papel que juegan los elementos de las superdiagonales en cada caso.

Sea $A \in \mathbb{C}^{3 \times 3}$ una matriz con valores propios $\lambda_1, \lambda_2, \lambda_3$, no necesariamente distintos. Seguiremos un proceso de reducción de la matriz dada $A \in \mathbb{C}^{3 \times 3}$ para encontrar su forma canónica $c(A)$. Primero, utilizaremos el Lema de Schur, el cual asegura que podemos encontrar una matriz unitaria U tal que A sea unitariamente semejante a

$$U^*AU = \begin{pmatrix} \lambda_1 & b_{12} & b_{13} \\ 0 & \lambda_2 & b_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix}. \quad (8.2.13)$$

A continuación expresamos cada uno de los elementos por encima de la diagonal en forma exponencial:

$$b_{12} = |b_{12}|e^{i\theta_{12}}, \quad b_{13} = |b_{13}|e^{i\theta_{13}}, \quad b_{23} = |b_{23}|e^{i\theta_{23}}. \quad (8.2.14)$$

Mediante otra transformación de semejanza unitaria, podemos considerar la matriz (8.2.13) de modo que los elementos de la superdiagonal sean números

reales no negativos. En concreto, sea

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i\theta_{12}} & 0 \\ 0 & 0 & e^{-i(\theta_{12}+\theta_{23})} \end{pmatrix},$$

entonces

$$V^* \begin{pmatrix} \lambda_1 & b_{12} & b_{13} \\ 0 & \lambda_2 & b_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} V = \begin{pmatrix} \lambda_1 & |b_{12}| & |b_{13}|e^{i\alpha} \\ 0 & \lambda_2 & |b_{23}| \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

donde

$$\alpha := \theta_{13} - \theta_{12} - \theta_{23}. \quad (8.2.15)$$

En el caso de que alguno de los elementos b_{12} o b_{23} sean nulos, se puede llevar el elemento de la posición $(1, 3)$ a un número real no negativo mediante otra transformación unitaria análoga a la anterior. En efecto,

$$\text{si } b_{12} = 0, \tilde{V} = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \text{si } b_{23} = 0, \tilde{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i\alpha} & 0 \\ 0 & 0 & e^{-i\alpha} \end{pmatrix}.$$

Una vez expresada la matriz en esa forma, podemos dar la correspondiente forma canónica. Percy [59] demostró que existen distintas formas canónicas para la semejanza unitaria de matrices $\mathbb{C}^{3 \times 3}$ en función del número de valores propios distintos. Más precisamente,

Lema 8.2.1. [59] *Dada una matriz $A \in \mathbb{C}^{3 \times 3}$, existe una matriz unitaria $U \in \mathbb{C}^{3 \times 3}$, de manera que:*

(1) *Si A tiene un valor propio triple α , entonces*

$$U^*AU = \begin{pmatrix} \alpha & a & b \\ 0 & \alpha & c \\ 0 & 0 & \alpha \end{pmatrix}, \quad (8.2.16)$$

siendo $a, c \geq 0, b \in \mathbb{C}$; además, si uno de los números a, c es igual a 0, el otro también lo es, y en este caso $b \geq 0$.

(2) *Si A tiene un valor propio doble α_1 y otro simple α_2 , entonces*

$$U^*AU = \begin{pmatrix} \alpha_1 & a & b \\ 0 & \alpha_1 & c \\ 0 & 0 & \alpha_2 \end{pmatrix}, \quad (8.2.17)$$

siendo $a, c \geq 0, b \in \mathbb{C}$; si $a = 0$ entonces $c = 0$; y si $c = 0$ entonces $b \geq 0$.

(3) *Si A tiene tres valores propios distintos $\alpha_1, \alpha_2, \alpha_3$, entonces*

$$U^*AU = \begin{pmatrix} \alpha_1 & a & b \\ 0 & \alpha_2 & c \\ 0 & 0 & \alpha_3 \end{pmatrix}, \text{ siendo } \begin{cases} a, c \geq 0, \\ \text{si } a = 0 \text{ o } c = 0 \Rightarrow b \geq 0. \end{cases} \quad (8.2.18)$$

Una vez determinadas las formas canónicas, sería deseable poder encontrar unos invariantes para esta relación de semejanza, tal que se pueda expresar los elementos de la matriz mediante esos invariantes. Murnaghan [56] dio un conjunto de 6 invariantes formados por las trazas de matrices siguientes:

$$\operatorname{tr}(A), \operatorname{tr}(A^2), \operatorname{tr}(A^3), \operatorname{tr}(A^*A), \operatorname{tr}(A^*A^2), \operatorname{tr}((A^*)^2A^2).$$

Pearcy [59] añadió estos tres invariantes:

$$\operatorname{tr}(A^*AA^*A), \operatorname{tr}(A^*A^2A^*A), \operatorname{tr}(A^*A^2(A^*)^2A),$$

finalmente, Sibirskii [68] probó que el conjunto siguiente es un sistema minimal completo de invariantes para la semejanza unitaria:

$$\operatorname{tr}(A), \operatorname{tr}(A^2), \operatorname{tr}(A^3), \operatorname{tr}(A^*A), \operatorname{tr}(A^*A^2), \operatorname{tr}((A^*)^2A^2), \operatorname{tr}(A^*A^2(A^*)^2A). \quad (8.2.19)$$

Los invariantes $\operatorname{tr}(A^*AA^*A)$, $\operatorname{tr}(A^*A^2A^*A)$ se pueden expresar como polinomios en las trazas del anterior sistema, por lo tanto, no los tomaremos en cuenta. Por otro lado, el sistema propuesto por Murnaghan no es completo, véanse ejemplos 5), 6), 7) y 8) en [68]).

Un lema adicional que nos resultará útil para la ilustración del ejemplo al final de esta introducción es el siguiente. Antes damos algo de notación necesaria.

$$\operatorname{Tri}(a_{11}, a_{22}, a_{33}, a_{12}, a_{23}, a_{13}) := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Lema 8.2.2 ([59]). Sean $A_1, A_2 \in \mathbb{C}^{3 \times 3}$ las matrices

$$A_i := \operatorname{Tri}(\lambda_1, \lambda_2, \lambda_3, d_i, g_i, f_i),$$

con $d_i, g_i \geq 0$, $f_i \in \mathbb{C}$, $i = 1, 2$, y supongamos que las trazas correspondientes $\operatorname{tr}(A_i)$, $\operatorname{tr}(A_i^2)$, $\operatorname{tr}(A_i^3)$, $\operatorname{tr}(A_i^*A_i)$, $\operatorname{tr}(A_i^*A_i^2)$, $\operatorname{tr}((A_i^*)^2A_i^2)$, $i = 1, 2$, son iguales. Entonces, los tres números

$$\begin{aligned} & d_i^2 + g_i^2 + |f_i|^2 \\ & |\lambda_3 - \lambda_2|^2 d_i^2 + |\lambda_2 - \lambda_1|^2 g_i^2 + d_i^2 g_i^2 \\ & (\lambda_3 - \lambda_2)^2 d_i^2 + (\lambda_2 - \lambda_1)^2 g_i^2 - d_i g_i f_i^* \end{aligned}$$

son iguales.

Las trazas de las potencias de A hasta grado igual al orden de la matriz determinan los valores propios de la matriz. En el caso particular de una matriz normal $A \in \mathbb{C}^{n \times n}$, la cual es unitariamente semejante a una matriz diagonal, es suficiente considerar $\operatorname{tr}(A^i)$, $i = 1, \dots, n$ para tener el sistema completo de invariantes para esta relación de semejanza. El modo de expresar los valores

propios como función de las trazas se puede hacer mediante las relaciones de las sumas de Newton, las funciones simétricas elementales y las fórmulas de Cardano-Ferrari para la expresión mediante radicales de las ecuaciones de grado 3. En concreto, las sumas de Newton de $\lambda_1, \lambda_2, \lambda_3$, hasta tercer orden, se definen como:

$$s_0 = \lambda_1^0 + \lambda_2^0 + \lambda_3^0 = 3, \quad s_1 = \lambda_1 + \lambda_2 + \lambda_3, \quad s_2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad s_3 = \lambda_1^3 + \lambda_2^3 + \lambda_3^3.$$

Si identificamos $\lambda_1, \lambda_2, \lambda_3$ con los valores propios de la matriz A , entonces:

$$s_1 = \text{tr}(A), \quad s_2 = \text{tr}(A^2), \quad s_3 = \text{tr}(A^3). \quad (8.2.20)$$

Por otro lado, las tres primeras funciones simétricas elementales en las variables $\lambda_1, \lambda_2, \lambda_3$ se definen como:

$$e_1 = \lambda_1 + \lambda_2 + \lambda_3, \quad e_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3, \quad e_3 = \lambda_1\lambda_2\lambda_3. \quad (8.2.21)$$

Las identidades de Newton nos relacionan las expresiones (8.2.20) y (8.2.21); además, conocidas las segundas, podemos expresarlas en función de las primeras:

$$\begin{cases} s_k - e_1 s_{k-1} + e_2 s_{k-2} - \cdots + (-1)^{k-1} e_{k-1} s_1 + (-1)^k k e_k = 0 & \text{si } k \leq n, \\ s_k - e_1 s_{k-1} + e_2 s_{k-2} - \cdots + (-1)^n e_n s_{k-n} = 0 & \text{si } k > n. \end{cases}$$

En nuestro caso, tenemos:

$$\begin{cases} s_1 - e_1 = 0 \\ s_2 - e_1 s_1 + 2e_2 = 0 \\ s_3 - e_1 s_2 + e_2 s_1 - 3e_3 = 0 \end{cases} \Rightarrow \begin{cases} e_1 = s_1 = \text{tr}(A) \\ e_2 = \frac{s_1^2 - s_2}{2} = \frac{\text{tr}(A)^2 - \text{tr}(A^2)}{2} \\ e_3 = \frac{-2s_3 + 3s_1 s_2 - s_1^3}{3} \\ = \frac{-2\text{tr}(A^3) + 3\text{tr}(A)\text{tr}(A^2) - \text{tr}(A)^3}{6}. \end{cases}$$

Finalmente las fórmulas de Cardano-Ferrari nos dan las raíces del polinomio característico de A , $p(\lambda) = \lambda^3 + a\lambda^2 + b\lambda + c$ en términos de sus coeficientes por medio de radicales. Lo que nos permite obtener una expresión explícita de los valores propios de A en función de las trazas $\text{tr}(A), \text{tr}(A^2), \text{tr}(A^3)$.

Cabría esperar que las restantes trazas del Criterio de Sibirskiĭ determinen los elementos de las superdiagonales, de forma análoga a lo que ocurre en el caso 2×2 . Sin embargo, Paulsen [58] demostró que *no existe una forma canónica continua* para la semejanza unitaria en $\mathbb{C}^{n \times n}$ si $n \geq 3$; por lo tanto, en general, no podemos escribir los elementos de las superdiagonales como funciones algebraicas de los invariantes dados por las trazas (8.2.19). No obstante, podemos analizar el caso particular en el que la matriz tiene un único valor propio que, sin pérdida de generalidad, se puede tomar nulo. Dada una matriz del tipo $\text{Tri}(\lambda, \lambda, \lambda, d, g, a + bi)$, siendo $\lambda = 0$; d, g números reales no negativos, y a, b reales, intentaremos expresar los elementos de las superdiagonales en función de

las cuatro últimas trazas dadas en (8.2.19). Si denotamos por $p_i, i = 4, \dots, 7$, el valor de las correspondientes trazas, tenemos el siguiente sistema de ecuaciones

$$d^2 + a^2 + b^2 + g^2 - p_4 = 0,$$

$$(a - bi)dg - p_5 = 0,$$

$$d^2g^2 - p_6 = 0,$$

$$d^4g^2 + d^2g^2a^2 + d^2g^2b^2 - p_7 = 0.$$

MAPLE 13 nos da las expresiones siguientes de a, b, c, d en función de las trazas:

$$a = \frac{1}{2} \frac{p_4p_6p_5^2 + p_7p_4p_6 - p_5^2p_7 - p_7^2 - p_6^3}{\alpha\beta p_5p_6(p_4p_6 - p_7)} \quad (8.2.22)$$

$$b = \frac{1}{2} \frac{(p_4p_6p_5^2 - p_7p_4p_6 - p_5^2p_7 + p_7^2 + p_6^3)}{\alpha\beta p_5p_6(p_4p_6 - p_7)} i, \quad (8.2.23)$$

$$d = \alpha p_6, \quad (8.2.24)$$

$$g = \beta. \quad (8.2.25)$$

siendo

$$\alpha := \sqrt{\frac{1}{p_4p_6 - p_7}}, \quad \beta := \sqrt{p_4 - \frac{p_7}{p_6}}.$$

Observación 8.2.3. Enseguida veremos cómo se pueden transformar estas matrices canónicas en otras aún más reducidas, o sea más simples, mediante transformaciones afines, las cuales no tienen por qué ser semejantes unitariamente a las matrices dadas. Es decir, que a la hora de estudiar la geometría de los pseudoespectros de las matrices A de $\mathbb{C}^{3 \times 3}$ podemos servirnos de la forma canónica de Pearcy, más de transformaciones del tipo $A \mapsto \alpha I_3 + \beta A$, con $\alpha, \beta \in \mathbb{C}$.

Dada una matriz $A \in \mathbb{C}^{3 \times 3}$, para $(x, y) \in \mathbb{R}^2$ y $\varepsilon > 0$ definimos:

$$\begin{aligned} \mathbf{A}(x, y) &:= (x + iy)I_3 - A, \\ f_{\varepsilon, A}(x, y) &:= \det \begin{pmatrix} \varepsilon I_3 & -\mathbf{A}(x, y) \\ -\mathbf{A}(x, y)^* & \varepsilon I_3 \end{pmatrix}. \end{aligned} \quad (8.2.26)$$

Con estas notaciones, vamos a considerar la familia de curvas algebraicas

$$V_{\mathbb{R}}(f_{\varepsilon, A}) := \{(x, y) \in \mathbb{R}^2 : f_{\varepsilon, A}(x, y) = 0\},$$

que, recordemos, incluye a la frontera del pseudoespectro estricto de A de nivel ε , denotada por $\partial\Lambda'_\varepsilon(A)$. Para todo valor de $\varepsilon \in [0, \infty)$ excepto para un número finito de valores de ε la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ es *lisa*, i.e. no tiene puntos singulares. Nuestro objetivo principal es estudiar cuáles de las curvas de la familia $V_{\mathbb{R}}(f_{\varepsilon, A})$ al variar $\varepsilon \in [0, \infty)$ tienen puntos singulares (x_0, y_0) ; y después averiguar cuáles de estos puntos (x_0, y_0) están contenidos en $\partial\Lambda'_\varepsilon(A)$; es decir, cuáles cumplan que $\sigma_3(\mathbf{A}(x_0, y_0)) = \varepsilon$. Para ello, vamos a utilizar las siguientes propiedades:

Lema 8.2.4. *Con las notaciones anteriores y sobrentendiendo las operaciones de \mathbb{C} como las correspondientes transformaciones afines en \mathbb{R}^2 , se tiene:*

- (1) *Si $U \in \mathbb{C}^{3 \times 3}$ es una matriz unitaria, entonces $V_{\mathbb{R}}(f_{\varepsilon, U^*AU}) = V_{\mathbb{R}}(f_{\varepsilon, A})$.*
- (2) *Dados $\alpha, \beta \in \mathbb{C}$ con $\beta \neq 0$, $V_{\mathbb{R}}(f_{\varepsilon, \alpha I_3 + \beta A}) = \alpha + \beta V_{\mathbb{R}}(f_{\varepsilon/|\beta|, A})$.*
- (3) *Las mismas propiedades se cumplen para $\partial\Lambda'_\varepsilon(A)$.*

Nuestro objetivo es ver qué clase de puntos singulares puede tener la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$. El caso de matrices normales es muy simple, puesto que la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ está compuesta por circunferencias de radio ε con centros en los valores propios de A . Luego, de aquí en adelante, supondremos que la matriz A es no normal. Además, vamos considerar tres casos, dependiendo del número de valores propios distintos de la matriz A .

A diferencia del caso de las matrices 2×2 , aquí no vamos a desarrollar de una forma general toda la casuística para las matrices 3×3 . Nos limitaremos a ejemplos significativos que a nuestro parecer cubren todos las posibilidades de puntos singulares de una curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ que pertenezcan a la frontera del ε -pseudoespectro estricto. Obviamente, no analizaremos los posibles puntos crunodos para el caso de matrices con dos o tres valores propios distintos, que es un caso trivial.

En las siguientes subsecciones utilizaremos la siguiente notación: Consideremos una curva de \mathbb{R}^2 definida por $f(x, y) = 0$, con $f \in \mathbb{R}[x, y]$. Supongamos que el punto (x_0, y_0) es singular para la curva f , es decir

$$\begin{aligned} f(x_0, y_0) &= 0, \\ f'_x(x_0, y_0) &= 0, \\ f'_y(x_0, y_0) &= 0. \end{aligned}$$

La curva f es una abreviatura de $V_{\mathbb{R}}(f)$. Denotemos por

$$\Delta(x_0, y_0) := f''_{xy}(x_0, y_0)^2 - f''_{xx}(x_0, y_0)f''_{yy}(x_0, y_0).$$

Supongamos que el punto (x_0, y_0) es *doble*. Entonces por el Teorema A.5.1 se tiene que: (i) Si $\Delta < 0$ el punto (x_0, y_0) es aislado (acnodo) para la curva f . (ii) Si $\Delta > 0$ el punto (x_0, y_0) tiene dos tangentes (crunodo). (iii) Si $\Delta = 0$, el punto (x_0, y_0) tiene una tangente doble, T , y puede ser un tacnodo o una cúspide; además, si el número de intersección de la curva f con la recta T en (x_0, y_0) , $I((x_0, y_0), f, T) = 3$ entonces es una cúspide ordinaria. *Sin embargo*, del hecho de que $I((x_0, y_0), f, T) \geq 4$ no se infiere nada: hay ejemplos en los que el punto (x_0, y_0) es una cúspide y otros en los que es un tacnodo. En este caso, se debe recurrir al *número de ramas centradas* en (x_0, y_0) de la curva f para decidirlo.

8.2.1. Caso de un valor propio triple.

Teniendo en cuenta el Lema 8.2.1 no se pierde generalidad si asumimos que A es de la forma (8.2.16). Además, por el Lema 8.2.4, podemos suponer que $\alpha = 0$. Es decir, podemos suponer que

$$A := \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix},$$

con $a, c \geq 0$. Para este caso, vamos a dar ejemplos de matrices tales que la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ tiene puntos singulares que pueden ser: o aislados (acnodos), o crunodos, o de retroceso de primera especie. Seguimos la Definición Definitiva A.2.22 y las notaciones (8.2.26) sobre el polinomio de Wielandt $f_{\varepsilon}(x, y)$. Omitiremos la referencia a la matriz A siempre que esto no cause confusión.

Ya hemos probado en la Proposición 7.4.6, que tomando $a = c = 1, b = 4$, el punto $(x_0, y_0) = (-4/15, 0)$ es un punto doble para la curva $f_{1/15}(x, y)$, y además es aislado. También se tiene que $1/15$ es el menor valor singular de $A(-4/15, 0)$.

De otro lado, eligiendo la misma matriz y

$$(x_1, y_1, \varepsilon_1) = \left(-\frac{2 + \sqrt{22}}{24}, 0, \frac{11\sqrt{22} - 50}{24} \right)$$

se tiene que (x_1, y_1) es un punto doble de la curva f_{ε_1} y como

$$\Delta(x_1, y_1) = \frac{74990311\sqrt{22}}{432} - \frac{43966967}{54} \approx 0'001269414575,$$

resulta que es un punto crunodo. Nuevamente, se tiene que ε_1 es el menor valor singular de $A(x_1, y_1)$.

Finalmente, si

$$a = c = 1, b = 2 + \sqrt{3}; (x_2, y_2, \varepsilon_2) = \left(-\frac{\sqrt{3}}{6}, 0, \frac{2\sqrt{3} - 3}{6} \right),$$

entonces $p_2 := (x_2, y_2)$ es un punto doble de la curva f_{ε_2} , $\Delta(x_2, y_2) = 0$, y además ε_2 es el menor valor singular de $A(x_2, y_2)$. Este es un caso de punto de retroceso de primera especie. Ya que calculadas las singularidades de la curva f_{ε_2} con `Maple 17` nos da:

`singularities(f,x,y);`

$$\left\{ \left[\left[-1/6\sqrt{3}, 0, 1 \right], 2, 1, 1 \right], \left[\left[\text{RootOf}(Z^2 + 1), 1, 0 \right], 3, 4, 1 \right] \right\}$$

lo que nos dice que sólo hay una rama de la curva f_{ε_2} centrada en $\left(-\frac{\sqrt{3}}{6}, 0\right)$. Como $\Delta = 0$ en dicho punto, este punto doble es un punto cuspidal ordinario pues $I(p_2, f_{\varepsilon_2}, T) = 3$ siendo T la recta tangente a $V_{\mathbb{R}}(f_{\varepsilon_2})$ en p_2 dada por las ecuaciones paramétricas $x = -\frac{\sqrt{3}}{6} + t, y = 0$ en virtud del Teorema A.5.1 III (a). Para ver que es de primera especie, utilizaremos la parametrización que nos da `puiseux`. Una parametrización de la rama de f_{ε_2} en el punto p_2 viene dada por

$$\tilde{x}(T) = -\frac{4}{3}(1 + 2\sqrt{3})T^2 - \frac{\sqrt{3}}{6}, \quad \tilde{y}(T) = \frac{208 + 64\sqrt{3}}{9}T^3 + \dots$$

Vemos que $\tilde{x}(0) = -\sqrt{3}/6$, $\tilde{y}(0) = 0$. Las derivadas sucesivas nos dan:

$$\begin{aligned} \tilde{x}'(T) &= -\frac{8}{3}(1 + 2\sqrt{3})T, & \tilde{y}'(T) &= \frac{208 + 64\sqrt{3}}{3}T^2 + \dots, \\ \tilde{x}'(0) &= 0, & \tilde{y}'(0) &= 0, \\ \tilde{x}''(T) &= -\frac{8 + 16\sqrt{3}}{3}, & \tilde{y}''(T) &= \frac{2(208 + 64\sqrt{3})}{3}T + \dots, \\ \tilde{x}''(0) &= -\frac{8 + 16\sqrt{3}}{3}, & \tilde{y}''(0) &= 0, \\ \tilde{x}'''(0) &= 0, & \tilde{y}'''(0) &= \frac{2(208 + 64\sqrt{3})}{3}. \end{aligned}$$

Por lo tanto, al considerar las derivadas sucesivas de $\mathbf{r}(T) := (\tilde{x}(T), \tilde{y}(T))$ en $T = 0$ la primera derivada, $\mathbf{r}'(0)$, se anula, la segunda derivada, $\mathbf{r}''(0)$, no se anula, y la tercera derivada, $\mathbf{r}'''(0)$ no depende linealmente de $\mathbf{r}''(0)$. Así pues, por el Teorema A.2.17 y con sus notaciones se tiene que $p = 2$ y $q = 3$. Como p es par y q es impar, la curva f_{ε_2} tiene en (x_2, y_2) un *punto de retroceso* (o *cúspide*) de primera especie.

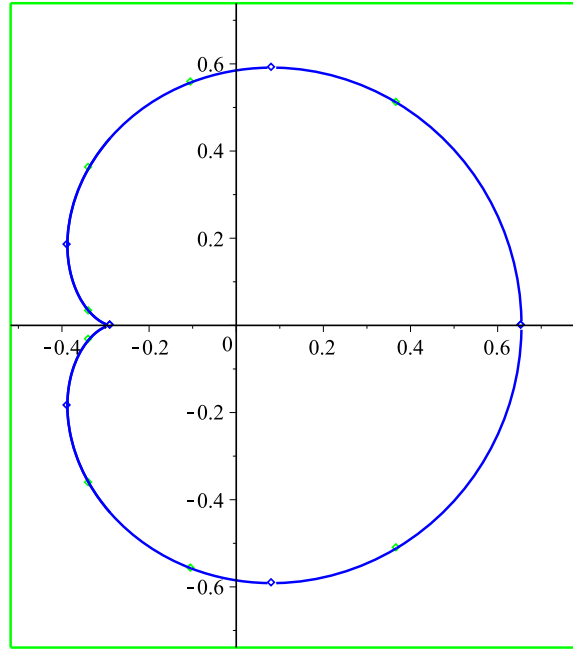


Figura 8.1: Retroceso de primera especie en $(x_2, y_2) = \left(-\frac{\sqrt{3}}{6}, 0\right)$

8.2.2. Caso de un valor propio doble y uno simple

Nuevamente, por el Lema 8.2.1, no se pierde generalidad A es de la forma (8.2.17). Además, por el Lema 8.2.4, podemos suponer que $\alpha_1 = 0, \alpha_2 = 1$. Es decir, podemos suponer que

$$A := \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 1 \end{pmatrix},$$

con $a, c, \geq 0, b \in \mathbb{C}$. Para este caso, vamos a dar ejemplos de matrices concretas tales que para cierto valor de ε la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ tiene un punto doble que es un tacnodo.

