



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Problemas de perturbación de objetos espectrales discontinuos en haces matriciales

Autor:

Fernando de Terán Vergara

Directores:

Froilán Martínez Dopico

Julio Moro Carreño

DEPARTAMENTO DE MATEMÁTICAS

Leganés, Octubre de 2007

TESIS DOCTORAL

Problemas de perturbación de objetos espectrales discontinuos en haces matriciales

Autor: **Fernando de Terán Vergara**

Directores: **Froilán Martínez Dopico**
Julio Moro Carreño

Firma del Tribunal Calificador:

	Firma
Presidente:	
Vocal:	
Vocal:	
Vocal:	
Secretario:	

Calificación:

Leganes,

de

de

Agradecimientos

Aún a riesgo de que suene a tópico, algo me obliga a comenzar diciendo que hay muchas personas sin las que este trabajo no habría sido posible. Por un lado, están las que han contribuído a él de manera indirecta; por otro, quienes lo han hecho de forma directa y, por último, están los que ya no están, que no pertenecen estrictamente a ninguno de los dos grupos anteriores pero quizá deban incluirse en ambos.

El primero de los grupos, bastante extenso, comprende a casi todas aquellas personas que han supuesto algo importante para mí, pero, como no me parece que la sección de agradecimientos de una tesis doctoral sea el lugar apropiado para narrar la historia de una vida (y menos de la mía), lo más conveniente será destacar unos cuantos nombres que me es imposible obviar. Empezando por mis padres, que han hecho tanto por mí, y en consecuencia por este trabajo, que seré incapaz de agradecerles en toda mi existencia, y siguiendo por el resto de mis familiares y amigos y, en general, todos aquellos que me aprecian y me apoyan, consciente o inconscientemente. También quiero testimoniar mi gratitud especial hacia Mario Salas, entre otras cosas, por haberme dado el empujón que necesitaba para embarcarme en esta aventura.

En el grupo de personas que han colaborado directamente en este trabajo se encuentran aquellas que, aparte de mí mismo, conocen y entienden todo su contenido. Se trata, naturalmente, de los dos directores de esta tesis, Julio Moro y Froilán M. Dopico, cuya aportación ha sido fundamental, tanto en la elección de los problemas que tratamos como en su desarrollo, revisión y presentación. Quiero agradecerles no sólo su confianza, su ayuda y sus enseñanzas en la materia de que es objeto esta tesis sino, sobre todo, su apoyo moral y humano que ha sido incondicional a lo largo de estos años.

Finalmente, quiero recordar a aquellos que me ayudan desde un lugar desconocido. En especial, a mi abuelo y a mi tío Quique, de los que me acuerdo todos los días de mi vida.

Siempre me había parecido un tanto pretencioso el desfile habitual de agradecimientos que suelen contener las tesis como esta. Quizá no entendí, como lo entiendo ahora, que, al margen de la calidad o relevancia científica del trabajo, está el gran esfuerzo que se ha destinado a él, y que se traduce en mucho, mucho tiempo invertido. En mi caso, buena parte de los últimos cinco años de mi vida, dedicados a una relación no siempre amistosa y que en determinados momentos ha tenido trazas de obsesión. Por eso, por el tiempo que le debo, por su paciencia y por su amor constante, no puedo terminar sin nombrar a una de las dos personas que dan más sentido a todo esto. Gracias, Natalia.

A Lishé

Índice general

1. Introducción	1
1.1. Teoría de perturbación de primer orden en haces singulares cuadrados . . .	5
1.2. Perturbaciones de rango bajo de formas canónicas	6
1.3. Órbitas de haces matriciales	9
1.4. Organización de la memoria	10
1.5. Símbolos y notación	11
2. Herramientas básicas	13
2.1. Definiciones básicas	14
2.2. La forma de Weierstrass	16
2.3. La forma de Kronecker	17
2.4. La forma de Smith	20
2.4.1. La forma de Smith, los índices minimales y el autovalor infinito . .	23
2.5. El espacio de las n-uplas de funciones racionales. Bases minimales	24
2.6. Subespacios reductores	28
2.7. Aplicaciones	32
3. El Polígono de Newton	37
3.1. El polígono de Newton	37
3.2. Desarrollos en torno a un punto	38
3.2.1. El Polígono de Newton y los desarrollos asintóticos	39
3.2.2. Puntos críticos	42
4. Teoría espectral de primer orden	47
4.1. Introducción	47
4.2. Preliminares	51
4.3. Existencia de desarrollos	54
4.4. De la forma de Kronecker a la forma local de Smith	57
4.5. Desarrollos de Puiseux para los autovalores del haz perturbado	61
4.5.1. Haces auxiliares	62
4.5.2. Desarrollos de primer orden para autovalores finitos	64
4.5.3. El autovalor infinito	67
4.5.4. Desarrollos para autovalores simples	68
4.6. Autovectores del haz perturbado	70
4.6.1. El caso de autovalores simples	73

5. Perturbaciones de rango bajo de la Forma de Weierstrass	75
5.1. Introducción	75
5.2. Cotas inferiores	79
5.3. Comportamiento genérico para el autovalor cero	82
5.4. Perturbaciones de rango bajo de haces regulares	93
5.4.1. Notación	93
5.4.2. El resultado principal	95
6. Perturbaciones de rango bajo de la Forma de Kronecker	99
6.1. Introducción	99
6.2. Desarrollo de un haz como suma de haces de rango 1	101
6.3. Hipótesis de rango bajo: significado y genericidad	103
6.3.1. Genericidad de la hipótesis $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$. Número de índices minimales de $H + M$	104
6.4. La estructura regular del haz perturbado	106
6.5. Los índices minimales del haz perturbado	108
6.5.1. Polinomios conexión y matrices mosaico Toeplitz asociadas	109
6.5.2. Propiedades de las matrices mosaico Toeplitz	112
6.5.3. Los índices minimales derechos genéricos de $H + M$	117
6.5.4. Ilustración del Teorema 6.5.10 con un ejemplo	121
6.5.5. Sobre la genericidad de las hipótesis	122
6.5.6. Cuando la única información disponible de la perturbación es su rango	122
6.6. La forma canónica de Kronecker de haces perturbados sin rango completo .	123
6.6.1. Relación con la teoría de órbitas	128
6.6.2. Cuando H y M son ambos genéricos	129
6.7. El caso límite $\rho = \min\{p, q\}$	132
6.7.1. El caso rectangular	133
6.7.2. El caso cuadrado	142
6.8. Problemas relacionados	143
6.9. Apéndice: Programas Matlab	144
6.9.1. Programas Matlab	144
6.9.2. Experimentos	149
7. Órbitas de haces matriciales de rango fijo	151
7.1. Introducción	151
7.2. Preliminares y resultados previos	152
7.2.1. Órbitas y la forma canónica de Kronecker	152
7.2.2. Particiones enteras	152
7.2.3. Relaciones de inclusión entre cierres de órbitas	153
7.3. El conjunto de haces singulares de rango menor o igual que r	154
7.3.1. Irreducibilidad en la topología de Zariski	157
7.4. Notas históricas	158
8. Conclusiones, publicaciones y problemas abiertos	161

Capítulo 1

Introducción

La presente memoria se enmarca en el ámbito de la teoría espectral de haces matriciales. Entendemos por *haz matricial* un polinomio de grado 1 en una variable escalar cuyos coeficientes son matrices con entradas complejas o, equivalentemente, una matriz cuyas entradas son polinomios de grado 1 en una variable escalar (λ en esta memoria) con coeficientes complejos. En general, utilizaremos la primera definición, lo que permite escribir el haz en la forma $A_0 + \lambda A_1$, donde A_0 y A_1 son matrices de la misma dimensión. Diremos que el haz es *regular* si la matriz es cuadrada y su determinante no es el polinomio idénticamente nulo. En caso contrario, es decir, cuando la matriz no es cuadrada o cuando es cuadrada con determinante idénticamente nulo, se dice que el haz es *singular*. A partir de esta definición, tiene sentido hablar del *rango* de un haz matricial como la dimensión del mayor menor del haz no idénticamente nulo. Este concepto se encuentra habitualmente en la literatura bajo el nombre de *rango normal*. En particular, el rango de un haz regular coincide con la dimensión de dicho haz. Los objetos fundamentales de la teoría espectral son los *autovalores* y *autovectores*. Un autovalor *finito* de un haz es un valor complejo, λ_0 , tal que el rango de la matriz $A_0 + \lambda_0 A_1$ es menor que el rango del haz. Esta definición es válida tanto para haces regulares como para singulares. En el caso regular, es equivalente a la igualdad $\det(A_0 + \lambda_0 A_1) = 0$. Esto implica que existe un vector no nulo, x_0 , tal que $(A_0 + \lambda_0 A_1)x_0 = 0$. A este vector se le denomina *autovector* del haz asociado al autovalor λ_0 . Por otra parte, en los haces matriciales, puede también aparecer el *autovalor infinito*. Diremos que un haz $A_0 + \lambda A_1$ tiene un autovalor infinito si el *haz dual* $A_1 + \lambda A_0$ tiene un autovalor nulo.

Las formas canónicas más comunes para haces de matrices, y que son objeto de esta memoria, son la extensión de la *forma canónica de Jordan* para matrices cuadradas. Dicha forma canónica es invariante por *semejanza*, mientras que las formas canónicas de haces matriciales son invariantes por *equivalencia estricta*, esto es, por multiplicación (a la izquierda o a la derecha) por matrices constantes invertibles. Las formas canónicas de haces matriciales fueron introducidas por Weierstrass [92] para haces regulares y por Kronecker [39], que extendió la definición de Weierstrass a haces rectangulares. Dichas formas contienen toda la información espectral de los haces matriciales. Esto significa, en particular, que los autovalores de un haz matricial no cambian cuando éste se multiplica por una matriz invertible. Además, y al igual que ocurre en el caso matricial con la relación de semejanza, hay una única forma canónica por cada clase de equivalencia con respecto a la relación de equivalencia estricta.

Los haces matriciales surgen de manera natural en diversos contextos, por ejemplo, a

partir de los sistemas lineales de ecuaciones diferenciales ordinarias de orden 1:

$$A_1 x'(t) + A_0 x(t) = f(t), \quad (1.1)$$

donde A_0 y A_1 son matrices $m \times n$ con entradas complejas, $x(t)$ es un vector de incógnitas con n entradas y $f(t)$ es un vector cuyas m entradas son funciones de la variable t . El sistema anterior tiene asociado el haz

$$H(\lambda) = A_0 + \lambda A_1,$$

y las soluciones del sistema (1.1) están íntimamente relacionadas con la información espectral del haz $H(\lambda)$. Aunque en el apartado §2.7 expondremos con detalle esta relación, consideremos, a título ilustrativo, una solución del sistema homogéneo $A_1 x'(t) + A_0 x(t) = 0$ de la forma:

$$x(t) = e^{\lambda_0 t} x_0,$$

donde x_0 es un vector constante. Dicha solución ha de cumplir:

$$A_1(x_0 \lambda_0 e^{\lambda_0 t}) + A_0 e^{\lambda_0 t} x_0 = 0,$$

o bien:

$$(\lambda_0 A_1 + A_0)x_0 = 0.$$

Es decir, en el caso de que el haz sea regular, λ_0 es un *autovalor* de $H(\lambda)$ y x_0 es un *autovector* de $H(\lambda)$ asociado a λ_0 .

La relevancia de los haces matriciales en este contexto se extiende a los sistemas lineales de ecuaciones diferenciales ordinarias de orden superior

$$A_\ell x^{(\ell)}(t) + A_{\ell-1} x^{(\ell-1)}(t) + \dots + A_1 x'(t) + A_0 x(t) = f(t), \quad (1.2)$$

(donde $x^{(k)}$ denota la derivada k -ésima del vector $x(t)$ respecto a la variable t y A_0, \dots, A_ℓ son matrices $m \times n$), a través de las llamadas *linealizaciones* (véase la Sección 2.7). La solución de tales sistemas (cuando el polinomio matricial asociado, $A_0 + \lambda A_1 + \dots + \lambda^\ell A_\ell$, es *regular*) también se puede obtener mediante la información espectral de una linealización adecuada, como veremos en §2.7. Por otra parte, en las ciencias físicas es de especial importancia el sistema (1.2) con $\ell = 2$, que comprende el caso de los *sistemas vibratorios* y que ha sido tratado ampliamente en la literatura (citaremos, como muestra, las referencias [25, 41, 43, 44, 54, 56, 77, 88, 89]). Otro contexto en el que las formas canónicas de haces matriciales tienen especial relevancia es el de la Teoría de Control. Como ejemplo, citaremos uno de los resultados más elementales en este contexto, conocido como el criterio de Hautus [65, p. 272], según el cual, un par (A, B) es controlable si y sólo si el haz matricial $[A - \lambda I \ B]$ tiene rango completo para cualquier $\lambda \in \mathbb{C} \cup \{\infty\}$, lo que equivale a decir que el haz anterior no tiene autovalores.

En este trabajo nos ocuparemos de la teoría de perturbaciones de ciertas propiedades espectrales de haces matriciales. La teoría de perturbaciones de autovalores ha sido objeto de estudio en el campo de la matemática aplicada desde que Lord Rayleigh estableciera sus orígenes en [57]. En uno de sus cálculos pretendía determinar tanto las frecuencias como los modos de vibración de una cuerda oscilante con elasticidad constante y cuya masa y densidad eran una pequeña desviación de un valor fijo para el que se conocían dichas frecuencias y modos de oscilación. Este problema particular ilustra perfectamente una situación típica de la teoría de perturbaciones de autovalores: una matriz, u operador,

de partida se supone una pequeña desviación de una determinada matriz (u operador) próximo para el que el problema espectral está completamente resuelto (y, en la mayoría de los casos, de modo más sencillo). El operador dado, \tilde{A} , se sustituye por un operador cercano A cuyos autovalores y autovectores son conocidos y, a partir de ahí, se emplea una teoría de perturbaciones apropiada para analizar la influencia de la desviación $B = \tilde{A} - A$ sobre los objetos espectrales. En uno de sus trabajos sobre mecánica cuántica, Schrödinger [64] dio un paso importante en el desarrollo formal de esta teoría: describió un observable de un sistema mecánico cuántico como un operador autoadjunto U (el Hamiltoniano en el caso de la energía) definido en un cierto espacio de Hilbert. Un autovalor aislado λ_0 y su correspondiente autovector normalizado x_0 se interpretan como un estado x_0 con nivel de energía λ_0 si el observable es el Hamiltoniano. Si el sistema está influido por algún campo externo, o si ha de tenerse en cuenta alguna interacción, al operador U ha de añadirse una perturbación U_1 . El tratamiento de perturbaciones autoadjuntas de operadores autoadjuntos es, por tanto, uno de los primeros problemas tratados y resueltos en el marco de la teoría de perturbaciones de problemas espectrales [4, §3.5]. La teoría de perturbación de autovalores y autovectores (o subespacios invariantes) de matrices y operadores lineales ha gozado de un desarrollo sistemático en excelentes tratados como los de Rellich [58], Kato, [38], Baumgärtel [4] o Stewart y Sun [69]. Por otro lado, la teoría de perturbación espectral en haces matriciales tiene una historia más reciente, aunque también ha recibido bastante atención a partir de los trabajos de Stewart [66, 67].

En todos los trabajos citados en el párrafo anterior se estudia la variación de objetos continuos, como son los autovalores o los subespacios invariantes de operadores lineales y de haces matriciales regulares. Por el contrario, existen otros objetos espectrales que pueden sufrir, en determinados casos, cambios severos incluso bajo la influencia de pequeñas perturbaciones. El objetivo de esta memoria es precisamente el estudio de estos objetos como, por ejemplo, bloques de Jordan de tamaño mayor que 1 (que corresponden necesariamente a autovalores *mal condicionados*) o autovalores en haces matriciales singulares. El primer tipo de objetos discontinuos que tratamos son los autovalores de haces matriciales singulares. Estos objetos son no genéricos (lo que se traduce en que los haces singulares no tienen, genéricamente, autovalores) y discontinuos. Con respecto a esto último, la situación es, a priori, mucho más dramática debido a que existen perturbaciones arbitrariamente pequeñas de un haz singular que pueden llevar sus autovalores a cualquier punto del plano complejo. En cambio, una de las aportaciones más relevantes de esta tesis consiste en demostrar que estas perturbaciones son muy particulares y que, por el contrario, perturbaciones infinitesimales provocan, en general, variaciones infinitesimales de los autovalores. En otras palabras: los autovalores de haces singulares son continuos para la mayoría de las perturbaciones. Esto nos permitirá desarrollar una teoría de primer orden para autovalores que será válida para perturbaciones genéricas.

El segundo tipo de objetos discontinuos que tratamos son las formas canónicas de haces matriciales. Por una lado, la *forma canónica de Weierstrass* (FCW) para haces regulares y, por otro, la *forma canónica de Kronecker* (FCK) para haces singulares. Nuestro interés, como muestra el excelente trabajo de I. de Hoyos [8], se centra precisamente en aquellos puntos en los que dichas formas son discontinuas. La fragilidad de las estructuras en que nos centraremos se pone de manifiesto en el hecho de que desaparecen completamente tras perturbaciones genéricas. Por ejemplo, los bloques de Jordan de tamaño mayor que 1 asociados a un autovalor λ_0 de un haz regular se destruyen por completo al añadir una pequeña perturbación, para la mayor parte de estas perturbaciones, dando lugar a una nube de autovalores simples en torno a λ_0 . También, la parte singular de un haz matri-

cial singular cuadrado desaparece, en general, al introducir una pequeña perturbación, puesto que los haces cuadrados son genéricamente regulares. Dicho de otra manera, en esta memoria estudiamos el comportamiento de estructuras no genéricas. Aunque estas estructuras son poco habituales en la práctica, precisamente por no ser genéricas, han atraído la atención de la comunidad científica del álgebra lineal numérica debido a que su presencia implica serios problemas desde el punto de vista computacional.

En cualquier caso, el intenso desarrollo que ha experimentado la teoría de perturbaciones de objetos espectrales continuos contrasta con la escasez de resultados para los objetos espectrales discontinuos (las formas canónicas de Weierstrass o Kronecker o los autovalores de haces singulares). Nuestro trabajo es un humilde aporte que pretende ayudar a la superación de estas carencias.

Desde otra perspectiva, podemos también decir que existen dos ámbitos distintos en la teoría espectral de perturbaciones atendiendo al tipo de perturbaciones que se consideren. El primero de ellos se centra en perturbaciones *no estructuradas*, cuyo estudio cobra sentido cuando las perturbaciones son de norma pequeña, es decir: pequeñas desviaciones de la matriz (o función matricial) de partida, y es, sin duda, el que más interés ha despertado desde el punto de vista científico no solamente en el contexto aplicado que hemos mencionado al principio de esta Introducción, sino también porque este tipo de perturbaciones son las que aparecen debidas a errores de redondeo. El segundo campo se ocupa de perturbaciones *estructuradas* es decir, con una cualidad concreta, como la de ser de *rango bajo* o una estructura determinada, como puede ser la simetría, la antisimetría, etc. También en este campo se incluyen las perturbaciones que afectan sólo a una parte de las funciones matriciales que se estén tratando (matrices, haces, polinomios...), como puedan ser algunas entradas concretas, o ciertas filas o columnas. Naturalmente, existen conexiones e intersecciones entre estos dos campos o entre problemas distintos de un mismo campo. Por ejemplo, una perturbación que sólo modifique la entrada (i, j) -ésima de una matriz es una perturbación de rango 1 de la forma $ae_i e_j^T$ (donde e_k es el k -ésimo vector canónico y a es un escalar). En este segundo tipo las perturbaciones pueden tener norma pequeña o arbitraria dependiendo del objeto inicial o del problema que se quiera estudiar.

Cada uno de los dos grandes problemas que abordamos en esta memoria pertenece a uno de los dos ámbitos distintos que acabamos de mencionar. El problema de describir el cambio en la estructura canónica que contiene la información espectral de los haces matriciales (la forma canónica de Weierstrass en el caso regular y la forma canónica de Kronecker en el singular) tras perturbaciones de *rango bajo* pertenece al segundo de ellos porque las perturbaciones que consideramos en este caso son de **norma arbitraria**. Por el contrario, las perturbaciones que consideramos para describir el cambio que experimentan los autovalores de haces matriciales singulares cuadrados son de **norma pequeña**.

A la hora de obtener el cambio genérico de la forma canónica de Kronecker de un haz cuando es sometido a perturbaciones de rango bajo, surge de modo natural la cuestión de conocer la forma canónica genérica, si existe, de los haces singulares de un rango prefijado, ρ . La descripción de dicho conjunto nos ayudará a entender las condiciones de genericidad sobre las que se sustentan algunos de los resultados fundamentales acerca de este tipo de perturbaciones. El conjunto de haces matriciales de rango menor o igual que un valor determinado es una variedad que se descompone en una serie de subvariedades (llamadas *órbitas*) formadas por haces que tienen la misma forma canónica y entre las que existe una jerarquía determinada por relaciones de inclusión entre sus cierres topológicos.

A lo largo del presente trabajo se utilizará con frecuencia el término *genérico*. Aunque esta palabra aparece en numerosos trabajos matemáticos, no es un término unívocamente

definido, y su significado preciso no es siempre el mismo en la literatura. En esta memoria, utilizamos *genérico* en el siguiente sentido: *diremos que una propiedad es genérica en un conjunto \mathcal{C} si la propiedad se cumple en un subconjunto abierto denso de \mathcal{C}* . En nuestro contexto, \mathcal{C} será el conjunto de perturbaciones admisibles e identificamos el conjunto de haces matriciales complejos de tamaño $m \times n$ con el espacio \mathbb{C}^{2mn} , dotado de la topología usual. Por lo tanto, todo subconjunto \mathcal{C} de haces matriciales puede interpretarse como un subconjunto de \mathbb{C}^{2mn} , y un conjunto $\mathcal{G} \subset \mathcal{C}$ será denso en \mathcal{C} si y sólo si cada elemento de \mathcal{C} es el límite de una sucesión de elementos de \mathcal{G} , mientras que \mathcal{G} será *abierto* en \mathcal{C} si \mathcal{G} es la intersección de \mathcal{C} con un subconjunto abierto de \mathbb{C}^{2mn} . En otras palabras, consideramos en \mathcal{C} la *topología del subespacio* inducida por la topología usual de \mathbb{C}^{2mn} . En determinados momentos de esta memoria se justificará que una propiedad es genérica aduciendo que el conjunto en que no se satisface es una subvariedad algebraica propia. Ha de notarse que esta condición implica la genericidad tal y como la hemos definido previamente.

1.1. Teoría de perturbación de primer orden en haces singulares cuadrados

Como ya hemos mencionado antes, dado un haz matricial singular, $H(\lambda) = A_0 + \lambda A_1$, es bien sabido que hay perturbaciones arbitrariamente pequeñas de $H(\lambda)$ que pueden llevar sus autovalores a cualquier lugar del plano complejo, es decir, los autovalores son funciones discontinuas de las entradas de A_0 y A_1 . Por lo tanto, no es posible desarrollar una teoría de perturbación de autovalores para perturbaciones arbitrarias de $H(\lambda)$. En cambio, veremos que se pueden restringir las perturbaciones a un conjunto particular en el que los autovalores cambien de manera continua. Demostraremos que este conjunto de perturbaciones es genérico si $H(\lambda)$ es cuadrado, es decir, que contiene a casi todos los haces matriciales, y presentaremos condiciones suficientes para que un haz pertenezca a este conjunto.

Las perturbaciones que consideramos pueden ser descritas mediante un parámetro pequeño, ϵ , de manera que nuestro interés se centra en estudiar los autovalores del haz perturbado

$$P(\lambda, \epsilon) = H(\lambda) + \epsilon M(\lambda) = A_0 + \lambda A_1 + \epsilon(B_0 + \lambda B_1) \quad (1.3)$$

en relación con los del haz de partida, $H(\lambda)$. Para las perturbaciones, $M(\lambda)$, que se encuentren en el conjunto genérico antes aludido, obtendremos el término de primer orden de los desarrollos en ϵ en torno a un autovalor, λ_0 , de $H(\lambda)$ en función del haz de perturbación $M(\lambda)$ y de determinadas bases de los espacios nulos derecho e izquierdo de $H(\lambda_0)$, tanto para autovalores simples como para autovalores múltiples. El caso simple será tratado en una sección adicional por su especial relevancia (los haces matriciales cuadrados tienen, genéricamente, todos sus autovalores simples) y, también, porque las fórmulas que se derivan para este caso son independientes de las bases de los núcleos que se consideren. Hemos de hacer notar que el autovalor infinito estará también incluido en nuestro estudio mediante el procedimiento habitual de considerar el autovalor cero del haz dual.

También obtendremos desarrollos asintóticos para los autovectores. El concepto de autovector en un haz singular no aparece en la literatura por las razones que expondremos en §2.6, pero sí está perfectamente definido el núcleo, $\ker H(\lambda_0)$, asociado con un autovalor λ_0 . Cuando un haz singular se perturba como en (1.3) se obtiene, genéricamente, un

haz regular, para el que están bien definidos los autovectores. Cada uno de los desarrollos de autovalores de (1.3), digamos $\lambda(\epsilon)$, en torno a λ_0 tiene asociado un autovector en $\ker P(\lambda(\epsilon), \epsilon)$. Cuando ϵ tiende a cero, dicho autovector se aproxima a un vector de $\ker H(\lambda_0)$. Lo que haremos en la Sección 4.6 es determinar dicho vector.

La teoría espectral de perturbaciones de haces matriciales singulares contrasta con la de haces regulares, en que los autovalores son siempre funciones continuas de las entradas. Esto último se deriva del hecho de que los autovalores son las raíces de un polinomio (el determinante del haz) y que las raíces de un polinomio dependen continuamente de sus coeficientes y éstos, a su vez, de las entradas del haz. Este argumento no es válido en el caso singular porque los autovalores no son las raíces del determinante del haz, que ahora es idénticamente nulo. No obstante, es posible reducir el caso singular a un problema regular, como veremos, utilizando las nociones de geometría algebraica que subyacen a la teoría de perturbaciones. Podemos ver la ecuación polinómica $f(\lambda, \epsilon) = \det P(\lambda, \epsilon) = 0$ como una curva algebraica. Si $f(\lambda, 0)$ no es idénticamente nulo, en torno a cada una de las raíces, λ_0 , de $f(\lambda, 0)$ hay un determinado número de *ramas* (llamadas de *Puiseux*), es decir, desarrollos, $\lambda(\epsilon)$, en serie de potencias de ϵ , de manera que $\lambda(0) = \lambda_0$ y $f(\lambda(\epsilon), \epsilon) \equiv 0$ [32, §12.1]. Para el caso en que $f(\lambda, 0) \equiv 0$ esto sigue siendo cierto una vez *regularizado* el problema, es decir: cuando $f(\lambda, 0)$ es idénticamente nulo, existe un número natural k de modo que el polinomio $f(\lambda, \epsilon)$ puede expresarse de la forma

$$f(\lambda, \epsilon) = \epsilon^k \tilde{f}(\lambda, \epsilon),$$

donde, para perturbaciones genéricas, $\tilde{f}(\lambda, 0)$ no es idénticamente nulo. El problema, por tanto, se regulariza considerando, para $\epsilon \neq 0$, la ecuación $\tilde{f}(\lambda, \epsilon) = 0$ en lugar de $f(\lambda, \epsilon) = 0$. Concretamente, el conjunto de las raíces del polinomio $\tilde{f}(\lambda, 0)$ incluye a los autovalores de $H(\lambda)$ y las ramas de la curva anterior en torno a ellos son las ramas de $f(\lambda, \epsilon) = 0$ en torno a estos autovalores. El conjunto de perturbaciones para las que los autovalores cambian de forma continua será el formado por aquellas perturbaciones que hacen que el polinomio $\tilde{f}(\lambda, 0)$ sea no idénticamente nulo para un valor determinado del exponente k .

La teoría espectral de perturbaciones de haces singulares que acabamos de describir es el objeto del Capítulo 4, y se encuentra en [19].

1.2. Perturbaciones de rango bajo de formas canónicas

Como hemos mencionado, las formas canónicas espectrales, cuya relevancia abarca tanto el ámbito teórico como el práctico, son estructuras matemáticas muy frágiles bajo perturbaciones. En cambio, hay perturbaciones que nos permiten garantizar que una parte de la forma canónica espectral del haz original permanece en la forma canónica espectral del haz perturbado. A este tipo de perturbaciones pertenecen las *perturbaciones de rango bajo*, esto es: perturbaciones con un rango fijado, que es pequeño en un sentido que viene especificado por alguna propiedad de la matriz o el haz no perturbados.

Las perturbaciones de rango bajo de propiedades espectrales han aparecido en diversos problemas aplicados. Por ejemplo, en el área de modificaciones estructurales de sistemas dinámicos, es de particular relevancia estudiar cómo debe modificarse un sistema a fin de fijar determinados autovalores del nuevo sistema. Este problema es conocido, genéricamente, como de “asignación de ceros y polos” (citamos, a modo de ejemplo, las referencias [56] y [54], aunque la bibliografía es más amplia y puede consultarse en las obras

citadas). Por otro lado, en [88] y [89] se consideran perturbaciones de rango 1 de la matriz de amortiguamiento en sistemas vibratorios, una situación “*físicamente no desdeñable que ocurre cuando un sistema vibratorio elástico se controla pasivamente con unos pocos amortiguadores*” [88, §1]. En el contexto de modificaciones estructurales las perturbaciones de rango bajo representan modificaciones que afectan a pocos grados de libertad o a una parte aislada del sistema. Este tipo de modificaciones, aunque de tamaño arbitrario, son, en ocasiones, preferibles a pequeñas modificaciones que afectan a todo el sistema. También en [25], se consideran perturbaciones de rango bajo de la matriz de amortiguamiento de un sistema vibratorio con el fin de obtener sistemas defectivos.

Las perturbaciones de rango bajo de formas canónicas espectrales han sido objeto de atención desde los años 80. Se pueden considerar, al menos, dos tipos de contribuciones en esta área. Dados un haz de tamaño $m \times n$, $H(\lambda)$, y perturbaciones $M(\lambda)$ de un rango determinado, la primera clase de trabajos trata de clasificar todas las formas canónicas espectrales de $(H + M)(\lambda)$ que son compatibles con la forma canónica de $H(\lambda)$ y con el rango de las perturbaciones $M(\lambda)$. Hasta donde llega nuestro conocimiento, esto se ha hecho únicamente para perturbaciones de rango 1. Citamos las referencias [6] y [76] en este contexto. Una segunda clase de trabajos caracterizan las propiedades genéricas de la forma canónica espectral de $(H + M)(\lambda)$, es decir: las propiedades que se dan para todas las perturbaciones en un subconjunto abierto denso del conjunto de haces matriciales de un rango determinado. El cambio genérico en la forma de Jordan de matrices cuadradas se ha descrito en [34] y, posteriormente, en [52] y [60, 61]. El estudio de las condiciones necesarias y suficientes para que se dé el comportamiento genérico se ha desarrollado en [52]. Quedaba, así, abierto el problema de extender los resultados al caso de haces matriciales. Este problema ha sido resuelto en [18] para haces matriciales regulares y en [16] para haces matriciales singulares. En los capítulos 5 y 6 de esta memoria presentamos los resultados de estos dos trabajos, precedidos de una sucinta descripción de lo que ocurre en el caso matricial, que consideramos muy ilustrativa.

La distinción entre los casos regular y singular es tradicional en la literatura relacionada con la teoría espectral de haces (y, en general, polinomios) matriciales, y en nuestro trabajo es inevitable debido a la presencia de los bloques singulares (la que denominaremos *estructura singular*) en el segundo caso. La descripción del cambio genérico de tales bloques concentra la mayor parte del esfuerzo dedicado en el Capítulo 6. A continuación, resumimos los resultados fundamentales que hemos obtenido en cada uno de estos dos casos.

El caso de haces regulares

Si A es una matriz $n \times n$, λ_0 es un autovalor de A con g bloques de Jordan asociados en su forma canónica de Jordan y B es una matriz que satisface $\text{rg}(B) < g$, la matriz perturbada $A+B$ sigue teniendo a λ_0 como autovalor, en virtud de la conocida desigualdad

$$\text{rg}(A + B - \lambda_0 I) \leq \text{rg}(A - \lambda_0 I) + \text{rg}(B),$$

donde $\text{rg} B$ denota el rango de B e I es la matriz identidad de tamaño $n \times n$. Para matrices genéricas, B , que satisfacen la cota anterior sobre el rango, se sabe que los bloques de Jordan de $A + B$ asociados al autovalor λ_0 son, exactamente, **los $g - \text{rg}(B)$ bloques de Jordan de A de menor tamaño asociados a λ_0** [34, 52, 60, 61, 62].

Sea ahora $A_0 + \lambda A_1$ un haz matricial regular y λ_0 uno de sus autovalores finitos con g bloques de Jordan asociados en la forma canónica de Weierstrass. El resultado

que demostraremos en el Capítulo 5 afirma que, para un par de matrices genéricas B_0 y B_1 con $\text{rg}(B_0 + \lambda_0 B_1) < g$, hay $g - \text{rg}(B_0 + \lambda_0 B_1)$ bloques de Jordan asociados al autovalor λ_0 en la forma de Weierstrass del haz perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$. Si $\text{rg}(B_0 + \lambda_0 B_1) + \text{rg}(B_1)$ no excede el número de bloques de Jordan asociados a λ_0 en $A_0 + \lambda A_1$ con dimensión mayor que 1, entonces los bloques de Jordan asociados a λ_0 en el haz perturbado son **los $g - \text{rg}(B_0 + \lambda_0 B_1) - \text{rg}(B_1)$ bloques de menor tamaño asociados a λ_0 en $A_0 + \lambda A_1$, junto con $\text{rg}(B_1)$ bloques de dimensión 1**. En caso contrario, **todos los $g - \text{rg}(B_0 + \lambda_0 B_1)$ bloques de Jordan asociados a λ_0 en el haz perturbado son de dimensión 1**. Esto ocurre para cualquier par de matrices B_0 y B_1 , excepto para aquellas que se encuentran en una subvariedad algebraica propia, definida por una ecuación determinantal escalar. Esta ecuación involucra las entradas de las matrices B_0 y B_1 , y algunos de los autovectores de $A_0 + \lambda A_1$ asociados a λ_0 . Si $A_0 + \lambda A_1$ tiene un autovalor infinito, el resultado correspondiente se sigue de considerar el autovalor cero en los *haces duales* $A_1 + \lambda A_0$ y $A_1 + B_1 + \lambda(A_0 + B_0)$.

El caso de haces singulares

Ahora $H(\lambda) = A_0 + \lambda A_1$ es un haz singular de tamaño $m \times n$ **sin rango completo** cuya forma canónica de Kronecker es conocida. El *rango* de un haz singular H es, como ya se ha mencionado, el orden del mayor menor de H no idénticamente nulo (como función de la variable λ), y lo denotaremos por $\text{rg}(H)$. Sea ρ un número entero positivo tal que $\rho \leq \min\{m, n\} - \text{rg}(H)$. En el Capítulo 6 estudiamos el cambio de la forma canónica de Kronecker de $H(\lambda)$ frente a perturbaciones aditivas, $M(\lambda)$, que satisfacen $\text{rg}(M) = \rho$. Como en el caso regular, nos centramos en el comportamiento genérico de la forma canónica de Kronecker de $(H + M)(\lambda)$, es decir, el comportamiento que se da para perturbaciones $M(\lambda)$ en un subconjunto abierto denso del conjunto de haces con rango ρ . Las propiedades más destacables de la forma canónica de Kronecker genérica del haz perturbado $(H + M)(\lambda)$ son las siguientes:

- (i) si λ_0 es un autovalor de $H(\lambda)$, finito o infinito, entonces λ_0 es un autovalor de $(H + M)(\lambda)$;
- (ii) si λ_0 es un autovalor de $H(\lambda)$ entonces el número de bloques de Jordan asociados a λ_0 en la forma canónica de Kronecker de $(H + M)(\lambda)$ es mayor o igual que el número de bloques de Jordan asociados a λ_0 en la forma canónica de Kronecker de $H(\lambda)$;
- (iii) si λ_0 es un autovalor de $H(\lambda)$ entonces las dimensiones de los bloques de Jordan asociados a λ_0 en la forma canónica de Kronecker de $(H + M)(\lambda)$ son mayores o iguales que las dimensiones de los bloques de Jordan asociados a λ_0 en la forma canónica de Kronecker de $H(\lambda)$;
- (iv) el número de índices minimales derechos (resp. izquierdos) de $(H + M)(\lambda)$ es menor en ρ unidades que el número de índices minimales derechos (resp. izquierdos) de $H(\lambda)$;
- (v) los índices minimales derechos (resp. izquierdos) de $(H + M)(\lambda)$ son mayores o iguales que los mayores índices minimales derechos (resp. izquierdos) de $H(\lambda)$.

Más aún, si la suma de los índices minimales derechos (resp. izquierdos) de las perturbaciones $M(\lambda)$ es conocida, aparte de su rango, entonces todos los índices minimales derechos (resp. izquierdos) de $(H + M)(\lambda)$ se pueden determinar genéricamente, y en el caso en que $\rho < \min\{m, n\} - \text{rg}(H)$ y ambas sumas son conocidas, se puede determinar, genéricamente, la **forma canónica de Kronecker completa** del haz perturbado $(H + M)(\lambda)$.

Las condiciones de genericidad, en este caso, consisten en que determinadas matrices, construídas a partir de las entradas de una cierta descomposición del haz $M(\lambda)$ en suma de haces de rango 1 y unas determinadas bases de los núcleos izquierdo y derecho de $H(\lambda)$ (considerado como una matriz con entradas en el cuerpo de funciones racionales en la variable λ), sean de rango completo. Si se dan estas condiciones, la parte regular del haz perturbado $(H + M)(\lambda)$ contiene **exactamente los bloques de Jordan de $H(\lambda)$ y los de $M(\lambda)$** (es decir, es la suma directa de las partes regulares de H y de M). En cambio, la descripción de la parte singular genérica de $H + M$ no es tan sencilla, como se verá en la Sección 6.5.3.

Por último, dedicamos una sección a describir el cambio de la estructura de Kronecker en lo que denominamos el *caso límite*, es decir, el caso en que $\rho = \min\{m, n\}$. En este caso, la perturbación hace que desaparezcan todos los índices minimales izquierdos (si $m \leq n$) o derechos (si $n \leq m$), en virtud del apartado *iv*) anterior. Esta estructura singular que desaparece contribuye a la creación de una nueva estructura regular adicional que consiste, genéricamente, en autovalores simples. Dichos autovalores son las raíces de un polinomio cuyos coeficientes son descritos en la Sección 6.7.

El hecho de que, como hemos mencionado anteriormente, el rango de la perturbación no sea suficiente para determinar genéricamente los índices minimales del haz perturbado, y que sea preciso conocer, además, las sumas de los índices minimales de la perturbación, está relacionado con los resultados que presentamos en el Capítulo 7.

1.3. Órbitas de haces matriciales

Como hemos mencionado en la sección anterior, la suma de los índices minimales de una perturbación de rango bajo de un haz singular es esencial a la hora de determinar los índices minimales genéricos del haz perturbado. La descripción de una cierta estructura topológica en el conjunto de haces matriciales de rango menor o igual que r nos permitirá descubrir que este hecho no es tan sorprendente como parece a primera vista.

Aunque en los últimos años se han producido avances muy significativos en el desarrollo de algoritmos para el cálculo de la forma canónica de Kronecker (véase [22] y las referencias allí incluídas), el cálculo de dicha forma canónica es una tarea delicada y muy costosa en términos computacionales. Por lo tanto, aquellos resultados teóricos que vayan en la línea de describir la forma canónica de Kronecker genérica de algunos subconjuntos de haces matriciales son interesantes desde el punto de vista aplicado. Como se afirma en [21, p. 656]: “*si pudiéramos entender mejor la geometría del espacio de matrices, nuestro conocimiento de los algoritmos numéricos y sus fallos también mejoraría.*” Por tanto, el problema geométrico en el espacio de haces de matrices que tratamos en la última parte de esta memoria no sólo tiene sentido por sí mismo, sino que además está relacionado con la forma canónica de Kronecker y, como se ha mencionado en el párrafo anterior, con las perturbaciones de rango bajo.

La *forma de Kronecker genérica* de los haces matriciales complejos de tamaño $m \times n$,

con $m \neq n$, se ha descrito explícitamente en [11, Cor. 7.1] (véase también [22, §3.3]). Para los haces matriciales singulares de tamaño $n \times n$ existen n formas canónicas de Kronecker genéricas, cada una de las cuales corresponde a una órbita de codimensión $n + 1$ formada por haces matriciales estrictamente equivalentes. Estas estructuras de Kronecker se han descrito explícitamente en [91, Th. 1] (véanse también [11, Cor. 7.2] y [22, §3.3]). En términos matemáticos rigurosos, se puede decir, en el lenguaje de la geometría algebraica [91], que el conjunto de los haces matriciales singulares de tamaño $n \times n$ tiene exactamente n componentes irreducibles de codimensión $n + 1$, o, en el lenguaje introducido en [22], que el conjunto de haces matriciales singulares de tamaño $n \times n$ es la unión de los cierres de n órbitas maximales de haces matriciales estrictamente equivalentes. Otro resultado relevante en este contexto es que el conjunto de matrices de tamaño $m \times n$ con rango r es una variedad en \mathbb{C}^{mn} de codimensión $(m - r)(n - r)$ [11, Lemma 3.3]. En cambio, hasta la aparición de los resultados que contiene esta memoria, no existían resultados similares para el conjunto de haces matriciales con rango r . El desarrollo de este tipo de resultados es el objeto del Capítulo 7.

Demostremos, en dicho capítulo, que hay exactamente $r + 1$ formas de Kronecker genéricas del conjunto de haces de tamaño $m \times n$ con rango r . Más concretamente, para $r = 1, \dots, \min\{m, n\} - 1$, mostramos que el conjunto de haces matriciales con rango menor o igual que r es la unión de los cierres de las órbitas correspondientes a estas formas de Kronecker, y que estos cierres son maximales en el sentido de que no están contenidos en el cierre de ninguna otra órbita de haces con rango a lo sumo r . Además, calculamos las dimensiones de estas órbitas. Las formas de Kronecker genéricas de los haces con rango r no tienen parte regular, al igual que ocurre con las formas de Kronecker genéricas de los haces de tamaño $m \times n$ con rango completo y con los haces singulares $n \times n$ [11, Corolarios 7.1 y 7.2], y tienen tanto bloques singulares derechos como izquierdos. *Cada una de las formas de Kronecker genéricas con rango r depende de la suma de los índices minimales derechos, que pueden tomar los valores $0, 1, \dots, r$, o, equivalentemente, de la suma de los índices minimales izquierdos.* Es importante destacar que las órbitas correspondientes a estas $r + 1$ formas de Kronecker genéricas tienen dimensiones distintas en el caso $m \neq n$. Para demostrar estos resultados, utilizamos técnicas introducidas en [11, 22] (véase también [8]). Finalmente, presentamos un resultado adicional sobre la irreducibilidad de los cierres de las órbitas de las formas de Kronecker genéricas en la topología de Zariski [91, §1]. Nuestros resultados incluyen, como caso particular, las formas de Kronecker genéricas de los haces singulares $n \times n$.

El hecho de que las formas de Kronecker genéricas de los haces matriciales de rango r dependan no solamente del rango, sino también de la suma de los índices minimales derechos (o equivalentemente, izquierdos) explica por qué la forma de Kronecker tras una perturbación genérica de rango bajo depende de la suma de los índices minimales derechos e izquierdos de la perturbación, como ya hemos mencionado en §1.2, y, en particular, por qué no se puede determinar dicha forma de Kronecker completa cuando la única información disponible de la perturbación es su rango.

1.4. Organización de la memoria

La presente memoria está organizada de la siguiente manera.

En el Capítulo 2 introducimos las formas canónicas objeto de nuestro estudio, junto con otras definiciones básicas que necesitaremos posteriormente. También dedicamos una

sección, la §2.7, a exponer brevemente algunas de las aplicaciones más relevantes de estas formas canónicas.

El Capítulo 3 está dedicado a la descripción del *Polígono de Newton*, una herramienta utilizada en teoría de funciones algebraicas para obtener los desarrollos asintóticos de ramas locales y que aquí emplearemos para determinar los términos de primer orden de los desarrollos de autovalores.

Los tres primeros capítulos contienen principalmente resultados conocidos, mientras que en los capítulos siguientes presentamos resultados originales.

El objetivo más importante del Capítulo 4 es obtener el coeficiente y el exponente director de los desarrollos asintóticos de autovalores de una perturbación uniparamétrica de un haz matricial cuadrado y singular. El caso regular, que ha sido ampliamente tratado en la literatura, puede considerarse como un caso particular del anterior.

El Capítulo 5 está dedicado a describir el cambio genérico de la forma canónica de Weierstrass de haces matriciales regulares sometidos a perturbaciones de rango bajo, mientras que en el Capítulo 6 hacemos lo mismo para la forma canónica de Kronecker de haces singulares.

En el Capítulo 7 describimos las relaciones de dominancia entre los cierres de las órbitas de los haces singulares de rango menor o igual que r , lo que nos permite dar una descripción del conjunto de dichos haces como una variedad cerrada compuesta por la unión de $r + 1$ componentes irreducibles. También describimos explícitamente estas componentes.

Finalmente, a modo de conclusión, ofrecemos, en el Capítulo 8, un resumen de los resultados más relevantes de esta memoria, junto con una relación de los trabajos a los que ha dado lugar. También incluimos una serie de problemas abiertos que pueden abordarse como continuación natural de los aquí tratados.

1.5. Símbolos y notación

Aunque la mayoría de los símbolos y notaciones que empleamos en esta memoria son convencionales en la literatura existente, hemos creído conveniente introducir una lista de las notaciones más empleadas en el trabajo con el fin de agilizar su lectura.

Con carácter general, utilizaremos el símbolo ■ para indicar la conclusión de una demostración, mientras que □ aparecerá al final de cada ejemplo. Por otra parte, la abreviatura "ker", habitual en la literatura para designar núcleos de aplicaciones, se utilizará exclusivamente para núcleos de matrices constantes (correspondientes a aplicaciones lineales en un espacio vectorial sobre el cuerpo de los números complejos) mientras que el símbolo \mathcal{N} denotará el núcleo de matrices racionales, que corresponden a aplicaciones lineales sobre el cuerpo de funciones racionales con coeficientes complejos.

Símbolo	Significado
\mathbb{C}	Cuerpo de los números complejos
$\mathbb{C}^{m \times n}$	Espacio vectorial de las matrices $m \times n$ con coeficientes complejos
$\mathbb{C}[\lambda]$	Anillo de los polinomios en la variable λ con coeficientes complejos
$\mathbb{C}(\lambda)$	Cuerpo de las funciones racionales en la variable λ con coeficientes complejos
$\mathbb{C}^n(\lambda)$	Espacio vectorial de las n -uplas con entradas en $\mathbb{C}(\lambda)$
$\dim(V)$	Dimensión del espacio vectorial V
$\text{rg}(H)$	Rango de la matriz (o el haz) H
H^T	Traspuesta de la matriz (o el haz) H
Q^H	Traspuesta conjugada de la matriz Q
I_n	Matriz identidad de tamaño n
$\det(H)$	Determinante de la matriz (o el haz) H
$\deg(p)$	Grado del polinomio $p(\lambda)$
$\text{diag}(A_1, \dots, A_n)$	Matriz diagonal por bloques, que son A_1, \dots, A_n
$V + W$	Suma de los espacios vectoriales V y W
$\ker A$	Núcleo de la matriz A (sobre \mathbb{C})
$\mathcal{N}(H)$	Núcleo del haz H (sobre $\mathbb{C}(\lambda)$)
$A\mathcal{X}$	Espacio imagen del subespacio \mathcal{X} por la matriz A
$\mathbb{C}\langle S \rangle$	Espacio vectorial sobre \mathbb{C} generado por el conjunto S
$\text{col}(A)$	Espacio generado por las columnas de la matriz A
$P^\#$	Polinomio dual del polinomio P

Capítulo 2

Herramientas básicas

En este capítulo introducimos los conceptos y herramientas básicos sobre formas canónicas de haces matriciales que utilizaremos en los capítulos posteriores.

En primer lugar, un **polinomio matricial** es una matriz con entradas en el anillo de polinomios en la variable λ . Denotaremos a este anillo por $\mathbb{C}[\lambda]$. Un **haz matricial** es un polinomio matricial de grado 1, es decir: aquel cuyas entradas son polinomios de grado menor o igual que 1.

Antes de introducir las definiciones de equivalencia que manejaremos a lo largo de toda la memoria definamos los polinomios matriciales unimodulares: un polinomio matricial $n \times n$ es **unimodular** si su determinante es una constante no nula (en otras palabras: si la inversa del polinomio es también un polinomio matricial).

Dos haces matriciales $H_1(\lambda)$ y $H_2(\lambda)$ de tamaño $m \times n$ son **equivalentes** si son equivalentes por matrices unimodulares, es decir, si existen dos polinomios matriciales $E(\lambda)$ y $F(\lambda)$ unimodulares de tamaños $m \times m$ y $n \times n$, respectivamente, de manera que

$$E(\lambda)H_1(\lambda)F(\lambda) = H_2(\lambda).$$

Por otro lado, diremos que los haces $H_1(\lambda)$ y $H_2(\lambda)$ son **estrictamente equivalentes** si existen dos matrices $E \in \mathbb{C}^{m \times m}$ y $F \in \mathbb{C}^{n \times n}$ invertibles de modo que

$$EH_1(\lambda)F = H_2(\lambda).$$

Por último, diremos que los haces $H_1(\lambda)$ y $H_2(\lambda)$ son **localmente equivalentes en** $\lambda_0 \in \mathbb{C}$ si existen dos funciones analíticas matriciales $E(\lambda)$ y $F(\lambda)$ de tamaños respectivos $m \times m$ y $n \times n$, invertibles en λ_0 , de modo que

$$E(\lambda)H_1(\lambda)F(\lambda) = H_2(\lambda).$$

Nótese que, por continuidad, las matrices $E(\lambda)$ y $F(\lambda)$ anteriores serán invertibles en todo un entorno de λ_0 .

En la Sección 2.1 definiremos los conceptos básicos como *autovalor*, *autovector* (finito e infinito) o *bloque de Jordan*. En la Sección 2.2 describiremos la *forma canónica de Weierstrass* de haces regulares que es invariante por equivalencia estricta. En la Sección 2.3 realizamos un tratamiento análogo con los haces singulares y describimos la *forma canónica de Kronecker*, que también se obtiene del haz original por equivalencia estricta. Finalmente, en la Sección 2.4 presentamos la última de las formas canónicas que utilizaremos en este trabajo: la *forma canónica de Smith*, en sus versiones global y local. La

primera se alcanza por equivalencia de matrices unimodulares, mientras que la segunda se obtiene por equivalencia de matrices localmente invertibles.

El tratamiento que suele darse a la teoría de haces (o, en general, polinomios) matriciales en el campo del álgebra lineal parte de la interpretación de éstos como polinomios cuyos coeficientes son matrices. La definición de haz (polinomio) matricial que nosotros hemos utilizado los considera, en cambio, como matrices cuyas entradas son polinomios, es decir: como matrices con entradas en el cuerpo de las fracciones algebraicas (o funciones racionales). En la Sección 2.5 explotaremos este punto de vista, que nos permitirá observar algunos conceptos fundamentales relacionados con los polinomios matriciales a la luz de la teoría de espacios vectoriales sobre el cuerpo de las fracciones algebraicas.

En la Sección 2.6 recordaremos el concepto de *subespacio reductor* que fue introducido por Van Dooren [81] como el equivalente para haces singulares del concepto de *subespacio deflactor*, que había definido Stewart [66] años atrás para haces regulares. Estos conceptos son de capital importancia en el desarrollo de algoritmos estables para obtener la forma de Schur generalizada de haces regulares y singulares y, a partir de ella, la información relevante de las formas canónicas de Weierstrass y de Kronecker [14, 15]. Nosotros emplearemos el primero de ellos en el Capítulo 4 para dar una descripción adecuada de los coeficientes directores de los desarrollos asintóticos de autovalores de haces matriciales singulares cuadrados.

2.1. Definiciones básicas

Diremos que el haz matricial

$$H(\lambda) = A_0 + \lambda A_1. \quad (2.1)$$

es **regular** si las matrices de coeficientes A_0 , A_1 son cuadradas (pongamos de tamaño $n \times n$) y, además, $\det H(\lambda)$ es un polinomio no idénticamente nulo en λ .

Cuando las matrices coeficientes, A_0 , A_1 , del haz matricial (2.1) son rectangulares de tamaño $m \times n$ (con $m \neq n$), o bien cuando son cuadradas $n \times n$ pero $\det H(\lambda)$ es idénticamente nulo, el haz (2.1) es **singular**.

Definición 2.1.1 *El rango de un haz es el orden del mayor menor no idénticamente nulo (como función de λ).*

Nótese que, de acuerdo a la definición anterior, el rango de un haz regular $n \times n$ es exactamente n . En este apartado nos centraremos en el caso regular. En este caso, tiene sentido la siguiente definición.

Definición 2.1.2 *Un autovalor finito del haz regular (2.1) es un valor $\lambda_0 \in \mathbb{C}$ que verifica*

$$\det H(\lambda_0) = 0.$$

La multiplicidad geométrica del autovalor λ_0 es el valor

$$g = \dim(\ker H(\lambda_0)),$$

mientras que la multiplicidad algebraica de λ_0 es el valor a tal que:

$$\det H(\lambda) = (\lambda - \lambda_0)^a q(\lambda), \quad q(\lambda_0) \neq 0.$$

Diremos que un autovalor de $H(\lambda)$ es **simple** cuando su multiplicidad algebraica es igual a 1.

Diremos que un autovalor de $H(\lambda)$ es **semisimple** cuando su multiplicidad geométrica es igual a su multiplicidad algebraica.

Una **cadena de Jordan derecha de longitud $k + 1$ del haz $A_0 + \lambda A_1$ asociada al autovalor λ_0** del haz (2.1) es un conjunto ordenado de $k + 1$ vectores

$$x_0, \dots, x_k$$

que satisfacen las $k + 1$ igualdades:

$$(A_0 + \lambda_0 A_1)x_0 = 0, \quad A_1 x_0 + (A_0 + \lambda A_1)x_1 = 0, \dots, \quad A_1 x_{k-1} + (A_0 + \lambda A_1)x_k = 0.$$

Análogamente, una **cadena de Jordan izquierda de longitud $k + 1$ asociada al autovalor λ_0** del haz (2.1) es un conjunto ordenado de $k + 1$ vectores

$$y_0, y_1, \dots, y_k$$

de modo que

$$y_0^T (A_0 + \lambda_0 A_1) = 0, \quad y_0^T A_1 + y_1^T (A_0 + \lambda A_1) = 0, \dots, \quad y_{k-1}^T A_1 + y_k^T (A_0 + \lambda A_1) = 0.$$

En particular, un **autovector derecho** (resp. **izquierdo**) asociado al autovalor λ_0 es un vector x_0 (resp. y_0) que constituye una cadena derecha (resp. izquierda) de longitud 1, es decir:

$$(A_0 + \lambda_0 A_1)x_0 = 0, \quad (\text{resp. } y_0^T (A_0 + \lambda_0 A_1) = 0)$$

Todas las cadenas de Jordan comienzan por un autovector. Llamaremos **longitud** de un autovector a la longitud de la cadena de Jordan más larga que comienza en ese autovector.

Un **bloque de Jordan** de tamaño d asociado al autovalor λ_0 es una matriz de dimensión $d \times d$ de la forma:

$$J_d(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda_0 \end{bmatrix}.$$

Una **matriz de Jordan asociada a λ_0** , que denotaremos J_{λ_0} , es una suma directa de bloques de Jordan asociados a λ_0 :

$$J_{\lambda_0} = \bigoplus_{i=1}^k J_{d_i}(\lambda_0).$$

El **haz dual** del haz (2.20) es el haz matricial

$$H^\sharp(\lambda) = \lambda H(1/\lambda) = A_1 + \lambda A_0.$$

Este haz tiene relevancia en la siguiente definición.

Definición 2.1.3 *El haz regular (2.1) tiene un autovalor infinito si su haz dual, $H^\sharp(\lambda)$, tiene un autovalor 0. La multiplicidad geométrica del autovalor infinito es el valor*

$$g = \dim(\ker H_d(0)),$$

mientras que su multiplicidad algebraica es el valor a tal que:

$$\det H_d(\lambda) = \lambda^a q(\lambda), \quad q(0) \neq 0$$

(es decir: son, respectivamente, las multiplicidades geométrica y algebraica del autovalor 0 en el haz dual).

El autovalor infinito aparece cuando la matriz A_1 en (2.1) es singular. Más aún, la multiplicidad algebraica del autovalor infinito es, por definición, la dimensión del núcleo de A_1 .

Siguiendo la definición anterior, una **cadena de Jordan derecha de longitud $k+1$ asociada al autovalor infinito** es una cadena de Jordan derecha de longitud $k+1$ asociada al autovalor 0 en el haz dual. Análogamente se definen las cadenas de Jordan izquierdas asociadas al autovalor infinito.

2.2. La forma de Weierstrass

La primera de las formas canónicas que trataremos en esta memoria contiene la información espectral de un haz matricial regular, $H(\lambda)$, y es invariante por equivalencia estricta.

Teorema 2.2.1 (forma canónica de Weierstrass) *Dado el haz regular $H(\lambda) = A_0 + \lambda A_1$ con autovalores finitos $\lambda_0, \lambda_1, \dots, \lambda_k$, que tienen multiplicidades algebraicas respectivas a_0, a_1, \dots, a_k , existen dos matrices invertibles, Y, X , tales que*

$$Y^T H(\lambda) X = \mathcal{W}_H(\lambda) = \text{diag}(\lambda I_{a_0} - J_{\lambda_0}, \dots, \lambda I_{a_k} - J_{\lambda_k}, N_{m_1}, \dots, N_{m_s}) \quad (2.2)$$

donde J_{λ_i} es una matriz de Jordan de dimensión $a_i \times a_i$ asociada al autovalor λ_i , para $i = 0, 1, \dots, k$, y N_{m_i} es un bloque de Jordan asociado al autovalor infinito con multiplicidad m_i ,

$$N_{m_i} = \begin{bmatrix} -1 & \lambda & & \\ & -1 & \ddots & \\ & & \ddots & \lambda \\ & & & -1 \end{bmatrix} = \lambda J_{m_i}(0) - I_{m_i}.$$

La matriz derecha de (2.2) es la que se conoce como **forma canónica de Weierstrass (FCW) de $H(\lambda)$** . Es única salvo permutación de los bloques de Jordan diagonales y muestra explícitamente toda la información espectral del haz regular $H(\lambda)$.

La forma en que hemos enunciado el Teorema 2.2.1 procede de [10, Th. 4.10]. Para una demostración, véase [26, XII, §2, Th. 3].

La matriz

$$J = \text{diag}(J_{\lambda_0}, \dots, J_{\lambda_k}, J_{m_1}(0), \dots, J_{m_s}(0))$$

que se obtiene a partir de (2.2) se denomina la **matriz de Jordan de $H(\lambda)$** .

Las filas de la matriz Y^T en (2.2) son cadenas de Jordan izquierdas del haz, mientras que las columnas de X son cadenas de Jordan derechas. Concretamente, las columnas de X que corresponden a un bloque diagonal de la forma $\lambda I_d - J_d(\lambda_0)$ son una cadena de Jordan derecha longitud d asociada al autovalor λ_0 . Algo similar ocurre con las filas, y también con el autovalor infinito considerando los bloques correspondientes a dicho autovalor. Además, como puede verse, estas cadenas están *normalizadas* de modo que:

$$Y_F^T A_1 X_F = I, \quad Y_\infty^T A_0 X_\infty = -I \quad (2.3)$$

donde $X = [X_F \ X_\infty]$, $Y = [Y_F \ Y_\infty]$ están particionadas separando las cadenas correspondientes a los autovalores finitos y las correspondientes al autovalor infinito.

Lo que nos muestra la forma de Weierstrass (2.2) es que la estructura espectral de un haz matricial regular caracteriza la clase de equivalencia estricta de ese haz. Esto es lo que se resume en el siguiente Teorema.

Teorema 2.2.2 [26, Ch. XII, §2] *Dos haces matriciales regulares son estrictamente equivalentes si y solamente si tienen los mismos autovalores (finitos e infinitos) con los mismos bloques de Jordan asociados en la FCW.*

2.3. La forma de Kronecker

En esta sección vamos a recordar las definiciones básicas que necesitaremos posteriormente en lo que concierne a los haces matriciales singulares. En § 2.1 hemos definido el rango de un haz matricial. Algunos autores [6, 8, 22, 82] se refieren a este último concepto como el *rango normal*. Nosotros, en cambio, preferimos utilizar el nombre clásico de *rango* porque el concepto se corresponde con el rango usual si los haces se consideran como matrices con entradas en el cuerpo de las funciones racionales con coeficientes complejos.

Si $H(\lambda)$ es un haz singular cuadrado, cualquier valor $\lambda_0 \in \mathbb{C}$ satisface

$$\det H(\lambda_0) = 0,$$

lo que ha llevado a algunos autores a afirmar que cualquier valor es un autovalor de $H(\lambda)$. Por el contrario, en la presente memoria manejaremos las siguientes definiciones, que son válidas para haces no necesariamente cuadrados:

Definición 2.3.1 *Un autovalor finito de un haz singular $H(\lambda)$ es un número $\lambda_0 \in \mathbb{C}$ tal que*

$$\text{rg } H(\lambda_0) < \text{rg } H(\lambda). \quad (2.4)$$

La multiplicidad geométrica de un autovalor finito λ_0 es el valor $g(\lambda_0) = \text{rg } H(\lambda) - \text{rg } H(\lambda_0)$.

La anterior definición fue introducida en [71], y coincide con la definición de autovalor introducida en la Sección 2.1 para el caso de haces regulares y matrices. Nótese que en el término que figura a la izquierda de (2.4) aparece el rango de una matriz constante, mientras que en el término derecho aparece el rango de un haz matricial.

En lo sucesivo, por simplicidad y cuando no haya lugar a confusión, omitiremos la referencia al autovalor en la multiplicidad geométrica, y escribiremos simplemente g en lugar de $g(\lambda_0)$.

Definición 2.3.2 *Un haz singular $H(\lambda)$ tiene un autovalor infinito con multiplicidad geométrica g si y sólo si el haz dual $H^\sharp(\lambda)$ tiene un autovalor cero con multiplicidad geométrica g .*

Las definiciones de cadenas de Jordan derecha e izquierda asociadas a un autovalor en haces singulares se introducirán en la Sección 2.6, donde expondremos el problema que acarrea una definición análoga a la del caso regular y cómo evitarlo.

Por otro lado, la información relativa a la *parte singular* de un haz está contenida en los llamados **bloques singulares derechos**, $L_\sigma(\lambda)$, e **izquierdos**, $L_\sigma^T(\lambda)$, que son de la forma

$$L_\sigma(\lambda) = \begin{bmatrix} \lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & -1 \end{bmatrix} \in \mathbb{C}^{\sigma \times (\sigma+1)}.$$

Teorema 2.3.3 (Forma de Kronecker) [26, Ch. XII] *Sea $H(\lambda) = A_0 + \lambda A_1$ un haz singular, con $A_0, A_1 \in \mathbb{C}^{m \times n}$. Entonces existen dos matrices invertibles $Y \in \mathbb{C}^{m \times m}$ y $X \in \mathbb{C}^{n \times n}$ tales que*

$$Y^T H(\lambda) X = K_H(\lambda) = \text{diag}(L_{\varepsilon_1}(\lambda), \dots, L_{\varepsilon_p}(\lambda), L_{\eta_1}^T(\lambda), \dots, L_{\eta_q}^T(\lambda), \mathcal{J}_H(\lambda)) \quad (2.5)$$

donde

$$\mathcal{J}_H(\lambda) = \text{diag}(\lambda I_{a_0} - J_{\lambda_0}, \dots, \lambda I_{a_k} - J_{\lambda_k}, N_{m_1}, \dots, N_{m_s})$$

es un haz cuadrado regular en forma de Jordan que contiene la **estructura regular** de H . Los valores $\lambda_0, \dots, \lambda_k$ son los autovalores finitos de H y los números a_0, \dots, a_k son sus **multiplicidades algebraicas** respectivas. Los bloques N_{m_1}, \dots, N_{m_s} son bloques de Jordan asociados al autovalor infinito.

La matriz derecha en (2.5) es la **Forma Canónica de Kronecker** (FCK) de $H(\lambda)$ y es única salvo permutación de los bloques diagonales.

Los números $\varepsilon_1, \dots, \varepsilon_p$ se conocen como **índices minimales columna** (o **derechos**) de $H(\lambda)$, mientras que los números η_1, \dots, η_q se conocen como los **índices minimales fila** (o **izquierdos**) [26, Ch. XII]. En esta memoria, los supondremos ordenados en orden no decreciente, es decir:

$$0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p, \quad 0 \leq \eta_1 \leq \dots \leq \eta_q.$$

Los números de bloques singulares derechos e izquierdos, p y q respectivamente, satisfacen las siguientes relaciones, que se obtienen de manera inmediata a partir de (2.5):

$$\text{rg } H = m - q = n - p. \quad (2.6)$$

Por otra parte, las sumas

$$\varepsilon = \varepsilon_1 + \dots + \varepsilon_p, \quad \eta = \eta_1 + \dots + \eta_q$$

de los índices minimales satisfacen

$$\varepsilon + \eta = \text{rg } H - \text{rg } \mathcal{J}_H.$$

Nótese que, al ser \mathcal{J}_H un haz regular, su rango coincide con su dimensión.

Las sumas anteriores son cantidades fundamentales en los desarrollos del Capítulo 6, por lo que introducimos la siguiente:

Definición 2.3.4 Dado un haz singular $H(\lambda)$ con índices minimales derechos $\varepsilon_1 \leq \dots \leq \varepsilon_p$ e índices minimales izquierdos $\eta_1 \leq \dots \leq \eta_q$, las cantidades

$$\varepsilon = \varepsilon_1 + \dots + \varepsilon_p, \quad \eta = \eta_1 + \dots + \eta_q \quad (2.7)$$

se llamarán, respectivamente, **orden singular derecho** y **orden singular izquierdo** de $H(\lambda)$.

El análogo al Teorema 2.2.2 para haces singulares es el siguiente

Teorema 2.3.5 [26, Ch. XII, §5] *Dos haces matriciales singulares son estrictamente equivalentes si y solamente si tienen los mismos índices minimales (izquierdos y derechos) y los mismos autovalores (finitos e infinitos) con los mismos bloques de Jordan asociados.*

Es bien sabido que una matriz de rango r con entradas complejas puede expresarse como suma de r matrices de rango 1. Esto puede generalizarse a haces matriciales. La forma de Kronecker nos permite una demostración sencilla de este hecho, que incluimos en un lema técnico que usaremos posteriormente.

Lema 2.3.6 *Todo haz matricial de rango r puede expresarse como suma de r haces singulares de rango 1.*

Demostración. Sea $H(\lambda)$ un haz singular de rango r y denotemos por $K_H(\lambda)$ a su forma canónica de Kronecker. Podemos expresar $K_H(\lambda)$ como la suma de los siguientes haces matriciales (prescindimos, por simplicidad, de la variable λ en los bloques singulares):

1. Para cada bloque singular L_k de dimensión $k \times (k + 1)$ que aparece en $K_H(\lambda)$, tenemos que $L_k = L_k^{(1)} + \dots + L_k^{(k)}$, donde $L_k^{(j)}$ tiene la fila j -ésima igual a la fila j -ésima de L_k y el resto de las filas de $L_k^{(j)}$ son nulas. Por lo tanto, L_k es la suma de k haces singulares de rango 1.
2. Una expresión análoga se tiene para cualquier bloque singular L_p^T de dimensión $(p + 1) \times p$ que aparece en $K_H(\lambda)$.
3. Finalmente, para la parte regular de $K_H(\lambda)$, $F(\lambda) = F_0 + \lambda F_1$, que supondremos de tamaño $m \times m$, tenemos de nuevo que $F(\lambda) = F^{(1)} + \dots + F^{(m)}$, donde $F^{(j)}$ tiene la fila j -ésima igual a la fila j -ésima de $F(\lambda)$ y el resto de las filas de $F^{(j)}$ son cero. Por tanto, $F(\lambda)$ es la suma de m haces singulares de rango 1.

El haz $K_H(\lambda)$ se puede expresar como la suma de r haces singulares de rango 1 combinando simplemente los desarrollos de sus bloques singulares y de su parte regular. Lo mismo se tiene para $H(\lambda)$ dado que es estrictamente equivalente a $K_H(\lambda)$. ■

El Lema 2.3.6 será posteriormente ampliado en el Lema 6.2.1, en el que se exhiben determinadas propiedades de los sumandos (haces de rango 1) de la descomposición.

2.4. La forma de Smith

En esta sección introducimos la última de las formas canónicas que emplearemos en esta memoria: la *forma de Smith*. La definiremos para polinomios matriciales de grado arbitrario, dado que en la Sección 6.4 (concretamente en el Teorema 6.4.5) será inevitable el uso de dichos polinomios. Los conceptos básicos que aparecen en esta forma canónica, como son el de autovalor y su multiplicidad geométrica asociada, se definen de manera análoga a como se ha hecho para haces matriciales. Por otra parte, describiremos la forma canónica de Smith para el caso general de polinomios de dimensión arbitraria (no necesariamente cuadrados), como se hace, por ejemplo, en [26, Ch. VI, §2].

Teorema 2.4.1 (Forma de Smith) *Dado un polinomio matricial $m \times n$, $L(\lambda)$, existen dos polinomios matriciales unimodulares, $P(\lambda)$ y $Q(\lambda)$, de tamaños respectivos $m \times m$ y $n \times n$, de manera que*

$$P(\lambda)L(\lambda)Q(\lambda) = \Delta(\lambda), \quad (2.8)$$

donde

$$\Delta(\lambda) = \begin{bmatrix} d_1(\lambda) & & & & 0 \\ & \ddots & & & \\ & & d_r(\lambda) & & \vdots \\ & & & 0 & \\ & & & & \ddots \\ 0 & \dots & & & 0 \end{bmatrix}$$

es un polinomio matricial diagonal con polinomios mónicos $d_i(\lambda)$ en la diagonal tales que $d_i(\lambda)$ divide a $d_{i+1}(\lambda)$.

La matriz derecha de (2.8) es la que se conoce como la **forma de Smith** del polinomio $L(\lambda)$, y es única, como lo demuestra el siguiente resultado, que aporta una descripción de las entradas diagonales.

Teorema 2.4.2 *Sea $L(\lambda)$ un polinomio matricial $m \times n$. Sea $p_k(\lambda)$ el máximo común divisor (mónico) de los menores de orden k de $L(\lambda)$, si no son todos cero, y sea $p_k(\lambda) = 0$ si todos los menores de orden k de $L(\lambda)$ son cero. Sea $p_0(\lambda) = 1$ y*

$$\Delta(\lambda) = \text{diag}(d_1(\lambda), \dots, d_r(\lambda), 0 \dots, 0)$$

la forma de Smith de $L(\lambda)$. Entonces r es el rango de $L(\lambda)$ y

$$d_i(\lambda) = p_i(\lambda)/p_{i-1}(\lambda), \quad i = 1, \dots, r.$$

La manera en que hemos enunciado los dos Teoremas anteriores procede de [28, Th. S1.1], donde pueden encontrarse asimismo las demostraciones.

Los polinomios diagonales $d_1(\lambda), \dots, d_r(\lambda)$ de la forma de Smith se conocen con el nombre de **factores** (o **polinomios**) **invariantes** de $L(\lambda)$. Los factores invariantes se descomponen en producto de factores irreducibles de la siguiente manera:

$$d_i(\lambda) = (\lambda - \lambda_0)^{n_{i,0}} \dots (\lambda - \lambda_k)^{n_{i,k}}, \quad i = 1, \dots, r,$$

donde los exponentes $n_{i,j}$ son enteros no negativos (eventualmente, algunos son cero). Los factores $(\lambda - \lambda_j)^{n_{i,j}}$, con $n_{i,j} > 0$, se denominan **divisores elementales** de $L(\lambda)$. Nótese

que los autovalores de un haz son las raíces de sus polinomios invariantes o, equivalentemente, de sus divisores elementales. Si ahora reunimos los exponentes de los factores invariantes correspondientes a un valor concreto, λ_0 , tendremos, por el Teorema 2.4.1, que

$$0 \leq n_{1,0} \leq \dots \leq n_{r,0}.$$

Cuando estos exponentes son positivos se denominan **multiplicidades parciales** de $L(\lambda)$ en λ_0 . Si λ_0 es un autovalor de $L(\lambda)$, habrá, al menos, un exponente $n_{i,0}$ positivo. Si reunimos todos los que son no nulos y los ordenamos en una sucesión completando con ceros, llegamos al siguiente concepto:

Definición 2.4.3 *Dado un autovalor λ_0 del polinomio $L(\lambda)$ con multiplicidad geométrica g y multiplicidades parciales $0 < n_1 \leq \dots \leq n_g$, la **característica de Segre** de λ_0 relativa a $L(\lambda)$ es la sucesión infinita*

$$\mathcal{S}_H(\lambda_0) = (n_g, n_{g-1}, \dots, n_1, 0, \dots).$$

Si $\mu \in \mathbb{C}$ no es un autovalor de $L(\lambda)$, la característica de Segre de μ relativa a $L(\lambda)$ es la sucesión infinita nula:

$$\mathcal{S}_L(\mu) = (0, 0, \dots).$$

Observación 2.4.4 *La característica de Segre de λ_0 relativa a L es la partición conjugada (como se definirá más adelante en la Sección 7.2.2) de la **característica de Weyr** de λ_0 relativa a L :*

$$\mathcal{W}_L(\lambda_0) = (m_g, m_{g-1}, \dots),$$

donde m_i es el número de multiplicidades parciales de $L(\lambda)$ en λ_0 que son mayores o iguales que $g - i + 1$. En esta memoria utilizaremos la característica de Segre en algunos resultados, aunque en todos los casos puede hacerse, por conjugación, un tratamiento paralelo con la característica de Weyr.

Dados dos polinomios matriciales $L(\lambda)$ y $M(\lambda)$, escribiremos

$$\mathcal{S}_L(\lambda_0) \geq \mathcal{S}_M(\lambda_0) \quad \text{si} \quad (\mathcal{S}_L(\lambda_0))_i \geq (\mathcal{S}_M(\lambda_0))_i \quad \text{para todo } i > 0,$$

es decir: si la desigualdad se cumple para cada entrada de las correspondientes características de Segre.

Si expresamos los polinomios diagonales de la forma de Smith (2.8) en torno a λ_0 , tendremos $d_i(\lambda) = (\lambda - \lambda_0)^{n_i} \tilde{d}_i(\lambda)$, con $\tilde{d}_i(\lambda_0) \neq 0$ para $i = 1, \dots, g$ y $d_i(\lambda_0) \neq 0$ para $i = g + 1, \dots, r$. Premultiplicando en (2.8) por la matriz de dimensión $m \times m$:

$$\text{diag}(1, \dots, 1, 1/\tilde{d}_1(\lambda), \dots, 1/\tilde{d}_g(\lambda), 1/d_{g+1}(\lambda), \dots, 1/d_r(\lambda)),$$

que es invertible en λ_0 , llegaremos al siguiente resultado.

Teorema 2.4.5 *Sea $L(\lambda)$ un haz matricial $m \times n$. Entonces, para cada $\lambda_0 \in \mathbb{C}$ existen dos matrices $P_{\lambda_0}(\lambda)$ y $Q_{\lambda_0}(\lambda)$ invertibles en un entorno de λ_0 de modo que*

$$P_{\lambda_0}(\lambda)L(\lambda)Q_{\lambda_0}(\lambda) = \Delta_{\lambda_0}(\lambda) := \text{diag}((\lambda - \lambda_0)^{n_1}, \dots, (\lambda - \lambda_0)^{n_g}, 1, \dots, 1, 0, \dots, 0), \quad (2.9)$$

donde $n_1 \leq \dots \leq n_g$ son las multiplicidades parciales de $H(\lambda)$ en λ_0 .

La matriz derecha de (2.9) se conoce como la **forma local de Smith de $L(\lambda)$ en λ_0** , y es única. Dicha forma canónica es invariante por equivalencia local en λ_0 .

Existe una estrecha relación entre los divisores elementales y las dimensiones de los bloques de Jordan asociados a los autovalores finitos en la estructura regular de la FCK de un haz $H(\lambda)$. Este hecho se pone de manifiesto en el siguiente resultado, que es una simple consecuencia de la teoría desarrollada en [26, Ch. VI]

Lema 2.4.6 *Sea $H(\lambda)$ un haz matricial complejo de tamaño $m \times n$. Para cada divisor elemental $(\lambda - \lambda_j)^{n_{ij}}$ de $H(\lambda)$ existe un bloque de Jordan de dimensión n_{ij} asociado al autovalor finito λ_j en la estructura regular de la FCK de $H(\lambda)$. A la inversa, por cada bloque de Jordan de dimensión n_{ij} asociado al autovalor finito λ_j en la FCK de $H(\lambda)$ hay un divisor elemental $(\lambda - \lambda_j)^{n_{ij}}$ de $H(\lambda)$.*

Una consecuencia inmediata del Lema 2.4.6 es que las multiplicidades parciales de un haz $H(\lambda)$ en λ_0 coinciden con las dimensiones de los bloques de Jordan asociados a λ_0 en la FCK de $H(\lambda)$. En esta memoria ambos conceptos serán utilizados sin distinción.

Debe observarse que el Lema 2.4.6 no proporciona información sobre el autovalor infinito del haz $H(\lambda)$. Esta información se puede obtener a partir del autovalor cero del haz dual a través del Lema 2.4.7, cuya demostración es trivial.

Lema 2.4.7 *Sean A y B dos matrices complejas $m \times n$. Los haces $A + \lambda B$ y $B + \lambda A$ tienen los mismos índices minimales derechos e izquierdos. Además, el número y las dimensiones de los bloques de Jordan correspondientes al autovalor infinito en la FCK de $A + \lambda B$ son iguales al número y las dimensiones de los bloques de Jordan correspondientes al autovalor cero en la FCK de $B + \lambda A$, y viceversa.*

Si definimos las **multiplicidades parciales del haz $H(\lambda)$ en el infinito** como las multiplicidades parciales del cero en el haz dual, el Lema anterior nos permite generalizar el Lema 2.4.6 al autovalor infinito de un haz matricial.

Nótese que, para cualquier autovalor λ_0 sus multiplicidades parciales relativas a H coinciden con las dimensiones de los bloques de Jordan asociados a λ_0 en la estructura regular de la FCK de $H(\lambda)$. Esto es también válido para el autovalor infinito.

En el Lema 2.3.6 hemos visto que todo haz matricial de rango r puede expresarse como suma de r haces de rango 1. Por otro lado, la forma canónica de Smith (2.8) nos permite expresar todo haz matricial $H(\lambda)$ de rango r como

$$H(\lambda) = d_1(H) a_1(\lambda) z_1^T(\lambda) + \dots + d_r(H) a_r(\lambda) z_r^T(\lambda), \quad (2.10)$$

donde $d_1(H), \dots, d_r(H)$ son los factores invariantes, y $a_i(\lambda)$ y $z_i(\lambda)$ son vectores polinómicos (véase el primer párrafo de la Sección 2.5). Además, de acuerdo con (2.8), los vectores $a_i(\lambda)$ y $z_i(\lambda)$ son, respectivamente, las columnas de $P^{-1}(\lambda)$ y $Q^{-1}(\lambda)$. Esto implica que ni $a_i(\lambda)$ ni $z_i(\lambda)$ se pueden escribir como el producto de un polinomio escalar de grado mayor que cero por un vector polinómico, porque las matrices $P^{-1}(\lambda)$ y $Q^{-1}(\lambda)$ son unimodulares. Por tanto, (2.10) no es una descomposición como las que se describen en el Lema 2.3.6.

Aunque en esta memoria solamente utilizaremos la forma de Smith de haces matriciales, citaremos algunos trabajos en los que se considera la forma de Smith de matrices más generales, como matrices analíticas (cuyas entradas son series analíticas en una variable) (véase [45, 46]) o matrices racionales (cuyas entradas son fracciones algebraicas en una

variable) (véase [82]). De hecho, la forma de Smith es empleada con frecuencia en Teoría de Control, y su definición se extiende, en este contexto, a matrices racionales [82, 59], en lo que conoce como la **forma de Smith-MacMillan**, y cuyas entradas diagonales pueden contener exponentes negativos (incluso en el caso de haces). Tales exponentes, que equivalen a las multiplicidades parciales de la forma de Smith, se denominan **índices estructurales** [82, §2.4]. A diferencia de lo que ocurre con la forma de Smith en el infinito, la forma de Smith-MacMillan local en el infinito se puede obtener mediante *equivalencia en el infinito* (véase, por ejemplo, [85]). Las clases de equivalencia de esta relación están formadas por todas las matrices racionales que se generan multiplicando una matriz fija por matrices racionales cuyas entradas son fracciones algebraicas que no tienen “ceros” en el infinito (es decir: aquellas en las que el grado del numerador es menor o igual que el grado del denominador, y que son llamadas fracciones **propias**) y que, además, son “unimodulares en el infinito”, lo que significa que su determinante es una fracción algebraica cuyos numerador y denominador tienen el mismo grado. Esta definición permite unificar las relaciones de equivalencia para autovalores finitos e infinitos de matrices racionales sobre un sustento algebraico coherente. Debe notarse que la forma de Smith-MacMillan de un polinomio en el infinito no coincide, en general, con la forma de Smith local del polinomio dual en el cero, aunque existe una correspondencia biunívoca entre los exponentes de las entradas diagonales de una y otra (véanse [82, Th. 3.7] o [85, Prop. 1] para haces matriciales y [84, Cor. 3.1] para polinomios de grado arbitrario). Todas estas cuestiones pueden consultarse, además, en [83].

2.4.1. La forma de Smith, los índices minimales y el autovalor infinito

La forma de Smith (2.8) de un haz matricial $H(\lambda)$ no refleja la estructura del autovalor infinito ni da información acerca de los índices minimales, como demuestra el siguiente ejemplo elemental.

Ejemplo 2.4.8 Los haces de matrices,

$$L_1(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad L_2(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tienen la misma forma de Smith

$$\Delta(\lambda) = L_2(\lambda),$$

pero, mientras el primero tiene índices minimales derechos $\varepsilon_1 = 0, \varepsilon_2 = 1$ y no tiene ningún bloque de Jordan asociado al infinito, el segundo tiene índices minimales derechos $\hat{\varepsilon}_1 = 0 = \hat{\varepsilon}_2$ y un bloque de Jordan de orden 1 asociado al infinito. \square

Esto muestra, en particular, que la equivalencia unimodular no respeta la estructura espectral del autovalor infinito, y su explicación radica en el hecho de que las transformaciones unimodulares no son invertibles en el infinito.

No obstante, para obtener la información espectral del autovalor infinito en un haz matricial podemos usar la forma (local) de Smith del haz dual en torno al autovalor 0. En el ejemplo anterior, los duales de los haces dados son

$$L_1^\sharp(\lambda) = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad L_2^\sharp(\lambda) = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

La forma de Smith local en 0 de $L_1^\sharp(\lambda)$ es

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

mientras que la de $L_2^\sharp(\lambda)$ es

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

De hecho, dado que no hay autovalores finitos, éstas son también las formas de Smith globales.

2.5. El espacio de las n -uplas de funciones racionales. Bases minimales

Es bien conocido que los índices minimales derechos (respectivamente, izquierdos) de un haz singular $H(\lambda) = A_0 + \lambda A_1$, con $A_0, A_1 \in \mathbb{C}^{m \times n}$, están relacionados con los grados de ciertas soluciones polinómicas de la ecuación $(A_0 + \lambda A_1)x(\lambda) = 0$ (resp. $(A_0 + \lambda A_1)^T y(\lambda) = 0$) [26, Ch. XII], donde $x(\lambda)$ (resp. $y(\lambda)$) es una n -upla (resp. m -upla) cuyas entradas son polinomios. El vector $x(\lambda)$ se llamará **vector polinómico**. El comentario anterior pone de manifiesto que los vectores polinómicos pueden aparecer de forma natural cuando se trabaja con haces singulares. El conjunto de polinomios con coeficientes complejos es un *anillo* pero no un *cuerpo*. Esto significa que para extender muchas de las ideas elementales del álgebra lineal a los vectores polinómicos hay que considerar el cuerpo de las funciones racionales con coeficientes complejos. Por ejemplo, sean $v_1 = [1 + \lambda, 1 + \lambda]^T$ y $v_2 = [1 + \lambda^2, 1 + \lambda^2]^T$ dos vectores polinómicos. El determinante de la matriz $[v_1 | v_2]$ es obviamente cero, pero es necesario usar una función racional como coeficiente para expresar v_2 como combinación lineal de v_1 : $v_2 = \frac{1+\lambda^2}{1+\lambda} v_1$. El *cuerpo de funciones racionales* con coeficientes complejos será denotado por $\mathbb{C}(\lambda)$, y el *espacio vectorial* sobre $\mathbb{C}(\lambda)$ de las n -uplas de funciones racionales será denotado por $\mathbb{C}^n(\lambda)$.

En algún momento de esta memoria haremos uso del siguiente resultado:

Lema 2.5.1 *Sea $L(\lambda)$ un polinomio matricial de rango 1. Entonces existen dos vectores polinómicos, $x(\lambda)$ e $y(\lambda)$, de modo que*

$$L(\lambda) = x(\lambda)y(\lambda)^T.$$

Demostración. Dado que el polinomio tiene rango 1, su forma de Smith es $\text{diag}(d_1(\lambda), 0, \dots, 0)$, donde $d_1(\lambda)$ es un cierto polinomio no nulo. Por tanto

$$L(\lambda) = P(\lambda)^{-1} \text{diag}(d_1(\lambda), 0, \dots, 0) Q(\lambda)^{-1} = d_1(\lambda) p_1(\lambda) q_1(\lambda)^T,$$

donde $p_1(\lambda)$ es la primera columna de $P(\lambda)^{-1}$ y $q_1(\lambda)^T$ es la primera fila de $Q(\lambda)^{-1}$. Como $P(\lambda)$ y $Q(\lambda)$ son unimodulares, sus inversas son polinómicas, luego el resultado es inmediato. ■

Las siguientes definiciones están tomadas de [24] (véase también [37]). El **grado**, $\text{deg}(x)$, de un vector polinómico $x(\lambda)$ es el mayor grado de sus componentes. Todo subespacio vectorial \mathcal{V} de $\mathbb{C}^n(\lambda)$ tiene, al menos, una base formada por vectores polinómicos.

Se puede obtener a partir de una base cualquiera simplemente multiplicando cada vector por el producto de los denominadores de sus entradas. El **orden** de dicha base polinómica está definido como la suma de los grados de sus vectores. Una **base minimal** de \mathcal{V} es una base polinómica de \mathcal{V} que tiene el menor orden entre todas las bases polinómicas de \mathcal{V} .

A continuación introducimos algunos conceptos adicionales que usaremos frecuentemente. Dado un haz matricial $m \times n$, $H(\lambda) = A_0 + \lambda A_1$, el **espacio nulo derecho** (resp. **izquierdo**) de $H(\lambda)$ es el subespacio de $\mathbb{C}^n(\lambda)$ (resp. $\mathbb{C}^m(\lambda)$)

$$\begin{aligned} \mathcal{N}(H) &= \{x(\lambda) \in \mathbb{C}^n(\lambda) : H(\lambda)x(\lambda) = 0\} \quad (\text{resp. } \mathcal{N}(H^T) = \\ &= \{y(\lambda) \in \mathbb{C}^m(\lambda) : H^T(\lambda)y(\lambda) = 0\}). \end{aligned}$$

Un **vector nulo derecho** (resp. **izquierdo**) de $H(\lambda)$ es un *vector polinómico* contenido en $\mathcal{N}(H)$ (resp. $\mathcal{N}(H^T)$). Una **base minimal derecha ordenada de $H(\lambda)$ (ROMB)** es una base minimal, $\{x_1(\lambda), \dots, x_p(\lambda)\}$, de $\mathcal{N}(H)$ con $\deg(x_1) \leq \deg(x_2) \leq \dots \leq \deg(x_p)$. Una **base minimal izquierda ordenada de $H(\lambda)$ (LOMB)** es una base minimal, $\{y_1(\lambda), \dots, y_q(\lambda)\}$, de $\mathcal{N}(H^T)$ con $\deg(y_1) \leq \deg(y_2) \leq \dots \leq \deg(y_q)$.

Es bien sabido [26, Ch. XII, p. 38] que las dimensiones de $\mathcal{N}(H)$ y $\mathcal{N}(H^T)$ son, respectivamente, el número de bloques singulares derechos y el número de bloques singulares izquierdos de la forma de Kronecker de $H(\lambda)$. Por tanto, podemos reescribir las igualdades de (2.6) de la siguiente forma:

$$\begin{aligned} \dim \mathcal{N}(H) &= n - \text{rg}(H) \\ \dim \mathcal{N}(H^T) &= m - \text{rg}(H). \end{aligned}$$

El siguiente resultado nos dice que los coeficientes de cualquier ROMB de un mismo haz generan el mismo espacio vectorial sobre \mathbb{C} . Lo enunciaremos con carácter general para bases polinómicas de un espacio vectorial cualquiera sobre el cuerpo de funciones racionales sobre \mathbb{C} .

Lema 2.5.2 Sean $\{v_1(\lambda), \dots, v_n(\lambda)\}$ y $\{w_1(\lambda), \dots, w_n(\lambda)\}$ dos bases polinómicas de un subespacio vectorial \mathcal{V} de $\mathbb{C}^p(\lambda)$ sobre el cuerpo $\mathbb{C}(\lambda)$. Sean

$$\begin{aligned} V &= \{v_1^0, \dots, v_1^{d_1}; \dots; v_n^0, \dots, v_n^{d_n}\} \\ W &= \{w_1^0, \dots, w_1^{e_1}; \dots; w_n^0, \dots, w_n^{e_n}\} \end{aligned}$$

los coeficientes de los vectores de dichas bases. Entonces

$$\mathbb{C}\langle V \rangle = \mathbb{C}\langle W \rangle.$$

Demostración. Veremos únicamente que $\mathbb{C}\langle V \rangle \subseteq \mathbb{C}\langle W \rangle$, ya que la otra inclusión se obtiene de manera similar.

Fijemos un índice $i \in \{1, \dots, n\}$. Como $v_i(\lambda) \in \langle w_1(\lambda), \dots, w_n(\lambda) \rangle$, existen $n+1$ polinomios en λ , $a_0(\lambda), a_1(\lambda), \dots, a_n(\lambda)$, de modo que $a_0(\lambda)$ no es el polinomio nulo y

$$a_0(\lambda)v_i(\lambda) = a_1(\lambda)w_1(\lambda) + \dots + a_n(\lambda)w_n(\lambda). \quad (2.11)$$

Sin pérdida de generalidad, podemos suponer que $a_0(0) \neq 0$ ya que, de lo contrario, cancelamos la potencia de λ correspondiente en ambos lados de la igualdad (2.11) y argumentamos de la misma forma. Expresemos

$$a_0(\lambda) = a_0 + a_1\lambda + \dots + a_k\lambda^k, \quad a_0 \neq 0.$$

Igualando términos de grado 0 en (2.11) obtenemos

$$a_0 v_i^0 \in \mathbb{C} \langle W \rangle,$$

de donde se sigue que $v_i^0 \in \mathbb{C} \langle W \rangle$.

Igualando ahora los términos de grado 1 en (2.11) llegamos a

$$a_0 v_i^1 + a_1 v_i^0 \in \mathbb{C} \langle W \rangle,$$

luego, por el paso anterior

$$a_0 v_i^1 \in \mathbb{C} \langle W \rangle,$$

de donde se sigue que $v_i^1 \in \mathbb{C} \langle W \rangle$. Procediendo de igual forma con los grados superiores llegamos a que $v_i^0, \dots, v_i^{d_i} \in \mathbb{C} \langle W \rangle$, lo que demuestra el resultado. ■

El siguiente Lema nos muestra que los grados de los vectores en una ROMB (LOMB) de $H(\lambda)$ son iguales a los índices minimales derechos (izquierdos) de $H(\lambda)$. Aunque es un hecho bien conocido, incluimos una demostración por completitud.

Lema 2.5.3 Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ y $\eta_1 \leq \dots \leq \eta_q$, respectivamente, los índices minimales derechos e izquierdos del haz $H(\lambda)$. Sean $\{x_1(\lambda), \dots, x_p(\lambda)\}$ e $\{y_1(\lambda), \dots, y_q(\lambda)\}$, respectivamente, una ROMB y una LOMB de $H(\lambda)$. Entonces $\deg(x_i) = \varepsilon_i$, para $i = 1, \dots, p$, y $\deg(y_j) = \eta_j$, para $j = 1, \dots, q$.

Demostración. Demostraremos el resultado para los índices minimales derechos. Para los índices minimales izquierdos simplemente se usa $H^T(\lambda)$ y se invoca el resultado para los índices minimales derechos. Recordemos [26, Ch. XII, p. 38] la relación que hay entre los índices minimales derechos de $H(\lambda)$ y las soluciones polinómicas de $H(\lambda)x(\lambda) = 0$. Entre todas las soluciones polinómicas del sistema de ecuaciones elegimos una solución no nula, $z_1(\lambda)$, del menor grado posible. Este grado es ε_1 . Entre todas las soluciones polinómicas que son linealmente independientes de $z_1(\lambda)$ tomamos una solución, $z_2(\lambda)$, del menor grado posible. Este grado es ε_2 . Continuamos este proceso hasta obtener una *serie fundamental de soluciones* $\{z_1(\lambda), \dots, z_p(\lambda)\}$, es decir, $p = \dim \mathcal{N}(H)$ soluciones polinómicas linealmente independientes de la ecuación $H(\lambda)x(\lambda) = 0$ cuyos grados son $\varepsilon_1 \leq \dots \leq \varepsilon_p$. Una serie fundamental de soluciones no está unívocamente determinada, pero los grados de sus vectores sí lo están (véase [29, Prop. 1.7.1]) y, como demostraremos en el siguiente párrafo, toda serie fundamental de soluciones es una ROMB y viceversa.

Supongamos que existe algún índice j tal que $\deg(x_j) < \varepsilon_j$. Sea j_0 el menor de estos índices, es decir, $\deg(x_{j_0}) < \varepsilon_{j_0}$ y $\deg(x_k) \geq \varepsilon_k$ para $k = 1, \dots, j_0 - 1$. Obviamente $j_0 > 1$. Por lo tanto, $\varepsilon_{j_0-1} \leq \deg(x_{j_0-1}) \leq \deg(x_{j_0}) < \varepsilon_{j_0}$. La definición de los índices minimales implica que los vectores linealmente independientes $\{x_1(\lambda), \dots, x_{j_0}(\lambda)\}$ son combinaciones lineales de $\{z_1(\lambda), \dots, z_{j_0-1}(\lambda)\}$. Esto es imposible. En consecuencia, $\deg(x_j) \geq \varepsilon_j$ para todo $j = 1, \dots, p$, y, de hecho, $\deg(x_j) = \varepsilon_j$ para todo j , porque $\{x_1(\lambda), \dots, x_p(\lambda)\}$ es una ROMB. ■

Observación 2.5.4 Lo que nosotros, por simplicidad, hemos llamado orden de una base polinómica minimal de un espacio vectorial \mathcal{V} sobre $\mathbb{C}(\lambda)$ está definido en [24, p. 497] con el nombre de **orden dinámico invariante** de \mathcal{V} . Como sugiere esta denominación, se trata de un invariante del espacio \mathcal{V} . En la presente memoria \mathcal{V} es el espacio $\mathcal{N}(H)$, en cuyo caso el orden es la suma de los índices minimales derechos de H , o bien $\mathcal{N}(H^T)$, en cuyo caso es la suma de los índices minimales izquierdos.

También usaremos en el futuro el siguiente Lema relacionado con lo anterior:

Lema 2.5.5 Sean $H(\lambda)$ un haz cuya FCK está dada por (2.5), y $\{x_1(\lambda), \dots, x_p(\lambda)\}$ una ROMB de $H(\lambda)$. Entonces todo vector nulo derecho de $H(\lambda)$ de grado a lo sumo ε_i es una combinación lineal de $\{x_1(\lambda), \dots, x_j(\lambda)\}$ con coeficientes polinómicos, donde j es el mayor índice tal que $\deg(x_j) \leq \varepsilon_i$. En particular, todo vector nulo derecho de $H(\lambda)$ es una combinación lineal de $\{x_1(\lambda), \dots, x_p(\lambda)\}$ con coeficientes polinómicos. Un resultado similar se tiene para los vectores nulos izquierdos y una LOMB.

Demostración. El hecho de que todo vector nulo derecho es una combinación lineal de los vectores mencionados es una consecuencia inmediata de la definición de índice minimal. El hecho de que los coeficientes son polinomios se sigue de [24, Main Theorem en p. 495]. ■

En el Capítulo 6 necesitaremos garantizar la independencia lineal de ciertos conjuntos de vectores polinómicos de $\mathbb{C}^n(\lambda)$. En algunas situaciones, este problema se puede resolver a través de un problema de independencia lineal en \mathbb{C}^n . Esto se muestra en el Lema 2.5.6.

Lema 2.5.6 Sea $\{v_1(\lambda), \dots, v_r(\lambda)\}$ un conjunto de vectores polinómicos de $\mathbb{C}^n(\lambda)$. Expresemos estos vectores como

$$v_i(\lambda) = v_{i0} + \lambda v_{i1} + \dots + \lambda^{d_i} v_{id_i}, \quad \text{para } 1 \leq i \leq r,$$

donde $v_{ij} \in \mathbb{C}^n$, para todos los subíndices i, j , y $d_i = \deg(v_i(\lambda))$.

1. Si $\{v_{10}, \dots, v_{r0}\}$ es un conjunto linealmente independiente en \mathbb{C}^n entonces $\{v_1(\lambda), \dots, v_r(\lambda)\}$ es un conjunto linealmente independiente en $\mathbb{C}^n(\lambda)$.
2. Si $\{v_{1d_1}, \dots, v_{rd_r}\}$ es un conjunto linealmente independiente en \mathbb{C}^n entonces $\{v_1(\lambda), \dots, v_r(\lambda)\}$ es un conjunto linealmente independiente en $\mathbb{C}^n(\lambda)$.

Demostración. Para demostrar el primer apartado, consideramos la combinación lineal

$$\alpha_1(\lambda) v_1(\lambda) + \dots + \alpha_r(\lambda) v_r(\lambda) = 0, \quad (2.12)$$

donde $\alpha_i(\lambda)$, para $1 \leq i \leq r$, pueden elegirse polinomios, porque si fueran funciones racionales, se podría multiplicar por sus denominadores y argumentar de la misma forma. Expresemos estos polinomios como

$$\alpha_i(\lambda) = \alpha_{i0} + \lambda \alpha_{i1} + \dots + \lambda^{t_i} \alpha_{it_i}, \quad \text{para } 1 \leq i \leq r,$$

donde $\alpha_{ij} \in \mathbb{C}$, para todos los subíndices i, j . En esta situación, el vector coeficiente del término de grado cero en (2.12) es

$$\sum_{i=1}^r \alpha_{i0} v_{i0} = 0.$$

Si $\{v_{10}, \dots, v_{r0}\}$ es un conjunto linealmente independiente en \mathbb{C}^n entonces $\alpha_{10} = \alpha_{20} = \dots = \alpha_{r0} = 0$. Ahora, el vector coeficiente del término de grado uno en (2.12) es $\sum_{i=1}^r \alpha_{i1} v_{i0} = 0$, lo que implica que $\alpha_{11} = \alpha_{21} = \dots = \alpha_{r1} = 0$. Un simple argumento inductivo completa la demostración del primer apartado. Para demostrar el segundo apartado, comenzamos con el coeficiente del término de mayor grado, y efectuamos hacia abajo el argumento inductivo anterior. ■

Finalizamos este apartado con otro resultado técnico sobre la independencia lineal de vectores polinómicos que usaremos en el Capítulo 6.

Lema 2.5.7 Sean $\{z_1(\lambda), \dots, z_k(\lambda)\}$, con $k < n$, un conjunto linealmente independiente de vectores polinómicos en $\mathbb{C}^n(\lambda)$ y $\{z'_1(\lambda), \dots, z'_l(\lambda)\}$ otro conjunto de vectores polinómicos en $\mathbb{C}^n(\lambda)$ tal que $k + l \leq n$ y $\text{rg}[z_1(\lambda) | \dots | z_k(\lambda) | z'_1(\lambda) | \dots | z'_l(\lambda)] = k$. Denotemos por $\{u_1, \dots, u_n\}$ la base canónica de \mathbb{C}^n , es decir, las entradas de estos vectores son $(u_i)_j = \delta_{ij}$. Entonces existen l vectores de la base canónica, u_{j_1}, \dots, u_{j_l} , tales que

$$\{z_1(\lambda), \dots, z_k(\lambda), z'_1(\lambda) + \alpha_1 u_{j_1}, \dots, z'_l(\lambda) + \alpha_l u_{j_l}\}$$

es un conjunto linealmente independiente en $\mathbb{C}^n(\lambda)$ para cualesquiera números complejos $\alpha_1, \dots, \alpha_l$ no nulos.

Demostración. Existe, al menos, un u_{j_1} tal que el conjunto $\{z_1(\lambda), \dots, z_k(\lambda), u_{j_1}\}$ es linealmente independiente porque, de lo contrario, todos los vectores en $\{u_1, \dots, u_n\}$ serían combinaciones lineales de $\{z_1(\lambda), \dots, z_k(\lambda)\}$. Esto es imposible porque $\{u_1, \dots, u_n\}$ es también una base de $\mathbb{C}^n(\lambda)$ y $k < n$. Este argumento puede ser sucesivamente aplicado para demostrar que existen u_{j_1}, \dots, u_{j_l} vectores de la base canónica tales que el conjunto $\{z_1(\lambda), \dots, z_k(\lambda), u_{j_1}, \dots, u_{j_l}\}$ es linealmente independiente. Por tanto, el conjunto $\{z_1(\lambda), \dots, z_k(\lambda), \alpha_1 u_{j_1}, \dots, \alpha_l u_{j_l}\}$ es linealmente independiente para cualesquiera números complejos $\alpha_1, \dots, \alpha_l$ no nulos. Nótese que la hipótesis

$$\text{rg}[z_1(\lambda) | \dots | z_k(\lambda) | z'_1(\lambda) | \dots | z'_l(\lambda)] = k$$

implica que los vectores $z'_i(\lambda)$ son combinaciones lineales de $\{z_1(\lambda), \dots, z_k(\lambda)\}$ con coeficientes en $\mathbb{C}(\lambda)$. Por lo tanto, podemos efectuar operaciones elementales de columnas para transformar la matriz $[z_1(\lambda) | \dots | z_k(\lambda) | \alpha_1 u_{j_1} | \dots | \alpha_l u_{j_l}]$ en $[z_1(\lambda) | \dots | z_k(\lambda) | z'_1(\lambda) + \alpha_1 u_{j_1} | \dots | z'_l(\lambda) + \alpha_l u_{j_l}]$. Esto no altera el rango de la matriz, lo que demuestra el resultado. ■

2.6. Subespacios reductores

Es bien sabido que la definición de autovector que se utiliza en el caso de haces regulares carece de sentido cuando el haz es singular. A modo ilustrativo, consideremos el siguiente ejemplo.

Ejemplo 2.6.1 Consideremos el haz

$$A_0 + \lambda A_1 = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}$$

dado en FCK, con un autovalor simple igual a $\lambda_0 = 0$. A primera vista, parece adecuado decir que $[1 \ 0]^T$ es el autovector derecho asociado a $\lambda_0 = 0$. En cambio, nótese que

$$\begin{bmatrix} 1/\alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ \beta & 1 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}$$

para cualquier par de números α, β tales que $\alpha \neq 0$. Esto significa que cualquier vector $[\alpha \ \beta]^T$ podría considerarse autovector de $A_0 + \lambda A_1$ asociado al autovalor $\lambda_0 = 0$, es decir, cualquier vector de \mathbb{C}^2 sería un autovector asociado a $\lambda_0 = 0$, lo que implica que las matrices de cambio de base a la forma canónica de Kronecker no estarían bien definidas, a pesar de que λ_0 es un autovalor simple. □

La definición adecuada para haces singulares es la de *espacio reductor*, introducida por Van Dooren en [81].

Definición 2.6.2 [81] *Dado un haz singular $H(\lambda) = A_0 + \lambda A_1$, con $A_0, A_1 \in \mathbb{C}^{m \times n}$, un par de subespacios, $\mathcal{X} \subseteq \mathbb{C}^n$ e $\mathcal{Y} = (A_0\mathcal{X} + A_1\mathcal{X}) \in \mathbb{C}^m$ son **subespacios reductores** del haz si $\dim \mathcal{Y} = \dim \mathcal{X} - \dim \mathcal{N}(H)$ donde $\mathcal{N}(H)$ es el espacio nulo derecho de $H(\lambda)$.*

Los espacios reductores permiten “reducir” el problema de obtener la información espectral de un haz singular al problema de obtener la información espectral de haces de menor dimensión. Más concretamente, siguiendo a Van Dooren [81], denotemos por \mathcal{X} e \mathcal{Y} un par de subespacios tales que $\mathcal{Y} = A_0\mathcal{X} + A_1\mathcal{X}$. Sean U, V dos matrices unitarias particionadas de la forma

$$U = [U_1 \mid U_2], \quad V = [V_1 \mid V_2],$$

tales que $\mathcal{X} = \mathbb{C}\langle U_1 \rangle$ e $\mathcal{Y} = \mathbb{C}\langle V_1 \rangle$. Puesto que $\mathcal{Y} = (A_0\mathcal{X} + A_1\mathcal{X})$, se tiene que $V_2^H A_0 U_1 = V_2^H A_1 U_1 = 0$ (donde V^H es la traspuesta conjugada de V) y, por lo tanto,

$$V^H(A_0 + \lambda A_1)U = \widehat{A}_0 + \lambda \widehat{A}_1 := \left[\begin{array}{c|c} \widehat{A}_{01} & \widehat{A}_{02} \\ \hline 0 & \widehat{A}_{03} \end{array} \right] + \lambda \left[\begin{array}{c|c} \widehat{A}_{11} & \widehat{A}_{12} \\ \hline 0 & \widehat{A}_{13} \end{array} \right]. \quad (2.13)$$

La propiedad relevante de los espacios reductores que hemos aludido previamente se enuncia el siguiente resultado.

Teorema 2.6.3 [81, Th. 2.3] *Sean \mathcal{X}, \mathcal{Y} un par de subespacios reductores de $A_0 + \lambda A_1$. Si consideramos la partición (2.13), entonces los haces diagonales tienen orden singular izquierdo y derecho, respectivamente, igual a cero, y el espectro de $A_0 + \lambda A_1$ es la unión de los espectros de dichos haces.*

Esta propiedad, en cambio, no se satisface si en lugar de subespacios reductores consideramos otro tipo de subespacios \mathcal{X} e \mathcal{Y} que satisfagan la relación $\mathcal{Y} = A_0\mathcal{X} + A_1\mathcal{X}$. Esto es lo que se pone de manifiesto en el siguiente ejemplo.

Ejemplo 2.6.4 *Sea el haz*

$$H(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

El par de subespacios reductores

$$\mathcal{X}_1 = \left\langle \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right] \right\rangle, \quad \mathcal{Y}_1 = A_0\mathcal{X} + A_1\mathcal{X} = \left\langle \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right] \right\rangle$$

permiten particionar el haz de la forma

$$H(\lambda) = \left[\begin{array}{cc|c} \lambda & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right],$$

que nos da los haces “reducidos” $[\lambda \ 1]$ y $[0 \ 0]^T$. Tal y como establece el Teorema anterior, el primero de ellos tiene orden singular izquierdo nulo y el segundo tiene orden

singular derecho nulo. Por otra parte, la unión de los autovalores de ambos haces es vacía, lo que coincide con el espectro de $H(\lambda)$. En cambio, el par de subespacios

$$\mathcal{X}_2 = \left\langle \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle, \quad \mathcal{Y}_2 = A_0\mathcal{X} + A_1\mathcal{X} = \left\langle \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle$$

tienen asociada la partición

$$H(\lambda) = \left[\begin{array}{c|cc} \lambda & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right],$$

cuyo primer bloque diagonal es un haz regular que contiene un autovalor igual a cero, lo que no ocurre con el haz $H(\lambda)$. Nótese que el par de subespacios \mathcal{X}_2 e \mathcal{Y}_2 no es reductor, aunque $H(0) \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T = 0$, es decir, $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \in \ker H(0)$. \square

Los subespacios reductores de un haz singular permiten obtener una forma canónica triangular superior (GUPTRI), análoga a la forma de Schur de matrices [13]. Esta forma canónica puede ser calculada de manera estable [79, 14, 15], debido a que las transformaciones, como acabamos de ver, se llevan a cabo a través de matrices unitarias. Esto, en cambio, no es posible para la FCK. Esta es una de las razones por las que los subespacios reductores son tan importantes desde el punto de vista numérico [13].

Entre todos los subespacios reductores de un haz singular, hay uno que permite separar exactamente la parte singular derecha del resto del haz. La existencia de ese subespacio está garantizada por el siguiente resultado.

Teorema 2.6.5 [81] *Dado un haz singular $H(\lambda) = A_0 + \lambda A_1$, existe un subespacio reductor \mathcal{X}_{min} de modo que para cualquier otro subespacio reductor, \mathcal{X} , se tiene*

$$\{0\} \subseteq \mathcal{X}_{min} \subseteq \mathcal{X} \tag{2.14}$$

Llamaremos al espacio \mathcal{X}_{min} del teorema anterior **subespacio reductor minimal del haz $H(\lambda)$** .

Aunque no se utilizará en el resto de la memoria, consideramos ilustrativo describir, en este apartado, la relación que existe entre el espacio reductor minimal y las bases minimales de un haz singular arbitrario.

Teorema 2.6.6 *El subespacio reductor minimal, \mathcal{X}_{min} , de un haz singular $H(\lambda)$ está generado por los vectores de coeficientes de una base minimal derecha cualquiera de $H(\lambda)$*

Demostración. El espacio reductor minimal es aquel que permite separar la estructura singular derecha del haz (es decir: la que contiene los bloques singulares derechos) del resto del haz en una descomposición triangular de la forma (2.13). Por lo tanto, dada la forma de Kronecker (2.5) de $H(\lambda)$, las columnas de la matriz X que corresponden a los bloques singulares derechos son los coeficientes de una base minimal derecha de $H(\lambda)$ y generan \mathcal{X}_{min} . Como, por el Lema 2.5.2, los vectores de coeficientes de una base minimal cualquiera generan el mismo subespacio, se tiene el resultado. \blacksquare

Por otro lado, los vectores de coeficientes de una base minimal de un haz tienen la siguiente propiedad.

Lema 2.6.7 *Los vectores de coeficientes de una base minimal (derecha o izquierda) de un haz singular son linealmente independientes sobre \mathbb{C} .*

Demostración. Nos centraremos en bases minimales derechas. El otro caso es similar. En primer lugar, veamos que un haz singular arbitrario $H(\lambda)$ tiene una base minimal derecha cuyos vectores de coeficientes son linealmente independientes sobre \mathbb{C} . Si el haz está dado en forma de Kronecker, el resultado es obvio, pues los vectores de la base minimal son de la forma $[\dots 0 \ 1 \ \lambda \ \dots \ \lambda^k \ 0 \ \dots]^T$, cuyos coeficientes son linealmente independientes. Si el haz no está dado en forma de Kronecker, consideramos la transformación $Y^T H(\lambda) X = K_H(\lambda)$ que lo lleva a la forma de Kronecker. Ahora, una base minimal de $H(\lambda)$ está formada por los vectores de la forma $x_i + \lambda x_{i+1} + \dots + \lambda^k x_{i+k}$, donde los vectores de coeficientes x_i, \dots, x_{i+k} son las columnas de la matriz X que corresponden a cada uno de los bloques singulares derechos que aparecen en la forma de Kronecker. Como X es invertible, estos vectores son linealmente independientes.

Ahora veamos que cualquier base minimal tiene vectores coeficientes linealmente independientes. La afirmación es consecuencia inmediata de dos hechos: por un lado, dos bases minimales cualesquiera tienen, una vez ordenadas, el mismo número de vectores (polinómicos) con los mismos grados, lo que significa que tienen el mismo número de vectores de coeficientes. Por otro lado, los vectores de coeficientes de dos bases minimales generan el mismo subespacio, en virtud del Lema 2.5.2. Como, por el párrafo anterior, dicho subespacio tiene dimensión igual al número de vectores coeficientes de la base, se concluye el resultado. ■

Nótese que el resultado anterior no es cierto para una base minimal de un subespacio vectorial cualquiera de $\mathbb{C}^p(\lambda)$. Por ejemplo, el vector $\begin{bmatrix} 1 + \lambda \\ 1 + \lambda^2 \end{bmatrix}$ es una base minimal del subespacio de $\mathbb{C}^2(\lambda)$

$$\mathcal{V} = \left\langle \begin{bmatrix} 1 + \lambda \\ 1 + \lambda^2 \end{bmatrix} \right\rangle$$

y sus vectores de coeficientes $[1 \ 1]^T$, $[1 \ 0]^T$ y $[0 \ 1]^T$ son linealmente dependientes sobre \mathbb{C} .

Finalmente, incluimos una definición que emplearemos más adelante, concretamente en §6.7.1. Para ello, nótese que si particionamos las matrices X e Y que dan paso a la forma de Kronecker (2.5) de la siguiente manera:

$$X = [X_1 \mid X_2 \mid X_3], \quad Y = [Y_1 \mid Y_2 \mid Y_3], \quad (2.15)$$

donde

$$X_1 \in \mathbb{C}^{n \times (\varepsilon+p)}, \quad X_2 \in \mathbb{C}^{n \times \eta}, \quad Y_1 \in \mathbb{C}^{m \times \varepsilon}, \quad Y_2 \in \mathbb{C}^{m \times (\eta+q)}, \quad (2.16)$$

con ε, η como en (2.7), y denotamos

$$\mathcal{X}_i = \text{col}(X_i), \quad \mathcal{Y}_i = \text{col}(Y_i), \quad i = 1, 2,$$

entonces \mathcal{X}_1 es el subespacio reductor minimal de H , mientras que \mathcal{Y}_2 es el subespacio reductor minimal de H^T . En particular, y en virtud del Teorema 2.6.5, \mathcal{X}_1 es una base del espacio singular derecho de $H(\lambda)$ e \mathcal{Y}_2 es una base del espacio singular izquierdo de $H(\lambda)$. Esto nos sugiere la siguiente definición.

Definición 2.6.8 Dado un haz matricial $H(\lambda)$ de tamaño $m \times n$, una **base de Kronecker del espacio singular derecho** de H es un conjunto de $\varepsilon + p$ vectores de \mathbb{C}^n , $x_1, \dots, x_{\varepsilon+p}$, tales que si

$$X_1 = [x_1 \quad \dots \quad x_{\varepsilon+p}]$$

entonces existen dos matrices X, Y como en (2.15) y (2.16) que dan paso a la forma de Kronecker de $H(\lambda)$ como en (2.5).

Análogamente, una **base de Kronecker del espacio singular izquierdo** de H es un conjunto de $\eta + q$ vectores de \mathbb{C}^m , $y_1, \dots, y_{\eta+q}$, tales que si

$$Y_2 = [y_1 \quad \dots \quad y_{\eta+q}]$$

entonces existen dos matrices X, Y como en (2.15) y (2.16) que dan paso a la forma de Kronecker de $H(\lambda)$ como en (2.5).

Obsérvese que no toda base del espacio singular derecho (izquierdo) de un haz matricial singular es una base de Kronecker del espacio singular derecho (izquierdo) del haz.

2.7. Aplicaciones

Una de las aplicaciones más relevantes de la forma canónica de Weierstrass de haces regulares reside en la resolución de sistemas de ecuaciones diferenciales y de ecuaciones en diferencias. Por ejemplo, un sistema lineal como (1.1) puede resolverse utilizando cierta información espectral, que incluye la forma de Weierstrass, del haz asociado (2.1). Dicha información está contenida en una terna (X, J, Y) , donde X e Y son las matrices que dan paso a la forma de Weierstrass (2.2) y J es la matriz de Jordan de H . Si particionamos $J = J_F \oplus J_\infty$, donde J_F contiene los bloques correspondientes a los autovalores finitos y J_∞ los bloques correspondientes al autovalor infinito, y las matrices $X = [X_F \ X_\infty]$ e $Y = [Y_F \ Y_\infty]$ conforme a la partición anterior, y denotamos por ν el menor entero tal que $J_\infty^\nu = 0$, entonces, si $f(t)$ es una función derivable con continuidad en \mathbb{C} hasta el orden ν , la solución general del sistema (1.1) está dada por:

$$x(t) = X_F e^{J_F t} x + \int_{t_0}^t X_F e^{J_F(t-s)} Y_F f(s) ds - \sum_{i=0}^{\nu-1} X_\infty J_\infty^i Y_\infty f^{(i)}(t), \quad (2.17)$$

donde $x \in \mathbb{C}^m$ es un vector arbitrario (y m es la dimensión de J_F) [28, Th. 8.1].

La condición de diferenciabilidad de f está justificada teniendo en cuenta la forma que adopta la solución (2.17). No obstante, en algunos casos, no es necesario imponer un orden de diferenciabilidad sobre f tan elevado [50, Ex. 1]. Esto ocurre, por ejemplo, en el caso en que el término $X_\infty J_\infty^{\nu-1} Y_\infty f^{(\nu-1)}(t)$ en el último sumatorio de (2.17) es nulo. Se plantea, entonces, la siguiente cuestión inversa: suponiendo que existe una solución continua al sistema (1.1), ¿qué orden de diferenciabilidad ha de tener esta solución?. En [50] se resuelve esta cuestión para sistemas de orden superior mediante la obtención de formas canónicas “condensadas” que reducen el sistema a otro más elemental y con las mismas soluciones.

Cuando el sistema lineal

$$A_1 x'(t) + A_0 x(t) = f(t)$$

está definido por un haz singular $A_0 + \lambda A_1$, el propio sistema se denomina **singular**. En esta situación, en la que el sistema puede no tener solución, o incluso puede tener infinitas

soluciones, la parte singular de la FCK de $A_0 + \lambda A_1$ está íntimamente relacionada con la solubilidad del sistema. En concreto [93, p. 249],

- Por cada bloque singular derecho la solución contiene una función arbitraria. Este bloque no impone ninguna restricción en la condición inicial o en las coordenadas de la solución.
- Por cada bloque singular izquierdo, L_η^T , quedan completamente determinadas η variables y no se pueden asignar valores iniciales arbitrarios. Se requiere una condición de compatibilidad entre las entradas de f y sus derivadas, concretamente, entre

$$f_i, f'_{i+1}, f''_{i+2}, \dots, f^{(\eta)}_{i+\eta},$$

donde L_η^T ocupa las filas desde la i -ésima hasta la $(i + \eta)$ -ésima.

El segundo de los terrenos en que son de gran relevancia las formas canónicas que hemos descrito es en el vasto territorio de la Teoría de Control. El caso más elemental de un sistema lineal invariante en el tiempo:

$$\begin{aligned} x'(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \quad ,$$

tiene asociado el haz singular rectangular de rango completo $[A - \lambda I \ B]$. Su FCK es determinante en relación con las propiedades de controlabilidad y observabilidad del sistema. Por ejemplo, como es bien sabido [95, p. 45], el sistema es controlable si y sólo si la FCK del haz asociado no tiene parte regular.

Finalmente, la forma canónica de Jordan ha sido utilizada con frecuencia en el contexto de la Teoría de Sistemas Vibratorios. Concretamente, en [43, 44] se usa la información espectral de un sistema para caracterizar determinadas propiedades de dicho sistema o de las matrices que lo componen, como puede ser la simetría o la positividad, o también para recuperar por completo el sistema. Estas cuestiones forman parte de lo que se conoce, genéricamente, como el *Problema Inverso* de los sistemas vibratorios. Como referencia introductoria (y breve) a la aplicación de la forma de Jordan en sistemas vibratorios puede consultarse [77, §3].

La terna (X, J, Y) que hemos descrito anteriormente se conoce con el nombre de *triple de Jordan* y es, para haces matriciales, un caso particular de lo que se conoce como **triple resolvente** de un polinomio matricial de grado arbitrario [28, p. 219]. Estos triples permiten obtener las soluciones de un sistema lineal de ecuaciones diferenciales ordinarias de orden superior asociado a dicho polinomio. En el apéndice que viene a continuación tratamos brevemente estos sistemas.

Apéndice: Linealizaciones

Consideremos el siguiente sistema lineal de orden ℓ con coeficientes constantes

$$A_\ell x^{(\ell)} + A_{\ell-1} x^{(\ell-1)}(t) + \dots + A_1 x'(t) + A_0 x(t) = f(t), \quad (2.18)$$

donde $x(t)$ es un vector de incógnitas con n coordenadas que son funciones de la variable t , $x^{(k)}(t)$ denota la k -ésima derivada del vector $x(t)$ con respecto a la variable t , A_k son matrices de tamaño $n \times n$ con entradas complejas y $f(t)$ es un vector dado de n coordenadas

que son funciones de t . En la teoría clásica de ecuaciones diferenciales ordinarias, es frecuente la transformación del sistema (2.18) en un sistema de ecuaciones de primer orden mediante la introducción de nuevas variables. La manera de llevar a cabo esta transformación no es única, pero una elección natural es la siguiente:

$$x_0(t) = x(t), \quad x_1(t) = x'(t), \quad \dots, \quad x_{\ell-1}(t) = x^{(\ell-1)}(t),$$

con lo que el sistema (2.18) se transforma en

$$\begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & A_\ell \end{bmatrix} \mathbf{x}'(t) + \begin{bmatrix} 0 & -I & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & -I \\ A_0 & A_1 & \cdots & A_{\ell-1} \end{bmatrix} \mathbf{x}(t) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(t) \end{bmatrix}, \quad (2.19)$$

donde

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{\ell-1}(t) \end{bmatrix},$$

e I es la matriz identidad $n \times n$.

La ecuación (2.18) está asociada al polinomio matricial de grado ℓ

$$L(\lambda) = A_0 + \lambda A_1 + \dots + \lambda^\ell A_\ell, \quad (2.20)$$

donde los coeficientes, A_i , son matrices de tamaño $n \times n$ con entradas complejas, mientras que la ecuación (2.19) tiene asociado el haz

$$C_L(\lambda) = \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & A_\ell \end{bmatrix} \lambda + \begin{bmatrix} 0 & -I & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & -I \\ A_0 & A_1 & \cdots & A_{\ell-1} \end{bmatrix}. \quad (2.21)$$

Dicho haz recibe el nombre de **primer haz compañero** del polinomio (2.20).

La información espectral de dicho haz compañero, contenida en su forma de Weierstrass, permite resolver el sistema (2.18) de un modo similar a (2.17), aunque entran en juego los triples resolventes que hemos mencionado antes.

En general, un haz $A + \lambda B$ de tamaño $n\ell \times n\ell$ es una **linealización** del polinomio (2.20) con coeficientes cuadrados $A_i \in \mathbb{C}^{n \times n}$ si $A + \lambda B$ es equivalente a

$$\begin{bmatrix} L(\lambda) & 0 \\ 0 & I \end{bmatrix}.$$

La relevancia de linealizaciones no se limita al contexto de los sistemas de ecuaciones diferenciales de orden superior. Un procedimiento habitual para resolver el Problema Polinómico de Autovalores (PPE)

$$L(\lambda)x = \sum_{i=1}^{\ell} A_i x = 0,$$

donde $A_i \in \mathbb{C}^{n \times n}$, consiste en linealizar el problema. Existen infinidad de formas de hacerlo debido a que hay infinidad de linealizaciones de un polinomio matricial arbitrario. A la hora de elegir una de ellas hay que tener en cuenta los problemas inherentes al método numérico que se va a emplear para revelar la información espectral del haz. En particular, el condicionamiento y el error backward del problema de autovalores generalizado

$$L(\lambda)x = 0$$

asociado a la linealización L . El problema de encontrar linealizaciones bien condicionadas ha sido tratado en [30] para la familia de linealizaciones introducida en [48], mientras que el error backward de dichas linealizaciones se ha estudiado en [31]. Una buena introducción a estas cuestiones se puede encontrar en [77, §5].

Más en general, las linealizaciones pueden emplearse para obtener toda la información espectral de un polinomio regular, es decir, los divisores elementales y las multiplicidades parciales asociadas al infinito. Todas las linealizaciones tienen los mismos divisores elementales (por tanto, los mismos autovalores finitos con las mismas multiplicidades parciales asociadas) que el polinomio que linealizan. En otras palabras, tienen la misma información espectral finita que dicho polinomio. En cambio, no ocurre lo mismo con las multiplicidades parciales en el infinito ni con los índices minimales (si se trata de polinomios singulares). Las linealizaciones que conservan la estructura en el infinito se conocen con el apelativo de *linealizaciones fuertes* [27], y se caracterizan porque, además de ser linealizaciones del polinomio $L(\lambda)$, su haz dual es equivalente al polinomio

$$\begin{bmatrix} L^\sharp(\lambda) & \\ & I \end{bmatrix},$$

donde $L^\sharp(\lambda) = \lambda^\ell L(1/\lambda)$ es el polinomio dual de $L(\lambda)$.

Por otra parte, las linealizaciones de polinomios singulares han sido escasamente tratadas en la literatura. En concreto, no existe aún una caracterización de las linealizaciones que conservan los índices minimales del polinomio o que, en su defecto, permitan recuperarlos a través de una fórmula sencilla. Actualmente, estamos trabajando en este problema para el caso de polinomios singulares cuadrados, sobre un conjunto de linealizaciones que incluyen las introducidas en [48] y, entre ellas, el primer haz compañero. Hemos demostrado que los índices minimales de estas linealizaciones se obtienen a partir de los índices minimales del polinomio mediante la adición de una cantidad fija [20].

Finalmente, queremos mencionar otra aportación relevante de las linealizaciones al terreno computacional, y que incluye tanto a polinomios singulares como regulares. Se trata del trabajo de P. Van Dooren y P. Dewilde [83], en el que, mediante linealizaciones construidas a partir de matrices habituales en Teoría de Control [86, 87], proponen un algoritmo estable para obtener la estructura espectral de un polinomio matricial arbitrario.

Capítulo 3

El Polígono de Newton

Los autovalores de un haz matricial $H(\lambda) = A_0 + \lambda A_1$ (regular o singular) son las raíces de unos determinados polinomios (el determinante de $H(\lambda)$ en el caso regular o ciertos menores del haz en el caso singular), cuyos coeficientes dependen continuamente de las entradas de A_0 y A_1 . Por tanto, una manera de abordar problemas de perturbación de autovalores consiste en estudiar la perturbación de raíces de polinomios. Este problema, más general, ha sido ampliamente tratado en el ámbito de la Teoría de Funciones Algebraicas. En este contexto, una herramienta fundamental es el Polígono de Newton, que es el objeto fundamental del presente Capítulo. Los resultados que expondremos serán utilizados, en el Capítulo 4, para la obtención de los desarrollos de primer orden de los autovalores de perturbaciones de haces matriciales singulares cuadrados.

3.1. El polígono de Newton

En la presente sección vamos a describir el procedimiento del *Polígono* (o *diagrama*) *de Newton*, que se emplea para obtener desarrollos asintóticos en torno a un punto de una curva algebraica. Nosotros lo usaremos para obtener los términos de primer orden de dichos desarrollos.

Dado un polinomio en dos variables

$$F(\lambda, \epsilon) = \sum_{i,j=0}^n a_{ij} \lambda^i \epsilon^j, \quad (3.1)$$

la *nube de puntos* asociada a (3.1) es el conjunto

$$\mathcal{N}(P) := \{(i, j) : a_{ij} \neq 0\}.$$

Ejemplo 3.1.1 Consideremos el polinomio

$$F(\lambda, \epsilon) = 4\lambda^9 \epsilon^3 + 5\lambda^7 \epsilon^2 + \lambda^6(1 + \epsilon) - \lambda^4 - 5\lambda^3(\epsilon^2 + 2\epsilon^3) + 9\lambda^2(\epsilon^2 - \epsilon^5) + 6\lambda \epsilon^4 + \epsilon^6.$$

La nube de puntos asociada a $P(\lambda, \epsilon)$ está representada en la siguiente figura:

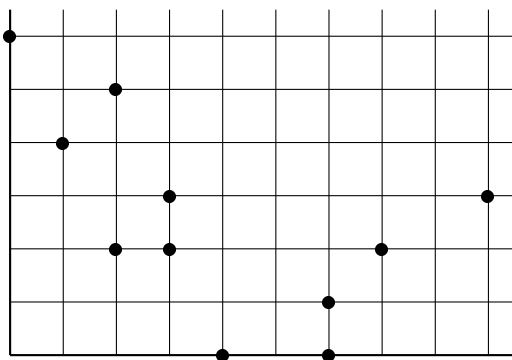


Figura 1: Nube de puntos asociada a $F(\lambda, \epsilon)$

Asociado a la nube de puntos, se define el Polígono de Newton:

Definición 3.1.2 (Polígono de Newton) Dado un polinomio en dos variables, $F(\lambda, \epsilon)$, su Polígono de Newton es el borde inferior de la envolvente convexa de la nube de puntos de $F(\lambda, \epsilon)$.

Ejemplo 3.1.3 El polígono de Newton asociado al polinomio del ejemplo anterior es la poligonal dibujada en la siguiente figura:

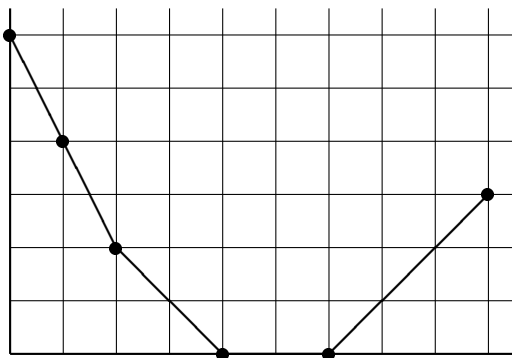


Figura 2: Polígono de Newton asociado a $F(\lambda, \epsilon)$

3.2. Desarrollos en torno a un punto

En este apartado vamos a enunciar los resultados básicos de la teoría local de funciones algebraicas complejas que necesitaremos en el análisis de las perturbaciones de norma pequeña de haces matriciales. Todos los resultados de esta sección son, por tanto, bien conocidos.