Tomemos

$$a = \frac{\sqrt{5}}{10}, b = -\frac{\sqrt{3}}{6}, c = \frac{\sqrt{15}}{15}; (x_3, y_3, \varepsilon_3) = \left(\frac{1}{2}, 0, \frac{\sqrt{15}}{10}\right).$$

Para esta elección, se tiene que $\Delta(x_3, y_3) = 0$, y además ε_3 es el menor valor singular de $A(x_3, y_3)$. Vamos a probar que el punto (x_3, y_3) es un tacnodo de la curva f_{ε_3} . En efecto, el polinomio f_{ε_3} viene dado por

$$\begin{aligned} f_{\varepsilon_3}(x, y) = f &= \frac{1}{40}x - \frac{1}{64} + \frac{21}{80}y^2 - y^6 + \frac{21}{80}x^2 - \frac{11}{20}x^4 - 3/5x^3 - x^6 + 2x^5 \\ &- 3y^2x^4 + 4y^2x^3 - \frac{11}{10}y^2x^2 - \frac{3}{5}y^2x - \frac{11}{20}y^4 - 3y^4x^2 + 2y^4x. \end{aligned}$$

Con el paquete Maple[algcures] obtenemos los puntos singulares y sus invariantes m, δ, r , así como las ramas en los puntos de la forma (x_3, y_3) :

`singularities(f,x,y);`

$$\{[[1/2, 0, 1], 2, 2, 2], [[RootOf(Z^2 + 1), 1, 0], 3, 4, 3]\}.$$

De aquí se deduce que $p_3 := (1/2, 0)$ es un punto singular con $m = 2, \delta = 2, r = 2$. Es decir, corroboramos que es un punto doble y que $f_{\varepsilon_3} = f$ tiene dos ramas R_1 y R_2 que pasan por él. Como $\Delta(1/2, 0) = 0$, el punto p_3 es un *tacnodo*. Además, sea τ la recta de ecuaciones paramétricas $x = 1/2, y = t$, entonces el número de intersección de f_{ε_3} con τ en el punto p_3 , $I(p_3, f_{\varepsilon_3}, \tau)$ es igual a 4, pues

$$f_{\varepsilon_3}(1/2, t) = -t^6 - \frac{3}{10}t^4;$$

mas de este hecho, no se deduce nada sobre si p_3 es un tacnodo o una cúspide. Sólo que τ es tangente a $V_{\mathbb{R}}(f_{\varepsilon_3})$ en p_3 . También queremos ilustrar que $\partial\Lambda'_{\varepsilon_2}(A) \subset V_{\mathbb{R}}(f_{\varepsilon_3})$, pero el contenido puede ser estricto y en este caso lo es: de hecho para el punto $(0,352564517516, 0)$ capturado con el ratón como el vértice de la curva interior más próximo a $(1/2, 0)$ tenemos que

$$\sigma_2(A(0,352564517516, 0)) = 0,3873 \simeq \frac{\sqrt{15}}{10};$$

véanse las Figuras 8.2 y 8.3.

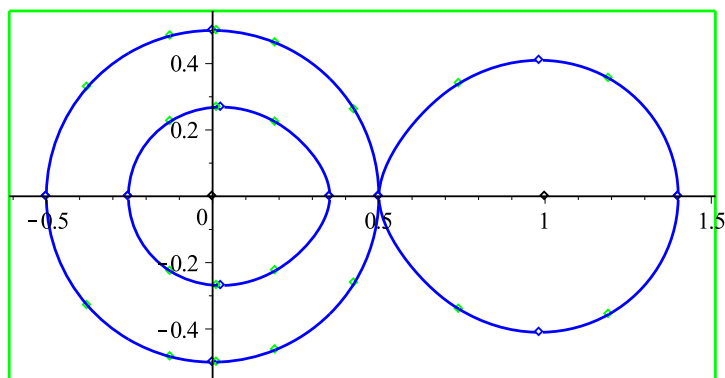


Figura 8.2: Tacnodo en $(x_3, y_3) = (1/2, 0)$ a la curva algebraica $V_{\mathbb{R}}(f_{\epsilon_3})$.

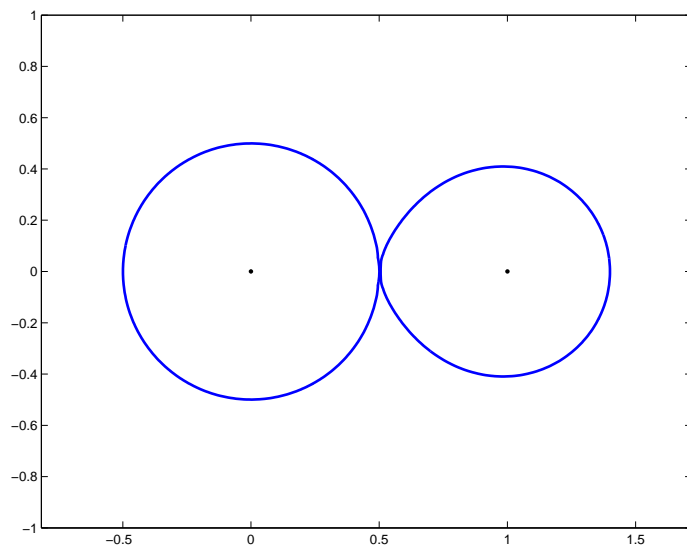


Figura 8.3: Conjunto frontera $\partial \Lambda'_{\epsilon_3}(A)$

También obtenemos mediante
`puiseux(f, x=1/2, y, 0, T);`

$$[x = -T^2 + \frac{1}{2}, y = -T], [x = \frac{1}{3}T^2 + \frac{1}{2}, y = \frac{1}{3}T],$$

$$[x = T + \frac{1}{2}, y = \text{RootOf}(10Z^2 + 3)],$$

de donde se deduce que las sumas parciales de los pares de series formales de potencias en T que parametrizan las ramas R_1 y R_2 son $(x = -T^2 + \frac{1}{2}, y = -T)$ y $(x = \frac{1}{3}T^2 + \frac{1}{2}, y = \frac{1}{3}T)$, respectivamente.

El procedimiento utilizado para demostrar que el punto $(1/2, 0)$ es un *tacnodo* para la curva algebraica $C := V_{\mathbb{R}}(f_{\varepsilon_3})$ sigue la definición que aparece en el Ejercicio 5.4.3 página 115 del libro de Walker [73]: “un tacnodo es el centro de dos ramas (*places*) lineales que se cortan en él con multiplicidad dos”. Véase más precisión en la Proposición A.7.9 del Apéndice A. Procedamos a hacer estos cálculos con ayuda de **Maple** 17, usando las notaciones de la Definición A.7.8.

Sean $\eta := \zeta := e^{\frac{2\pi i}{2}} = -1$;

$$\bar{y}_1 := -\eta X^{1/2}, \bar{y}_2 := -\eta^2 X^{1/2}, \bar{z}_1 := \frac{1}{3}\zeta X^{1/2}, \bar{z}_2 := \frac{1}{3}\zeta^2 X^{1/2}.$$

De donde tenemos que,

$$\begin{aligned} \varphi(R_1, R_2) &:= (\bar{y}_1 - \bar{z}_1)(\bar{y}_1 - \bar{z}_2)(\bar{y}_2 - \bar{z}_1)(\bar{y}_2 - \bar{z}_2) \\ &= (-\eta - \zeta/3)X^{1/2}(-\eta - \zeta^2/3)X^{1/2}(-\eta^2 - \zeta/3)X^{1/2}(-\eta^2 - \zeta^2/3)X^{1/2} \\ &= (-\eta - \zeta/3)(-\eta - \zeta^2/3)(-\eta^2 - \zeta/3)(-\eta^2 - \zeta^2/3)X^2 = \frac{64}{81}X^2. \end{aligned}$$

Por lo tanto $\varphi(R_1, R_2) \in \mathbb{C}[[X]]$ y su orden $\text{ord}(\frac{64}{81}X^2) = 2$; es decir,

$$I((1/2, 0), R_1, R_2) = 2.$$

Obviamente R_1 y R_2 son ramas de orden uno (lineales). Hemos acabado de probar, según la Definición A.7.8, que f_{ε_3} tiene un tacnodo en el punto $(1/2, 0)$.

Mas nos asalta una duda. Si tomamos sumas parciales más largas de las series formales $b + a_1T + a_2T^2 + \dots$ y $b + b_1T + b_2T^2 + \dots$ ¿llegaremos a la misma conclusión en lo que al orden de $\varphi(R_1, R_2)$ respecta? Veamos, con la instrucción `puiseux` con otros parámetros obtenemos:

`puiseux(f, x=1/2, y, 5, T);`

$$\begin{aligned} [x = T + \frac{1}{2}, y = -\frac{649175}{17496}\text{RootOf}(10Z^2 + 3)T^4 - \frac{7075}{486}\text{RootOf}(10Z^2 + 3)T^3 \\ + \frac{1045}{162}\text{RootOf}(10Z^2 + 3)T^2 + \frac{5}{9}\text{RootOf}(10Z^2 + 3)T + \text{RootOf}(10Z^2 + 3)], \\ [x = -T^2 + \frac{1}{2}, y = \frac{5}{128}T^9 + \frac{1}{16}T^7 + \frac{1}{8}T^5 + \frac{1}{2}T^3 - T], \\ [x = \frac{1}{3}T^2 + \frac{1}{2}, y = \frac{70793995}{204073344}T^9 + \frac{77359}{314928}T^7 - \frac{9209}{17496}T^5 + \frac{53}{162}T^3 + \frac{1}{3}T] \end{aligned}$$

La componente y de la rama R_1 es

$$y := \frac{5}{128}T^9 + \frac{1}{16}T^7 + \frac{1}{8}T^5 + \frac{1}{2}T^3 - T;$$

de donde se tiene que,

$$\bar{y}_1 := \eta \left(\frac{5}{128}X^{9/2} + \frac{1}{16}X^{7/2} + \frac{1}{8}X^{5/2} + \frac{1}{2}X^{3/2} - \sqrt{X} \right),$$

$$\bar{y}_2 := \eta^2 \left(\frac{5}{128} X^{9/2} + \frac{1}{16} X^{7/2} + \frac{1}{8} X^{5/2} + \frac{1}{2} X^{3/2} - \sqrt{X} \right),$$

La componente y de la rama R_2 es

$$\frac{70793995}{204073344} T^9 + \frac{77359}{314928} T^7 - \frac{9209}{17496} T^5 + \frac{53}{162} T^3 + \frac{1}{3} T;$$

con lo cual tenemos que

$$\bar{z}_1 := \zeta \left(\frac{70793995}{204073344} X^{9/2} + \frac{77359}{314928} X^{7/2} - \frac{9209}{17496} X^{5/2} + \frac{53}{162} X^{3/2} + \frac{1}{3} \sqrt{X} \right),$$

$$\bar{z}_2 := \zeta^2 \left(\frac{70793995}{204073344} X^{9/2} + \frac{77359}{314928} X^{7/2} - \frac{9209}{17496} X^{5/2} + \frac{53}{162} X^{3/2} + \frac{1}{3} \sqrt{X} \right);$$

De aquí que

$$\begin{aligned} \psi(R_1, R_2) &:= (\bar{y}_1 - \bar{z}_1)(\bar{y}_1 - \bar{z}_2)(\bar{y}_2 - \bar{z}_1)(\bar{y}_2 - \bar{z}_2) = \\ &= X^2 c(X), \end{aligned}$$

siendo $c(X)$ el elemento de $\mathbb{C}[[X]]$ dado por

$$\begin{aligned} &\frac{382579837531913164947171300625}{27099741612310992724404694155264} X^{16} + \frac{411295441490796522449626375}{10455147226971833612810452992} X^{15} \\ &- \frac{28046371136444992247656025}{580841512609546311822802944} X^{14} - \frac{677446491035175543090145}{5378162153792095479840768} X^{13} \\ &+ \frac{11546742711555729555509}{49797797720297180368896} X^{12} + \frac{17948174163609910579}{76848453272063549952} X^{11} \\ &- \frac{2290584833705753399}{4269358515114641664} X^{10} + \frac{5490615733383025}{26354064908115072} X^9 + \frac{2419338057094745}{2928229434235008} X^8 \\ &- \frac{8607835828897}{6778308875544} X^7 + \frac{7828308911}{125524238436} X^6 + \frac{2175086699}{1162261467} X^5 - \frac{22528000}{14348907} X^4 \\ &- \frac{435200}{1594323} X^3 + \frac{113216}{59049} X^2 - \frac{4736}{2187} X + \frac{64}{81} \end{aligned}$$

con $c(0) = 64/81$. Por tanto, $I((1/2, 0), R_1, R_2) = 2$ y se corrobora que $(1/2, 0)$ es un tacnodo, ya que obviamente las ramas R_1 y R_2 son lineales.

8.2.3. Caso de tres valores propios distintos

Nuevamente, por el Lema 8.2.1, no se pierde generalidad A es de la forma (8.2.18). Además, por el Lema 8.2.4, podemos suponer que $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = \beta + \gamma i$. Es decir, podemos suponer que

$$A := \begin{pmatrix} 0 & a & b \\ 0 & 1 & c \\ 0 & 0 & \beta + \gamma i \end{pmatrix},$$

con $a, c \geq 0$. Para este caso, vamos a dar ejemplos de matrices tales que la curva $V_{\mathbb{R}}(f_{\varepsilon, A})$ tiene puntos dobles que son tacnodos. Vamos a analizar este ejemplo con mayor detalle.

Tomemos

$$a = b = 0, c = 1, \beta = \frac{1}{2}, \gamma = 1;$$

es decir nuestra matriz será

$$A := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & \frac{1}{2} + i \end{pmatrix}.$$

Si definimos su polinomio de Wielandt (1.2.2) por

$$q(x, y, \varepsilon) := \begin{vmatrix} \varepsilon I_n & (x + yi)I_n - A \\ (x - yi)I_n - A^* & \varepsilon I_n \end{vmatrix}$$

tenemos que

$$q(x, y, \varepsilon) = \det \left(\varepsilon^2 I_n - ((x + yi)I_n - A)^* ((x + yi)I_n - A) \right).$$

Y haciendo $\lambda := \varepsilon^2$ para simplificar nos queda que:

$$\begin{aligned} p(x, y, \lambda) &= \det \left(\lambda I_n - ((x + yi)I_n - A)^* ((x + yi)I_n - A) \right) = \frac{1}{4} (-x^2 - y^2 + \lambda) \\ &\quad (4x^4 + 8x^2y^2 + 4y^4 - 8\lambda x^2 - 8\lambda y^2 - 12x^3 - 8yx^2 - 12xy^2 - 8y^3 \\ &\quad + 4\lambda^2 + 12\lambda x + 8y\lambda + 17x^2 + 16yx + 9y^2 - 13\lambda - 14x - 8y + 5) \end{aligned}$$

Vemos que $p(x, y, \lambda) = f(x, y, \lambda)g(x, y, \lambda)$, llamando

$$f(x, y, \lambda) := \frac{1}{4}(-x^2 - y^2 + \lambda)$$

y

$$\begin{aligned} g(x, y, \lambda) &:= 4x^4 + 8x^2y^2 + 4y^4 - 8\lambda x^2 - 8\lambda y^2 - 12x^3 - 8yx^2 - 12xy^2 - 8y^3 \\ &\quad + 4\lambda^2 + 12\lambda x + 8y\lambda + 17x^2 + 16yx + 9y^2 - 13\lambda - 14x - 8y + 5. \end{aligned}$$

Deseamos hallar los valores de $\lambda \geq 0$ para los que la curva algebraica $p(x, y, \lambda) = 0$ tiene valores singulares en los que la curva tiene una única tangente real. Esto nos lleva a buscar los valores de x, y, λ para los que el sistema de ecuaciones

$$\begin{cases} p(x, y, \lambda) = 0, \\ p'_x(x, y, \lambda) = 0, \\ p'_y(x, y, \lambda) = 0, \\ \Delta(x, y, \lambda) = 0, \end{cases}$$

donde $\Delta(x, y, \lambda) := p''_{xy}(x, y, \lambda) - p''_{x^2}(x, y, \lambda)p''_{y^2}(x, y, \lambda)$, tiene soluciones en \mathbb{R}^3 . Abreviando los argumentos, este sistema queda reducido a $p = 0, p'_x = 0, p'_y = 0, \Delta = 0$. Como $p = fg$, este sistema es equivalente al sistema de ecuaciones

$$\begin{cases} f = 0, \\ g = 0, \\ 4f'_x g'_x f'_y g'_y - (f'_x g'_y + f'_y g'_x)^2 = 0. \end{cases}$$

Resuelto este sistema con **Maple** 17 nos da que una de sus soluciones es

$$x = \frac{3}{8}, y = \frac{1}{8}, \lambda = \frac{5}{32}.$$

Lo que nos da lugar el punto singular

$$p_4 := (x_4, y_4) = \left(\frac{3}{8}, \frac{1}{8}\right)$$

de la curva algebraica $f_{\varepsilon_4}(x, y) := q(x, y, \varepsilon_4) = 0$, siendo $\varepsilon_4 := \sqrt{\frac{5}{32}}$. Además, se tiene que $\Delta(p_4) = 0$, y además ε_4 es el menor valor singular de $\mathbf{A}(x_4, y_4)$.

`singularities(p(x,y,5/32),x,y);`

`[[3/8, 1/8, 1], 2, 2, 2], [[RootOf (-Z^2 + 1), 1, 0], 3, 3, 3],`
`[[RootOf (64 _Z^2 - 48 _Z + 13), 3/4 - RootOf (64 _Z^2 - 48 _Z + 13), 1], 2, 1, 2]`

`puiseux(q(x,y,5/32),x=3/8,y,0,T);`

`[x = T + 3/8, y = -1/8], [x = T + 3/8, y = RootOf (512 _Z^3 - 960 _Z^2 + 440 _Z - 185)]`
`[x = T + 3/8, y = 1/8 - 3T - 40T^2], [x = T + 3/8, y = 1/8 - 3T - 40/9 T^2],`

Las ramas que pasan por p_4 , $R_1 := [x = T + 3/8, y = 1/8 - 3T - 40T^2]$ y $R_2 := [x = T + 3/8, y = 1/8 - 3T - \frac{40}{9}T^2]$ son lineales. Además, como

$$(1/8 - 3T - 40T^2) - \left(1/8 - 3T - \frac{40}{9}T^2\right) = \left(\frac{40}{9} - 40\right)T^2$$

se tiene que $I(p_4, R_1, R_2) = 2$. De donde se deduce que el punto p_4 es un tacnodo. Véase la Figura 8.4.

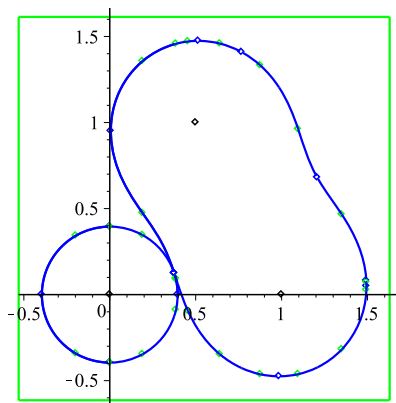


Figura 8.4: Tacnodo en $(x_4, y_4) = (3/8, 1/8)$ a la curva algebraica $V_{\mathbb{R}}(f_{\varepsilon_4})$.

Ejemplo final de matriz 3×3 con tres valores propios distintos. Si tomamos

$$a = b = 0, c = 2, \beta = 0, \gamma = 1; (x_5, y_5, \varepsilon_5) = \left(\frac{\sqrt{2}-1}{2}, \frac{\sqrt{2}-1}{2}, \frac{3-2\sqrt{2}}{2} \right).$$

nos queda que $p_5 := (x_5, y_5) = \left(\frac{\sqrt{2}-1}{2}, \frac{\sqrt{2}-1}{2} \right)$ es un punto doble de la curva algebraica $f_{\varepsilon_5}(x, y) := q(x, y, \varepsilon_5)$, siendo

$$\varepsilon_5 := \sqrt{\frac{3-2\sqrt{2}}{2}},$$

con una tangente única y dos ramas reales. Por tanto, es también un tacnodo. Véase la Figura 8.5

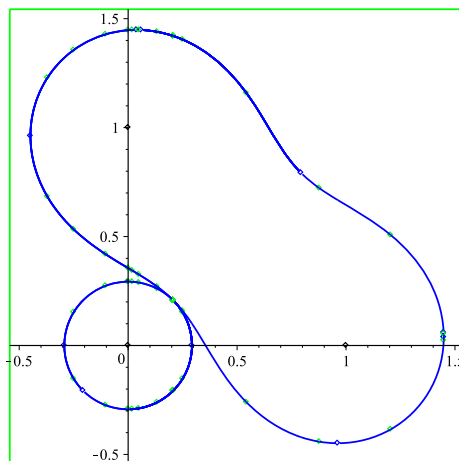


Figura 8.5: Tacnodo en $(x_5, y_5) = \left(\frac{\sqrt{2}-1}{2}, \frac{\sqrt{2}-1}{2} \right)$ a la curva algebraica $V_{\mathbb{R}}(f_{\varepsilon_5})$.

Apéndice A

Curvas algebraicas reales

Al estudiar la frontera del pseudoespectro de una matriz vemos que está contenida en una curva algebraica real plana. Esto nos lleva a considerar los subconjuntos de \mathbb{R}^2 que satisfacen una ecuación de la forma $f(x, y) = 0$ donde $f(x, y)$ es un polinomio de coeficientes reales en las variables x e y . Los resultados mostrados en las dos primeras secciones este apéndice han sido recogidos en su mayor parte del libro de C. G. Gibson [37]. También han sido consultados el libro y los artículos de M.J. de la Puente [23], [24] y [25]. Para las secciones tercera y siguientes véanse allí las referencias consultadas: en especial el libro de Walker [73].

A.1. Introducción

Sean K un cuerpo (conmutativo) y $n \in \mathbb{N}^*$. Recordemos que $\mathbb{N} := \{0, 1, 2, \dots\}$ y $\mathbb{N}^* := \{1, 2, \dots\}$. En lo que sigue, denotaremos por $K[x_1, \dots, x_n]$ el anillo de los polinomios en las variables x_1, \dots, x_n con coeficientes en el cuerpo K .

Definición A.1.1. Un **conjunto algebraico** en el espacio afín K^n se define como

$$\{(x_1, \dots, x_n) \in K^n : f_1(x_1, \dots, x_n) = 0, \dots, f_s(x_1, \dots, x_n) = 0\},$$

para algún $s \in \mathbb{N}^*$ y algunos polinomios $f_1, \dots, f_s \in K[x_1, \dots, x_n]$.

Se denota por $V_K(f_1, \dots, f_s)$. También es llamado el conjunto de los ceros de f_1, \dots, f_s en K^n .

Observación A.1.2. En el caso de $n = 2$, es usual denotar las variables como x, y ; mientras que para $n = 3$, se suelen utilizar x, y, z .

Una **curva algebraica plana** afín sobre K es el conjunto de los ceros en K^2 de un polinomio no constante $f \in K[x, y]$. En nuestro caso, el cuerpo es $K = \mathbb{R}$, por lo tanto, se puede definir una **curva algebraica real plana** como

$$V_{\mathbb{R}}(f) = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$$

donde $f \in \mathbb{R}[x, y]$ es un polinomio de grado ≥ 1 . A veces, se habla de la *curva algebraica* f como abreviatura de $V_{\mathbb{R}}(f)$. En el estudio de las curvas algebraicas reales planas, estamos interesados en las dos siguientes cuestiones:

1. Caracterización de los puntos de autointersección.
2. Irreducibilidad/reducibilidad de las curvas algebraicas reales.

Estas cuestiones serán planteadas al estudiar las fronteras de los pseudoespectros de matrices de tamaño 2×2 y 3×3 . Más en concreto, para las matrices de tamaño 2×2 estudiaremos ambas cuestiones; mientras que para las de tamaño 3×3 , sólo la primera.