Consideraremos la ecuación polinómica

$$F(\lambda, \epsilon) = \sum_{k=0}^n F_k(\epsilon)\lambda^k = 0, \quad (3.2)$$

donde $F_k(\epsilon)$ son polinomios en la variable ϵ , que escribiremos

$$F_k(\epsilon) = F_k\epsilon^{\beta_k} + O(\epsilon^{\beta_k+1}),$$

con $F_k \in \mathbb{C}$ y $F_n(\epsilon)$ no es idénticamente nulo. El resultado fundamental, en torno al que gira el resto de la sección, es el siguiente:

Teorema 3.2.1 ([78, Teorema 5.1]) *La ecuación (3.2) tiene, para cada ϵ fijo, n soluciones, $\lambda_1(\epsilon), \dots, \lambda_n(\epsilon)$, que pueden desarrollarse en serie de potencias*

$$\lambda_i(\epsilon) = a_i \epsilon^{\alpha_i} + o(\epsilon^{\alpha_i}), \quad (3.3)$$

donde α_i es un exponente fraccionario (que puede ser nulo). La serie (3.3) es convergente en un entorno de cualquier valor de λ salvo en $\lambda = 0$ si $\alpha_i < 0$.

Lo que ahora nos interesa es determinar los valores a_i y α_i de los desarrollos (3.3). En particular, cuántos exponentes α_i son negativos. Las dos subsecciones siguientes están dedicadas a ello.

3.2.1. El Polígono de Newton y los desarrollos asintóticos

Para obtener los valores del exponente α_i de (3.3) usaremos el Polígono de Newton (PN) asociado al polinomio $F(\lambda, \epsilon)$ que define la variedad (3.2) en la forma que explicaremos a continuación.

Consideremos uno de los desarrollos anteriores. Por simplicidad, eliminamos los subíndices:

$$\lambda(\epsilon) = a \epsilon^\alpha + \dots \quad (3.4)$$

Al sustituir (3.4) en la ecuación (3.2), cada término $F_k(\epsilon)\lambda^k$ produce un término en ϵ de orden $\epsilon^{\alpha k + \beta_k}$ y términos de orden superior. Para que (3.4) sea una raíz de (3.2) los términos del mismo grado (en particular, los de menor grado) deben cancelarse. Esto significa que todos los coeficientes de los términos con el mismo exponente, $\alpha k + \beta_k$, han de sumar 0. En particular, tendrá que haber, al menos, dos valores $j, k \in \{0, 1, \dots, n\}$ con

$$\alpha k + \beta_k = \alpha j + \beta_j, \quad \alpha k + \beta_k \leq \alpha s + t,$$

para cualquier punto (s, t) de la nube asociada a $F(\lambda, \epsilon)$. Si consideramos los correspondientes puntos $P_j = (j, \beta_j)$ y $P_k = (k, \beta_k)$ de la nube de $F(\lambda, \epsilon)$, tendremos:

A1. P_j y P_k forman parte del Polígono de Newton asociado a $F(\lambda, \epsilon)$.

A2. P_j y P_k están en un segmento de pendiente

$$-\alpha = \frac{\beta_j - \beta_k}{j - k}.$$

La afirmación **A1** es consecuencia de la desigualdad anterior:

$$\alpha k + \beta_k \leq \alpha s + t,$$

pues cualquier punto (s, t) bajo el segmento que une P_j y P_k satisface

$$\alpha s + t < \alpha k + \beta_k.$$

Por otra parte, invirtiendo el argumento, cualesquiera dos puntos que formen parte del Polígono de Newton asociado a $F(\lambda, \epsilon)$ nos darán un valor α para el exponente de (3.4),

que es el opuesto de la pendiente del segmento que los une. Por tanto, **los valores de los exponentes α en (3.4) son los opuestos de las pendientes de los segmentos que forman el Polígono de Newton asociado a $F(\lambda, \epsilon)$** . Es importante hacer constar que algunos de estos exponentes serán negativos (los que corresponden a pendientes positivas). A estos exponentes corresponden las ramas que se van a infinito, y que trataremos en la Sección 3.2.2.

Para obtener el coeficiente director, a , de los desarrollos (3.4), hemos de considerar todos los términos que corresponden a puntos del Polígono de Newton de $F(\lambda, \epsilon)$ pertenecientes a un mismo segmento, \mathcal{S} , de pendiente $-\alpha$. Denotemos por $\mathcal{I}_{\mathcal{S}}$ el conjunto de índices, k , para los cuales el punto $P_k = (k, \beta_k)$ está en \mathcal{S} , es decir:

$$\mathcal{I}_{\mathcal{S}} = \{k : (k, \beta_k) \in \mathcal{S}\}.$$

Estos puntos corresponden a términos del mismo orden de ϵ en $F(\lambda(\epsilon), \epsilon)$. Igualando a 0 la suma de los coeficientes correspondientes obtenemos:

$$\sum_{k \in \mathcal{I}_{\mathcal{S}}} a^k F_k = 0. \quad (3.5)$$

Las soluciones de esta ecuación son los valores de los coeficientes, a , de los desarrollos (3.4). Si denotamos

$$\begin{aligned} k_{max} &= \max \mathcal{I}_{\mathcal{S}} \\ k_{min} &= \min \mathcal{I}_{\mathcal{S}} \end{aligned}$$

la ecuación (3.5) tiene k_{min} soluciones nulas y un total de $k_{max} - k_{min}$ soluciones no nulas, que corresponden a desarrollos de orden α . Tenemos, por tanto, el:

Teorema 3.2.2 *El número de desarrollos de orden α de la ecuación $F(\lambda, \epsilon) = 0$ coincide con la longitud de la proyección horizontal del segmento \mathcal{S} del polígono de Newton de $F(\lambda, \epsilon)$ con pendiente $-\alpha$.*

El procedimiento para determinar el exponente y el coeficiente director de los desarrollos (3.4) es, por lo tanto, el siguiente [53]:

Procedimiento del Polígono de Newton: Dada una ecuación polinómica (3.2) de grado n en λ :

1. Se traza el Polígono de Newton de $F(\lambda, \epsilon)$.
2. Los exponentes directores, α , de los n desarrollos (3.4) son los opuestos de las pendientes de los segmentos que forman el Polígono de Newton.
3. El número de desarrollos de orden α es la longitud de la proyección sobre el eje horizontal del segmento, \mathcal{S} , de pendiente $-\alpha$.
4. Los coeficientes directores, a , de los desarrollos de orden ϵ^α son las raíces no nulas de la ecuación (3.5).

Hemos de advertir que, debido a la expresión (3.4), el procedimiento anterior sólo nos permite obtener los términos de primer orden de los desarrollos en torno al valor $\lambda = 0$. Si $F(0, 0) \neq 0$, el punto $(0, 0)$ forma parte de la nube de puntos de $F(\lambda, \epsilon)$ y el

Polígono de Newton no contiene ningún segmento de pendiente negativa. En este caso no hay desarrollos en torno a $\lambda = 0$. Por otro lado, si $\lambda_0 \in \mathbb{C}$ es un valor para el cual

$$F(\lambda_0, 0) = 0$$

(es decir: $(\lambda_0, 0)$ es un punto de la curva algebraica definida por la ecuación (3.2)), existirán desarrollos en torno a λ_0 de la forma:

$$\lambda(\epsilon) = \lambda_0 + a\epsilon^\alpha + \dots$$

En este caso los valores a , α se pueden obtener expresando $F(\lambda, \epsilon)$ en torno a λ_0 :

$$F(\lambda, \epsilon) = \sum_{k=0}^n \bar{F}_k(\epsilon)(\lambda - \lambda_0)^k = \sum_{k=0}^n \bar{F}_k(\epsilon)\mu^k = \bar{F}(\mu, \epsilon),$$

y aplicando el procedimiento del Polígono de Newton al polinomio $\bar{F}(\mu, \epsilon)$, donde $\mu = \lambda - \lambda_0$. Nótese que el coeficiente, $\bar{F}_k(\epsilon)$, de μ^k en $\bar{F}(\mu, \epsilon)$ coincide con el de λ^k en $F(\lambda + \lambda_0, \epsilon)$.

Ejemplo 3.2.3 Consideremos la ecuación

$$F(\lambda, \epsilon) = \lambda^2(1 + \epsilon) - 2\lambda(1 + 2\epsilon) + 1 + \epsilon^2 = 0.$$

La nube de puntos de $F(\lambda, \epsilon)$ contiene al punto $(0, 0)$. En cambio, si desarrollamos el polinomio en torno a $\lambda - 1$ tendremos:

$$F(\lambda, \epsilon) = (\lambda - 1)^2(1 + \epsilon) + (\lambda - 1)(-2\epsilon) - 3\epsilon + \epsilon^2 = \mu^2(1 + \epsilon) + \mu(-2\epsilon) - 3\epsilon + \epsilon^2 = \bar{F}(\mu, \epsilon).$$

El Polígono de Newton asociado a $\bar{F}(\mu, \epsilon)$ consta de un solo segmento que une $(0, 1)$ y $(2, 0)$:

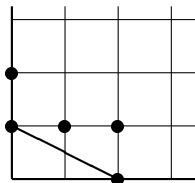


Figura 3: Nube de puntos y Polígono de Newton asociado a $\bar{F}(\mu, \epsilon)$

lo que significa que hay dos desarrollos en torno a $\lambda = 1$ de orden $\epsilon^{1/2}$

$$\begin{aligned} \lambda_1(\epsilon) &= 1 + a_1 \epsilon^{1/2} + o(\epsilon^{1/2}) \\ \lambda_2(\epsilon) &= 1 + a_2 \epsilon^{1/2} + o(\epsilon^{1/2}), \end{aligned}$$

y cuyos coeficientes directores son las soluciones de la ecuación:

$$a^2 - 3 = 0,$$

es decir: $a_1 = \sqrt{3}$, $a_2 = -\sqrt{3}$. □

3.2.2. Puntos críticos

Si λ_0 es una raíz de multiplicidad m de la ecuación

$$F(\lambda, 0) = 0, \quad (3.6)$$

entonces habrá un total de m desarrollos (3.3) del total de n mencionados en el Teorema 3.2.1 que convergen a λ_0 cuando ϵ tiende a 0. En este caso, $\epsilon = 0$ es un **punto crítico de segunda especie** de (3.2) [32, p. 93]. Los desarrollos en torno a λ_0 se agrupan en **ciclos** [32, p. 99] de k elementos, $\lambda_1(\epsilon), \dots, \lambda_k(\epsilon)$, de la forma

$$\lambda_i(\epsilon) = \lambda_0 + \sum_{j=1}^{\infty} a_{ij} \epsilon^{j/k}, \quad i = 1, \dots, k,$$

uno por cada determinación de $\epsilon^{1/k}$. Además, habrá un número ν de ciclos que constan de k_1, \dots, k_ν elementos, respectivamente, de manera que

$$k_1 + \dots + k_\nu = m.$$

Por otra parte, tal y como se indica en el enunciado del Teorema 3.2.1, algunos de los desarrollos (3.3) pueden tener exponente α_i negativo, lo que significa que la serie (3.3) tiende a infinito cuando ϵ se aproxima a 0. En este caso, el valor $\epsilon = 0$ es un **punto crítico de primera especie** de (3.2) [32, p. 93]. Eso significa que $\epsilon = 0$ es una raíz de la ecuación:

$$F_n(\epsilon) = 0.$$

En general, si

$$F_n(0) = 0, \quad F_{n-1}(0) = 0, \quad \dots, \quad F_{n-\kappa+1}(0) = 0, \quad F_{n-\kappa}(0) \neq 0$$

el punto crítico $\epsilon = 0$ tiene **multiplicidad** κ (si $F_j(0) = 0$ para todo $j = 0, 1, \dots, n$, diremos que tiene multiplicidad ∞).

Lema 3.2.4 *Si $\epsilon = 0$ es un punto crítico de primera especie de (3.2) con multiplicidad $\kappa \leq n$, hay exactamente κ desarrollos (3.3) con exponente α_i negativo.*

Demostración:

Como, por hipótesis, el punto crítico $\epsilon = 0$ no tiene multiplicidad ∞ , podemos escribir

$$F(\lambda, 0) = f_{n-\kappa} \lambda^{n-\kappa} + \dots + f_1 \lambda + f_0,$$

con $f_i \in \mathbb{C}$, para $i = 0, \dots, n-\kappa$, y $f_{n-\kappa} \neq 0$. En esta situación, los puntos $(n-\kappa, 0)$ y (n, s) (para algún s positivo) forman parte del Polígono de Newton de $F(\lambda, \epsilon)$. Además, no hay ningún punto de dicho Polígono situado a la derecha de $(n-\kappa, 0)$ sobre el eje horizontal. Esto significa que la proyección sobre el eje horizontal de la parte del Polígono de Newton con pendiente positiva tiene longitud κ . El resultado es consecuencia del Teorema 3.2.2. ■

Los κ desarrollos del Lema 3.2.4, junto con los desarrollos que se obtienen para cada raíz de la ecuación (3.6) conforman el total de n desarrollos mencionados en el Teorema 3.2.1. Estas observaciones están incluidas en el siguiente Teorema, a modo de resumen de la presente sección.

Teorema 3.2.5 *Sea el polinomio*

$$F(\lambda, \epsilon) = \sum_{k=0}^n F_k(\epsilon) \lambda^k,$$

donde $F_k(\epsilon)$ es un polinomio en ϵ , para todo $k = 0, \dots, n$, y $F_n(\epsilon)$ no es idénticamente nulo. La ecuación polinómica

$$F(\lambda, \epsilon) = 0$$

tiene exactamente n soluciones, $\lambda_1(\epsilon), \dots, \lambda_n(\epsilon)$, para cada ϵ , que pueden desarrollarse en series de potencias y satisfacen:

- i) Si $\epsilon = 0$ es un punto crítico de primera especie de multiplicidad κ , hay exactamente κ soluciones que tienden a infinito cuando ϵ tiende a 0.
- ii) Si $F(\lambda, 0) = \prod_{i=1}^q (\lambda - \lambda_i)^{a_i}$, entonces hay exactamente a_i soluciones que tienden a λ_i cuando ϵ tiende a 0.

En ambos casos, los coeficientes y los exponentes directores de $\lambda_1(\epsilon), \dots, \lambda_n(\epsilon)$ se obtienen a través del procedimiento del Polígono de Newton.

Demostración: *i)* es consecuencia del Lema 3.2.4.

Para demostrar *ii)*, podemos suponer, sin pérdida de generalidad, que $F(\lambda, \epsilon)$ es irreducible pues, en caso contrario, razonamos de igual forma con cada uno de sus factores irreducibles. El hecho de que $F(\lambda, \epsilon)$ es irreducible implica, en particular, que el Polígono de Newton de $\widehat{F}(\lambda, \epsilon) = F(\lambda + \lambda_i, \epsilon)$ contiene un punto del eje vertical y también el punto $(a_i, 0)$, para cualquier $i = 1, \dots, q$. Esto significa que la proyección sobre el eje horizontal del Polígono de Newton de $\widehat{F}(\lambda, \epsilon)$ tiene longitud a_i . En consecuencia, hay exactamente a_i desarrollos en torno a λ_i .

Finalmente, nótese que, por definición de multiplicidad del punto crítico de primera especie, se tiene

$$\kappa + a_1 + \dots + a_q = n,$$

es decir: los desarrollos de *i)* y *ii)* son los n desarrollos del Teorema 3.2.1. ■

Ejemplo 3.2.6 El polinomio del Ejemplo 3.1.1

$$F(\lambda, \epsilon) = 4\lambda^9 \epsilon^3 + 5\lambda^7 \epsilon^2 + \lambda^6(1 + \epsilon) - \lambda^4 - 5\lambda^3(\epsilon^2 + 2\epsilon^3) + 9\lambda^2(\epsilon^2 - \epsilon^5) + 6\lambda\epsilon^4 + \epsilon^6,$$

satisface

$$F(\lambda, 0) = \lambda^4(\lambda^2 - 1).$$

Por tanto, habrá

- Cuatro desarrollos en torno a 0,
- un desarrollo en torno a 1,
- un desarrollo en torno a -1 ,
- tres desarrollos en torno a infinito.

Por el Polígono de Newton de $F(\lambda, \epsilon)$ sabemos que los desarrollos en torno a 0 tienen grados respectivos 2, 2, 1 y 1. Además, los coeficientes directores de los dos primeros son las soluciones, a , de la ecuación

$$9a^2 + 6a + 1 = 0,$$

mientras que los de los dos últimos son las soluciones de

$$-a^2 + 9 = 0.$$

Por lo tanto:

$$\lambda_1(\epsilon) = -\frac{1}{3}\epsilon^2 + o(\epsilon^2)$$

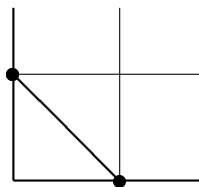
$$\lambda_2(\epsilon) = -\frac{1}{3}\epsilon^2 + o(\epsilon^2)$$

$$\lambda_3(\epsilon) = 3\epsilon + o(\epsilon)$$

$$\lambda_4(\epsilon) = -3\epsilon + o(\epsilon).$$

Para hallar el orden del desarrollo en torno a 1 consideramos el polinomio

$\widehat{F}(\lambda, \epsilon) = F(\lambda + 1, \epsilon) = 4\lambda^9\epsilon^3 + 36\lambda^8\epsilon^3 + \lambda^7(5\epsilon^2 + 144\epsilon^3) + \lambda^6(1 + \epsilon + 35\epsilon^2 + 336\epsilon^3) + \lambda^5(6 + 6\epsilon + 105\epsilon^2 + 504\epsilon^3) + \lambda^4(14 + 15\epsilon + 175\epsilon^2 + 504\epsilon^3) + \lambda^3(16 + 20\epsilon + 170\epsilon^2 + 326\epsilon^3) + \lambda^2(9 + 15\epsilon + 99\epsilon^2 + 114\epsilon^3 - 9\epsilon^5) + \lambda(2 + 6\epsilon + 38\epsilon^2 + 6\epsilon^3 + 6\epsilon^4 - 18\epsilon^5) + \epsilon + 9\epsilon^2 - 6\epsilon^3 + 6\epsilon^4 - 9\epsilon^5 + \epsilon^6$ y su Polígono de Newton:



lo que indica que hay un desarrollo en torno a 1 de orden 1 cuyo coeficiente director es la raíz de la ecuación:

$$2a + 1 = 0.$$

Así pues:

$$\lambda_5(\epsilon) = 1 - \frac{1}{2}\epsilon + o(\epsilon).$$

Procediendo de igual forma con la raíz -1 , llegaremos a un desarrollo:

$$\lambda_6(\epsilon) = -1 + \frac{1}{2}\epsilon + o(\epsilon).$$

Por último, hay un punto crítico de multiplicidad 3 en $\epsilon = 0$, lo que significa que hay tres desarrollos que se van a infinito. Dichos puntos corresponden a un segmento de pendiente 1. Por otra parte, los coeficientes directores son las raíces de la ecuación

$$8a^3 + 1 = 0.$$

Por lo tanto, los tres desarrollos que tienden a infinito son de la forma:

$$\lambda_7(\epsilon) = -\frac{1}{\sqrt[3]{2}}\epsilon^{-1} + o(\epsilon^{-1})$$

$$\lambda_8(\epsilon) = -\frac{\zeta}{\sqrt[3]{2}}\epsilon^{-1} + o(\epsilon^{-1}) ,$$

$$\lambda_9(\epsilon) = -\frac{\zeta^2}{\sqrt[3]{2}}\epsilon^{-1} + o(\epsilon^{-1})$$

donde ζ es una raíz cúbica compleja de la unidad. □

Aunque nosotros hemos tratado únicamente el caso en que la función (3.1) es un polinomio en λ y ϵ , es frecuente encontrar en la literatura el caso, más general, en que es una serie analítica en ambas variables [46]:

$$\sum_{k=0}^{\infty} F_k(\epsilon)\lambda^k .$$

En el siguiente Capítulo, nos encontraremos con un caso particular del anterior. Concretamente, con una expresión de la forma

$$\tilde{F}(\lambda, \epsilon) = \sum_{k=0}^n \tilde{f}_k(\lambda)\epsilon^k ,$$

con $\tilde{F}(0, 0) = 0$, donde $\tilde{f}_k(\lambda) = f_k(\lambda)\delta^{-1}(\lambda)$, siendo $f_k(\lambda)$, para $k = 0, \dots, n$, un polinomio en λ , y $\delta(\lambda)$ un polinomio en λ tal que $\delta(0) \neq 0$. Si denotamos $F(\lambda, \epsilon) = \sum_{k=0}^n f_k(\lambda)\epsilon^k$, es evidente que $F(\lambda, \epsilon)$ puede expresarse en la forma $\sum_{k=0}^m F_k(\epsilon)\lambda^k$, para ciertos polinomios en ϵ , $F_k(\epsilon)$, y, además, si $\lambda(\epsilon)$ es una función continua de ϵ con $\lambda(0) = 0$, puesto que $\delta(0) \neq 0$, en un entorno de $\epsilon = 0$ se tiene que

$$F(\lambda(\epsilon), \epsilon) \equiv 0 \Leftrightarrow \tilde{F}(\lambda(\epsilon), \epsilon) \equiv 0 .$$

Por otro lado, el Polígono de Newton de $\tilde{F}(\lambda, \epsilon)$ se puede definir de igual manera a como se ha hecho para $F(\lambda, \epsilon)$, es decir: como el borde inferior de la envolvente convexa de la nube de puntos de $\tilde{F}(\lambda, \epsilon)$, donde un punto (α, β) forma parte de la nube si y sólo si $\lambda^\alpha \epsilon^\beta$ tiene coeficiente no nulo en $\tilde{F}(\lambda, \epsilon)$. A partir de la definición de $\tilde{F}(\lambda, \epsilon)$, es evidente que las partes izquierdas (es decir: las de pendiente negativa) de los Polígonos de Newton de $\tilde{F}(\lambda, \epsilon)$ y de $F(\lambda, \epsilon)$ coinciden. Esto significa que los desarrollos, $\lambda(\epsilon)$, en torno a $\epsilon = 0$ de la ecuación $\tilde{F}(\lambda, \epsilon) = 0$ para los que $\lambda(0) = 0$ coinciden con los desarrollos en torno a $\epsilon = 0$ de la ecuación polinómica $F(\lambda, \epsilon) = 0$ y se pueden obtener usando el Polígono de Newton. Así pues, no habrá inconveniente en aplicar la técnica del Polígono de Newton para funciones del tipo $\tilde{F}(\lambda, \epsilon)$.

Capítulo 4

Teoría espectral de primer orden en perturbaciones de haces singulares cuadrados

4.1. Introducción

Es bien sabido que la mayoría de los haces matriciales singulares, cuadrados o rectangulares, no tienen autovalores (véase [11, Section 7] y también la Sección 7.3 de esta memoria). A pesar de ello, cuando existen, los autovalores de haces matriciales singulares tienen un papel importante en un buen número de aplicaciones como, por ejemplo, las ecuaciones algebraico-diferenciales [93] y la teoría de control [80]. En particular, los autovalores de ciertos haces singulares son los modos no controlables y no observables de los sistemas lineales invariantes en el tiempo [13].

En [94] se puso de manifiesto que los autovalores de los haces matriciales singulares son funciones discontinuas de las entradas de las matrices de coeficientes. Por ejemplo, el haz

$$A_0 + \lambda A_1 = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}$$

tiene un único autovalor igual a 0. En cambio, el haz perturbado

$$\widehat{A}_0 + \lambda \widehat{A}_1 = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} + \epsilon \left(\begin{bmatrix} 6 & -3 \\ -10 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) = \begin{bmatrix} \lambda + 6\epsilon & \epsilon(\lambda - 3) \\ \epsilon(\lambda - 10) & 0 \end{bmatrix}, \quad (4.1)$$

satisface $\det(\widehat{A}_0 + \lambda \widehat{A}_1) = -\epsilon^2(\lambda - 3)(\lambda - 10)$, y, por lo tanto, es regular y tiene dos autovalores, 3 y 10, para cualquier $\epsilon \neq 0$. Nótese que si el ejemplo anterior se modifica reemplazando -3 y -10 por cualquier par de números $-a$ y $-b$, entonces los autovalores de $\widehat{A}_0 + \lambda \widehat{A}_1$ son a y b . Así pues, perturbaciones arbitrariamente pequeñas pueden situar los autovalores en cualquier punto del plano complejo. La situación es incluso peor en el caso de haces rectangulares. Por ejemplo, el haz

$$A_0 + \lambda A_1 = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tiene de nuevo un solo autovalor $\lambda_0 = 0$, pero el haz perturbado

$$\widehat{A}_0 + \lambda \widehat{A}_1 = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \epsilon \begin{bmatrix} 1 & 2 + \lambda & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad (4.2)$$

no tiene autovalores para $\epsilon \neq 1/2$ (y ϵ no nulo), ya que, en este caso, $\text{rg}(\widehat{A}_0 + \lambda_0 \widehat{A}_1) = \text{rg}(\widehat{A}_0 + \lambda \widehat{A}_1) = 2$ para cualquier número complejo λ_0 (incluido el infinito).

Los anteriores ejemplos muestran que existen perturbaciones arbitrariamente pequeñas que pueden cambiar o destruir completamente los autovalores y la estructura espectral de un haz singular. Esto significa que no podemos esperar una teoría de perturbación de autovalores razonable para perturbaciones arbitrarias de un haz singular, y que, por tanto, habremos de restringir el conjunto de perturbaciones admisibles antes de desarrollar tal teoría. En este contexto, los casos de haces cuadrados y rectangulares son muy diferentes entre sí, porque, dado un haz singular cuadrado $A_0 + \lambda A_1$, casi todas las perturbaciones pequeñas convierten al haz perturbado en regular y, además, algunos de los autovalores del haz perturbado están muy próximos a los autovalores de $A_0 + \lambda A_1$, en el caso de que este haz tenga autovalores. Este hecho fue observado en [94]. Por lo tanto, para un haz cuadrado singular que tiene autovalores, es esperable que se pueda desarrollar una teoría de perturbación de autovalores *para casi todas las perturbaciones pequeñas*. La situación en el caso de haces rectangulares es justamente la contraria, porque, dado un haz rectangular, casi todas las perturbaciones pequeñas producen un haz que no tiene autovalores. La razón es que, genéricamente, los haces rectangulares no tienen autovalores [11, Cor. 7.1]. Por lo tanto, una teoría de perturbaciones para un haz rectangular que tiene autovalores es posible sólo para perturbaciones muy especiales que forman parte de una variedad particular en el conjunto de haces matriciales. Como consecuencia de la anterior discusión, el estudio de la variación de los autovalores de un haz singular para casi todas las perturbaciones pequeñas sólo tiene sentido si los haces son cuadrados.

El principal objetivo de este capítulo es, dado un haz matricial complejo cuadrado singular $H(\lambda) = A_0 + \lambda A_1$ que tiene autovalores, encontrar condiciones suficientes sobre el haz $M(\lambda) = B_0 + \lambda B_1$ que permitan desarrollar una teoría de perturbación de autovalores de primer orden para los autovalores de

$$H(\lambda) + \epsilon M(\lambda), \quad (4.3)$$

en términos del parámetro ϵ , y, asimismo, desarrollar esta teoría de perturbación de primer orden. Estas condiciones suficientes sobre $M(\lambda) = B_0 + \lambda B_1$ implicarán que el haz (4.3) es regular para todo $\epsilon \neq 0$ suficientemente pequeño, y son *genéricas*, es decir, se cumplen para todos los haces salvo aquellos que se encuentran en una variedad algebraica de codimensión mayor que cero. Esto implica que se cumplen para todos los haces excepto aquellos que están en una subvariedad propia en el conjunto de haces matriciales. Bajo estas condiciones genéricas, obtendremos desarrollos de primer orden para aquellos autovalores de (4.3) cuyo límite, cuando ϵ tiende a cero, es un autovalor del haz no perturbado $H(\lambda)$. Haremos esto tanto para autovalores simples como para autovalores múltiples. Los resultados que presentamos se encuentran en [19] y, por lo que sabemos, es la primera vez que se obtienen desarrollos de primer orden de autovalores para perturbaciones de haces matriciales singulares. Merece la pena destacar que estos desarrollos siguen siendo válidos cuando $H(\lambda)$ es regular, y los que obtendremos en este caso para autovalores múltiples en términos de la forma canónica de Weierstrass son también novedosos.

Con mayor precisión, sea λ_0 un autovalor finito de $H(\lambda)$ con divisores elementales $(\lambda - \lambda_0)^{m_1}, \dots, (\lambda - \lambda_0)^{m_g}$, o, equivalentemente, con bloques de Jordan de dimensiones m_1, \dots, m_g en la forma canónica de Kronecker de $H(\lambda)$. En esta situación, demostraremos que, genéricamente, hay $m_1 + \dots + m_g$ autovalores de $H(\lambda) + \epsilon M(\lambda)$ con desarrollos

$$\lambda(\epsilon) = \lambda_0 + c \epsilon^{1/p} + o(\epsilon^{1/p}), \quad (4.4)$$

y que, para m_1 de estos desarrollos se tiene que $p = m_1$, para m_2 de estos desarrollos $p = m_2$, y así sucesivamente hasta llegar a que $p = m_g$ para otros m_g desarrollos. Además, encontraremos expresiones explícitas para los coeficientes directores, c , de estos desarrollos (4.4). Nótese que los exponentes genéricos de los desarrollos están determinados por los grados de los divisores elementales de λ_0 del mismo modo que en el caso regular [46]. En particular, si el autovalor λ_0 es simple, entonces $p = 1$ y se puede escribir $\lambda(\epsilon) = \lambda_0 + c\epsilon + O(\epsilon^2)$, porque en este caso $\lambda(\epsilon)$ es una serie de potencias en ϵ , convergente en un entorno de $\epsilon = 0$. Todas las series en (4.4) son convergentes para ϵ suficientemente pequeño, y cuando contienen exponentes fraccionarios son llamadas *desarrollos de Puiseux*.

Demostremos que los coeficientes, c , de los desarrollos (4.4) están determinados por $M(\lambda_0)$ y por ciertas bases de los espacios nulos derecho e izquierdo de la matriz $H(\lambda_0)$. En el caso de autovalores múltiples estas bases tienen que estar cuidadosamente seleccionadas y normalizadas de un modo no trivial. Esta dificultad no está relacionada con el hecho de que $H(\lambda)$ es singular, y aparece también en la teoría de perturbaciones de autovalores múltiples de matrices y de haces regulares [90, 47, 46, 51]. En cambio, en el caso más frecuente de que λ_0 sea un autovalor simple, la normalización no es necesaria, cualesquiera bases pueden ser utilizadas, y los resultados adoptan una forma más sencilla: si denotamos por W (respectivamente, Z) una matriz cuyas columnas forman una base del espacio nulo izquierdo (resp., derecho) de $H(\lambda_0)$, entonces el haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es genéricamente regular y tiene sólo un autovalor finito. Si este autovalor se denota por ξ , existe un único autovalor de $H(\lambda) + \epsilon M(\lambda)$ tal que

$$\lambda(\epsilon) = \lambda_0 + \xi \epsilon + O(\epsilon^2),$$

cuando ϵ tiende a cero. Debe señalarse que en el caso simple las condiciones genéricas son, precisamente, que el haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es regular y tiene sólo un autovalor finito. Si $H(\lambda)$ es regular entonces $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es 1×1 , y es regular con sólo un autovalor finito para cualquier perturbación $M(\lambda)$. Así pues, en el caso regular,

$$\xi = -\frac{W^T M(\lambda_0) Z}{W^T A_1 Z},$$

lo que permite recuperar un resultado bien conocido (véase, por ejemplo, [69, Th. VI.2.2]).

No es posible desarrollar una teoría *genérica* de perturbaciones de autovectores, porque los autovectores no están definidos para haces singulares, ni siquiera para autovalores simples (véase la Introducción a esta memoria). El concepto correcto que usaremos en haces singulares es el de *subespacio reductor* [81] que hemos discutido en la Sección 2.6. Teniendo en cuenta que el haz perturbado (4.3) es genéricamente regular, no tiene subespacios reductores y, por tanto, tampoco es posible una teoría *genérica* de perturbaciones para subespacios reductores. En cambio, cuando (4.3) es regular, sus autovectores están perfectamente definidos, y es natural preguntarse cómo están relacionados estos autovectores con determinadas propiedades del haz no perturbado $H(\lambda)$ cuando ϵ tiende a cero. Contestaremos a esta pregunta describiendo el vector al que se aproximan dichos autovectores.

Como afirman Stewart y Sun [69, Preface], el análisis de perturbaciones puede dar lugar a desarrollos o a cotas de dichas perturbaciones. Como ya hemos mencionado, el trabajo que presentamos en este capítulo se centra en lo primero, y está en la línea de una serie de trabajos previos, en los que se han tratado tanto perturbaciones uniparamétricas como perturbaciones multiparamétricas. Por un lado, la teoría de perturbaciones uniparamétricas de matrices (o, más en general, operadores) ha sido tratada en profundidad

en [4, 38, 58]. El término de primer orden de los desarrollos de los autovalores fue descrito por Lancaster [40] para autovalores simples y por Lidskii [47] para autovalores múltiples (en particular, defectivos) a partir del trabajo de Vishik y Lyusternik [90]. Los resultados de este último trabajo fueron posteriormente reformulados en [51] utilizando la técnica del *diagrama de Newton*. Por otra parte, las perturbaciones de haces matriciales regulares han sido tratadas en una serie de artículos posteriores. Stewart obtuvo en [67] el término de primer orden para perturbaciones de autovalores simples. El problema, más general, de perturbaciones de matrices analíticas (regulares) ha sido tratado por Langer y Najman en [45] y [46], donde se describe el término de primer orden de los desarrollos de autovalores múltiples (en particular, defectivos). En cuanto a las perturbaciones multiparamétricas, el caso de autovalores semisimples ha sido el más tratado. En este caso, los desarrollos son analíticos, y el problema de primer orden equivale al problema de hallar las derivadas (parciales) de los autovalores. Sun ha estudiado las perturbaciones de matrices en [72] y de haces matriciales en [74]. También se ha estudiado el caso semisimple para perturbaciones de haces matriciales en [96] y de matrices analíticas en [42]. Por último, el caso de autovalores arbitrarios (defectivos o no) ha sido tratado por Chu en [7] para perturbaciones de matrices.

La teoría de perturbación de autovalores de haces singulares se ha estudiado en una serie de trabajos previos. Sun [71, 73] considera haces singulares $A_0 + \lambda A_1$ de tamaño $n \times n$ que son estrictamente equivalentes a haces diagonales y de modo que $\text{rg}(A_0 + \lambda A_1) = n - 1$, y obtiene cotas finitas de perturbaciones de tipo Gerschgorin, Hoffman-Wielandt y Bauer-Fike en sentido probabilístico, es decir, suponiendo que los haces perturbadores satisfacen cierta distribución aleatoria. Así, los haces perturbadores se pueden considerar genéricos. Comparados con los resultados de [71, 73], los desarrollos que presentamos en esta memoria no presuponen ninguna estructura especial sobre el haz de partida, y no son de naturaleza probabilística, sino que son, como ya hemos mencionado, resultados de primer orden.

Demmel y Kågström [12] han estudiado permutaciones *no genéricas* muy específicas de haces singulares cuadrados y rectangulares, y han obtenido cotas para la variación de los autovalores y los subespacios reductores. Estas perturbaciones particulares son muy útiles para acotar los errores en los algoritmos que calculan la forma generalizada de Schur (GUPTRI) de haces singulares [79, 14, 15]. Finalmente, Stewart [68], considera sólo haces rectangulares y perturbaciones *no genéricas* específicas que pueden aparecer en la práctica.

Un planteamiento común en la teoría de haces singulares [12, 68, 71, 73] es que el problema original se reduce a un problema de perturbación de autovalores de una haz regular usando el hecho de que las perturbaciones que se consideran tienen propiedades específicas. Nosotros seguiremos también este enfoque utilizando la forma canónica local de Smith de polinomios matriciales [26, 28] para transformar el problema original de perturbación de un haz singular en un problema regular de perturbaciones. Una vez hecho esto, aplicaremos la teoría de perturbaciones para problemas regulares desarrollada en [46]. Además, llevaremos a cabo un considerable trabajo algebraico para presentar los desarrollos de perturbaciones en términos de magnitudes espectrales intrínsecas de los haces singulares, esto es, espacios nulos asociados con autovalores, espacios reductores y la forma canónica de Kronecker.

El capítulo está organizado de la siguiente manera: en la Sección 4.2 recuperamos, de secciones anteriores, la notación que vamos a utilizar y la adaptamos a las necesidades de este capítulo. Asimismo, identificamos las bases de los espacios nulos de $H(\lambda_0)$ que contienen la información necesaria para obtener los desarrollos de Puiseux (4.4). En la

Sección 4.3, presentamos condiciones de genericidad suficientes para la existencia de una teoría de primer orden de perturbación de autovalores de haces singulares, y mostramos cómo transformar este problema de perturbaciones en uno regular. En la Sección 4.4, establecemos una conexión entre la forma local de Smith y la forma de Kronecker de haces. La Sección 4.5 presenta los desarrollos de autovalores que hemos anunciado: el Teorema 4.5.3 para autovalores finitos múltiples, el Corolario 4.5.5 para el autovalor infinito, y en el Teorema 4.5.6 el resultado, libre de normalizaciones, para los autovalores simples. Finalmente, en la Sección 4.6 estudiamos los autovectores del haz perturbado.

4.2. Preliminares

Dado un autovalor finito λ_0 del haz singular $H(\lambda) = A_0 + \lambda A_1$ de tamaño $n \times n$, los subespacios nulos izquierdo y derecho de $H(\lambda_0)$ serán esenciales en la teoría de perturbación de autovalores, como lo son en el caso de haces regulares. Denotaremos a estos subespacios, respectivamente, por $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$. También será necesario considerar las intersecciones de estos subespacios con los subespacios reductores minimales, es decir (recuérdese la notación introducida en §2.5), $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$. Para este fin, agruparemos las columnas de la matriz X que da paso a la forma de Kronecker (2.5) en bloques correspondientes a los bloques de $K_H(\lambda)$ como sigue

$$X = [X_{\lambda_0} | \widehat{X} | X_\infty | X_{\varepsilon_1} | \dots | X_{\varepsilon_d} | X_{\eta_1} | \dots | X_{\eta_d}], \quad (4.5)$$

y las de Y como

$$Y = [Y_{\lambda_0} | \widehat{Y} | Y_\infty | Y_{\varepsilon_1} | \dots | Y_{\varepsilon_d} | Y_{\eta_1} | \dots | Y_{\eta_d}]. \quad (4.6)$$

Así, por ejemplo, $Y_{\varepsilon_i}^T H(\lambda) X_{\varepsilon_i} = L_{\varepsilon_i}(\lambda)$ e $Y_{\eta_j}^T H(\lambda) X_{\eta_j} = L_{\eta_j}^T(\lambda)$. A partir de estas particiones, definimos los vectores polinómicos

$$\pi_i(\lambda) \equiv Y_{\eta_i} \begin{bmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{\eta_i} \end{bmatrix}, \quad \psi_i(\lambda) \equiv X_{\varepsilon_i} \begin{bmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{\varepsilon_i} \end{bmatrix}, \quad i = 1, \dots, d. \quad (4.7)$$

Estos vectores polinómicos satisfacen ciertas propiedades que se recogen en el Lema 4.2.1. La definición de *base minimal*, que hemos introducido en §2.5, será empleada únicamente para demostrar el segundo apartado del Lema 4.2.2.

Lema 4.2.1 Sean $\{\pi_1(\lambda), \dots, \pi_d(\lambda)\}$ y $\{\psi_1(\lambda), \dots, \psi_d(\lambda)\}$, respectivamente, los vectores polinómicos definidos en (4.7). Entonces

1. Si el número μ no es un autovalor del haz cuadrado singular $H(\lambda)$, entonces los conjuntos $\{\pi_1(\mu), \dots, \pi_d(\mu)\}$ y $\{\psi_1(\mu), \dots, \psi_d(\mu)\}$ son, respectivamente, bases de los espacios nulos izquierdo y derecho de la matriz $H(\mu)$. Además, estos espacios nulos son, respectivamente, subconjuntos de \mathcal{R}_T y \mathcal{R} .
2. Si el número λ_0 es un autovalor del haz cuadrado singular $H(\lambda)$, entonces los conjuntos $\{\pi_1(\lambda_0), \dots, \pi_d(\lambda_0)\}$ y $\{\psi_1(\lambda_0), \dots, \psi_d(\lambda_0)\}$ son, respectivamente, bases de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$.

3. Los conjuntos $\{\pi_1(\lambda), \dots, \pi_d(\lambda)\}$ y $\{\psi_1(\lambda), \dots, \psi_d(\lambda)\}$ son, respectivamente, bases minimales del espacio singular izquierdo y derecho (sobre el cuerpo de las funciones racionales en λ) del polinomio matricial $H(\lambda)$.

Demostración. Los primeros dos apartados se siguen trivialmente de (2.5). Para el tercero: es fácil demostrar que los conjuntos que se consideran son bases. El hecho de que son minimales es una simple consecuencia de la teoría de haces singulares, véase el Lema 2.5.3. ■

Los subespacios considerados en el Lema 4.2.1 admiten muchas otras bases. El Lema 4.2.2 muestra algunas otras que aparecerán en las siguientes secciones. Antes de enunciarlo, queremos hacer constar que, por abuso de lenguaje, de aquí en adelante haremos referencia en algunas ocasiones a que un determinado conjunto de filas es una base del espacio nulo de una matriz A (o de un subespacio de este espacio nulo) o de un haz $H(\lambda)$. Con esto queremos decir que estamos considerando, implícitamente, el producto “a la izquierda”, $y^T A$, en lugar del producto “a la derecha”, $A^T y$.

Lema 4.2.2 *Sea $H(\lambda)$ un haz singular $n \times n$ con forma de Smith dada por (2.8). Recordemos que $d = n - r$. Entonces*

1. Si el número μ no es un autovalor de $H(\lambda)$ las últimas d filas de $P(\mu)$ y las últimas d columnas de $Q(\mu)$ son, respectivamente, bases de los espacios nulos izquierdo y derecho de la matriz $H(\mu)$.
2. Si el número λ_0 es un autovalor de $H(\lambda)$, las últimas d filas de $P(\lambda_0)$ y las últimas d columnas de $Q(\lambda_0)$ son, respectivamente, bases de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$.
3. Las últimas d filas de $P(\lambda)$ y las últimas d columnas de $Q(\lambda)$ son, respectivamente, bases del espacio nulo izquierdo y derecho (sobre el cuerpo de funciones racionales en λ) del polinomio matricial $H(\lambda)$.

Demostración. Los polinomios matriciales $P(\lambda)$ y $Q(\lambda)$ tienen determinante constante no nulo luego, para cualquier número μ las filas y columnas de las matrices constantes $P(\mu)$ y $Q(\mu)$ son linealmente independientes. El primer apartado se sigue directamente de combinar este hecho con la igualdad (2.8). El tercer apartado es trivial. Para demostrar el segundo apartado, necesitamos trabajar un poco más. Sólo demostraremos el enunciado para las últimas d columnas de $Q(\lambda_0)$, pues el resultado para las filas de $P(\lambda_0)$ es similar. Por el apartado 3 y el Lema 4.2.1, las últimas d columnas de $Q(\lambda)$ son combinaciones lineales de $\{\psi_1(\lambda), \dots, \psi_d(\lambda)\}$ con coeficientes polinómicos, porque $\{\psi_1(\lambda), \dots, \psi_d(\lambda)\}$ es una base minimal del espacio nulo derecho de $H(\lambda)$ [24, p. 495]. Si estas combinaciones lineales se particularizan para $\lambda = \lambda_0$, obtenemos que las últimas d columnas de $Q(\lambda_0)$ son combinaciones lineales de $\{\psi_1(\lambda_0), \dots, \psi_d(\lambda_0)\}$. Como ambos conjuntos son linealmente independientes, generan el mismo subespacio de \mathbb{C}^n . ■

Nótese que, para cualquier autovalor λ_0 de $H(\lambda)$, se sigue de (2.5) o (2.8) que

$$\dim \mathcal{N}_T(H(\lambda_0)) = \dim \mathcal{N}(H(\lambda_0)) = d + g, \quad (4.8)$$

donde g es la multiplicidad geométrica de λ_0 , es decir, el número de bloques de Jordan asociados con λ_0 en la FCK de $H(\lambda)$. Hasta el momento, solamente hemos determinado las bases de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$. Ahora completamos estas bases para conseguir bases de los subespacios $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$. Es esencial destacar que

cualquier base de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$ se puede utilizar en los desarrollos de perturbaciones que presentamos, pero para los *autovalores múltiples* se deben añadir unos vectores muy particulares para conseguir las bases de $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$ que necesitamos. Estos vectores están relacionados con la FCK de $H(\lambda)$ (2.5), y son descritos en el resto de esta sección.

Vamos a especificar un poco más la estructura espectral asociada con el autovalor finito λ_0 en la FCK (2.5) del haz singular $H(\lambda)$. La notación que emplearemos volverá a aparecer en la Sección 5.4. Sea la matriz J_{λ_0} de la forma

$$J_{\lambda_0} = \text{diag}(J_{n_1}^1(\lambda_0), \dots, J_{n_1}^{r_1}(\lambda_0), \dots, J_{n_q}^1(\lambda_0), \dots, J_{n_q}^{r_q}(\lambda_0)), \quad (4.9)$$

donde, para cada $i = 1, \dots, q$, las matrices $J_{n_i}^k(\lambda_0)$, $k = 1, \dots, r_i$, son bloques de Jordan de dimensión $n_i \times n_i$ asociados con λ_0 . Suponemos los bloques de Jordan, $J_{n_i}^k(\lambda_0)$, ordenados de manera que

$$n_1 < n_2 < \dots < n_q. \quad (4.10)$$

Como se ha mencionado en §2.4, las dimensiones n_i son las *multiplicidades parciales* de λ_0 , y nos referiremos a la partición (4.9) como la *estructura espectral* de λ_0 en $H(\lambda)$.

Sea a la multiplicidad algebraica de λ_0 y sean Y_{λ_0} (resp. X_{λ_0}) las matrices que aparecen en (4.6) (resp. (4.5)), es decir, la matriz cuyas columnas son las primeras a columnas de la matriz Y (resp. X) de (2.5), y particionemos

$$\begin{aligned} Y_{\lambda_0} &= \left[\begin{array}{c|c|c|c|c|c|c} Y_{n_1}^1 & \dots & Y_{n_1}^{r_1} & \dots & Y_{n_q}^1 & \dots & Y_{n_q}^{r_q} \\ \hline \hline \hline \hline \hline \hline \hline \end{array} \right], \\ X_{\lambda_0} &= \left[\begin{array}{c|c|c|c|c|c|c} X_{n_1}^1 & \dots & X_{n_1}^{r_1} & \dots & X_{n_q}^1 & \dots & X_{n_q}^{r_q} \\ \hline \hline \hline \hline \hline \hline \hline \end{array} \right], \end{aligned} \quad (4.11)$$

de manera conforme con (4.9). Denotaremos por x_i^k a la primera columna de $X_{n_i}^k$, y por y_i^k a la última columna de $Y_{n_i}^k$. Con esta elección, cada x_i^k es un elemento de $\mathcal{N}(H(\lambda_0))$ pero no un elemento de \mathcal{R} , y cada y_i^k es un elemento de $\mathcal{N}_T(H(\lambda_0))$ pero no un elemento de \mathcal{R}_T . Ahora, para cada $i = 1, \dots, q$ construimos las matrices

$$L_i = [y_i^1 \ \dots \ y_i^{r_i}], \quad R_i = [x_i^1 \ \dots \ x_i^{r_i}]$$

y

$$W_i = [L_i \ \dots \ L_q], \quad Z_i = [R_i \ \dots \ R_q]. \quad (4.12)$$

En esta situación, las dos cantidades

$$a = \sum_{i=1}^q r_i n_i, \quad g = \sum_{i=1}^q r_i,$$

son, respectivamente, la multiplicidad algebraica y geométrica de λ_0 . Finalmente, para cada $j = 1, \dots, q$, definimos

$$f_j = \sum_{i=j}^q r_i, \quad f_{q+1} = 0 \quad (4.13)$$

luego, tanto W_i como Z_i son matrices de tamaño $n \times f_i$. En particular, $f_1 = g$.

Si $H(\lambda)$ es regular, las matrices W_1 y Z_1 contienen, respectivamente, bases de los autovectores izquierdos y derechos asociados con λ_0 , es decir, bases de los espacios nulos izquierdo y derecho de $H(\lambda_0)$. Cuando $H(\lambda)$ es singular, necesitamos añadir a W_1 y Z_1 , respectivamente, bases de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$ para obtener las bases de $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$ que necesitamos.

4.3. Existencia de desarrollos

Esta sección está dedicada a caracterizar las perturbaciones genéricas, $M(\lambda) = B_0 + \lambda B_1$, para las que todos los autovalores del haz perturbado (4.3) son series de potencias en ϵ (eventualmente con exponentes racionales), y tal que tomando los límites de estas series cuando ϵ tiende a cero se obtienen *todos* los autovalores, finitos o infinitos, del haz cuadrado singular $H(\lambda) = A_0 + \lambda A_1$, junto a ciertos números (o infinitos) que quedan determinados por $M(\lambda)$ y no son autovalores de $H(\lambda)$. En el proceso, mostraremos cómo transformar el problema original de perturbación del haz singular $H(\lambda)$ en un problema de perturbación regular.

La forma canónica de Smith (2.8) será fundamental en esta sección. En aras de la simplicidad, expresaremos (2.8) en bloques como

$$P(\lambda) H(\lambda) Q(\lambda) \equiv \begin{bmatrix} P_1(\lambda) \\ P_2(\lambda) \end{bmatrix} H(\lambda) [Q_1(\lambda) Q_2(\lambda)] \equiv \begin{bmatrix} D_S(\lambda) & 0 \\ 0 & 0_{d \times d} \end{bmatrix}, \quad (4.14)$$

donde $D_S(\lambda) = \text{diag}(h_1(\lambda), \dots, h_r(\lambda))$, y las dimensiones de $P_1(\lambda)$ y $Q_1(\lambda)$ se toman de acuerdo a esta partición. Veremos que las condiciones genéricas sobre las perturbaciones están relacionadas con los bloques de la matriz (4.14). El siguiente lema expresa de varias formas equivalentes estas condiciones.

Lema 4.3.1 *Sea $H(\lambda) = A_0 + \lambda A_1$ un haz singular $n \times n$ con forma canónica de Smith dada por (4.14), y $M(\lambda) = B_0 + \lambda B_1$ otro haz $n \times n$. Entonces las siguientes afirmaciones son equivalentes*

1. $\det(P_2(\lambda) M(\lambda) Q_2(\lambda)) \neq 0$.
2. Existe un número μ , que no es un autovalor de $H(\lambda)$, y tal que

$$\det(\tilde{P}_2 M(\mu) \tilde{Q}_2) \neq 0,$$

para cualquier par de matrices $\tilde{P}_2 \in \mathbb{C}^{d \times n}$ y $\tilde{Q}_2 \in \mathbb{C}^{n \times d}$ cuyas filas y columnas, respectivamente, son bases de $\mathcal{N}_T(H(\mu))$ y $\mathcal{N}(H(\mu))$.

3. Existe un número μ , que no es autovalor de $A_1 + \lambda A_0$, y tal que

$$\det(\tilde{P}_2 (B_1 + \mu B_0) \tilde{Q}_2) \neq 0,$$

para cualquier par de matrices $\tilde{P}_2 \in \mathbb{C}^{d \times n}$ y $\tilde{Q}_2 \in \mathbb{C}^{n \times d}$ cuyas filas y columnas, respectivamente, son bases de $\mathcal{N}_T(A_1 + \mu A_0)$ y $\mathcal{N}(A_1 + \mu A_0)$.

Demostración. Nótese que $p(\lambda) = \det(P_2(\lambda) M(\lambda) Q_2(\lambda))$ es un polinomio en λ , por lo que será el polinomio nulo si y sólo si $p(\mu) = 0$ para cualquier μ . Nótese también que $p(\mu) \neq 0$ para algún μ si y sólo si $p(\mu) \neq 0$ para algún μ que no es un autovalor de

$H(\lambda)$. Por tanto, la primera afirmación es equivalente a la existencia de un valor μ que no es autovalor de $H(\lambda)$ y tal que $\det(P_2(\mu) M(\mu) Q_2(\mu)) \neq 0$, y la equivalencia con la segunda afirmación se sigue del Lema 4.2.2, porque $\tilde{Q}_2 = Q_2(\mu)S$ y $\tilde{P}_2 = T P_2(\mu)$ con S y T ciertas matrices invertibles. La equivalencia entre la segunda y la tercera afirmación se sigue de los hechos de que μ se puede tomar distinto de cero, que los espacios nulos de $A_0 + \mu A_1$ y de $(1/\mu)A_0 + A_1$ son iguales, y que $\tilde{P}_2(B_0 + \mu B_1)\tilde{Q}_2$ es regular si y sólo si $\tilde{P}_2((1/\mu)B_0 + B_1)\tilde{Q}_2$ es regular. ■

Observemos que, si tanto el haz $H(\lambda)$ como las matrices $P(\lambda)$ y $Q(\lambda)$ de (4.14) están fijados, entonces

$$\det(P_2(\lambda) M(\lambda) Q_2(\lambda)) \neq 0$$

es una condición genérica sobre el conjunto de haces perturbadores $B_0 + \lambda B_1$, porque no se cumple únicamente en la variedad algebraica definida igualando a cero todos los coeficientes del polinomio $p(\lambda) = \det(P_2(\lambda) M(\lambda) Q_2(\lambda))$. Estos coeficientes son polinomios multivariados en las entradas de B_0 y B_1 . Nótese también que el tercer apartado del Lema 4.3.1 significa que la condición se cumple simultáneamente para los haces duales.

El Teorema 4.3.2 traduce el problema original de perturbación de autovalores de un haz singular (4.3) en un problema regular de perturbación de las raíces de un cierto polinomio. Combinando este hecho con algunos resultados clásicos de la Teoría de Funciones Algebraicas (véase, por ejemplo, [32, Ch. 12]) se obtienen las siguientes conclusiones.

Teorema 4.3.2 *Sea $H(\lambda)$ un haz singular $n \times n$ con forma canónica de Smith dada por (4.14), y $M(\lambda)$ otro haz $n \times n$ tal que $\det(P_2(\lambda) M(\lambda) Q_2(\lambda)) \neq 0$. Entonces*

1. *Existe una constante $b > 0$ tal que el haz $H(\lambda) + \epsilon M(\lambda)$ es regular siempre que $0 < |\epsilon| < b$.*
2. *Para $0 < |\epsilon| < b$ los autovalores finitos de $H(\lambda) + \epsilon M(\lambda)$ son las raíces de un polinomio en λ , $p_\epsilon(\lambda)$, cuyos coeficientes son polinomios en ϵ . Además, cuando $\epsilon = 0$,*

$$p_0(\lambda) = \det(D_S(\lambda)) \det(P_2(\lambda) M(\lambda) Q_2(\lambda)). \quad (4.15)$$

3. *Sea ϵ tal que $0 < |\epsilon| < b$, entonces los n autovalores¹, $\{\lambda_1(\epsilon), \dots, \lambda_n(\epsilon)\}$, de $H(\lambda) + \epsilon M(\lambda)$ pueden desarrollarse en series de potencias (fraccionarias) de ϵ . Además, algunas de estas series pueden contener términos con exponentes negativos y tienden a ∞ cuando ϵ tiende a cero. El resto de las series convergen en un entorno de $\epsilon = 0$.*
4. *Si los autovalores finitos de $H(\lambda)$ son $\{\mu_1, \dots, \mu_k\}$, donde los elementos comunes se repiten de acuerdo a su multiplicidad algebraica, entonces existe un subconjunto $\{\lambda_{i_1}(\epsilon), \dots, \lambda_{i_k}(\epsilon)\}$ tal que*

$$\lim_{\epsilon \rightarrow 0} \lambda_{i_j}(\epsilon) = \mu_j \quad j = 1, \dots, k.$$

5. *Si el haz $H(\lambda)$ tiene un autovalor infinito con multiplicidad algebraica p , entonces existe un subconjunto $\{\lambda_{i_1}(\epsilon), \dots, \lambda_{i_p}(\epsilon)\}$ tal que*

$$\lim_{\epsilon \rightarrow 0} \lambda_{i_j}(\epsilon) = \infty \quad j = 1, \dots, p.$$

¹Es bien sabido que cualquier haz regular $n \times n$ tiene exactamente n autovalores, si se cuentan los finitos e infinitos [69, Ch. VI].

Demostración. Particionemos el producto $P(\lambda) M(\lambda) Q(\lambda)$ conforme con (4.14) en la forma

$$P(\lambda) M(\lambda) Q(\lambda) = \begin{bmatrix} B_{11}(\lambda) & B_{12}(\lambda) \\ B_{21}(\lambda) & B_{22}(\lambda) \end{bmatrix}.$$

Esto significa que $B_{22}(\lambda) = P_2(\lambda) M(\lambda) Q_2(\lambda)$. Entonces

$$\det(H(\lambda) + \epsilon M(\lambda)) = C \det \begin{bmatrix} D_S(\lambda) + \epsilon B_{11}(\lambda) & \epsilon B_{12}(\lambda) \\ \epsilon B_{21}(\lambda) & \epsilon B_{22}(\lambda) \end{bmatrix},$$

donde C es la constante no nula $C = 1/\det(P(\lambda) Q(\lambda))$. Entonces

$$\det(H(\lambda) + \epsilon M(\lambda)) = C \epsilon^d \det \begin{bmatrix} D_S(\lambda) + \epsilon B_{11}(\lambda) & B_{12}(\lambda) \\ \epsilon B_{21}(\lambda) & B_{22}(\lambda) \end{bmatrix}.$$

Definamos el polinomio en λ cuyos coeficientes son polinomios en ϵ

$$p_\epsilon(\lambda) \equiv \det \begin{bmatrix} D_S(\lambda) + \epsilon B_{11}(\lambda) & B_{12}(\lambda) \\ \epsilon B_{21}(\lambda) & B_{22}(\lambda) \end{bmatrix},$$

para escribir

$$\det(H(\lambda) + \epsilon M(\lambda)) = C \epsilon^d p_\epsilon(\lambda). \quad (4.16)$$

Es obvio que cuando $\epsilon = 0$

$$p_0(\lambda) = \det(D_S(\lambda)) \det(B_{22}(\lambda)). \quad (4.17)$$

Sabemos que $\det(D_S(\lambda)) \neq 0$, y, por lo tanto, $\det(B_{22}(\lambda)) \neq 0$ implica que $H(\lambda) + \epsilon M(\lambda)$ es regular en un disco punteado $0 < |\epsilon| < b$. Esto es obvio por continuidad: si $\det(D_S(\mu)) \cdot \det(B_{22}(\mu)) \neq 0$ para algún número fijo μ , entonces $p_\epsilon(\mu) \neq 0$ para ϵ suficientemente pequeño, porque $p_\epsilon(\mu)$ es una función continua de ϵ . Además, siempre que $0 < |\epsilon| < b$, la ecuación (4.16) implica que z es un autovalor finito de $H(\lambda) + \epsilon M(\lambda)$ si y sólo si $p_\epsilon(z) = 0$. Luego, el primer y segundo apartados del Teorema 4.3.2 quedan demostrados.

Nótese que hemos reducido el problema original de perturbación de autovalores al estudio de la variación de las raíces del polinomio $p_\epsilon(\lambda)$ cuando ϵ tiende a cero. Pero como los coeficientes de este polinomio son, a su vez, polinomios en ϵ este es un problema clásico resuelto por la Teoría de Funciones Algebraicas, véase, por ejemplo, [32, Sections 12.1–12.3]. En particular, el tercer apartado es una consecuencia de esta teoría (para autovalores infinitos se pueden aplicar argumentos similares al autovalor cero de los haces duales). Simplemente comentamos que si el grado de $p_\epsilon(\lambda)$ en λ es δ_1 y el grado de $\det(D_S(\lambda)) \cdot \det(B_{22}(\lambda))$ es $\delta_2 < \delta_1$, entonces un total de $\delta_1 - \delta_2$ raíces de $p_\epsilon(\lambda)$ tienden a infinito cuando ϵ tiende a cero. El cuarto apartado es, de nuevo, una consecuencia de la Teoría Algebraica de Funciones y (4.17), dado que las raíces que permanecen finitas tienen como límites las raíces de $\det(D_S(\lambda)) \cdot \det(B_{22}(\lambda))$, y las raíces de $\det(D_S(\lambda))$ son precisamente los autovalores finitos de $H(\lambda)$.

El último apartado se puede demostrar aplicando los resultados previos al autovalor cero del haz dual de $H(\lambda) + \epsilon M(\lambda)$, y teniendo en cuenta que $\lambda_i(\epsilon)$ es un autovalor de $H(\lambda) + \epsilon M(\lambda)$ si y sólo si $1/\lambda_i(\epsilon)$ es un autovalor del haz dual. ■

El Teorema 4.3.2 nos da una condición suficiente para la existencia *simultánea* de desarrollos de perturbaciones para *todos* los autovalores de $H(\lambda) + \epsilon M(\lambda)$. Algunos de estos

desarrollos tienen por límites las raíces de $\det(P_2(\lambda)M(\lambda)Q_2(\lambda))$, que quedan completamente determinadas por la perturbación $M(\lambda)$, y el resto de los desarrollos tienen por límites los autovalores de $H(\lambda)$. La condición $\det(P_2(\lambda)M(\lambda)Q_2(\lambda)) \neq 0$ se puede relajar si estamos únicamente interesados en la existencia de algunos de estos desarrollos. Además, el Teorema 4.3.2 es un resultado muy simple que no nos dice cuáles son estos desarrollos, o cuáles son sus coeficientes y exponentes directores. Obtendremos esta información en la Sección 4.5, con el coste adicional de imponer condiciones más específicas. El punto principal del Teorema 4.3.2 y su demostración es que, genéricamente, la teoría de perturbación de primer orden de autovalores de haces singulares cuadrados es simplemente un problema usual de perturbaciones de las raíces de un polinomio cuyos coeficientes son polinomios en el parámetro de la perturbación.

Ejemplo 4.3.3 Apliquemos los resultados de esta sección al primer ejemplo (4.1) que hemos presentado en la Introducción. Nótese que, en este caso, ,

$$H(\lambda) = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}, \quad M(\lambda) = \begin{bmatrix} 6 & -3 \\ -10 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

y $P(\lambda)$ y $Q(\lambda)$ son, ambas, la matriz identidad 2×2 . Por lo tanto,

$$P_2(\lambda)M(\lambda)Q_2(\lambda) = 0$$

(es la entrada $(2, 2)$ de la perturbación) y el Teorema 4.3.2 no puede aplicarse. Si la perturbación $M(\lambda)$ se modifica con $M_{22}(\lambda) = c_{22} + \lambda d_{22} \neq 0$, entonces puede comprobarse que los límites cuando ϵ tiende a cero de las raíces del polinomio $\det(H(\lambda) + \epsilon M(\lambda)) = 0$ son precisamente las raíces de $p_0(\lambda) = \lambda(c_{22} + \lambda d_{22})$, que es (4.15) en este ejemplo. Luego, para ϵ suficientemente pequeño, hay siempre una raíz próxima a cero. Otras observaciones interesantes que pueden comprobarse fácilmente son: (i) si $d_{22} = 0$ entonces una de las raíces tiende a infinito; (ii) si $c_{22} = 0$ ambas raíces se aproximan a cero como $\pm c \epsilon^{1/2} + o(\epsilon^{1/2})$. En este último caso la perturbación hace que el autovalor simple $\lambda = 0$ del haz $H(\lambda)$ se comporte como un autovalor doble desde el punto de vista de las perturbaciones. La teoría que desarrollaremos no cubre este tipo de situaciones no genéricas. Finalmente, nótese que, aunque la perturbación

$$M(\lambda) = \begin{bmatrix} 6 & \lambda \\ \lambda - 3 & 0 \end{bmatrix}$$

no satisface las hipótesis del Teorema 4.3.2, se tiene que $\det(H(\lambda) + \epsilon M(\lambda)) = -\epsilon^2 \lambda(\lambda - 3)$, luego $\lambda_0 = 0$ es un autovalor simple del haz perturbado para cualquier valor de ϵ . Por lo tanto, la hipótesis genérica $\det(P_2(\lambda)M(\lambda)Q_2(\lambda)) \neq 0$ del Teorema 4.3.2 es suficiente pero no necesaria para la existencia de desarrollos. \square

4.4. De la forma de Kronecker a la forma local de Smith

Los resultados de la Sección 4.3 muestran que la forma canónica de Smith tiene un papel relevante en la teoría de perturbaciones genéricas de autovalores de haces matriciales singulares. No obstante, la forma canónica de Smith no revela toda la información espectral de los haces singulares, al contrario de lo ocurre con la FCK. De hecho, como

hemos visto en §2.4.1, es fácil encontrar ejemplos de haces con la misma forma de Smith, pero diferentes formas de Kronecker. El propósito de esta sección es relacionar las matrices que transforman un haz en su FCK con las matrices que transforman el mismo haz en su forma local de Smith que, como sabemos, revela el rango del haz y los divisores elementales correspondientes a un autovalor concreto λ_0 del haz.

Recordemos que la forma local de Smith en λ_0 (2.9) se obtiene por equivalencia local en λ_0 a partir de dos matrices P_{λ_0} y Q_{λ_0} . Estas matrices pueden obtenerse a partir de las matrices de paso a la forma de Smith (2.8), $P(\lambda)$ y/o $Q(\lambda)$ (sólo una es necesaria) multiplicándolas por las inversas de matrices polinómicas diagonales cuyas entradas diagonales no se anulan en λ_0 . Las matrices $P_{\lambda_0}(\lambda)$ y $Q_{\lambda_0}(\lambda)$ de (2.9) no son únicas. En esta sección, relacionaremos la forma de Kronecker con la forma local de Smith mostrando que es posible transformar las matrices constantes X e Y de la FCK (2.5) para obtener matrices racionales específicas $P_{\lambda_0}(\lambda)$ y $Q_{\lambda_0}(\lambda)$ que satisfacen la igualdad (2.9). El procedimiento será el siguiente:

- i) Transformar $H(\lambda)$ en su FCK, $K_H(\lambda)$, por medio de X e Y como en (2.5).
- ii) Transformar $K_H(\lambda)$ en $\Delta_{\lambda_0}(\lambda)$ por medio de matrices racionales $P_1(\lambda)$ y $Q_1(\lambda)$, tales que $\det(P_1(\lambda)) = 1/p(\lambda)$, $\det(Q_1(\lambda)) = 1/q(\lambda)$, donde $p(\lambda)$ y $q(\lambda)$ son polinomios que satisfacen $p(\lambda_0) \neq 0$ y $q(\lambda_0) \neq 0$:

$$P_1(\lambda)K_H(\lambda)Q_1(\lambda) = \Delta_{\lambda_0}(\lambda).$$

- iii) Tomar $P_{\lambda_0}(\lambda) = P_1(\lambda)Y^T$ y $Q_{\lambda_0}(\lambda) = XQ_1(\lambda)$.

Estas matrices evaluadas en λ_0 , esto es, $P_{\lambda_0}(\lambda_0)$ y $Q_{\lambda_0}(\lambda_0)$, están estrechamente relacionadas con las matrices W_1 y Z_1 definidas en (4.12).

Comenzaremos especificando las transformaciones dependientes de λ que se emplearán en el paso ii).

Lema 4.4.1 *Sea λ_0 un número complejo. Entonces*

- a) *Para cada entero positivo k se tiene*

$$P_k(\lambda - \lambda_0)(\lambda I_k - J_k(\lambda_0))Q_k(\lambda - \lambda_0) = \text{diag}((\lambda - \lambda_0)^k, 1, \dots, 1),$$

donde las matrices

$$P_k(\lambda) = \begin{bmatrix} \lambda^{k-1} & \dots & \lambda & 1 \\ \vdots & \ddots & \ddots & \\ \lambda & 1 & & \\ 1 & & & \end{bmatrix} \in \mathbb{C}^{k \times k}, \quad Q_k(\lambda) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda & 0 & \dots & -1 \\ \vdots & \vdots & \ddots & \\ \lambda^{k-1} & -1 & & \end{bmatrix} \in \mathbb{C}^{k \times k}$$

son matrices polinómicas con determinante constante no nulo igual a ± 1 , es decir, invertibles para todo λ .

- b) *Para todo $\lambda_i \neq \lambda_0$ y cada entero positivo k existen dos matrices $k \times k$, $\widehat{P}_k^i(\lambda)$ y $\widehat{Q}_k^i(\lambda)$, tales que una de ellas es una matriz polinomial con determinante constante no nulo, y la otra tiene entradas racionales cuyos denominadores son $(\lambda - \lambda_i)^k$ ó 1 y determinante $\pm 1/(\lambda - \lambda_i)^k$, y*

$$\widehat{P}_k^i(\lambda)(\lambda I_k - J_k(\lambda_i))\widehat{Q}_k^i(\lambda) = I_k.$$

c) Para cada entero positivo k se tiene

$$(I - \lambda J_k(0))Q_k^\infty(\lambda) = I_k,$$

donde la matriz

$$Q_k^\infty(\lambda) = \begin{bmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{k-1} \\ & 1 & \lambda & \dots & \lambda^{k-2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \lambda \\ & & & & 1 \end{bmatrix} \in \mathbb{C}^{k \times k}$$

tiene determinante constante no nulo igual a 1.

d) Para cada entero positivo σ se tiene

$$L_\sigma(\lambda)C_{\sigma+1}(\lambda) := B_\sigma := \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & 0 & 1 & \\ & & & & \end{bmatrix} \in \mathbb{C}^{\sigma \times (\sigma+1)},$$

donde la matriz

$$C_{\sigma+1}(\lambda) = \begin{bmatrix} 1 & & & & \\ \lambda & -1 & & & \\ \lambda^2 & -\lambda & -1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \lambda^\sigma & -\lambda^{\sigma-1} & \dots & -\lambda & -1 \end{bmatrix} \in \mathbb{C}^{(\sigma+1) \times (\sigma+1)}$$

tiene determinante constante no nulo igual a ± 1 .

Demostración. Los apartados a), c) y d) pueden comprobarse fácilmente. Para demostrar b), nótese que una transformación análoga a la que se ha descrito en a) transforma cada bloque $(\lambda I_k - J_k(\lambda_i))$ en $\text{diag}((\lambda - \lambda_i)^k, 1, \dots, 1)$, para $\lambda_i \neq \lambda_0$. Multiplicando a la derecha por $\text{diag}((\lambda - \lambda_i)^{-k}, 1, \dots, 1)$ llegamos a I_k . Nótese que $\det(\text{diag}((\lambda - \lambda_i)^{-k}, 1, \dots, 1)) = (\lambda - \lambda_i)^{-k}$. ■

Ahora podemos especificar las matrices mencionadas $P_{\lambda_0}(\lambda_0)$ y $Q_{\lambda_0}(\lambda_0)$. Esto involucra a los espacios singulares izquierdo y derecho del haz (o subespacios reductores minimales), y a los espacios nulos izquierdo y derecho asociados con un autovalor finito λ_0 , es decir, $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$, que aparecen en la Sección 4.2.

Lema 4.4.2 Sea $H(\lambda)$ un haz singular $n \times n$ con FCK dada por (2.5), con subespacio singular derecho \mathcal{R} , y subespacio singular izquierdo \mathcal{R}_T . Sea λ_0 un autovalor finito de $H(\lambda)$ con estructura espectral (4.9). Sean W_1 y Z_1 las matrices definidas en (4.12), denotemos por $W_{\mathcal{R}} \in \mathbb{C}^{d \times n}$ a una matriz cuyas columnas forman una base de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$, y por $Z_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ a una matriz cuyas columnas forman una base de $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$. Sea $\Delta_{\lambda_0}(\lambda)$ la forma local de Smith de $H(\lambda)$ en λ_0 definida en (2.9). Entonces, existen dos matrices $P_{\lambda_0}(\lambda)$ y $Q_{\lambda_0}(\lambda)$, cuyas entradas son funciones racionales con denominadores no nulos en λ_0 , $\det(P_{\lambda_0}(\lambda)) = 1/p(\lambda)$, $\det(Q_{\lambda_0}(\lambda)) = 1/q(\lambda)$, donde $p(\lambda)$ y $q(\lambda)$ son polinomios que satisfacen $p(\lambda_0) \neq 0 \neq q(\lambda_0)$, de modo que

$$P_{\lambda_0}(\lambda)H(\lambda)Q_{\lambda_0}(\lambda) = \Delta_{\lambda_0}(\lambda),$$

y

$$P_{\lambda_0}(\lambda_0) = [W_1 \ * \ W_{\mathcal{R}}]^T, \quad Q(\lambda_0) = [Z_1 \ * \ Z_{\mathcal{R}}], \quad (4.18)$$

donde las columnas denotadas con * no se especifican.

Demostración. Primero, agrupamos en dos matrices diagonales por bloques, $\widehat{P}(\lambda)$ y $\widehat{Q}(\lambda)$, todas las transformaciones $\widehat{P}_k^i(\lambda)$ y $\widehat{Q}_k^i(\lambda)$, como en el Lema 4.4.1 b), correspondientes a los bloques de Jordan asociados con los autovalores finitos $\lambda_i \neq \lambda_0$. También construimos una matriz diagonal por bloques, $Q_\infty(\lambda)$, de dimensión $a_\infty \times a_\infty$ (a_∞ es la multiplicidad algebraica del autovalor infinito) que incluye todas las matrices $Q_k^\infty(\lambda)$ del Lema 4.4.1 c) correspondientes a los bloques de Jordan asociados con el autovalor infinito.

Ahora, si definimos

$$P_0(\lambda) = \text{diag}(P_{n_1}(\lambda - \lambda_0), \dots, P_{n_q}(\lambda - \lambda_0), \widehat{P}(\lambda), I_{a_\infty}, I_\varepsilon, C_{\eta_1+1}^T(\lambda), \dots, C_{\eta_d+1}^T(\lambda)),$$

con ε dado por (2.7), y

$$Q_0(\lambda) = \text{diag}(Q_{n_1}(\lambda - \lambda_0), \dots, Q_{n_q}(\lambda - \lambda_0), \widehat{Q}(\lambda), Q_\infty(\lambda), C_{\varepsilon_1+1}(\lambda), \dots, C_{\varepsilon_d+1}(\lambda), I_\eta),$$

donde η está dado por (2.7), y los bloques diagonales $P_{n_i}(\cdot)$ y $Q_{n_i}(\cdot)$ son como se han definido en el Lema 4.4.1 a), y donde C_{ε_j+1} , $C_{\eta_j+1}^T$ son como en el Lema 4.4.1 d), entonces

$$P_0(\lambda)K_H(\lambda)Q_0(\lambda) = \text{diag}(\overbrace{((\lambda - \lambda_0)^{n_1}, 1, \dots, 1)}^{n_1}, \dots, \overbrace{((\lambda - \lambda_0)^{n_q}, 1, \dots, 1)}^{n_q}, I, B_{\varepsilon_1}, \dots, \\ \dots, B_{\varepsilon_d}, B_{\eta_1}^T, \dots, B_{\eta_d}^T),$$

y cada bloque diagonal de la forma $\text{diag}((\lambda - \lambda_0)^{n_i}, 1, \dots, 1)$, para $i = 1, \dots, q$, aparece repetido r_i veces. Por tanto, hay $g = \sum_{i=1}^q r_i$ de estos bloques.

Una permutación final de las filas y columnas de esta matriz nos lleva a la forma local de Smith en λ_0 . Esta permutación lleva la primera fila y la primera columna correspondientes a un bloque diagonal, $\text{diag}((\lambda - \lambda_0)^{n_i}, 1, \dots, 1)$, a las primeras g filas y columnas de $\Delta_{\lambda_0}(\lambda)$. Por otro lado, las últimas d filas nulas (resp. las últimas d columnas nulas) de $\Delta_{\lambda_0}(\lambda)$ proceden de las primeras filas (resp. las primeras columnas) de cada uno de los d bloques singulares $B_{\eta_j}^T$ (resp. B_{ε_i}) anteriores. Si denotamos por Π_l y Π_r las correspondientes matrices de permutación izquierdas y derechas, podemos definir

$$P_1(\lambda) = \Pi_l P_0(\lambda), \quad Q_1(\lambda) = Q_0(\lambda) \Pi_r,$$

y

$$\widetilde{P}(\lambda) = P_1(\lambda) Y^T, \quad \widetilde{Q}(\lambda) = X Q_1(\lambda).$$

Las matrices $\widetilde{P}(\lambda_0)$ y $\widetilde{Q}(\lambda_0)$ son como las que se han descrito en (4.18) para una elección específica de $W_{\mathcal{R}}$ y $Z_{\mathcal{R}}$. Para ver esto, solamente necesitamos seguir la pista a las columnas de Y (resp., a las de X) tras multiplicar a la izquierda por $P_1(\lambda_0)$ (resp., a la derecha por $Q_1(\lambda_0)$). Primero, nótese que, para cada $k = n_1, \dots, n_q$, la matriz de permutación $P_k(0)$ incluye una transposición de las filas, mientras que la multiplicación a la derecha por $Q_k(0)$ mantiene la primera columna fija. Así pues, usando la notación del párrafo que sigue a la ecuación (4.11), la multiplicación a la izquierda por $P_0(\lambda_0)$ mueve cada vector fila $(y_i^j)^T$ a la primera fila en su bloque correspondiente, mientras que la multiplicación por $Q_0(\lambda_0)$ a la derecha mantiene los vectores columna x_i^j inalterados.

La multiplicación final por Π_l y Π_r lleva los vectores $(y_i^j)^T$ (resp., x_i^j) a las primeras g filas de $\tilde{P}(\lambda_0)$ (resp., a las primeras g columnas de $\tilde{Q}(\lambda_0)$). Por lo tanto, obtenemos que

$$\tilde{P}(\lambda_0) = [W_1 \mid *]^T, \quad \tilde{Q}(\lambda_0) = [Z_1 \mid *].$$

Igual que para las últimas d filas de $\tilde{P}(\lambda_0)$, tomemos las filas de Y^T correspondientes a un bloque $Y_{\eta_i}^T$ de los que aparecen en (4.6). La multiplicación a la izquierda por $P_0(\lambda_0)$, restringida a estas filas, nos da el producto

$$C_{\eta_i+1}^T(\lambda_0) Y_{\eta_i}^T = [\pi_i(\lambda_0) \quad * \quad \dots \quad *]^T,$$

de acuerdo con (4.7), y las entradas denotadas con $*$ carecen de relevancia en nuestro argumento. La permutación final Π_l mueve las filas $\pi_1(\lambda_0)^T, \dots, \pi_d(\lambda_0)^T$ a las últimas d filas de $\tilde{P}(\lambda_0)$. Un argumento similar con las columnas de X nos da el correspondiente resultado para $\tilde{Q}(\lambda_0)$. Hemos obtenido, hasta el momento, que

$$\tilde{P}(\lambda_0) = [W_1 \quad * \quad \pi_1(\lambda_0) \quad \dots \quad \pi_d(\lambda_0)]^T, \quad \tilde{Q}(\lambda_0) = [Z_1 \quad * \quad \psi_1(\lambda_0) \quad \dots \quad \psi_d(\lambda_0)],$$

con los vectores polinómicos $\pi_i(\lambda)$ y $\psi_i(\lambda)$ como se definen en (4.7). Estas matrices son, por el Lema 4.2.1, del tipo que aparece en (4.18). Finalmente, para obtener cualquier base $W_{\mathcal{R}}$ de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$, y cualquier base $Z_{\mathcal{R}}$ de $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$, multiplicamos por las matrices diagonales por bloques

$$P(\lambda) = \text{diag}(I_{n-d}, E) P_1(\lambda) Y^T, \quad \tilde{Q}(\lambda) = X Q_1(\lambda) \text{diag}(I_{n-d}, F),$$

donde E y F son matrices constantes invertibles de tamaño $d \times d$ ■

4.5. Desarrollos de Puiseux para los autovalores del haz perturbado

Dado un autovalor finito λ_0 de un haz cuadrado arbitrario $H(\lambda)$, regular o singular, regresamos a nuestro problema central: obtener, bajo ciertas condiciones de genericidad en las perturbaciones $M(\lambda)$, desarrollos, al primer orden, de los autovalores de la perturbación en términos del parámetro ϵ para aquellos autovalores del haz perturbado (4.3) cuyo límite es λ_0 cuando ϵ tiende a cero. Los coeficientes directores de estos desarrollos son los autovalores de cierto haz regular auxiliar que se construye a partir de la matriz $M(\lambda_0)$ y de ciertas bases de los espacios nulos izquierdo y derecho de $H(\lambda_0)$. Para los autovalores simples veremos que podemos utilizar cualquiera de estas bases, pero para autovalores múltiples tendrán que utilizarse unas bases normalizadas de un modo no trivial con el fin de construir estos haces auxiliares. En la Sección 4.5.1, definimos estos haces auxiliares y demostramos algunas de sus propiedades básicas. En la Sección 4.5.2, presentamos los desarrollos de las perturbaciones para autovalores finitos. Los desarrollos para el autovalor infinito se obtienen a partir de los desarrollos del autovalor cero en el haz dual. Estos desarrollos se presentan en la Sección 4.5.3. Finalmente, los desarrollos para autovalores simples se estudian en la Sección 4.5.4.

4.5.1. Haces auxiliares

Recordemos algunas de las matrices que hemos definido previamente. Dado un autovalor finito λ_0 del haz cuadrado $H(\lambda)$ con FCK dada por (2.5) y estructura espectral de λ_0 dada por (4.9), consideramos las matrices W_i y Z_i , $i = 1, \dots, q$, definidas en (4.12). Denotemos por $W_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ y $Z_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ dos matrices cuyas columnas forman una base de, respectivamente, $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$, donde, recordamos, \mathcal{R}_T y \mathcal{R} son los subespacios singulares izquierdo y derecho de $H(\lambda)$ (véase la Sección 2.5), y $\mathcal{N}_T(H(\lambda_0))$ y $\mathcal{N}(H(\lambda_0))$ son los espacios nulos izquierdo y derecho de $H(\lambda_0)$. Denotaremos por Φ_1 a la matriz de tamaño $(g+d) \times (g+d)$

$$\Phi_1 = \begin{bmatrix} W_1^T \\ W_{\mathcal{R}}^T \end{bmatrix} M(\lambda_0) \begin{bmatrix} Z_1 & Z_{\mathcal{R}} \end{bmatrix}. \quad (4.19)$$

Recordemos que las columnas de $[W_1 \ W_{\mathcal{R}}]$ son una base de $\mathcal{N}_T(H(\lambda_0))$, y las columnas de $[Z_1 \ Z_{\mathcal{R}}]$ son una base de $\mathcal{N}(H(\lambda_0))$. Recordemos también las dimensiones f_j definidas en (4.13) y, para cada $j = 1, \dots, q$, definamos

$$\Phi_j = \Phi_1(g - f_j + 1, \dots, g, g + 1, \dots, g + d) = \begin{bmatrix} W_j^T \\ W_{\mathcal{R}}^T \end{bmatrix} M(\lambda_0) \begin{bmatrix} Z_j & Z_{\mathcal{R}} \end{bmatrix} \quad (4.20)$$

como la submatriz principal inferior derecha de tamaño $(f_j+d) \times (f_j+d)$ de Φ_1 . Finalmente, definimos

$$\Phi_{q+1} = \Phi_1(g + 1, \dots, g + d) = W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}. \quad (4.21)$$

Nótese que cada Φ_j está anidada como una submatriz principal inferior derecha de Φ_{j-1} . Nótese también que si $H(\lambda)$ es regular, entonces Φ_j es, simplemente,

$$\Phi_j = W_j^T M(\lambda_0) Z_j. \quad (4.22)$$

Estas notaciones se ilustran en el siguiente ejemplo.

Ejemplo 4.5.1 Consideremos el haz

$$H(\lambda) = \begin{bmatrix} \lambda - 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda - 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & \lambda - 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

que está dado en forma de Kronecker. Tiene un único autovalor finito igual a $\lambda_0 = 1$ con multiplicidad algebraica 3, y dos bloques singulares, uno derecho y el otro izquierdo, con índices minimales iguales a 1. De acuerdo con nuestra notación de (4.9) y de la Sección 2.3, tenemos

$$r_1 = 1, \ n_1 = 1, \ r_2 = 1, \ n_2 = 2, \ d = 1, \ \varepsilon_1 = \eta_1 = 1.$$

Si tomamos el haz perturbación

$$M(\lambda) = \begin{bmatrix} 2 & 1 & -1 & 1 & 1 & 2 \\ 0 & 1 & -2 & 3 & 1 & 4 \\ 3 & 2 & 1 & -1 & 2 & 1 \\ 3 & 0 & 2 & 5 & 1 & 0 \\ 0 & 3 & 1 & 1 & 1 & 2 \\ 5 & 1 & 0 & 0 & -2 & -2 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 2 & 3 & 1 & 2 & 0 \\ 1 & 1 & -1 & 3 & 2 & 1 \\ 1 & 0 & 0 & 2 & 3 & 1 \\ -4 & 1 & 2 & 6 & 1 & -2 \\ 0 & 0 & 0 & 1 & 2 & -2 \\ -5 & -1 & 2 & 1 & -3 & 0 \end{bmatrix},$$

entonces la matriz Φ_1 es

$$\Phi_1 = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & & & \\ 0 & 0 & 1 & & & \\ \hline & & & 0 & 1 & 1 \end{array} \right] M(1) \left[\begin{array}{cc|c} 1 & 0 & \\ 0 & 1 & \\ 0 & 0 & \\ \hline & & 1 \\ & & 1 \\ & & 0 \end{array} \right] = \begin{bmatrix} 2 & 3 & 5 \\ 4 & 2 & 6 \\ 0 & 3 & 1 \end{bmatrix},$$

donde se ha hecho uso del Lema 4.2.1 para construir las matrices $W_{\mathcal{R}}$ y $Z_{\mathcal{R}}$. Además,

$$\Phi_2 = \begin{bmatrix} 2 & 6 \\ 3 & 1 \end{bmatrix} \quad \text{y} \quad \Phi_3 = 1.$$

□

Asociados con las matrices Φ_j , $j = 1, \dots, q$, definimos

$$E_j = \text{diag}(I_{r_j}, 0_{(f_{j+1}+d) \times (f_{j+1}+d)}), \quad j = 1, \dots, q. \quad (4.23)$$

Nótese que E_j es una matriz $(f_j+d) \times (f_j+d)$. Los haces que se necesitan en los desarrollos de las perturbaciones que expondremos son $\Phi_j + \zeta E_j$, $j = 1, \dots, q$. Algunas propiedades de estos haces se presentan en el siguiente Lema 4.5.2.

Lema 4.5.2 Sean Φ_j y E_j , $j = 1, \dots, q$, las matrices definidas, respectivamente, en (4.20) y (4.23), y Φ_{q+1} la matriz definida en (4.21). Si la matriz Φ_{j+1} es invertible, entonces

1. El haz $\Phi_j + \zeta E_j$ es regular y tiene exactamente r_j autovalores finitos.
2. Los autovalores finitos de $\Phi_j + \zeta E_j$ son los opuestos de los autovalores del complemento de Schur de Φ_{j+1} en Φ_j .
3. Si, además, Φ_j es invertible, entonces los r_j autovalores finitos de $\Phi_j + \zeta E_j$ son todos distintos de cero.

Demostración. Expresemos

$$\Phi_j = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & \Phi_{j+1} \end{bmatrix}.$$

Entonces

$$\Phi_j + \zeta E_j = \begin{bmatrix} C_{11} + \zeta I_{r_j} & C_{12} \\ C_{21} & \Phi_{j+1} \end{bmatrix},$$

y

$$(\Phi_j + \zeta E_j) \begin{bmatrix} I_{r_j} & 0 \\ -\Phi_{j+1}^{-1} C_{21} & I \end{bmatrix} = \begin{bmatrix} C_{11} - C_{12} \Phi_{j+1}^{-1} C_{21} + \zeta I_{r_j} & C_{12} \\ 0 & \Phi_{j+1} \end{bmatrix}.$$

El haz que aparece en el término derecho de la ecuación anterior es estrictamente equivalente a $(\Phi_j + \zeta E_j)$, luego el primer apartado se sigue fácilmente. Si Φ_j es invertible, entonces $C_{11} - C_{12} \Phi_{j+1}^{-1} C_{21}$ es también invertible, lo que significa que todos sus autovalores son distintos de cero. ■

4.5.2. Desarrollos de primer orden para autovalores finitos

Ahora estamos en disposición de demostrar el Teorema 4.5.3, que es el resultado más importante de este capítulo. La demostración de este teorema tiene dos partes: la primera utiliza la forma local de Smith para transformar el problema original de perturbación de autovalores de un haz que puede ser singular en un problema de perturbación regular. La segunda parte aplica a este problema de perturbación regular las técnicas desarrolladas en [46] para obtener el primer orden de los desarrollos de las perturbaciones de los autovalores. Esta segunda parte es muy similar a la demostración que aparece en [46, pp. 798-801]. Después de la demostración del Teorema 4.5.3, discutiremos la genericidad de las condiciones impuestas sobre las perturbaciones y compararemos estas condiciones con las del Teorema 4.3.2. Nótese que si el haz $H(\lambda)$ es regular entonces la matriz Φ_{q+1} no existe, y las condiciones sobre esta matriz no son necesarias.

Teorema 4.5.3 *Sea $H(\lambda)$ un haz matricial arbitrario (singular o no) de tamaño $n \times n$ con forma de Kronecker (2.5), y $M(\lambda)$ otro haz de la misma dimensión. Sea λ_0 un autovalor finito de $H(\lambda)$ con estructura espectral dada por (4.9) y (4.10). Sean Φ_j y E_j , $j = 1, \dots, q$, las matrices definidas en (4.20) y (4.23), y Φ_{q+1} la matriz definida en (4.21). Si $\det \Phi_{j+1} \neq 0$ para algún $j \in \{1, 2, \dots, q\}$, sean ξ_1, \dots, ξ_{r_j} los r_j autovalores finitos del haz $\Phi_j + \zeta E_j$, y $(\xi_r)_s^{1/n_j}$, $s = 1, \dots, n_j$, las n_j determinaciones complejas de la n_j -ésima raíz de la unidad. Entonces, en un entorno de $\epsilon = 0$, el haz $H(\lambda) + \epsilon M(\lambda)$ tiene $r_j n_j$ autovalores que satisfacen*

$$\lambda_j^{rs}(\epsilon) = \lambda_0 + (\xi_r)_s^{1/n_j} \epsilon^{1/n_j} + o(\epsilon^{1/n_j}), \quad r = 1, 2, \dots, r_j, \quad s = 1, 2, \dots, n_j, \quad (4.24)$$

donde ϵ^{1/n_j} es la determinación principal² de la n_j -ésima raíz de ϵ . Más aún, el haz $H(\lambda) + \epsilon M(\lambda)$ es regular en el mismo entorno para $\epsilon \neq 0$. Si, además, $\det \Phi_j \neq 0$, entonces todos los valores ξ_r de (4.24) son no nulos, y los desarrollos (4.24) son todos los desarrollos en torno a λ_0 con exponente director $1/n_j$.

Demostración. La demostración se basa en la forma local de Smith del Lema 4.4.2. Nos restringimos al caso en que $\lambda_0 = 0$. Si $\lambda_0 \neq 0$, basta con hacer la traslación $\mu = \lambda - \lambda_0$ en la forma local de Smith: $P_{\lambda_0}(\lambda - \lambda_0 + \lambda_0)H(\lambda - \lambda_0 + \lambda_0)Q_{\lambda_0}(\lambda - \lambda_0 + \lambda_0) = \Delta_{\lambda_0}(\lambda - \lambda_0 + \lambda_0)$, definir $\tilde{P}(\mu) := P_{\lambda_0}(\mu + \lambda_0)$, $\tilde{Q}(\mu) := Q_{\lambda_0}(\mu + \lambda_0)$, $\tilde{H}(\mu) := H(\mu + \lambda_0)$, y $\tilde{\Delta}_{\lambda_0}(\mu) := \Delta_{\lambda_0}(\mu + \lambda_0)$, y, finalmente, considerar $\tilde{P}(\mu)\tilde{H}(\mu)\tilde{Q}(\mu) = \tilde{\Delta}_{\lambda_0}(\mu)$. Nótese que $\tilde{P}(0) = P_{\lambda_0}(\lambda_0)$ y $\tilde{Q}(0) = Q_{\lambda_0}(\lambda_0)$, y que estas matrices están dadas por (4.18).

Suponiendo que $\lambda_0 = 0$, consideramos la transformación a la forma local de Smith en $\lambda_0 = 0$,

$$P_{\lambda_0}(\lambda)(H(\lambda) + \epsilon M(\lambda))Q_{\lambda_0}(\lambda) = \Delta_{\lambda_0}(\lambda) + \epsilon P_{\lambda_0}(\lambda)M(\lambda)Q_{\lambda_0}(\lambda) \equiv \hat{\Delta}_{\lambda_0}(\lambda) + G(\lambda, \epsilon), \quad (4.25)$$

donde

$$\hat{\Delta}_{\lambda_0}(\lambda) = \begin{bmatrix} D(\lambda) & & \\ & 0 & \\ & & 0_{d \times d} \end{bmatrix} \quad \text{y} \quad G(\lambda, \epsilon) = \begin{bmatrix} \epsilon G_{11}(\lambda) & \epsilon G_{12}(\lambda) & \epsilon G_{13}(\lambda) \\ \epsilon G_{21}(\lambda) & I + \epsilon G_{22}(\lambda) & \epsilon G_{23}(\lambda) \\ \epsilon G_{31}(\lambda) & \epsilon G_{32}(\lambda) & \epsilon G_{33}(\lambda) \end{bmatrix}$$

están particionadas de manera conforme, y $[G_{ij}(\lambda)]_{i,j=1}^3 = P_{\lambda_0}(\lambda)M(\lambda)Q_{\lambda_0}(\lambda)$. Por lo tanto, si $H(\lambda) + \epsilon M(\lambda)$ es regular, sus autovalores finitos son las raíces de

$$f(\lambda, \epsilon) = \det(H(\lambda) + \epsilon M(\lambda)) = \delta(\lambda)\epsilon^d \tilde{f}(\lambda, \epsilon),$$

²De hecho, es fácil ver que cualquier determinación puede ser utilizada.

donde

$$\tilde{f}(\lambda, \epsilon) = \det(\widehat{\Delta}_{\lambda_0}(\lambda) + \tilde{G}(\lambda, \epsilon))$$

y

$$\tilde{G}(\lambda, \epsilon) = \begin{bmatrix} \epsilon G_{11}(\lambda) & \epsilon G_{12}(\lambda) & G_{13}(\lambda) \\ \epsilon G_{21}(\lambda) & I + \epsilon G_{22}(\lambda) & G_{23}(\lambda) \\ \epsilon G_{31}(\lambda) & \epsilon G_{32}(\lambda) & G_{33}(\lambda) \end{bmatrix}.$$

Además, la función $\delta(\lambda)$ está dada por $\delta(\lambda) = p(\lambda)q(\lambda)$, donde $\det(P_{\lambda_0}(\lambda)) = 1/p(\lambda)$ y $\det(Q_{\lambda_0}(\lambda)) = 1/q(\lambda)$. Luego $\delta(\lambda)$ es un polinomio tal que $\delta(0) \neq 0$ y que no depende de la perturbación $M(\lambda)$. Estos hechos implican que, para $\epsilon \neq 0$, el haz $H(\lambda) + \epsilon M(\lambda)$ es regular si y sólo si $\tilde{f}(\lambda, \epsilon) \neq 0$, y que, en este caso, los autovalores de $H(\lambda) + \epsilon M(\lambda)$ cuyo límite es $\lambda_0 = 0$ cuando ϵ tiende a cero son las raíces, $\lambda(\epsilon)$, de $\tilde{f}(\lambda, \epsilon)$ cuyo límite es 0. Obviamente (véase (4.16)), $\tilde{f}(\lambda, \epsilon)$ es una función racional en λ , donde los coeficientes del numerador son polinomios en ϵ , y el denominador es precisamente $\delta(\lambda)$. Luego, $\tilde{f}(\lambda, \epsilon)$ puede ser visto como un polinomio en ϵ cuyos coeficientes son funciones racionales en λ . Estudiemos más detenidamente la función $\tilde{f}(\lambda, \epsilon)$.

En primer lugar, nótese que, de acuerdo con el Lema 4.4.2 y las definiciones (4.19) y (4.21),

$$\Phi_1 = \begin{bmatrix} G_{11}(0) & G_{13}(0) \\ G_{31}(0) & G_{33}(0) \end{bmatrix}, \quad \text{y} \quad \Phi_{q+1} = G_{33}(0). \quad (4.26)$$

En segundo lugar, renombramos las dimensiones de los bloques de Jordan asociados con $\lambda_0 = 0$

$$\underbrace{\{n_1, \dots, n_1\}}_{r_1}, \dots, \underbrace{\{n_q, \dots, n_q\}}_{r_q} \equiv \{m_1, \dots, m_g\}.$$

Ahora hacemos uso del Lema 6.4.1 sobre los determinantes del tipo $\det(D + G)$ con D diagonal, para desarrollar $\tilde{f}(\lambda, \epsilon)$ como

$$\tilde{f}(\lambda, \epsilon) = \det \tilde{G}(\lambda, \epsilon) + \sum \lambda^{m_{\nu_1}} \dots \lambda^{m_{\nu_r}} \det \tilde{G}(\lambda, \epsilon)(\{\nu_1, \dots, \nu_r\}'), \quad (4.27)$$

donde, para cada matriz C , $C(\{\nu_1, \dots, \nu_r\}')$ denota la matriz obtenida tras eliminar de C las filas y columnas con índices ν_1, \dots, ν_r . La suma varía entre todos los índices $r \in \{1, \dots, g\}$ y todos los ν_1, \dots, ν_r tales que $1 \leq \nu_1 < \dots < \nu_r \leq g$. Finalmente, nótese que

$$\det \tilde{G}(\lambda, \epsilon) = \epsilon^g (\det \Phi_1 + Q_0(\lambda, \epsilon)), \quad (4.28)$$

para $Q_0(\lambda, \epsilon)$ racional con $Q_0(0, 0) = 0$, y

$$\det \tilde{G}(\lambda, \epsilon)(\{\nu_1, \dots, \nu_r\}') = \epsilon^{g-r} (\det \Phi_1(\{\nu_1, \dots, \nu_r\}') + Q_{\nu_1 \dots \nu_r}(\lambda, \epsilon)), \quad (4.29)$$

con $Q_{\nu_1 \dots \nu_r}$ racional y $Q_{\nu_1 \dots \nu_r}(0, 0) = 0$. A partir de aquí, la demostración sigue los argumentos de [46, pp. 799-800] y está basada en la técnica del Polígono de Newton que hemos expuesto en el Capítulo 3. Nótese, en primer lugar, que las ecuaciones (4.27-4.28-4.29) muestran que $\tilde{f}(\lambda, \epsilon) \neq 0$, ya que $\det \Phi_{j+1} = \det \Phi_1(\{1, \dots, \sum_{i=1}^j r_i\}') \neq 0$ es el coeficiente de $\epsilon^{f_{j+1}} \lambda^{r_1 n_1 + \dots + r_j n_j}$ en el desarrollo de Taylor en dos variables de $\tilde{f}(\lambda, \epsilon)$ (con f_{j+1} como se ha definido en (4.13)). Recordemos, por otra parte, que el punto (a, b) aparece en la nube de puntos de $\tilde{f}(\lambda, \epsilon)$ si y sólo si el coeficiente de $\lambda^a \epsilon^b$ en $\tilde{f}(\lambda, \epsilon)$ es no nulo. Si definimos los puntos

$$P_j = \left(\sum_{i=1}^{j-1} r_i n_i, f_j \right), \quad j = 1, \dots, q,$$

junto con $P_{q+1} = (a, 0)$, el razonamiento empleado en [46, pp. 800] nos permite concluir que el Polígono de Newton de $\tilde{f}(\lambda, \epsilon)$ no puede quedar bajo el Polígono formado por los puntos P_j (fijado un valor, f_j , para el exponente de ϵ , el menor valor posible para el exponente de λ en $\tilde{f}(\lambda, \epsilon)$ es, precisamente, la primera coordenada de P_j). Esto significa que si P_j forma parte de la nube de puntos, será necesariamente un punto del Polígono de Newton de $\tilde{f}(\lambda, \epsilon)$. Por otro lado, P_j aparece en la nube de puntos si y sólo si $\det \Phi_j \neq 0$. Ahora los resultados son consecuencia inmediata de la teoría expuesta en el Capítulo 3 y el Lema 4.5.2. Obsérvese que la pendiente del segmento que une P_j y P_{j+1} es, precisamente, $-1/n_j$. ■

Obviamente, la hipótesis $\det \Phi_{j+1} \neq 0$ del Teorema 4.5.3 es una condición genérica sobre el conjunto de perturbaciones $M(\lambda) = B_0 + \lambda B_1$, porque si se fija $H(\lambda)$, entonces $\det \Phi_{j+1}$ es un polinomio de varias variables en las entradas de B_0 y B_1 . En cambio, la hipótesis $\det \Phi_{j+1} \neq 0$ es distinta de la hipótesis del Teorema 4.3.2, es decir: $\det(U_2(\lambda) M(\lambda) V_2(\lambda)) \neq 0$. La razón es que el Teorema 4.5.3 se refiere únicamente a un autovalor del haz no perturbado $H(\lambda)$, mientras que el Teorema 4.3.2 hace referencia simultáneamente a todos los autovalores de $H(\lambda)$. Nótese también que, aunque la multiplicidad algebraica de λ_0 en $H(\lambda)$ es $r_1 n_1 + \dots + r_q n_q$, la condición $\det \Phi_{j+1} \neq 0$ del Teorema 4.5.3 sólo garantiza la existencia de $r_j n_j$ desarrollos con coeficientes y exponentes directores como los de (4.24). Para concluir esta breve discusión, destacamos que la condición $\det \Phi_{q+1} \neq 0$ implica la condición $\det(U_2(\lambda) M(\lambda) V_2(\lambda)) \neq 0$. Esto se sigue fácilmente de (4.21) y el Lema 4.2.2. Por lo tanto, $\det \Phi_{q+1} \neq 0$ para un solo autovalor garantiza la existencia de desarrollos para todos los autovalores, aunque no necesariamente del tipo (4.24).

Ilustramos el Teorema 4.5.3 con el siguiente ejemplo.

Ejemplo 4.5.4 Continuamos con el Ejemplo 4.5.1. El hecho de que $\det \Phi_3 \neq 0$ garantiza la existencia de dos desarrollos con exponente director $1/2$ y límite 1 cuando ϵ tiende a cero. Para obtener los coeficientes directores de estos desarrollos, hemos de resolver la ecuación

$$\det \begin{bmatrix} 2 + \zeta & 6 \\ 3 & 1 \end{bmatrix} = 0.$$

Las dos raíces cuadradas de su solución, $\zeta = 16$, nos dan los coeficientes directores de los desarrollos con exponente director $1/2$:

$$\begin{aligned} \lambda_1(\epsilon) &= 1 + 4\epsilon^{1/2} + o(\epsilon^{1/2}) \\ \lambda_2(\epsilon) &= 1 - 4\epsilon^{1/2} + o(\epsilon^{1/2}). \end{aligned}$$

De modo similar, $\det \Phi_2 \neq 0$ garantiza la existencia de un desarrollo con exponente director 1 y límite 1 cuando ϵ tiende a cero. El coeficiente director del desarrollo es la raíz de la ecuación

$$\det \begin{bmatrix} 2 + \zeta & 3 & 5 \\ 4 & 2 & 6 \\ 0 & 3 & 1 \end{bmatrix} = 0,$$

luego

$$\lambda_3(\epsilon) = 1 + \epsilon + o(\epsilon).$$

A modo de comparación, hemos calculado los autovalores del haz $H(\lambda) + \epsilon M(\lambda)$, para $\epsilon = 10^{-4}$, 10^{-6} , 10^{-8} , resolviendo la ecuación polinómica $\det(H(\lambda) + \epsilon M(\lambda)) = 0$ con la precisión aritmética del programa MATLAB 7.0 (64 dígitos decimales de precisión), y redondeando los resultados a diez dígitos. Las tres raíces más próximas a 1 son:

	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-8}$
λ_1	1,053399042	1,004079394	1,000400768
λ_2	0,9657365454	0,9960738628	0,9996007623
λ_3	1,000099915	1,000001000	1,000000010

El lector puede comprobar que los resultados coinciden con los predichos por la teoría de perturbación, hasta el orden correspondiente. \square

4.5.3. El autovalor infinito

Aunque el autovalor infinito ha sido excluido de nuestro análisis previo, puede ser incluido fácilmente considerando el autovalor cero del haz dual

$$H^\sharp(\lambda) + \epsilon M^\sharp(\lambda) := A_1 + \lambda A_0 + \epsilon(B_1 + \lambda B_0).$$

A partir de la FCK podemos comprobar de manera sencilla una serie de propiedades: si $\mu(\epsilon) \neq 0$ es un autovalor de $H^\sharp(\lambda) + \epsilon M^\sharp(\lambda)$, entonces $\lambda(\epsilon) = \mu(\epsilon)^{-1}$ es un autovalor del haz original $H(\lambda) + \epsilon M(\lambda)$. Recíprocamente, cada autovalor finito, $\lambda(\epsilon)$, de $H(\lambda) + \epsilon M(\lambda)$ es de la forma $\mu(\epsilon)^{-1}$ para un autovalor $\mu(\epsilon)$ de $H^\sharp(\lambda) + \epsilon M^\sharp(\lambda)$, y los datos espectrales (autovectores, número de bloques de Jordan, multiplicidades parciales, etc.) son los mismos en ambos casos. Los subespacios reductores minimales de un haz y su dual son iguales. Dada la FCK (2.5) de $H(\lambda)$, las columnas de X e Y correspondientes a los bloques de Jordan “infinitos” son las columnas asociadas con los bloques de Jordan del autovalor cero en la FCK del haz dual.

Si el autovalor cero $\mu_0 = 0$ de $H^\sharp(\lambda)$ tiene estructura espectral (4.9) en $H^\sharp(\lambda)$, entonces podemos definir las matrices Φ_j^∞ , $j = 1, \dots, q+1$, para el autovalor infinito de $H(\lambda)$ como las matrices Φ_j correspondientes al autovalor cero de $H^\sharp(\lambda)$. Además, podemos usar las matrices X e Y de la FCK de $H(\lambda)$ para construir estas matrices Φ_j^∞ .

Por lo tanto, para obtener los desarrollos de Puiseux de los autovalores $\lambda(\epsilon)$ que proceden del infinito, simplemente aplicamos el Teorema 4.5.3 anterior a los autovalores $\mu(\epsilon)$ de $H^\sharp(\lambda) + \epsilon M^\sharp(\lambda)$ con $\mu(0) = 0$, y calculamos el término director de $\mu(\epsilon)^{-1}$. Esto nos lleva al siguiente resultado.

Corolario 4.5.5 *Sea $H(\lambda)$ un haz matricial $n \times n$ con forma de Kronecker (2.5) y $M(\lambda)$ otro haz de la misma dimensión. Sea $\mu_0 = 0$ un autovalor de $H^\sharp(\lambda)$ con estructura espectral dada por (4.9) y (4.10). Sean Φ_j^∞ y E_j , $j = 1, \dots, q$, las matrices definidas en (4.20) y (4.23), y Φ_{q+1}^∞ la matriz definida en (4.21), para el autovalor cero del haz dual $H^\sharp(\lambda)$. Si $\det \Phi_{j+1}^\infty \neq 0$ para algún $j \in \{1, 2, \dots, q\}$, sean ξ_1, \dots, ξ_{r_j} los r_j autovalores finitos del haz $\Phi_j^\infty + \zeta E_j$, y $(\xi_r)_s^{1/n_j}$, $s = 1, \dots, n_j$, las n_j determinaciones complejas de la raíz n_j -ésima. Entonces, en un entorno de $\epsilon = 0$, el haz $H(\lambda) + \epsilon M(\lambda)$ tiene $r_j n_j$ autovalores que satisfacen*

$$\lambda_j^{rs}(\epsilon) = (\xi_r)_s^{-1/n_j} \epsilon^{-1/n_j} + o(\epsilon^{-1/n_j}), \quad r = 1, 2, \dots, r_j, \quad s = 1, 2, \dots, n_j, \quad (4.30)$$

donde ϵ^{1/n_j} es la determinación principal de la raíz n_j -ésima de ϵ . Si, además, $\det \Phi_j^\infty \neq 0$, entonces todos los valores ξ_r de (4.30) son no nulos, y los desarrollos (4.30) son todos los desarrollos con exponente director $-1/n_j$.

4.5.4. Desarrollos para autovalores simples

Los desarrollos del Teorema 4.5.3 y del Corolario 4.5.5 dependen de las matrices Φ_j definidas en (4.20), y estas matrices se construyen utilizando unos vectores específicos de los espacios nulos de $H(\lambda_0)$. Estos vectores son construídos fácilmente a partir de las matrices X e Y que transforman $H(\lambda)$ en su FCK (2.5), pero *las matrices X e Y (o los bloques que necesitamos) son muy difíciles de obtener en presencia de autovalores múltiples defectivos*. Este no es el caso para autovalores simples, porque entonces podemos usar cualquier base de los espacios nulos derecho e izquierdo de $H(\lambda_0)$ para construir las correspondientes matrices. Esto es lo que se muestra en esta sección.

Teorema 4.5.6 Sean $H(\lambda) = A_0 + \lambda A_1$ un haz matricial arbitrario $n \times n$ (singular o no), $M(\lambda) = B_0 + \lambda B_1$ otro haz con la misma dimensión, y λ_0 un autovalor finito simple de $H(\lambda)$. Denotemos por W a una matriz cuyas columnas forman una base del espacio nulo izquierdo de $H(\lambda_0)$ y por Z a una matriz cuyas columnas forman una base del espacio nulo derecho de $H(\lambda_0)$. Entonces

1. El haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es genéricamente regular y tiene un único autovalor finito, es decir, esto es cierto para todos los haces $M(\lambda)$ excepto aquellos que se encuentran en una variedad algebraica de codimensión positiva.
2. Si el haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es regular y tiene un único autovalor finito igual a ξ , entonces hay un único autovalor de $H(\lambda) + \epsilon M(\lambda)$ tal que

$$\lambda(\epsilon) = \lambda_0 + \xi \epsilon + O(\epsilon^2),$$

cuando ϵ tiende a cero.

3. Además, si $H(\lambda)$ es regular, entonces $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es 1×1 y es regular con un único autovalor finito, para todas las perturbaciones $M(\lambda)$. Por lo tanto, $\xi = -(W^T M(\lambda_0) Z) / W^T A_1 Z$.

Demostración. Las propiedades espectrales, en particular los autovalores, del haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ son los mismos para cualquier par de bases W y Z de los espacios nulos derecho e izquierdo de $H(\lambda_0)$, porque el cambio de bases implica simplemente una equivalencia estricta de dicho haz. Por lo tanto, podemos elegir un par de bases específicas para demostrar el teorema. Para este propósito, sean \mathcal{R} y \mathcal{R}_T , respectivamente, los subespacios singulares derecho e izquierdo de $H(\lambda)$, y sean $\mathcal{N}(H(\lambda_0))$ y $\mathcal{N}_T(H(\lambda_0))$ los espacios nulos derecho e izquierdo de la matriz $H(\lambda_0)$. Denotemos por $W_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ a una matriz cuyas columnas forman una base de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$, y por $Z_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ a una matriz cuyas columnas forman una base de $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$. Ahora, consideremos la FCK (2.5) de $H(\lambda)$ y las particiones (4.6) y (4.5) de X e Y , y observemos que X_{λ_0} e Y_{λ_0} tienen solamente una columna porque λ_0 es simple. A partir de la FCK y el Lema 4.2.1, es fácil ver que las columnas de $[Y_{\lambda_0} W_{\mathcal{R}}]$ forman una base de $\mathcal{N}_T(H(\lambda_0))$, y las columnas de $[X_{\lambda_0} Z_{\mathcal{R}}]$ forman una base de $\mathcal{N}(H(\lambda_0))$. Además, nótese que la estructura espectral (4.9) es, simplemente: $q = 1$, $n_1 = 1$, y $r_1 = 1$, y que, en este caso, las matrices $\Phi_1, \dots, \Phi_{q+1}$ definidas en (4.20-4.21) son exactamente dos, concretamente

$$\Phi_1 = [Y_{\lambda_0} \quad W_{\mathcal{R}}]^T M(\lambda_0) [X_{\lambda_0} \quad Z_{\mathcal{R}}], \quad \text{y} \quad \Phi_2 = W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}.$$

Si el haz es regular, entonces Φ_1 es 1×1 y Φ_2 no existe.

Elijamos $W = [Y_{\lambda_0} \ W_{\mathcal{R}}]$ y $Z = [X_{\lambda_0} \ Z_{\mathcal{R}}]$. De nuevo, por (2.5) y el Lema 4.2.1,

$$W^T A_1 Z = \begin{bmatrix} 1 & 0 \\ 0 & 0_{d \times d} \end{bmatrix}. \quad (4.31)$$

Nótese que esta matriz es E_1 , de acuerdo con (4.23). Por tanto,

$$W^T M(\lambda_0) Z + \zeta W^T A_1 Z = \begin{bmatrix} Y_{\lambda_0}^T M(\lambda_0) X_{\lambda_0} + \zeta & Y_{\lambda_0}^T M(\lambda_0) Z_{\mathcal{R}} \\ W_{\mathcal{R}}^T M(\lambda_0) X_{\lambda_0} & W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}} \end{bmatrix}. \quad (4.32)$$

El desarrollo de Laplace a lo largo de la primera columna nos da

$$\det(W^T M(\lambda_0) Z + \zeta W^T A_1 Z) = (Y_{\lambda_0}^T M(\lambda_0) X_{\lambda_0} + \zeta) \det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}) + b,$$

donde b es una constante independiente de ζ . Esta ecuación muestra que $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es regular y tiene un único autovalor finito si y sólo si $\det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}) \neq 0$. Claramente, esta condición es genérica porque $\det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}})$ es un polinomio de varias variables en las entradas de B_0 y B_1 . Esto demuestra el primer apartado del teorema. En el caso regular se tiene que $W^T A_1 Z = 1$, luego el haz es 1×1 , regular y con un único autovalor para cualquier $M(\lambda)$.

Para demostrar el segundo apartado simplemente ha de notarse que la condición $\det \Phi_2 = \det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}) \neq 0$ permite aplicar el Teorema 4.5.3, y que (4.32) es $\Phi_1 + \zeta E_1$. El único punto que hay que discutir es que ahora tenemos $O(\epsilon^2)$ (O grande!) mientras que en (4.24) se tiene $o(\epsilon)$ (o pequeña!). Esto es una simple consecuencia de la Teoría de Funciones Algebraicas. Nótese que, haciendo uso del Lema 4.2.2, $\det \Phi_2 = \det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}) \neq 0$ implica que $\det(P_2(\lambda_0) M(\lambda_0) Q_2(\lambda_0)) \neq 0$, de donde se tiene que $\det(P_2(\lambda) M(\lambda) Q_2(\lambda)) \neq 0$ en (4.15). Luego λ_0 es una raíz simple de (4.15), y $\lambda(\epsilon)$ es analítica en ϵ y única en un entorno de $\epsilon = 0$.

Finalmente, el apartado 3 es una simple consecuencia de los comentarios previos. ■

El Teorema 4.5.6 nos permite obtener los desarrollos de primer orden de la perturbación de los autovalores, y comprobar su existencia, usando bases arbitrarias de los espacios nulos derecho e izquierdo de la matriz $H(\lambda_0)$. El cálculo de estas bases es una tarea elemental de álgebra lineal. Con una elección particular de las bases puede obtenerse una expresión explícita de ξ . Esto es lo que se hace en el Corolario 4.5.7. En cambio, debe observarse que esta expresión requiere conocer los subespacios $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$ y $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$, algo que sólo es posible con un trabajo adicional.

Corolario 4.5.7 Sean $H(\lambda) = A_0 + \lambda A_1$ un haz matricial arbitrario $n \times n$ (singular o no), $M(\lambda) = B_0 + \lambda B_1$ otro haz de la misma dimensión y λ_0 un autovalor finito simple de $H(\lambda)$. Sean \mathcal{R} y \mathcal{R}_T , respectivamente, los subespacios singulares derecho e izquierdo de $H(\lambda)$, y sean $\mathcal{N}(H(\lambda_0))$ y $\mathcal{N}_T(H(\lambda_0))$ los subespacios nulos derecho e izquierdo de la matriz $H(\lambda_0)$. Denotemos por $W_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ a una matriz cuyas columnas son una base de $\mathcal{N}_T(H(\lambda_0)) \cap \mathcal{R}_T$, y por $Z_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ a una matriz cuyas columnas forman una base de $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$, y construyamos, a partir de estas matrices, las matrices

i) $W = [w \ W_{\mathcal{R}}]$ cuyas columnas forman una base de $\mathcal{N}_T(H(\lambda_0))$, y

ii) $Z = [z \ Z_{\mathcal{R}}]$ cuyas columnas forman una base de $\mathcal{N}(H(\lambda_0))$.

Si $\det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}) \neq 0$ entonces hay un único autovalor de $H(\lambda) + \epsilon M(\lambda)$ tal que

$$\lambda(\epsilon) = \lambda_0 - \frac{\det(W^T M(\lambda_0) Z)}{(w^T A_1 z) \cdot \det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}})} \epsilon + O(\epsilon^2),$$

cuando ϵ tiende a cero.

Demostración. Usando las matrices que aparecen en la demostración del Teorema 4.5.6, es obvio que

$$\begin{bmatrix} w & W_{\mathcal{R}} \end{bmatrix}^T = \begin{bmatrix} s_{11} & S_{12} \\ 0 & I_d \end{bmatrix} \begin{bmatrix} Y_{\lambda_0} \\ W_{\mathcal{R}} \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} z & Z_{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} X_{\lambda_0} & Z_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} t_{11} & 0 \\ T_{21} & I_d \end{bmatrix}.$$

Así, de (4.31)

$$\begin{bmatrix} w & W_{\mathcal{R}} \end{bmatrix}^T A_1 \begin{bmatrix} z & Z_{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} w^T A_1 z & 0 \\ 0 & 0_{d \times d} \end{bmatrix},$$

y, en este caso,

$$W^T M(\lambda_0) Z + \zeta W^T A_1 Z = \begin{bmatrix} w^T M(\lambda_0) z + \zeta(w^T A_1 z) & w^T M(\lambda_0) Z_{\mathcal{R}} \\ W_{\mathcal{R}}^T M(\lambda_0) z & W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}} \end{bmatrix}.$$

Finalmente,

$$\begin{aligned} & (W^T M(\lambda_0) Z + \zeta W^T A_1 Z) \begin{bmatrix} 1 & 0 \\ -(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}})^{-1} W_{\mathcal{R}}^T M(\lambda_0) z & I_d \end{bmatrix} = \\ & \begin{bmatrix} w^T M(\lambda_0) z - w^T M(\lambda_0) Z_{\mathcal{R}} (W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}})^{-1} W_{\mathcal{R}}^T M(\lambda_0) z + \zeta(w^T A_1 z) & w^T M(\lambda_0) Z_{\mathcal{R}} \\ 0 & W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}} \end{bmatrix}. \end{aligned}$$

El resultado se sigue igualando el determinante a cero y observando que

$$\det(W^T M(\lambda_0) Z) = (w^T M(\lambda_0) z - w^T M(\lambda_0) Z_{\mathcal{R}} (W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}})^{-1} W_{\mathcal{R}}^T M(\lambda_0) z) \det(W_{\mathcal{R}}^T M(\lambda_0) Z_{\mathcal{R}}).$$

■

4.6. Autovectores del haz perturbado

Hemos comentado ya que los autovectores no están definidos para haces singulares, incluso en el caso de autovalores simples. Así pues, una teoría de perturbación de autovectores carece de sentido. En cambio, para $\epsilon \neq 0$, el haz perturbado (4.3) es genéricamente regular, tiene autovalores simples, y autovectores perfectamente definidos. Para valores pequeños de ϵ , es natural esperar que los autovectores correspondientes a estos autovalores (4.3) cuyos límites son los autovalores de $H(\lambda)$ están relacionados con ciertas propiedades de $H(\lambda)$. Dado un autovalor finito λ_0 de $H(\lambda)$, en esta sección mostraremos que, genéricamente, los autovectores de (4.3) correspondientes a autovalores $\lambda(\epsilon)$ tales que $\lambda(0) = \lambda_0$ satisfacen tres propiedades: (i) pueden ser desarrollados en serie de Puiseux, $v(\epsilon)$, de manera que el vector $v(0) \neq 0$; (ii) $v(0)$ está en el espacio nulo de $H(\lambda_0)$; y (iii) dentro de este subespacio nulo, $v(0)$ está completamente determinado por la perturbación $M(\lambda)$. Además, mostraremos cómo determinar $v(0)$. Así, $v(0)$ es un autovector aproximado de (4.3) para valores pequeños de $\epsilon \neq 0$, pero no tiene significado especial en $H(\lambda)$ excepto que se trata de un vector de $\mathcal{N}(H(\lambda_0))$. Hablando de modo informal, puede decirse que cada perturbación $M(\lambda)$ selecciona una dirección distinta en el espacio nulo de $H(\lambda_0)$ como un autovector aproximado de $\lambda(\epsilon)$. En aras de la brevedad, nos centraremos en autovectores derechos. Es inmediato deducir resultados similares para los autovectores izquierdos. Como en el caso de autovalores, los resultados cuando λ_0 es un autovalor simple de $H(\lambda)$ son más sencillos e independientes de la normalización de las bases.

Ha de ser observado que nos encontramos en una situación diferente a la de los desarrollos de los autovalores (4.24): en (4.24) el término de orden cero, λ_0 , era conocido y nuestra tarea consistía en determinar el término siguiente, mientras que en el caso de autovectores queremos determinar, precisamente, el término de orden cero. De hecho, los resultados que presentamos carecen de sentido para autovalores simples de haces regulares, porque entonces el término de orden cero es obvio.

En los desarrollos de esta sección supondremos que se da la condición genérica $\det \Phi_{q+1} \neq 0$. Esta condición se puede relajar con el inconveniente de complicar la demostración del siguiente lema de existencia.

Lema 4.6.1 *Consideremos las mismas notación e hipótesis que en el Teorema 4.5.3 junto con las condiciones $\det \Phi_{q+1} \neq 0$, $\det \Phi_j \cdot \det \Phi_{j+1} \neq 0$, y que los r_j autovalores finitos del haz $\Phi_j + \zeta E_j$, ξ_1, \dots, ξ_{r_j} , son distintos. Entonces, cada autovalor perturbado de la forma (4.24) definido en un entorno de $\epsilon = 0$, tiene asociado, en el mismo entorno y para $\epsilon \neq 0$, un autovector derecho del haz regular $H(\lambda) + \epsilon M(\lambda)$ que es de la forma*

$$v_j^{rs}(\epsilon) = v_j^{rs} + \sum_{k=1}^{\infty} u_{jk}^{rs} \epsilon^{k/n_j}. \quad (4.33)$$

Demostración. Haremos simplemente un esbozo de la demostración. Nótese que las hipótesis $\det \Phi_j \cdot \det \Phi_{j+1} \neq 0$ y la de que ξ_1, \dots, ξ_{r_j} son distintos implican que los autovalores de (4.24) son simples para $\epsilon \neq 0$ suficientemente pequeño. Consideremos, sin pérdida de generalidad, que $\lambda_0 = 0$ como en la demostración del Teorema 4.5.3. Procederemos a partir de la igualdad (4.25). Para $\epsilon \neq 0$, los autovalores y autovectores derechos de $\widehat{\Delta}_{\lambda_0}(\lambda) + G(\lambda, \epsilon)$ coinciden con los autovalores y autovectores derechos de

$$F(\lambda, \epsilon) = \text{diag}(I_g, I, (1/\epsilon)I_d) (\widehat{\Delta}_{\lambda_0}(\lambda) + G(\lambda, \epsilon)),$$

y

$$F(\lambda, 0) = \begin{bmatrix} D(\lambda) & & \\ & I & \\ G_{31}(\lambda) & G_{32}(\lambda) & G_{33}(\lambda) \end{bmatrix},$$

que satisface $\det F(\lambda, 0) \neq 0$ por (4.26). Por lo tanto, $F(\lambda, \epsilon)$ es una función matricial analítica que es regular en $\epsilon = 0$, luego la variación de los autovalores de $F(\lambda, \epsilon)$ con el parámetro ϵ es un problema de perturbación regular de una función matricial analítica. Teniendo en cuenta que $P_{\lambda_0}(\lambda)$ y $Q_{\lambda_0}(\lambda)$ son invertibles y analíticas en un entorno de λ_0 , los autovalores de $H(\lambda) + \epsilon M(\lambda)$ y $F(\lambda, \epsilon)$ son los mismos en un entorno de λ_0 para $\epsilon \neq 0$ y, en particular, los desarrollos de (4.24) son autovalores de $F(\lambda, \epsilon)$. El Lema 2 de [42] se puede aplicar para demostrar que $F(\lambda, \epsilon)$ tiene autovectores derechos correspondientes, $w_j^{rs}(\epsilon)$, del tipo (4.33). Finalmente, para $\epsilon \neq 0$ los autovectores derechos (4.33) de $H(\lambda) + \epsilon M(\lambda)$ correspondientes a los autovalores (4.24) son $Q(\lambda_j^{rs}(\epsilon))w_j^{rs}(\epsilon)$. ■

A continuación, presentamos el resultado principal de esta sección, el Teorema 4.6.2, que determina el término de orden cero, v_j^{rs} , de los desarrollos (4.33). Ha de observarse que este teorema, de hecho, muestra que v_j^{rs} no depende de s , es decir, fijado ξ_r en (4.24), los n_j autovectores de los autovalores correspondientes a las determinaciones de las raíces n_j -ésimas $(\xi_r)_s^{1/n_j}$ tienen el mismo término de orden cero. Nótese también que en la ecuación (4.34) aparece un símbolo O grande.

Teorema 4.6.2 Sean $H(\lambda)$ un haz matricial arbitrario $n \times n$ (singular o no) con forma de Kronecker (2.5), y $M(\lambda)$ otro haz con la misma dimensión. Sea λ_0 un autovalor finito de $H(\lambda)$ con estructura espectral dada por (4.9) y (4.10). Sean Z_j , $j = 1, \dots, q$, las matrices definidas en (4.12), y $Z_{\mathcal{R}} \in \mathbb{C}^{n \times d}$ una matriz cuyas columnas forman una base de $\mathcal{N}(H(\lambda_0)) \cap \mathcal{R}$. Sean Φ_j y E_j , $j = 1, \dots, q$, las matrices definidas en (4.20) y (4.23), y Φ_{q+1} la matriz definida en (4.21). Si $\det \Phi_{q+1} \neq 0$, $\det \Phi_j \cdot \det \Phi_{j+1} \neq 0$ para algún $j \in \{1, 2, \dots, q\}$, y los r_j autovalores finitos del haz $\Phi_j + \zeta E_j$, ξ_1, \dots, ξ_{r_j} , son distintos y tienen autovectores c_1, \dots, c_{r_j} , entonces, en un entorno punteado $0 < |\epsilon| < b$, los autovectores del haz regular $H(\lambda) + \epsilon M(\lambda)$ correspondientes a sus $r_j n_j$ autovalores (4.24) satisfacen

$$v_j^{rs}(\epsilon) = [Z_j \ Z_{\mathcal{R}}] c_r + O(\epsilon^{1/n_j}), \quad r = 1, 2, \dots, r_j, \quad s = 1, 2, \dots, n_j. \quad (4.34)$$

Demostración. Para cada autovalor $\lambda_j^{rs}(\epsilon)$ de (4.24), consideramos, para $\epsilon \neq 0$, el correspondiente autovector $v_j^{rs}(\epsilon)$ dado por (4.33). Por brevedad, omitimos los superíndices y escribimos λ_j y v_j en lugar de λ_j^{rs} y v_j^{rs} . También tomamos $\lambda_0 = 0$ como en la demostración del Teorema 4.5.3. De nuevo, la demostración se basa en la forma local de Smith del Lema 4.4.2, que está bien definida y es analítica en un entorno de $\lambda_0 = 0$. Para hacer uso de esta forma local de Smith reemplazamos $v_j(\epsilon)$ por

$$w_j(\epsilon) = Q(\lambda_j(\epsilon))^{-1} v_j(\epsilon), \quad (4.35)$$

que satisface

$$[\Delta_{\lambda_0}(\lambda_j(\epsilon)) + \epsilon \widetilde{M}(\lambda_j(\epsilon))] w_j(\epsilon) = 0, \quad (4.36)$$

donde

$$\widetilde{M}(\lambda_j(\epsilon)) = P(\lambda_j(\epsilon)) M(\lambda_j(\epsilon)) Q(\lambda_j(\epsilon)).$$

Nótese que se puede recuperar fácilmente el vector $v_j = v_j(0)$ a partir de $w_j(0)$, ya que $v_j(0) = Q(0)w_j(0)$. Particionamos $\widetilde{M}(\lambda_j(\epsilon))$ por bloques como una matriz 3×3 de acuerdo con los tres bloques diagonales de $\Delta_{\lambda_0}(\lambda)$ (véase la ecuación (2.9)), y denotamos, como en la demostración del Teorema 4.5.3, $[G_{ik}(\lambda_j(\epsilon))]_{i,k=1}^3 \equiv \widetilde{M}(\lambda_j(\epsilon))$. Si particionamos el vector $w_j(\epsilon)$ de acuerdo a la partición anterior, la ecuación (4.36) puede escribirse como

$$\left(\begin{bmatrix} D(\lambda_j(\epsilon)) & & \\ & I & \\ & & 0_{d \times d} \end{bmatrix} + \epsilon \begin{bmatrix} G_{11}(\lambda_j(\epsilon)) & G_{12}(\lambda_j(\epsilon)) & G_{13}(\lambda_j(\epsilon)) \\ G_{21}(\lambda_j(\epsilon)) & G_{22}(\lambda_j(\epsilon)) & G_{23}(\lambda_j(\epsilon)) \\ G_{31}(\lambda_j(\epsilon)) & G_{32}(\lambda_j(\epsilon)) & G_{33}(\lambda_j(\epsilon)) \end{bmatrix} \right) \begin{bmatrix} w_j^{(1)}(\epsilon) \\ w_j^{(2)}(\epsilon) \\ w_j^{(3)}(\epsilon) \end{bmatrix} = 0. \quad (4.37)$$

Para $\epsilon = 0$ esta ecuación se reduce a $w_j^{(2)}(0) = 0$. Las filas correspondientes a los primeros bloques son

$$D(\lambda_j(\epsilon)) w_j^{(1)}(\epsilon) + \epsilon(G_{11}(\lambda_j(\epsilon)) w_j^{(1)}(\epsilon) + G_{12}(\lambda_j(\epsilon)) w_j^{(2)}(\epsilon) + G_{13}(\lambda_j(\epsilon)) w_j^{(3)}(\epsilon)) = 0. \quad (4.38)$$

Nótese que los términos de orden inferior en ϵ de las entradas de $D(\lambda_j(\epsilon))$ son del tipo $c \epsilon^{n_i/n_j}$, para $i = 1, \dots, q$, con $c \neq 0$, porque $\det \Phi_j \cdot \det \Phi_{j+1} \neq 0$. Luego, teniendo en cuenta (4.10), podemos dividir las primeras r_1 ecuaciones de (4.38) entre ϵ^{n_1/n_j} , tomar el límite $\epsilon \rightarrow 0$, y concluir que $w_{j,k}(0) = 0$ para $k \leq r_1$ (aquí $w_{j,k}(0)$ denota la k -ésima entrada de $w_j(0)$). Dividiendo entre ϵ^{n_2/n_j} las siguientes r_2 ecuaciones de (4.38) y tomando límites llegamos a que $w_{j,k}(0) = 0$ para $k \leq r_1 + r_2$. Este proceso continúa dividiendo sucesivamente entre $\epsilon^{n_3/n_j}, \dots, \epsilon^{n_{j-1}/n_j}$ para demostrar que $w_{j,k}(0) = 0$ para $k \leq r_1 + \dots + r_{j-1}$.