A.2. Caracterización de los puntos de autointersección

Recordamos que el grado de un polinomio es el mayor exponente de los términos con coeficiente no nulo.

A.2.1. Aplicaciones afines

Sea K un cuerpo.

Definición A.2.1. Decimos que $\varphi : K^2 \rightarrow K^2$ es una **aplicación afín** si

$$\varphi = \tau \circ \lambda,$$

siendo

$\lambda : K^2 \rightarrow K^2$ una aplicación invertible, y
 τ una traslación, $\tau(x) = x + t$, donde $t \in K^2$ es un vector fijo.

Más brevemente,

$$\varphi(x) = Lx + t, \quad L \in \text{GL}_2(K).$$

Podemos escribir $\varphi(x, y) = (X, Y)$, donde cada componente viene dada por

$$\begin{cases} X = px + qy + u, \\ Y = rx + sy + v, \end{cases}$$

con $ps - qr \neq 0$. El conjunto de las aplicaciones afines de K^2 , $\text{Af}(2)$, forma un grupo multiplicativo con la operación composición de aplicaciones.

Lema A.2.2. Sean $a, b \in K^2$ distintos. Sea φ una aplicación afín de K^2 , entonces $\forall t \in K$,

$$\varphi(a + t(b - a)) = \varphi(a) + t(\varphi(b) - \varphi(a)),$$

$$\varphi((1 - t)a + tb) = (1 - t)\varphi(a) + t\varphi(b).$$

Observación A.2.3. Se pueden interpretar las anteriores expresiones del siguiente modo: φ aplica la recta parametrizada que pasa por a y b en la recta parametrizada que pasa por $\varphi(a)$ y $\varphi(b)$, y conserva el parámetro.

A.2.2. Número de intersección. Recta tangente

Sea f una curva de grado $d > 0$ en K^2 ($f \in K[x, y]$, $\deg(f) = d > 0$); ℓ una recta que no es componente de $V_K(f) := \{(x, y) \in K^2 \mid f(x, y) = 0\}$; y p un punto de ℓ .

Queremos definir el número de intersección en p de f respecto de la recta ℓ , $I(p, f, \ell)$. Para ello, precisamos de dos puntos $a, b \in \ell$ distintos, la función $\phi(t) := f((1-t)a + tb)$, llamada polinomio de intersección, y el punto $p = (1-t_0)a + t_0b$.

Si $a = (a_1, a_2)$ y $b = (b_1, b_2)$, entonces la parametrización de ℓ es

$$\begin{cases} x(t) = (1-t)a_1 + tb_1, \\ y(t) = (1-t)a_2 + tb_2, \end{cases}$$

y el polinomio de intersección a lo largo de la recta ℓ será $\phi(t) = f(x(t), y(t))$, siendo $\deg(\phi) \leq d$.

Definición A.2.4. Llamamos **número (o multiplicidad) de intersección** de una curva f con una recta ℓ en un punto p , $I(p, f, \ell)$, a la multiplicidad de t_0 como raíz del polinomio $\phi(t)$.

Si ℓ no es una componente de $V_K(f)$, entonces $I(p, f, \ell)$ es un entero no negativo; $I(p, f, \ell) = 0$ si y sólo si $p \notin V_K(f) \cap \ell$. Cuando ℓ es una componente de $V_K(f)$ y $p \in \ell$, convenimos que $I(p, f, \ell) = \infty$.

Llamaremos cambio afín de parámetro a la aplicación

$$\begin{aligned} u : K &\longrightarrow K \\ t &\longmapsto u(t) = at + b. \end{aligned}$$

Lema A.2.5. *Supongamos que dos polinomios en una variable, $\phi(t)$ y $\psi(u)$, están relacionados por un cambio afín de parámetro, i.e. $\phi(t) = \psi(u(t))$. Entonces t_0 es una raíz de multiplicidad m de $\phi(t)$ si y sólo si $u(t_0)$ es una raíz de multiplicidad m de $\psi(u)$.*

Lema A.2.6. *$I(p, f, \ell)$ es independiente de la parametrización elegida para ℓ .*

El siguiente lema nos garantiza que los números de intersección son invariantes afines:

Lema A.2.7. *Sea φ una transformación afín de K^2 , p un punto en una recta ℓ , $f(x, y) \in K[x, y]$. Denotamos por $p' = \varphi(p)$, $\ell' = \varphi(\ell)$, $\varphi(V(f)) = V(g)$ con $g(x, y) = f(\varphi^{-1}(x, y))$. Entonces*

$$I(p, f, \ell) = I(p', g, \ell').$$

Definición A.2.8. La **multiplicidad de un punto** $p \in \mathbb{R}^2$ en f , $m_p(f)$, se define como el mínimo número de intersección $I(p, f, \ell)$ cuando ℓ varía en el haz de rectas que pasan por p ; es decir,

$$m_p(f) := \min_{p \in \ell} I(p, f, \ell).$$

Si $m_p(f) = 1$, el punto es **simple**; si $m_p(f) = 2$, **doble**; si $m_p(f) = 3$, **triple**. En general, si $m_p(f) \geq 2$, el punto se llama **singular**. En este caso, se satisface que $f'_x(a, b) = 0$ y $f'_y(a, b) = 0$, donde $p = (a, b)$.

Definición A.2.9. Llamamos **recta tangente** a f en p a toda recta ℓ que pasa por p tal que

$$I(p, f, \ell) > m_p(f).$$

Se dice que una recta ℓ , tangente en p a f , es una **tangente ordinaria** si $I(p, f, \ell) = m_p(f) + 1$, y que es una **tangente de inflexión** si $I(p, f, \ell) \geq m_p(f) + 2$.

A.2.3. Determinación de $m_p(f)$

Se dice que un polinomio $F \in \mathbb{K}[x, y]$ es **homogéneo** si todos sus monomios son de igual grado. Por ejemplo, $F(x, y) = 3x^2y + y^3 + x^3$ es un polinomio homogéneo, pero $G(x, y) = x^4 + xy^2$ no lo es. Una **forma** $F \in K[x, y]$ es un polinomio homogéneo. Todo polinomio $f \in K[x, y]$, con $\deg(f) = d$, se puede expresar de forma única como suma de formas

$$f = F_m + F_{m+1} + \cdots + F_d, \text{ con } m < m+1 < \cdots < d,$$

donde cada F_k , $k = m+1, \dots, d$, es cero o una forma de grado k ; mientras que la forma $F_m \neq 0$, a la cual se suele llamar *forma de términos de grado inferior*, también se denota por TGI.

Podemos determinar la multiplicidad de un punto $p = (a, b)$ en una curva $f(x, y) = 0$ de dos formas:

1. mediante los términos de grado inferior, TGI.
2. mediante derivadas parciales.

Veamos cómo se procede con cada uno de ellos

1. Mediante los términos de grado inferior

Lema A.2.10. Sea $p = (a, b)$ un punto de una curva $f \in K^2$. Sea $g(x, y) := f(a + x, b + y)$. Entonces, $m_p(f)$ es el grado de la forma TGI de g . En particular, $m_{(0,0)}(f)$ es el grado de la forma TGI de f .

2. Mediante derivadas parciales

Primero recordamos la expresión de la fórmula de Taylor par una función polinómica $f \in K[x, y]$, con $\deg(f) = d$, en un punto (a, b)

$$f(a + x, b + y) = \sum_{\substack{(i,j) \in \mathbb{N}^2 \\ 0 \leq i+j \leq d}} \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(a, b) x^i y^j. \quad (\text{A.2.1})$$

El primer lema nos permitirá caracterizar la multiplicidad de puntos pertenecientes a curvas algebraicas mediante sus derivadas parciales. El segundo nos dará una condición necesaria y suficiente para caracterizar los puntos singulares de una curva algebraica; y, finalmente, el último lema, consecuencia del anterior, pondrá de manifiesto que los puntos aislados son necesariamente puntos singulares.

Lema A.2.11. *Sea $p = (a, b)$ un punto de la curva f en K^2 . Entonces $m_p(f) = m$ si y sólo si*

$$\frac{\partial^{i+j} f}{\partial x^i \partial y^j}(a, b) = 0, \text{ si } i + j \leq m,$$

y $\exists i, j$ tales que

$$\frac{\partial^{i+j} f}{\partial x^i \partial y^j}(a, b) \neq 0, \text{ si } i + j = m.$$

Lema A.2.12. *Una condición necesaria y suficiente para que un punto $p = (a, b)$ de una curva f sea singular es que $f(a, b) = 0, f'_x(a, b) = 0, f'_y(a, b) = 0$*

DEMOSTRACIÓN.

Es una consecuencia del Lema A.2.11, ya que

$$m_p(f) \geq 2 \iff \begin{cases} f(a, b) = 0, \frac{\partial f}{\partial x}(a, b) = 0, \frac{\partial f}{\partial y}(a, b) = 0, \\ \text{existe } \frac{\partial^2 f}{\partial x^i \partial y^j}(a, b) \neq 0, \text{ siendo } i + j = 2. \end{cases}$$

□

Lema A.2.13. *Sea f una curva en \mathbb{R}^2 . Si p es un punto aislado de la curva f entonces es un punto singular.*

DEMOSTRACIÓN.

Si p fuera un punto simple, como $f(p) = 0$, por el Lema A.2.12, se tendría que $f'_x(p) \neq 0$ o $f'_y(p) \neq 0$. Por el teorema de la función implícita, se seguiría que la curva $f(x, y) = 0$ podría ser parametrizada en un entorno de p . Con lo cual, p no podría ser aislado. □

A.2.4. Tangentes en un punto múltiple

Sean $p = (a, b) \in \mathbb{R}^2$, $f \in \mathbb{R}[x, y]$, con $m_p(f) = m$. Definimos $g(x, y) := f(a + x, b + y)$, y lo escribimos como suma de formas

$$g = G_m + G_{m+1} + \cdots + G_d,$$

donde cada G_k es una forma de grado k y $G_m \neq 0$. Sea ℓ la recta que pasa por (a, b) en la dirección del vector $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

Nuestro objetivo es estudiar las condiciones que debe satisfacer (α, β) para que la recta ℓ sea una recta tangente a f en (a, b) . Las ecuaciones paramétricas de ℓ son las siguientes

$$\begin{cases} x = a + \alpha t, \\ y = b + \beta t. \end{cases}$$

El polinomio de intersección, $\phi(t)$, se define como

$$\begin{aligned} \phi(t) &:= f(a + \alpha t, b + \beta t) = g(\alpha t, \beta t) = \\ &G_m(\alpha t, \beta t) + G_{m+1}(\alpha t, \beta t) + \cdots + G_d(\alpha t, \beta t) = \\ &t^m G_m(\alpha, \beta) + t^{m+1} G_{m+1}(\alpha, \beta) + \cdots + t^d G_d(\alpha, \beta). \end{aligned} \quad (\text{A.2.2})$$

Podemos ver que 0 es raíz de (A.2.2) con multiplicidad $\geq m+1$ si $G_m(\alpha, \beta) = 0$. Es decir, $\beta x - \alpha y$ es un factor lineal de $G_m(x, y)$. La tangente a f en (a, b) es la recta $\beta(x - a) - \alpha(y - b) = 0$.

A.2.5. Clasificación de puntos dobles

La terminología sobre curvas algebraicas planas cambió drásticamente a lo largo del siglo XX. En el transcurso de la elaboración de esta memoria encontramos una serie de puntos dobles que recibían los nombres de cúspides, cúspides ranfoideas, tacnodos, oscnodos, ... Véase, por ejemplo, la página 131 del libro de Gibson [36]. De gran utilidad nos ha sido la consulta de la página de Wikipedia [74] que reúne un glosario de geometría algebraica clásica. Damos la clasificación de los puntos singulares que estudiamos en esta sección, a saber, los puntos dobles. No es este un asunto que los libros aclaren mucho. Suponemos que ello es debido a la complejidad que puede tener una curva algebraica en el entorno de un punto singular. Citamos varias definiciones, pero no bajamos al detalle de armonizarlas. Cuando a una *definición* la llamemos *concepto* queremos decir que es una descripción intuitiva o heurística de la noción.

Concepto A.2.14. Dada una curva algebraica real C definida como $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ donde $f \in \mathbb{R}[x, y]$ tiene grado mayor o igual que uno. Decimos que un punto doble $p \in C$ es un **crunodo**, o **nodo ordinario**, si C tiene en p dos tangentes reales distintas.

En la Figura A.1 puede verse la curva $C : y^2 - x^3 - x^2 = 0$, que tiene a $(0, 0)$ como punto doble. Este punto es un crunodo para esta curva. Su forma de grado inferior es $y^2 - x^2$; dado que ésta se factoriza así: $y^2 - x^2 = (y+x)(y-x)$, las rectas tangentes a la curva C en $(0, 0)$ son $y = x$ e $y = -x$.

Concepto A.2.15. Dada una curva algebraica real C definida como $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ donde $f \in \mathbb{R}[x, y]$ tiene grado mayor o igual que uno. Decimos que un punto $p \in C$ es un **punto aislado** si existe un entorno $D(p, \varepsilon_0)$ de p tal que $D(p, \varepsilon_0) \cap C = \{p\}$. Cuando el punto aislado p es un **punto doble**, se dice que es un **acnodo**.

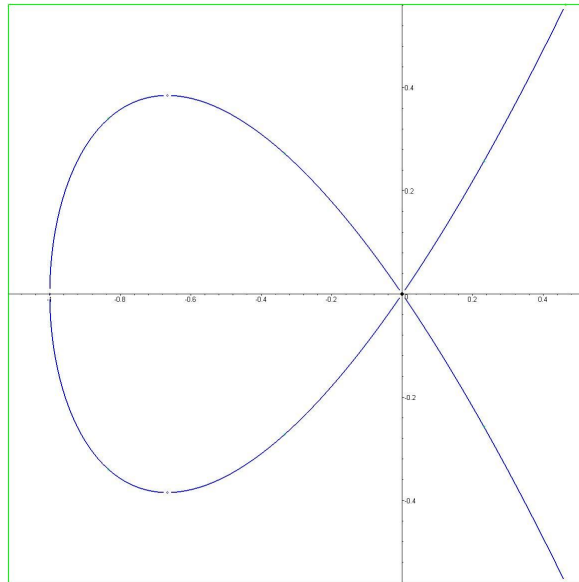


Figura A.1: Curva $y^2 - x^3 - x^2 = 0$; crunodo en $(0, 0)$.

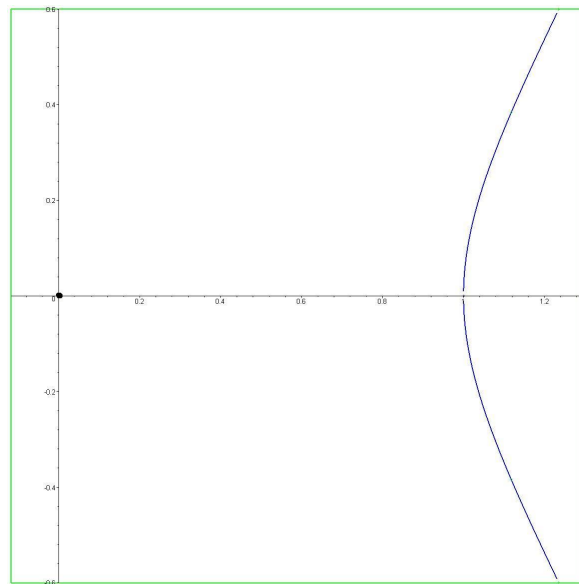


Figura A.2: Curva $y^2 - x^3 + x^2 = 0$; acnodo en $(0, 0)$.

En la Figura A.2 puede verse la curva $y^2 - x^3 + x^2 = 0$ que tiene a $(0, 0)$ como punto doble. Este punto es un acnodo para esta curva.

Concepto A.2.16. Dada una curva algebraica real C definida como $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ donde $f \in \mathbb{R}[x, y]$ tiene grado mayor o igual que uno. Decimos que un punto doble $p \in C$ es un **punto de retroceso, o punto cuspidal** si C tiene en p una tangente real T única y tal que si r es la recta perpendicular a T que pasa por p , existe un valor $\varepsilon_0 > 0$ tal que para todo entorno $D(p, \varepsilon)$ de p con $0 < \varepsilon \leq \varepsilon_0$, la intersección $D(p, \varepsilon) \cap C$ está contenida en uno solo de los dos semiplanos abiertos $S_{r,1}, S_{r,2}$ que r determina.

Llamemos $S_{T,1}, S_{T,2}$ a los semiplanos abiertos determinados por T . Quedan establecidos de esta forma cuatro cuadrantes alrededor del punto p :

$$S_{r,1} \cap S_{T,1}, \quad S_{r,1} \cap S_{T,2}, \quad S_{r,2} \cap S_{T,1}, \quad S_{r,2} \cap S_{T,2}. \quad (\text{A.2.3})$$

Se dice que un punto de retroceso p es de **primera especie** si $D(p, \varepsilon) \cap C$, para todo $\varepsilon \in (0, \varepsilon_0]$, tiene puntos en exactamente dos de estos cuadrantes. Se dice que es de **segunda especie** si $D(p, \varepsilon) \cap C$, para todo $\varepsilon \in (0, \varepsilon_0]$, tiene puntos en exactamente uno de estos cuadrantes. Véase la Figura A.3. Por otra parte, en la Figura A.4 pueden verse: a la izquierda el punto de retroceso de primera especie $(2, 3/2)$ para la curva $5 + 4y - 2yx^2 + 4y^2 - 6x + 2x^3 - 4xy = 0$; y a la derecha el punto de retroceso de segunda especie $(0, 0)$ para la curva $y^2 - 2yx^2 + x^4 - x^5 = 0$. En esta última curva la tangente T a $V_{\mathbb{R}}(f)$ en el punto de retroceso $(0, 0)$ es el eje x y la recta perpendicular r a T que pasa por $(0, 0)$ es el eje y . Los cuatro cuadrantes de (A.2.3) son los cuatro cuadrantes ordinarios cartesianos: se ve que los dos “trozos” de $V_{\mathbb{R}}(f)$ que tienen sus puntos suficientemente cerca del origen, están contenidos en el primer cuadrante ($x > 0, y > 0$).

En el libro de Dixmier [29] con ayuda del desarrollo de Taylor de una función vectorial $\mathbf{r} : (a, b) \rightarrow \mathbb{R}^2$ definida en un intervalo (a, b) de \mathbb{R} y suficientes veces derivable en él, se pueden caracterizar los puntos de retroceso de primera y de segunda especies en un punto $\mathbf{r}(t_0), t_0 \in (a, b)$ de la curva paramétrica $x = r_1(t), y = r_2(t)$, siendo $\mathbf{r}(t) = (r_1(t), r_2(t)), t \in (a, b)$, y en el que $\mathbf{r}'(t_0) = \mathbf{0}$.

Teorema A.2.17. Sea $\mathbf{r}^{(p)}(t_0)$ la primera derivada de \mathbf{r} que no se anula en t_0 ; es decir, que

$$\mathbf{r}'(t_0) = \mathbf{r}''(t_0) = \dots = \mathbf{r}^{(p-1)}(t_0) = \mathbf{0}, \mathbf{r}^{(p)}(t_0) \neq \mathbf{0}, \quad p \geq 2,$$

y sea $\mathbf{r}^{(q)}(t_0)$ el primero de los vectores $\mathbf{r}^{(p+1)}(t_0), \mathbf{r}^{(p+2)}(t_0), \dots$ que no depende linealmente de $\mathbf{r}^{(p)}(t_0)$. Entonces $\mathbf{r}^{(p)}(t_0)$ es un **vector tangente** a la curva en $\mathbf{r}(t_0)$ y se dan los dos casos siguientes.

CASO p PAR, q IMPAR.- Hay en $\mathbf{r}(t_0)$ un punto de retroceso de primera especie.

CASO p PAR, q PAR.- Hay en $\mathbf{r}(t_0)$ un punto de retroceso de segunda especie.

Concepto A.2.18. Dada una curva algebraica real C definida como $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$. Decimos que un punto $p \in C$ es un **tacnodo** si es un punto

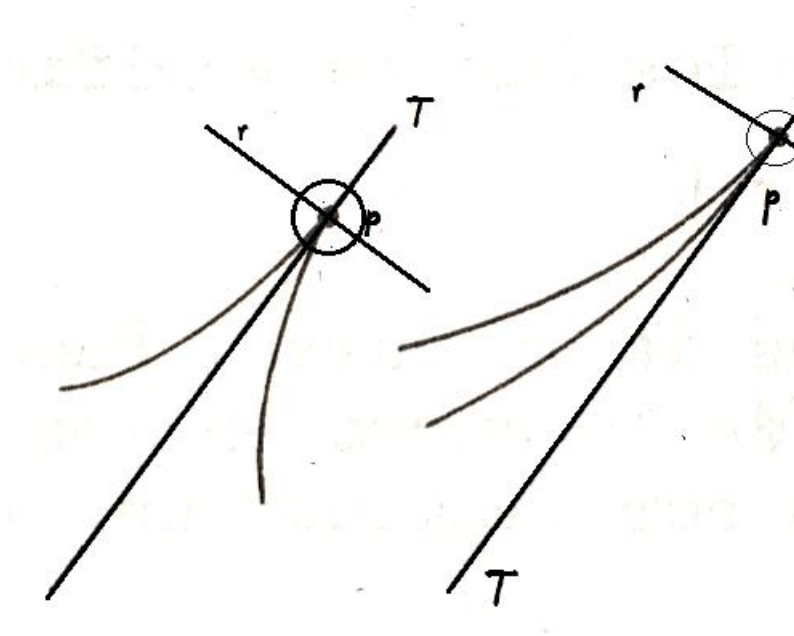


Figura A.3: Puntos de retroceso: 1ª especie (izquierda), 2ª especie (derecha).

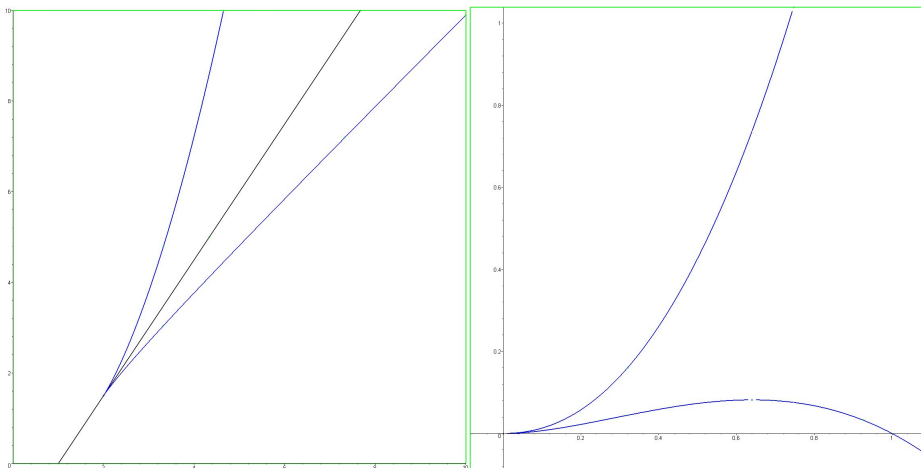


Figura A.4: Puntos de retroceso de primera y segunda especies.

doble con tangente real T única y si r es la recta perpendicular a T que pasa por p , existe un valor $\varepsilon_0 > 0$ tal que en todo entorno $D(p, \varepsilon)$ de p , con $\varepsilon \in (0, \varepsilon_0]$, la curva tiene puntos en la intersección de $D(p, \varepsilon)$ con cada uno de los dos semiplanos abiertos que r determina.

Dicho en términos más intuitivos, p es un tacnodo de C si la curva tiene puntos a ambos lados de r tan cerca como queramos de p . Los tacnodos también son llamados **puntos de autoosculación**.

En la Figura A.5 (Parte Izquierda) puede verse que la curva algebraica real $C = V_{\mathbb{R}}(f) = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$, donde

$$f(x, y) := (x^2 + y^2 - 1) \left((x - 3/2)^2 + (y - 3/2)^2 - \frac{(3 - \sqrt{2})^2}{2} \right),$$

tiene un tacnodo en $p = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Asimismo, la curva algebraica real

$$g(x, y) := \left((x - 5/2)^2 + (y - 5)^2 - \frac{45}{4} \right) \left((x - 2)^2 + (y - 4)^2 - 5 \right)$$

tiene un tacnodo en $p = (1, 2)$ (Parte Derecha).