Denotemos por $\bar{w}_j(0)$ el vector obtenido eliminando de $w_j(0)$ las entradas nulas correspondientes a $w_j^{(2)}(0) = 0$ y a $w_{j,k}(0) = 0$ para $k \leq r_1 + \dots + r_{j-1}$. Dividiendo la parte de (4.37) correspondiente a \bar{w}_j entre ϵ , haciendo $\epsilon = 0$, y teniendo en cuenta (4.26), se llega a

$$(\xi_r E_j + \Phi_j) \bar{w}_j(0) = 0.$$

Ahora, el resultado se sigue de (4.18) y (4.35). ■

4.6.1. El caso de autovalores simples

Destacamos en este apartado el caso en que λ_0 es un autovalor simple de $H(\lambda)$. El siguiente resultado complementa al Teorema 4.5.6.

Corolario 4.6.3 *Con las mismas notación e hipótesis que en el Teorema 4.5.6. Si el haz $W^T M(\lambda_0) Z + \zeta W^T A_1 Z$ es regular y tiene un único autovalor finito igual a ξ con autovector asociado c , entonces existe un único autovector de $H(\lambda) + \epsilon M(\lambda)$ tal que $\lambda(\epsilon) = \lambda_0 + \xi \epsilon + O(\epsilon^2)$, cuando ϵ tiende a cero, y para $\epsilon \neq 0$ el correspondiente autovector satisface*

$$v(\epsilon) = Zc + O(\epsilon).$$

Demostración. El resultado es consecuencia directa del Teorema 4.6.2, las igualdades (4.31) y (4.32) y un cambio elemental de bases. ■

Capítulo 5

Perturbaciones de rango bajo de la Forma de Weierstrass

5.1. Introducción

Este capítulo y el siguiente están dedicados al estudio y la descripción del cambio que producen las perturbaciones aditivas de rango bajo en las formas canónicas introducidas en el Capítulo 2. En este capítulo trataremos el cambio de la estructura de Weierstrass de haces regulares y en el siguiente el cambio de la estructura de Kronecker de haces singulares de rango no completo. No trataremos explícitamente el caso matricial, que ya ha sido estudiado en profundidad en [34, 52, 60, 61, 62, 63], pero los resultados de esos trabajos, que son el punto de partida de los que aquí expondremos, se pueden obtener como caso particular de estos últimos. Concretamente, el cambio de la forma de Jordan de una matriz A coincide con el cambio de la forma de Weierstrass del haz regular $A - \lambda I$.

Al perturbar un haz por otro de rango bajo, la forma canónica de Weierstrass puede cambiar de múltiples maneras. A modo de ilustración, incluimos el siguiente ejemplo elemental, que es sólo una muestra de lo que puede ocurrir.

Ejemplo 5.1.1 Sea el haz $A_0 + \lambda A_1 = (\lambda I_2 - J_2(1)) \oplus (\lambda I_2 - J_2(1)) \oplus (\lambda - J_1(0))$. Consideremos las dos perturbaciones siguientes de rango 1:

$$\begin{aligned}
 A_0 + \lambda A_1 + E_1 &= \left[\begin{array}{cc|cc|c} \lambda - 1 & -1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda - 1 & -1 & 0 \\ 0 & 0 & 0 & \lambda - 1 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array} \right] + \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] = \\
 &= \left[\begin{array}{cccc|c} \lambda - 1 & -1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & -1 & 0 & 0 \\ 0 & 0 & \lambda - 1 & -1 & 0 \\ 0 & 0 & 0 & \lambda - 1 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array} \right] \\
 A_0 + \lambda A_1 + E_2 &= \left[\begin{array}{cc|cc|c} \lambda - 1 & -1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda - 1 & -1 & 0 \\ 0 & 0 & 0 & \lambda - 1 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array} \right] + \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{array} \right] =
 \end{aligned}$$

$$= \left[\begin{array}{cc|cc|c} \lambda - 1 & -1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda - 1 & -1 & 0 \\ 0 & 0 & 0 & \lambda - 1 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda - 1 \end{array} \right]$$

En el primer caso la parte correspondiente al autovalor 1 ha pasado de tener dos bloques 2×2 a tener un solo bloque de mayor tamaño (4×4), mientras que la parte correspondiente al autovalor 0 permanece inalterada. Por contra, en el segundo caso, la parte correspondiente al autovalor 0 ha desaparecido, mientras que la parte correspondiente al autovalor 1 ha pasado de tener dos bloques a tener tres. \square

Nuestra intención no es hacer una descripción de todos los cambios posibles, sino obtener el cambio *genérico*, es decir: aquel que se da para la mayoría de las perturbaciones. Por otro lado, la relevancia de nuestros resultados no radica exclusivamente en la descripción de dicho cambio genérico sino también en la caracterización del conjunto de perturbaciones para las que dicho cambio se produce.

Como se ha mencionado en la Introducción de esta memoria, el trabajo que presentamos en este capítulo está motivado por el de J. Moro y F. M. Dopico [52] en el que se describe el cambio genérico de la forma de Jordan de matrices cuadradas tras perturbaciones de rango bajo (y, asimismo, el conjunto de perturbaciones para las que se da este cambio). Dado que será una referencia fundamental a la hora de comparar con los resultados equivalentes para los casos de haces regulares y singulares que se incluyen en esta memoria, enunciamos dicho resultado a continuación:

Teorema 5.1.2 [52] *Sea A una matriz compleja de tamaño $n \times n$ y $\lambda_0 \in \mathbb{C}$ un autovalor de A con multiplicidad geométrica g . Sea B una matriz compleja $n \times n$ con $\text{rg } B \leq g$. Entonces*

“los bloques de Jordan de $A + B$ asociados con λ_0 son exactamente los $g - \text{rg } B$ bloques más pequeños de A asociados con λ_0 si y sólo si $C_0(B) \neq 0$ ”

donde $C_0(X)$ es un cierto polinomio de n^2 variables no idénticamente nulo.

Es importante observar que la frase entrecomillada del Teorema 5.1.2 describe completamente la forma de Jordan asociada con el autovalor λ_0 en la matriz $A + B$ para todas las perturbaciones, B , salvo aquellas que están en una subvariedad propia, y que está definida por la igualdad $C_0(B) = 0$. En [52] se describe el polinomio $C_0(X)$ como una suma de menores de una matriz \tilde{B} que se obtiene a partir de B multiplicándola a la izquierda y a la derecha por unos determinados autovectores de A asociados al autovalor λ_0 . Una construcción muy similar se empleará para describir el conjunto de perturbaciones de un haz matricial regular para el que se da el cambio genérico de la forma de Weierstrass.

Las técnicas que se utilizan en este capítulo para describir el comportamiento genérico son distintas a las empleadas para el caso matricial en [52]. Aquí nos basaremos en resultados previos obtenidos por R. C. Thompson en [76] que permiten acotar las dimensiones de los bloques de Jordan del haz perturbado y sugieren las dimensiones que tendrán, genéricamente, dichos bloques. En cambio, para obtener la descripción de la variedad de perturbaciones en que se da el cambio genérico emplearemos técnicas determinantaes análogas a las de [52], aunque, en el caso que nos ocupa, dicha descripción es bastante más compleja y difícil de obtener. Como se verá posteriormente, las técnicas empleadas

en este capítulo para el caso regular son completamente distintas a las que utilizaremos en el capítulo siguiente para estudiar el caso singular.

A lo largo de este capítulo $A_0 + \lambda A_1$ (*haz no perturbado*) denotará un haz regular de matrices $n \times n$ cuya forma de Weierstrass es conocida y $B_0 + \lambda B_1$ (*haz perturbación*) será otro haz de la misma dimensión y *rango bajo*. La suma de ambos, $A_0 + B_0 + \lambda(A_1 + B_1)$ se denominará *haz perturbado*. Recordemos (Sección 2.1) que el haz $A_0 + \lambda A_1$ puede tener un autovalor *infinito* cuyos bloques de Jordan en la forma canónica de Weierstrass son precisamente los bloques de Jordan asociados al autovalor *cero* en la forma de Weierstrass del haz *dual* $A_1 + \lambda A_0$. Por lo tanto, nos centraremos en los autovalores finitos de $A_0 + \lambda A_1$. Los resultados correspondientes al autovalor infinito se siguen de los resultados para el autovalor cero considerando tanto el dual del haz no perturbado como el del haz perturbado.

Antes de nada, debemos establecer qué entendemos por *rango bajo*. Sea λ_0 un autovalor finito del haz regular $A_0 + \lambda A_1$ con multiplicidad geométrica g . Las desigualdades elementales $\text{rg}(C + D) \leq \text{rg}(C) + \text{rg}(D)$ y $\text{rg}(C) \leq \text{rg}(C + D) + \text{rg}(D)$ se cumplen para cualquier par de matrices C y D de la misma dimensión, y de ellas se deducen fácilmente las desigualdades

$$\begin{aligned} \text{rg}(A_0 + \lambda_0 A_1 + B_0 + \lambda_0 B_1) &\leq \text{rg}(A_0 + \lambda_0 A_1) + \text{rg}(B_0 + \lambda_0 B_1) \\ \text{rg}(A_0 + \lambda_0 A_1) &\leq \text{rg}(A_0 + \lambda_0 A_1 + B_0 + \lambda_0 B_1) + \text{rg}(B_0 + \lambda_0 B_1). \end{aligned}$$

Combinando estas dos desigualdades se obtiene

$$g - \text{rg}(B_0 + \lambda_0 B_1) \leq \dim \ker(A_0 + \lambda_0 A_1 + B_0 + \lambda_0 B_1) \leq g + \text{rg}(B_0 + \lambda_0 B_1). \quad (5.1)$$

La conclusión más importante que extraemos de estas desigualdades es que si

$$\text{rg}(B_0 + \lambda_0 B_1) < g, \quad (5.2)$$

entonces λ_0 es también un autovalor del haz perturbado

$$A_0 + B_0 + \lambda(A_1 + B_1). \quad (5.3)$$

En consecuencia, una perturbación de **rango bajo** será, en toda esta sección, aquella para la cual las matrices coeficientes B_0 y B_1 satisfacen (5.2). Ha de ser notado que esta condición depende del autovalor particular λ_0 que estemos considerando. Es un hecho bien conocido que para un haz regular $H_0 + \lambda H_1$ el número de bloques de Jordan asociados a λ_0 en su forma canónica de Weierstrass es igual a $\dim \ker(H_0 + \lambda_0 H_1)$. Por lo tanto, asumiendo que (5.3) continúa siendo regular, las dos desigualdades de (5.1) implican, por un lado, que la perturbación $B_0 + \lambda B_1$ puede destruir, a lo sumo, $\text{rg}(B_0 + \lambda_0 B_1)$ de los bloques de Jordan de $A_0 + \lambda A_1$ y, por otro, que puede crear, como mucho, $\text{rg}(B_0 + \lambda_0 B_1)$ nuevos bloques de Jordan asociados a cada autovalor finito λ_0 de $A_0 + \lambda A_1$. Esto permite múltiples posibilidades distintas para el número y las dimensiones de los bloques de Jordan que aparecen en la forma de Weierstrass de $A_0 + B_0 + \lambda(A_1 + B_1)$. El objetivo de este capítulo es encontrar el comportamiento más habitual a este respecto.¹

¹La hipótesis de que $A_0 + B_0 + \lambda(A_1 + B_1)$ continúa siendo regular se cumple excepto para elecciones muy particulares de B_0 y B_1 . Estas elecciones están excluidas de nuestro resultado principal, el Teorema 5.4.2, por las condiciones de genericidad. Hablaremos más acerca de esto en la Observación 5.4.3 de la Sección 5.4.

Los resultados que presentamos en este capítulo dependen de dos cantidades *para cada autovalor* λ_0 : $\rho_0 = \text{rg}(B_0 + \lambda_0 B_1)$ y $\rho_1 = \text{rg}(B_1)$. Asumiendo que la condición (5.2) se satisface, demostraremos que, *para la mayoría de las matrices* B_0 y B_1 , se tiene la igualdad

$$\dim \ker(A_0 + \lambda_0 A_1 + B_0 + \lambda_0 B_1) = g - \text{rg}(B_0 + \lambda_0 B_1),$$

y que hay, precisamente, $g - \text{rg}(B_0 + \lambda_0 B_1)$ bloques de Jordan asociados a λ_0 en la forma canónica de Weierstrass del haz perturbado (5.3). Más aún, si denotamos por d_0 el número de bloques de Jordan de $A_0 + \lambda A_1$ asociados al autovalor λ_0 de dimensión mayor que 1, demostraremos que, siempre que $\rho_0 + \rho_1 \leq d_0$, los mayores ρ_0 bloques de Jordan de $A_0 + \lambda A_1$ asociados a λ_0 desaparecen y los siguientes ρ_1 bloques mayores asociados a λ_0 se transforman en bloques de tamaño 1×1 , mientras que el resto de los bloques de Jordan de $A_0 + \lambda A_1$ asociados a λ_0 permanecen como bloques de Jordan del haz perturbado (5.3). En el otro caso, $\rho_0 + \rho_1 > d_0$, habrá sólo bloques 1×1 asociados a λ_0 en la forma de Weierstrass del haz (5.3). Este comportamiento genérico coincide con el que se ha descrito anteriormente para perturbaciones de rango bajo de la forma canónica de Jordan de matrices si $B_1 = 0$, mientras que es ligeramente distinto cuando $B_1 \neq 0$.

La desigualdad (5.1) pone de manifiesto que $B_0 + \lambda_0 B_1$ tiene un papel relevante en la perturbación de la estructura de Weierstrass porque determina la multiplicidad geométrica de λ_0 en (5.3). Para entender por qué, además, B_1 tiene un papel distinguido, debemos recordar que una cadena de Jordan de $A_0 + \lambda A_1$ de longitud s asociada a λ_0 es aquella que satisface las ecuaciones $(A_0 + \lambda_0 A_1)v_1 = 0$ y $(A_0 + \lambda_0 A_1)v_k = A_1 v_{k-1}$ para $2 \leq k \leq s$. Por lo tanto, es de esperar que una perturbación de A_1 afecte por sí misma a la longitud de las cadenas de Jordan. De manera informal, podemos decir que el comportamiento genérico descrito anteriormente corresponde a la acción conjunta de $B_0 + \lambda_0 B_1$ y B_1 para destruir algunos de los bloques y hacer decrecer la dimensión del mayor número posible de los bloques de Jordan más grandes, mientras que se siga cumpliendo la desigualdad (5.1) sobre la multiplicidad geométrica. Una forma de resumir el comportamiento genérico es el siguiente: la matriz $B_0 + \lambda_0 B_1$ destruye los ρ_0 bloques de Jordan más grandes asociados al autovalor λ_0 y la matriz B_1 transforma los siguientes ρ_1 bloques mayores en bloques de tamaño 1×1 .

La segunda parte de este capítulo está dedicada a obtener, para cada haz regular $A_0 + \lambda A_1$ y cada autovalor finito λ_0 , una caracterización explícita del conjunto de perturbaciones $B_0 + \lambda B_1$ para los que ocurre el comportamiento genérico antes descrito. La condición necesaria y suficiente para que esto ocurra es simplemente que una cierta cantidad, denotada por K_0 , sea distinta de cero. El escalar K_0 está definido a través de una suma de productos de determinantes de matrices que involucran a B_0 , B_1 y algunos de los λ_0 -autovectores de $A_0 + \lambda A_1$. Utilizando el mismo punto de vista que para el caso matricial comentado en la sección previa, el escalar K_0 puede escribirse como la evaluación, $K_0(B_0, B_1)$, de un determinado polinomio de $2n^2$ variables, $K_0(X, Y)$, en el par (B_0, B_1) (identificando el conjunto de haces matriciales $n \times n$ con el espacio \mathbb{C}^{2n^2}). Como consecuencia inmediata, el conjunto de perturbaciones $B_0 + \lambda B_1$ para las cuales el comportamiento genérico no se da (es decir: aquellas para las que $K_0 = 0$) es una subvariedad propia en el conjunto de haces de rango bajo $B_0 + \lambda B_1$.

La obtención del comportamiento genérico y la descripción del conjunto correspondiente de perturbaciones son las dos contribuciones originales más importantes de este capítulo.

Finalmente, introducimos la notación que usaremos a lo largo de este capítulo. Consideraremos un autovalor λ_0 del haz regular $A_0 + \lambda A_1$ con multiplicidad geométrica g y

multiplicidad algebraica a . Esto significa que la parte correspondiente al autovalor λ_0 en la forma de Weierstrass (2.2) del haz contiene g bloques asociados a λ_0 que ocupan un tamaño total de $a \times a$. Es decir, siguiendo la notación de la Sección 2.1,

$$J_{\lambda_0} = \bigoplus_{i=1}^g J_{n_i}(\lambda_0), \quad (5.4)$$

de donde

$$a = n_1 + \dots + n_g. \quad (5.5)$$

Sin pérdida de generalidad, supondremos las dimensiones n_i ordenadas en forma no decreciente, es decir,

$$n_1 \leq n_2 \leq \dots \leq n_g, \quad (5.6)$$

para todas las estructuras de Jordan que aparecerán en este capítulo.

El capítulo está organizado de la siguiente manera: la Sección 5.2 contiene uno de los principales resultados (Teorema 5.2.4) concerniente al cambio de la forma de Weierstrass. Nos da una cota inferior de los tamaños de los bloques de Jordan asociados a un autovalor particular en el haz perturbado y nos predispone a la búsqueda del conjunto de perturbaciones para el que esta cota se alcanza. Este conjunto es descrito en la Sección 5.3.2 para perturbaciones de una clase particular de haces: aquellos que muestran explícitamente la estructura de Weierstrass correspondiente al autovalor cero, y es descrito en la Sección 5.4 para perturbaciones de haces matriciales arbitrarios. También en §5.4, unificamos los resultados previos con el fin de enunciar de manera compacta nuestra contribución más importante relativa al cambio de la estructura de Weierstrass por perturbaciones de rango bajo: el Teorema 5.4.2.

5.2. Cotas inferiores de las multiplicidades algebraicas y las dimensiones de los bloques de Jordan en el haz perturbado

Nuestra intención es determinar la estructura de Weierstrass del autovalor λ_0 en el haz matricial perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$ partiendo de la estructura de este autovalor en el haz no perturbado $A_0 + \lambda A_1$. Para ello necesitamos conocer, en primer lugar, la multiplicidad algebraica, que denotaremos por a_p , de λ_0 en el haz perturbado, y también cómo está distribuída esta multiplicidad algebraica entre los bloques de Jordan de λ_0 en dicho haz. Al menos dos procedimientos son posibles para resolver este problema. Primero, se puede, partiendo de las cadenas de Jordan de $A_0 + \lambda A_1$ asociadas a λ_0 , construir nuevas cadenas de Jordan de $A_0 + B_0 + \lambda(A_1 + B_1)$ asociadas a λ_0 que completen la multiplicidad algebraica a_p . Este procedimiento ha sido el seguido en [52, 60, 61] para el problema de los autovalores de matrices. Tiene la ventaja de calcular explícitamente las nuevas cadenas de Jordan y la desventaja de ser muy intrincado en el caso de haces. En esta memoria, seguiremos un procedimiento más sencillo: primero, determinaremos, por métodos indirectos, una cota inferior del número y los tamaños de los bloques de Jordan asociados a λ_0 en el haz perturbado. Estos métodos indirectos están basados en un resultado de R.C. Thompson [76] e implican el uso de las multiplicidades parciales en λ_0 . Después, demostraremos que, cuando esta cota inferior se alcanza, (el *caso genérico*) los tamaños de estos bloques se pueden determinar. Este ha sido el procedimiento que

se ha utilizado en [61] para el problema de autovalores de matrices. En la Subsección 5.4.1 describiremos explícitamente el conjunto de perturbaciones para el que esta cota se alcanza.

Comenzamos recordando, una vez más, que el rango de un haz matricial arbitrario, regular o singular, $T(\lambda) = T_0 + \lambda T_1$, es r si todos los menores de $T(\lambda)$ de dimensión mayor que r son idénticamente nulos pero $T(\lambda)$ contiene menores de orden r que, como polinomios en λ , no son idénticamente nulos. En consecuencia, el rango de un haz regular $n \times n$ es igual a n .

Incluimos, por completitud, el resultado de R. C. Thompson aludido arriba, que establece relaciones (de divisibilidad) necesarias y suficientes entre los factores invariantes de un haz y los de una perturbación de rango 1 de ese haz. Lo enunciaremos en el caso más general de perturbaciones de polinomios matriciales.

Teorema 5.2.1 [76, Teorema 1] *Sea $H(\lambda)$ un polinomio matricial $n \times n$ con factores invariantes $h_1(H), \dots, h_n(H)$, ordenados de manera que*

$$h_1(H) | h_2(H) | \dots | h_n(H).$$

Sea el polinomio matricial

$$M(\lambda) = H(\lambda) + x(\lambda)y(\lambda)^T$$

donde $x(\lambda)$, $y(\lambda)$ son vectores polinómicos. Entonces todos los posibles factores invariantes

$$h_1(M) | h_2(M) | \dots | h_n(M)$$

de $M(\lambda)$, cuando $x(\lambda)$ e $y(\lambda)$ varían entre todos los vectores polinómicos sobre \mathbb{C} , son precisamente aquellos polinomios que cumplen

$$\begin{aligned} h_1(H) | h_2(M) | h_3(H) | h_4(M) \dots \\ h_1(M) | h_2(H) | h_3(M) | h_4(H) \dots \end{aligned}$$

Observación 5.2.2 1. *Hemos enunciado el Lema ajustándonos lo más posible al enunciado del artículo de referencia [76] pero adaptándolo al caso de polinomios matriciales sobre \mathbb{C} . Por esa razón hemos expresado el haz perturbación como un producto de un vector columna por un vector fila, en lugar de escribir un haz en la forma usual y enunciar como hipótesis que tiene rango 1. El Lema 2.5.1 nos dice que ambas expresiones son equivalentes, es decir: el resultado del Teorema 5.2.1 es cierto para cualquier perturbación de $H(\lambda)$ de rango 1.*

2. *El resultado que nosotros enunciamos es, en realidad, un caso particular del enunciado por R. C. Thompson, que considera matrices sobre un dominio de ideales principales. En nuestro caso, dicho dominio es el anillo de polinomios en la variable λ con coeficientes complejos.*

El siguiente lema auxiliar es una consecuencia del Teorema 5.2.1. Establece cotas inferiores en el número y las dimensiones de los bloques de Jordan de la forma de Weierstrass de un haz matricial regular $(R + T)(\lambda)$, donde $R(\lambda)$ es un haz regular y $T(\lambda)$ es un haz de rango r .

Teorema 5.2.4 *Sea λ_0 un autovalor del haz matricial regular $A_0 + \lambda A_1$, y $n_1 \leq \dots \leq n_g$ los tamaños de los bloques de Jordan asociados a λ_0 en su forma canónica de Weierstrass. Sea $B_0 + \lambda B_1$ un haz tal que el haz perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$ es también regular. Supongamos que $g \geq \text{rg}(B_0 + \lambda_0 B_1)$. Sean $\rho = \text{rg}(B_0 + \lambda_0 B_1) + \text{rg} B_1$ y $n_m = 1$ para todo $m \leq 0$. Entonces*

$$a_p \geq a + \text{rg} B_1 - n_{g-\rho+1} - \dots - n_g. \quad (5.7)$$

Si, además, se da la igualdad, los tamaños de los bloques de Jordan asociados a λ_0 en la forma de Weierstrass de $A_0 + B_0 + \lambda(A_1 + B_1)$ se obtienen eliminando los primeros ρ términos de la secuencia $n_g, \dots, n_1, \underbrace{1, \dots, 1}_{\text{rg} B_1}$.

Demostración. Nótese que

$$\text{rg}(B_0 + \lambda B_1) = \text{rg}(B_0 + \lambda_0 B_1 + (\lambda - \lambda_0) B_1) \leq \text{rg}(B_0 + \lambda_0 B_1) + \text{rg} B_1 = \rho.$$

El Lema 5.2.3 garantiza la existencia de $g - \rho$ bloques de Jordan de tamaños $\beta_1 \geq n_1, \dots, \beta_{g-\rho} \geq n_{g-\rho}$ en la forma canónica de Weierstrass del haz perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$. Más aún, el término de la izquierda de la desigualdad (5.1) implica que hay, al menos, $\rho_1 = \text{rg} B_1$ bloques de Jordan adicionales de tamaños $\alpha_1 \geq 1, \dots, \alpha_{\rho_1} \geq 1$. Por tanto,

$$a_p \geq \beta_1 + \dots + \beta_{g-\rho} + \alpha_1 + \dots + \alpha_{\rho_1} \geq n_1 + \dots + n_{g-\rho} + \rho_1.$$

Obviamente, esta desigualdad es equivalente a (5.7). Si $g \leq \rho$, la desigualdad (5.7) se convierte en

$$a_p \geq g - \text{rg}(B_0 + \lambda_0 B_1).$$

Este hecho es trivial en virtud de (5.1) y la desigualdad espectral evidente

$$a_p \geq \dim \ker (A_0 + \lambda_0 A_1 + B_0 + \lambda_0 B_1).$$

■

Observación 5.2.5 *Nótese que la definición antinatural $n_m = 1$ para $m \leq 0$, nos permite expresar la desigualdad (5.7) de una manera unificada en los dos casos $\rho \leq g$ y $\rho > g$.*

El Teorema 5.2.4 nos da todos los tamaños de los bloques de Jordan asociados a λ_0 en el haz perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$ cuando la desigualdad de (5.7) es una igualdad. Como veremos en la siguiente sección, éste es el caso para la mayoría de las perturbaciones $B_0 + \lambda B_1$. Nuestra intención es describir el conjunto de estas perturbaciones.

5.3. El conjunto de perturbaciones con comportamiento genérico para el autovalor cero

La cantidad $\tilde{a} = a + \text{rg} B_1 - n_{g-\rho+1} - \dots - n_g$ de (5.7) es una cota inferior de la multiplicidad algebraica de λ_0 como autovalor del haz perturbado $A_0 + B_0 + \lambda(A_1 + B_1)$. Esto significa que, para cada perturbación $B_0 + \lambda B_1$ de $A_0 + \lambda A_1$, tenemos

$$\det (A_0 + B_0 + \lambda(A_1 + B_1)) = (\lambda - \lambda_0)^{\tilde{a}} q(\lambda), \quad (5.8)$$

para un cierto polinomio $q(\lambda)$, donde $\rho = \text{rg}(B_0 + \lambda_0 B_1) + \text{rg} B_1$, como en el enunciado del Teorema 5.2.4. Esta igualdad también incluye el caso en que $A_0 + B_0 + \lambda(A_1 + B_1)$ es singular (en este caso, $q(\lambda) \equiv 0$). De cualquier forma, la multiplicidad algebraica de λ_0 en el haz perturbado es exactamente \tilde{a} si y sólo si el coeficiente $q(\lambda_0)$ de $\lambda^{\tilde{a}}$ en (5.8) es distinto de cero. El único objetivo de esta sección es expresar explícitamente este coeficiente en términos de la perturbación, algo que demandará un esfuerzo importante, así como el uso de una notación más detallada. Esta sección es, por tanto, de naturaleza técnica.

Las demostraciones de este apartado hacen uso del siguiente Lema, cuya demostración es una consecuencia directa de igualdades determinantaes bien conocidas [35].

Lema 5.3.1 Sean A y B dos matrices $n \times n$. Entonces, para cada $r = 0, 1, \dots, n$, el coeficiente de λ^{n-r} en $\det(A + \lambda B)$ está dado por la suma

$$\sum_{\substack{1 \leq i_1 < \dots < i_r \leq n \\ 1 \leq j_1 < \dots < j_r \leq n}} (-1)^{\sum_{i=1}^r i_i + j_i} \det A(\{i_1, \dots, i_r\}, \{j_1, \dots, j_r\}) \cdot \det B(\{i_1, \dots, i_r\}^c, \{j_1, \dots, j_r\}^c), \quad (5.9)$$

donde $A(\{i_1, \dots, i_r\}, \{j_1, \dots, j_r\})$ denota la submatriz de A que se obtiene tomando las filas indexadas por i_1, \dots, i_r y las columnas indexadas por j_1, \dots, j_r y $B(\{i_1, \dots, i_r\}^c, \{j_1, \dots, j_r\}^c)$ denota la submatriz de B obtenida eliminando las filas i_1, \dots, i_r y las columnas j_1, \dots, j_r .

Como veremos en la demostración del Teorema 5.4.1, el problema de caracterizar el conjunto de perturbaciones de rango bajo para las que se da el cambio genérico de la estructura de Weierstrass puede reducirse al estudio del autovalor cero de una clase especial de haces regulares, concretamente aquellas que exhiben explícitamente la estructura de Weierstrass asociada a dicho autovalor cero. En aras de la brevedad, el conjunto de haces con tal estructura se llamará *la clase* \mathcal{W}_0 . Dedicamos el resto de esta sección a obtener el coeficiente del término de menor grado en el determinante del haz perturbado cuando el haz no perturbado es regular y está en la clase \mathcal{W}_0 . Pero antes definamos con más precisión la clase \mathcal{W}_0 .

Definición 5.3.2 Un haz regular cuadrado $M_0 + \lambda M_1$ pertenece a la clase \mathcal{W}_0 si

$$M_0 = \text{diag}(-J_0, \hat{J}) \quad \text{y} \quad M_1 = \text{diag}(I, R),$$

con \hat{J} regular, I la matriz identidad de la misma dimensión que J_0 , y,

$$J_0 = \text{diag}(J_{n_1}(0), \dots, J_{n_g}(0)). \quad (5.10)$$

Nótese que no se asume ninguna estructura especial sobre las matrices \hat{J} y R . Nótese también que la estructura de Weierstrass asociada al autovalor cero de $M_0 + \lambda M_1$ está exhibida explícitamente en la matriz J_0 .

A lo largo de esta sección, consideraremos el haz $M_0 + C_0 + \lambda(M_1 + C_1)$, donde $M_0 + \lambda M_1$ está en la clase \mathcal{W}_0 y $C_0 + \lambda C_1$ es un haz tal que $\text{rg}(C_0) < g$. Además, usaremos la notación

$$\rho_0 := \text{rg}(C_0), \quad \rho_1 := \text{rg}(C_1), \quad \rho := \rho_0 + \rho_1.$$

El Lema 5.3.1 nos dice que, si nos restringimos al caso $\lambda_0 = 0$, el coeficiente, $q(0)$, del término $\lambda^{\tilde{a}}$ en (5.8) es la suma de ciertos menores complementarios con signo de $A_0 + B_0$ y de $A_1 + B_1$. Si denotamos por \mathcal{M} un menor arbitrario de $A_0 + B_0$ y \mathcal{M}_{comp} a su menor complementario en $A_1 + B_1$, entonces \mathcal{M} tiene orden $n - \tilde{a}$ y \mathcal{M}_{comp} tiene orden \tilde{a} . Muchos de estos menores son siempre nulos, independientemente de la perturbación $B_0 + \lambda B_1$, básicamente como consecuencia de los siguientes hechos triviales:

F1 Todo menor de $A_0 + B_0$ con más de $\text{rg } B_0$ columnas con sólo elementos de B_0 es cero.

F2 Todo menor de $A_1 + B_1$ con más de $\text{rg } B_1$ columnas con sólo elementos de B_1 es cero.

Nótese que, si $A_0 + \lambda A_1$ está en la clase \mathcal{W}_0 , hay algunas filas y columnas de ceros en A_0 (resp. A_1), luego las filas y columnas con sólo elementos de B_0 (resp. B_1) aparecen en la suma $A_0 + B_0$ (resp. $A_1 + B_1$). Nuestro principal esfuerzo en esta sección es determinar cómo están contruídos aquellos menores \mathcal{M} para los cuales $\mathcal{M} \cdot \mathcal{M}_{comp}$ no es necesariamente cero y calcular su suma como en (5.9). Para este fin, hemos de considerar separadamente los casos $\rho \leq d_0$ y $\rho > d_0$, donde d_0 es el número de bloques de Jordan en (5.10) de dimensión mayor que 1.

Supondremos también que en la estructura de Weierstrass (5.10) los tamaños de los bloques de Jordan asociados al autovalor cero satisfacen:

$$n_1 = \dots = n_{r_1} < n_{r_1+1} = \dots = n_{r_1+r_2} < \dots < n_{r_1+\dots+r_{q-1}+1} = \dots = n_{r_1+\dots+r_q}, \quad (5.11)$$

es decir: hay r_1 bloques de dimensión $\bar{n}_1 \equiv n_1$, r_2 bloques de dimensión $\bar{n}_2 \equiv n_{r_1+1}, \dots$, y r_q bloques de dimensión $\bar{n}_q \equiv n_{r_1+\dots+r_{q-1}+1}$. Nótese que, en este caso, la multiplicidad algebraica del autovalor cero es $a = r_1 \bar{n}_1 + \dots + r_q \bar{n}_q$ y la multiplicidad geométrica es $g = r_1 + \dots + r_q$.

Dada la estructura de Weierstrass (5.10) situada en el bloque superior izquierdo de M_0 , denotamos por ψ_0 (resp. ψ_1) a la submatriz $g \times g$ de C_0 (resp. de C_1) que contiene a las filas y columnas correspondientes al elemento inferior izquierdo de cada bloque de Jordan² en J_0 , es decir, ψ_0 (resp. ψ_1) contiene los elementos de las filas y columnas de C_0 (resp. de C_1) que no contienen -1 s supradiagonales en J_0 . También denotamos por

$$f_j = \sum_{i=j}^q r_i, \quad j = 1, \dots, q, \quad f_{q+1} = 0, \quad (5.12)$$

a las sumas parciales de las cantidades r_i y definimos, para cada $j = 1, \dots, q-1$, la matriz $\psi_{0,j}$ (resp. $\psi_{1,j}$) como la submatriz principal (inferior derecha) de tamaño $f_j \times f_j$ de ψ_0 (resp. de ψ_1). Análogamente, definimos $\psi_{0,1} := \psi_0$, pero para definir $\psi_{1,1}$ debemos distinguir: si $\rho \leq d_0$ entonces $\psi_{1,1} := \psi_1$, pero si $\rho > d_0$, entonces

$$\psi_{1,1} := E_1 + \psi_1, \quad \text{donde} \quad E_1 = \begin{bmatrix} I_{r_1} & 0 \\ 0 & 0 \end{bmatrix}. \quad (5.13)$$

Nótese que cada $\psi_{i,j}$ es el bloque inferior derecho de tamaño $f_j \times f_j$ de $\psi_{i,j-1}$.

En el caso $\rho \leq d_0$, sea $s \in \{1, \dots, q\}$ tal que $\rho = r_q + \dots + r_{q-s+2} + \gamma = f_{s-1} + \gamma$, para γ cumpliendo $0 < \gamma \leq r_{q-s+1}$, y sea $I = \{i_1, \dots, i_\gamma\}$ cualquier subconjunto de $\{1, \dots, r_{q-s+1}\}$ de cardinal γ . Denotamos por

$$\psi_{0,q-s+1}^{(I)} \quad (\text{resp.} \quad \psi_{1,q-s+1}^{(I)}) \quad (5.14)$$

la submatriz principal de dimensión ρ de $\psi_{0,q-s+1}$ (resp. de $\psi_{1,q-s+1}$) que contiene a $\psi_{0,q-s}$ (resp. $\psi_{1,q-s}$) junto con las γ filas y columnas de $\psi_{0,q-s+1}$ (resp. de $\psi_{1,q-s+1}$) correspondientes a los γ bloques de tamaño \bar{n}_{q-s+1} en J_0 indexados por i_1, \dots, i_γ . Nótese que si $\gamma = r_{q-s+1}$ entonces $\psi_{0,q-s+1}^{(I)} = \psi_{0,q-s+1}$ (resp. $\psi_{1,q-s+1}^{(I)} = \psi_{1,q-s+1}$).

²Obviamente, para los bloques de Jordan de dimensión 1×1 , el elemento inferior izquierdo es el único elemento del bloque.

Si $J = (J_1, J_2)$, donde $J_1 = \{j_1, \dots, j_{\rho_0}\}$, $J_2 = \{k_1, \dots, k_{\rho_0}\}$ son subconjuntos de $\{1, \dots, \rho\}$ de cardinal ρ_0 , denotamos por

$$\nabla_{0, q-s+1}^{(I, J)} \quad (\text{resp. } \nabla_{1, q-s+1}^{(I, J^c)}) \quad (5.15)$$

a los menores $\rho_0 \times \rho_0$ de $\psi_{0, q-s+1}^{(I)}$ correspondientes a las filas indexadas por j_1, \dots, j_{ρ_0} y las columnas indexadas por k_1, \dots, k_{ρ_0} (resp. los menores $\rho_1 \times \rho_1$ de $\psi_{1, q-s+1}^{(I)}$ complementarios a $\nabla_{0, q-s+1}^{(I, J)}$, esto es, aquéllos que contienen todas las filas y columnas excepto las indexadas por j_1, \dots, j_{ρ_0} y k_1, \dots, k_{ρ_0} , respectivamente).

En el caso $\rho > d_0$, serán necesarios otros menores: denotamos por

$$\nabla_0^{(J)} \quad (\text{resp. } \nabla_1^{(J^c)}) \quad (5.16)$$

a los menores $\rho_0 \times \rho_0$ de ψ_0 (resp. los menores $(g - \rho_0) \times (g - \rho_0)$ de $\psi_{1,1}$) que contienen las filas y columnas indexadas por J (resp. que contienen todas las filas y columnas excepto las indexadas por J). Obsérvese que, en este caso, no hay necesidad de usar ningún conjunto de índices I .

Ilustraremos las definiciones previas con un ejemplo.

Ejemplo 5.3.3 Supongamos que

$$M_0 = \left[\begin{array}{c|c|c|c|c|c} 0 & & & & & \\ \hline & 0 & & & & \\ \hline & & 0 & 1 & & \\ & & 0 & 0 & & \\ \hline & & & & 0 & 1 \\ & & & & 0 & 0 \\ \hline & & & & & & 0 & 1 & 0 \\ & & & & & & 0 & 0 & 1 \\ & & & & & & 0 & 0 & 0 \end{array} \right],$$

i.e., $\bar{n}_1 = 1$, $\bar{n}_2 = 2$, $\bar{n}_3 = 3$, $r_1 = 2$, $r_2 = 3$, $r_3 = 1$, y

$$C_0 = \left[\begin{array}{c|c|c|c|c|c|c|c|c} a_1 & a_2 & a_3 & * & a_4 & * & a_5 & * & a_6 & * & * \\ \hline a_7 & a_8 & a_9 & * & a_{10} & * & a_{11} & * & a_{12} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline a_{13} & a_{14} & a_{15} & * & a_{16} & * & a_{17} & * & a_{18} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline a_{19} & a_{20} & a_{21} & * & a_{22} & * & a_{23} & * & a_{24} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline a_{25} & a_{26} & a_{27} & * & a_{28} & * & a_{29} & * & a_{30} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline a_{31} & a_{32} & a_{33} & * & a_{34} & * & a_{35} & * & a_{36} & * & * \end{array} \right],$$

$$C_1 = \left[\begin{array}{c|c|c|c|c|c|c|c|c} b_1 & b_2 & b_3 & * & b_4 & * & b_5 & * & b_6 & * & * \\ \hline b_7 & b_8 & b_9 & * & b_{10} & * & b_{11} & * & b_{12} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline b_{13} & b_{14} & b_{15} & * & b_{16} & * & b_{17} & * & b_{18} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline b_{19} & b_{20} & b_{21} & * & b_{22} & * & b_{23} & * & b_{24} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline b_{25} & b_{26} & b_{27} & * & b_{28} & * & b_{29} & * & b_{30} & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline * & * & * & * & * & * & * & * & * & * & * \\ \hline b_{31} & b_{32} & b_{33} & * & b_{34} & * & b_{35} & * & b_{36} & * & * \end{array} \right].$$

En esta situación,

$$\psi_0 = \left[\begin{array}{c|c|c|c|c|c} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ \hline a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} \\ \hline a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} \\ \hline a_{19} & a_{20} & a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{25} & a_{26} & a_{27} & a_{28} & a_{29} & a_{30} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \end{array} \right], \quad \psi_1 = \left[\begin{array}{c|c|c|c|c|c} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ \hline b_7 & b_8 & b_9 & b_{10} & b_{11} & b_{12} \\ \hline b_{13} & b_{14} & b_{15} & b_{16} & b_{17} & b_{18} \\ \hline b_{19} & b_{20} & b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{25} & b_{26} & b_{27} & b_{28} & b_{29} & b_{30} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \end{array} \right]$$

mientras que

$$\psi_{0,3} = [a_{36}], \quad \psi_{0,2} = \left[\begin{array}{c|c|c|c} a_{15} & a_{16} & a_{17} & a_{18} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{27} & a_{28} & a_{29} & a_{30} \\ \hline a_{33} & a_{34} & a_{35} & a_{36} \end{array} \right],$$

y análogamente para $\psi_{1,j}$ reemplazando la letra a por la b .

Ahora, supongamos que el haz perturbación satisface $\rho_0 = 2$, $\rho_1 = 1$. En esta situación, $\rho = 3 \leq d_0 = 4$, $s = 2$ y $\gamma = 2$. Entonces, $\psi_{0,3} = \psi_0$, $\psi_{1,3} = \psi_1$ y, por ejemplo,

$$\psi_{0,2}^{\{\{1,3\}\}} = \left[\begin{array}{c|c|c} a_{15} & a_{17} & a_{18} \\ \hline a_{27} & a_{29} & a_{30} \\ \hline a_{33} & a_{35} & a_{36} \end{array} \right], \quad \psi_{0,2}^{\{\{2,3\}\}} = \left[\begin{array}{c|c|c} a_{22} & a_{23} & a_{18} \\ \hline a_{28} & a_{29} & a_{30} \\ \hline a_{34} & a_{35} & a_{36} \end{array} \right],$$

y análogamente para $\psi_{1,2}^{(\cdot)}$ reemplazando la letra a por la b . Así pues,

$$\begin{aligned}\nabla_{0,2}^{\{1,3\},\{1,2;2,3\}} &= \det \begin{bmatrix} a_{17} & a_{18} \\ a_{29} & a_{30} \end{bmatrix}, & \nabla_{1,2}^{\{1,3\},\{1,2;2,3\}^c} &= b_{33} \\ \nabla_{0,2}^{\{2,3\},\{1,3;1,2\}} &= \det \begin{bmatrix} a_{22} & a_{23} \\ a_{34} & a_{35} \end{bmatrix}, & \nabla_{1,2}^{\{2,3\},\{1,3;1,2\}^c} &= b_{30}.\end{aligned}$$

En cambio, si consideramos una perturbación tal que $\rho_0 = 3$ y $\rho_1 = 2$, entonces $\rho = 5 > d_0 = 4$, luego

$$\psi_{0,1} = \psi_0, \quad \psi_{1,1} = E_1 + \psi_1 = \begin{bmatrix} 1 + b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ b_7 & 1 + b_8 & b_9 & b_{10} & b_{11} & b_{12} \\ b_{13} & b_{14} & b_{15} & b_{16} & b_{17} & b_{18} \\ b_{19} & b_{20} & b_{21} & b_{22} & b_{23} & b_{24} \\ b_{25} & b_{26} & b_{27} & b_{28} & b_{29} & b_{30} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \end{bmatrix},$$

mientras que, por ejemplo,

$$\nabla_0^{\{1,3,4;1,4,5\}} = \det \begin{bmatrix} a_1 & a_4 & a_5 \\ a_{13} & a_{16} & a_{17} \\ a_{19} & a_{22} & a_{23} \end{bmatrix}, \quad \nabla_1^{\{1,3,4;1,4,5\}^c} = \det \begin{bmatrix} 1 + b_8 & b_9 & b_{12} \\ b_{26} & b_{27} & b_{30} \\ b_{32} & b_{33} & b_{36} \end{bmatrix}.$$

□

Con esta notación a mano, estamos en situación de enunciar el siguiente resultado.

Teorema 5.3.4 *Sea $M_0 + \lambda M_1$ un haz en la clase \mathcal{W}_0 , y supongamos que los tamaños de los bloques de Jordan asociados al cero satisfacen (5.11). Sean a y g , respectivamente, las multiplicidades algebraica y geométrica del cero como autovalor de $M_0 + \lambda M_1$. Sea d_0 el número de bloques de Jordan asociados al cero cuya dimensión es mayor que 1. Sea $C_0 + \lambda C_1$ cualquier haz matricial de la misma dimensión que $M_0 + \lambda M_1$, y sea*

$$\rho_0 := \text{rg}(C_0), \quad \rho_1 := \text{rg}(C_1), \quad \rho := \rho_0 + \rho_1.$$

Sea

$$\tilde{a} = a + \rho_1 - n_1 - \dots - n_\rho,$$

y denotemos $n_m = 1$ para cada $m \leq 0$.

Si $\rho_0 < g$ entonces el coeficiente de $\lambda^{\tilde{a}}$ en el polinomio perturbado (5.8) es

$$q(0) = K_0 \det \hat{J},$$

donde:

(i) si $\rho \leq d_0$ entonces

$$K_0 = \sum_I \sum_J (-1)^{\alpha_{I,J}} \nabla_{0,q-s+1}^{(I,J)} \nabla_{1,q-s+1}^{(I,J^c)},$$

donde $s \in \{1, \dots, q\}$ es tal que $\rho = r_q + \dots + r_{q-s+2} + \gamma$ con $0 < \gamma \leq r_{q-s+1}$, I varía entre todos los subconjuntos de $\{1, \dots, r_{q-s+1}\}$ con γ elementos y $J = (J_1, J_2)$, donde J_1 y J_2 varían, independientemente, entre todos los subconjuntos de $\{1, \dots, \rho\}$ con ρ_0 elementos. El signo $(-1)^{\alpha_{I,J}}$ depende sólo de los conjuntos I y J . Nótese que si $\gamma = r_{q-s+1}$ entonces la suma en I no aparece.

(ii) si $\rho > d_0$ entonces

$$K_0 = \sum_J (-1)^{\alpha_J} \nabla_0^{(J)} \nabla_1^{(J^c)},$$

donde $J = (J_1, J_2)$ y J_1 y J_2 varían, independientemente, entre todos los subconjuntos $\{1, \dots, g\}$ con ρ_0 elementos. El signo $(-1)^{\alpha_J}$ depende sólo del conjunto J .

Demostración. De acuerdo con el Lema 5.3.1, el coeficiente de $\lambda^{\tilde{a}}$ en el polinomio (5.8) es una suma de productos con signo, digamos $\mathcal{D}_0 \mathcal{D}_1$, de menores, donde cada factor \mathcal{D}_0 es un menor de $M_0 + C_0$ y el correspondiente factor \mathcal{D}_1 es su menor complementario en $M_1 + C_1$. También, la dimensión de \mathcal{D}_0 es $n - \tilde{a}$, mientras que la dimensión de \mathcal{D}_1 es \tilde{a} . Ahora buscamos los menores complementarios de estos tamaños cuyo producto no es necesariamente cero.

(i) Suponemos, en primer lugar, que $\rho \leq d_0$. Primero nótese que en \mathcal{D}_0 sólo pueden aparecer filas y columnas que intersecan, a lo sumo, con ρ bloques diferentes de J_0 : si tomamos sólo filas y columnas correspondientes a entradas diagonales del bloque, entonces en este menor aparece una fila y una columna con entradas sólo en C_0 , mientras que, en otro caso, aparecen en \mathcal{D}_1 una fila y una columna con entradas sólo en C_1 . Así, la afirmación es una consecuencia de los hechos **F1** y **F2** al comienzo de esta subsección. Por otro lado, de nuevo usando **F1**, podemos tomar en \mathcal{D}_0 sólo las filas y columnas de, a lo más, ρ_0 bloques completos. Ahora, contamos el máximo orden de un menor de $M_0 + C_0$ con estas restricciones. Obviamente, este máximo se alcanza si tomamos las $n - a$ últimas filas y columnas (correspondientes a los autovalores no nulos de M_0), junto con ρ_0 bloques completos entre los ρ mayores. De nuevo, es fácil ver que podemos tomar también los restantes ρ_1 bloques completos entre los ρ mayores excepto por la última fila y la primera columna (esto es: tomamos las filas y columnas que contienen los -1 s supradiagonales). Y no más. Todas estas filas (o columnas) hacen un total de

$$n - a + n_{g-\rho+1} + \dots + n_g - \rho_1 = n - \tilde{a}.$$

Esto significa que el menor \mathcal{D}_0 debe construirse del modo anterior porque, de lo contrario, sería de orden menor que $n - \tilde{a}$. Así, cada menor de \mathcal{D}_0 de tamaño $n - \tilde{a}$ está formado con las filas y columnas de $M_0 + C_0$ seleccionadas mediante los siguientes tres pasos:

Paso 1 Incluir en \mathcal{D}_0 las $n - a$ últimas filas y columnas de $M_0 + C_0$.

Paso 2 Incluir en \mathcal{D}_0 las filas y columnas en que están situados los -1 s supradiagonales de los ρ bloques de Jordan de J_0 con mayor tamaño. Si $\gamma = r_{q-s+1}$, estas posiciones están unívocamente determinadas, pero en otro caso existen diversas elecciones de los ρ bloques de Jordan mayores de J_0 , cada una de ellas especificada por un subconjunto I de $\{1, \dots, r_{q-s+1}\}$ con γ elementos. Los γ elementos de I determinan cuáles de los bloques de Jordan \bar{n}_{q-s+1} -dimensionales de (5.10) están incluidos entre los ρ mayores. Conforme I varía sobre todos los subconjuntos de $\{1, \dots, r_{q-s+1}\}$ con γ elementos, se cubren todas las posibles elecciones de las posiciones de los -1 s supradiagonales de \mathcal{D}_0 .

Paso 3 De acuerdo con la discusión previa, las ρ_0 filas y ρ_0 columnas de \mathcal{D}_0 con sólo elementos de C_0 tienen que ser elegidas entre aquellas que corresponden a los

elementos inferiores de los ρ bloques de Jordan mayores de J_0 seleccionados en el **Paso 2** a través del subconjunto I . Una forma de especificar estas filas y columnas es como sigue: formamos la matriz $\psi_{0,q-s+1}^{(I)}$ y elegimos cualquiera de sus menores ρ_0 -dimensionales $\nabla_{0,q-s+1}^{(I,J)}$ (no necesariamente principal), e incluimos en \mathcal{D}_0 las correspondientes filas y columnas de $M_0 + C_0$. Cuando J_1 y J_2 varían, independientemente, sobre todos los subconjuntos de $\{1, \dots, \rho\}$ con ρ_0 elementos, todos los menores $\nabla_{0,q-s+1}^{(I,J)}$ de tamaño $\rho_0 \times \rho_0$ son considerados.

Esto completa la descripción de todos los menores admisibles, \mathcal{D}_0 , de $M_0 + C_0$. Cada menor admisible \mathcal{D}_0 está completamente especificado mediante un par de conjuntos de índices I y J .

A continuación, demostramos que, dados dos conjuntos I y J , el menor correspondiente \mathcal{D}_0 es $\mathcal{D}_0 = \pm(\det \hat{J}) \nabla_{0,q-s+1}^{(I,J)}$, donde \hat{J} es la matriz que aparece en la Definición 5.3.2, y el signo depende exclusivamente de los conjuntos I y J . Denotemos por l el número total de -1 s supradiagonales de J_0 incluidos en \mathcal{D}_0 , i.e., $l = r_q(\bar{n}_q - 1) + \dots + r_{q-s+2}(\bar{n}_{q-s+2} - 1) + \gamma(\bar{n}_{q-s+1} - 1)$, y sean $1 = m_1 < \dots < m_l$ los índices de las filas de \mathcal{D}_0 que contienen las posiciones de los -1 s supradiagonales en las columnas $p_1 < \dots < p_l$ de \mathcal{D}_0 . La m_k -ésima fila de \mathcal{D}_0 es la suma de dos filas: una es la p_k -ésima fila e_{p_k} de la matriz identidad y la otra es una parte de una fila de C_0 . Usando este hecho, \mathcal{D}_0 puede expresarse como la suma de 2^l determinantes cuya m_k -ésima fila, para $1 \leq k \leq l$, es bien e_{p_k} o una fila con sólo elementos de C_0 . A excepción del determinante cuyas filas son $e_{p_1}, e_{p_2}, \dots, e_{p_l}$, el resto de estos determinantes es igual a cero, puesto que cada uno contiene más de ρ_0 filas con elementos sólo de C_0 . Un argumento similar en las últimas $n - a$ filas de \mathcal{D}_0 nos permite reemplazar por cero cada elemento de C_0 en estas filas sin alterar el valor de \mathcal{D}_0 . El desarrollo por cofactores del determinante que queda usando las filas $1 = m_1 < \dots < m_l$ lleva a un valor para \mathcal{D}_0 igual a $\pm(\det \hat{J}) \nabla_{0,q-s+1}^{(I,J)}$. Nótese que el signo depende sólo de las posiciones en \mathcal{D}_0 correspondientes a los -1 s supradiagonales, y éstas están completamente determinadas por I y J .

En este párrafo demostramos que el menor complementario correspondiente, \mathcal{D}_1 , de $M_1 + C_1$ es $\mathcal{D}_1 = \pm \nabla_{1,q-s+1}^{(I,J^c)}$, y, nuevamente, el signo depende sólo de los conjuntos I y J . El menor \mathcal{D}_1 se obtiene eligiendo entre las primeras a filas y columnas de $M_1 + C_1$ aquellas que no han sido elegidas en \mathcal{D}_0 . Esto implica que las ρ_1 filas y las ρ_1 columnas de \mathcal{D}_1 con sólo elementos en C_1 tienen que elegirse entre las filas y columnas correspondientes a los elementos inferiores izquierdos de los ρ bloques de Jordan más grandes de J_0 especificados por I . Nótese que \mathcal{D}_1 se completa con todas las filas y columnas del bloque superior izquierdo $a \times a$ de $M_1 + C_1$ que no están incluidas en estos ρ bloques de Jordan mayores. Todas estas últimas filas y columnas incluyen posiciones de los 1 s diagonales del bloque superior izquierdo $a \times a$ de M_1 , aunque quizá en \mathcal{D}_1 no estén en posiciones diagonales porque \mathcal{D}_1 no es necesariamente un menor principal de $M_1 + C_1$. Supongamos que $\alpha_1 < \dots < \alpha_t$ y $\beta_1 < \dots < \beta_t$ son los índices de las filas y columnas de \mathcal{D}_1 correspondientes a posiciones de los 1 s diagonales del bloque superior izquierdo $a \times a$ de M_1 . Ahora podemos argumentar como antes: la fila α_k -ésima de \mathcal{D}_1 es la suma de dos filas: una es la β_k -ésima fila e_{β_k} de la matriz identidad y la otra es una parte de una fila de C_1 . Descomponiendo \mathcal{D}_1 en pasos sucesivos como la suma de dos menores y cancelando los menores con

más de ρ_1 filas en C_1 , el menor resultante es precisamente $\pm \nabla_{1,q-s+1}^{(I,J^c)}$ con el mismo conjunto de índices J que en \mathcal{D}_0 . Nótese que el signo depende exclusivamente de las posiciones en \mathcal{D}_1 correspondientes a los 1s diagonales del bloque superior izquierdo $a \times a$ de M_1 , y éstas están completamente determinadas por I y J .

(ii) En el caso $\rho > d_0$, cualquier menor \mathcal{D}_0 de orden $n - \tilde{a}$ se forma con las filas y columnas de $M_0 + C_0$ seleccionadas en los siguientes tres pasos:

Paso 1 Incluir en \mathcal{D}_0 las $n - a$ últimas filas y columnas de $M_0 + C_0$.

Paso 2 Incluir en \mathcal{D}_0 las filas y columnas correspondientes a las posiciones de *todos* los -1 s supradiagonales de J_0 .

Paso 3 Ahora, la única restricción para las ρ_0 filas y ρ_0 columnas de \mathcal{D}_0 con elementos sólo de C_0 es que deben ser elegidas entre las g filas y las g columnas que no corresponden a todos los -1 s supradiagonales de J_0 . Una forma de especificar estas filas y columnas es como sigue: formamos la matriz ψ_0 y escogemos cualquiera de sus menores ρ_0 -dimensionales $\nabla_0^{(J)}$ (no necesariamente principal), e incluimos en \mathcal{D}_0 las correspondientes filas y columnas de $M_0 + C_0$. Cuando J_1 y J_2 varían, independientemente, sobre todos los subconjuntos de $\{1, \dots, \rho\}$ con ρ_0 elementos, todos los $\rho_0 \times \rho_0$ menores $\nabla_0^{(J)}$ son considerados. Esto completa la descripción de todos los posibles menores de dimensión k_{\max} , \mathcal{D}_0 , de $M_0 + C_0$. Cada posible menor \mathcal{D}_0 está completamente especificado por el conjunto J .

Un argumento similar al utilizado arriba muestra que, para el menor \mathcal{D}_0 definido por J , tenemos $\mathcal{D}_0 = \pm(\det \widehat{J}) \nabla_0^{(J)}$. El correspondiente menor complementario \mathcal{D}_1 es, simplemente, $\mathcal{D}_1 = \nabla_1^{(J^c)}$, porque, en este caso, la condición del rango sobre C_1 no permite tomar todas las filas y columnas correspondientes a los 1s diagonales en el bloque superior izquierdo de M_1 . De aquí las definiciones diferentes para $\psi_{1,1}$.

■

De acuerdo con el Teorema 5.3.4, la multiplicidad algebraica del cero como autovalor del haz perturbado $M_0 + C_0 + \lambda(M_1 + C_1)$ es \tilde{a} si y sólo si $K_0 \neq 0$. Ponemos de manifiesto que, como puede verse en la demostración del Teorema 5.3.4, los signos $(-1)^{\alpha_{I,J}}$ no dependen en forma alguna del haz perturbación $C_0 + \lambda C_1$. Dependen exclusivamente de la estructura de Weierstrass del haz no perturbado $M_0 + \lambda M_1$ y, por supuesto, de la elección de I y J . Una vez que se ha hecho esta elección, cada signo se obtiene de manera directa aplicando el Lema 5.3.1 y usando el desarrollo por cofactores descrito en el Teorema 5.3.4.

Ejemplo 5.3.5 Consideremos el haz no perturbado perteneciente a la clase \mathcal{W}_0 :

$$A_0 + \lambda A_1 = \text{diag}(\lambda I_8 - J_0, \widehat{J}(\lambda)),$$

donde

$$\begin{aligned} J_0 &= J_1(0) \oplus J_2(0) \oplus J_2(0) \oplus J_3(0) \\ \widehat{J}(\lambda) &= \lambda J_2(0) - I_2 \end{aligned}$$

Por tanto:

$$\begin{aligned} \bar{n}_1 &= 1, & \bar{n}_2 &= 2, & \bar{n}_3 &= 3 \\ r_1 &= 1, & r_2 &= 2, & r_3 &= 1. \end{aligned}$$

Perturbemos por otro haz $B_0 + \lambda B_1$ con

$$B_0 = \begin{bmatrix} 0 & 6 & 4 & 4 & -6 & 2 & -2 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 30 & 20 & 20 & -30 & 10 & -10 & 0 & 0 & -10 \\ 0 & 3 & 2 & 2 & -3 & 1 & -1 & 0 & 0 & -1 \\ \hline 0 & 51 & 34 & 34 & -51 & 17 & -17 & 0 & 0 & -17 \\ 0 & -3 & 2 & -2 & 3 & -1 & 1 & 0 & 0 & 1 \\ 0 & 51 & 34 & 34 & -51 & 17 & -17 & 0 & 0 & -17 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & -6 & -6 & 9 & -3 & 3 & 0 & 0 & 3 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} -3 & 1 & -1 & -1 & 0 & -2 & -2 & 3 & 0 & -1 \\ 45 & -15 & 15 & 15 & 0 & 30 & 30 & -45 & 0 & 15 \\ 6 & -2 & 2 & 2 & 0 & 4 & 4 & -6 & 0 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & -1 & -1 & 0 & -2 & -2 & 3 & 0 & -1 \\ \hline 36 & -12 & 12 & 12 & 0 & 24 & 24 & -36 & 0 & 12 \\ 9 & -3 & 3 & 3 & 0 & 6 & 6 & -9 & 0 & 3 \\ 3 & -1 & 1 & 1 & 0 & 2 & 2 & -3 & 0 & 1 \\ \hline 3 & -1 & 1 & 1 & 0 & 2 & 2 & -3 & 0 & 1 \\ 135 & -45 & 45 & 45 & 0 & 90 & 90 & -135 & 0 & 45 \end{bmatrix},$$

Puede verse fácilmente que $\rho_0 = \rho_1 = 1$, es decir: $\rho = 2$. Además, $\gamma = 1$, de modo que las posibilidades para los subconjuntos I , J_1 y J_2 son $\{1\}$ ó $\{2\}$ en cada uno de los tres casos. En consecuencia, las matrices que nos interesan son

$$\psi_{0,2}^1 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 51 & 17 \end{array} \right], \quad \psi_{0,2}^2 = \left[\begin{array}{c|c} 2 & 1 \\ \hline 34 & 17 \end{array} \right], \quad \psi_{1,2}^1 = \left[\begin{array}{c|c} -2 & 4 \\ \hline -1 & 2 \end{array} \right], \quad \psi_{1,2}^2 = \left[\begin{array}{c|c} -1 & -2 \\ \hline 1 & 2 \end{array} \right],$$

de donde obtenemos

$$\begin{aligned} \nabla_{0,2}^{\{\{1\},\{1;1\}\}} &= 0, & \nabla_{1,2}^{\{\{1\},\{1;1\}^c\}} &= 2 \\ \nabla_{0,2}^{\{\{1\},\{1;2\}\}} &= 0, & \nabla_{1,2}^{\{\{1\},\{1;2\}^c\}} &= -1 \\ \nabla_{0,2}^{\{\{1\},\{2;1\}\}} &= 51, & \nabla_{1,2}^{\{\{1\},\{2;1\}^c\}} &= 4 \\ \nabla_{0,2}^{\{\{1\},\{2;2\}\}} &= 17, & \nabla_{1,2}^{\{\{1\},\{2;2\}^c\}} &= -2 \\ \nabla_{0,2}^{\{\{2\},\{1;1\}\}} &= 2, & \nabla_{1,2}^{\{\{2\},\{1;1\}^c\}} &= 2 \\ \nabla_{0,2}^{\{\{2\},\{1;2\}\}} &= 1, & \nabla_{1,2}^{\{\{2\},\{1;2\}^c\}} &= 1 \\ \nabla_{0,2}^{\{\{2\},\{2;1\}\}} &= 34, & \nabla_{1,2}^{\{\{2\},\{2;1\}^c\}} &= -2 \\ \nabla_{0,2}^{\{\{2\},\{2;2\}\}} &= 17, & \nabla_{1,2}^{\{\{2\},\{2;2\}^c\}} &= -1. \end{aligned}$$

Por otra parte, el signo $(-1)^{\{\{1\},\{2;1\}\}}$ es igual al producto $\sigma_1 \cdot \sigma_2 \cdot \sigma_3$ donde σ_1 es el signo que resulta al desarrollar por los -1 s el determinante

$$\begin{vmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 51 & 0 & 0 & 0 \end{vmatrix},$$

que es la parte del menor \mathcal{D}_0 correspondiente al autovalor 0. El signo σ_2 es el que resulta al desarrollar por los 1s el menor complementario, \mathcal{D}_1 ,

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix},$$

y σ_3 es el signo que corresponde según la fórmula (5.9). Como el primer factor se obtiene tomando las filas del haz perturbado cuyos índices son 2, 6, 7 y 8 y las columnas cuyos índices son 2, 3, 7 y 8, se tiene que $\sigma_3 = -1$. Como, por otro lado, $\sigma_1 = 1$ y $\sigma_2 = 1$, se concluye que $(-1)^{(\{1\}, \{2;1\})} = -1$.

Para obtener el menor $\nabla_{0,2}^{(\{1\}, \{2;2\})}$ tomamos las filas del haz perturbado cuyos índices son 2, 6, 7 y 8 y las columnas cuyos índices son 3, 6, 7 y 8, por lo que, en este caso, $\sigma_3 = -1$. Por otro lado, la parte correspondiente al autovalor 0 del menor \mathcal{D}_0 es

$$\begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 17 & 0 & 0 \end{vmatrix},$$

lo que significa que $\sigma_1 = -1$, mientras que el menor complementario, \mathcal{D}_1 , es

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix},$$

luego $\sigma_2 = 1$. Por lo tanto, $(-1)^{(\{1\}, \{2;2\})} = 1$. El resto de signos se obtiene de manera análoga. Finalmente:

$$K_0 = [-51 \cdot 4 + 17 \cdot (-2)] + [2 \cdot 2 - 1 \cdot 1 - 34 \cdot (-2) + 17 \cdot (-1)] = -184$$

$$\det \widehat{J}(0) = \det(-I_2) = 1.$$

De hecho, el determinante completo del haz perturbado es

$$\det(A_0 + B_0 + \lambda(A_1 + B_1)) = -184\lambda^4 - 2075\lambda^5 - 4314\lambda^6 - 190\lambda^7 - 1647\lambda^8 - 45\lambda^9.$$

En virtud del Teorema 5.3.4, como $K_0 \neq 0$, la forma de Weierstrass de $A_0 + B_0 + \lambda(A_1 + B_1)$ asociada al autovalor 0 es

$$J_2(0) \oplus J_1(0) \oplus J_1(0).$$

□

Obsérvese que la ecuación determinantal $K_0 = 0$ define una variedad algebraica, formada por aquellas perturbaciones $C_0 + \lambda C_1$, donde C_0 y C_1 tienen rangos predeterminados, para las cuales la multiplicidad algebraica del cero como autovalor del haz perturbado no es \tilde{a} . En otras palabras: las perturbaciones para las cuales no se da el comportamiento genérico descrito en la introducción a este capítulo.

5.4. Perturbaciones de rango bajo de haces regulares

En esta sección daremos una respuesta completa al problema originariamente planteado en la Introducción: Dado un haz regular $A_0 + \lambda A_1$ con un autovalor λ_0 , perturbado por un segundo haz $B_0 + \lambda B_1$, determinar la estructura de Weierstrass más habitual asociada a λ_0 como autovalor del haz perturbado (5.3) cuando se da la condición de rango bajo (5.2) para la perturbación, así como describir el conjunto de perturbaciones para las que se da este cambio. Necesitaremos, únicamente, generalizar el Teorema 5.3.4 a un haz no perturbado arbitrario $A_0 + \lambda A_1$ y aplicar el Teorema 5.2.4. Comenzamos transformando este problema en otro relacionado con la clase \mathcal{W}_0 introducida en la Definición 5.3.2. Usando las matrices Y y X de paso a la forma de Weierstrass como en la identidad (5.17), podemos escribir la siguiente igualdad:

$$Y^T(A_0 + \lambda A_1 + B_0 + \lambda B_1)X = \text{diag}(-J_{\lambda_0}, \tilde{J}, I_\infty) + \lambda \text{diag}(I_a, I_b, N) + \tilde{B}_0 + \lambda \tilde{B}_1, \quad (5.17)$$

donde J_{λ_0} es como en (5.4), con bloques cuyos tamaños satisfacen (5.11), \tilde{J} contiene la estructura de Jordan correspondiente a los restantes autovalores finitos, $\tilde{B}_0 := Y^T B_0 X$ y $\tilde{B}_1 := Y^T B_1 X$. La matriz N es suma directa de bloques de Jordan asociados al autovalor 0 y contiene la estructura de Jordan del autovalor infinito. El término derecho de esta igualdad puede escribirse como:

$$\text{diag}(-J_{\tilde{n}_1}^{r_1}(0), \dots, -J_{\tilde{n}_q}^{r_q}(0), \tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N) + (\lambda - \lambda_0) \text{diag}(I_a, I_b, N) + \tilde{B}_0 + \lambda_0 \tilde{B}_1 + (\lambda - \lambda_0) \tilde{B}_1,$$

donde $\det(\text{diag}(\tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N)) \neq 0$. Definamos, en aras de la brevedad,

$$\mu := \lambda - \lambda_0,$$

luego el término derecho de la igualdad (5.17) puede expresarse como

$$\text{diag}(-J_0, \tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N) + \mu \text{diag}(I_a, I_b, N) + \tilde{B}_0 + \lambda_0 \tilde{B}_1 + \mu \tilde{B}_1, \quad (5.18)$$

con J_0 como en (5.10). Esta ecuación muestra que podemos restringirnos al problema original de estudiar el cambio de la estructura de Weierstrass del autovalor *cero* del haz regular $\text{diag}(-J_0, \tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N) + \mu \text{diag}(I_a, I_b, N)$ perteneciente a la clase \mathcal{W}_0 cuando es perturbado por el haz $(\tilde{B}_0 + \lambda_0 \tilde{B}_1) + \mu \tilde{B}_1$. Es importante tener en mente que $\text{rg}(\tilde{B}_0 + \lambda_0 \tilde{B}_1) = \text{rg}(B_0 + \lambda_0 B_1)$ y $\text{rg} \tilde{B}_1 = \text{rg} B_1$. Obviamente, el Teorema 5.3.4 se cumple para el haz en la variable μ que aparece en (5.18), con $(\tilde{B}_0 + \lambda_0 \tilde{B}_1)$ haciendo el papel de C_0 , y \tilde{B}_1 el papel de C_1 . En este caso es conveniente expresar los menores (5.15) y (5.16) que aparecen en el Teorema 5.3.4 en términos de las matrices B_0 , B_1 , Y y X . Hacer esto requiere una cierta notación preliminar.

5.4.1. Notación

Particionemos las matrices Y y X que aparecen en (5.17) como sigue:

$$Y = \left[Y_1 \mid \hat{Y} \right], \quad X = \left[X_1 \mid \hat{X} \right], \quad (5.19)$$

donde X_1 e Y_1 tienen dimensiones $n \times a$, y

$$X_1 = \left[X_1^1 \mid \dots \mid X_1^{r_1} \mid \dots \mid X_q^1 \mid \dots \mid X_q^{r_q} \right], \quad (5.20)$$

$$Y_1 = \left[\begin{array}{c|c|c|c|c|c} Y_1^1 & \dots & Y_1^{r_1} & \dots & Y_q^1 & \dots & Y_q^{r_q} \end{array} \right], \quad (5.21)$$

donde las particiones son conformes con los bloques de Jordan de λ_0 en (5.4).

Las columnas de cada X_j^k forman una cadena de Jordan derecha de $A_0 + \lambda A_1$ de longitud \bar{n}_j correspondiente al autovalor λ_0 . La l -ésima columna de X_j^k es un vector de Jordan derecho de orden l . En particular, si denotamos por x_j^k a la primera columna de X_j^k , cada x_j^k es un autovector derecho de $A_0 + \lambda A_1$ asociado a λ_0 . Análogamente, las columnas de cada Y_j^k forman una cadena de Jordan izquierda de $A_0 + \lambda A_1$ de longitud \bar{n}_j correspondiente a λ_0 . La l -ésima columna, contando de derecha a izquierda, de Y_j^k es un vector de Jordan izquierdo de orden l . Por tanto, si denotamos por y_j^k a la última (i.e. la \bar{n}_j -ésima) columna de Y_j^k , cada y_j^k es un autovector izquierdo correspondiente a λ_0 . Con estos autovectores construimos las matrices

$$L_j = [y_j^1, \dots, y_j^{r_j}], \quad R_j = [x_j^1, \dots, x_j^{r_j}],$$

para $j = 1, \dots, q$, y

$$W_i = [L_i, \dots, L_q], \quad Z_i = [R_i, \dots, R_q],$$

para $i = 1, \dots, q$. Se puede ver fácilmente que las matrices

$$\Phi_{0,j} = W_j^T (B_0 + \lambda_0 B_1) Z_j, \quad j = 1, \dots, q \quad (5.22)$$

$$\Phi_{1,j} = W_j^T B_1 Z_j, \quad j = 2, \dots, q, \quad (5.23)$$

corresponden a las matrices $\psi_{0,j}$ y $\psi_{1,j}$ definidas en el párrafo que sigue a (5.12), tomando las matrices $\tilde{B}_0 + \lambda_0 \tilde{B}_1$ y \tilde{B}_1 que aparecen en (5.18) como C_0 y C_1 . Nótese que, debido a las definiciones acumulativas de W_j y Z_j , cada $\Phi_{i,j}$ es el bloque inferior derecho de tamaño $f_j \times f_j$ de $\Phi_{i,j-1}$. La definición de $\Phi_{1,1}$ depende, como en (5.13), del caso específico que consideremos. Definamos

$$\rho_0 := \text{rg}(B_0 + \lambda_0 B_1) = \text{rg}(\tilde{B}_0 + \lambda_0 \tilde{B}_1), \quad \rho_1 := \text{rg}(B_1) = \text{rg}(\tilde{B}_1), \quad \rho := \rho_0 + \rho_1,$$

análogamente a como se hizo en el Teorema 5.3.4, y, de nuevo, definamos d_0 como el número de bloques de Jordan en J_0 de dimensión mayor que 1, donde J_0 es la matriz que aparece en (5.18). Obviamente, d_0 es también el número de bloques de Jordan de dimensión mayor que 1 asociados al autovalor λ_0 en la forma canónica de Weierstrass del haz $A_0 + \lambda A_1$. Ahora podemos definir

$$\Phi_{1,1} := W_q^T B_1 Z_q \quad \text{si } \rho \leq d_0 \quad \text{y} \quad \Phi_{1,1} := W_q^T B_1 Z_q + E_q \quad \text{si } \rho > d_0, \quad (5.24)$$

donde E_q es la matriz que aparece en (5.13).

Ahora procedemos como en (5.14–5.16). Si $\rho \leq d_0$, sea $s \in \{1, \dots, q\}$ tal que $\rho = r_q + \dots + r_{q-s+2} + \gamma$ con $0 < \gamma \leq r_{q-s+1}$, y sea $I = \{i_1, \dots, i_\gamma\}$ cualquier subconjunto de $\{1, \dots, r_{q-s+1}\}$ con cardinal γ . Denotamos por

$$\Phi_{0,q-s+1}^{(I)} \quad (\text{resp. } \Phi_{1,q-s+1}^{(I)})$$

la submatriz principal de dimensión ρ de $\Phi_{0,q-s+1}$ (resp. de $\Phi_{1,q-s+1}$) que contiene a $\Phi_{0,q-s}$ (resp. $\Phi_{1,q-s}$) junto con las γ filas y columnas de $\Phi_{0,q-s+1}$ (resp. de $\Phi_{1,q-s+1}$) correspondientes a los γ autovectores de las cadenas de Jordan de longitud \bar{n}_{q-s+1} del autovalor λ_0

que están indexados por i_1, \dots, i_γ . Nótese que si $\gamma = r_{q-s+1}$ entonces $\Phi_{0,q-s+1}^{(I)} = \Phi_{0,q-s+1}$ (resp. $\Phi_{1,q-s+1}^{(I)} = \Phi_{1,q-s+1}$).

Si $J = (J_1, J_2)$, donde $J_1 = \{j_1, \dots, j_{\rho_0}\}$, $J_2 = \{k_1, \dots, k_{\rho_0}\}$ son subconjuntos de $\{1, \dots, \rho\}$, denotamos por

$$\Delta_{0,q-s+1}^{(I,J)} \quad (\text{resp. } \Delta_{1,q-s+1}^{(I,J^c)}) \quad (5.25)$$

el menor de orden $\rho_0 \times \rho_0$ de $\Phi_{0,q-s+1}^{(I)}$ correspondiente a las filas y columnas indexadas por, respectivamente, j_1, \dots, j_{ρ_0} y k_1, \dots, k_{ρ_0} (resp. el menor $\rho_1 \times \rho_1$ de $\Phi_{1,q-s+1}^{(I)}$ complementario a $\Delta_{0,q-s+1}^{(I,J)}$, i.e., el que contiene todas las filas y columnas excepto las indexadas por j_1, \dots, j_{ρ_0} y k_1, \dots, k_{ρ_0} , respectivamente).

En el caso $\rho > d_0$, denotamos por

$$\Delta_0^{(J)} \quad (\text{resp. } \Delta_1^{(J^c)})$$

el menor de orden $\rho_0 \times \rho_0$ de $\Phi_{0,q}$ (resp. el menor $(g - \rho_0) \times (g - \rho_0)$ de $\Phi_{1,q}$) que contiene las filas y columnas indexadas por J (resp. el que contiene todas las filas y columnas salvo las indexadas por J).

5.4.2. El resultado principal

Como consecuencia inmediata de la notación y los desarrollos introducidos anteriormente, podemos enunciar el siguiente teorema, que es una versión del Teorema 5.3.4 para haces de matrices.

Teorema 5.4.1 *Sea λ_0 un autovalor del haz matricial complejo regular de tamaño $n \times n$ $A_0 + \lambda A_1$ con forma canónica de Weierstrass dada por (5.17), y supongamos que los tamaños de los bloques de Jordan asociados a λ_0 satisfacen (5.11). Sean a y g , respectivamente, las multiplicidades algebraica y geométrica de λ_0 como autovalor de $A_0 + \lambda A_1$. Sea d_0 el número de bloques de Jordan asociados a λ_0 en (5.17) cuya dimensión es mayor que 1. Sea $B_0 + \lambda B_1$ otro haz $n \times n$, y*

$$\rho_0 := \text{rg}(B_0 + \lambda_0 B_1), \quad \rho_1 := \text{rg}(B_1), \quad \rho := \rho_0 + \rho_1.$$

Sea

$$\tilde{a} = a + \rho_1 - n_1 - \dots - n_\rho,$$

y sea $n_m = 1$ para cada $m \leq 0$. Si $\rho_0 < g$ entonces

$$\det(A_0 + \lambda A_1 + B_0 + \lambda B_1) = (\lambda - \lambda_0)^{\tilde{a}} q(\lambda - \lambda_0), \quad (5.26)$$

donde $q(\lambda - \lambda_0)$ es un polinomio de grado, a lo sumo, $n - \tilde{a}$.

Más aún, el coeficiente de $(\lambda - \lambda_0)^{\tilde{a}}$ en (5.26) está dado por:

$$q(0) = \det(X^{-1} Y^{-T}) \cdot K_0 \cdot \det(\text{diag}(\tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N)), \quad (5.27)$$

donde $\det(\text{diag}(\tilde{J} + \lambda_0 I_b, I_\infty + \lambda_0 N)) \neq 0$ y K_0 está dado por:

(i) si $\rho \leq d_0$ entonces

$$K_0 = \sum_I \sum_J (-1)^{\alpha_{I,J}} \Delta_{0,q-s+1}^{(I,J)} \Delta_{1,q-s+1}^{(I,J^c)},$$

donde I varía sobre todos los subconjuntos de $\{1, \dots, r_{q-s+1}\}$ con γ elementos y $J = (J_1, J_2)$, donde J_1 y J_2 varían, independientemente, entre todos los subconjuntos de $\{1, \dots, \rho\}$ con ρ_0 elementos. El signo $(-1)^{\alpha_{I,J}}$ depende sólo de los conjuntos I y J . Nótese que si $\gamma = r_{q-s+1}$ entonces la suma en I no aparece.

(ii) si $\rho > d_0$ entonces

$$K_0 = \sum_J (-1)^{\alpha_J} \Delta_0^{(J)} \Delta_1^{(J^c)},$$

donde $J = (J_1, J_2)$ y J_1 y J_2 varían, independientemente, entre todos los subconjuntos $\{1, \dots, g\}$ con ρ_0 elementos. El signo $(-1)^{\alpha_J}$ depende sólo del conjunto J .

Demostración. Aplíquese el Teorema 5.3.4 al término de la izquierda en (5.18). ■

De acuerdo con el Teorema 5.4.1 la multiplicidad algebraica de λ_0 como autovalor del haz no perturbado $A_0 + \lambda A_1 + B_0 + \lambda B_1$ es $\tilde{\alpha}$ si y sólo si $K_0 \neq 0$. Recordemos que $\tilde{\alpha}$ es precisamente la cantidad de la derecha en (5.7), luego la igualdad en (5.7) se da si y sólo si $K_0 \neq 0$. Esta observación es la clave para el resultado más importante de esta sección.

Teorema 5.4.2 Sea λ_0 un autovalor del haz matricial complejo regular $A_0 + \lambda A_1$ de tamaño $n \times n$ cuya forma canónica de Weierstrass está dada por (5.17), y sea g la multiplicidad geométrica de λ_0 en $A_0 + \lambda A_1$. Sea $B_0 + \lambda B_1$ un haz $n \times n$, y sean

$$\rho_0 := \text{rg}(B_0 + \lambda_0 B_1), \quad \rho_1 := \text{rg}(B_1), \quad \rho := \rho_0 + \rho_1.$$

Sea K_0 como en el enunciado del Teorema 5.4.1 y $n_1 \leq \dots \leq n_g$ las dimensiones de los bloques de Jordan asociados con λ_0 en $A_0 + \lambda A_1$. Si $\rho_0 < g$ entonces λ_0 es un autovalor del haz perturbado

$$A_0 + B_0 + \lambda(A_1 + B_1). \quad (5.28)$$

Más aún, en este caso, $K_0 \neq 0$ si y sólo si $\dim \ker(A_0 + B_0 + \lambda(A_1 + B_1)) = g - \rho_0$ y las dimensiones de los bloques de Jordan de λ_0 en la forma canónica de Weierstrass de $A_0 + B_0 + \lambda(A_1 + B_1)$ se obtienen eliminando los primeros ρ miembros en la serie $n_g, \dots, n_1, \underbrace{1, \dots, 1}_{\rho_1}$.

Observación 5.4.3 1. Para el autovalor infinito de $A_0 + \lambda A_1$ se tiene un resultado análogo aplicando el Teorema 5.4.2 a los haces duales $A_1 + \lambda A_0$ y $A_1 + B_1 + \lambda(A_0 + B_0)$.

2. La condición $\dim \ker(A_0 + B_0 + \lambda_0(A_1 + B_1)) = g - \rho_0$ que aparece en el Teorema 5.4.2, junto con el enunciado sobre los bloques de Jordan, garantiza que (5.28) es regular y, por lo tanto, $K_0 \neq 0$: si apareciera algún bloque singular en su forma canónica de Kronecker junto con los $g - \rho_0$ bloques de Jordan de λ_0 entonces tendríamos $\dim \ker(A_0 + B_0 + \lambda_0(A_1 + B_1)) > g - \rho_0$. Sin esta condición, en cambio, es posible que (5.28) sea singular (lo que implicaría que $K_0 = 0$), y tenga al mismo tiempo los bloques de Jordan que se han descrito asociados a λ_0 en la parte regular de su forma canónica de Kronecker. El siguiente ejemplo es una muestra de esta situación.

Ejemplo 5.4.4 Sea el haz

$$H(\lambda) = (\lambda I_2 - J_2(0)) \oplus (J_1(\infty)) = \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Si lo perturbamos por

$$M(\lambda) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix},$$

obtenemos

$$H(\lambda) + M(\lambda) = \begin{bmatrix} \lambda & 1 & -1 \\ 0 & \lambda & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

cuya forma de Kronecker es

$$K_{H+M}(\lambda) = \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

es decir: $H + M$ es singular y en la parte regular se da el cambio genérico descrito en el Teorema 5.4.2. \square

Capítulo 6

Perturbaciones de rango bajo de la Forma de Kronecker

6.1. Introducción

La continuación natural de los problemas planteados en el capítulo anterior consiste en tratar de describir el cambio genérico de la forma de Kronecker tras perturbaciones de rango bajo, así como el conjunto de perturbaciones para las que dicho cambio se produce. A ello está dedicado el presente capítulo.

A lo largo de este capítulo, $H(\lambda)$ es un haz singular $m \times n$ que no tiene *rango completo*, es decir, $\text{rg}(H) < \min\{m, n\}$. En esta situación, dado un número entero ρ tal que

$$0 < \rho \leq \min\{m, n\} - \text{rg}(H), \quad (6.1)$$

el conjunto de perturbaciones está restringido a los haces $M(\lambda)$ con $\text{rg}(M) = \rho$. Nótese que (6.1) es una condición de rango bajo impuesta a las perturbaciones.

Para el conjunto de perturbaciones definidas en el párrafo previo, el primer problema que abordamos es el de obtener información sobre la estructura regular genérica del haz perturbado $(H + M)(\lambda)$. La descripción del cambio en la estructura regular se encuentra en el apartado §6.4, donde se demuestra que, genéricamente, si λ_0 es un autovalor de $H(\lambda)$, finito o infinito, entonces λ_0 es también un autovalor de $(H + M)(\lambda)$ con multiplicidades parciales mayores o iguales que las correspondientes multiplicidades parciales de λ_0 relativas a $H(\lambda)$. Estos resultados son consecuencia del Teorema 6.4.5, que es nuestra primera contribución original relevante de este capítulo, y contrastan con lo que se vio en el capítulo anterior para el cambio de la forma de Weierstrass. En aquella ocasión, las multiplicidades parciales mayores desaparecen, y, de las restantes, algunas pueden verse reducidas. La razón de esta diferencia está en que las condiciones de rango bajo que impondremos en el caso singular son muy distintas a las que se imponían en el caso regular. Ahora se trata de condiciones sobre el rango de un haz (es decir: el rango para cualquier valor de λ) en lugar de referirse al rango de una matriz (la matriz $A_0 + \lambda_0 A_1$ en el caso regular).

El segundo problema que abordamos es el de obtener información acerca de los índices minimales genéricos de $(H + M)(\lambda)$. En aras de la brevedad, desarrollamos los resultados solamente para los índices minimales derechos (o de columnas), pero se tienen resultados análogos para los índices minimales izquierdos (o de filas). Es sabido que el número de índices minimales derechos de $H(\lambda)$ es $n - \text{rg}(H)$. El resultado inicial que presentamos

es que, genéricamente, $H + M$ tiene $n - \text{rg}(H) - \rho$ índices minimales derechos. Esto implica, en particular, que si $\rho = n - \text{rg}(H)$ entonces $H + M$ no tiene índices minimales derechos, es decir, tiene rango completo por columnas. Estos resultados se siguen del Teorema 6.3.1 y su consecuencia inmediata, el Corolario 6.3.2. El caso $\rho < n - \text{rg}(H)$ es mucho más difícil y se incluye en el Teorema 6.5.10, donde determinamos todos los índices minimales derechos genéricos de $H + M$ si, además del rango, se conoce el orden singular derecho de la perturbación $M(\lambda)$ (véase la Definición 2.3.4). Como corolario, el Teorema 6.5.12 presenta información parcial sobre los índices minimales derechos genéricos del haz perturbado $H + M$ en el caso en que $\rho = \text{rg}(M)$ es la única información conocida de las perturbaciones. Hablando de manera informal, se puede decir que los índices minimales derechos genéricos de $H + M$ son mayores o iguales que los mayores $n - \text{rg}(H) - \rho$ índices minimales derechos de H . Los Teoremas 6.5.10 y 6.5.12 constituyen nuestra segunda contribución original relevante de este capítulo. Todos los resultados descritos previamente continúan siendo válidos en el caso límite

$$\rho = \text{mín}\{m, n\} - \text{rg}(H). \quad (6.2)$$

Si se asume la desigualdad estricta en (6.1), es decir: $\text{rg}(H) + \text{rg}(M) < \text{mín}\{m, n\}$, es posible determinar completamente la FCK genérica de $(H + M)(\lambda)$ en términos de los órdenes singulares derecho e izquierdo y de la estructura regular de la FCK de $M(\lambda)$. En el caso de que $\text{rg}(M)$ sea la única información conocida de las perturbaciones, la FCK genérica de $(H + M)(\lambda)$ sólo puede ser determinada parcialmente. Estos resultados aparecen en los Teoremas 6.6.2 y 6.6.5, que son nuestra tercera contribución original relevante de este capítulo. Debe ser notado que todos los resultados genéricos sobre la FCK de $(H + M)(\lambda)$ que presentamos son muy sencillos de describir, aunque demostrar que ocurren bajo ciertas condiciones suficientes de genericidad es una ardua tarea que requiere técnicas muy diferentes a las utilizadas en el capítulo anterior.

La última sección de este capítulo está dedicada a describir el cambio de la FCK en el que denominamos el “caso límite”, es decir, aquel en el que $\text{rg}(H) + \text{rg}(M) = \text{mín}\{m, n\}$. Como ya hemos comentado, en este caso, el haz perturbado $H + M$ no tiene índices minimales derechos (si $n \leq m$) o izquierdos (si $m \leq n$). Esto significa, hablando sin rigor, que la estructura singular derecha (si $n \leq m$) o izquierda (si $m \leq n$) de H se “convierte” en regular. En la Sección 6.7 determinamos cuál es esa nueva parte regular, lo que nos permite obtener la FCK genérica del haz perturbado $H + M$. Esta forma genérica está compuesta de:

- 1) La estructura regular de H ,
- 2) la estructura regular que se crea a partir de la estructura singular de H ,
- 3) la estructura singular derecha (si $m \leq n$) o izquierda (si $n \leq m$) cuyos índices minimales describe el Teorema 6.5.10.

El comportamiento genérico (descrito en el Teorema 6.5.10) del cambio en la FCK de haces singulares depende del rango de la perturbación, como es natural, y también de los órdenes singulares de dicha perturbación. Este hecho, a priori sorprendente, se volverá más natural a la luz de los resultados del Capítulo 7, donde demostraremos que el conjunto de haces matriciales de rango menor o igual que r se descompone en $r + 1$ variedades irreducibles, asociada cada una de ellas a uno de los órdenes singulares que pueden darse en haces de rango r (es decir: los números naturales desde 0 hasta r).

En este capítulo seguiremos las definiciones y notaciones de los apartados § 2.3 y § 2.4. En las definiciones correspondientes a este último, que está dedicado esencialmente a polinomios de grado arbitrario, hemos de considerar el caso particular de un polinomio de grado 1, lo que supondrá, en general, sustituir la notación $L(\lambda)$ (polinomio) por $H(\lambda)$ (haz).

La FCK de una suma directa, $H(\lambda) \oplus M(\lambda)$, de dos haces matriciales $H(\lambda)$ y $M(\lambda)$ es la suma directa de las FCKs de $H(\lambda)$ y $M(\lambda)$, salvo permutaciones de los bloques diagonales. Por lo tanto, la característica de Segre de λ_0 relativa a $L(\lambda) \oplus M(\lambda)$ se obtiene simplemente uniendo las características de Segre de λ_0 relativas a $L(\lambda)$ y a $M(\lambda)$ y reordenando la sucesión resultante. Esto se sigue de [26, Th. 5, Ch. VI].

6.2. Desarrollo de un haz como suma de haces de rango 1

El desarrollo que se presenta en el Lema 6.2.1 tiene un papel clave en lo que seguirá a continuación.

Lema 6.2.1 *Sea $M(\lambda)$ un haz matricial $m \times n$ de rango ρ , y sea $\tilde{\varepsilon}$ su orden singular derecho. Entonces $M(\lambda)$ se puede expresar en la forma*

$$M(\lambda) = v_1(\lambda)w_1^T(\lambda) + \dots + v_\rho(\lambda)w_\rho^T(\lambda) \quad (6.3)$$

donde

- (i) $\{v_1(\lambda), \dots, v_\rho(\lambda)\}$ es un conjunto linealmente independiente de vectores polinómicos de $\mathbb{C}^m(\lambda)$ cuyos grados son, a lo sumo, 1;
- (ii) $\{w_1(\lambda), \dots, w_\rho(\lambda)\}$ es un conjunto linealmente independiente de vectores polinómicos de $\mathbb{C}^n(\lambda)$ cuyos grados son, a lo sumo, 1;
- (iii) cada sumando $v_i(\lambda)w_i^T(\lambda)$, $1 \leq i \leq \rho$, es un haz $m \times n$ de rango 1, es decir, $v_i(\lambda)$ y $w_i(\lambda)$ tienen ambos grado cero, o bien uno de ellos tiene grado 0 y el otro grado 1;
- (iv) hay $\tilde{\varepsilon}$ vectores entre $w_1(\lambda), \dots, w_\rho(\lambda)$ con grado exactamente 1, y los restantes vectores $w_i(\lambda)$ tienen grado 0.

Una descomposición (6.3) que satisface estas condiciones se llamará una **descomposición derecha de $M(\lambda)$** . Cualquier otra descomposición de $M(\lambda)$ como suma de ρ haces matriciales de rango 1 contiene, al menos, $\tilde{\varepsilon}$ vectores entre $w_1(\lambda), \dots, w_\rho(\lambda)$ de grado exactamente 1.

Demostración. El resultado es una consecuencia directa de la FCK. Sea

$$\mathcal{K}_M(\lambda) = \text{diag}(L_{\tilde{\varepsilon}_1}, \dots, L_{\tilde{\varepsilon}_h}, L_{\tilde{\eta}_1}^T, \dots, L_{\tilde{\eta}_l}^T, \mathcal{J}_M),$$

la FCK de M , donde \mathcal{J}_M es la estructura regular de $M(\lambda)$, y existen dos matrices invertibles, X e Y , tales que $M(\lambda) = (Y^T)^{-1}\mathcal{K}_M(\lambda)X^{-1}$. Ahora, nótese que un bloque $L_{\tilde{\varepsilon}_i}$ se puede expresar como una suma de $\tilde{\varepsilon}_i$ haces de rango 1,

$$\begin{bmatrix} \lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [\lambda \quad -1 \quad \dots \quad 0] + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} [0 \quad \dots \quad \lambda \quad -1],$$

donde los vectores fila (resp. columna) tienen grado igual a 1 (resp. 0) (véase el Lema 2.3.6). Se puede obtener un desarrollo análogo para un bloque $L_{\tilde{\eta}_j}^T$ por transposición, pero ahora los vectores columna (resp. fila) tienen grado igual a 1 (resp. 0). Para los bloques de Jordan de \mathcal{J}_M , correspondientes a los autovalores finitos e infinitos, se tienen unos desarrollos similares con vectores fila de grado 0. Todos estos desarrollos se pueden combinar con $M(\lambda) = (Y^T)^{-1}\mathcal{K}_M(\lambda)X^{-1}$ para demostrar de forma directa los cuatro apartados del Lema.

Demostremos ahora el hecho de que cualquier otra descomposición de $M(\lambda)$ como suma de ρ haces de rango 1 contiene al menos $\tilde{\varepsilon}$ vectores entre $w_1(\lambda), \dots, w_\rho(\lambda)$ con grado exactamente 1. Nótese que el conjunto de soluciones de $M(\lambda)x(\lambda) = 0$ coincide con el conjunto de soluciones de $[w_1(\lambda), \dots, w_\rho(\lambda)]^T x(\lambda) = 0$ y, por lo tanto, los índices minimales derechos de los haces $M(\lambda)$ y $D_0 + \lambda D_1 \equiv [w_1(\lambda), \dots, w_\rho(\lambda)]^T$ son iguales. Si hubiera menos de $\tilde{\varepsilon}$ vectores entre $w_1(\lambda), \dots, w_\rho(\lambda)$ con grado exactamente 1, entonces $\text{rg}(D_1) < \tilde{\varepsilon}$. Esto implicaría que la matriz coeficiente de λ en la FCK de $[w_1(\lambda), \dots, w_\rho(\lambda)]^T$ tiene rango menor que $\tilde{\varepsilon}$, lo que está en contradicción con el hecho de que $\tilde{\varepsilon}$ es el orden singular derecho. ■

Observación 6.2.2 *Se puede obtener un resultado similar al enunciado en el Lema 6.2.1 considerando el orden singular izquierdo, $\tilde{\eta}$, del haz $M(\lambda)$ y escribiendo los vectores columna de los desarrollos de los bloques de Jordan de \mathcal{J}_M como vectores de grado 0. En este caso, obtendremos una **descomposición izquierda** de $M(\lambda)$:*

$$M(\lambda) = \hat{v}_1(\lambda)\hat{w}_1^T(\lambda) + \dots + \hat{v}_\rho(\lambda)\hat{w}_\rho^T(\lambda), \quad (6.4)$$

donde los vectores $\{\hat{v}_1(\lambda), \dots, \hat{v}_\rho(\lambda)\}$ y $\{\hat{w}_1(\lambda), \dots, \hat{w}_\rho(\lambda)\}$ tienen las propiedades que se enuncian en los apartados (i), (ii), y (iii) del Lema 6.2.1, mientras que (iv) se sustituye por la condición: “hay $\tilde{\eta}$ vectores entre $\{\hat{v}_1(\lambda), \dots, \hat{v}_\rho(\lambda)\}$ con grado exactamente 1, y los restantes vectores tienen grado 0”.

Observación 6.2.3 *Las descomposiciones derecha e izquierda no son únicas. Además, una descomposición izquierda de $M(\lambda)$ puede no ser simultáneamente una descomposición derecha. De hecho, como puede verse siguiendo la demostración del Lema 6.2.1, una descomposición de $M(\lambda)$ de la forma (6.4) será al mismo tiempo una descomposición derecha e izquierda si y sólo si $M(\lambda)$ no tiene estructura regular.*

Ejemplo 6.2.4 Vamos a mostrar una descomposición derecha y otra izquierda de un haz con $\tilde{\varepsilon} = 1$, $\tilde{\eta} = 0$ y $\rho = 2$. Dicho haz ha de tener estructura regular 1×1 . Tomemos, por ejemplo:

$$M(\lambda) = \begin{bmatrix} \lambda & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Tenemos las descomposiciones derecha e izquierda:

$$\begin{bmatrix} \lambda & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \lambda & -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \lambda \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} \lambda & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \lambda & -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \lambda \end{bmatrix}.$$

Es imposible obtener una descomposición que sea, al mismo tiempo, descomposición derecha e izquierda de $M(\lambda)$. \square

Ejemplo 6.2.5 La siguiente descomposición de un haz sin estructura regular es, al mismo tiempo, descomposición derecha e izquierda:

$$\begin{bmatrix} \lambda & -1 & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} [\lambda \quad -1 \quad 0] + \begin{bmatrix} 0 \\ \lambda \\ -1 \end{bmatrix} [0 \quad 0 \quad 1].$$

\square

Observamos, finalmente, que el desarrollo (2.10) no es, en general, un desarrollo del tipo (6.3) para el caso en que $L(\lambda)$ es un haz, porque los sumandos $h_i(L) a_i(\lambda) z_i^T(\lambda)$ tienen, en general, grado mayor que 1.

6.3. Hipótesis de rango bajo: significado y genericidad

A lo largo de esta sección trataremos, como en el capítulo anterior, con tres haces complejos de tamaño $m \times n$: el haz (fijo) *no perturbado* lo denotaremos $H(\lambda)$, el haz *perturbación* se denotará $M(\lambda)$, y el haz *perturbado* es, como en el capítulo anterior, la suma $(H+M)(\lambda)$. El haz $H(\lambda)$ no tiene rango completo y su FCK se supone conocida y la denotaremos por $\mathcal{K}_H(\lambda)$, como en (2.5). Omitiremos con frecuencia la variable λ cuando no haya riesgo de confusión.

Como se ha anunciado al comienzo del capítulo, el conjunto de perturbaciones que consideramos es el conjunto de haces

$$\mathcal{C} = \{M(\lambda) : \text{rg}(M) = \rho\}, \quad (6.5)$$

donde $\rho > 0$ es un número entero tal que

$$\text{rg}(H) + \rho \leq \min\{m, n\}. \quad (6.6)$$

Esta es la condición de rango bajo que imponemos al conjunto de perturbaciones. Nótese, además, que $\rho > 0$ y (6.6) implican que $\text{rg}(H) < \min\{m, n\}$, es decir, que $H(\lambda)$ no tiene rango completo.

Un resultado clave en esta sección es que la propiedad

$$\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M), \quad (6.7)$$

es genérica en el conjunto \mathcal{C} . Esto se demuestra de manera rigurosa en el Teorema 6.3.1 más adelante. Combinando este resultado con la identidad (2.6), se puede decir que, para perturbaciones que pertenecen al conjunto \mathcal{C} , los haces perturbados $(H+M)(\lambda)$ tienen, genéricamente, $n - \text{rg}(H) - \rho$ índices minimales derechos y $m - \text{rg}(H) - \rho$ índices minimales izquierdos. Véase el Corolario 6.3.2 posterior acerca de este particular.

Nótese que, teniendo en cuenta que $H+M$ es un haz $m \times n$, la condición (6.7) implica que $\text{rg}(H) + \text{rg}(M) \leq \min\{m, n\}$, es decir: la hipótesis (6.6), y también que $H(\lambda)$ no tiene rango completo si $M(\lambda)$ no es el haz nulo, es decir, para perturbaciones no triviales.

Estos hechos y la genericidad de (6.7) en el conjunto \mathcal{C} nos llevan a imponer la condición $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$ en la mayoría de los lemas y teoremas que demostraremos, sin mencionar explícitamente la condición de rango bajo (6.6).

El apartado §6.5 está dedicado al estudio de los índices minimales del haz $(H + M)(\lambda)$. En dicho apartado, la condición

$$\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n \quad (6.8)$$

se asume como hipótesis en muchos de los resultados. Nótese que (6.8) es consecuencia de (6.7) sólo si $\min\{m, n\} = m < n$. De lo contrario, (6.8) es una hipótesis adicional. El motivo de suponer (6.8) en la Sección 6.5 es que, de acuerdo con (2.6), el número de índices minimales derechos de $H + M$ es $n - \text{rg}(H + M)$, que es cero si $\text{rg}(H + M) = n$. Por tanto, el estudio de los índices minimales derechos genéricos de $H + M$ sólo tiene sentido si (6.8) se da. Para el caso de los índices minimales izquierdos se usará m en lugar de n en la condición (6.8).

6.3.1. Genericidad de la hipótesis $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$. Número de índices minimales de $H + M$

El propósito de este apartado es presentar una demostración rigurosa de la genericidad de la hipótesis más crucial de nuestra teoría de perturbaciones de rango bajo de haces singulares. Esta hipótesis determina el número genérico de índices minimales derechos e izquierdos del haz perturbado $(H + M)(\lambda)$.

Teorema 6.3.1 Sean $H(\lambda)$ un haz matricial complejo de tamaño $m \times n$ y ρ un número entero positivo tal que $\text{rg}(H) + \rho \leq \min\{m, n\}$. Entonces, el conjunto de haces matriciales complejos $m \times n$

$$\mathcal{G} = \{M(\lambda) \text{ haz } m \times n : \text{rg}(M) = \rho \text{ y } \text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)\}$$

es denso y abierto en el conjunto de haces matriciales complejos $m \times n$ de rango ρ .

Demostración. Primero, demostraremos que \mathcal{G} es denso en el conjunto de haces de rango ρ . Nótese que $\text{rg}(H + E) \leq \text{rg}(H) + \text{rg}(E)$ para cualquier haz $E(\lambda)$. Por lo tanto, tenemos que demostrar que, para cada haz $E(\lambda)$ con rango ρ y $\text{rg}(H + E) < \text{rg}(H) + \text{rg}(E)$, existe una sucesión $\{M^{(t)}(\lambda)\}_{t=1}^{\infty} \subset \mathcal{G}$ cuyo límite es $E(\lambda)$. Sea $r \equiv \text{rg}(H)$. De acuerdo con (6.3), podemos escribir

$$\begin{aligned} H(\lambda) &= v_1(\lambda)w_1^T(\lambda) + \dots + v_r(\lambda)w_r^T(\lambda) \\ E(\lambda) &= a_1(\lambda)b_1^T(\lambda) + \dots + a_\rho(\lambda)b_\rho^T(\lambda), \end{aligned}$$

y

$$(H + E)(\lambda) = [v_1 | \dots | v_r | a_1 | \dots | a_\rho] [w_1 | \dots | w_r | b_1 | \dots | b_\rho]^T,$$

donde hemos omitido, en algunos casos, la variable λ por simplicidad. Con argumentos elementales se puede probar que $\text{rg}(H + E) < \text{rg}(H) + \text{rg}(E) = r + \rho$ si y sólo si $\text{rg}[v_1 | \dots | v_r | a_1 | \dots | a_\rho] < r + \rho$ ó $\text{rg}[w_1 | \dots | w_r | b_1 | \dots | b_\rho] < r + \rho$. Supongamos que $\text{rg}[v_1 | \dots | v_r | a_1 | \dots | a_\rho] < r + \rho$. Debido al hecho de que el conjunto $\{v_1, \dots, v_r\}$ es linealmente independiente, esto implica que

$$(i) \text{ rg}[v_1 | \dots | v_r | a_1 | \dots | a_\rho] = r + \hat{\rho} \text{ con } 0 \leq \hat{\rho} < \rho; \text{ y}$$

- (ii) los vectores a_1, \dots, a_ρ se pueden reordenar como $a_{i_1}, \dots, a_{i_{\hat{\rho}}}, a_{k_1}, \dots, a_{k_{\rho-\hat{\rho}}}$, donde $\{v_1, \dots, v_r, a_{i_1}, \dots, a_{i_{\hat{\rho}}}\}$ es un conjunto linealmente independiente.

Ahora, el Lema 2.5.7 se utiliza para demostrar que existen $\rho - \hat{\rho}$ vectores, $u_{j_1}, \dots, u_{j_{\rho-\hat{\rho}}}$, de la base canónica de \mathbb{C}^m tales que, para cada $t = 1, 2, \dots$,

$$\operatorname{rg} \left[v_1 | \dots | v_r | a_{i_1} | \dots | a_{i_{\hat{\rho}}} | a_{k_1} + \frac{1}{t} u_{j_1} | \dots | a_{k_{\rho-\hat{\rho}}} + \frac{1}{t} u_{j_{\rho-\hat{\rho}}} \right] = r + \rho.$$

Sea $\{a_1^{(t)}, \dots, a_\rho^{(t)}\}$ el conjunto de vectores que se obtienen de $\{a_1, \dots, a_\rho\}$ reemplazando $a_{k_1}, \dots, a_{k_{\rho-\hat{\rho}}}$ por $a_{k_1} + \frac{1}{t} u_{j_1}, \dots, a_{k_{\rho-\hat{\rho}}} + \frac{1}{t} u_{j_{\rho-\hat{\rho}}}$. Si $\operatorname{rg} [w_1 | \dots | w_r | b_1 | \dots | b_\rho] < r + \rho$ procedemos de una manera similar para obtener un conjunto de vectores $\{b_1^{(t)}, \dots, b_\rho^{(t)}\}$. Finalmente, definimos la sucesión de haces

$$M^{(t)}(\lambda) = a_1^{(t)}(\lambda)(b_1^{(t)}(\lambda))^T + \dots + a_\rho^{(t)}(\lambda)(b_\rho^{(t)}(\lambda))^T \quad t = 1, 2, \dots$$

Es inmediato comprobar que: (i) $\lim_{t \rightarrow \infty} M^{(t)}(\lambda) = E(\lambda)$; (ii) $\operatorname{rg}(M^{(t)}) = \rho$ para todo t ; y (iii) $\operatorname{rg}(H + M^{(t)}) = \operatorname{rg}(H) + \operatorname{rg}(M^{(t)})$ para todo t . Esto demuestra que \mathcal{G} es denso.

Ahora, demostraremos que \mathcal{G} es abierto en el conjunto, \mathcal{C} , de haces matriciales de rango ρ (definido en (6.5)). Para este fin, observamos, en primer lugar, que $\mathcal{G} \subset \mathcal{C} \subset \mathbb{C}^{2mn}$. Consideramos en \mathcal{C} la *topología del subespacio* inducida por la topología usual de \mathbb{C}^{2mn} , como ya se ha explicado anteriormente. Así pues, para demostrar que \mathcal{G} es abierto en \mathcal{C} , es suficiente demostrar que todo haz $M(\lambda) \in \mathcal{G}$ está incluido en un abierto \mathcal{X}_M de \mathbb{C}^{2mn} tal que

$$\operatorname{rg}(H + E) \geq \operatorname{rg}(H + M) = \operatorname{rg}(H) + \operatorname{rg}(M) \quad \text{para todo } E \in \mathcal{X}_M.$$

Si esto ocurre, entonces:

1. $\mathcal{X}_M \cap \mathcal{C}$ es abierto en \mathcal{C} ; y
2. el hecho de que $\operatorname{rg}(H) + \operatorname{rg}(M) \leq \operatorname{rg}(H + E) \leq \operatorname{rg}(H) + \operatorname{rg}(E)$ para todo $E \in \mathcal{X}_M$ implica que $\operatorname{rg}(H + E) = \operatorname{rg}(H) + \operatorname{rg}(E)$ para todo $E \in \mathcal{X}_M \cap \mathcal{C}$. Esto significa que $\mathcal{X}_M \cap \mathcal{C} \subset \mathcal{G}$ y que M es un punto interior de \mathcal{G} .

Veamos cómo se construye $\mathcal{X}_M \subset \mathbb{C}^{2mn}$. Dado $M \in \mathcal{G}$, la igualdad $\operatorname{rg}(H + M) = \operatorname{rg}(H) + \rho \equiv r + \rho$ implica que el haz $(H + M)(\lambda)$ tiene un menor de orden $(r + \rho) \times (r + \rho)$ que es un polinomio en λ con al menos un coeficiente no nulo. Sea $\det(H + M)(\alpha, \beta)$ este menor, donde los conjuntos $\alpha \subseteq \{1, \dots, m\}$ y $\beta \subseteq \{1, \dots, n\}$ denotan, respectivamente, las filas y columnas que definen el menor. Identificando cada haz $E(\lambda) = E_0 + \lambda E_1$ con un elemento de \mathbb{C}^{2mn} los coeficientes del menor $\det(H + E)(\alpha, \beta)$ definen una función continua $f(E)$, donde $f : \mathbb{C}^{2mn} \rightarrow \mathbb{C}^{r+\rho+1}$, porque estos coeficientes son polinomios en las entradas de las matrices complejas E_0 y E_1 . Teniendo en cuenta que $f(M) \neq 0$, existe una bola abierta, \mathcal{B} , en $\mathbb{C}^{r+\rho+1}$ cuyo centro es $f(M)$ y tal que $0 \notin \mathcal{B}$. Entonces podemos tomar $\mathcal{X}_M = f^{-1}(\mathcal{B})$, que es abierto, de modo que $f(E) \neq 0$ para todo $E \in \mathcal{X}_M$, y, por tanto, $\operatorname{rg}(H + E) \geq r + \rho$ para todo $E \in \mathcal{X}_M$. ■

Como consecuencia del Teorema 6.3.1 y la igualdad (2.6) podemos determinar el número genérico de índices minimales derechos e izquierdos de $H + M$.

Corolario 6.3.2 *Sea $H(\lambda)$ un haz matricial complejo de tamaño $m \times n$ con p índices minimales derechos y q índices minimales izquierdos, y sea ρ un número entero positivo tal que $\operatorname{rg}(H) + \rho \leq \min\{m, n\}$. Entonces, el conjunto de perturbaciones $M(\lambda)$ con $\operatorname{rg}(M) = \rho$ y tales que $H + M$ tiene $p - \rho$ índices minimales derechos y $q - \rho$ índices minimales izquierdos es un abierto denso en el conjunto de haces matriciales complejos $m \times n$ de rango ρ .*

6.4. La estructura regular del haz perturbado

En este apartado, que es el equivalente a §5.2 para haces matriciales singulares, obtendremos información sobre la estructura regular de la FCK del haz perturbado $(H + M)(\lambda)$ en términos de las estructuras regulares de $H(\lambda)$ y $M(\lambda)$, es decir: de \mathcal{J}_H y \mathcal{J}_M . Con la única hipótesis $\operatorname{rg}(H + M) = \operatorname{rg}(H) + \operatorname{rg}(M)$, demostramos que, para cada autovalor de H o de M , la estructura regular de $H + M$ tiene, al menos, tantos bloques como $\mathcal{J}_H \oplus \mathcal{J}_M$, con dimensiones mayores o iguales que las dimensiones de los bloques de $\mathcal{J}_H \oplus \mathcal{J}_M$. Esto significa, en particular, que pueden aparecer más bloques. Este resultado se presenta en el Teorema 6.4.5. En la Sección 6.6 veremos que la estructura regular genérica de $H + M$ es, precisamente, $\mathcal{J}_H \oplus \mathcal{J}_M$ si $\operatorname{rg}(H + M) = \operatorname{rg}(H) + \operatorname{rg}(M) < \min\{m, n\}$.

En esta sección, utilizaremos el lema auxiliar 6.4.1. Aparece sin demostrar en [46], y su demostración es elemental.

Lema 6.4.1 [46, p. 799] *Si $D = \operatorname{diag}(z_1, \dots, z_k)$ y G es una matriz arbitraria $k \times k$, entonces*

$$\det(D + G) = \det G + \sum z_{\nu_1} \dots z_{\nu_j} \cdot \det \check{G}(\nu_1, \dots, \nu_j),$$

donde la suma varía entre todos los $j \in \{1, \dots, k\}$ y todos los ν_1, \dots, ν_j tales que $1 \leq \nu_1 < \dots < \nu_j \leq k$, $\check{G}(\nu_1, \dots, \nu_j)$ denota la matriz obtenida a partir de G eliminando las filas y columnas cuyos índices son ν_1, \dots, ν_j , y $\det \check{G}(1, \dots, k) \equiv 1$.

El Lema 6.4.2 extiende el resultado [76, Th. 1] bajo hipótesis más restrictivas.

Lema 6.4.2 *Sea $L(\lambda)$ un polinomio matricial $m \times n$ con rango r , y $e_1 \leq \dots \leq e_r$ las multiplicidades parciales de λ_0 relativas a $L(\lambda)$. Sea $M(\lambda)$ un polinomio matricial de rango 1, y e la multiplicidad parcial de λ_0 relativa a $M(\lambda)$. Supongamos que $\operatorname{rg}(L + M) = \operatorname{rg}(L) + \operatorname{rg}(M)$, y que $e_i < e \leq e_{i+1}$, para algún $i \in \{0, 1, \dots, r\}$, donde definimos $e_0 \equiv -1$ y $e_{r+1} \equiv \infty$. Entonces, las multiplicidades parciales $f_1 \leq \dots \leq f_{r+1}$ de λ_0 relativas a $(L + M)(\lambda)$ satisfacen:*

$$f_1 = e_1, \dots, f_i = e_i, e \leq f_{i+1}, e_{i+1} \leq f_{i+2}, \dots, e_r \leq f_{r+1}.$$

Observación 6.4.3 *Nótese que, en el Lema 6.4.2, es posible que $e_1 = \dots = e_r = 0$ o que $e = 0$, es decir, λ_0 puede no ser un autovalor de $L(\lambda)$ ni de $M(\lambda)$.*

Demostración del Lema 6.4.2. El Teorema 5.2.1 implica que

$$e_1 \leq f_2, e_2 \leq f_3, \dots, e_r \leq f_{r+1}.$$

Luego, sólo necesitamos probar que $f_1 = e_1, \dots, f_i = e_i, e \leq f_{i+1}$. Sean $U(\lambda)$ y $V(\lambda)$ los polinomios matriciales, con determinante constante no nulo, que transforman L en

su forma normal de Smith, es decir, $U(\lambda)L(\lambda)V(\lambda) = \text{diag}((\lambda - \lambda_0)^{e_1}p_1(\lambda), \dots, (\lambda - \lambda_0)^{e_r}p_r(\lambda), 0, \dots, 0)$, donde los polinomios $p_1(\lambda), \dots, p_r(\lambda)$ satisfacen $p_j(\lambda_0) \neq 0$, para $j = 1, \dots, r$. Los polinomios invariantes y las multiplicidades parciales permanecen inalteradas por la multiplicación de $U(\lambda)$ y $V(\lambda)$, luego podemos centrarnos en las multiplicidades parciales del polinomio matricial:

$$U(\lambda)(L + M)(\lambda)V(\lambda) = \text{diag}((\lambda - \lambda_0)^{e_1}p_1(\lambda), \dots, (\lambda - \lambda_0)^{e_r}p_r(\lambda), 0, \dots, 0) + (\lambda - \lambda_0)^e x(\lambda)y^T(\lambda), \quad (6.9)$$

donde el segundo término del lado derecho es $U(\lambda)M(\lambda)V(\lambda)$ (véase (2.10)).

En el caso $e_0 < e \leq e_1$, i.e., $i = 0$, el exponente del factor $(\lambda - \lambda_0)$ del máximo común divisor de todos los menores 1×1 de (6.9) es mayor o igual que e , luego $e \leq f_1$ por definición de los polinomios invariantes [26, Ch. VI, §3], y el resultado queda demostrado. Supondremos, de ahora en adelante, que $i \geq 1$. En lo que resta de la demostración, probaremos que si c_k , para $k = 1, \dots, r + 1$, denota el exponente del factor $(\lambda - \lambda_0)$ del máximo común divisor de todos los menores $k \times k$ en (6.9), entonces

$$c_1 = e_1, c_2 = e_1 + e_2, \dots, c_i = e_1 + \dots + e_i, c_{i+1} \geq e_1 + \dots + e_i + e. \quad (6.10)$$

Esto y la definición de polinomios invariantes implican que $f_1 = e_1, \dots, f_i = e_i, e \leq f_{i+1}$.

Es fácil ver que la menor potencia de $(\lambda - \lambda_0)$ en un menor 1×1 de (6.9) es e_1 , luego $c_1 = e_1$. Para $k \geq 2$, nótese que todos los menores no nulos de tamaño $k \times k$ de (6.9) han de contener, al menos, $k - 1$ de las $(1, 1), \dots, (r, r)$ entradas diagonales. Entonces, un menor $k \times k$ no nulo de (6.9) ha de ser de uno de estos dos tipos:

(i)

$$\det(\text{diag}((\lambda - \lambda_0)^{e_{i_1}}p_{i_1}(\lambda), \dots, 0, \dots, (\lambda - \lambda_0)^{e_{i_{k-1}}}p_{i_{k-1}}(\lambda)) + (\lambda - \lambda_0)^e [x(\lambda)y^T(\lambda)]_k),$$

(ii)

$$\det(\text{diag}((\lambda - \lambda_0)^{e_{i_1}}p_{i_1}(\lambda), \dots, (\lambda - \lambda_0)^{e_{i_k}}p_{i_k}(\lambda)) + (\lambda - \lambda_0)^e [x(\lambda)y^T(\lambda)]_k), \quad (6.11)$$

donde $[x(\lambda)y^T(\lambda)]_k$ es una cierta submatriz $k \times k$ de $x(\lambda)y^T(\lambda)$. Si a estos menores les aplicamos el Lema 6.4.1, vemos que

(i) Todo menor del tipo (i) puede escribirse como

$$(\lambda - \lambda_0)^{e_1 + \dots + e_{k-1} + e} q(\lambda), \quad (6.12)$$

donde $q(\lambda)$ es un polinomio.

(ii) Todo menor del tipo (ii) puede escribirse como

$$(\lambda - \lambda_0)^{e_1 + \dots + e_{k-1} + \min\{e, e_k\}} t(\lambda), \quad (6.13)$$

donde $t(\lambda)$ es un polinomio.

En el caso $k = i + 1$, estos resultados implican directamente que $c_{i+1} \geq e_1 + \dots + e_i + e$. En el caso $k \leq i$, (6.12) y (6.13) implican que $c_k \geq e_1 + \dots + e_k$. Más aún, la igualdad se sigue tomando en (6.11) los índices particulares $i_1 = 1, i_2 = 2, \dots, i_k = k$, y aplicando el Lema 6.4.1. ■

Como consecuencia del Lema 6.4.2 se tiene:

Corolario 6.4.4 Sean $L(\lambda)$ y $M(\lambda)$ dos polinomios matriciales $m \times n$ tales que $\text{rg}(M) = 1$ y $\text{rg}(L + M) = \text{rg}(L) + \text{rg}(M)$. Sea $h(M)$ el único polinomio invariante de $M(\lambda)$. Entonces

$$\mathcal{S}_{L+M}(\lambda_0) \geq \mathcal{S}_{L \oplus h(M)}(\lambda_0) = \mathcal{S}_{L \oplus M}(\lambda_0), \quad \text{para todo número complejo } \lambda_0.$$

Demostración. Usaremos la notación del Lema 6.4.2 para las multiplicidades parciales de λ_0 . La multiplicidades parciales de λ_0 relativas a $L \oplus h(M)$ son $e_1 \leq \dots \leq e_i < e \leq e_{i+1} \leq \dots \leq e_r$ por el Teorema 5 de [26, p. 142, Ch. VI]. Éstas son también las multiplicidades parciales de λ_0 relativas a $L \oplus M$, por el mismo argumento. El Lema 6.4.2 implica la desigualdad $\mathcal{S}_{L+M}(\lambda_0) \geq \mathcal{S}_{L \oplus h(M)}(\lambda_0)$. ■

A continuación, enunciamos y demostramos el resultado más importante de este apartado.

Teorema 6.4.5 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales $m \times n$ tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$. Entonces, para cada número complejo λ_0 , incluido el infinito, $\mathcal{S}_{H+M}(\lambda_0) \geq \mathcal{S}_{H \oplus M}(\lambda_0)$. Esto significa, en particular, que si λ_0 es un autovalor de $H(\lambda)$ o si λ_0 es un autovalor de $M(\lambda)$ entonces λ_0 es un autovalor de $(H + M)(\lambda)$.

Demostración. Consideraremos solamente números finitos λ_0 . El resultado para el autovalor infinito se sigue de considerar el autovalor cero de los haces duales de $H(\lambda)$ y $M(\lambda)$. De acuerdo con (2.10), $M(\lambda)$ se puede expresar como

$$M(\lambda) = h_1(M) b_1(\lambda) c_1^T(\lambda) + \dots + h_\rho(M) b_\rho(\lambda) c_\rho^T(\lambda),$$

donde $\rho \equiv \text{rg}(M)$, y $h_1(M), \dots, h_\rho(M)$ son los polinomios invariantes de $M(\lambda)$. La propiedad $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$ implica que

$$\begin{aligned} \text{rg}(H + h_1(M) b_1 c_1^T + \dots + h_k(M) b_k c_k^T) & \quad (6.14) \\ = \text{rg}(H + h_1(M) b_1 c_1^T + \dots + h_{k-1}(M) b_{k-1} c_{k-1}^T) + \text{rg}(h_k(M) b_k c_k^T) & \quad , \end{aligned}$$

para $k = 1, \dots, \rho$. Hemos omitido la variable λ por simplicidad. Por lo tanto, el Corolario 6.4.4 se puede aplicar ρ veces para demostrar

$$\begin{aligned} \mathcal{S}_{H+M}(\lambda_0) & \geq \mathcal{S}_{(H+h_1(M)b_1c_1^T+\dots+h_{\rho-1}(M)b_{\rho-1}c_{\rho-1}^T) \oplus h_\rho(M)}(\lambda_0) \geq \dots \\ & \geq \mathcal{S}_{H \oplus h_1(M) \oplus \dots \oplus h_\rho(M)}(\lambda_0), \end{aligned}$$

donde hemos usado que $(A + B) \oplus C = (A \oplus C) + (B \oplus 0)$. Finalmente, el Teorema 5 de [26, p. 142, Ch. VI] implica que $\mathcal{S}_{H \oplus h_1(M) \oplus \dots \oplus h_\rho(M)}(\lambda_0) = \mathcal{S}_{H \oplus M}(\lambda_0)$. ■

6.5. Los índices minimales del haz perturbado

El propósito de este apartado es determinar los índices minimales del haz perturbado $(H + M)(\lambda)$ a partir de ciertos datos de $H(\lambda)$ y $M(\lambda)$. En aras de la brevedad, desarrollaremos los resultados solamente para los índices minimales derechos. Los resultados equivalentes para los índices minimales izquierdos pueden obtenerse de manera inmediata a partir de éstos considerando los índices minimales derechos del haz traspuesto $(H + M)^T(\lambda)$.

El principal resultado de esta sección es el Teorema 6.5.10, donde se determinan todos los índices minimales derechos de $H + M$ para la mayor parte de las perturbaciones M con un rango y un orden singular derecho determinados. La genericidad de las hipótesis del Teorema 6.5.10 se discutirá en la Subsección 6.5.5. El Teorema 6.5.12 presenta cierta información genérica sobre los índices minimales derechos de $H + M$ cuando sólo es conocido el rango de la perturbación.

De acuerdo con el Lema 2.5.3, determinar los índices minimales derechos de $H + M$ es equivalente a encontrar los grados de una ROMB de $H + M$. Esta ROMB no será construída explícitamente, pero los grados de sus vectores serán determinados con precisión. El Lema 6.5.1 es el resultado clave para abordar esta tarea, porque permite delimitar la búsqueda de esta base.

Lema 6.5.1 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$. Entonces $x(\lambda)$ es un vector nulo derecho de $H + M$ si y sólo si $x(\lambda)$ es, simultáneamente, un vector nulo derecho de $H(\lambda)$ y $M(\lambda)$. En otras palabras: $\mathcal{N}(H + M) = \mathcal{N}(H) \cap \mathcal{N}(M)$.

Demostración. Sea $\text{col}(H)$ el espacio de columnas de $H(\lambda)$ en $\mathbb{C}^m(\lambda)$. Entonces, $\dim(\text{col}(H + M)) \leq \dim(\text{col}[H \ M]) = \dim(\text{col}(H)) + \dim(\text{col}(M)) - \dim(\text{col}(H) \cap \text{col}(M))$.

La hipótesis $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$ implica que $\text{col}(H) \cap \text{col}(M) = \{0\}$. Si $x(\lambda)$ es un vector nulo derecho de $H + M$ entonces $H(\lambda)x(\lambda) = -M(\lambda)x(\lambda)$. Nótese que el vector $z(\lambda) \equiv H(\lambda)x(\lambda) = -M(\lambda)x(\lambda)$ es un vector de $\text{col}(H) \cap \text{col}(M)$, luego ha de ser $z(\lambda) = 0$ y, por tanto, $H(\lambda)x(\lambda) = M(\lambda)x(\lambda) = 0$. El opuesto es trivial. ■

Ya hemos observado en la Sección 6.3 que si $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = n$ entonces el haz $(H + M)(\lambda)$ no tiene índices minimales derechos. Por lo tanto, en el resto de esta sección, se supondrá que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n$. Esto implica que $\text{rg}(M) < p$, donde p es el número de índices minimales derechos de H .

6.5.1. Polinomios conexión y matrices mosaico Toeplitz asociadas

A partir del Lema 6.5.1, es posible obtener una caracterización más específica del espacio nulo derecho de $H + M$.

Lema 6.5.2 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales $m \times n$ tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n$, $\tilde{\varepsilon}$ el orden singular derecho de $M(\lambda)$, y $\{x_1(\lambda), \dots, x_p(\lambda)\}$ una ROMB de $H(\lambda)$. Consideremos una descomposición derecha de $M(\lambda)$ dada por (6.3), donde los primeros $\tilde{\varepsilon}$ vectores de $\{w_1(\lambda), \dots, w_\rho(\lambda)\}$ se suponen de grado 1 sin pérdida de generalidad. Entonces, todo vector nulo derecho de $(H + M)(\lambda)$ es una combinación lineal de los vectores $\{x_1(\lambda), \dots, x_p(\lambda)\}$ con coeficientes polinómicos. Más aún,

$$x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_p(\lambda)x_p(\lambda) \quad (6.15)$$

es un vector nulo derecho de $(H + M)(\lambda)$ si y sólo si los polinomios $\alpha_1(\lambda), \dots, \alpha_p(\lambda)$ satisfacen el sistema de ecuaciones lineales en $\mathbb{C}^p(\lambda)$

$$\begin{aligned} a_{11}(\lambda)\alpha_1(\lambda) + \dots + a_{1p}(\lambda)\alpha_p(\lambda) &= 0 \\ &\vdots \\ a_{\rho 1}(\lambda)\alpha_1(\lambda) + \dots + a_{\rho p}(\lambda)\alpha_p(\lambda) &= 0 \end{aligned} \quad (6.16)$$

donde

$$a_{ij}(\lambda) = w_i^T(\lambda)x_j(\lambda), \quad i = 1, \dots, \rho, \quad j = 1, \dots, p. \quad (6.17)$$

Demostración. Un vector polinómico, $x(\lambda)$, es un vector nulo derecho de $H + M$ si y sólo si es un vector nulo derecho de H y M , por el Lema 6.5.1. $H(\lambda)x(\lambda) = 0$ es equivalente al hecho de que $x(\lambda)$ sea una combinación lineal de $\{x_1(\lambda), \dots, x_p(\lambda)\}$ con coeficientes polinómicos, por el Lema 2.5.5. $M(\lambda)x(\lambda) = 0$ es equivalente, teniendo en cuenta (6.3), a $w_1^T(\lambda)x(\lambda) = \dots = w_\rho^T(\lambda)x(\lambda) = 0$, y éste es precisamente el sistema de ecuaciones (6.16). ■

El sistema de ecuaciones (6.16) es de capital importancia en esta sección, porque el conjunto de sus soluciones nos permite obtener el espacio nulo derecho de $H + M$ a través de (6.15), y nuestra intención es conocer los grados de una ROMB de $\mathcal{N}(H + M)$. Así, los coeficientes $a_{ij}(\lambda)$ del sistema (6.16) tienen un papel esencial. Se trata de polinomios en la variable λ y relacionan los haces H y M . Serán utilizados con frecuencia, por lo que introducimos la siguiente definición.