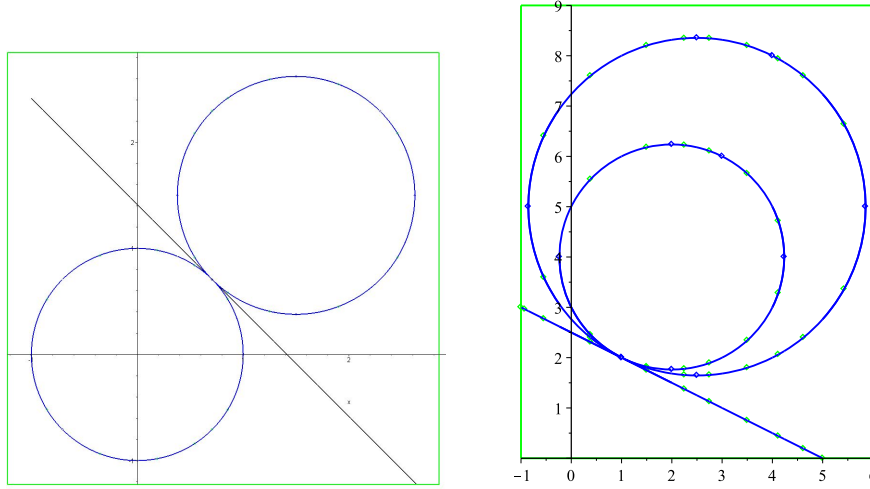


Figura A.5: Ejemplo de puntos tacnodos.

Definición A.2.19 (Véanse M.J. de la Puente[25], Definición 5.37, pág. 211, y el artículo de Wikipedia [75]). Sea $C = V_{\mathbb{R}}(f)$, con $f \in \mathbb{R}[x, y]$, una curva algebraica real y sea $p \in C$. Se dice que p es un **punto cuspidal** si existe un cambio afín de coordenadas

$$\begin{cases} x = aX + bY + \alpha, \\ y = cX + dY + \beta, \end{cases} \quad \text{con } a, b, c, d, \alpha, \beta \in \mathbb{R} \text{ y } ad - bc \neq 0$$

tal que p se transforma en $(0, 0)$ y el polinomio f en

$$g(X, Y) = Y^2 - X^{2n+1} + \text{términos de grado } \geq 2n + 2.$$

donde n es un entero ≥ 1 . Como $g'_X(0, 0) = 0$, $g'_Y(0, 0) = 0$ y $g''_{Y^2}(0, 0) = 2 \neq 0$, el punto $(0, 0)$ es *doble* para la curva g . Por lo tanto, en virtud del Lema A.2.7, el punto cuspidal p es necesariamente un punto doble. Cuando $n = 1$, el punto p se llama **cuspidal ordinario**.

Definición A.2.20 (Gibson [37] Ejercicio 7.4.3, página 93). Sea $C = V_{\mathbb{R}}(f)$, con $f \in \mathbb{R}[x, y]$, una curva algebraica real y sea $p \in C$ un punto doble. Se dice que p es un **punto cuspidal ordinario** si C tiene una tangente real única T tal que $I(p, f, T) = 3$.

Como p es un punto doble, se tiene que $m_p(f) = 2$. Por tanto, si $I(p, f, T) = 3$, se tiene que

$$I(p, f, T) = m_p(f) + 1;$$

esto es, C tiene en p una tangente *ordinaria*.

Definición A.2.21 (Wikipedia [75]). Sea $C = V_{\mathbb{R}}(f)$, con $f \in \mathbb{R}[x, y]$, una curva algebraica real y sea $p \in C$. Se dice que p es un **tacnodo** si existe un cambio afín de coordenadas

$$\begin{cases} x = aX + bY + \alpha, \\ y = cX + dY + \beta, \end{cases} \quad \text{con } ad - bc \neq 0$$

tal que p se transforma en $(0, 0)$ y el polinomio f en

$$g(X, Y) = Y^2 - X^{2n} + \text{términos de grado } \geq 2n + 1.$$

donde n es un entero ≥ 2 . Como $(0, 0)$ es un punto doble de la curva g , por el Lema A.2.7 el punto p es *doble*. En el caso en que $n = 2$, el punto p se llama **tacnodo ordinario**.

DEFINICIONES DEFINITIVAS

Vamos a establecer unas definiciones definitivas de a qué puntos dobles llamamos **crunodos**, **acnodos**, **cúspides (puntos de retroceso)** y **tacnodos** en función de dos parámetros: el signo de $\Delta(p)$ y el número $r(p)$ de **ramas** de la curva $V_{\mathbb{C}}(f)$ **que tienen por centro al punto** $p \in C := V_{\mathbb{R}}(f)$, donde $f \in \mathbb{R}[x, y]$, $\deg(f) \geq 1$, y $p = (a, b) \in \mathbb{R}^2$ un **punto doble** de f : es decir, que $f(p) = 0$, $f'_x(p) = 0$, $f'_y(p) = 0$ y alguna de las derivadas parciales segundas de f en p es no nula. Definimos las siguientes cantidades:

$$\begin{aligned} A &:= f''_{x^2}(a, b), & B &:= f''_{xy}(a, b), & C &:= f''_{y^2}(a, b), \\ \Delta(a, b) &:= B^2 - AC. \end{aligned} \tag{A.2.4}$$

Definición A.2.22. Sean $f \in \mathbb{R}[x, y]$, $\deg(f) \geq 1$, y $p = (a, b) \in \mathbb{R}^2$ un *punto doble* de la curva *complejificada*¹ $V_{\mathbb{C}}(f)$. Se dice que $p \in C := V_{\mathbb{R}}(f)$ es: un **crunodo** de C si $\Delta(p) > 0$; un **acnodo** si $\Delta(p) < 0$; una **cúspide** o **punto de retroceso** si $\Delta(p) = 0$ y $V_{\mathbb{C}}(f)$ tiene una rama centrada en p ; un **tacnodo** si $\Delta(p) = 0$ y $V_{\mathbb{C}}(f)$ tiene dos ramas centradas en p .

En resumen, si

$$\left\{ \begin{array}{ll} \Delta(p) > 0 & \text{el punto } p \text{ es un crunodo;} \\ \Delta(p) < 0 & \text{el punto } p \text{ es un acnodo;} \\ \Delta(p) = 0, r(p) = 1 & \text{el punto } p \text{ es una cúspide;} \\ \Delta(p) = 0, r(p) = 2 & \text{el punto } p \text{ es un tacnodo.} \end{array} \right.$$

Hay un inconveniente serio en estas definiciones y es que exigen definir qué se entiende por **rama** de una curva $V_{\mathbb{C}}(f)$ para un $f \in \mathbb{R}[x, y]$ **que tiene por centro a un punto** de la misma. Este es un concepto elusivo que se refiere a los “segmentos” o “trozos” de la curva que pasan por el punto considerado p y que son homeomorfos a un intervalo $(-\varepsilon, \varepsilon)$ de \mathbb{R} con $\varepsilon > 0$ y donde p se corresponde con el número real 0. Para ello precisamos de los conceptos de series formales de potencias, de Laurent y de Puiseux, así como del Teorema de Newton-Puiseux. Las principales referencias que seguimos para escribir este resumen contextual son los libros de M.J. de la Puente [25], de R.J. Walker [73], de S.S. Abhyankar [1] y de G. Fisher [34].

A.3. Series formales de potencias

El conjunto de aplicaciones de un conjunto cualquiera X en otro conjunto cualquiera Y se denota por Y^X . Sea K un cuerpo cualquiera. Siempre entendemos que un cuerpo es conmutativo, i.e. que su multiplicación es conmutativa. Llamamos **serie formal de potencias** con coeficientes en K a cualquier sucesión

$$a = (a_0, a_1, \dots, a_i, \dots)$$

con cada $a_i \in K, i = 0, 1, \dots$. Dicho de otro modo, a es una aplicación de \mathbb{N} en K . Dada cualquier otra serie formal de potencias

$$b = (b_0, b_1, \dots, b_j, \dots)$$

se definen su suma $a + b := (a_0 + b_0, a_1 + b_1, \dots)$ y su producto $ab := (c_0, c_1, \dots)$ por medio de

$$c_k := \sum_{\substack{(i,j) \in \mathbb{N}^2 \\ i+j=k}} a_i b_j$$

¹Si $f \in \mathbb{R}[x, y]$ por la curva complejificada de f entendemos el subconjunto de \mathbb{C}^2 formado por los puntos $(z, w) \in \mathbb{C}^2$ tales que $f(z, w) = 0$.

para todo $k = 0, 1, \dots$. Con estas dos operaciones el conjunto $K^{\mathbb{N}}$ de aplicaciones de \mathbb{N} en K llega a ser un anillo.

Si a tiene la forma

$$a = (a_0, a_1, \dots, a_n, 0, 0, \dots)$$

identificaremos a con un polinomio en una variable con coeficientes en K . En particular, las series “constantes”, i.e. los elementos c de K son también series formales haciendo la identificación $c = (c, 0, 0, \dots)$. Una de estas series es el elemento unidad $1 := (1, 0, 0, \dots)$. De ahora en adelante, dejaremos este formalismo y escribiremos una serie formal de potencias por el símbolo

$$a = a(T) = \sum_{n=0}^{\infty} a_n T^n$$

como una serie de potencias en una variable T . Cuando no haya confusión abreviaremos **serie formal de potencias** por **serie**. No se considera ninguna cuestión de *convergencia* de estas series. No se tiene el concepto de grado como en los polinomios, sino el de **orden** de a , $\text{ord}(a)$. Se llama así a

$$\text{ord}(a) := \begin{cases} \text{mín}\{i \in \mathbb{N}: a_i \neq 0\}, & \text{si } a \neq 0 := (0, 0, \dots) \\ \infty & \text{si } a = 0. \end{cases}$$

Dicho de otro modo, si $a \neq 0$, que $\text{ord}(a) = p$ quiere decir que $a_0 = a_1 = \dots = a_{p-1} = 0$ y $a_p \neq 0$. Así pues, *el orden de una serie formal de potencias no nula es un número natural*. Denotaremos por $K[[T]] \equiv K^{\mathbb{N}}$ el conjunto de las series formales de potencias con coeficientes en K en la variable T . Es sabido que con las operaciones de suma y producto antes definidas $K[[T]]$ tiene estructura de anillo. Más aun, $K[[T]]$ no tiene divisores de cero; es decir, $a, b \in K[[T]]$, $ab = 0$ implica que $a = 0$ o $b = 0$. Además, es fácil probar que las **unidades** o **elementos invertibles** de $K[[T]]$ son las series de orden 0; es decir, $a(T) = a_0 + a_1 T + a_2 T^2 + \dots$ es invertible si y sólo si $a_0 \neq 0$. Una familia de series formales $\{s_i(T)\}_{i \in I}$, donde I designa un conjunto de índices, se dice **sumable** si para cada $k \in \mathbb{N}$ existe a lo más un número finito de índices i tales que $\text{ord}(s_i) \leq k$. Por definición, la **suma** de una familia sumable de series formales

$$s_i(T) = \sum_{n=0}^{\infty} a_{ni} T^n$$

es la serie

$$s(T) = \sum_{n=0}^{\infty} a_n T^n,$$

donde, para cada n , $a_n = \sum_{i \in I} a_{ni}$. Esto tiene sentido puesto que por hipótesis los a_{ni} para un n dado, son nulos salvo para un número finito de valores de i . En particular, dados $a_0, a_1, \dots \in K$, la familia $\{a_n T^n\}_{n \in \mathbb{N}}$ es sumable y las dos

posibles interpretaciones de la expresión

$$\sum_{n=0}^{\infty} a_n T^n$$

coinciden.

A.3.1. Sustitución de una serie en otra

Para esta subsección seguimos estrechamente el Capítulo 1 del libro de Cartan [18]. Sean las series formales de potencias

$$a(Z) = \sum_{n=0}^{\infty} a_n Z^n \in K[[Z]], \quad b(T) = \sum_{n=0}^{\infty} b_n T^n \in K[[T]].$$

Queremos darle sentido a la sustitución de Z por $b(T)$ en $a(Z)$ para obtener una serie $a(b(T))$ que pertenezca a $K[[T]]$. La **sustitución de T por b por Z en a** solo tiene sentido cuando sea sumable la familia $\{c_n\}_{n \in \mathbb{N}}$, donde

$$c_n = a_n \left(\sum_{j=0}^{\infty} T^j \right)^n = a_n b_0^n + \dots$$

Se presentan dos casos:

- si a es un polinomio, entonces la familia $\{c_n\}_{n \in \mathbb{N}}$ es finita; por tanto sumable.
- Si a no es polinomio, entonces la familia $\{c_n\}_{n \in \mathbb{N}}$ es sumable si y sólo si $b_0 = 0$, si y sólo si $\text{ord}(b) \geq 1$; lo que es esencial.

En particular la **traslación** (i.e. sustituir T por $T - b_0$, con $b_0 \neq 0$) tiene sentido solo cuando a sea un polinomio.

El resultado de la sustituir la serie formal $b(T)$ por Z en $a(Z)$ se puede denotar por

$$a(b(T)) = \sum_{n=0}^{\infty} a_n (b(T))^n,$$

en la que se reagruparán los términos en T . A la serie $a(b(T))$ también la denotaremos por $a \circ b$ cuando no se especifica la variable o indeterminada T . Esta “composición” \circ goza de las siguientes propiedades en relación con la suma y el producto ordinario de series:

$$\begin{cases} (a_1 + a_2) \circ b = a_1 \circ b + a_2 \circ b, \\ (a_1 a_2) \circ b = (a_1 \circ b)(a_2 \circ b), \quad 1 \circ b = 1. \end{cases}$$

para todas series formales $a_1, a_2 \in K[[Z]]$ y $b \in K[[T]]$ tales que $\text{ord}(b) \geq 1$. Pero, no debemos creer que $a \circ (b_1 + b_2)$ sea igual a $a \circ b_1 + a \circ b_2$ aunque $\text{ord}(b_1) \geq 1, \text{ord}(b_2) \geq 1$ y $\text{ord}(b_1 + b_2) \geq 1$.

Observación A.3.1. Si se sustituye 0 en $a(T) = \sum_{n=0}^{\infty} a_n T^n$, se encuentra la serie formal a_0 reducida a su “término constante”. Notemos que $\text{ord}(0) := \infty$; por tanto $\text{ord}(0) \geq 1$ y la operación de sustitución $a \circ 0 = a_0$ tiene sentido. Preferimos escribir, $a(0) = a_0$. Con estas notaciones: *Para que $a(T) = \sum_{n=0}^{\infty} a_n T^n$ posea un elemento inverso para la multiplicación ordinaria de $K[[T]]$ es necesario y suficiente que $a_0 \neq 0$, es decir $a(0) \neq 0$*

Asociatividad de la sustitución. Se tiene que $(a \circ b) \circ c = a \circ (b \circ c)$ siempre que $\text{ord}(b) \geq 1$, $\text{ord}(c) \geq 1$.

DERIVADA DE UNA SERIE FORMAL

Para definir la derivada de una serie formal no hacen faltan límites. He aquí la definición: Sea $a(T) = \sum_{n=0}^{\infty} a_n T^n$, su **serie derivada** $a'(T)$ viene dada por la fórmula

$$a'(T) := \sum_{n=0}^{\infty} (n+1)a_{n+1}T^n.$$

Se escribe también

$$\frac{d a}{d T}, \text{ o incluso } \frac{d}{d T} a,$$

para la derivada a' . Se tienen las propiedades $(a+b)' = a' + b'$; $(ab)' = a'b + ab'$. Si $a(0) \neq 0$, se tiene que

$$\frac{d}{d T} \left(\frac{1}{a} \right) = -\frac{1}{a^2} \frac{d a}{d T}.$$

Por recurrencia, se definen las **derivadas sucesivas** de una serie formal. Si $a(T) = \sum_{n=0}^{\infty} a_n T^n$, la derivada de orden n es

$$a^{(n)}(T) = n!a_n + \text{términos de grado } \geq 1.$$

Por tanto, $a^{(n)}(0) = n!a_n$, denotando por $a^{(n)}(0)$ el resultado de la sustitución de la serie 0 por la variable T en $a^{(n)}(T)$.

SERIES RECÍPROCAS

La serie $I(T)$ definida por $I(T) = T$ es el **elemento neutro** para la composición de series formales:

$$a \circ I = a = I \circ a.$$

Proposición A.3.2 (Función inversa formal). *Sea dada una serie formal a ; para que exista una serie formal b tal que*

$$b(0) = 0, \quad a \circ b = I,$$

es necesario y suficiente que

$$a(0) = 0, \quad a'(0) \neq 0.$$

Si es así, b es única, y $b \circ a = I$; dicho de otro modo, b es la inversa o recíproca de a para la ley de composición \circ .

No se debe confundir esta *serie recíproca* de a con la inversa $1/a$ para el producto ordinario. De hecho, como $a(0) \neq 0$ es una condición necesaria y suficiente para que tenga sentido $1/a$, cuando existe ésta, la serie recíproca no está definida.

A.3.2. Series de Laurent y de Puiseux

El cuerpo de fracciones del anillo de integridad $K[[T]]$, lo denotaremos por $K((T))$ y está formado por los cocientes $\frac{a}{b}$, donde $a, b \in K[[T]]$ y b no es nulo. Por el Teorema 1.2, página 88 de Walker [73] se tiene que cualquier elemento f de $K((T))$ puede expresarse en la forma

$$f = \frac{a_0 + a_1T + a_2T^2 + \dots}{T^r}$$

con r entero ≥ 0 y $a_0 \neq 0$. Esta representación es única. Como Corolario se obtiene que toda $f \in K((T))$ se puede representar de manera única

$$f = \sum_{k=-\infty}^{\infty} a_k T^k$$

con coeficientes $a_k \in K$, con la condición de que sólo haya un conjunto finito de números enteros *negativos* k tales que $a_k \neq 0$. Podemos pensar que f es la sucesión doblemente infinita

$$f \equiv (\dots, 0, 0, a_{-r}, \dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots, a_i, \dots)$$

con $r \geq 0$. Así pues, podemos mirar a $K((T))$ como el subconjunto del conjunto $K^{\mathbb{Z}}$ de las aplicaciones f de \mathbb{Z} en K tales que existe un entero $s = O(f) \in \mathbb{Z}$ (*orden de f*) tal que $f(s) \neq 0$ y $f(k) = 0$ para todo $k < O(f)$. Esto también lo expresaremos diciendo que $f(k)$ es nulo para *casi todo* $k \in \mathbb{Z}_- := \{n \in \mathbb{Z} : n < 0\}$. A cada f se la llama **serie formal de Laurent en T** . Sean

$$f := \sum_{k=-\infty}^{\infty} f(k)T^k, \quad g := \sum_{k=-\infty}^{\infty} g(k)T^k.$$

La suma y el producto en $K((T))$ coinciden con

$$f + g = \sum_{k=-\infty}^{\infty} (f(k) + g(k))T^k$$

y el producto de *convolución* $h = fg$ que viene dado por

$$\forall k \in \mathbb{Z}, \quad h(k) = \sum_{\substack{(i,j) \in \mathbb{Z}^2 \\ i+j=k}} f(i)g(j).$$

Continuando con esta línea formalista, definiremos las series de Puiseux con coeficientes en K en una variable T , según se indica en los Ejercicios 7 y 8 de la pág. 595 del libro de N. Jacobson[43]. Sean K un cuerpo cualquiera, \mathbb{Q} el cuerpo de los números racionales y sea $K^{\mathbb{Q}}$ el conjunto de aplicaciones de \mathbb{Q} en K . Para cada $f \in K^{\mathbb{Q}}$ definamos el **soporte** de f como

$$\text{Sop}(f) := \{\alpha \in \mathbb{Q} : f(\alpha) \neq 0\}.$$

Denotemos por $\mathbb{P}(K)$ el subconjunto de $K^{\mathbb{Q}}$ formado por los f tales que

- (1) $\text{Sop}(f) \subset \mathbb{Z}n^{-1}$ para algún entero positivo n (dependiente de f);
- (2) $\text{Sop}(f)$ está acotado inferiormente.

Los elementos de $\mathbb{P}(K)$ pueden ser representados como series formales

$$f = \sum_{\alpha \in \text{Sop}(f)} f(\alpha)T^\alpha$$

y éstas son llamadas **series formales de Newton-Puiseux**. Si $f, g \in \mathbb{P}(K)$ definimos $f + g$ y fg mediante

$$(f + g)(\alpha) := f(\alpha) + g(\alpha),$$

$$(fg)(\alpha) := \sum_{\substack{(\beta, \gamma) \in \mathbb{Q}^2 \\ \beta + \gamma = \alpha}} f(\beta)g(\gamma), \text{ (que está bien definida)}$$

para todo $\alpha \in \mathbb{Q}$. Definimos 0 por $0(\alpha) = 0$ para todo α y 1 por

$$1(\alpha) = \begin{cases} 1 & \text{si } \alpha = 0, \\ 0 & \text{si } \alpha \neq 0. \end{cases}$$

Antes de continuar adelante, vamos a introducir una presentación más intuitiva de las series de Puiseux. Una serie de Puiseux en X es una suma formal de elementos de la forma $aX^{m/n}$ donde $a \in K$ y m/n es un número racional. Por ejemplo, $2 + 3X^{-3/2} + 2X^{1/2} - 6X^{5/2}$, y $1 + X^{1/3} + X^{1/5} + X^{4/3} + X^{2/5}$ son series de Puiseux con coeficientes en \mathbb{Q} , pero

$$1 + X^{1/2} + X^{2/3} + X^{3/4} + \dots$$

no lo es, ya que los exponentes no pueden tener un denominador común. En una serie de Puiseux puede haber potencias de X con exponentes fraccionarios: exponentes que pueden ser negativos; pero sólo puede haber un número finito de exponentes negativos; potencias con exponentes positivos m/n puede haber infinitas, pero sólo ha de haber un número finito de distintos denominadores n . Es decir, una serie de Puiseux tiene la forma

$$p := \sum_{i=1}^q \sum_{j=1}^{\infty} a_{ij} X^{m_{ij}/n_i}$$

donde $a_{ij} \in K$, n_1, \dots, n_q son enteros > 0 y $m_{ij} \in \mathbb{Z}$. Se supone que los símbolos $X^{m/n}$ están sujetos a operaciones “coherentes”:

$$X^{1/1} = X, \quad (X^{1/rn})^r = X^{1/n}, \quad X^{m/n} = (X^{1/n})^m, \quad X^{rm/rn} = X^{m/n}.$$

Como en una serie de Puiseux sólo aparece un número finito de diferentes denominadores: n_1, \dots, n_q , podemos escribir toda serie de Puiseux como potencias con exponentes que tengan *común* denominador: digamos $N := n_1 \cdots n_q$

$$p := a_{-r} X^{-r/N} + \dots + a_{-1} X^{-1/N} + a_0 + a_1 X^{1/N} + a_2 X^{2/N} + \dots + a_k X^{k/N} + \dots,$$

donde $r > 0$ y $a_i \in K$, $i = -r, -(r-1), \dots, -1, 0, 1, 2, \dots, k, \dots$. Los coeficientes $a_i \in K$ pueden ser algunos iguales a cero. Si nos fijamos bien, podemos considerar la serie de Puiseux p como una serie de Laurent en $X^{1/N}$: De este modo, el conjunto de las series de Puiseux es el conjunto

$$K\{\{X\}\} := \bigcup_{N=1}^{\infty} K((X^{1/N})).$$

Se puede demostrar que con estas operaciones $K\{\{X\}\} = \mathbb{P}(K)$ es un cuerpo. Recordemos que un cuerpo \mathbb{F} es *algebraicamente cerrado* si todo polinomio $f(X) \in \mathbb{F}[X]$ de grado ≥ 1 admite raíces en \mathbb{F} . Asimismo se dice que un cuerpo \mathbb{F} tiene *característica 0* si toda suma de un número finito de veces el elemento 1 nunca es igual a 0: para todo $n \geq 1$, $1 + \dots + 1 \neq 0$. Con estos preliminares se tiene el importante

Teorema A.3.3 (Newton-Puiseux(1676-1850)). *Supongamos que K es un cuerpo algebraicamente cerrado de característica 0. Entonces el cuerpo de las series de Puiseux $K\{\{X\}\} = \mathbb{P}(K)$ es algebraicamente cerrado.*

Una demostración constructiva de este teorema fundamental puede verse en el libro clásico de Walker [73]. Versiones más rigurosas aparecen en el libro de Fischer [34]. Así pues, designaremos el cuerpo de las series formales de Puiseux con coeficientes en K mediante una doble llave: $K\{\{T\}\}$, remedando el uso de $K[[T]]$ y $K((T))$ para las series formales, de potencias y de Laurent en T .²

Lo que el Teorema A.3.3 significa es que si $f(X, Y)$ es un elemento de $K\{\{X\}\}[Y]$ que no pertenece a $K\{\{X\}\}$, entonces existe un $\bar{y} \in K\{\{X\}\}$ tal que $f(X, \bar{y}) = 0$ en $K\{\{X\}\}$. Una serie cualquiera de Puiseux también puede escribirse así

$$\bar{a}(X) = a_1 X^{m_1/n_1} + a_2 X^{m_2/n_2} + \dots \in K\{\{X\}\},$$

donde

$$a_i \neq 0, \quad m_1/n_1 < m_2/n_2 < \dots.$$

Definimos el **orden** de \bar{a} como $O(\bar{a}) := m_1/n_1$; es un número racional. El conjunto de elementos de $K\{\{X\}\}$ de orden ≥ 0 es un dominio de integridad: a saber

$$\bigcup_{n=1}^{\infty} K[[X^{1/n}]] = \{\bar{a} \in K\{\{X\}\} \mid O(\bar{a}) \geq 0\}.$$

Si $\bar{a} \in \bigcup_{n=1}^{\infty} K[[X^{1/n}]]$, el coeficiente de X^0 en $\bar{a}(X)$ será denotado por $\bar{a}(0)$. Obviamente $\bar{a}(0) \neq 0$ si y sólo si $O(\bar{a}) = 0$. Si 0 es la serie de Puiseux cero, definimos $O(0) := \infty$.