Definición 6.5.3 *Un conjunto de polinomios $\{a_{ij}(\lambda) : i = 1, \dots, \rho, j = 1, \dots, p\}$ como los que aparecen en (6.17) se denominará un **conjunto completo de polinomios conexión derechos** de $H(\lambda)$ y $M(\lambda)$.*

Dado que, ni una ROMB de H ni una descomposición derecha (6.3) de M son únicas, un conjunto completo de polinomios conexión derechos de H y M no es necesariamente único.

Observación 6.5.4 *Una descomposición izquierda (6.4) de $M(\lambda)$ y una LOMB de $H(\lambda)$, dada por $\{y_1(\lambda), \dots, y_q(\lambda)\}$, definen un **conjunto completo de polinomios conexión izquierdos** de $H(\lambda)$ y $M(\lambda)$. Se trata de los polinomios*

$$b_{ij}(\lambda) = \hat{v}_i^T(\lambda)y_j(\lambda), \quad i = 1, \dots, \rho, \quad j = 1, \dots, q.$$

Estos polinomios son necesarios para obtener los índices minimales izquierdos de $(H + M)(\lambda)$.

Denotemos por $\varepsilon_1 \leq \dots \leq \varepsilon_p$ los índices minimales derechos del polinomio no perturbado $H(\lambda)$, y por $\tilde{\varepsilon}$ el orden singular derecho de $M(\lambda)$, como en la Sección 6.2. Entonces los grados de los polinomios conexión derechos de H y M están acotados de la siguiente manera

$$\deg(a_{ij}(\lambda)) \leq \begin{cases} \varepsilon_j + 1 & , \quad i = 1, \dots, \tilde{\varepsilon} \\ \varepsilon_j & , \quad i = \tilde{\varepsilon} + 1, \dots, \rho \end{cases} . \quad (6.18)$$

Para la mayoría de las perturbaciones $M(\lambda)$ estas desigualdades son, de hecho, igualdades, aunque esto no se supondrá en los siguientes desarrollos. No obstante, el comportamiento genérico de los índices minimales del haz perturbado $H + M$ se cumple bajo ciertas condiciones que limitan el número de polinomios conexión derechos con grado estrictamente menor que el término de la derecha en (6.18). Estas condiciones genéricas involucran ciertas de las matrices mosaico Toeplitz que aparecen en la siguiente definición.

Definición 6.5.5 *Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ los índices minimales derechos del haz $H(\lambda)$, $\tilde{\varepsilon}$ el orden singular derecho del haz $M(\lambda)$, y $\{a_{ij}(\lambda) : i = 1, \dots, \rho, j = 1, \dots, p\}$ un conjunto*

completo de polinomios conexión derechos de H y M . Expresemos estos polinomios como sigue¹

$$a_{ij}(\lambda) = a_{ij}^0 + \lambda a_{ij}^1 + \dots + \lambda^{\epsilon_{ij}} a_{ij}^{\epsilon_{ij}}, \quad \text{donde} \quad \epsilon_{ij} = \begin{cases} \epsilon_j + 1 & , \quad i = 1, \dots, \tilde{\epsilon} \\ \epsilon_j & , \quad i = \tilde{\epsilon} + 1, \dots, \rho \end{cases} .$$

Sean k y d dos números enteros no negativos tales que $1 \leq k \leq p$ y $d \geq \epsilon_k - 1$. La k -ésima matriz mosaico Toeplitz de grado d asociada a los polinomios conexión $a_{ij}(\lambda)$ se denota por $A_k(d)$, y es la matriz particionada en ρ filas y k columnas de bloques cuyo bloque (s, t) , para $s = 1, \dots, \rho$ y $t = 1, \dots, k$, es la matriz de Toeplitz

$$(A_k(d))_{st} = \underbrace{\begin{bmatrix} a_{st}^0 & & & & \\ \vdots & \ddots & & & \\ a_{st}^{\epsilon_{st}} & & \ddots & & \\ & \ddots & & \ddots & \\ & & & & a_{st}^0 \\ & & & & \vdots \\ & & & & a_{st}^{\epsilon_{st}} \end{bmatrix}}_{d - \epsilon_t + 1},$$

con $d - \epsilon_t + 1$ columnas, y, por lo tanto, un número de filas igual a

$$\epsilon_{st} + 1 + d - \epsilon_t = \begin{cases} d + 2 & , \quad s = 1, \dots, \tilde{\epsilon} \\ d + 1 & , \quad s = \tilde{\epsilon} + 1, \dots, \rho \end{cases} .$$

Observación 6.5.6 Nótese que, en el caso $d = \epsilon_k - 1$, la k -ésima columna de bloques de $A_k(d)$ está formada por matrices que tienen un “número de columnas igual a cero”, es decir, por matrices vacías. Esto también ocurre para aquellas columnas cuyo índice j satisface $d = \epsilon_j - 1$. Entendemos que en este caso $A_k(d)$ tiene menos de k columnas de bloques. Este convenio simplificará la notación y los enunciados de nuestros resultados.

La importancia de la familia de matrices mosaico Toeplitz $A_k(d)$ se pondrá de manifiesto en el Lema 6.5.7. Este resultado extiende y complementa el Lema 6.5.2 caracterizando los vectores nulos derechos de $H + M$ de un grado determinado a través de sistemas de ecuaciones lineales con coeficientes constantes, es decir, sistemas de ecuaciones en \mathbb{C}^n y no en $\mathbb{C}^n(\lambda)$ como (6.16). Los grados son las cantidades fundamentales en esta sección, porque nuestro objetivo es obtener los grados de los vectores de una ROMB de $H + M$, esto es, los índices minimales derechos de $H + M$. De acuerdo con el Lema 6.5.2 todos los vectores nulos derechos, $x(\lambda)$, de $H + M$ son de la forma (6.15), y tienen $\deg(x) = \max_{1 \leq i \leq p} \{\epsilon_i + \deg(\alpha_i) : \alpha_i(\lambda) \neq 0\}$ [24, Main Theorem en p. 495]. Esto implica que si j es el mayor índice tal que $\alpha_j(\lambda) \neq 0$, entonces $\deg(x) \geq \epsilon_j$. Para encontrar los grados más pequeños, hay que considerar necesariamente combinaciones lineales $x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_k(\lambda)x_k(\lambda)$, con $k < j$. Véase también el Lema 2.5.5.

Lema 6.5.7 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos $m \times n$ tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n$, $\{x_1(\lambda), \dots, x_p(\lambda)\}$ una ROMB de H , y $\epsilon_1 \leq \dots \leq \epsilon_p$ los índices minimales derechos de H , i.e., $\deg(x_i) = \epsilon_i$. Entonces:

¹El lector ha de notar que la notación de superíndices a_{ij}^k no significa la k -ésima potencia de a_{ij} .

1. Todo vector nulo derecho, $x(\lambda)$, de $(H + M)(\lambda)$ de grado d se puede expresar en la forma $x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_k(\lambda)x_k(\lambda)$ para el mayor número k tal que $1 \leq k \leq p$ y $d \geq \varepsilon_k$, y un conjunto único de polinomios $\{\alpha_1(\lambda), \dots, \alpha_k(\lambda)\}$.

2. Más aún,

$$x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_k(\lambda)x_k(\lambda), \quad (6.19)$$

es un vector nulo derecho de $(H + M)(\lambda)$ de grado $d \geq \varepsilon_k$ si y sólo si los polinomios $\alpha_1(\lambda), \dots, \alpha_k(\lambda)$ satisfacen las siguientes dos condiciones:

(i) Los polinomios $\alpha_i(\lambda)$ se pueden expresar como

$$\alpha_i(\lambda) = \alpha_{i0} + \lambda \alpha_{i1} + \dots + \lambda^{d-\varepsilon_i} \alpha_{i,d-\varepsilon_i}, \quad (6.20)$$

para todo $i = 1, \dots, k$, con $\alpha_{j,d-\varepsilon_j} \neq 0$ para, al menos, un índice j .

(ii) Si $A_k(d)$ es la k -ésima matriz mosaico Toeplitz de grado d asociada a un conjunto completo de polinomios conexión derechos de H y M definido por $\{x_1(\lambda), \dots, x_p(\lambda)\}$, los coeficientes α_{il} satisfacen el sistema de ecuaciones lineales constantes

$$A_k(d) \begin{bmatrix} \alpha_{10} \\ \vdots \\ \alpha_{1,d-\varepsilon_1} \\ \vdots \\ \vdots \\ \alpha_{k0} \\ \vdots \\ \alpha_{k,d-\varepsilon_k} \end{bmatrix} = 0. \quad (6.21)$$

Nótese que, si todas las soluciones del sistema (6.21) son tales que $\alpha_{1,d-\varepsilon_1} = \dots = \alpha_{k,d-\varepsilon_k} = 0$, entonces todos los vectores nulos derechos de $H + M$ de la forma (6.19) tienen grado menor que d .

Demostración. La demostración del primer apartado es una consecuencia directa del Lema 6.5.2 y el hecho de que en (6.15) $\deg(x) = \max_{1 \leq i \leq p} \{\varepsilon_i + \deg(\alpha_i) : \alpha_i(\lambda) \neq 0\}$ de acuerdo con [24, Main Theorem en p. 495]. Una vez que el índice k se ha elegido, la unicidad de $\{\alpha_1(\lambda), \dots, \alpha_k(\lambda)\}$ se sigue de la independencia lineal de $\{x_1(\lambda), \dots, x_k(\lambda)\}$.

El segundo apartado se sigue de (6.15) y (6.16) definiendo $\alpha_{k+1}(\lambda) = \dots = \alpha_p(\lambda) = 0$. Nótese que (6.20) establece simplemente que no hay índices i , con $1 \leq i \leq k$, de modo que $\deg(\alpha_i) > d - \varepsilon_i$, porque esto implicaría que $\deg(x) > d$. La condición $\alpha_{j,d-\varepsilon_j} \neq 0$ para al menos un índice j garantiza que $\deg(x) = d$. Por otro lado, el sistema lineal (6.21) es el sistema obtenido de (6.16) desarrollando los productos y las sumas de los polinomios e igualando a cero los coeficientes. Con estas observaciones en mente, la demostración es trivial. ■

6.5.2. Propiedades de las matrices mosaico Toeplitz

El Lema 6.5.8 recoge las propiedades de las matrices mosaico Toeplitz que utilizaremos para deducir los índices minimales genéricos del haz $H + M$.

Lema 6.5.8 Sea $\mathcal{T} = \{A_k(d) : 1 \leq k \leq p, d \geq \varepsilon_k - 1\}$ el conjunto de matrices mosaico Toeplitz de la Definición 6.5.5. Entonces:

1. El número de filas de $A_k(d)$ es igual a $\rho(d+1) + \tilde{\varepsilon}$.
2. El número de columnas de $A_k(d)$ es igual a $k(d+1) - \sum_{j=1}^k \varepsilon_j$.
3. $A_k(d)$ tiene más columnas que filas si y sólo si

$$k > \rho \quad \text{y} \quad d > \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} - 1.$$

4. Si las columnas de $A_k(d)$ son linealmente independientes, i.e., $A_k(d)$ tiene rango completo por columnas, entonces toda matriz $A_{k'}(d')$ de \mathcal{T} con $k' \leq k$ y $d' \leq d$ tiene rango completo por columnas.
5. Si las filas de $A_k(d)$ son linealmente independientes, i.e., $A_k(d)$ tiene rango completo por filas, entonces toda matriz $A_{k'}(d')$ de \mathcal{T} con $k' \geq k$ y $d' \geq d$ tiene rango completo por filas.
6. Si las filas de $A_k(d)$ son linealmente independientes entonces

$$\text{rg} \begin{bmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \dots & a_{1j}(\lambda) \\ \vdots & \vdots & & \vdots \\ a_{\rho 1}(\lambda) & a_{\rho 2}(\lambda) & \dots & a_{\rho j}(\lambda) \end{bmatrix} = \rho \quad \text{para} \quad j \geq k. \quad (6.22)$$

Demostración. Los tres primeros apartados son consecuencia inmediata del número de filas y columnas de los bloques que aparecen en la Definición 6.5.5.

4. Nótese que $A_{k-1}(d)$ se obtiene de $A_k(d)$ eliminando la última columna de bloques. En consecuencia, las columnas de $A_{k-1}(d)$ son un subconjunto de las columnas de $A_k(d)$. Entonces, $A_{k-1}(d)$ tiene rango completo por columnas si $A_k(d)$ tiene rango completo por columnas, y, por inducción, $A_{k'}(d)$ tiene rango completo por columnas siempre que $k' \leq k$.

Si $d-1 \geq \varepsilon_k - 1$ entonces $A_k(d-1)$ es un elemento de \mathcal{T} , y es obtenido a partir de $A_k(d)$ eliminando la última columna de cada bloque de $A_k(d)$ para dar una cierta matriz $A'_k(d)$ y, tras esto, eliminar la última fila de cada bloque de $A'_k(d)$ para llegar a $A_k(d-1)$. Pero, nótese que $A'_k(d)$ es de rango completo por columnas, y que las últimas filas de los bloques de $A'_k(d)$ son filas de ceros, luego $A_k(d-1)$ tiene también rango completo por columnas.

Si $d-1 < \varepsilon_k - 1$ entonces $A_k(d-1)$ no está en \mathcal{T} . Sea $k' < k$ el mayor índice tal que $d-1 \geq \varepsilon_{k'} - 1$. Entonces $A_{k'}(d-1)$ es un elemento de \mathcal{T} , y $A_{k'}(d)$ tiene rango completo por filas. El argumento del último párrafo se puede aplicar para demostrar que $A_{k'}(d-1)$ tiene rango completo por columnas.

Finalmente, los resultados anteriores se pueden combinar inductivamente para demostrar el apartado 4.

5. Sea $k' \geq k$. Entonces la submatriz de $A_{k'}(d)$ que contiene las primeras $k(d+1) - \sum_{j=1}^k \varepsilon_j$ columnas es precisamente $A_k(d)$. Por tanto, si $A_k(d)$ tiene rango completo por filas entonces $A_{k'}(d)$ tiene también rango completo por filas. Para completar la demostración, probaremos que $A_k(d+t)$ tiene rango completo por filas para cualquier número entero $t > 0$ siempre que $A_k(d)$ tenga rango completo por filas. Es suficiente demostrar este

resultado para $t = 1$ y después aplicar un argumento inductivo. Nótese que la submatriz de $A_k(d)$ que contiene la última fila de cada fila de bloques de $A_k(d)$ tiene filas linealmente independientes. Esto significa que la matriz

$$B \equiv \begin{bmatrix} a_{11}^{\varepsilon_{11}} & \cdots & a_{1k}^{\varepsilon_{1k}} \\ \vdots & & \vdots \\ a_{\rho 1}^{\varepsilon_{\rho 1}} & \cdots & a_{\rho k}^{\varepsilon_{\rho k}} \end{bmatrix}$$

tiene rango ρ . Obsérvese que la matriz B es también la submatriz $\rho \times k$ de $A_k(d+1)$ que contiene las últimas filas y columnas de los bloques de $A_k(d+1)$. Si las últimas filas y columnas de los bloques de $A_k(d+1)$ se llevan mediante permutaciones a las últimas posiciones (hacia abajo y hacia la derecha), el rango no varía, y la matriz que obtenemos tiene la estructura,

$$\begin{bmatrix} A_k(d) & * \\ 0 & B \end{bmatrix}.$$

El rango de esta matriz es claramente igual a $\text{rg}(A_k(d)) + \rho$, que a su vez es igual a $\rho(d+2) + \tilde{\varepsilon}$, es decir, el número de filas de $A_k(d+1)$. En consecuencia, $A_k(d+1)$ tiene rango completo por filas.

6. Es suficiente demostrar esta propiedad para $j = k$. Si las filas de $A_k(d)$ son linealmente independientes entonces la submatriz de $A_k(d)$ que contiene la primera fila de cada fila de bloques de $A_k(d)$ tiene filas linealmente independientes. Esto significa que

$$\text{rg} \begin{bmatrix} a_{11}^0 & \cdots & a_{1k}^0 \\ \vdots & & \vdots \\ a_{\rho 1}^0 & \cdots & a_{\rho k}^0 \end{bmatrix} = \rho.$$

El resultado se sigue aplicando el primer apartado del Lema 2.5.6 a las filas de la matriz (6.22) para $j = k$. ■

De acuerdo con [26, Ch. XII, p. 38] (véase también la demostración del Lema 2.5.3), el menor índice minimal derecho de $H + M$ es el menor grado entre todos aquellos grados de vectores nulos derechos de $H + M$ (distintos del vector cero). Teniendo en cuenta el segundo apartado del Lema 6.5.7, dicho índice minimal más pequeño corresponde al menor d para el que un sistema lineal de la familia (6.21) ($1 \leq k \leq p$) tiene soluciones no nulas con $\alpha_{j,d-\varepsilon_j} \neq 0$ para al menos un índice j . Nuestra intuición en este punto es que las soluciones de este tipo no existen, genéricamente, si $A_k(d)$ tiene un número de filas mayor o igual que el número de columnas, y que existen, genéricamente, en el caso contrario. Esta intuición está basada en la idea de que si los coeficientes de los polinomios conexión son aleatorios para perturbaciones aleatorias M , entonces las columnas de $A_k(d)$ deberían ser linealmente independientes si $A_k(d)$ tiene más filas que columnas o el mismo número de filas y columnas. Basándose en esta intuición, uno puede pensar a priori que el valor más habitual del menor índice minimal de $H + M$ para perturbaciones aleatorias M es el menor valor d tal que alguna de las matrices $A_k(d)$ tiene más columnas que filas. Por supuesto, estos argumentos informales tienen que ser respaldados por hipótesis rigurosas, pero ellos, junto con el tercer apartado del Lema 6.5.8, hacen que sea natural considerar la siguiente sucesión de números enteros

$$d_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p,$$

donde $\lfloor x \rfloor$ denota la parte entera de x , es decir, el mayor entero que es menor o igual que x . Nótese que $A_k(d_k)$ existe sólo si $d_k \geq \varepsilon_k - 1$, y en este caso el tercer apartado del Lema 6.5.8 garantiza que $A_k(d_k)$ tiene más columnas que filas. En cambio, no es difícil encontrar ejemplos para los cuales $d_k < \varepsilon_k - 1$ para algún k . El candidato natural para el menor índice minimal derecho de $H + M$ es $\min_{\rho+1 \leq k \leq p} d_k$. Para demostrar que este es el caso bajo ciertas hipótesis de genericidad, y también para encontrar el resto de índices minimales derechos, es necesario estudiar las propiedades de la sucesión $\{d_k\}$.

Lema 6.5.9 *Sean $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$ un total de p números enteros, y $\rho, \tilde{\varepsilon}$ otros dos números enteros tales que $0 < \rho < p$ y $0 \leq \tilde{\varepsilon} \leq \rho$. Consideremos la sucesión de números enteros:*

$$d_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p. \quad (6.23)$$

Entonces:

1. $d_{\rho+1} \geq \varepsilon_{\rho+1} \geq \dots \geq \varepsilon_1$.
2. Si $d_k < d_{k-1}$ entonces $d_k \geq \varepsilon_k \geq \dots \geq \varepsilon_1$.
3. Si $d_k < d_{k+1}$ entonces $d_k < d_{k+1} \leq d_{k+2} \leq \dots \leq d_p$.
4. Si $d_k < d_{k+1}$ entonces $d_i < \varepsilon_i$ para todo $i \geq (k + 1)$.
5. Sea

$$d_{\min} \equiv \min_{\rho+1 \leq k \leq p} d_k.$$

Entonces todos los índices j tales que $d_j = d_{\min}$ son consecutivos, i.e., si $d_j = d_{\min}$ se cumple para más de un índice, entonces existen dos índices $j_1 < j_2$ tales que

$$d_j = d_{\min} \quad \text{si } j_1 \leq j \leq j_2, \quad \text{y } d_j > d_{\min} \quad \text{si } j < j_1 \text{ o } j_2 < j.$$

Además $d_{j_1} \geq \varepsilon_{j_1}$.

6. Sea s el mayor índice tal que $d_s = d_{\min}$ y $d_s \geq \varepsilon_s$. Entonces

$$\varepsilon_k > d_k \geq d_s \quad \text{para todo } k > s.$$

7. Sea s el índice definido en el apartado anterior, $A_s(d_{\min})$ y $A_s(d_{\min} - 1)$ las matrices mosaico Toeplitz introducidas en la Definición 6.5.5. Entonces

- (i) $A_s(d_{\min})$ tiene más columnas que filas, y contiene exactamente s columnas de bloques.
- (ii) $A_s(d_{\min} - 1)$ tiene un número de filas mayor o igual que el número de columnas, o es la matriz vacía.
- (iii) Para todo $k > s$, $A_k(d_{\min})$ no está definida o $A_k(d_{\min}) = A_s(d_{\min})$.

Antes de demostrar este Lema, debemos resaltar que el índice s que aparece en el apartado 6 tendrá un papel esencial a la hora de determinar los índices minimales derechos genéricos de $H + M$.

Demostración del Lema 6.5.9. El primer apartado es trivial.

2. Consideremos las divisiones enteras

$$\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon} = (k - \rho)d_k + r_k, \quad \text{donde } 0 \leq r_k < k - \rho, \quad (6.24)$$

$$\sum_{i=1}^{k-1} \varepsilon_i + \tilde{\varepsilon} = (k - 1 - \rho)d_{k-1} + r_{k-1}, \quad \text{donde } 0 \leq r_{k-1} < k - 1 - \rho. \quad (6.25)$$

Restando (6.25) de (6.24) llegamos a

$$\varepsilon_k = (k - \rho - 1)(d_k - d_{k-1}) + d_k + r_k - r_{k-1} \leq (k - \rho - 1)(d_k - d_{k-1}) + d_k + k - \rho - 1.$$

De aquí, $\varepsilon_k \leq (k - \rho - 1)(d_k - d_{k-1} + 1) + d_k \leq d_k$. El último paso es consecuencia de las desigualdades $(d_k - d_{k-1} + 1) \leq 0$ y $(k - 1) > \rho$.

3. Consideremos la división entera

$$\sum_{i=1}^{k+1} \varepsilon_i + \tilde{\varepsilon} = (k + 1 - \rho)d_{k+1} + r_{k+1}, \quad \text{donde } 0 \leq r_{k+1} < k + 1 - \rho. \quad (6.26)$$

Si restamos (6.24) de (6.26) llegamos a

$$\varepsilon_{k+1} = (k - \rho)(d_{k+1} - d_k) + d_{k+1} + r_{k+1} - r_k \geq (k - \rho) + d_{k+1} + r_{k+1} - r_k > d_{k+1},$$

donde hemos usado que $r_{k+1} - r_k > -(k - \rho)$. Por lo tanto, hemos demostrado que

$$d_k < d_{k+1} \quad \text{implica} \quad \varepsilon_{k+1} > d_{k+1}. \quad (6.27)$$

Consideremos ahora un índice l tal que $l \geq (k + 2)$, y la división entera

$$\sum_{i=1}^l \varepsilon_i + \tilde{\varepsilon} = (l - \rho)d_l + r_l, \quad \text{donde } 0 \leq r_l < l - \rho. \quad (6.28)$$

Restando (6.26) de (6.28) se llega a

$$\varepsilon_l + \varepsilon_{l-1} + \dots + \varepsilon_{k+2} = (d_l - d_{k+1})(l - \rho) + d_{k+1}(l - (k + 1)) + r_l - r_{k+1},$$

y entonces,

$$(\varepsilon_l - d_{k+1}) + (\varepsilon_{l-1} - d_{k+1}) + \dots + (\varepsilon_{k+2} - d_{k+1}) = (d_l - d_{k+1})(l - \rho) + r_l - r_{k+1}.$$

La desigualdad (6.27) implica que $(d_l - d_{k+1})(l - \rho) + r_l - r_{k+1} > 0$, y por tanto, $(d_l - d_{k+1} + 1)(l - \rho) > 0$. Luego hemos demostrado que

$$d_k < d_{k+1} \quad \text{implica} \quad d_l \geq d_{k+1}, \quad \text{para todo } l \geq (k + 2). \quad (6.29)$$

Este resultado nos permite demostrar el resultado más general que aparece en el apartado 3. Procederemos por reducción al absurdo. Supongamos que $d_{k+1} \leq d_{k+2} \leq \dots \leq d_p$ es falso. Esto significa que existe un índice $l \geq (k + 2)$ tal que $d_{k+1} \leq d_{k+2} \leq \dots \leq d_{l-1} > d_l$. Sea j el menor número entero tal que $(k + 1) \leq j \leq (l - 1)$ y $d_j = d_{j+1} = \dots = d_{l-1}$. Nótese que este entero es al menos $k + 1$, porque $d_k < d_{l-1}$ por (6.29). Entonces $d_{j-1} < d_j$, y se

puede aplicar (6.29) con $k = j - 1$ para ver que $d_j \leq d_l$. Pero, por otro lado, $d_j = d_{l-1} > d_l$, lo cual es absurdo.

4. Demostraremos el resultado por inducción. En (6.27), ya hemos demostrado el caso de partida de la inducción: $d_{k+1} < \varepsilon_{k+1}$. Supongamos que $d_i < \varepsilon_i$, para algún $i \geq (k + 1)$. Por otra parte, $d_i \leq d_{i+1}$ debido al resultado del apartado 3. Si $d_i < d_{i+1}$, se puede aplicar (6.27) con $k = i$ para ver que $d_{i+1} < \varepsilon_{i+1}$. En otro caso, $d_i = d_{i+1}$ y $d_{i+1} < \varepsilon_i \leq \varepsilon_{i+1}$.

5. El hecho de que los índices son consecutivos es una consecuencia directa del apartado 3. El hecho de que $d_{j_1} \geq \varepsilon_{j_1}$ es una consecuencia de los apartados 1 y 2.

6. Si sólo hay un índice s tal que $d_s = d_{\min}$ el resultado es una simple consecuencia de los apartados 4 y 5. En otro caso, sean j_1 y j_2 los dos índices que aparecen en el apartado 5. Si $s = j_2 \leq p$, el resultado se sigue nuevamente del apartado 4. Si $s < j_2$, entonces, por definición, $d_s = d_{s+1} < \varepsilon_{s+1} \leq \dots \leq \varepsilon_{j_2}$. Por lo tanto, $d_k < \varepsilon_k$ para $s + 1 \leq k \leq j_2$. También, por definición, $d_{j_2} < d_{j_2+1}$ y el apartado 4 implica que $d_k < \varepsilon_k$ para $k \geq (j_2 + 1)$.

7. Las afirmaciones sobre el número de filas y columnas de $A_s(d_{\min})$ y $A_s(d_{\min} - 1)$ se siguen del apartado 3 del Lema 6.5.8. Recordemos que la j -ésima columna de bloques de $A_s(d_{\min})$ tiene $d_{\min} - \varepsilon_j + 1$ columnas, luego, $d_{\min} \geq \varepsilon_s$ garantiza que todos los bloques de $A_s(d_{\min})$ tienen, al menos, una columna. Nótese que $d_{\min} \geq \varepsilon_s$ también implica que $d_{\min} - 1 \geq \varepsilon_s - 1$, luego $A_s(d_{\min} - 1)$ está definida, pero alguno de (o todos) sus bloques pueden ser vacíos. Finalmente, para $j > s$, sabemos que $\varepsilon_j > d_{\min}$, es decir, $\varepsilon_j - 1 \geq d_{\min}$. Esto significa que $A_k(d_{\min})$, con $k > s$, no está definida salvo que $\varepsilon_j - 1 = d_{\min}$ para $s + 1 \leq j \leq k$, pero en este caso los bloques j -ésimos de $(A_k(d_{\min}))_{ij}$ son matrices vacías. ■

6.5.3. Los índices minimales derechos genéricos de $H + M$

Ahora, estamos en disposición de obtener los índices minimales derechos de $(H + M)(\lambda)$ para una perturbación genérica suponiendo que, además del rango, conocemos el orden singular derecho de la perturbación.

Teorema 6.5.10 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos de tamaño $m \times n$ tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n$ y $\rho \equiv \text{rg}(M)$. Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ los índices minimales derechos de H , $\tilde{\varepsilon}$ el orden singular derecho de M , $\{d_{\rho+1}, \dots, d_p\}$ la sucesión de números enteros definida en (6.23), d_{\min} el mínimo de esta sucesión, y s el mayor índice tal que $d_s = d_{\min}$ y $d_s \geq \varepsilon_s$. Finalmente, sean $A_s(d_{\min} - 1)$ y $A_s(d_{\min})$ las matrices mosaico Toeplitz s -ésimas de grados $d_{\min} - 1$ y d_{\min} , respectivamente, asociadas a un conjunto completo de polinomios conexión de H y M . Si

$$\begin{aligned} A_s(d_{\min} - 1) & \text{ tiene rango completo por columnas o es la matriz vacía, y} & (6.30) \\ A_s(d_{\min}) & \text{ tiene rango completo por filas,} & (6.31) \end{aligned}$$

entonces $(H + M)(\lambda)$ tiene los siguientes $p - \rho$ índices minimales derechos

$$\underbrace{d_{\min} = \dots = d_{\min}}_{s - \rho - \gamma_s} < \underbrace{(d_{\min} + 1) = \dots = (d_{\min} + 1)}_{\gamma_s} \leq \varepsilon_{s+1} \leq \dots \leq \varepsilon_p, \quad (6.32)$$

donde γ_s es el resto de la división entera de $\sum_{i=1}^s \varepsilon_i + \tilde{\varepsilon}$ entre $s - \rho$.

Demostración. En primer lugar, observemos que el orden que aparece en (6.32) es una consecuencia del apartado 6 del Lema 6.5.9. Nótese también que, como consecuencia

de (2.6) y de la igualdad $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$, el número de índices minimales derechos de $H + M$ es $p - \rho$. Para el resto de la demostración es conveniente tener en mente el Lema 2.5.3 aplicado a $H + M$, y el modo en que se construyen las ROMBs de $H + M$ (véase el primer párrafo de la demostración del Lema 2.5.3).

Comenzaremos demostrando que no hay índices minimales derechos de $H + M$ menores que $d_{\text{mín}}$. El primer apartado del Lema 6.5.7 y el apartado 6 del Lema 6.5.9 garantizan que todo vector nulo derecho de $H + M$ de grado $d < d_{\text{mín}}$ es una combinación lineal del tipo $x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_k(\lambda)x_k(\lambda)$ para algún $k \leq s$. En esta situación, la matriz $A_k(d)$ que aparece en (6.21) tiene rango completo por columnas en el caso de que $A_s(d_{\text{mín}} - 1)$ tenga rango completo por columnas, debido al apartado 4 del Lema 6.5.8 y, si esto ocurre, el sistema (6.21) sólo tiene la solución trivial $x(\lambda) = 0$. En el caso de que $A_s(d_{\text{mín}} - 1)$ sea la matriz vacía, $d_{\text{mín}} = \varepsilon_k$ siempre que $1 \leq k \leq s$, y no hay combinaciones lineales no nulas de $\{x_1(\lambda), \dots, x_p(\lambda)\}$ con grado menor que $\varepsilon_1 = d_{\text{mín}}$, porque de lo contrario el menor índice minimal derecho de H sería menor que ε_1 .

Nuestro siguiente paso es demostrar que $d_{\text{mín}}$ es un índice minimal derecho de $H + M$. El sistema (6.21) con matriz de coeficientes $A_s(d_{\text{mín}})$ tiene necesariamente soluciones no nulas porque $A_s(d_{\text{mín}})$ tiene más columnas que filas por el apartado 7 del Lema 6.5.9. Además, no hay soluciones no nulas con $\alpha_{1, d_{\text{mín}} - \varepsilon_1} = \dots = \alpha_{s, d_{\text{mín}} - \varepsilon_s} = 0$, porque de lo contrario las soluciones de (6.21) corresponderían a vectores nulos derechos de $H + M$ de grado menor que $d_{\text{mín}}$, y ya sabemos que éstos no existen. Esto demuestra que $d_{\text{mín}}$ es el menor índice minimal derecho de $H + M$.

Para ver que hay exactamente $s - \rho - \gamma_s$ índices minimales derechos de $H + M$ iguales a $d_{\text{mín}}$, necesitamos encontrar $s - \rho - \gamma_s$ vectores nulos derechos de $H + M$ linealmente independientes y de grado $d_{\text{mín}}$, y demostrar que no hay más. De nuevo, el primer apartado del Lema 6.5.7 junto con el apartado 6 del Lema 6.5.9 garantizan que todo vector nulo derecho de $H + M$ de grado $d_{\text{mín}}$ es una combinación lineal del tipo

$$x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_s(\lambda)x_s(\lambda). \quad (6.33)$$

Nótese que el conjunto de soluciones de (6.21) con $A_s(d_{\text{mín}})$ como matrices de coeficientes se puede describir en términos de un número de parámetros libres igual a la diferencia entre el número de columnas y el número de filas de $A_s(d_{\text{mín}})$, i.e.,

$$s(d_{\text{mín}} + 1) - \sum_{i=1}^s \varepsilon_i - \rho(d_{\text{mín}} + 1) - \tilde{\varepsilon} = s - \rho - \gamma_s,$$

donde se han usado los dos primeros apartados del Lema 6.5.8. Esto significa que el sistema de ecuaciones lineales (6.21) con $A_s(d_{\text{mín}})$ tiene $s - \rho - \gamma_s$ soluciones linealmente independientes y, por el Lema 6.5.7, que corresponden a $s - \rho - \gamma_s$ vectores nulos derechos de $H + M$ de la forma (6.33) con grado $d_{\text{mín}}$. Denotemos a estos vectores por,

$$\{z_1(\lambda), z_2(\lambda), \dots, z_{\beta_s}(\lambda)\}, \quad \text{con } \beta_s \equiv s - \rho - \gamma_s. \quad (6.34)$$

Es claro que cualquier otra solución de (6.21) corresponde a vectores nulos derechos de grado $d_{\text{mín}}$ que son combinaciones lineales de (6.34) con coeficientes constantes, pero en cambio aún tenemos que probar que los vectores $\{z_1(\lambda), z_2(\lambda), \dots, z_{\beta_s}(\lambda)\}$ se pueden elegir linealmente independientes en $\mathbb{C}^n(\lambda)$. Para ver esto, nótese que los β_s parámetros libres de (6.21) con $A_s(d_{\text{mín}})$ se pueden tomar entre las variables $\alpha_{1, d_{\text{mín}} - \varepsilon_1}, \dots, \alpha_{s, d_{\text{mín}} - \varepsilon_s}$, porque las columnas de $A_s(d_{\text{mín}})$ que no corresponden a estas variables son linealmente

independientes, como ya hemos visto en el párrafo en que probamos que d_{\min} es el menor índice minimal de $H + M$. Tomando las primeras l de estas β_s variables iguales a 1 y el resto iguales a 0, y repitiendo este proceso para $l = 1, \dots, \beta_s$, se puede obtener un conjunto, \mathcal{S} , de β_s soluciones de (6.21) linealmente independientes. Denotemos por

$$a_l = [\alpha_{1,d_{\min}-\varepsilon_1}^l, \dots, \alpha_{s,d_{\min}-\varepsilon_s}^l]^T, \quad l = 1, \dots, \beta_s, \quad (6.35)$$

un vector que contiene las entradas que acabamos de mostrar de la solución l -ésima de (6.21) en \mathcal{S} . Los vectores $\{a_1, \dots, a_{\beta_s}\}$ son, de manera obvia, linealmente independientes. Si recordamos (6.20) y (6.33), los coeficientes de los términos de mayor grado de los vectores (6.34) correspondientes a las β_s soluciones de (6.21) en \mathcal{S} son

$$z_{l,d_{\min}} = \alpha_{1,d_{\min}-\varepsilon_1}^l x_{1,\varepsilon_1} + \dots + \alpha_{s,d_{\min}-\varepsilon_s}^l x_{s,\varepsilon_s}, \quad \text{para } l = 1, \dots, \beta_s, \quad (6.36)$$

donde x_{i,ε_i} es el coeficiente de mayor grado de $x_i(\lambda)$. Los vectores $\{x_{1,\varepsilon_1}, \dots, x_{s,\varepsilon_s}\}$, son linealmente independientes en \mathbb{C}^n , porque $x_1(\lambda), \dots, x_s(\lambda)$ forman parte de una ROMB y se puede aplicar [24, Apartado 2-Main Theorem en p. 495]. Por lo tanto, $\{z_{1,d_{\min}}, \dots, z_{\beta_s,d_{\min}}\}$ es un conjunto linealmente independiente, porque

$$[z_{1,d_{\min}}, \dots, z_{\beta_s,d_{\min}}] = [x_{1,\varepsilon_1}, \dots, x_{s,\varepsilon_s}][a_1, \dots, a_{\beta_s}]$$

y las dos matrices del lado derecho tienen rango completo por columnas. Finalmente, el segundo apartado del Lema 2.5.6 implica que $\{z_1(\lambda), z_2(\lambda), \dots, z_{\beta_s}(\lambda)\}$ son linealmente independientes, y que hay precisamente $\beta_s \equiv s - \rho - \gamma_s$ índices minimales derechos de $H + M$ iguales a d_{\min} .

En este párrafo demostramos que hay γ_s índices minimales derechos de $H + M$ iguales a $d_{\min} + 1$. Hasta el momento, hemos encontrado un conjunto $\mathcal{C}_1 = \{z_1(\lambda), z_2(\lambda), \dots, z_{\beta_s}(\lambda)\}$ de $s - \rho - \gamma_s$ vectores nulos derechos de $H + M$ linealmente independientes de la forma (6.33) y con grado d_{\min} . Por otro lado, el hecho de que $A_s(d_{\min})$ tiene rango completo por filas (apartado 6 del Lema 6.5.8 con $j = s$) y el Lema 6.5.2 implican que un conjunto maximal linealmente independiente de vectores nulos derechos de $H + M$ de la forma (6.33) tiene $s - \rho$ vectores. Demostraremos que los restantes γ_s vectores se pueden elegir de grado $d_{\min} + 1$. Consideremos el sistema (6.21) con matriz de coeficientes $A_s(d_{\min} + 1)$. La matriz $A_s(d_{\min} + 1)$ tiene rango completo por filas porque $A_s(d_{\min})$ lo tiene, y podemos aplicar el apartado 5 del Lema 6.5.8. Esto significa que $\text{rg}(A_s(d_{\min} + 1)) = \text{rg}(A_s(d_{\min})) + \rho$. Recordemos que $A_s(d_{\min} + 1)$ se puede obtener de $A_s(d_{\min})$ añadiendo una fila y una columna en las últimas posiciones de cada bloque. Por lo tanto, entre las s columnas de $A_s(d_{\min} + 1)$ que están en las últimas posiciones de los bloques de columnas, $s - \rho$ son combinaciones lineales de las restantes columnas de $A_s(d_{\min} + 1)$. Luego, las correspondientes variables del sistema (6.21) con $A_s(d_{\min} + 1)$ se pueden tomar como algunos de los parámetros libres para resolver este sistema². Esto implica que se pueden tomar $s - \rho$ parámetros libres para resolver (6.21) con $A_s(d_{\min} + 1)$ entre las $\alpha_{1,(d_{\min}+1)-\varepsilon_1}, \dots, \alpha_{s,(d_{\min}+1)-\varepsilon_s}$ variables. Si procedemos con estos parámetros como en el párrafo previo (argumentos en torno a (6.34-6.36)), podemos encontrar un conjunto \mathcal{C}_2 de $(s - \rho)$ vectores nulos derechos de $H + M$ linealmente independientes, de grado

²El lector debe notar que la diferencia entre el número de columnas y de filas de $A_s(d_{\min} + 1)$ es $2(s - \rho) - \gamma_s$. Por tanto, el sistema (6.21) con matriz $A_s(d_{\min} + 1)$ tiene $2(s - \rho) - \gamma_s$ soluciones linealmente independientes, mientras que hay sólo $s - \rho$ vectores nulos derechos linealmente independientes de la forma (6.33). Esto significa que no siempre soluciones linealmente independientes de (6.21) corresponden a vectores nulos derechos linealmente independientes de la forma (6.19).

exactamente $d_{\min} + 1$ y de la forma (6.33). Por lo tanto, podemos unir el conjunto \mathcal{C}_1 de $s - \rho - \gamma_s$ vectores de grado d_{\min} con algunos de los γ_s vectores de \mathcal{C}_2 para obtener un conjunto maximal linealmente independiente de vectores nulos derechos de $H + M$ de la forma (6.33). Esto demuestra que existen γ_s índices minimales derechos de $H + M$ iguales a $d_{\min} + 1$.

Nuestra última tarea para demostrar el Teorema 6.5.10 es demostrar que los restantes índices minimales derechos de $H + M$ son $\varepsilon_{s+1} \leq \dots \leq \varepsilon_p$. Los vectores nulos derechos de $H + M$ que ya hemos encontrado, correspondientes a índices minimales iguales a d_{\min} y a $d_{\min} + 1$, constituyen un conjunto maximal linealmente independiente de vectores nulos derechos de $H + M$ de la forma (6.33). Este hecho implica que cualquier vector nulo derecho $x(\lambda)$ de $H + M$ correspondiente al siguiente índice minimal más pequeño ha de ser necesariamente de la forma (6.15) con, al menos, uno de los coeficientes $\alpha_{s+1}(\lambda), \dots, \alpha_p(\lambda)$ distinto de cero. De lo contrario, dependería linealmente de los vectores nulos derechos correspondientes a los índices minimales d_{\min} y $d_{\min} + 1$. Así, de acuerdo con [24, Main Theorem en p. 495], $\deg(x) = \max_{1 \leq i \leq p} \{\varepsilon_i + \deg(\alpha_i) : \alpha_i(\lambda) \neq 0\} \geq \varepsilon_{s+1} \geq d_{\min} + 1$, donde la última desigualdad es consecuencia del apartado 6 del Lema 6.5.9. Entonces, el menor candidato al siguiente índice minimal es ε_{s+1} . Para mostrar que, de hecho, ε_{s+1} es el siguiente índice minimal derecho, demostraremos que existe un vector nulo derecho de $H + M$ de la forma

$$x(\lambda) = \alpha_1(\lambda)x_1(\lambda) + \dots + \alpha_{s+1}(\lambda)x_{s+1}(\lambda), \quad (6.37)$$

con $\alpha_{s+1}(\lambda) \neq 0$ y $\deg(x) = \varepsilon_{s+1}$, i.e., con $\alpha_{s+1}(\lambda)$ una constante no nula. Esto es equivalente a demostrar que el sistema lineal (6.21) con matriz de coeficientes $A_{s+1}(\varepsilon_{s+1})$ tiene soluciones cuya última entrada es distinta de cero. Nótese que $A_{s+1}(\varepsilon_{s+1})$ tiene rango completo por filas porque $A_s(d_{\min})$ tiene rango completo por filas, $\varepsilon_{s+1} > d_{\min}$ por el apartado 6 del Lema 6.5.9, y se puede aplicar el apartado 5 del Lema 6.5.8. Además, las matrices de las últimas columnas de bloques de $A_{s+1}(\varepsilon_{s+1})$ tienen sólo una columna. Por lo tanto, si se elimina la última columna de $A_{s+1}(\varepsilon_{s+1})$, se obtiene $A_s(\varepsilon_{s+1})$. Pero

$$\text{número de filas de } A_{s+1}(\varepsilon_{s+1}) = \text{número de filas de } A_s(\varepsilon_{s+1}),$$

y $A_s(\varepsilon_{s+1})$ tiene también rango completo por filas, por el mismo argumento, luego

$$\text{rg}(A_{s+1}(\varepsilon_{s+1})) = \text{rg}(A_s(\varepsilon_{s+1})).$$

Esto implica que la última columna de $A_{s+1}(\varepsilon_{s+1})$ es una combinación lineal de las restantes columnas. Como consecuencia, la última variable del sistema lineal (6.21) con matriz de coeficientes $A_{s+1}(\varepsilon_{s+1})$ se puede considerar como parámetro libre, y por tanto puede ser distinta de cero. Esto demuestra que ε_{s+1} es el siguiente índice minimal.

Nótese que la hipótesis (6.31) y el apartado 6 del Lema 6.5.8 implican que un conjunto maximal linealmente independiente de vectores nulos derechos de $H + M$ de la forma (6.37) tiene $s + 1 - \rho$ vectores. Por tanto, en los párrafos anteriores hemos encontrado un conjunto maximal de vectores linealmente independientes de este tipo. Con esta observación en mente, la demostración de que ε_{s+2} es el siguiente índice minimal derecho se sigue paso por paso de la demostración presentada en el párrafo previo para ε_{s+1} con los cambios de índices correspondientes. Lo mismo se cumple para demostrar que $\varepsilon_{s+3}, \dots, \varepsilon_p$ son los restantes índices minimales derechos de $H + M$. ■

6.5.4. Ilustración del Teorema 6.5.10 con un ejemplo

Vamos a ilustrar con un ejemplo cómo aplicar el Teorema 6.5.10. Sea $H(\lambda)$ el haz matricial 5×5

$$H(\lambda) = \text{diag}(L_0, L_1, L_1, L_0^T, L_0^T, L_0^T) = \begin{bmatrix} 0 & \lambda & -1 & 0 & 0 \\ 0 & 0 & 0 & \lambda & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

con $\varepsilon_1 = 0$, $\varepsilon_2 = \varepsilon_3 = 1$. Una ROMB de H está dada por

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ \lambda \\ 0 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \lambda \end{bmatrix}.$$

Consideremos una perturbación arbitraria, M , de H con $\rho = 2$ y $\tilde{\varepsilon} = 1$. Esto significa que una descomposición derecha de M , véase (6.3), es de la forma

$$M(\lambda) = v_1 w_1^T + v_2 w_2^T,$$

donde

$$w_1 = \begin{bmatrix} b_1 + \lambda c_1 \\ b_2 + \lambda c_2 \\ b_3 + \lambda c_3 \\ b_4 + \lambda c_4 \\ b_5 + \lambda c_5 \end{bmatrix}, \quad w_2 = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{bmatrix},$$

y $b_i, c_i, d_i \in \mathbb{C}$ para $i = 1, \dots, 5$. Además, $\deg(v_1) = 0$ y $\deg(v_2) \leq 1$. Nótese que en este ejemplo $p - \rho = 3 - 2 = 1$ y, así, la sucesión $\{d_{\rho+1}, \dots, d_p\}$ tiene un único elemento d_3 . Esto significa que, en las condiciones del Teorema 6.5.10, *el haz $H + M$ tiene un único índice minimal derecho, que es*

$$d_{\min} = d_3 = \left\lfloor \frac{0 + 1 + 1 + 1}{3 - 2} \right\rfloor = 3.$$

Las matrices $A_s(d_{\min} - 1)$ y $A_s(d_{\min})$ son, en este caso, $A_3(2)$ y $A_3(3)$, respectivamente. Los polinomios conexión derechos asociados a los datos anteriores están dados por

$$\begin{aligned} a_{11}(\lambda) &= w_1^T(\lambda)x_1 = b_1 + \lambda c_1, & a_{21}(\lambda) &= w_2^T x_1 = d_1 \\ a_{12}(\lambda) &= w_1^T(\lambda)x_2 = b_2 + \lambda(c_2 + b_3) + \lambda^2 c_3, & a_{22}(\lambda) &= w_2^T x_2 = d_2 + \lambda d_3, \\ a_{13}(\lambda) &= w_1^T(\lambda)x_3 = b_4 + \lambda(c_4 + b_5) + \lambda^2 c_5, & a_{23}(\lambda) &= w_2^T x_3 = d_4 + \lambda d_5 \end{aligned}$$

y, en consecuencia, la matriz mosaico Toeplitz $A_3(3)$ es la matriz 9×10

$$A_3(3) = \left[\begin{array}{cccc|ccc|ccc} b_1 & 0 & 0 & 0 & b_2 & 0 & 0 & b_4 & 0 & 0 \\ c_1 & b_1 & 0 & 0 & c_2 + b_3 & b_2 & 0 & c_4 + b_5 & b_4 & 0 \\ 0 & c_1 & b_1 & 0 & c_3 & c_2 + b_3 & b_2 & c_5 & c_4 + b_5 & b_4 \\ 0 & 0 & c_1 & b_1 & 0 & c_3 & c_2 + b_3 & 0 & c_5 & c_4 + b_5 \\ 0 & 0 & 0 & c_1 & 0 & 0 & c_3 & 0 & 0 & c_5 \\ \hline d_1 & 0 & 0 & 0 & d_2 & 0 & 0 & d_4 & 0 & 0 \\ 0 & d_1 & 0 & 0 & d_3 & d_2 & 0 & d_5 & d_4 & 0 \\ 0 & 0 & d_1 & 0 & 0 & d_3 & d_2 & 0 & d_5 & d_4 \\ 0 & 0 & 0 & d_1 & 0 & 0 & d_3 & 0 & 0 & d_5 \end{array} \right].$$

Puede comprobarse numéricamente con el programa MATLAB que esta matriz tiene rango completo por filas para b_i, c_i y d_i elegidos aleatoriamente. La matriz 7×7 $A_3(2)$ se construye de un modo similar, y puede ser comprobado numéricamente que tiene también rango completo por filas.

6.5.5. Sobre la genericidad de las hipótesis

La relevancia del Teorema 6.5.10 depende de la *genericidad* de sus hipótesis, es decir: de si son satisfechas en un conjunto abierto denso del conjunto de perturbaciones que se considera. El significado y la genericidad de la condición $\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M) < n$ ya se ha discutido en profundidad en la Sección 6.3. Las otras dos hipótesis esenciales del Teorema 6.5.10 son (6.30-6.31). Ambas hacen referencia al rango de las matrices mosaico Toeplitz $A_s(d_{\min-1})$ y $A_s(d_{\min})$. Aunque no disponemos de una demostración rigurosa que garantice que las condiciones (6.30-6.31) se satisfacen para la mayoría de las perturbaciones M , hemos realizado suficientes experimentos numéricos para convencernos de que así es. Concretamente, hemos comprobado, a través del programa MATLAB testcomp.m, sobre un conjunto de 60.000 casos, que las matrices Toeplitz $A_s(d_{\min})$ y $A_s(d_{\min-1})$ tienen rango máximo. En dichos casos se han considerado distintos tamaños para los bloques singulares del haz de partida y también distintos valores para el rango y el orden singular de la perturbación. Estos experimentos, junto a los programas empleados, se detallan en la Sección 6.9.

Por otra parte, en [16] se incluye la restricción

$$\text{rg}(M) \leq \text{rg}(H) \quad (6.38)$$

con el fin de garantizar que los polinomios conexión son aleatorios para perturbaciones aleatorias (véase [16, §5.5]). Esta hipótesis, aunque bastante natural dado que estamos considerando perturbaciones de rango bajo, no es necesaria, como demuestran los experimentos a los que acabamos de aludir. En ellos hemos considerado tanto perturbaciones $M(\lambda)$ que satisfacen (6.38) como perturbaciones cuyo rango es mayor que el rango del haz no perturbado, $H(\lambda)$. Como se dice en [16, §5.5], si se da esta última condición (es decir: $\text{rg}(M) > \text{rg}(H)$), los vectores fila $w_i(\lambda)$ de una descomposición derecha (6.3) de $M(\lambda)$ tendrán necesariamente entradas linealmente dependientes en el espacio de coeficientes \mathbb{C}^n , pero estas relaciones de dependencia no tienen por qué implicar necesariamente otras relaciones de dependencia entre las columnas (resp. las filas) de la matriz $A_s(d_{\min-1})$ (resp. $A_s(d_{\min})$). Los experimentos aludidos muestran que, de hecho, estas relaciones no existen genéricamente.

6.5.6. Cuando la única información disponible de la perturbación es su rango

El Teorema 6.5.10 determina completamente el conjunto de índices minimales derechos del haz perturbado $(H+M)(\lambda)$. Este conjunto depende de $\tilde{\varepsilon}$, es decir, del orden singular derecho de la perturbación $M(\lambda)$. La causa de esta dependencia se encuentra en el desarrollo (6.3), puesto que las propiedades de (6.3) dependen de $\tilde{\varepsilon}$. Este hecho está relacionado con los resultados de la Sección 7, concretamente con el Teorema 7.3.4, en virtud del cual el conjunto de haces matriciales de rango menor o igual que ρ tiene exactamente $\rho + 1$ componentes maximales, cada una de las cuales corresponde precisamente a un valor de

$\tilde{\varepsilon}$, con $\tilde{\varepsilon} = 0, \dots, \rho$ [17]. No obstante, se puede obtener información parcial en el caso en que el único dato conocido de la perturbación $M(\lambda)$ es su rango. Esta información parcial se presenta en el Teorema 6.5.12. Para demostrar dicho teorema es necesario el siguiente lema.

Lema 6.5.11 Sean $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$ un total de p números enteros, y $\rho, \tilde{\varepsilon}$ dos números enteros tales que $0 < \rho < p$ y $0 \leq \tilde{\varepsilon} \leq \rho$. Consideremos, para cada valor de $\tilde{\varepsilon}$, la sucesión de números enteros:

$$d_k(\tilde{\varepsilon}) = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p. \quad (6.39)$$

Sean $d_{\min}(\tilde{\varepsilon}) \equiv \min_{\rho+1 \leq k \leq p} d_k(\tilde{\varepsilon})$, y $s(\tilde{\varepsilon})$ el mayor índice tal que $d_{s(\tilde{\varepsilon})}(\tilde{\varepsilon}) = d_{\min}(\tilde{\varepsilon})$ y $d_{s(\tilde{\varepsilon})}(\tilde{\varepsilon}) \geq \varepsilon_{s(\tilde{\varepsilon})}$. Entonces

$$d_{\min}(\rho) \geq d_{\min}(\rho - 1) \geq \dots \geq d_{\min}(0) \quad \text{y} \quad s(\rho) \geq s(\rho - 1) \geq \dots \geq s(0).$$

Demostración. Demostraremos que $d_{\min}(\tilde{\varepsilon}') \geq d_{\min}(\tilde{\varepsilon})$ y $s(\tilde{\varepsilon}') \geq s(\tilde{\varepsilon})$, siempre que $\tilde{\varepsilon}' > \tilde{\varepsilon}$. Nótese que $d_k(\tilde{\varepsilon}') \geq d_k(\tilde{\varepsilon})$, para $k = \rho + 1, \dots, p$. Por lo tanto $d_{\min}(\tilde{\varepsilon}') \geq d_{\min}(\tilde{\varepsilon})$. De acuerdo con el apartado 6 del Lema 6.5.9, $\varepsilon_{s(\tilde{\varepsilon}')} + 1 > d_{s(\tilde{\varepsilon}')}(s(\tilde{\varepsilon}')) \geq d_{s(\tilde{\varepsilon})}(\tilde{\varepsilon}) \geq \varepsilon_{s(\tilde{\varepsilon})}$. Esto significa que $s(\tilde{\varepsilon}') + 1 > s(\tilde{\varepsilon})$. ■

Combinando el Teorema 6.5.10 y el Lema 6.5.11, podemos establecer el Teorema genérico 6.5.12. Calificamos de genérico al Teorema 6.5.12 porque las condiciones necesarias son (6.30-6.31), que dependen del orden singular derecho de la perturbación M , una información de la que no se dispone. Para demostrar el Teorema 6.5.12 simplemente hay que darse cuenta de que si ρ es el rango de M y $\tilde{\varepsilon}$ es el orden singular derecho de M entonces $0 \leq \tilde{\varepsilon} \leq \rho$.

Teorema 6.5.12 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos de tamaño $m \times n$ tales que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) < n$. Definamos $\rho \equiv \text{rg}(M)$. Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ los índices minimales derechos de H , y

$$d'_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \rho}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p. \quad (6.40)$$

Sea d'_{\min} el mínimo de la serie $\{d'_k\}$, y s' el mayor índice tal que $d'_{s'} = d'_{\min}$ y $d'_{s'} \geq \varepsilon_{s'}$. Entonces, para un haz genérico $M(\lambda)$ de rango ρ , $(H + M)(\lambda)$ tiene exactamente $p - \rho$ índices minimales derechos y:

1. Los mayores $p - s'$ índices minimales derechos de $(H + M)(\lambda)$ son $\varepsilon_{s'+1} \leq \dots \leq \varepsilon_p$.
2. Los menores $s' - \rho$ índices minimales derechos de $(H + M)(\lambda)$, $\hat{\varepsilon}_1 \leq \dots \leq \hat{\varepsilon}_{s'-\rho}$, satisfacen $\varepsilon_{\rho+j} \leq \hat{\varepsilon}_j$, para $j = 1, \dots, s' - \rho$, y $\hat{\varepsilon}_1 \leq d'_{\min}$.

6.6. La forma canónica de Kronecker de haces perturbados sin rango completo

Los resultados que se han presentado hasta el momento en este capítulo son válidos siempre que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) \leq \min\{m, n\}$. Esta hipótesis incluye el caso

límite $\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M) = \min\{m, n\}$, esto es: el caso en que el haz perturbado $(H+M)(\lambda)$ tiene rango completo. En este caso de rango completo el haz $H+M$ no tiene índices minimales izquierdos si $\text{rg}(H+M) = m$, y $H+M$ no tiene índices minimales derechos si $\text{rg}(H+M) = n$, de acuerdo con (2.6). Si $\text{rg}(H+M) = m < n$ los índices minimales derechos genéricos de $H+M$ están descritos en el Teorema 6.5.10, y si $\text{rg}(H+M) = n < m$ los índices minimales izquierdos de $H+M$ están descritos en el mismo Teorema 6.5.10 pero aplicado a $(H+M)^T$. El Teorema 6.4.5 también se cumple en el caso de rango completo y nos ofrece información parcial sobre la estructura regular de $H+M$.

El propósito de esta sección es mostrar que se puede obtener información completa sobre la FCK genérica de $H+M$ para haces perturbados sin rango completo, es decir, en el caso en que

$$\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M) < \min\{m, n\}.$$

Recopilaremos la información obtenida por los Teoremas 6.4.5 y 6.5.10, junto con la correspondiente versión del Teorema 6.5.10 para los índices minimales izquierdos, con el fin de describir completamente la FCK de $H+M$ en términos de los órdenes singulares derechos e izquierdos y la estructura regular de la perturbación $M(\lambda)$. Dicha FCK se presentará en el Teorema 6.6.2. Además, el Teorema 6.6.5 presentará cierta información parcial genérica de la FCK de $H+M$ cuando $\text{rg}(M)$ es la única información disponible de la perturbación.

El Lema 6.6.1 nos permitirá evitar cierta redundancia en las hipótesis.

Lema 6.6.1 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales de tamaño $m \times n$ tales que $\text{rg}(H) + \text{rg}(M) < \min\{m, n\}$. Sea $A_s(d_{\min})$ la matriz mosaico Toeplitz asociada a un conjunto completo de polinomios conexión derechos de H y M que aparece en el Teorema 6.5.10, y $B_t(h_{\min})$ la correspondiente matriz asociada a un conjunto completo de polinomios conexión izquierdos, es decir, la matriz del Teorema 6.5.10 cuando se aplica a H^T y M^T . Si $A_s(d_{\min})$ y $B_t(h_{\min})$ tienen rango completo por filas, entonces $\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M)$.

Demostración. Del álgebra lineal elemental sabemos que $\text{rg}(H+M) \leq \text{rg}(H) + \text{rg}(M)$. Consideremos las descomposiciones derechas de H y M del tipo que aparece en (6.3):

$$\begin{aligned} H(\lambda) &= v'_1(\lambda)w'_1(\lambda)^T + \dots + v'_r(\lambda)w'_r(\lambda)^T, \\ M(\lambda) &= v_1(\lambda)w_1(\lambda)^T + \dots + v_\rho(\lambda)w_\rho(\lambda)^T, \end{aligned}$$

donde $r \equiv \text{rg}(H)$ y $\rho \equiv \text{rg}(M)$. Por lo tanto

$$H+M = [v'_1, \dots, v'_r, v_1, \dots, v_\rho] [w'_1, \dots, w'_r, w_1, \dots, w_\rho]^T,$$

donde la dependencia de λ se ha omitido. Esto significa que $H+M$ es el producto de una matriz de tamaño $m \times (r+\rho)$ por una matriz de tamaño $(r+\rho) \times n$, con $(r+\rho) < \min\{m, n\}$. Por tanto $\text{rg}(H+M) = \text{rg}(H) + \text{rg}(M)$ si y sólo si

$$\text{rg}[v'_1, \dots, v'_r, v_1, \dots, v_\rho] = r + \rho \quad \text{y} \quad \text{rg}[w'_1, \dots, w'_r, w_1, \dots, w_\rho] = r + \rho.$$

Demostremos que si $A_s(d_{\min})$ tiene rango completo por filas, entonces

$$\text{rg}[w'_1, \dots, w'_r, w_1, \dots, w_\rho] = r + \rho.$$

Si $\text{rg}[w'_1, \dots, w'_r, w_1, \dots, w_\rho] < r + \rho$ existe un índice i tal que $w_i(\lambda)$ es una combinación lineal de $\{w'_1(\lambda), \dots, w'_r(\lambda), w_1(\lambda), \dots, w_{i-1}(\lambda)\}$ en $\mathbb{C}^n(\lambda)$, es decir,

$$w_i(\lambda) = \beta'_1(\lambda)w'_1(\lambda) + \dots + \beta'_r(\lambda)w'_r(\lambda) + \beta_1(\lambda)w_1(\lambda) + \dots + \beta_{i-1}(\lambda)w_{i-1}(\lambda),$$

para ciertas funciones racionales $\beta'_1(\lambda), \dots, \beta_{i-1}(\lambda)$. Recordemos (6.17) y la igualdad $H(\lambda)x_j(\lambda) = 0$, es decir, $w'_k(\lambda)^T x_j(\lambda) = 0$ para todo k . En consecuencia, los polinomios conexión derechos de H y M satisfacen

$$a_{ij}(\lambda) = \beta_1(\lambda)a_{1j}(\lambda) + \dots + \beta_{i-1}(\lambda)a_{i-1,j}(\lambda), \quad \text{para } j = 1, \dots, p,$$

y la matriz $[a_{kl}(\lambda)]_{\substack{1 \leq l \leq p \\ 1 \leq k \leq \rho}}$ no tiene rango completo por filas. Pero el hecho de que $A_s(d_{\min})$ tiene rango completo por filas implica que $\text{rg}[a_{kl}(\lambda)]_{\substack{1 \leq l \leq p \\ 1 \leq k \leq \rho}} = \rho$, por el apartado 6 del Lema 6.5.8.

Un argumento análogo muestra que si $B_t(h_{\min})$ tiene rango completo por filas entonces $\text{rg}[v'_1, \dots, v'_r, v_1, \dots, v_\rho] = r + \rho$. ■

No impondremos explícitamente la condición $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M)$ en el Teorema 6.6.2 a causa del Lema 6.6.1.