Siguiendo a Walker [73], Teoremas 3.2 a 3.5 en la pág. 106, se tienen las siguientes extensiones simples del Teorema A.3.3.

² Aunque hay otros autores [9] que usan $K\{\{T\}\}$ para el *anillo (o dominio) de integridad* $\bigcup_{n=1}^{\infty} K[[T^{1/n}]]$.

Teorema A.3.4. *Supongamos que*

$$f(X, Y) = \bar{a}_0 + \bar{a}_1 Y + \cdots + \bar{a}_n Y^n, \quad \bar{a}_i \in K\{\{X\}\}, \quad \bar{a}_n \neq 0.$$

Entonces:

- *existe una única n -tupla desordenada $(\bar{y}_1, \dots, \bar{y}_n) \in K\{\{X\}\}^n$ tal que*

$$f(X, Y) = \bar{a}_n \prod_{j=1}^n (Y - \bar{y}_j);$$

- *si $O(\bar{a}_n) \leq O(\bar{a}_i), i = 0, \dots, n-1$, los \bar{y}_j del apartado anterior son elementos de $\bigcup_{n=1}^{\infty} K[[X^{1/n}]]$;*
- *si $O(\bar{a}_n) = 0, O(\bar{a}_0) > 0, O(\bar{a}_i) \geq 0, i = 1, \dots, n-1$, entonces para al menos uno de los \bar{y}_j , $O(\bar{y}_j) > 0$;*
- *si $f(X, Y) \in K[X, Y]$ no tiene divisores de la forma $d(X) \in K[X]$ de grado ≥ 1 , entonces f tiene un factor múltiple en $K[X, Y]$ si y sólo si $f(X, Y) = 0$ tiene una raíz múltiple en $K\{\{X\}\}$.*

Por si acaso no está claro, se tiene que: todo polinomio en X es una serie formal de potencias en X ; toda serie formal de potencias en X es una serie de Laurent en X ; toda serie de Laurent en X es una serie de Puiseux en X . Es decir, que

$$K \subset K[X] \subset K[[X]] \subset K((X)) \subset K\{\{X\}\}.$$

Finalmente, daremos una versión del Teorema de Newton que sigue el Teorema de la página 89 del libro intuitivo de Abhyankar [1] que viene muy bien a nuestros propósitos.

Teorema A.3.5. *Sea K un cuerpo algebraicamente cerrado de característica 0, y sea*

$$f(X, Y) = Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X) \in K((X))[Y]$$

un polinomio mónico de grado $n > 0$ en Y con coeficientes $a_1(X), \dots, a_n(X)$. Entonces existen un entero positivo m y n series de Laurent $\eta_i(T) \in K((T))$, $i = 1, 2, \dots, n$ tales que

$$f(T^m, Y) = \prod_{i=1}^n (Y - \eta_i(T)). \tag{A.3.5}$$

También se tienen los dos suplementos siguientes.

Suplemento 1. *Si $a_i(X) \in K[[X]]$ para $i = 1, 2, \dots, n$, entonces $\eta_i(T) \in K[[T]]$ para $i = 1, 2, \dots, n$.*

Suplemento 2. *Si $f(X, Y)$ es irreducible en $K((X))[Y]$, entonces $m = n$. De hecho, en este caso, n es el único m permisible. Además, las raíces $\eta_1(T), \dots, \eta_n(T)$ pueden ser renombradas de manera que $\eta_i(T) = \eta_1(\omega^i T)$ donde ω es una raíz n -ésima primitiva de 1. (Una raíz n -ésima ω de 1 se dice **primitiva** si todas las otras raíces n -ésimas de 1 son potencias de ella. Por ejemplo, en el caso de que $K = \mathbb{C}$ podemos tomar $\omega = e^{2\pi i/n}$.)*

La ecuación (A.3.5) implica que para todo $i = 1, 2, \dots, n$,

$$f(T^m, \eta_i(T)) = 0.$$

Habíamos visto que el conjunto de series de Puiseux de orden ≥ 0 forman un dominio de integridad, denotado: $\cup_{n=1}^{\infty} K[[X^{1/n}]]$. Pues bien, cuando K es un cuerpo algebraicamente cerrado de característica 0, el Teorema de Newton-Puiseux garantiza que el cuerpo de fracciones de este dominio es algebraicamente cerrado. De hecho,

$$\left\{ \frac{p(X)}{q(X)} \mid p(X) = \sum_{i=0}^{\infty} a_i X^{i/N}, q(X) = \sum_{j=0}^{\infty} b_j X^{j/M}, \text{ con } q \neq 0 \right\} = K\{\{X\}\}.$$

A.4. Ramas de una curva algebraica

Sea $f \in \mathbb{R}[X, Y]$ un polinomio mónico en Y . Una **parametrización (local)** de la curva algebraica $C := V_{\mathbb{R}}(f)$ en el espacio afín \mathbb{R}^2 es un par de series formales de potencias no ambas constantes $(\tilde{x}(T), \tilde{y}(T)) \in \mathbb{C}[[T]] \times \mathbb{C}[[T]]$ tales que $f(\tilde{x}(T), \tilde{y}(T)) = 0$ como elemento de $\mathbb{C}[[T]]$. El punto $(\tilde{x}(0), \tilde{y}(0)) \in \mathbb{C}^2$ se llama el **centro** de la parametrización, y claramente pertenece a la curva compleja definida por f . Una **parametrización local** se dice **irreducible** si no existe $k \in \mathbb{N}, k > 1$ tal que

$$\tilde{x}(T), \tilde{y}(T) \in \mathbb{C}[[T^k]].$$

EJEMPLO. La parametrización

$$(T^2, T^2 - \frac{1}{2}T^6 - \frac{1}{8}T^{10} - \frac{1}{16}T^{14} - \frac{5}{128}T^{18} + \dots)$$

de la *lemniscata de Huygens* $X^2 - Y^2 - X^4 = 0$ es reducible.

La prueba del teorema siguiente puede verse en Walker [73], Teorema 2.1, págs. 94 y 95.

Teorema A.4.1. *La parametrización (\tilde{x}, \tilde{y}) dada por*

$$\tilde{x}(T) = T^n, \quad \tilde{y}(T) = a_1 T^{n_1} + a_2 T^{n_2} + \dots,$$

donde

$$0 < n, \quad 0 < n_1 < n_2 < \dots, a_i \neq 0,$$

es irreducible si y sólo si el máximo común divisor de los enteros n, n_1, n_2, \dots es 1.

Para que la definición siguiente tenga sentido recordemos que sólo tiene sentido sustituir la variable T de una serie de potencias por otra serie de potencias

$\alpha(T)$, cuando el orden de α , $\text{ord}(\alpha)$, es positivo. Se dice que dos **parametrizaciones locales** $(\tilde{x}(T), \tilde{y}(T))$, $(\tilde{x}_1(T), \tilde{y}_1(T))$ de una curva $C \subset \mathbb{R}^2$ son **equivalentes** si existe una serie $\alpha(T) \in \mathbb{C}[[T]]$ de orden 1 tal que

$$\tilde{x}_1(T) = \tilde{x}(\alpha(T)), \quad \tilde{y}_1(T) = \tilde{y}(\alpha(T)).$$

Como $\text{ord}(\alpha) = 1$, se tiene que $\alpha(0) = 0$ y $\alpha'(0) \neq 0$. Por tanto, en virtud de la Proposición A.3.2 existe la serie recíproca de α respecto de \circ y la relación anterior es una relación de equivalencia. Se deduce de $\alpha(0) = 0$ que el **centro** $(\tilde{x}(0), \tilde{y}(0))$ de la parametrización $(\tilde{x}(T), \tilde{y}(T))$ es igual al centro $(\tilde{x}_1(0), \tilde{y}_1(0))$ de la parametrización $(\tilde{x}_1(T), \tilde{y}_1(T))$. Tiene sentido, pues, la definición siguiente.

Definición A.4.2. Sea p el centro de la parametrización $(\tilde{x}(T), \tilde{y}(T))$. Una **rama de la curva C con centro el punto p** es una clase de equivalencia de tales parametrizaciones locales irreducibles con centro p . A veces, cuando digamos **rama de la curva C que pasa por el punto p** , sobrentenderemos que quiere decir *centrada en p* , por comodidad del lenguaje.

A.5. Naturaleza de los puntos dobles

El teorema siguiente nos permite una primera clasificación de los puntos dobles $p \in C$ atendiendo al número 0, 1, 2 de tangentes reales a la curva C en p , y a los números de intersección en p de la curva C con su recta tangente en el caso de una sola tangente.

Teorema A.5.1. Sean $f \in \mathbb{R}[x, y]$, $\deg(f) \geq 1$, y $p = (a, b) \in \mathbb{R}^2$ un punto doble de f . Definimos las siguientes cantidades:

$$\begin{aligned} A &:= f''_{x^2}(a, b), & B &:= f''_{xy}(a, b), & C &:= f''_{y^2}(a, b), \\ \Delta &:= B^2 - AC. \end{aligned} \tag{A.5.6}$$

- I. Si $\Delta > 0$, entonces (a, b) es un *crunodo*, con dos rectas tangentes reales.
- II. Si $\Delta < 0$, entonces (a, b) es un *punto acnodo*.
- III. Si $\Delta = 0$, entonces (a, b) es un punto con una única recta tangente (real) T : puede ser una *cúspide* o un *tacnodo*.
 - (a) Si $I((a, b), f, T) = 3$, la curva f tiene una sola rama centrada en (a, b) . Por tanto este punto es una *cúspide (ordinaria)*.

DEMOSTRACIÓN.

Como (a, b) es un punto doble de f , al menos uno de los números A, B, C es distinto de cero. Supongamos que $C \neq 0$. Por la fórmula (A.2.1),

$$g(x, y) := f(a + x, b + y) = \frac{1}{2!} f''_{x^2}(a, b)x^2 + f''_{xy}(a, b)xy + \frac{1}{2!} f''_{y^2}(a, b)y^2 + \text{TGS} = G_2(x, y) + \text{TGS},$$

donde TGS denota términos de grado superior.

Como las rectas tangentes a g y $2g$ en $(0, 0)$ son las mismas, no se pierde generalidad si suponemos que

$$G_2(x, y) = f''_{x^2}(a, b)x^2 + 2f''_{xy}(a, b)xy + f''_{y^2}(a, b)y^2 = Ax^2 + 2Bxy + Cy^2.$$

Podemos descomponer el trinomio de la siguiente forma

$$G_2(x, y) = x^2 \left(A + 2B\frac{y}{x} + C\frac{y^2}{x^2} \right) \quad (\text{A.5.7})$$

Si denotamos $\frac{y}{x}$ por z , tenemos que la expresión entre paréntesis en (A.5.6) se puede escribir como $A + 2Bz + Cz^2$. Calculamos sus raíces

$$z = \frac{-2B \pm \sqrt{4B^2 - 4AC}}{2C} = \frac{-B \pm \sqrt{\Delta}}{C}. \quad (\text{A.5.8})$$

Con lo cual,

1. Si $\Delta > 0$, hay dos raíces reales r_1, r_2 . Se tiene que

$$G_2(x, y) = Cx^2 \left(\frac{y}{x} - r_1 \right) \left(\frac{y}{x} - r_2 \right) = C(y - r_1x)(y - r_2x);$$

por lo tanto, (a, b) es un nodo con dos tangentes reales, a saber,

$$y - b = r_1(x - a) \quad y - b = r_2(x - a).$$

2. Si $\Delta = 0$, hay una única raíz real $r = -\frac{B}{C}$, que es doble. La tangente única es la recta real

$$G_2(x, y) = Cx^2 \left(\frac{y}{x} + \frac{B}{C} \right)^2 = C \left(y + \frac{B}{C}x \right)^2;$$

por lo tanto, (a, b) es un punto cuspidal con una única tangente real, a saber,

$$y - b = -\frac{B}{C}(x - a). \quad (\text{A.5.9})$$

Esta recta, T , es ortogonal al vector (B, C) ; con lo cual, está dirigida a lo largo de la dirección del vector $(-C, B)$. Por ello, sus ecuaciones paramétricas son

$$T : \begin{cases} x = a - Ct, \\ y = b + Bt. \end{cases} \quad (\text{A.5.10})$$

Vamos a analizar el número de intersección en $p = (a, b)$ de f con la recta T , $I(p, f, T)$. Para ello, damos la expresión del polinomio de intersección, así como sus dos primeras derivadas, y evaluamos en $t = 0$.

$$\begin{aligned}
\phi(t) &= f(a - Ct, b + Bt), \\
\phi'(t) &= -Cf'_x(a - Ct, b + Bt) + Bf'_y(a - Ct, b + Bt), \\
\phi''(t) &= C^2f''_{x^2}(a - Ct, b + Bt) - 2BCf''_{xy}(a - Ct, b + Bt) \\
&\quad + B^2f''_{y^2}(a - Ct, b + Bt).
\end{aligned}$$

$$\begin{aligned}
\phi(0) &= 0, \\
\phi'(0) &= 0 \\
\phi''(0) &= C^2f''_{x^2}(a, b) - 2BCf''_{xy}(a, b) + B^2f''_{y^2}(a, b) \\
&= C^2A - 2B^2C + B^2C = C(AC - B^2) = 0.
\end{aligned}$$

Por tanto,

$$I(p, f, T) \geq 3.$$

Como la multiplicidad $m_{(a,b)}(f) = 2$, la recta T es tangente.

3. Si ocurre que $I(p, f, T) = 3$, entonces el punto p es una cúspide y es ordinaria. En efecto, por un lado sabemos que el número de ramas $r(p)$ en un punto singular p es menor o igual que su multiplicidad $m_p(f)$. Como aquí $m_p(f) = 2$, el número de ramas, $r(p)$, de la curva f en p es 1 ó 2. Veamos que no puede ser 2. Si hubiera dos ramas R_1 y R_2 de f que pasaran por p se tendría que p sería un punto liso o simple de cada una de las dos ramas. Por tanto, la recta tangente T en p a cada rama tendría un número de intersección, $I(p, R_i, T) \geq 2, i = 1, 2$. Como (véase la página 59, línea 4 de Kending [46])

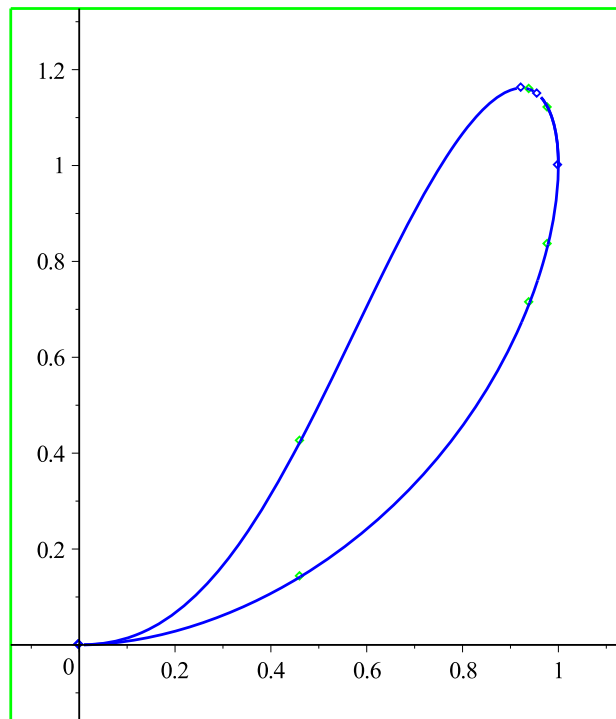
$$I(p, f, T) = I(p, R_1, T) + I(p, R_2, T)$$

se tendría que $I(p, f, T) \geq 4$. Lo que contradice la suposición. Así pues, la curva f tiene una sola rama que pasa por p y como $\Delta(p) = 0$, el punto p es una cúspide o punto de retroceso. Dado que $I(p, f, T) = m_p(f) + 1$ la recta T es una tangente ordinaria; o sea, la cúspide en p es *ordinaria*.

4. Si $\Delta < 0$, no hay raíces reales en (A.5.8). Por lo tanto, p es un punto aislado, ya que no habrá tangentes reales.

□

Observación A.5.2. Cuando $p := (a, b)$, $\Delta(p) = 0$ e $I(p, f, T) \geq 4$ no se puede asegurar que el punto p sea un tacnodo para la curva f . Puede serlo o no. Por ejemplo, el punto $(0, 0)$ es un punto doble de la curva $f(x, y) = x^4 + x^2y^2 - 2x^2y - xy^2 + y^2$ pues $f''_{y^2}(0, 0) = 2$ y el polinomio $f(x, 0) = x^4$ que tiene a $x = 0$ como raíz cuádruple. De hecho, $(0, 0)$ es un punto cuspidal de segunda especie (o cúspide ranfoidea) de la curva f . Con ayuda de Maple 17 y de la instrucción `plot_real_curve` hemos dibujado la curva f que aparece en la Figura A.6. Es decir que este punto no es un tacnodo, a pesar de que $I((0, 0), f, y = 0)$ es igual a 4.

Figura A.6: Cúspide ranfoidea en $(0,0)$

Daremos una explicación intuitiva sobre los distintos puntos dobles de una curva algebraica real $f(x, y) = 0$. Es posible resolver esta ecuación en y en función de x mediante series de potencias de exponentes números racionales, llamadas series de Puiseux. Hay un algoritmo, llamado de Newton-Puiseux, para hacer esto, y el programa `Maple` calcula tales desarrollos en series de Puiseux: basta usar la instrucción `puiseux` del paquete `algebraiccurves`. Si (a, b) es el punto doble que se está considerando la instrucción `puiseux(f, x=a, y, n)` nos da la suma parcial de términos de exponentes racionales $\leq n$ de cada una de las series de Puiseux que parametrizan la curva $V_{\mathbb{C}}(f)$ en un entorno del punto $x = a$. Dicho de otra forma, tendremos tantas *ramas locales* en $x = a$ como sumas parciales produzca la instrucción. Si utilizamos la instrucción con un quinto parámetro, de nombre `T`; a saber, `puiseux(f, x=a, y, 0, T)` el programa devuelve dos series formales de Laurent en T : $(\varphi_1(T), \varphi_2(T))$ con $\varphi_1(T) = \alpha T^k + \beta$ y $\varphi_2(T) \in \mathbb{C}((T))$, $\alpha, \beta \in \mathbb{C}$ por cada *rama*. De hecho, habiendo puesto $n = 0$, `puiseux` produce sumas parciales (de las series de Laurent) lo suficientemente largas para *separar* las “ramas”. En otras palabras, las parejas $(\varphi_1(T), \varphi_2(T))$ de series de Laurent en T son únicas. Si $f(X, Y) = a_0(X) + a_1(X)Y + a_2(X)Y^2 + \dots + a_n(X)Y^n$, mirado como polinomio en Y con coeficientes en $\mathbb{R}[X]$, es tal que $a_n(X) \neq 0$, entonces también $\varphi_2(T) \in \mathbb{C}[[T]]$. Si todos los coeficientes de las series $\varphi_1(T)$ y $\varphi_2(T)$ son números reales decimos que tenemos *ramas reales*.

La diferencia entre cúspides de primera especie y de segunda especie se podría establecer mediante la presencia o no de puntos de la curva tan cercanos como se quisiera a p y situados exactamente en uno o dos cuadrantes de los que determinan las rectas T y r perpendicular a T por p .

A.6. Maple

Resumiendo, el programa de cálculo simbólico `Maple 17` implementa el *algoritmo de Newton* y con las instrucciones `singularities` y `puiseux` de su paquete `algebraiccurves` nos produce con

(1) `>singularities(f, x, y);`

los puntos singulares p de la curva f en el plano proyectivo complejo $\mathbb{P}^2(\mathbb{C})$ y sus invariantes $(m_p, \delta_p, r(p))$, donde m_p es la multiplicidad del punto, δ_p es el llamado invariante delta del punto, y $r(p)$ es el **número de ramas distintas** que tienen por centro al punto. Los puntos vienen dados por sus coordenadas homogéneas $(x : y : z)$, siendo $z = 1$ ó 0 , según que se trate de un punto finito o un punto infinito. Las salidas producidas son de las forma

$$[[x, y, z], m, \delta, r];$$

y con

(2) `>puiseux(f,x=a,y,0,T);`

obtenemos una parametrización irreducible para cada una de las distintas ramas cuyo centro tiene coordenadas (a, β) para algún $\beta \in \mathbb{C}$ que satisfacen $f(a, \beta) = 0$. En particular si $\beta = b$ (a, b) es un punto singular de la curva $V_{\mathbb{C}}(f)$, este procedimiento nos permite confirmar el **número de ramas distintas** que tienen por centro a (a, b) ; el cual ya sería conocido desde el número de ramas asignadas al punto $(a : b : 1)$ por **singularities**.

RAMAS REALES. Cuando todos los coeficientes de las ramas que nos da **puiseux** son números reales, diremos que las ramas son **ramas reales**. Véase al respecto lo que se comenta en la página 287 del libro de M.J. de la Puente [25].

Ejemplo A.6.1. *Sea el polinomio*

$$\begin{aligned} f(x, y) := & -y^6 + 2y^5 + \left(-3x^2 - 3\sqrt{2} + 2x + \frac{5}{2}\right)y^4 + \left(4x^2 + 4\sqrt{2} - 4x - 4\right)y^3 \\ & + \left(\sqrt{2} - 4x + 4\sqrt{2}x - 3x^4 - 6\sqrt{2}x^2 + 5x^2 + 4x^3 - \frac{7}{4}\right)y^2 \\ & + \left(4\sqrt{2}x^2 + \frac{11}{2} - 4x^2 - 4\sqrt{2} + 2x^4 - 4x^3 - 4\sqrt{2}x + 6x\right)y \\ & + \frac{11}{2}x - 4x^3 + \frac{33}{4}\sqrt{2} + \sqrt{2}x^2 - 4x\sqrt{2} - \frac{93}{8} - 3\sqrt{2}x^4 \\ & + 4\sqrt{2}x^3 + 2x^5 + \frac{5}{2}x^4 - \frac{7}{4}x^2 - x^6 \end{aligned}$$

Sabemos que $(a, b) := ((\sqrt{2} - 1)/2, (\sqrt{2} - 1)/2)$ es un punto singular de f ; de hecho, es un punto doble: esto es, $f(a, b) = f'_x(a, b) = f'_y(a, b) = 0$ y alguna de las derivadas parciales segundas $f''_{x^2}, f''_{xy}, f''_{y^2}$ no se anula en (a, b) . También habíamos calculado previamente $\Delta(a, b) := f''_{xy}(a, b)^2 - f''_{x^2}(a, b)f''_{y^2}(a, b) = 0$. Luego sabemos que la curva $V_{\mathbb{R}}(f)$ tiene una única tangente real en (a, b) . Veamos qué devuelve la instrucción

`>singularities(f,x,y);`

$$\{[(\sqrt{2} - 1)/2, (\sqrt{2} - 1)/2, 1], 2, 2, 2\},$$

$$\begin{aligned} & [[\text{RootOf}(4_Z^2 + (4\sqrt{2} - 8)_Z - 6\sqrt{2} + 9), 2 - \sqrt{2} - \\ & \quad \text{RootOf}(4_Z^2 + (4\sqrt{2} - 8)_Z - 6\sqrt{2} + 9), 1], 2, 1, 2], \\ & \quad [[\text{RootOf}(-Z^2 + 1), 1, 0], 3, 3, 3] \} \end{aligned}$$

Como las raíces de $4Z^2 + (4\sqrt{2} - 8)Z - 6\sqrt{2} + 9 = 0$ son

$$-(1/2)\sqrt{2} + (i/2)\sqrt{2} + 1 - i/2, -(1/2)\sqrt{2} - (i/2)\sqrt{2} + 1 + i/2,$$

vemos que nos ha devuelto $1 + 4 + 2$ puntos singulares en $\mathbb{P}^2(\mathbb{C})$: un punto real $((\sqrt{2} - 1)/2, (\sqrt{2} - 1)/2) \in \mathbb{R}^2$ que es doble ($m = 2$) y tiene dos ramas

($r = 2$); estos datos sugieren que este punto es un tacnodo. Además también se han obtenido cuatro puntos singulares en \mathbb{C}^2 y dos puntos infinitos singulares en $\mathbb{P}^2(\mathbb{C})$.

Procedamos ahora a ejecutar la instrucción

```
>puiseux(f,x=(sqrt(2)-1)/2,y,0,T);
```

Nos devuelve cuatro ramas centradas en cuatro puntos $((\sqrt{2}-1)/2, \beta)$ de la curva. Cada rama está representada por una parametrización formal:

$$\begin{aligned} & [x = T - 1/2 + \sqrt{2}/2, y = 1/2 - \sqrt{2}/2], \\ & [x = T - 1/2 + \sqrt{2}/2, y = -1/2 + \sqrt{2}/2 - T + T^2(-2 - 2\sqrt{2})], \\ & [x = T - 1/2 + \sqrt{2}/2, y = -1/2 + \sqrt{2}/2 - T + T^2(-1 - \sqrt{2})], \\ & [x = T - 1/2 + \sqrt{2}/2, y = \text{RootOf}(8Z^3 + (4\sqrt{2} - 20)Z^2 + (-12\sqrt{2} + 26)Z \\ & \quad + 27\sqrt{2} - 45)] \end{aligned}$$

Tras un minucioso examen de los resultados vemos que estas cuatro ramas son reales. Están centradas en los valores obtenidos al hacer $T = 0$ en las correspondientes sumas parciales de las series formales en T . Esos cuatro centros están en la recta vertical $x = -1/2 + \sqrt{2}/2$. Dicho de otra forma, estas sumas parciales dan parametrizaciones aproximadas de estas cuatro ramas. Las ramas que corresponden a nuestro punto de estudio son la 2ª y la 3ª. Confirmado, pues, que este punto doble es un tacnodo. Véase la Figura A.7 teniendo en cuenta que $(-1 + \sqrt{2})/2 = 0,2071067810$.

A.7. Intersección de curvas

Aunque en los casos más interesantes las ramas serán “ramas de una curva algebraica plana f ”, puede interesarnos partir de un concepto más general de ramas. Las llamaremos ramas abstractas y en principio no tienen por que ser ramas de una determinada curva f . En el conjunto $\mathbb{C}[[T]]^2 \setminus \mathbb{C}^2$ definimos una relación de equivalencia entre sus elementos (\tilde{x}, \tilde{y}) y (\bar{x}, \bar{y}) mediante la existencia de una serie $\varphi \in \mathbb{C}[[T]]$ de orden uno y tal que $\tilde{x} = \bar{x} \circ \varphi$, $\tilde{y} = \bar{y} \circ \varphi$. Llamaremos **rama abstracta** a toda clase de equivalencia $R = [(\tilde{x}, \tilde{y})]$ en el conjunto $\mathbb{C}[[T]]^2 \setminus \mathbb{C}^2$ donde la parametrización (\tilde{x}, \tilde{y}) es irreducible. Se llamará **centro de la rama abstracta** al elemento $(\tilde{x}(0), \tilde{y}(0))$, el cual no depende del par representante elegido. Por lo dicho en los párrafos anteriores de este Apéndice toda *rama de una curva algebraica* $f \in \mathbb{C}[x, y]$ (f libre de cuadrados) es una rama abstracta $[(\tilde{x}, \tilde{y})]$; sólo que en este caso, $f(\tilde{x}(T), \tilde{y}(T)) = 0$ en $\mathbb{C}[[T]]$. Lo que no sabemos es si para todo para $(\tilde{x}, \tilde{y}) \in \mathbb{C}[[T]]^2 \setminus \mathbb{C}^2$ irreducible existe o no un polinomio $f \in \mathbb{C}[x, y]$ tal que $f(\tilde{x}(T), \tilde{y}(T)) = 0$ en $\mathbb{C}[[T]]$; es decir, no sabemos si toda rama abstracta es *rama de una curva algebraica*. A veces, en inglés a las ramas de una curva algebraica se las denomina **places**.

Esta Sección sigue estrechamente los apuntes de Arrondo [9] sobre curvas algebraicas, así como el Capítulo IV, §5, páginas 108–115 de Walker [73].

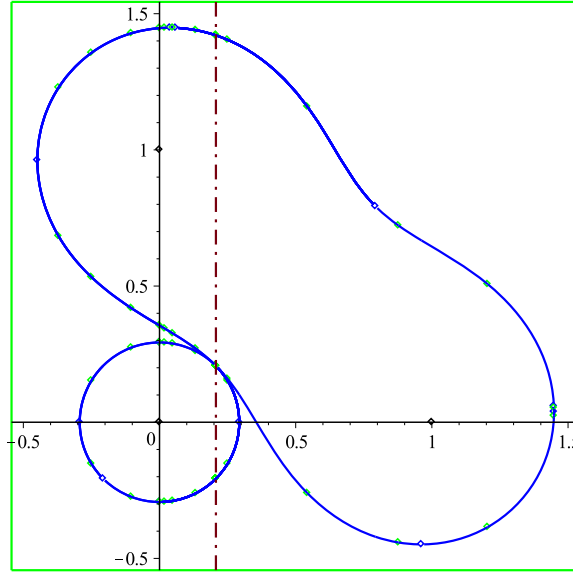


Figura A.7: Cuatro ramas reales centradas en la vertical $x = (-1 + \sqrt{2})/2$.

Definición A.7.1. Se llama **parametrización formal en el punto** (a_0, b_0) de una curva afín $C := V_{\mathbb{C}}(f)$ donde el polinomio f es minimal para C a un par $(\tilde{x}(T), \tilde{y}(T))$ de series formales de $\mathbb{C}[[T]]$ no ambas constantes tales que $f(\tilde{x}(T), \tilde{y}(T)) = 0$ en $\mathbb{C}[[T]]$ y $\tilde{x}(0) = a_0, \tilde{y}(0) = b_0$.

La **parametrización** (\tilde{x}, \tilde{y}) es **irreducible** si no existe $k \in \mathbb{N}, k > 1$ tal que $\tilde{x}, \tilde{y} \in \mathbb{C}[[T^k]]$.

Teorema A.7.2 (Arrondo, Teorema 6.10). *Toda parametrización formal (\tilde{x}, \tilde{y}) en (a_0, b_0) de una curva afín es equivalente a una parametrización de la forma $(a_0 + T^r, q(T))$, con $q(T) \in \mathbb{C}[[T]]$ y $q(0) = b_0$. Además, necesariamente $r = \text{ord}(\tilde{x} - a_0)$ y cualquier otra parametrización de esa forma se escribe como $(a_0 + T^r, q(\omega T))$, donde ω es una raíz r -ésima de la unidad.*

A.7.1. Multiplicidad de intersección de una rama con una curva

Definición A.7.3. Dadas dos curvas $V_{\mathbb{C}}(f)$ y $V_{\mathbb{C}}(g)$ que pasen por un punto $p = (a_0, b_0)$ donde $V_{\mathbb{C}}(g)$ tiene una rama S dada por la parametrización $(\tilde{x}(T), \tilde{y}(T))$ se llama **multiplicidad de intersección de la rama S y la curva $V_{\mathbb{C}}(f)$ en el punto p** al orden de la serie formal $f(\tilde{x}(T), \tilde{y}(T))$:

$$I(p, S, f) := \text{ord}(f(\tilde{x}(T), \tilde{y}(T))).$$

La definición del número de multiplicidad de intersección de dos curvas algebraicas en un punto está basada en el teorema siguiente.

Teorema A.7.4 (Walker, Teorema 5.1, p. 109). *Si $f(x, y)$ y $g(x, y)$ son los polinomios minimales de dos curvas algebraicas, respectivamente, que tienen un punto común $p = (a_0, b_0)$, entonces la suma de las multiplicidades de intersección de las ramas $S_1 = [(a_0 + T^{s_1}, q_1(T))], \dots, S_u = [(a_0 + T^{s_u}, q_u(T))]$ de $V_{\mathbb{C}}(g)$ centradas en p con la curva $V_{\mathbb{C}}(f)$ es igual a la suma de las multiplicidades de intersección de las ramas $R_1 = [(a_0 + T^{r_1}, p_1(T))], \dots, R_v = [(a_0 + T^{r_v}, p_v(T))]$ de $V_{\mathbb{C}}(f)$ centradas en p con la curva $V_{\mathbb{C}}(g)$. Es decir,*

$$I(p, f, S_1) + \dots + I(p, f, S_u) = I(p, g, R_1) + \dots + I(p, g, R_v).$$

O bien, con más detalle,

$$\sum_{i=1}^u \text{ord}(f(a_0 + T^{s_i}, q_i(T))) = \sum_{j=1}^v \text{ord}(g(a_0 + T^{r_j}, p_j(T))).$$

Esta suma común se llama **número de intersección en p de las curvas $C := V_{\mathbb{C}}(f)$ y $D := V_{\mathbb{C}}(g)$** y se denota por

$$I(p, f, g) = I(p, C, D).$$

También se lo conoce como la **multiplicidad de la intersección de C y D en el punto p** y se denota por $\text{mult}_p(C, D)$. Cuando una de las curvas f ó g es una recta ℓ , el concepto vuelve a coincidir con el número de intersección, $I(p, f, \ell)$, de f con ℓ en el punto p . Nótese que el número $I(p, f, g)$ puede ser infinito.

A.7.2. Tangente a una rama y multiplicidad de una rama

Si p es un punto de una curva algebraica f , en la Subsección A.2.2 hemos definido los conceptos de *multiplicidad* de curva f en el punto p , y de *recta tangente* a la curva f en el punto p . Vamos a extender estos conceptos cuando sustituycamos el *punto* p de la curva f por una *rama* R de f . Primero demostraremos el lema siguiente.

Lema A.7.5 (Teorema 5.7, pág. 113, Walker [73]). *Si R es una rama con centro $p = (a_0, b_0)$, de todas las rectas $L = 0$ que pasan por p hay una y sólo una, $L_0 = 0$, tal que*

$$I(p, R, L_0) > \min_{p \in L} I(p, R, L).$$

DEMOSTRACIÓN. Sea

$$\tilde{x} = \sum_{i=0}^{\infty} a_i T^i, \quad \tilde{y} = \sum_{i=0}^{\infty} b_i T^i$$

una parametrización de R . Ya que $L = 0$ pasa por p , tenemos $L = a(x - a_0) + b(y - b_0)$, y $L(\tilde{x}, \tilde{y}) = \sum_{i=1}^{\infty} (aa_i + bb_i)T^i$. Ahora si r es el menor entero positivo

para el que al menos uno de a_r, b_r es no nulo, entonces $\text{ord}(L(\tilde{x}, \tilde{y})) = r$ si y sólo si $aa_r + bb_r \neq 0$. Por otra parte, si $aa_r + bb_r = 0$ entonces $\text{ord}(L(\tilde{x}, \tilde{y})) > r$. De aquí que la condición $aa_r + bb_r = 0$ determina una única L_0 cuya multiplicidad de intersección $I(p, R, L_0)$ es mayor que el mínimo. \square

Definición A.7.6. Sea R una rama abstracta con centro p parametrizada por $(\tilde{x}(T), \tilde{y}(T))$. Sea L una recta cualquiera que pasa por p . Llamaremos **multiplicidad de la rama R** al mínimo

$$r := \min_{p \in L} I(p, R, L) = \min_{p \in L} \text{ord}(L(\tilde{x}(T), \tilde{y}(T)))$$

cuando L recorre todas las rectas que pasan por p .

Una rama se dice **lineal** si su multiplicidad es uno. La única recta L_0 que satisface $I(p, R, L_0) > r$ se llama la **tangente** a R .

Ejemplo A.7.1. La rama $R_1 := [(T^2, T^3 + \dots)]$ es de multiplicidad $r = 2$; su tangente es la recta $L_0 : y = 0$; $I((0, 0), R_1, y = 0) = 3$. Estos valores son característicos de una cúspide ordinaria. La rama $R_2 := [(T^2, T^4 + T^5 + \dots)]$ es de multiplicidad $r = 2$; tangente $y = 0$ e $I((0, 0), R_2, y = 0) = 4$. Estos valores los típicos de una cúspide ranfoidea.

En el libro de Shikin [66], página 52, puede leerse que una rama

$$[(aT^n, a_1T^{n_1} + a_2T^{n_2} + \dots)] \text{ con } 1 \leq n \leq n_1 < n_2 < \dots$$

tiene una cúspide ranfoidea en $(0, 0)$ si al menos uno de los números n_i no es divisible por n y si n y n_1 son pares.

La relación entre estos conceptos sobre ramas y los relativos a las multiplicidades y tangentes en un punto son mostrados en el teorema que sigue.

Teorema A.7.7 (Teorema 5.8, pág. 113, Walker [73]). (I) Si p es un punto de la curva f , la suma de las multiplicidades de las ramas de f con centros en p es $m_p(f)$.

(II) Un punto p de f es no singular si y sólo si es el centro de una única rama (lineal) de f .

(III) Las tangentes de f en el punto p coinciden con las tangentes a las ramas de f con centro p .

Hemos definido $I(p, f, R)$ cuando f es una curva y R es una rama; también lo hemos definido para dos curvas f y g : $I(p, f, g)$. Pero, no lo hemos definido para la intersección de dos ramas R y S con centro común p ; i.e. $I(p, R, S)$ está sin definir aún. Esto es advertido claramente en el libro de Abhyankar [1]: véanse la línea -10, página 125 y la línea última de la página 126, donde el asunto es resuelto mediante el orden de una serie formal de potencias en X de la resultante de Sylvester de dos polinomios mónicos de $\mathbb{C}[[X]][Y]$ mediante el Teorema de Preparación de Weierstrass. Nosotros vamos a dar la definición de $I(p, R, S)$ siguiendo los libros clásicos de Walker [73]: Ejercicio 2, p. 115; y de Lefschetz [50]: Capítulo V, § 5, p. 116.

Definición A.7.8. La multiplicidad de intersección, $I(p, R, S)$, de dos ramas R y S de parametrizaciones $(a + T^r, b + a_1T + a_2T^2 + \dots)$ y $(a + T^s, b + b_1T + b_2T^2 + \dots)$ con el mismo centro $p = (a, b)$ se define como el orden ν de

$$\varphi(R, S) := \prod_{(j,k) \in \underline{r} \times \underline{s}} (\bar{y}_j - \bar{z}_k),$$

donde

$$\begin{aligned} \bar{y}_j &= \eta^j \sum_{i=1}^{\infty} a_i X^{i/r}, & \eta &:= e^{2\pi i/r}, & j \in \underline{r}, \\ \bar{z}_k &= \zeta^k \sum_{i=1}^{\infty} b_i X^{i/s}, & \zeta &:= e^{2\pi i/s}, & k \in \underline{s}, \end{aligned}$$

con $\underline{n} := \{1, 2, \dots, n\}$ para cualquier entero positivo n .

Ya que $\varphi(R, S)$ es un polinomio simétrico tanto en las $\bar{y}_1, \dots, \bar{y}_r$ como en las $\bar{z}_1, \dots, \bar{z}_s$ sólo contiene potencias enteras positivas de X (i.e. los términos de exponente fraccionario $X^{m/n}$ no aparecen). De aquí que $\varphi(R, S) = X^\nu c(X)$ donde $c(X)$ es una serie de $\mathbb{C}[[X]]$ tal que $c(0) \neq 0$ y ν es un entero positivo: el orden de $\varphi(R, S) \in \mathbb{C}[[X]]$. Así pues, ν es un entero positivo bien definido que depende de R y S , e

$$I(p, R, S) := \nu.$$

Proposición A.7.9 (Ejercicio 3, pág. 115, Walker [73]). *Un tacnodo p de una curva f es el centro de dos ramas lineales R_1, R_2 tales que $I(p, R_1, R_2) = 2$.*

En literatura antigua [36] se llama *oscnodo* al punto $(0, 0)$ de la curva $y^2 - x^6 = 0$. La gráfica de esta curva puede verse en la Figura A.8. Como $y^2 - x^6 = (y - x^3)(y + x^3)$, se tiene que $I((0, 0), y - x^3, y + x^3) = 3$. Para nosotros este punto es un *tacnodo de orden superior*. Por eso, ampliaremos la definición de Walker de tacnodo diciendo que es un punto doble p centro de dos ramas lineales R_1, R_2 tales que $I(p, R_1, R_2) \geq 2$.

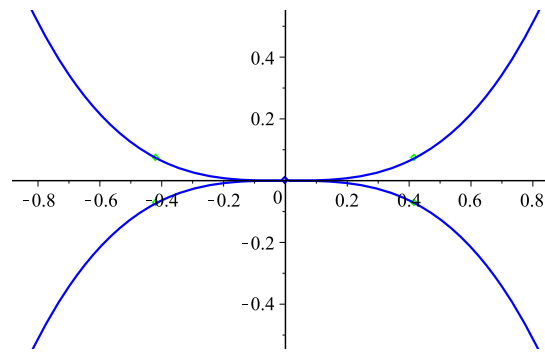
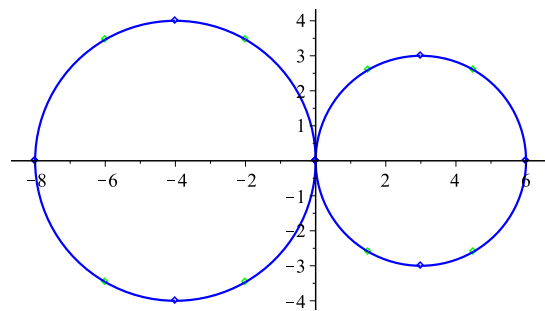
A.7.2.1. Tacnodo tangente a dos circunferencias

Comprobaremos que siempre que estemos en presencia de dos circunferencias tangentes en un punto, dicho punto es un tacnodo de la curva algebraica definida por la unión de las circunferencias. Sean $f := (x + 4)^2 + y^2 - 16$ y $g := (x - 3)^2 + y^2 - 9$; entonces la curva $V_{\mathbb{R}}(fg)$ viene representada en la Figura A.9

`singularities(f*g,x,y);)`

$$\{[[0, 0, 1], 2, 2, 2], [[RootOf (-Z^2 + 1), 1, 0], 2, 1, 2]\}$$

luego $(0, 0)$ es un punto afín donde fg tiene multiplicidad $m = 2$, $\delta = 2$ y número de ramas $r = 2$. Si sustituimos $x = 0$ en $(fg)(0, y) = y^4$; por tanto $y = 0$ es raíz cuádruple de y^4 —de donde— $I((0, 0), fg, x = 0) = 4$; por consiguiente, $x = 0$ es una tangente a fg en $(x, y) = (0, 0)$.

Figura A.8: Oscnodo de $y^2 - x^6 = 0$.Figura A.9: Tacnodo de $fg = 0$ en el punto $(0,0)$.

```
puiseux(f*g,x=0,y,0,T);)
```

$$R_1 := [x = -8T^2, y = -8T], R_2 := [x = 6T^2, y = 6T]$$

Teniendo en cuenta la Definición A.7.8 y llamando $\eta := \exp(2\pi i/2) = -1$, definimos

$$\bar{y}_1 := -8\eta X^{1/2}, \bar{z}_1 := 6\eta X^{1/2}, \bar{y}_2 := -8\eta^2 X^{1/2}, \bar{z}_2 := 6\eta^2 X^{1/2}.$$

Por tanto,

$$(\bar{y}_1 - \bar{z}_1)(\bar{y}_2 - \bar{z}_1)(\bar{y}_1 - \bar{z}_2)(\bar{y}_2 - \bar{z}_2) = 784X^2.$$

En consecuencia, $I((0,0), R_1, R_2) = 2$. Es decir, la curva fg tiene un tacnodo en el punto doble $(0,0)$. Y es obvio que un razonamiento así es trasladable a cualquier par de circunferencias tangentes, exteriores o interiores. El cual será un caso frecuente al considerar fronteras de pseudoespectros de matrices normales.

Consideremos otro ejemplo:

$$h(x, y) := \left((x - 5/2)^2 + (y - 5)^2 - \frac{45}{4} \right) \left((x - 2)^2 + (y - 4)^2 - 5 \right)$$

El punto $p = (1, 2)$ es doble para la curva h y ésta tiene una tangente única L en él. Esta recta viene dada por la ecuación $(x - 1) + 2(y - 2) = 0$. Además, Maple 17 nos permite obtener

```
>singularities(h,x,y);
```

$$[[1, 2, 1], [2, 2, 2], [[RootOf(Z^2 + 1), 1, 0], [2, 1, 2]]$$

```
>puiseux(h,x=1,y,0,T);
```

$$[x = T + 1, y = 6], [x = T + 1, y = 8], [x = T + 1, y = 2 - \frac{1}{2}T + \frac{5}{16}T^2],$$

$$[x = T + 1, y = 2 - \frac{1}{2}T + \frac{5}{24}T^2]$$

lo que nos proporciona las ramas $R_1 = [(T + 1, 2 - \frac{1}{2}T + \frac{5}{16}T^2)]$ y $R_2 = [(T + 1, 2 - \frac{1}{2}T + \frac{5}{24}T^2)]$ que pasan por $p = (1, 2)$; además,

$$\varphi(R_1, R_2) = \left(\left(2 - \frac{1}{2}X + \frac{5}{16}X^2 \right) - \left(2 - \frac{1}{2}X + \frac{5}{24}X^2 \right) \right) = \frac{5}{48}X^2;$$

por tanto, $I(p, R_1, R_2) = 2$. Lo que confirma que p es un tacnodo de h . Véase la Parte Derecha de la Figura A.5.

Apéndice B

Conexión

Como una parte sustancial de esta Tesis trata sobre las *componentes conexas* de un pseudoespectro añadimos este Apéndice sobre este concepto. Adelantamos que las componentes conexas de un pseudoespectro son siempre *conexas por arcos*, cómo así lo hicimos notar en la página 18.

B.1. Espacios topológicos

Aunque el principal espacio topológico que hemos usado es el plano \mathbb{C} , identificado con \mathbb{R}^2 , con la topología ordinaria, daremos unas definiciones generales de *espacio conexo* y *conjunto conexo*. Un **espacio topológico** es un par (E, \mathcal{T}) formado por un conjunto E y una colección \mathcal{T} de subconjuntos de E llamadas conjuntos **abiertos** (o abreviadamente “abiertos”) que satisfacen las tres propiedades siguientes:

- Toda unión (finita o infinita) de abiertos es abierta.
- Toda intersección finita de abiertos es abierta.
- El conjunto E y el conjunto vacío \emptyset son abiertos.

Se dice que el conjunto \mathcal{T} define una **topología** sobre E . Cuando esté claro por el contexto cuál es la topología, \mathcal{T} , que estamos utilizando hablaremos simplemente del espacio topológico E .

Definición B.1.1. Sea $x_0 \in E$. Se llama **entorno de x_0** a cualquier subconjunto V de E que contenga un abierto A al que pertenezca x_0 . Esto es, $V \subset E$ es un entorno de x_0 si existe un abierto A de E tal que

$$x_0 \in A \subset V.$$

De esta definición se deduce que un conjunto abierto es entorno de todos sus puntos. De hecho, esta propiedad caracteriza los conjuntos abiertos. Pero hay entornos de un punto que no son conjuntos abiertos.

Sea S un subconjunto de un espacio topológico E . Se llama **conjunto interior** de S a la unión de todos los abiertos del espacio E contenidos en S . Se denotará por \mathring{S} y es el mayor abierto de E contenido en S . Un punto $x_0 \in S$ se llama **punto interior a S** si existe un entorno de x_0 contenido en S . Así pues, el conjunto \mathring{S} es el conjunto de todos los puntos interiores de S . Se dice que $y_0 \in E$ es un **punto frontera** de S si en todo entorno de y_0 existen algún punto de S y algún punto que no pertenece a S . Al conjunto de todos los puntos frontera de S se le llamaremos el **conjunto frontera, o frontera** de S y lo denotaremos por $\partial(S)$ o ∂S (cuando la omisión del paréntesis no dé lugar a confusión). Se dice que el punto $x_0 \in E$ es un **punto adherente a S , o punto de adherencia de S** , si en todo entorno de x_0 hay puntos de S . Se dice que $x_0 \in E$ es un **punto de acumulación** de S si en todo entorno de x_0 existen puntos de S distintos de x_0 . Así pues, todo punto de acumulación de S es punto adherente a S ; pero el recíproco no es cierto. Al conjunto de puntos de adherencia de S se le llama la **adherencia, o clausura**, de S y se denota por \overline{S} . Un subconjunto B de E se llama **cerrado** si su complemento, $B^c := E \setminus B$, es un conjunto abierto. Es fácil ver que \overline{S} es un conjunto cerrado; de hecho, \overline{S} es el menor conjunto cerrado que contiene a S . Dicho de otro modo, \overline{S} es la intersección de todos los conjuntos cerrados que contienen a S . Es fácil ver que

$$\partial(S) = \overline{S} \cap \overline{S^c}.$$

Se dice que $x_0 \in E$ es un **punto exterior a S** si x_0 es un punto interior a S^c . Finalmente, señalemos que todo espacio topológico E admite la partición

$$E = \mathring{S} \dot{\cup} \partial(S) \dot{\cup} \widehat{S^c}$$

respecto de todo subconjunto S de E . Además, $\overline{S} = \mathring{S} \dot{\cup} \partial(S)$.

Para todo subconjunto S de un espacio topológico E se llama **subespacio S de E** al conjunto S dotado con la topología formada por las intersecciones de los abiertos de E con S . Se dice también que la **topología** de S es **inducida** por la de E o que es la **topología relativa**. Se debe prestar atención a que un abierto del subespacio S no es necesariamente un abierto de E . Así pues, $U \subset S$ es un abierto del subespacio S o un S -abierto si existe un abierto A de E tal que $U = S \cap A$.

Proposición B.1.2 (Transitividad de los subespacios). *Sean E un espacio topológico, S un subespacio de E , y T un subconjunto de S . Las topologías sobre T inducidas por las de E y S son idénticas.*

Véase la Proposición 9–3, página 26, del libro de Choquet [19].

Definición B.1.3. Sea A un subconjunto de un espacio topológico E . Se dice que A es **denso en E** si $\overline{A} = E$. Es decir, A es denso en E si para cualquier $x \in E$ en todo entorno V_x de x existen puntos de A .

B.2. Conexión

Definición B.2.1. Espacio topológico conexo.

Se dice que un espacio topológico E es **conexo** si no existe ninguna partición de E en dos subconjuntos abiertos no vacíos. Es decir, E es un *espacio conexo* si no existen abiertos disjuntos A, B tales que $E = A \cup B$, $A \neq \emptyset$, $B \neq \emptyset$. Esta propiedad es equivalente a las dos siguientes:

- No existen C, D subconjuntos cerrados en E , no vacíos, disjuntos tales que $E = C \cup D$.
- Los únicos conjuntos de E abiertos y cerrados a la vez son E y \emptyset .

Definición B.2.2. Subconjunto conexo. Se dice que un subconjunto S de un espacio topológico E es **conexo** si el subespacio topológico S de E es un espacio conexo.

Por lo tanto, el subconjunto S de E es conexo si y sólo si no existen abiertos A, B en E tales que $S \subset A \cup B$, $(S \cap A) \cap (S \cap B) = \emptyset$, $S \cap A \neq \emptyset$ y $S \cap B \neq \emptyset$.

La condición $S \subset A \cup B$ equivale a decir que $S = (S \cap A) \cup (S \cap B)$. En resumen, el **subconjunto** S de E es **conexo** si $S = U \cup V$ con U, V conjuntos S -abiertos implica que $U = \emptyset$ o $V = \emptyset$.

Teorema B.2.3 (Choquet [19], Teorema 13-4, página 43.). *Sea $(S_i)_{i \in I}$ una familia de conjuntos conexos de E . Si la intersección de esta familia no es vacía, su unión es conexa.*

DEMOSTRACIÓN. Sea

$$S = \bigcup_{i \in I} S_i.$$

Consideremos que existan dos abiertos A, B de E tales que

$$S = (S \cap A) \cup (S \cap B) \tag{B.2.1}$$

$$(S \cap A) \cap (S \cap B) = \emptyset \tag{B.2.2}$$

Por tanto, para cada $i \in I$, los conjuntos $S_i \cap A$ y $S_i \cap B$ son S_i -abiertos. Para todo $x \in S_i$ se tiene que $x \in S$. Por (B.2.1) $x \in S \cap A$ o $x \in S \cap B$; de donde $x \in S_i \cap A$ o $x \in S_i \cap B$: por lo que

$$S_i \subset (S_i \cap A) \cup (S_i \cap B); \tag{B.2.3}$$

como $S_i \cap A \subset S_i$ y $S_i \cap B \subset S_i$ se sigue que

$$(S_i \cap A) \cup (S_i \cap B) \subset S_i. \tag{B.2.4}$$

En consecuencia, (B.2.3) y (B.2.4) implican que $S_i = (S_i \cap A) \cup (S_i \cap B)$. Como S_i es conexo, uno de estos dos conjuntos es vacío. Por consiguiente, o bien $S_i \cap A = \emptyset$ o $S_i \cap B = \emptyset$.

Pero

$$\begin{array}{rcccl} S & = & (S \cap A) & \dot{\cup} & (S \cap B) \\ \cup & & \cup & \Downarrow & \cup \\ S_i & = & (S_i \cap A) & \dot{\cup} & (S_i \cap B) \end{array}$$

Si fuera $S_i \cap A = \emptyset$, se tendría que $S_i = S_i \cap B$; lo que implicaría que $S_i \subset S \cap B$ (pues $S_i \subset S$). Análogamente, si fuera cierto que $S_i \cap B = \emptyset$, se seguiría que $S_i \subset S \cap A$.

Por tanto todo S_i está contenido, ya sea en $S \cap A$, ya sea en $S \cap B$. Ahora bien los S_i tienen un punto en común x_0 al menos, que pertenece a S . Como $S = (S \cap A) \dot{\cup} (S \cap B)$ se tiene que $x_0 \in S \cap A$ o $x_0 \in S \cap B$. Sólo una de estas relaciones es cierta; por ejemplo $x_0 \in S \cap B$. Por lo tanto $S \cap B$ contiene a todos los S_i y $S \cap A$ es vacío. En consecuencia S es conexo. \square

Teorema B.2.4 (Choquet [19], Teorema 13–5, página 44). *Conexión de la clausura. La clausura de todo conjunto conexo es conexa.*

DEMOSTRACIÓN. Sea S un subconjunto conexo de E . Sean A, B abiertos en E tales que $\bar{S} \subset A \cup B$. De donde $S \subset A \cup B$. Como $S = (S \cap A) \cup (S \cap B)$, donde $S \cap A$ y $S \cap B$ son S -abiertos, y S es conexo, uno de estos dos últimos, por ejemplo $S \cap A$, es vacío. Como S es denso en \bar{S} , el conjunto A es vacío igualmente. Por lo que \bar{S} es conexo. \square

Una demostración análoga puede demostrar que todo conjunto T tal que $S \subset T \subset \bar{S}$ es conexo también.

Teorema B.2.5 (Choquet [19], Teorema 13–6, página 44). *Toda imagen continua de un espacio conexo es conexa.*

DEMOSTRACIÓN. Sea f una aplicación suprayectiva continua de un espacio conexo E sobre un espacio topológico F . Para todo subconjunto Y de F cerrado y abierto a la vez, $f^{-1}(Y)$ es cerrado y abierto en E ; por consiguiente es igual a E o a \emptyset ; ahora bien $Y = f(f^{-1}(Y))$, por ende $Y = F$ o $Y = \emptyset$; dicho de otro modo F es conexo. \square

B.2.1. Componentes conexas de un espacio

Podemos ahora estudiar la estructura de los espacios topológicos que no son conexos, precisando la noción vaga de “trozo” de un espacio tal.

Definición B.2.6. Para todo punto x de un espacio topológico E , se llama **componente conexa** de x a la unión $C(x)$ de todos los subconjuntos conexos de E que contienen a x .

El conjunto $C(x)$ es conexo por el Teorema B.2.3; por otra parte como $\overline{C(x)}$ es igualmente conexa y por construcción $C(x)$ es el mayor subconjunto conexo de E que contiene a x , se tiene que $C(x) = \overline{C(x)}$. Dicho de otro modo $C(x)$ es cerrada.

La relación binaria sobre E definida por $x_1 \sim x_2$ si x_1 y x_2 pertenecen al un mismo subconjunto conexo de E es una relación de equivalencia. Ahora bien

para todo x , $C(x)$ es la clase de equivalencia que contiene a x . Por consiguiente los conjuntos $C(x)$ pueden ser definidos también como las clases de equivalencia asociadas a la relación \sim .

Sea \mathcal{C} la colección de todos los subconjuntos conexos de E . Un elemento $C \in \mathcal{C}$ es *maximal* si $C \subset D$, $D \in \mathcal{C}$ implica que $C = D$. Entonces, las componentes conexas de E son exactamente los subconjuntos conexos maximales de E .

B.2.2. Componentes conexas de un subconjunto

Sea E un espacio topológico y T un subconjunto de E . Por definición, las componentes conexas de T son las componentes conexas del subespacio topológico T . Por lo tanto, una componente conexa de T es un subconjunto conexo maximal de T . Por la Proposición B.1.2 es obvio que un conjunto $U \subset T$ es conexo respecto del subespacio topológico T si y sólo si U es un conjunto conexo respecto del espacio topológico E .

B.3. Conexión por arcos

Se dice que un **espacio topológico E es conexo por arcos** si para todo par de puntos $a, b \in E$ existe una aplicación continua f de un intervalo $[\alpha, \beta]$ de \mathbb{R} en E tal que $f(\alpha) = a$ y $f(\beta) = b$. Un **subconjunto S de E se dice que es conexo por arcos** si para cada par de puntos $a, b \in S$ existe una aplicación continua f de un intervalo $[\alpha, \beta]$ de \mathbb{R} en E tal que $f(\alpha) = a$, $f(\beta) = b$ y para todo $t \in [\alpha, \beta]$ se tiene que $f(t) \in S$. Se dice que $f([\alpha, \beta])$ es un arco en S que conecta a con b .

La relación general entre los dos tipos de conexiones es

Teorema B.3.1 (Dugundji [30], Teorema 5.3, página 115). *Todo espacio topológico E (conjunto S) conexo por arcos es conexo. Pero un espacio (conjunto) conexo no necesita ser conexo por arcos.*

DEMOSTRACIÓN. Sea x_0 un punto fijo de E . Si E es conexo por arcos, para todo $x \in E$ existe un arco $A(x) := f([\alpha_0, \alpha]) \subset E$ que conecta x_0 con x . Como los arcos $A(x)$, $x \in E$ son conexos y tienen un punto en común: x_0 , su unión, que es E , es conexa por el Teorema B.2.3. \square

Para probar la parte segunda del Teorema se suele utilizar el ejemplo canónico siguiente (**La curva senoidal del topólogo**): Sea

$$S := \{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1, y = \sin(1/x)\}.$$

Como S es la imagen continua del intervalo $(0, 1]$, S es un conexo de \mathbb{R}^2 . Por tanto, por el Teorema B.2.4 su adherencia

$$\bar{S} = S \cup \{(0, y) \mid -1 \leq y \leq 1\}$$

es conexa. Sin embargo, \bar{S} no es conexa por arcos: no existe un arco que una el punto $(0, 0)$ con el punto $(1/\pi, 0)$. El detalle puede verse en Dugundji [30], Ej. 4, página 115.

B.3.1. Conexión en \mathbb{R}^n

En el espacio euclídeo ordinario \mathbb{R}^n se tienen los resultados siguientes. Recordemos que se llama bola abierta de centro $x \in \mathbb{R}^n$ y radio $r > 0$ al conjunto $B(x, r) := \{y \in \mathbb{R}^n : \|y - x\|_2 < r\}$, donde $\|\cdot\|_2$ es la norma euclídea en \mathbb{R}^n .

Teorema B.3.2. *Todo conjunto conexo por arcos de \mathbb{R}^n es conexo.*

Teorema B.3.3. *Todo conjunto abierto conexo de \mathbb{R}^n es conexo por arcos.*

Sus demostraciones pueden verse en Apostol [6], Teoremas 8–30 y 8–31, páginas 173 y 174. Se pueden hacer unas precisiones añadidas sobre las **componentes conexas** de un conjunto $S \subset \mathbb{R}^n$.

Definición B.3.4. Si T es un subconjunto conexo de S que no está contenido en ningún otro subconjunto conexo de S , se dice que T es una **componente conexa** de S .

Como ya se dijo, las componentes conexas de S son los subconjuntos conexos maximales de S . El siguiente teorema nos será de aplicación a las componentes conexas de un pseudoespectro estricto $\Lambda'_\varepsilon(A)$. Véase Apostol [6], Teorema 8–36, páginas 176 y 177.

Teorema B.3.5. *Todo punto de un conjunto $S \subset \mathbb{R}^n$ pertenece únicamente a una determinada componente conexa de S . Si S es abierto, cada componente conexa de S es un conjunto abierto conexo.*

DEMOSTRACIÓN. Si $x \in S$, la unión de todos los subconjuntos conexos de S que contienen a x debe ser una componente conexa que contiene a x . Dos componentes conexas distintas de S no pueden contener a x ; de otro modo, su unión sería un subconjunto conexo mayor que contendría a x . Esto demuestra la primera aserción.

Supongamos ahora que S es un conjunto abierto. Sea T una componente conexa de S . Veamos que cada punto $x \in T$ es interior a T . Como $T \subset S$, se tiene que $x \in S$. Como S es abierto existe una bola abierta $B(x, r)$ de radio $r > 0$ contenida en S . Pero toda bola abierta es un conjunto conexo (por arcos). Por tanto, $B(x, r)$ y T son dos subconjuntos conexos de S que contienen a x . Como T es el mayor de los subconjuntos conexos que contienen a x , se sigue que $B(x, r) \subset T$. Por lo tanto T es un conjunto abierto. \square

Apéndice C

Miscelánea algebraica

C.1. Resultante de Sylvester

Toda información concerniente a dos polinomios en una variable debe estar encerrada en los coeficientes de los mismos. Debiera ser posible extraer la información sobre cuantas raíces tienen comunes los polinomios simplemente examinando sus coeficientes. Esta idea la tuvo Sylvester y es el objeto del teorema siguiente.

Recordemos que el número complejo λ_0 se dice que es una raíz de multiplicidad m del polinomio de coeficientes complejos $f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$, o que es una raíz m -múltiple, si existe un polinomio $g(\lambda)$ tal que

$$f(\lambda) = (\lambda - \lambda_0)^m g(\lambda), \text{ y } g(\lambda_0) \neq 0.$$

Utilizando la fórmula de Taylor puede demostrarse fácilmente que λ_0 es una raíz de multiplicidad m de $f(\lambda)$ si y sólo si

$$f(\lambda_0) = f'(\lambda_0) = \dots = f^{(m-1)}(\lambda_0) = 0, \quad f^{(m)}(\lambda_0) \neq 0. \quad (\text{C.1.1})$$

Teorema C.1.1 (Resultante de Sylvester). *Sean*

$$\begin{aligned} f(\lambda) &= a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_n \\ g(\lambda) &= b_0\lambda^m + b_1\lambda^{m-1} + \dots + b_m \end{aligned}$$

dos polinomios de grados $n \geq 1, m \geq 1$, con $a_0 \neq 0, b_0 \neq 0$. Una condición necesaria y suficiente para que los polinomios $f(\lambda)$ y $g(\lambda)$ tengan al menos una

raíz común es que el determinante de orden $n + m$

$$\begin{vmatrix} a_0 & a_1 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{m-1} & b_m & 0 & \cdots & 0 \\ 0 & b_0 & \cdots & b_{m-2} & b_{m-1} & b_m & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_0 & b_1 & b_2 & \cdots & b_m \end{vmatrix} \quad (\text{C.1.2})$$

sea igual a cero.

El determinante (C.1.2) es llamado la **resultante de Sylvester** de f y g . Aun más puede probarse sobre este asunto.

Teorema C.1.2. *El número de raíces comunes a los polinomios*

$$\begin{aligned} f(\lambda) &= a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_n \\ g(\lambda) &= b_0\lambda^m + b_1\lambda^{m-1} + \cdots + b_m \end{aligned}$$

de grados $n \geq 1, m \geq 1$, con $a_0 \neq 0, b_0 \neq 0$, contando sus multiplicidades es igual a la nulidad de la matriz

$$R(f, g, \lambda) := \begin{pmatrix} a_0 & a_1 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{m-1} & b_m & 0 & \cdots & 0 \\ 0 & b_0 & \cdots & b_{m-2} & b_{m-1} & b_m & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_0 & b_1 & b_2 & \cdots & b_m \end{pmatrix}. \quad (\text{C.1.3})$$

C.2. Inversa de Moore-Penrose

Dada una matriz $A \in \mathbb{C}^{m \times n}$ el sistema de ecuaciones

$$\begin{cases} AXA = A, & (1), \\ XAX = X, & (2), \\ (AX)^* = AX, & (3), \\ (XA)^* = XA, & (4). \end{cases}$$

tiene una única solución $X \in \mathbb{C}^{n \times m}$. Véase el libro de Lancaster-Tismenetsky, Cap. 12. Esta única solución es llamada la **inversa Moore-Penrose** de A . Se denota por A^\dagger . Hay métodos sencillos para hallar A^\dagger , cuando A es una matriz de orden pequeño de números racionales, que están basados en factorizaciones de A

en dos factores B y C de rango completo. También existe el método basado en la descomposición de valores singulares. No pudiendo dejar de lado programas de álgebra por ordenador como *Derive*, *Maple* y *Mathematica*, siendo potente en métodos numéricos la aplicación *Matlab*.

C.3. Álgebras de conjuntos

Dado un conjunto E se llama σ -álgebra de subconjuntos de E a toda colección \mathcal{S} de subconjuntos de E tal que:

- (I) $E \in \mathcal{S}$.
- (II) Si $A \in \mathcal{S}$, entonces $A^c \in \mathcal{S}$ donde A^c denota el complementario de A respecto de E .
- (III) Si $A, B \in \mathcal{S}$, entonces $A \cup B, A \cap B \in \mathcal{S}$.
- (IV) Si $(A_i)_{i=1}^{\infty}$ es una sucesión de elementos de \mathcal{S} , entonces

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}.$$

Ejemplos.

- (1) La colección $\{\emptyset, E\}$ es una σ -álgebra de subconjuntos de E .
- (2) La colección $\mathcal{P}(E)$ formada por todos los subconjuntos de E es otra σ -álgebra de conjuntos de E .

Toda σ -álgebra de conjuntos de E es un álgebra de Boole de conjuntos de E , pero el recíproco no es cierto.

Si $\mathcal{G} \subset \mathcal{P}(E)$, entonces existen σ -álgebras de conjuntos de E que contienen a \mathcal{G} ; por ejemplo, la misma $\mathcal{P}(E)$. Como la intersección de σ -álgebras de conjuntos de E es una σ -álgebra de conjuntos de E , tiene sentido llamar σ -álgebra de conjuntos de E generada o engendrada por \mathcal{G} a la intersección de todas las σ -álgebras de conjuntos de E que contienen a \mathcal{G} . La denotamos por $\sigma(\mathcal{G})$. Así pues,

$$\sigma(\mathcal{G}) := \bigcap_{\substack{\mathcal{S} \subset \mathcal{P}(E) \\ \mathcal{S} \text{ } \sigma\text{-álgebra}}} \mathcal{S}.$$

Llamemos \mathcal{T} a la topología ordinaria de \mathbb{R}^2 ; es decir, \mathcal{T} es la colección de todos los subconjuntos abiertos de \mathbb{R}^2 . Se llama σ -álgebra de los conjuntos de Borel de \mathbb{R}^2 , \mathcal{B} , a la σ -álgebra de conjuntos de \mathbb{R}^2 engendrada por \mathcal{T} . A los elementos de $\mathcal{B} := \sigma(\mathcal{T})$ los llamaremos **conjuntos de Borel de \mathbb{R}^2** . Es conocido que los subconjuntos de \mathbb{R}^2 que son *medibles (en el sentido de) Lebesgue* forman una σ -álgebra de conjuntos de \mathbb{R}^2 —denotémosla por \mathcal{M} — y también que $\mathcal{B} \subset \mathcal{M}$. Como $\text{Card } \mathcal{B} = c$, donde c es la potencia del continuo o el cardinal de \mathbb{R} , y

$\text{Card } \mathcal{M} \geq 2^c$, se tiene que $\mathcal{B} \neq \mathcal{M}$. Véanse la Sección 2.5.3, página 46, de Rudin [63] y las Observaciones 10.11, página 247, de Rudin [62]. Resumamos lo que necesitamos:

Teorema C.3.1. *Todos los conjuntos abiertos de \mathbb{R}^2 son medibles Lebesgue.*

Teorema C.3.2 (Descomposición canónica de un abierto de \mathbb{R}^2). *Sea Ω un subconjunto abierto no vacío de \mathbb{R}^2 . Entonces Ω es igual a la unión de una sucesión de cuadrados Q_k , de lados paralelos a los ejes, cuyos interiores son mutuamente disjuntos, y cuyos diámetros son aproximadamente proporcionales a sus distancias a Ω^c . Más explícitamente:*

$$(I) \quad \Omega = \bigcup_{k=1}^{\infty} Q_k.$$

$$(II) \quad \overset{\circ}{Q}_j \cap \overset{\circ}{Q}_k = \emptyset \text{ si } j \neq k.$$

(III) *Existen constantes $c_1, c_2 > 0$ tales que*

$$c_1 \text{diam}(Q_k) \leq \text{dist}(Q_k, \Omega^c) \leq c_2 \text{diam}(Q_k).$$

Las constantes c_1 y c_2 son independientes de Ω . De hecho, pueden tomarse $c_1 = 1$ y $c_2 = 4$.

Para la demostración véanse el Teorema 1, p. 167 de Stein [69] y la Proposición 4.2.16, página 166, de Asplund-Bugart [10].

Epílogo

Uno de los objetivos principales, más ambicioso, de esta tesis fue extender la solución de Malyshev al *Problema de Wilkinson Generalizado*: Hallar la mínima distancia de A al conjunto de matrices que tienen una estructura de Jordan \mathcal{E} menos genérica que la de A , prefijada de antemano; lo podríamos llamar una estructura de Jordan *repelida* por la estructura de Jordan de A en el sentido de los Teoremas de Markus y Parilis (1983). En el camino aparecieron trabajos de Ikramov-Nazari (2003) y Mengi (2011) que hallaron fórmulas para la distancia $h_k(z)$ de A al conjunto de matrices que tiene a un complejo prefijado z como valor propio de multiplicidad $\geq k$, para $k = 3$ (Ikramov-Nazari) y $k > 3$ (Mengi). Sin embargo, sólo ha sido resuelto el *caso genérico* en que el valor singular σ_{kn-k+1} que se minimiza sobre $\mathbf{t} \in \mathbb{R}^{(k-1)^2}$ para hallar $h_k(z)$ es *simple* en el punto \mathbf{t}_0 en el que se alcanza el máximo. Hoy por hoy no se conoce una demostración válida para el caso general, aunque Mengi y Karow (2015) han anunciado que están trabajando sobre ello. Nosotros tampoco hemos resuelto este problema.

En esta tesis se inicia un camino para asociar a una matriz $A \in \mathbb{C}^{n \times n}$ un conjunto finito de números complejos que incluye a sus valores propios en pie de igualdad con los restantes números. Consideremos las funciones cuyos conjuntos de subnivel nos han dado los pseudoespectros algebraicos de primer, segundo, tercer, etc. órdenes: h_1, h_2, h_3, \dots , dadas por

$$h_1(x, y) := \sigma_n((x + yi)I_n - A)$$

$$h_2(x, y) := \max_{t \in \mathbb{R}} \sigma_{2n-1} \begin{pmatrix} (x + yi)I_n - A & tI_n \\ O & (x + yi)I_n - A \end{pmatrix}$$

$$h_3(x, y) := \max_{\mathbf{t} \in \mathbb{R}^4} \sigma_{3n-2} \begin{pmatrix} (x + yi)I_n - A & t_1I_n & (t_3 + t_4i)I_n \\ O & (x + yi)I_n - A & t_2I_n \\ O & O & (x + yi)I_n - A \end{pmatrix}$$

y más en general, por

$$h_k(x, y) := \max_{\mathbf{t} \in \mathbb{R}^{(k-1)^2}} \sigma_{kn-k+1}(I_k \otimes ((x + yi)I_n - A) + T_k(\mathbf{t}) \otimes I_n), k = 2, 3, \dots, n,$$

siendo

$$T_k(\mathbf{t}) := \begin{pmatrix} 0 & t_1 & t_k + t_{k+1}i & \cdots & & t_{(k-1)^2-1} + t_{(k-1)^2}i \\ 0 & 0 & t_2 & t_{k+2} + t_{k+3}i & \cdots & \vdots \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ \vdots & & & & & t_{k-1} \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix}.$$

Los números en cuestión vienen determinados por los puntos (x_0, y_0) donde alguna de las n funciones h_1, h_2 , etc. alcanza un mínimo relativo. A los correspondientes números complejos $\lambda_0 := x_0 + y_0i$ los llamaremos los **supervalores propios** de A . Véase la Figura C.1. El conjunto de supervalores propios de A está formado por los valores propios de A junto con los puntos de colisión de los pseudoespectros algebraicos de todos los órdenes de 1 a n de la matriz $A \in \mathbb{C}^{n \times n}$.

Los valores propios de A corresponderían a los puntos donde h_1 alcanza sus mínimos relativos. Así como las derivadas de las áreas de componentes conexas $\mathcal{K}_\lambda(\varepsilon)$ de $\Lambda_\varepsilon(A)$ nos dan los números de condición de los valores propios de A , esperamos que las derivadas de las áreas de las componentes conexas de los pseudoespectros $\Lambda_{\varepsilon,2}^a(A)$ alrededor del supervalor propio $\lambda_1 := x_1 + y_1i$ correspondiente a un punto (x_1, y_1) donde la función h_2 alcanza un mínimo relativo, den lugar a un concepto análogo.

Esperamos también que el teorema de la conservación de las multiplicidades de los pseudovalores propios en una componente conexa $\mathcal{K}_{\lambda_0}(\varepsilon)$ de $\Lambda_\varepsilon(A)$ tenga un correlato en la conservación de ciertos enteros asociados a los pseudosupervalores propios de segundo orden presentes en la componente conexa $\mathcal{K}_{\lambda_1,2}(\varepsilon)$ del pseudoespectro $\Lambda_{\varepsilon,2}^a(A)$.

Conjeturamos que si $(x_k, y_k) \in \mathbb{R}^2$ es un punto de coalescencia de dos componentes conexas de $\Lambda_{\varepsilon,k}^a(A)$, donde $\varepsilon := h_k(x_k, y_k)$, entonces $h_k(x_k, y_k) = h_{k+1}(x_k, y_k)$. Este resultado generalizaría la Proposición 3.3.1. También conjeturamos que si (x_k, y_k) es un primer punto de coalescencia entonces la función h_{k+1} alcanza su mínimo absoluto en $(x_k, y_k) \in \mathbb{R}^2$. Y los puntos de coalescencia de $\Lambda_{\varepsilon,k}^a(A)$ son puntos de ensilladura de h_k y son puntos en los que h_{k+1} alcanza un mínimo local. Véase la Figura C.2.

Ya hemos dicho que los pseudoespectros de una matriz $A \in \mathbb{C}^{n \times n}$ permanecen invariables al someter a la matriz A a una semejanza unitaria U^*AU . Las formas reducidas bajo semejanza unitaria han sido utilizadas por Boulton et al. [16] y [15] para el estudio de los puntos singulares de la frontera de los pseudoespectros de ciertos niveles. A propósito, queremos dejar constancia de la sorprendente analogía que presenta la matriz $T_k(\mathbf{t})$ con las formas reducidas bajo la semejanza unitaria de una matriz $M \in \mathbb{C}^{k \times k}$. Véase Shapiro [64]. Esto no debe ser una casualidad; debe haber una relación causal que la explique.

Todas las figuras que vienen a continuación se construyeron a partir de la

matriz

$$A := \begin{pmatrix} -1 & 5 & 6 \\ 0 & -2 & 0 \\ 0 & 0 & -1,8 + 0,2i \end{pmatrix}, \quad \varepsilon_1 < \varepsilon_2 < \varepsilon_3 < \varepsilon_4.$$

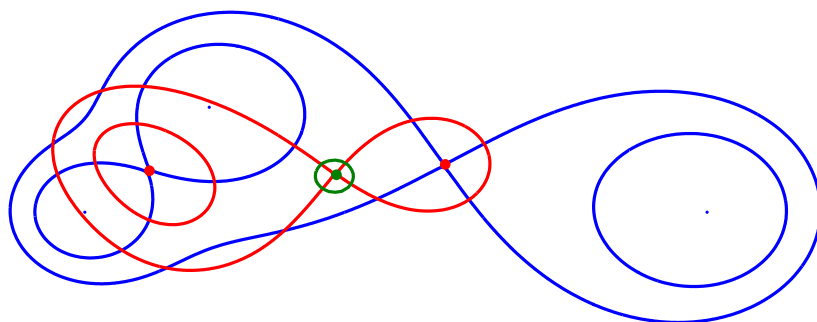


Figura C.1: Los seis supervalores propios de la matriz A .

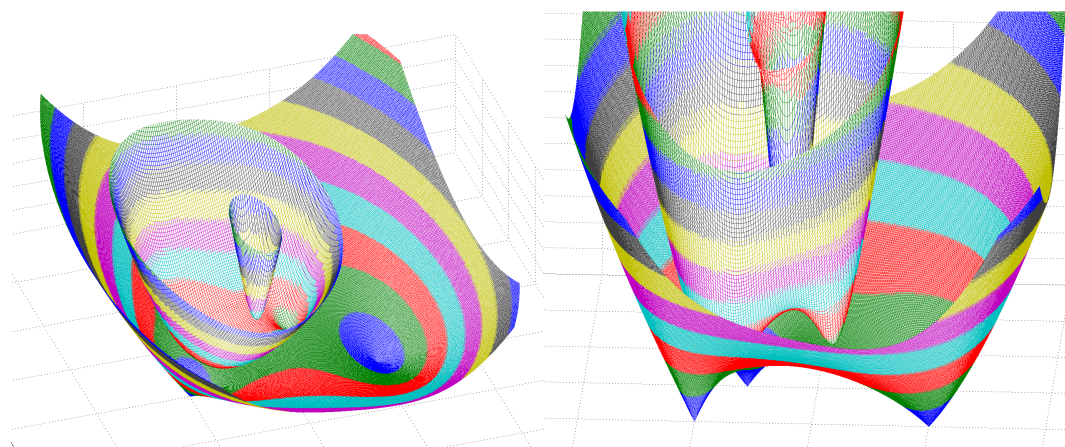


Figura C.2: Las funciones h_1, h_2 y h_3 para la matriz A .

Bibliografía

- [1] ABHYANKAR, S. S. *Algebraic geometry for scientists and engineers*. Amer. Math. Soc., Providence, Rhode Island, 1990. 7, 150, 157, 168
- [2] AHUES, M., AND HAMA, M. On pseudo spectra of matrices. *International Journal of Pure and Applied Mathematics* 89, 3 (2013), 305–322. 15
- [3] ALAM, R., AND BORA, S. On sensitivity of eigenvalues and eigendecompositions of matrices. *Linear Algebra Appl.* 396 (2005), 273–301. XIII, 11, 22, 23, 83
- [4] ALAM, R., AND BORA, S. Characterization and construction of the nearest matrix via coalescence of pseudospectral components. *Linear Algebra Appl.* 435, 3 (2011), 494–513. 23, 26
- [5] ALEKSIEJCZYK, M., AND SMOKTUNOWICZ, A. On properties of quadratic matrices. *Mathematica Pannonica* 11, 2 (2000), 239–248. 56
- [6] APOSTOL, T. M. *Análisis matemático*. Reverté, Barcelona, 1960. 51, 178
- [7] ARMENTIA, G., GRACIA, J.-M., AND VELASCO, F.-E. Derivatives of the diameter and the area of a connected component of the pseudospectrum. *Electronic Journal of Linear Algebra* 22 (October 2011), 1004–1019. 37
- [8] ARMENTIA, G., GRACIA, J.-M., AND VELASCO, F.-E. Identical pseudospectra of any geometric multiplicity. *Linear Algebra Appl.* 436, 6 (2012), 1683–1688. 42
- [9] ARRONDO, E. Apuntes de curvas algebraicas. <http://www.mat.ucm.es/~arrondo/curvas.html>, 2015. [Accedido el 9 junio 2015]. 156, 165
- [10] ASPLUND, E., AND BUNGART, L. *A first course in integration*. Holt, Rinehart and Winston, New York, 1966. 182
- [11] AUBIN, J.-P. *Mutational and morphological analysis*. Birkhäuser, Boston, 1999. 37
- [12] BHATIA, R. *Matrix analysis*. Springer, 1997. 17, 62

- [13] BOCHNAK, J., COSTE, M., AND ROY, M. *Real algebraic geometry*. Springer, 1998. 84, 86
- [14] BOLTE, J., DANILIDIS, A., AND LEWIS, A. Tame functions are semismooth. *Math. Program.* 117 (2009), 5–19. 91
- [15] BOULTON, L., AND LANCASTER, P. Schur decompositions of a matrix and the boundary of its pseudospectrum. *SIAM J. Matrix Anal. Appl.* 31, 5 (2010), 2921–2933. 184
- [16] BOULTON, L., LANCASTER, P., AND PSARRAKOS, P. On pseudospectra of matrix polynomials and their boundaries. *Math. Comp.* 77 (2008), 313–334. 184
- [17] BURKE, J. V., LEWIS, A. S., AND OVERTON, M. L. Optimization and pseudospectra, with applications to robust stability. *SIAM J. Matrix Anal. Appl.* 25, 1 (2003), 80–104. 11, 12, 18, 113
- [18] CARTAN, H. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Hermann, 115, boulevard Saint–Germain, Paris, VI, 1961. 152
- [19] CHOQUET, G. *Cours d'analyse. Topologie*. Masson, Paris, 1964. Tome II. 174, 175, 176
- [20] CIPRA, B. Are eigenvalues overvalued? *SIAM News* 28, 1 (January 1995), 8. x
- [21] COHEN, A. Is the polynomial so perfidious? *Numer. Math.* 68 (1994), 225–238. x
- [22] COURANT, R., AND JOHN, F. *Introduction to calculus and analysis II*. John Wiley and Sons, 1974. 53
- [23] DE LA PUENTE, M. J. Curvas algebraicas reales planas. In *Libro de homenaje al profesor D. Joaquín Arregui Fernández*. Editorial Complutense, Madrid, 2000, pp. 249–263. 139
- [24] DE LA PUENTE, M. J. Real plane algebraic curves. *Expo. Math.* 20 (2002), 291–314. 139
- [25] DE LA PUENTE, M. J. *Curvas algebraicas y planas*. Servicio de Publicaciones de la Universidad de Cádiz, Cádiz, 2007. 122, 139, 148, 150, 164
- [26] DE LA PUENTE, M. J. Comunicación privada, 2012. 122
- [27] DEMMEL, J. *A numerical analyst's Jordan canonical form*. PhD thesis, University of California, Berkeley, 1983. XIII
- [28] DEMMEL, J. W. Computing stable eigendecompositions of matrices. *Linear Algebra Appl.* 79 (1986), 163–193. XIII

- [29] DIXMIER, J. *Cours de mathématiques du premier cycle. Première année.* Gauthier-Villars, Paris, 1968. 146
- [30] DUGUNDJI, J. *Topology.* Allyn and Bacon, Inc., Boston, 1966. 177
- [31] EMBREE, M. Fact 5(b) in pseudospectra, chapter 23. In *Handbook of linear algebra. Second edition*, L. Hogben, Ed. Chapman and Hall/CRC, 2014, pp. “23–2”. 73
- [32] EMBREE, M., AND TREFETHEN, L. N. Generalizing eigenvalue theorems to pseudospectra theorems. *SIAM J. Sci. Comp.* 23, 2 (2001), 583–590. 18
- [33] EMBREE, M., AND TREFETHEN, N. Pseudospectra gateway. <http://web.comlab.ox.ac.uk/projects/pseudospectra/index.html>, 2000. IX, 12
- [34] FISCHER, G. *Plane algebraic curves.* Amer. Math. Soc., 2001. 150, 156
- [35] GEORGE, A., AND IKRAMOV, K. Unitary similarity of matrices with quadratic minimal polynomials. *Linear Algebra Appl.* 349 (2002), 11–16. 56
- [36] GIBSON, C. *Singular points of smooth mappings.* Pitman, London, 1079. 144, 169
- [37] GIBSON, C. *Elementary geometry of algebraic curves.* Cambridge University Press, 1998. 139, 149
- [38] GRACIA, J.-M. Multiplicities of the structured pseudoeigenvalues. arXiv:0907.1980v2 [math.SP], 2009. 20
- [39] GRACIA, J.-M. Directional derivatives of the singular values of matrices depending on several real parameters. arXiv:1504.01679v1 [math.SP], 2015. VIII, 106, 107, 108
- [40] HINRICHSSEN, D., AND PRITCHARD, A. J. *Mathematical systems theory I.* Springer, Berlin, 2005. IX, 28
- [41] HOFFMAN, J., MADDEN, J., AND ZHANG, H. Pseudozeros of multivariate polynomials. *Mathematics of computation* 72, 242 (2003), 975–1002. XIII
- [42] HOUSEHOLDER, A., AND CARPENTER, J. The singular values of involutory and of idempotent matrices. *Numer. Math.* 5 (1963), 234–237. 56
- [43] JACOBSON, N. *Basic algebra II. Second edition.* Dover, Mineola, New York, 1989. 154
- [44] KAROW, M. *Geometry of spectral value sets.* PhD thesis, Universität of Bremen, 2003. v, x, XIII, 9, 12, 18, 20, 37, 39, 40, 41, 42, 43, 45, 55, 56, 83, 84, 86, 121
- [45] KATO, T. *A short introduction to perturbation theory for linear operators.* Springer, 1982. 17

- [46] KENDIG, K. *A guide to plane algebraic curves*. MAA, Washington, 2011. [161](#)
- [47] KOUTIS, I. Spectrum through pseudospectrum. arXiv:math/0701368v1 [math.NA] 13 Jan 2007, 2008. [52](#)
- [48] LANCASTER, P., AND TISMENETSKY, M. *The theory of matrices with applications*, second ed. Academic Press, 1985. [5](#), [17](#), [35](#)
- [49] LANG, S. *Algebra*, revised third ed. Springer, 2002. [15](#)
- [50] LEFSCHETZ, S. *Algebraic geometry*. Princeton University Press, Princeton, New Jersey, 1953. [168](#)
- [51] LIPPERT, R. A. Fixing two eigenvalues by a minimal perturbation. *Linear Algebra Appl.* *406* (2005), 177–200. [111](#)
- [52] LORENZ, T. *Mutational analysis*. Springer, Berlin, 2010. [37](#)
- [53] MAGNUS, J., AND H.NEUDECKER. *Matrix differential calculus with applications in statistics and econometrics*, 2 ed. John Wiley and Sons, New York, 1999. [103](#)
- [54] MALYSHEV, A.-N. A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues. *Numer. Math.* *83*, 3 (1999), 443–454. [v](#), [28](#), [103](#)
- [55] MOSIER, R. G. Root neighborhoods of a polynomial. *Mathematics of Computation* *47*, 175 (July 1986), 265–273. [x](#)
- [56] MURNAGHAN, F. D. On the unitary invariants of a square matrix. *Anais da Academia Brasileira de Ciencias* *26*, 1 (1954), 1–7. [126](#)
- [57] ĐOKOVIĆ, D. Ž., AND JOHNSON, C. R. Unitarily achievable zero patterns and traces of words in A and A^* . *Linear Algebra Appl.* *421* (2007), 63–68. [116](#)
- [58] PAULSEN, V. Continuous canonical forms for matrices under unitary equivalence. *Pacific Journal of Mathematics* *76*, 1 (1978), 129–142. [127](#)
- [59] PEARCY, C. A complete set of unitary invariants for 3×3 complex matrices. *Trans. Amer. Math. Soc.* *104*, 3 (Sept. 1962), 425–429. [116](#), [125](#), [126](#)
- [60] QIU, L., BERNHARDSSON, B., RANTZER, A., DAVISON, E., YOUNG, P., AND DOYLE, J. A formula for computation of the real stability radius. *Automatica* *31*, 6 (1995), 879–890. [xii](#)
- [61] QUEIRÓ, J., AND SÁ, E. M. Singular values and invariant factors of matrix sums and products. *Linear Algebra Appl.* *225* (1995), 43–56. [77](#)
- [62] RUDIN, W. *Principios de análisis matemático*, 2 ed. Ediciones del Castillo, Madrid, 1966. [182](#)

- [63] RUDIN, W. *Análisis real y complejo*, 2 ed. Alhambra, Madrid, 1974. 182
- [64] SHAPIRO, H. A survey of canonical forms and invariants for unitary similarity. *Linear Algebra Appl.* 147 (1991), 101–167. 184
- [65] SHARGORODSKY, E. On the definition of pseudospectra. *Bull. London Math. Soc.* 41 (2009), 524–534. 16
- [66] SHIKIN, E. V. *Handbook and atlas of curves*. CRC Press, 1995. 168
- [67] SIBIRSKIĬ, K. Unitary and orthogonal invariants of matrices. *Soviet Math. Dokl.* 8 (1967), 36–40. 116
- [68] SIBIRSKIĬ, K. S. A minimal polynomial basis of unitary invariants of a square matrix of the third order. *Math. Notes* 3, 3 (1968), 185–187. 116, 126
- [69] STEIN, E. *Singular integrals and differentiability properties of functions*. Princeton Univ. Press, Princeton, New Jersey, 1970. 182
- [70] SUN, J.-G. A note on simple non-zero singular values. *Journal of Computational Mathematics* 6, 3 (1988), 258–266. 103
- [71] TREFETHEN, L. Pseudospectra of linear operators. *SIAM Review* 39, 3 (1997), 383–406. IX
- [72] TREFETHEN, L. N., AND EMBREE, M. *Spectra and pseudospectra*. Princeton University Press, 2005. IX
- [73] WALKER, R. J. *Algebraic curves*. Dover, New York, 1950, 1978. 134, 139, 150, 154, 156, 158, 165, 167, 168, 169
- [74] WIKIPEDIA. Glossary of classical algebraic geometry. http://en.wikipedia.org/wiki/Glossary_of_classical_algebraic_geometry, 2015. 144
- [75] WIKIPEDIA. Tacnode. <http://en.wikipedia.org/wiki/Tacnode>, 2015. 148, 149
- [76] WILKINSON, J. On neighbouring matrices with quadratic elementary divisors. *Numer. Math.* 44 (1984), 1–21. XIII
- [77] WILKINSON, J. Sensitivity of eigenvalues. *Utilitas Mathematica* 25 (1984), 5–76. XII
- [78] WILKINSON, J. Sensitivity of eigenvalues II. *Utilitas Mathematica* 30 (1986), 243–286. XII, XII
- [79] WILKINSON, J. H. The perfidious polynomial. In *Studies in Numerical Analysis*, G. H. Golub, Ed. Mathematical Association of America, 1984, pp. 1–28. X
- [80] ZHANG, F. *Matrix theory. Basic results and techniques*. Springer, 1999. 56

Índice alfabético

A

acnodo, 144
algcurses, paquete, 163
 σ -álgebra, 181
 engendada por una colección de subconjuntos, 181
álgebra de Boole, 84
algebraico
 conjunto, 139
algoritmo de Newton-Puiseux, 163
aplicación afín, 140

B

bloque de Jordan generalizado, 55
bola abierta en \mathbb{R}^n , 178

C

centro
 de rama abstracta, 165
 de una parametrización local, 159
coalescencia, 20
colisión, 20
 primer punto de, 20
componente conexa, 176
conexión de la clausura, 176
conjunto de Borel, 181
cuerpo de las series de Puiseux $K\{\{T\}\}$, 156
cuerpo de las series de Puiseux, 156
curva algebraica
 plana, 139
 real plana, 139
curva lisa, 128

cúspide ordinaria, 168
cúspide ranfoidea, 168

D

derivada de una serie formal, 153
dimensión infinita, 16
distancia de Hausdorff, 9

E

equivalentes según Hardy, 73
espacio de Banach, 16
espacio topológico
 conexo, 175
espectro
 de orden k , 1
 geométrico
 de orden k , 1

F

forma, 142
frontera de componente conexa, 39
funciones
 equivalentes según Hardy, 73

H

homotecia, 38

I

imagen conexa, 176
índice de un valor propio, 8
inflexión
 tangente de, 142
inversa Moore-Penrose, 180

M

Maple, [161](#), [163](#)
 Maple, algoritmo de Newton-Puiseux, [163](#)
 matrix
 simple, [23](#)
 multiplicidad
 de un punto, [141](#)
 de una rama, [168](#)
 multiplicidad de intersección
 de una curva con una recta, [141](#)
 de dos curvas, [167](#)
 de dos ramas, [169](#)
 de una rama con una curva, [166](#)

N

Newton
 algoritmo, [163](#)
 Teorema de, [156](#), [157](#)
 norma
 de Frobenius, [2](#)
 espectral, [1](#)
 número de condición
 de una matriz invertible, [35](#)
 número de condición de Hölder, [9](#)
 número de condición de un valor propio, [9](#)
 número de intersección
 de dos curvas, [167](#)
 de dos ramas, [169](#)
 de una curva con una recta, [141](#)

O

operador lineal acotado, [16](#)
 orden de una serie formal
 de Laurent, [154](#)
 de potencias, [151](#)
 de Puiseux, [156](#)
 ordinaria
 tangente, [142](#)

P

parametrización formal en un punto de una curva afín, [166](#)

parametrización local, [158](#)
 irreducible, [158](#)
place
 rama en inglés, [165](#)
 plot_real_curve, [161](#)
 polinomio
 mónico, [15](#)
 Pompeiu, *véase* distancia de Hausdorff
 primer punto de colisión, [20](#)
 pseudoespectro
 estricto
 algebraico, [11](#)
 ordinario
 algebraico, [11](#)
 pseudoespectro estricto, [11](#)
 pseudoespectro ordinario, [11](#)
 Puiseux
 serie, [155](#)
 Teorema de, [156](#)
 puiseux, [163](#)
 punto
 cuspidal, [148](#)
 punto
 tacnodo, [146](#)
 punto aislado, [144](#)
 punto cuspidal, [146](#), [148](#)
 ordinario, [149](#)
 punto de autoosculación, [148](#)
 punto de coalescencia, [20](#)
 punto de colisión, [20](#), [21](#)
 punto de retroceso, [146](#)
 de primera especie, [146](#)
 de segunda especie, [146](#)
 punto singular
 crunodo, [144](#)

R

radio
 máximo, [54](#)
 mínimo, [54](#)
 medio, [53](#)
 radio medio, [53](#)
 rama abstracta, [165](#)
 rama de una curva, [159](#)
 ramas reales, [164](#)

recta tangente a una curva algebraica,
142

S

segundo punto de colisión de pseudo-
espectros, 34

semejanza

unitaria, 3

serie formal

de Laurent, 154

de potencias, 151

de Puiseux, 154

singularities, 163

subconjunto conexo, 175

supervalores propios, 184

T

tacnodo, 149, 169

entre circunferencias, 169

ordinario, 149

oscnode, 169

tangente de inflexión, 142

tangente ordinaria, 142

Teorema

de Newton-Puiseux, 156

n -tupla desordenada, 62, 157

V

valores singulares, 2

variedad real analítica

unidimensional, 86