Teorema 6.6.2 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos de tamaño $m \times n$ tales que $\text{rg}(H) + \text{rg}(M) < \min\{m, n\}$ y $\rho \equiv \text{rg}(M)$. Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ y $\eta_1 \leq \dots \leq \eta_q$, respectivamente, los índices minimales derechos e izquierdos de H , y \mathcal{J}_H la estructura regular de la FCK de H . Sean $\tilde{\varepsilon}$ y $\tilde{\eta}$, respectivamente, el orden singular derecho y el orden singular izquierdo de M , y \mathcal{J}_M la estructura regular de la FCK de M . Consideremos las sucesiones

$$d_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p, \quad \text{y}$$

$$h_l = \left\lfloor \frac{\sum_{i=1}^l \eta_i + \tilde{\eta}}{l - \rho} \right\rfloor, \quad \text{para } l = \rho + 1, \dots, q.$$

Sea $d_{\min} = \min_{\rho+1 \leq k \leq p} \{d_k\}$, y s el mayor índice tal que $d_s = d_{\min}$ y $d_s \geq \varepsilon_s$. Sean $h_{\min} = \min_{\rho+1 \leq l \leq q} \{h_l\}$, y t el mayor índice tal que $h_t = h_{\min}$ y $h_t \geq \eta_t$. Finalmente, sean $A_s(d_{\min} - 1)$ y $A_s(d_{\min})$ ($B_t(h_{\min} - 1)$ y $B_t(h_{\min})$) las s -ésimas (t -ésimas) matrices mosaico Toeplitz de grados $d_{\min} - 1$ y d_{\min} ($h_{\min} - 1$ y h_{\min}), respectivamente, asociadas al conjunto completo derecho (izquierdo) de polinomios conexión de H y M . Si

$A_s(d_{\min} - 1)$ y $B_t(h_{\min} - 1)$ tienen rango completo por columnas o son matrices vacías, y
 $A_s(d_{\min})$ y $B_t(h_{\min})$ tienen rango completo por filas,

entonces

1. $(H + M)(\lambda)$ tiene exactamente $p - \rho$ índices minimales derechos, que son

$$\underbrace{d_{\min} = \dots = d_{\min}}_{s - \rho - \gamma_s} < \underbrace{(d_{\min} + 1) = \dots = (d_{\min} + 1)}_{\gamma_s} \leq \varepsilon_{s+1} \leq \dots \leq \varepsilon_p, \quad (6.41)$$

donde γ_s es el resto de la división entera de $\sum_{i=1}^s \varepsilon_i + \tilde{\varepsilon}$ entre $s - \rho$;

2. $(H + M)(\lambda)$ tiene exactamente $q - \rho$ índices minimales izquierdos, que son

$$\underbrace{h_{\min} = \dots = h_{\min}}_{t - \rho - \mu_t} < \underbrace{(h_{\min} + 1) = \dots = (h_{\min} + 1)}_{\mu_t} \leq \eta_{t+1} \leq \dots \leq \eta_q, \quad (6.42)$$

donde μ_t es el resto de la división entera de $\sum_{i=1}^t \eta_i + \tilde{\eta}$ entre $t - \rho$; y

3. $\mathcal{J}_H \oplus \mathcal{J}_M$ es la estructura regular de la FCK de $(H + M)(\lambda)$.

Esto determina completamente la FCK de $(H + M)(\lambda)$.

Demostración. El Teorema 6.5.10 aplicado a H y M demuestra (6.41), y aplicado a H^T y M^T demuestra (6.42). El Teorema 6.4.5 demuestra que, para cada número complejo λ_0 , incluido el infinito, $\mathcal{S}_{H+M}(\lambda_0) \geq \mathcal{S}_{H \oplus M}(\lambda_0)$. Para demostrar que, de hecho, $\mathcal{S}_{H+M}(\lambda_0) = \mathcal{S}_{H \oplus M}(\lambda_0)$, mostraremos simplemente que la suma directa de $\mathcal{J}_H \oplus \mathcal{J}_M$ con los bloques singulares derechos e izquierdos correspondientes a (6.41) y a (6.42) es un haz $m \times n$. Llamaremos a esta suma directa $Z(\lambda)$.

Supongamos que la matriz \mathcal{J}_H es de tamaño $r_1 \times r_1$, y la matriz \mathcal{J}_M es de tamaño $r_2 \times r_2$. Nótese que, en esta situación, son ciertas las siguientes identidades:

$$\varepsilon + \eta + q + r_1 = m, \quad \varepsilon + p + \eta + r_1 = n, \quad \tilde{\varepsilon} + \tilde{\eta} + r_2 = \rho,$$

donde ε (resp. η) es el orden singular derecho (resp. izquierdo) de H . Así, el número de filas de $Z(\lambda)$ es:

$$\begin{aligned} & [d_{\min}(s - \rho - \gamma_s) + (d_{\min} + 1)\gamma_s + \varepsilon_{s+1} + \dots + \varepsilon_p] + \\ & + [(h_{\min} + 1)(t - \rho - \mu_t) + (h_{\min} + 2)\mu_t + (\eta_{t+1} + 1) + \dots + (\eta_q + 1)] + r_1 + r_2 = \\ & = [\varepsilon + \tilde{\varepsilon}] + [\eta + q + \tilde{\eta} - \rho] + r_1 + r_2 = m + \tilde{\varepsilon} + \tilde{\eta} + r_2 - \rho = m. \end{aligned}$$

Un cálculo análogo muestra que el número de columnas de $Z(\lambda)$ es n , y, por tanto, que $Z(\lambda)$ es la FCK de $(H + M)(\lambda)$. ■

Ejemplo 6.6.3 Sea $H(\lambda)$ un haz cuadrado de tamaño 31×31 cuya FCK está dada por $\mathcal{K}_H(\lambda) = L_2(\lambda) \oplus L_3(\lambda) \oplus L_4(\lambda) \oplus L_4(\lambda) \oplus L_5(\lambda) \oplus L_0^T(\lambda) \oplus L_1^T(\lambda) \oplus L_1^T(\lambda) \oplus L_3^T(\lambda) \oplus L_3^T(\lambda)$.

Consideremos perturbaciones $M(\lambda)$ con

$$\text{rg}(M) = \rho = 2, \quad \tilde{\varepsilon} = 2, \quad \tilde{\eta} = 0.$$

En esta situación:

$$d_3 = \left\lfloor \frac{2 + 3 + 4 + 2}{1} \right\rfloor = 11, \quad d_4 = \left\lfloor \frac{2 + 3 + 4 + 4 + 2}{2} \right\rfloor = 7, \quad d_5 = \left\lfloor \frac{2 + 3 + 4 + 4 + 5 + 2}{3} \right\rfloor = 6,$$

$$h_3 = \left\lfloor \frac{0 + 1 + 1 + 0}{1} \right\rfloor = 2, \quad h_4 = \left\lfloor \frac{0 + 1 + 1 + 3 + 0}{2} \right\rfloor = 2, \quad h_5 = \left\lfloor \frac{0 + 1 + 1 + 3 + 3 + 0}{3} \right\rfloor = 2,$$

de donde se tiene que $s = 5$ y $t = 3$. Además, $\gamma_5 = 2$ y $\mu_3 = 0$. Como $5 - \rho - \gamma_5 = 1$, la estructura singular derecha de la FCK de $H + M$ estará formada, genéricamente, por

$$L_6(\lambda) \oplus L_7(\lambda) \oplus L_7(\lambda).$$

Por otra parte, como $3 - \rho - \mu_3 = 1$, la estructura singular izquierda estará formada por

$$L_2^T(\lambda) \oplus L_3^T(\lambda) \oplus L_3^T(\lambda),$$

donde los últimos dos bloques corresponden a los dos bloques mayores de la FCK de H , que permanecen inalterados. Así pues, la FCK genérica de $H + M$ es

$$\mathcal{K}_{H+M}(\lambda) = L_6(\lambda) \oplus L_7(\lambda) \oplus L_7(\lambda) \oplus L_2^T(\lambda) \oplus L_3^T(\lambda) \oplus L_3^T(\lambda).$$

□

Ejemplo 6.6.4 Apliquemos el Teorema 6.6.2 al haz $H(\lambda)$ del ejemplo de la Subsección 6.5.4. En dicho ejemplo, considerábamos una perturbación $M(\lambda)$ con $\rho = 2$ y $\tilde{\varepsilon} = 1$. Ahora, supondremos también que $\tilde{\eta} = 0$ y que M tiene un sólo autovalor $\mu = 1$. El índice minimal derecho genérico de $H + M$ predicho por el Teorema 6.6.2 fue calculado en la Subsección 6.5.4 y es igual a 3. Calculemos los índices minimales izquierdos genéricos de $H + M$. Tenemos que $\eta_1 = \eta_2 = \eta_3 = 0$, de manera que el número de índices minimales izquierdos de $H + M$ es $q - \rho = 3 - 2 = 1$. Por lo tanto, $h_{\min} = h_3 = 0$ es el índice minimal izquierdo genérico de $H + M$. La FCK genérica de $H + M$ es, por tanto,

$$L_3(\lambda) \oplus L_0^T \oplus (\lambda - J_1(1)) = \begin{bmatrix} \lambda & -1 & 0 & 0 & 0 \\ 0 & \lambda & -1 & 0 & 0 \\ 0 & 0 & \lambda & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda - 1 \end{bmatrix}.$$

□

En el caso de que la única información disponible acerca de la perturbación $M(\lambda)$ sea su rango, el Teorema 6.4.5 se puede combinar con el Teorema 6.5.12, y la correspondiente versión para los índices minimales izquierdos, para dar como resultado el Teorema 6.6.5, que aporta información parcial de la FCK de $(H + M)(\lambda)$.

Teorema 6.6.5 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos de tamaño $m \times n$ tales que $\text{rg}(H) + \text{rg}(M) < \min\{m, n\}$. Definamos $\rho \equiv \text{rg}(M)$. Sean $\varepsilon_1 \leq \dots \leq \varepsilon_p$ y $\eta_1 \leq \dots \leq \eta_q$, respectivamente, los índices minimales derechos e izquierdos de H , y \mathcal{J}_H la estructura regular de la FCK de H . Consideremos las sucesiones

$$d'_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \rho}{k - \rho} \right\rfloor, \quad \text{para } k = \rho + 1, \dots, p, \quad \text{y}$$

$$h'_l = \left\lfloor \frac{\sum_{i=1}^l \eta_i + \rho}{l - \rho} \right\rfloor, \quad \text{para } l = \rho + 1, \dots, q.$$

Sea d'_{\min} (h'_{\min}) el mínimo de la sucesión $\{d'_k\}$ ($\{h'_l\}$), y s' (t') el mayor índice tal que $d'_{s'} = d'_{\min}$ ($h'_{t'} = h'_{\min}$) y $d'_{s'} \geq \varepsilon_{s'}$ ($h'_{t'} \geq \eta_{t'}$). Entonces, para una perturbación genérica $M(\lambda)$ con rango ρ , $(H + M)(\lambda)$ tiene exactamente $p - \rho$ índices minimales derechos y $q - \rho$ índices minimales izquierdos y:

1. Los $p - s'$ mayores índices minimales derechos de $(H + M)(\lambda)$ son $\varepsilon_{s'+1} \leq \dots \leq \varepsilon_p$.

2. Los $s' - \rho$ menores índices minimales derechos de $(H + M)(\lambda)$, $\hat{\varepsilon}_1 \leq \dots \leq \hat{\varepsilon}_{s'-\rho}$, satisfacen $\varepsilon_{\rho+j} \leq \hat{\varepsilon}_j$, para $j = 1, \dots, s' - \rho$, y $\hat{\varepsilon}_1 \leq d'_{min}$.
3. Los $q - t'$ mayores índices minimales izquierdos de $(H + M)(\lambda)$ son $\eta_{t'+1} \leq \dots \leq \eta_q$.
4. Los $t' - \rho$ menores índices minimales izquierdos de $(H + M)(\lambda)$, $\hat{\eta}_1 \leq \dots \leq \hat{\eta}_{t'-\rho}$, satisfacen $\eta_{\rho+j} \leq \hat{\eta}_j$, para $j = 1, \dots, t' - \rho$, y $\hat{\eta}_1 \leq h'_{min}$.
5. La estructura regular de la FCK de $(H + M)(\lambda)$ contiene a \mathcal{J}_H .

6.6.1. Relación con la teoría de órbitas

Aunque las perturbaciones de rango bajo no son necesariamente perturbaciones de norma pequeña, es natural pensar, a priori, que, en el caso de que sus entradas sí sean de norma pequeña los resultados que hemos obtenido en los apartados anteriores habrán de ser coherentes con los resultados de la teoría clásica de perturbaciones. En concreto, una perturbación genérica $M(\lambda)$ de un haz $H(\lambda)$ debería ir a parar a una órbita genérica (véase [22]) (si existe) entre todas aquellas que satisfacen las restricciones derivadas de la condición de rango bajo, es decir: entre todas aquellas órbitas, $\mathcal{O}(\mathcal{K})$, para las cuales

$$\text{rg}(\mathcal{K} + \mathcal{K}_H) \leq \text{rg}(M) + \text{rg}(H) = \rho + r$$

(donde \mathcal{K} denota la forma de Kronecker que determina la órbita y \mathcal{K}_H es la FCK de H). Si nos restringimos a perturbaciones que satisfacen la condición (6.7), fundamental en nuestros desarrollos, una perturbación genérica debería llevar la FCK de H a alguna de las formas de Kronecker genéricas de rango $\rho + r$, es decir: debería llevar el haz H a un haz con FCK igual a \mathcal{K}_a (como se define en (7.1)) para algún $a \in \{0, 1, \dots, \rho + r\}$. En cambio, es fácil ver que esto no siempre es así. Basta con tener en cuenta el Lema 6.4.2. Según este resultado, y siempre que se satisfaga la igualdad (6.7), la estructura regular del haz no perturbado, H , permanece tras la perturbación.

Ejemplo 6.6.6 Consideremos un haz $H(\lambda)$ de tamaño 3×3 cuya forma de Kronecker es

$$\mathcal{K}_H(\lambda) = L_0 \oplus L_0 \oplus L_0^T \oplus L_0^T \oplus (\lambda - J_1(\lambda_0)) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda - \lambda_0 \end{bmatrix}.$$

Para una perturbación genérica $M(\lambda)$ de rango 1 la FCK del haz perturbado $(H + M)(\lambda)$ será

$$\begin{array}{ll} L_0 \oplus L_1^T(\lambda) \oplus (\lambda - J_1(\lambda_0)) & \text{si } \tilde{\varepsilon} = 0, \tilde{\eta} = 1 \\ L_1(\lambda) \oplus L_0^T \oplus (\lambda - J_1(\lambda_0)) & \text{si } \tilde{\varepsilon} = 1, \tilde{\eta} = 0 \\ L_0 \oplus L_0^T \oplus (\lambda - J_1(\lambda_0)) \oplus (\lambda - J_1(\lambda_1)) & \text{si } \tilde{\varepsilon} = 0 = \tilde{\eta} \end{array}$$

(donde $J_1(\lambda_1)$ es la estructura regular de $M(\lambda)$). Todas ellas tienen estructura regular. \square

Como veremos en el Capítulo 7, la forma de Kronecker genérica de los haces de tamaño $m \times n$, con $n > m$, es

$$\text{diag} \left(\underbrace{L_{\alpha+1}, \dots, L_{\alpha+1}}_s, \underbrace{L_{\alpha}, \dots, L_{\alpha}}_{n-m-s} \right),$$

donde $\alpha = \lfloor m/(n-m) \rfloor$ y $s \equiv m \pmod{m-n}$ [11, Cor. 7.1]. Podemos dar una justificación de este resultado utilizando el mismo argumento que ha inspirado todo el trabajo anterior

relativo al cambio en la estructura de Kronecker: Dado un haz, $H(\lambda)$, de tamaño $m \times n$, existe un vector $x(\lambda) \in \mathcal{N}(H)$ de grado l si y sólo si

$$H(\lambda)x(\lambda) \equiv 0,$$

y esta ecuación, igualando términos del mismo grado, es equivalente a un sistema lineal de ecuaciones con coeficientes complejos que consta de

$$\begin{cases} m(l+2) & \text{ecuaciones} \\ n(l+1) & \text{incógnitas} \end{cases}.$$

El número de grados de libertad del sistema anterior es $l(n-m) + n - 2m$, luego, para que exista solución (genéricamente) ha de ser $l(n-m) + n - 2m > 0$, o bien

$$l > \frac{2m-n}{n-m} = \frac{m}{n-m} - 1.$$

Así pues, el menor valor de l para el que, genéricamente, existe solución es

$$\alpha = \left\lceil \frac{m}{n-m} - 1 \right\rceil = \left\lfloor \frac{m}{n-m} \right\rfloor$$

(donde $\lceil x \rceil$ es el menor número entero mayor que x). Además, en este caso, hay un total de $\alpha(n-m) + n - 2m$ soluciones, lo que se traduce en la existencia de ese mismo número de vectores en $\mathcal{N}(H)$ de grado α . Si expresamos la división entera de m entre $n-m$ en la forma

$$m = (n-m)\alpha + s, \quad 0 \leq s < n-m,$$

tenemos que hay

$$\alpha(n-m) + n - 2m = n - m - s$$

vectores de grado α en $\mathcal{N}(H)$. Además, para $l = \alpha + 1$, el sistema de ecuaciones anterior tiene

$$(\alpha + 1)(n-m) + n - 2m = 2(n-m) - s$$

grados de libertad. Como $n-m > s$, la anterior cantidad es mayor que s , lo que significa que, genéricamente, podremos encontrar s vectores en $\mathcal{N}(H)$ de grado $\alpha + 1$ linealmente independientes con los anteriores.

6.6.2. Cuando H y M son ambos genéricos

A lo largo de toda esta sección hemos considerado un haz no perturbado H del que se conoce la forma de Kronecker completa y un haz perturbador M del que se conocen sólo su rango y los órdenes singulares derechos e izquierdos. Cabe plantearse la siguiente situación: si tanto H como M son dos haces “genéricos”, es decir: dos haces cuyas formas canónicas de Kronecker son genéricas (en el sentido del Lema 7.3.1), entonces el Teorema 6.6.2 debería arrojar el mismo resultado al intercambiar los papeles de H y M . En otras palabras: al considerar el haz H como haz no perturbado y el haz M como haz perturbador deberían obtenerse los mismos tamaños genéricos para los bloques singulares de $H + M$ que al tomar M como haz no perturbado y H como haz perturbador. Por otra parte, la FCK de $H + M$ deberá corresponder a una de las formas genéricas descritas en el Lema 7.3.1. Este apartado está dedicado a demostrar, de forma directa, que esto es así. Para ello necesitaremos, en primer lugar, el siguiente lema.

Lema 6.6.7 Sea $H(\lambda)$ un haz matricial complejo de rango r con índices minimales derechos $\varepsilon_1 \leq \dots \leq \varepsilon_p$ y cuya forma de Kronecker es genérica. Entonces, para cada par de números enteros no negativos $\tilde{\varepsilon}$ y $\rho \leq p$, la sucesión de números enteros

$$d_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad k = \rho + 1, \dots, p, \quad (6.43)$$

es decreciente. Un resultado análogo se tiene para los índices minimales izquierdos.

Demostración. Según el Lema 7.3.1, la forma de Kronecker de $H(\lambda)$ es

$$\mathcal{K}_a(\lambda) = \text{diag} \left(\underbrace{L_{\alpha+1}, \dots, L_{\alpha+1}}_s, \underbrace{L_\alpha, \dots, L_\alpha}_{n-r-s}, \underbrace{L_{\beta+1}^T, \dots, L_{\beta+1}^T}_t, \underbrace{L_\beta^T, \dots, L_\beta^T}_{m-r-t} \right),$$

para $a = 0, 1, \dots, r$, donde $\alpha = \lfloor a/(n-r) \rfloor$, $s = a \bmod (n-r)$, $\beta = \lfloor (r-a)/(m-r) \rfloor$ y $t = (r-a) \bmod (m-r)$. Por lo tanto, la sucesión (6.43) es de la forma

$$d_k = \left\lfloor \frac{\alpha k + j + \tilde{\varepsilon}}{k - \rho} \right\rfloor, \quad \text{si } k = n - r - s + j \quad (j \geq 0),$$

donde $j = 0$ si $k \leq n - r - s$.

Ahora, si $d_{k+1} > d_k$, para algún $k \in \{\rho + 1, \dots, p\}$, por el apartado 4 del Lema 6.5.9, se tiene que

$$\alpha > d_{k+1} > d_k, \quad \text{si } k < n - r - s,$$

y

$$\alpha + 1 > d_{k+1} \geq d_k + 1, \quad \text{si } k \geq n - r - s.$$

En cualquier caso, tendríamos que $\alpha > d_k$. Pero esto es imposible, ya que

$$d_k = \left\lfloor \frac{\alpha k + j + \tilde{\varepsilon}}{k - \rho} \right\rfloor \geq \left\lfloor \frac{\alpha k}{k - \rho} \right\rfloor \geq \left\lfloor \frac{\alpha k}{k} \right\rfloor = \alpha,$$

(donde $j = 0$ si $k \leq n - r - s$). ■

El resultado fundamental de este apartado nos dice que la forma de Kronecker del haz perturbado descrita en el Teorema 6.6.2 es la misma considerando H como haz no perturbado y M como haz perturbación y viceversa.

Teorema 6.6.8 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales complejos $m \times n$ de rangos r y \tilde{r} e índices minimales derechos $\varepsilon_1 \leq \dots \leq \varepsilon_p$ y $\tilde{\varepsilon}_1 \leq \dots \leq \tilde{\varepsilon}_{\tilde{p}}$, respectivamente. Sean las sucesiones de números enteros:

$$d_k = \left\lfloor \frac{\sum_{i=1}^k \varepsilon_i + \tilde{\varepsilon}}{k - \tilde{r}} \right\rfloor, \quad k = \tilde{r} + 1, \dots, p,$$

y

$$\tilde{d}_k = \left\lfloor \frac{\sum_{i=1}^k \tilde{\varepsilon}_i + \varepsilon}{k - r} \right\rfloor, \quad k = r + 1, \dots, \tilde{p},$$

con $\varepsilon = \varepsilon_1 + \dots + \varepsilon_p$ y $\tilde{\varepsilon} = \tilde{\varepsilon}_1 + \dots + \tilde{\varepsilon}_{\tilde{p}}$, y sean

$$\begin{aligned} d_{\min} &\equiv \min_{\tilde{r}+1 \leq k \leq p} d_k \\ \tilde{d}_{\min} &\equiv \min_{r+1 \leq k \leq \tilde{p}} \tilde{d}_k \end{aligned}.$$

Finalmente, sea s (resp. \tilde{s}) el mayor índice tal que $d_s = d_{\min}$ (resp. $\tilde{d}_{\tilde{s}} = \tilde{d}_{\min}$) y $d_s \geq \varepsilon_s$ (resp. $\tilde{d}_{\tilde{s}} \geq \tilde{\varepsilon}_s$), y γ (resp. $\tilde{\gamma}$) el resto de la división entera de $\sum_{i=1}^s \varepsilon_i + \tilde{\varepsilon}$ (resp. $\sum_{i=1}^{\tilde{s}} \tilde{\varepsilon}_i + \varepsilon$) entre $s - \tilde{r}$ (resp. $\tilde{s} - r$). Si las formas de Kronecker de H y M son genéricas, entonces

$$d_{\min} = \tilde{d}_{\min}$$

y

$$s - \tilde{r} - \gamma = \tilde{s} - r - \tilde{\gamma}.$$

Demostración. Sabemos, por el Lema 6.6.7, que las sucesiones $\{d_k\}_{\tilde{r} \leq k \leq p}$ y $\{\tilde{d}_k\}_{r \leq k \leq \tilde{p}}$ son decrecientes, de modo que el mínimo se alcanza, en ambos casos, en el último término. Para demostrar la primera igualdad del enunciado bastará comprobar que en ambos casos dicho término es el mismo, es decir:

$$d_p = \tilde{d}_{\tilde{p}}.$$

Como

$$d_p = \left\lfloor \frac{\varepsilon + \tilde{\varepsilon}}{p - \tilde{r}} \right\rfloor, \quad \tilde{d}_{\tilde{p}} = \left\lfloor \frac{\varepsilon + \tilde{\varepsilon}}{\tilde{p} - r} \right\rfloor,$$

será suficiente con demostrar que $p - \tilde{r} = \tilde{p} - r$. Pero esta igualdad es consecuencia de la igualdad 2.6.

Para demostrar la segunda identidad del enunciado, podemos suponer, en virtud del Lema 7.3.1, que la FCK de H tiene solamente bloques singulares derechos de tamaños α y $\alpha + 1$ (un total de p bloques) y que la FCK de M tiene solamente bloques singulares derechos de tamaños $\tilde{\alpha}$ y $\tilde{\alpha} + 1$ (un total de \tilde{p} bloques). Observamos, en primer lugar, que si $d_{\min} > \max\{\alpha, \tilde{\alpha}\}$, entonces el resultado es evidente, ya que, por definición, se tiene que $s = p$, $\tilde{s} = \tilde{p}$ y, además, $\gamma = \tilde{\gamma}$, pues en ambos casos se trata del resto de la misma división entera. Por lo tanto, únicamente hemos de analizar el caso en que $d_{\min} = \max\{\alpha, \tilde{\alpha}\}$. Sin pérdida de generalidad, podemos suponer que $\alpha \geq \tilde{\alpha}$ y considerar únicamente los casos:

Caso 1. $d_{\min} = \alpha > \tilde{\alpha}$.

En este caso, s es el número de bloques singulares de H de tamaño α , mientras que $\tilde{s} = \tilde{p}$. Por tanto, las divisiones enteras que definen γ y $\tilde{\gamma}$ son:

$$s\alpha + \tilde{\varepsilon} = \alpha(s - \tilde{r}) + \gamma$$

y

$$\varepsilon + \tilde{\varepsilon} = \alpha(\tilde{p} - r) + \tilde{\gamma}.$$

De la primera de ellas obtenemos que $\gamma = \tilde{\varepsilon} + \alpha\tilde{r}$, luego

$$s - \tilde{r} - \gamma = s - (\alpha + 1)\tilde{r} - \tilde{\varepsilon},$$

mientras que de la segunda, teniendo en cuenta que $\varepsilon = p(\alpha + 1) - s$ y que $\tilde{p} - r = p - \tilde{r}$, se obtiene que $\tilde{\gamma} = \tilde{\varepsilon} + \alpha\tilde{r} + p - s$, de donde

$$\tilde{p} - r - \tilde{\gamma} = p - \tilde{r} - \tilde{\gamma} = s - (\alpha + 1)\tilde{r} - \tilde{\varepsilon}.$$

Caso 2. $d_{\min} = \alpha = \tilde{\alpha}$.

Ahora, tenemos que

$$s\alpha + \tilde{\varepsilon} = \alpha(s - \tilde{r}) + \gamma$$

y

$$\tilde{s}\alpha + \varepsilon = \alpha(\tilde{s} - r) + \tilde{\gamma}.$$

Despejando, nos queda

$$\gamma = \tilde{\varepsilon} + \alpha\tilde{r}$$

y

$$\tilde{\gamma} = \varepsilon + \alpha r.$$

Por otro lado, como $\varepsilon = p(\alpha + 1) - s$ y $\tilde{\varepsilon} = \tilde{p}(\alpha + 1) - \tilde{s}$, se tiene que

$$\begin{aligned} s - \tilde{r} - \gamma &= s - \tilde{r} - \tilde{\varepsilon} - \alpha\tilde{r} = s + \tilde{s} - (\alpha + 1)(\tilde{p} + \tilde{r}) = \tilde{s} + s - (\alpha + 1)(p + r) = \\ &= \tilde{s} - \varepsilon - (\alpha + 1)r = \tilde{s} - r - \tilde{\gamma}, \end{aligned}$$

como queríamos demostrar.

■

6.7. El caso límite $\rho = \min\{p, q\}$

Buena parte de los resultados que hemos presentado a lo largo de este capítulo son válidos para el *caso límite*, es decir: aquél en el que $\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = \min\{m, n\}$. En el caso rectangular – supondremos, sin pérdida de generalidad, que $m < n$ – el haz $H + M$ no tiene índices minimales izquierdos, mientras que, por otro lado, los Teoremas 6.5.10 y 6.5.12 determinan los índices minimales derechos para perturbaciones genéricas. Además, el Teorema 6.4.5 establece una cota inferior en el número y el tamaño de los bloques de la estructura regular de $H + M$ que sigue siendo válida en el caso límite para cualquier perturbación que satisfaga la igualdad anterior. No obstante, en el caso cuadrado, estos resultados pueden producir información absolutamente irrelevante. Ilustraremos este punto mediante dos ejemplos. El primero es el siguiente:

$$\underbrace{\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & 0 \end{bmatrix}}_H + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda \end{bmatrix}}_M = \underbrace{\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 2 \\ 0 & 0 & \lambda \end{bmatrix}}_{H+M}. \quad (6.44)$$

Nótese que H tiene $\text{rg}(H) = 2$, índices minimales $\varepsilon_1 = 2$ y $\eta_1 = 0$, y no tiene ningún autovalor porque su rango es 2 para cualquier valor de λ . Lo mismo se cumple para el haz dual. Así pues, H no tiene estructura regular. El haz M tiene $\text{rg}(M) = 1$, índices minimales $\varepsilon_1 = \varepsilon_2 = 0$ y $\eta_1 = 0, \eta_2 = 1$, y no tiene ningún autovalor. Sin embargo, $H + M$ tiene $\text{rg}(H + M) = 3$, es decir: es un haz regular y no tiene índices minimales, ni derechos ni izquierdos. Nuestra teoría describe este comportamiento, véase el Corolario 6.3.2. Además, $H + M$ tiene al valor $\mu = 0$ como autovalor triple con multiplicidad geométrica igual a 1. Nótese que la información dada por el Teorema 6.4.5 es cierta – $\mathcal{S}_{H+M}(0) = (3, 0, \dots)$ mientras que $\mathcal{S}_{H \oplus M}(0) = (0, 0, \dots)$ – pero irrelevante, porque *no hay ninguna relación evidente entre las estructuras regulares (inexistentes) de H y M y la estructura regular de $H + M$* . Este primer ejemplo ilustra un tipo de perturbaciones que destruyen toda la estructura singular del haz no perturbado y crean una estructura regular en $H + M$ que no existía en H (ni en M). Por lo tanto, la *estructura regular* de $H + M$ se crea a partir de las *estructuras singulares* de H y M . Nótese que este ejemplo no

es un caso particular, porque una vez que H está fijado y el rango de las perturbaciones se fija igual a 1, es genérico que $\text{rg}(H + M) = 3$, véase el Teorema 6.3.1, de tal modo que $H + M$ no tiene índices minimales, sino únicamente estructura regular. El segundo ejemplo es el siguiente:

$$\underbrace{\begin{bmatrix} \lambda - 1 & 0 \\ 0 & 0 \end{bmatrix}}_H + \underbrace{\begin{bmatrix} \lambda & 1 + 2\lambda \\ 2\lambda & 2 + 4\lambda \end{bmatrix}}_M = \underbrace{\begin{bmatrix} 2\lambda - 1 & 1 + 2\lambda \\ 2\lambda & 2 + 4\lambda \end{bmatrix}}_{H+M}. \quad (6.45)$$

En este ejemplo, el haz H tiene $\text{rg}(H) = 1$, índices minimales $\varepsilon_1 = 0$ y $\eta_1 = 0$, y un autovalor simple igual a 1. El haz M tiene $\text{rg}(M) = 1$, índices minimales $\varepsilon_1 = 1$ y $\eta_1 = 0$, y no tiene estructura regular. El haz $H + M$ es regular con determinante $\det(H + M) = 2(1 + 2\lambda)(\lambda - 1)$, lo que significa que $H + M$ tiene dos autovalores simples iguales a $-1/2$ y 1. Nótese que, en este caso, $\mu = 1$ es un autovalor de H y también de $H + M$. Esto está garantizado por el Teorema 6.4.5 y no es una coincidencia. Pero el nuevo autovalor que aparece en $H + M$, es decir: $-1/2$, no está relacionado con la estructura regular de H . En ambos ejemplos, (6.44) y (6.45), parece difícil decir algo sobre la estructura regular de $H + M$ más allá del Teorema 6.4.5, excepto que los nuevos autovalores que aparecen en $H + M$ serán *genéricamente* distintos a los de H . En esta sección describimos cómo pueden obtenerse los nuevos autovalores de $H + M$ a partir de cierta información de los haces H y M . Concretamente, veremos que lo único necesario es una base de Kronecker del espacio singular derecho (resp., izquierdo) de H y las columnas (resp., filas) de una representación izquierda (resp., derecha) de M .

6.7.1. El caso rectangular

En este apartado supondremos, sin pérdida de generalidad, que $m \leq n$. Esto implica, en particular, que $q \leq p$, luego $\min\{p, q\} = q$. La notación que utilizaremos es la misma que en los apartados anteriores. En particular, $\rho = \text{rg}(M)$ y $\tilde{\varepsilon}, \tilde{\eta}$ denotan, respectivamente, el orden singular derecho y el orden singular izquierdo de $M(\lambda)$. La hipótesis básica ahora es:

$$\rho = \min\{p, q\} = q.$$

Igualmente, supondremos la hipótesis

$$\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = r + \rho,$$

lo que implica la igualdad

$$\text{rg}(H + M) = m.$$

Bajo esta hipótesis, el haz perturbado, $H + M$, no tiene bloques singulares izquierdos. En el caso cuadrado ($m = n$), también los bloques singulares derechos desaparecen. En el caso rectangular $m < n$, el cambio genérico de la estructura de Kronecker relativa a los bloques singulares derechos es la que se describe para el caso $\rho < n$ en el Teorema 6.5.10. Discutiremos el cambio que experimentan los bloques singulares izquierdos suponiendo que los haces son rectangulares, mientras en el apartado final trataremos separadamente el caso cuadrado.

Definición 6.7.1 Dadas una representación izquierda y una representación derecha del haz $M(\lambda)$ como en (6.4) y (6.3) (respectivamente) las matrices

$$V_l = [\hat{v}_1(\lambda) \mid \dots \mid \hat{v}_\rho(\lambda)] \quad \text{y} \quad W_r = [w_1(\lambda) \mid \dots \mid w_\rho(\lambda)]^T$$

se llamarán, respectivamente, la **matriz de columnas** de la representación izquierda de $M(\lambda)$ y la **matriz de filas** de la representación derecha de $M(\lambda)$. Nótese que

- V_l tiene exactamente $\tilde{\eta}$ columnas de grado exactamente 1 en λ ,
- W_r tiene exactamente $\tilde{\varepsilon}$ filas de grado exactamente 1 en λ .

Definición 6.7.2 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales singulares $m \times n$. Sea $X_1 \in \mathbb{C}^{n \times (\varepsilon+p)}$ ($Y_2 \in \mathbb{C}^{m \times (\eta+q)}$) una matriz cuyas columnas son una base de Kronecker del espacio singular derecho (izquierdo) de $H(\lambda)$, y V_l, W_r las matrices de columnas y de filas asociadas, respectivamente, a una descomposición izquierda y a una descomposición derecha de $M(\lambda)$. Entonces, las matrices $Y_2^T V_l$ y $W_r X_1$, de tamaños respectivos $(\eta+q) \times q$ y $\rho \times (p+\varepsilon)$, se llaman, respectivamente, la **matriz singular IC asociada a** (Y_2, V_l) y la **matriz singular DF asociada a** (W_r, X_1) .

Nótese que las entradas de $Y_2^T V_l$ y $W_r X_1$ son polinomios de grado, a lo sumo, 1 en λ . Más concretamente, $Y_2^T V_l$ tiene, en general, $\tilde{\eta}$ columnas con grado exactamente 1 y $q - \tilde{\eta}$ columnas de grado 0, y $W_r X_1$ tiene, en general, $\tilde{\varepsilon}$ filas de grado exactamente 1 y $\rho - \tilde{\varepsilon}$ filas de grado 0. En este apartado nos interesaremos exclusivamente por la matriz $Y_2^T V_l$, mientras que $W_r X_1$ cobrará relevancia en el siguiente apartado.

En lo sucesivo, nos referiremos a una *matriz IC* (o *DF*) asociada a H y M sin hacer referencia a la base de Kronecker de H ni a la representación de M cuando los resultados no dependan de estos dos elementos.

Definición 6.7.3 Sea

$$A = \begin{bmatrix} a_{11}^0 & \cdots & a_{1\nu}^0 \\ \vdots & & \vdots \\ a_{11}^{s_1} & \cdots & a_{1\nu}^{s_1} \\ \hline \vdots & \cdots & \vdots \\ \vdots & & \vdots \\ \hline a_{\nu 1}^0 & \cdots & a_{\nu\nu}^0 \\ \vdots & & \vdots \\ a_{\nu 1}^{s_\nu} & \cdots & a_{\nu\nu}^{s_\nu} \end{bmatrix}$$

una matriz con ν columnas particionada en ν bloques horizontales. Sea $s = s_1 + \dots + s_\nu$ el número de filas de A . El **k-ésimo coeficiente de A**, denotado por $\alpha_k(A)$, para $k = 0, 1, \dots, s$, es la suma de todos los menores de A obtenidos tomando una fila de cada bloque horizontal, respetando el orden de los bloques, de modo que la suma total de los superíndices de las filas es igual a k .

Como ejemplo, consideremos

$$A = \begin{bmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{11}^2 & a_{12}^2 & a_{13}^2 \\ \hline a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ \hline a_{31}^0 & a_{32}^0 & a_{33}^0 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{bmatrix}.$$

En este caso particular,

$$\begin{aligned} \alpha_0(A) &= \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix}, \quad \alpha_1(A) = \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix} + \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix}, \\ \alpha_2(A) &= \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix} + \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix}, \\ \alpha_3(A) &= \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^0 & a_{22}^0 & a_{23}^0 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix} + \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^0 & a_{12}^0 & a_{13}^0 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix}, \\ \alpha_4(A) &= \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix} + \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix} + \begin{vmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^0 & a_{32}^0 & a_{33}^0 \end{vmatrix}, \quad \alpha_5(A) = \begin{vmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 \end{vmatrix}. \end{aligned}$$

Obsérvese que tanto α_0 como α_η consisten siempre en un solo menor.

Para enunciar el resultado principal de esta sección necesitaremos las siguientes definiciones:

Definición 6.7.4 Sean k, m_1, \dots, m_q números naturales. Una **representación de k en (m_1, \dots, m_q)** es un vector $(k_1, \dots, k_q) \in (\mathbb{N} \cup \{0\})^q$ con $0 \leq k_i \leq m_i$ para $i = 1, \dots, q$, tal que $k = k_1 + \dots + k_q$.

Denotaremos por $\mu_{(m_1, \dots, m_q)}(k)$ el número de representaciones distintas de k en (m_1, \dots, m_q) .

Dados un número $k \in \mathbb{N}$, una representación (k_1, \dots, k_q) de k en (m_1, \dots, m_q) y una matriz arbitraria

$$G = \begin{bmatrix} f_1^0 \\ f_1^1 \\ \vdots \\ \hline f_1^{m_1} \\ \vdots \\ \vdots \\ \hline f_q^0 \\ f_q^1 \\ \vdots \\ f_q^{m_q} \end{bmatrix} \quad (6.46)$$

con $m_1 + \dots + m_q + q$ filas, f_j^i , de longitud q , para $i = 0, \dots, m_j$, $j = 1, \dots, q$, denotamos por $G(k_1, \dots, k_q)$ al menor de G

$$G(k_1, \dots, k_q) = \begin{vmatrix} f_1^{k_1} \\ \vdots \\ f_q^{k_q} \end{vmatrix}.$$

También necesitaremos el siguiente lema técnico, cuya demostración es directa:

Lema 6.7.5 *Sea*

$$L(\lambda) = A_0 + \lambda A_1 + \dots + \lambda^d A_d$$

un polinomio matricial con coeficientes cuadrados $A_i \in \mathbb{C}^{q \times q}$, para $i = 0, \dots, d$. Expresemos

$$A_i = \begin{bmatrix} f_1^i \\ \vdots \\ f_q^i \end{bmatrix}, \quad i = 0, \dots, d,$$

donde f_j^i , para $j = 1, \dots, q$, son vectores fila de longitud q . Sea G la matriz de (6.46), con $m_1 = \dots = m_q = d$. Entonces, para cada $k = 0, 1, \dots, qd$, el coeficiente de λ^k en el polinomio $\det(L(\lambda))$ es el número

$$c_k = \sum_{(k_1, \dots, k_q)} G(k_1, \dots, k_q),$$

donde la suma varía entre todas las representaciones de k en $(\overbrace{d, \dots, d}^q)$.

Ahora, estamos en disposición de demostrar el resultado principal de este apartado:

Teorema 6.7.6 *Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales singulares de tamaño $m \times n$ con $m \leq n$ tales que la FCK de $H(\lambda)$ contiene q bloques singulares izquierdos y $M(\lambda)$ tiene rango q . Sean η y $\tilde{\eta}$, respectivamente, los órdenes singulares izquierdos de H y M . Si se satisface la igualdad*

$$\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = r + q = m,$$

entonces la estructura regular de $H + M$ tiene por autovalores las $\eta + \tilde{\eta}$ soluciones de la ecuación

$$\alpha_\eta(x)x^\eta + \alpha_{\eta-1}(x)x^{\eta-1} + \dots + \alpha_1(x)x + \alpha_0(x) = 0, \quad (6.47)$$

(incluida la raíz infinito) donde α_j es el j -ésimo coeficiente de una matriz singular IC asociada a H y M .

Observación 6.7.7 *Por transposición, se tiene un resultado análogo para el caso en que $n \leq m$, cambiando q por p (el número de bloques singulares derechos de H), η y $\tilde{\eta}$ por, respectivamente, ε y $\tilde{\varepsilon}$ (los órdenes singulares derechos de H y M) y la ecuación (6.47) por la ecuación*

$$\beta_\varepsilon(x)x^\varepsilon + \beta_{\varepsilon-1}(x)x^{\varepsilon-1} + \dots + \beta_1(x)x + \beta_0(x) = 0$$

donde β_k es el k -ésimo coeficiente de una matriz singular DF asociada a H y M .

Demostración. Dada una matriz $Y_2 \in \mathbb{C}^{m \times (\eta+q)}$ cuyas columnas son una base de Kronecker del espacio singular izquierdo de H , la completamos hasta obtener una matriz

$$Y = [Y_1 \mid Y_2 \mid Y_3],$$

donde $Y_1 \in \mathbb{C}^{m \times \varepsilon}$, $Y_3 \in \mathbb{C}^{m \times k}$ (siendo k el tamaño de la estructura regular \mathcal{J}_H), de modo que

$$Y^T(H + M)X = \mathcal{K}_H(\lambda) + Y^T M(\lambda)X = \mathcal{K}_H(\lambda) + (Y^T V_l)(W_r X),$$

(para una cierta matriz invertible $X \in \mathbb{C}^{n \times n}$). Como la estructura regular de un haz es invariante por equivalencia, bastará demostrar el resultado para el término derecho de la igualdad anterior. Veremos que todo menor no nulo de $\mathcal{K}_H + Y^T M X$ de orden m es divisible por $\det(\mathcal{J}_H(\lambda)) \cdot (\lambda - \lambda_1) \dots (\lambda - \lambda_{\eta+\tilde{\eta}})$, donde $\lambda_1, \dots, \lambda_{\eta+\tilde{\eta}}$ son las raíces de (6.47). Haciendo lo mismo con los haces duales, se tiene el resultado.

Sea \mathcal{M} un menor no (necesariamente) nulo de $\mathcal{K}_H + Y^T M X$ de orden m (nótese que \mathcal{M} depende de λ). Este menor es el determinante de una matriz obtenida a partir de m filas y m columnas de \mathcal{K}_H y las correspondientes m filas y m columnas de $Y^T M X$. En lo sucesivo, y en aras de la simplicidad, identificaremos el menor \mathcal{M} con la matriz de la que procede. Haremos uso del siguiente hecho trivial: *Una submatriz $m \times m$ de $\mathcal{K}_H + Y^T M X$ cuyo primer sumando contiene menos de r filas o columnas linealmente independientes tiene rango menor o igual que m y, por tanto, su determinante es cero.* Como consecuencia, el menor \mathcal{M} ha de contener las siguientes columnas:

- Las columnas de la estructura regular $\mathcal{J}_H(\lambda)$ (k **columnas**).
- Las columnas de los bloques singulares izquierdos $L_{\eta_i}^T(\lambda)$ (η **columnas**).

Resulta ilustrativo para la demostración ver cuál es el origen de las restantes columnas de la primera matriz en el menor \mathcal{M} : puesto que el menor contiene m columnas, debemos añadir un número de columnas igual a $\varepsilon + q$ procedentes de los bloques singulares derechos. Esto es: todas excepto $p - q$ columnas procedentes de estos bloques. Por tanto, hemos de añadir las columnas correspondientes a q bloques completos. Por otro lado, no podemos añadir más de q bloques singulares derechos completos porque, de lo contrario, el número de columnas linealmente independientes sería menor que m : el número de columnas linealmente independientes de s bloques singulares derechos completos es $c - s$, donde c es el número total de columnas, luego, más de q bloques completos darían lugar a un menor \mathcal{M} de rango menor que $(m - q) + q = m$.

Ahora, el hecho básico que utilizaremos es el mismo que anteriormente: una matriz $m \times m$ de rango menor que t más otra matriz $m \times m$ con rango $m - t$ dan lugar a una matriz cuyo determinante es cero. Este hecho nos permitirá descomponer el menor \mathcal{M} en la suma de dos menores con todas las columnas (filas) en común salvo una de ellas, y de manera que uno de ellos es cero:

En primer lugar, consideramos la primera columna, c , de \mathcal{M} correspondiente a la estructura regular de \mathcal{K}_H y descomponemos $c = c_1 + c_2$, donde c_1 es una columna de \mathcal{K}_H y c_2 es una columna de $Y^T M X$. Entonces, $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$, donde \mathcal{M}_1 contiene las mismas filas y columnas que \mathcal{M} pero con c_1 en lugar de c , y \mathcal{M}_2 contiene las mismas filas y columnas que \mathcal{M} , pero con c_2 en lugar de c . Ahora, el menor \mathcal{M}_2 se obtiene de la suma de dos matrices: una de estas matrices tiene rango $m - q - 1$ y la otra tiene rango q . En consecuencia, el menor \mathcal{M}_2 es cero. Procediendo de esta manera con las restantes columnas de \mathcal{M} correspondientes a la estructura regular de \mathcal{K}_H , llegamos a un menor $\tilde{\mathcal{M}}$, igual a \mathcal{M} , cuyas últimas k columnas proceden exclusivamente de \mathcal{K}_H . Más concretamente, $\tilde{\mathcal{M}}$ tiene la forma:

$$\begin{aligned} \tilde{\mathcal{M}} &= \det \left[\begin{array}{c|c} \text{columnas de } H & 0 \\ + & \vdots \\ \text{columnas de } M & 0 \\ & \mathcal{J}_H(\lambda) \end{array} \right] = \det \left[\begin{array}{cc} \tilde{A}_1 & 0 \\ \tilde{A}_2 & \mathcal{J}_H(\lambda) \end{array} \right] = \\ &= \det(\mathcal{J}_H(\lambda)) \cdot \det(\tilde{A}_1) = \det(\mathcal{J}_H(\lambda)) \cdot \tilde{\mathcal{M}}_1. \end{aligned}$$

Ahora, basta con ver que el menor $\widetilde{\mathcal{M}}_1$ es divisible por $\alpha_\eta \lambda^\eta + \alpha_{\eta-1} \lambda^{\eta-1} + \dots + \alpha_1 \lambda + \alpha_0$.
Expresemos

$$\widetilde{A}_1 = \begin{bmatrix} \widetilde{A}_{11} \\ \widetilde{A}_{21} \end{bmatrix},$$

con \widetilde{A}_{11} y \widetilde{A}_{21} de tamaños respectivos $\varepsilon \times (\varepsilon + \eta + q)$ y $(\eta + q) \times (\varepsilon + \eta + q)$. La matriz \widetilde{A}_{21} es de la forma

$$\widetilde{A}_{21} = \text{diag}(L_{\eta_1}^T(\lambda), \dots, L_{\eta_q}^T(\lambda)) + (Y_2^T V_l) W_r^1.$$

donde $W_r X = [W_r^1 \ W_r^2]$ y $W_r^1 \in \mathbb{C}^{q \times (\varepsilon + \eta + q)}$, $W_r^2 \in \mathbb{C}^{q \times k}$. También, escribimos

$$Y_2^T V_l = \begin{bmatrix} f_1^0 \\ f_1^1 \\ \vdots \\ f_1^{\eta_1} \\ \vdots \\ \vdots \\ f_q^0 \\ f_q^1 \\ \vdots \\ f_q^{\eta_q} \end{bmatrix},$$

es decir: f_j^i son las filas de la matriz $Y_2^T V_l$. Ahora, fijamos nuestra atención en las primeras $\eta_1 + 1$ filas de \widetilde{A}_{21} :

$$\begin{bmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & & \ddots & \ddots & \\ & & & \ddots & \lambda \\ & & & & 1 \end{bmatrix}_{(\eta_1+1) \times \eta_1} + \begin{bmatrix} f_1^0 \\ f_1^1 \\ \vdots \\ f_1^{\eta_1} \end{bmatrix} W_r^1.$$

Mediante operaciones elementales por filas, llegamos a la matriz

$$\begin{bmatrix} 0 & & & & \\ -1 & 0 & & & \\ & & \ddots & \ddots & \\ & & & \ddots & 0 \\ & & & & -1 \end{bmatrix}_{(\eta_1+1) \times \eta_1} + \begin{bmatrix} f_1^0 + \lambda f_1^1 + \dots + \lambda^{\eta_1} f_1^{\eta_1} \\ * \\ \vdots \\ * \end{bmatrix} W_r^1,$$

donde las filas que están marcadas con un * no tienen interés en nuestro argumento. Podemos llevar a cabo operaciones análogas a las anteriores con las filas que corresponden

a los restantes bloques singulares izquierdos $L_{\eta_i}^T(\lambda)$ para $i = 2, \dots, q$, y obtenemos

$$\bar{A}_{21} = \text{diag}(L_{\eta_1}^T(0), \dots, L_{\eta_q}^T(0)) + \begin{array}{c} \left[\begin{array}{c} f_1^0 + \lambda f_1^1 + \dots + \lambda^{\eta_1} f_1^{\eta_1} \\ * \\ \vdots \\ * \\ \hline \vdots \\ \vdots \\ \hline f_q^0 + \lambda f_q^1 + \dots + \lambda^{\eta_q} f_q^{\eta_q} \\ * \\ \vdots \\ * \end{array} \right] W_r^1. \end{array}$$

Intercambiando filas, de tal modo que las primeras filas de cada bloque horizontal se llevan a los últimos lugares, llegamos a la matriz

$$\hat{A}_{21} = \begin{bmatrix} C \\ FW_r^1 \end{bmatrix},$$

donde C tiene η filas y

$$F = \begin{bmatrix} f_1^0 + \lambda f_1^1 + \dots + \lambda^{\eta_1} f_1^{\eta_1} \\ \vdots \\ f_q^0 + \lambda f_q^1 + \dots + \lambda^{\eta_q} f_q^{\eta_q} \end{bmatrix}$$

es cuadrada de tamaño q .

El hecho relevante es que el determinante de \tilde{A}_1 y el de

$$\hat{A}_1 = \begin{bmatrix} \tilde{A}_{11} \\ \hat{A}_{21} \end{bmatrix}$$

son iguales porque las matrices que corresponden a las operaciones elementales son unimodulares (aunque dependan de λ). Como

$$\hat{A}_1 = \begin{bmatrix} \tilde{A}_{11} \\ C \\ FW_r^1 \end{bmatrix} = \begin{bmatrix} I_{\varepsilon+\eta} & \\ & F \end{bmatrix} \begin{bmatrix} \tilde{A}_{11} \\ C \\ W_r^1 \end{bmatrix},$$

tenemos

$$\det(\hat{A}_1) = \det(F) \cdot \det \begin{bmatrix} \tilde{A}_{11} \\ C \\ W_r^1 \end{bmatrix},$$

y el resultado se sigue del Lema 6.7.5 aplicado a la matriz F .

Si procedemos de manera análoga con los haces duales $H^\sharp(\lambda)$ y $M^\sharp(\lambda)$ llegaremos a que todo menor de orden m de $(H + M)^\sharp(\lambda) = H^\sharp(\lambda) + M^\sharp(\lambda)$ es divisible por $p^\sharp(\lambda) = (\alpha^\sharp)_0(\lambda)\lambda^\eta + (\alpha^\sharp)_1(\lambda)\lambda^{\eta-1} + \dots + (\alpha^\sharp)_{\eta-1}(\lambda)\lambda + (\alpha^\sharp)_\eta(\lambda)$, donde $(\alpha^\sharp)_j(\lambda) = \lambda^{\tilde{\eta}}\alpha_j(1/\lambda)$ es el polinomio dual de $\alpha_j(\lambda)$, de tal modo que $p^\sharp(\lambda)$ es el polinomio dual de $p(\lambda) = \alpha_\eta(\lambda)\lambda^\eta + \dots + \alpha_1(\lambda)\lambda + \alpha_0(\lambda)$, lo que significa que $p^\sharp(\lambda)$ tiene una raíz cero de multiplicidad ν si y

sólo si $p(\lambda)$ tiene una raíz infinito de multiplicidad ν . Por tanto, dentro de las $\eta + \tilde{\eta}$ raíces de la ecuación (6.47) está incluida la raíz infinito con su multiplicidad correspondiente. ■

Ahora, añadiendo el teorema anterior a los resultados obtenidos en el Teorema 6.6.2, podemos establecer lo siguiente:

Teorema 6.7.8 *En las hipótesis del Teorema 6.7.6, la ecuación (6.47) tiene, para una perturbación genérica $M(\lambda)$, $\eta + \tilde{\eta}$ raíces distintas que no son autovalores de $H(\lambda)$ y, por tanto, la estructura de Kronecker genérica del haz perturbado $H + M$ es*

$$\mathcal{K}_{H+M}(\lambda) = \text{diag}(\mathcal{K}_{\text{sing}}(\lambda), \mathcal{J}_H(\lambda), \mathcal{J}_M(\lambda), \lambda - \mu_1, \dots, \lambda - \mu_{\eta+\tilde{\eta}}),$$

donde

- i) $\mathcal{K}_{\text{sing}}(\lambda)$ es la estructura singular de $\mathcal{K}_{H+M}(\lambda)$ y consta de $p - \rho$ bloques singulares derechos que ocupan un total de $\varepsilon + \tilde{\varepsilon}$ filas, y cuyos tamaños están dados por el Teorema 6.6.2,
- ii) $\mathcal{J}_H(\lambda)$ es la estructura regular de $H(\lambda)$,
- iii) $\mathcal{J}_M(\lambda)$ es la estructura regular de $M(\lambda)$,
- iv) $\mu_1, \dots, \mu_{\eta+\tilde{\eta}}$ son las $\eta + \tilde{\eta}$ soluciones de la ecuación (6.47).

Demostración. El hecho de que para una perturbación genérica las raíces de (6.47) son distintas es consecuencia del hecho, bien conocido, de que un polinomio genérico de coeficientes aleatorios tiene todas sus raíces distintas. La afirmación relativa a los bloques singulares derechos es una consecuencia del Teorema 6.6.2. Por otro lado, los apartados ii) y iii) son consecuencia del Lema 5.2.3. Por último, el Teorema 6.7.6 garantiza que las soluciones de (6.47) son autovalores de $H + M$. Si estos autovalores no se encuentran ni en \mathcal{J}_H ni en \mathcal{J}_M , podemos asegurar que se añaden a $\mathcal{J}_H \oplus \mathcal{J}_M$ en la estructura regular de $H + M$ (nótese que, en caso contrario, no podemos garantizar esto último).

Queda por demostrar, sin embargo, que si las raíces de la ecuación (6.47) son distintas y, además, los autovalores de la estructura regular de H , los de la estructura regular de M y el conjunto de soluciones de (6.47) son conjuntos disjuntos dos a dos, entonces la FCK de $H + M$ no contiene ningún elemento más de los que se describen en los apartados i) – iv) del enunciado. Esto puede verse calculando el número de filas y columnas que ocupan todos estos bloques. Primero, nótese que

$$\rho = q = \tilde{\varepsilon} + \tilde{\eta} + \tilde{k}.$$

Así, i) – iv) suman el siguiente número de filas y columnas:

■ **Filas:**

$$\underbrace{\varepsilon + \tilde{\varepsilon}}_{\text{bloques singulares}} + \underbrace{k}_{\text{estructura regular de H}} + \underbrace{\tilde{k}}_{\text{estructura regular de M}} + \underbrace{\eta + \tilde{\eta}}_{\text{raíces de (6.47)}} = \varepsilon + \eta + k + \tilde{\varepsilon} + \tilde{\eta} + \tilde{k} = m,$$

■ **Columnas:**

$$\underbrace{\varepsilon + \tilde{\varepsilon} + p - \rho}_{\text{bloques singulares}} + \underbrace{k}_{\text{estructura regular de H}} + \underbrace{\tilde{k}}_{\text{estructura regular de M}} + \underbrace{\eta + \tilde{\eta}}_{\text{raíces de (6.47)}} = \varepsilon + \eta + k + p + \tilde{\varepsilon} + \tilde{\eta} + \tilde{k} - q = n.$$

Esto significa que los elementos que se describen en el enunciado ocupan todas las filas y columnas de $H + M$, luego en la FCK de este haz no puede aparecer ningún elemento más. ■

Observación 6.7.9 *Se obtiene un resultado análogo al del Teorema 6.7.8 para el caso en que $n \leq m$ cambiando las raíces de (6.47) por las $\varepsilon + \tilde{\varepsilon}$ raíces de la ecuación que aparece en la Observación 6.7.7 y los $p - \rho$ bloques singulares derechos que describe el Teorema 6.6.2 por los $q - \rho$ bloques singulares izquierdos que describe el mismo Teorema (y que ahora ocuparán un total de $\eta + \tilde{\eta}$ filas).*

Observación 6.7.10 *El Teorema 6.7.8 sólo afirma los los valores que se mencionan en el apartado iv) son autovalores de $H + M$, pero en el caso de que alguno se repita, o coincida con algún autovalor de $H(\lambda)$ o $M(\lambda)$, no asegura que aparezca repetido. Por este motivo el Teorema se enuncia para perturbaciones genéricas.*

Cerramos esta sección con una aplicación del Lema 6.7.5 a la teoría de ecuaciones algebraicas:

Corolario 6.7.11 *Si $\lambda_0, x_1, \dots, x_q$ es una solución del siguiente sistema de ecuaciones:*

$$\begin{cases} \lambda^l(a_{l1}^1 x_1 + \dots + a_{lq}^1 x_q) + \dots + \lambda(a_{11}^1 x_1 + \dots + a_{1q}^1 x_q) + a_{01}^1 x_1 + \dots + a_{0q}^1 x_q = 0 \\ \vdots \\ \lambda^l(a_{l1}^q x_1 + \dots + a_{lq}^q x_q) + \dots + \lambda(a_{11}^q x_1 + \dots + a_{1q}^q x_q) + a_{01}^q x_1 + \dots + a_{0q}^q x_q = 0 \end{cases} \quad (6.48)$$

con $(x_1, \dots, x_q) \neq (0, \dots, 0)$, entonces λ_0 es una raíz del polinomio

$$p(\lambda) = c_{ql}\lambda^{ql} + c_{ql-1}\lambda^{ql-1} + \dots + c_1\lambda + c_0,$$

con c_i , para $i = 0, 1, \dots$, como en el enunciado del Lema 6.7.5 y

$$G = \begin{bmatrix} a_{01}^1 & \dots & a_{0q}^1 \\ \vdots & & \vdots \\ a_{l1}^1 & \dots & a_{lq}^1 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{01}^q & \dots & a_{0q}^q \\ \vdots & & \vdots \\ a_{l1}^q & \dots & a_{lq}^q \end{bmatrix}.$$

Demostración. Si $\lambda_0, x_1, \dots, x_q$ es una solución de (6.48) con (x_1, \dots, x_q) no nulo, entonces λ_0 es un autovalor del polinomio matricial

$$L(\lambda) = A_\ell \lambda^\ell + \dots + A_1 \lambda + A_0,$$

donde

$$A_i = \begin{bmatrix} a_{i1}^1 & \dots & a_{iq}^1 \\ \vdots & & \vdots \\ a_{i1}^q & \dots & a_{iq}^q \end{bmatrix}, \quad i = 1, \dots, \ell,$$

con autovector asociado (x_1, \dots, x_q) . Luego $\det L(\lambda_0) = 0$. Ahora, el resultado es una consecuencia inmediata del Lema 6.7.5. ■

6.7.2. El caso cuadrado

Los haces matriciales singulares cuadrados satisfacen la relación $p = q$, con p y q como en (2.5). Por tanto, en esta sección usaremos únicamente p y n en lugar de p, q y m, n . Ahora, las hipótesis básicas sobre el haz perturbación $M(\lambda)$ son

$$\rho = p,$$

y

$$\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = n.$$

Como se ha mencionado anteriormente, bajo estas hipótesis, el haz perturbado $H + M$ es regular. Esto significa que la FCK no contiene bloques singulares. Así, en lugar de referirnos a la FCK de $H + M$, nos referiremos a la FCW de $H + M$.

Ahora, el Teorema 5.4.2 se puede reescribir en la forma

Teorema 6.7.12 Sean $H(\lambda)$ y $M(\lambda)$ dos haces matriciales singulares de tamaño $n \times n$ tales que la FCK de $H(\lambda)$ contiene exactamente p bloques singulares derechos y otros p bloques singulares izquierdos y se satisface la igualdad

$$\text{rg}(H + M) = \text{rg}(H) + \text{rg}(M) = r + p = n.$$

Sean ε, η y $\tilde{\varepsilon}, \tilde{\eta}$, respectivamente, los órdenes singulares derechos e izquierdos de H y M . Sea β_k , para $k = 0, 1, \dots, \varepsilon$ el k -ésimo coeficiente de una matriz DF asociada a H y M y α_k , para $k = 0, 1, \dots, \eta$, el k -ésimo coeficiente de una matriz IC asociada a H y M . Si las $\varepsilon + \tilde{\varepsilon}$ soluciones, $\lambda_1, \dots, \lambda_{\varepsilon+\tilde{\varepsilon}}$, de la ecuación

$$\beta_\varepsilon(x)x^\varepsilon + \beta_{\varepsilon-1}(x)x^{\varepsilon-1} + \dots + \beta_1(x)x + \beta_0(x) = 0,$$

y las $\eta + \tilde{\eta}$ soluciones, $\mu_1, \dots, \mu_{\eta+\tilde{\eta}}$, de la ecuación

$$\alpha_\eta(x)x^\eta + \alpha_{\eta-1}(x)x^{\eta-1} + \dots + \alpha_1(x)x + \alpha_0(x) = 0,$$

son todas distintas y no son autovalores de H ni de M , la FCW de $H + M$ es

$$\mathcal{W}_{H+M}(\lambda) = \text{diag}(\mathcal{J}_H, \mathcal{J}_M, \lambda - \lambda_1, \dots, \lambda - \lambda_{\varepsilon+\tilde{\varepsilon}}, \lambda - \mu_1, \dots, \lambda - \mu_{\eta+\tilde{\eta}}),$$

donde

i) \mathcal{J}_H es la estructura regular $H(\lambda)$,

ii) \mathcal{J}_M es la estructura regular de $M(\lambda)$.

Demostración. Que las estructuras regulares de H y M permanecen en la estructura regular de $H + M$ es consecuencia del Teorema 5.2.1. Que las raíces de las ecuaciones del enunciado son autovalores de $H + M$ es consecuencia del Teorema 5.4.2 y su análogo para el caso $n \leq m$. Únicamente queda por ver que, en el caso de que estas raíces sean distintas entre sí y, al mismo tiempo, distintas a los autovalores de H y M , la forma canónica de Weierstrass de $H + M$ no contiene ningún autovalor más. Esto es fácil de comprobar, ya que el número de filas (columnas) ocupadas por los anteriores autovalores es

$$k + \tilde{k} + \varepsilon + \tilde{\varepsilon} + \eta + \tilde{\eta} = \varepsilon + \eta + k + \overbrace{\tilde{\varepsilon} + \tilde{\eta} + \tilde{k}}^p = n,$$

(donde k y \tilde{k} son las dimensiones de las estructuras regulares de H y M , respectivamente) luego no puede aparecer ningún autovalor más. ■

6.8. Problemas relacionados

La descripción del cambio genérico de la forma de Jordan de una matriz tras perturbaciones de rango bajo se publicó por vez primera en [34]. El resultado apareció nuevamente demostrado en [60] para perturbaciones de rango 1 y en [61] para perturbaciones de rango bajo. Ninguno de los trabajos anteriores incluía una caracterización del conjunto de perturbaciones para el que se da ese cambio genérico. Esta caracterización se dio en [52], donde también se incluía una nueva demostración del cambio genérico. Las técnicas utilizadas en este último trabajo son las que hemos empleado en el Capítulo 5 para generalizar los resultados al caso de haces matriciales regulares. Unas técnicas completamente distintas, descritas en este capítulo, permiten hacer lo propio con el cambio de la estructura de Kronecker en el caso singular, es decir: describir tanto el cambio genérico como el conjunto de perturbaciones para el que se da dicho cambio. Estos resultados han aparecido recientemente en [16].

Todos los trabajos que acabamos de citar en el párrafo anterior se centran en la descripción del cambio que experimentan las formas canónicas para la mayor parte de las perturbaciones. Otro problema distinto es el de caracterizar **todos los posibles cambios** en la forma canónica tras las perturbaciones que se estén considerando, en nuestro caso, las perturbaciones de rango bajo. Este problema ha sido tratado por R. C. Thompson en [75] y [76]. En el primero de estos trabajos se caracteriza el cambio de los factores invariantes al añadir filas a una matriz (no necesariamente cuadrada) cuyas entradas están en un Dominio de Ideales Principales (DIP) (por ejemplo: el anillo de polinomios en la variable λ con coeficientes complejos), y es la base del segundo de ellos, en el que se describen todos los posibles factores invariantes de las perturbaciones de rango 1 de una matriz con entradas en un DIP: dichos factores invariantes están dados por las relaciones de entrelazamiento que se describen en el enunciado del Teorema 5.2.1 de esta memoria y, aunque nosotros las hemos enunciado únicamente como condiciones de necesidad, son asimismo condiciones suficientes [76, Th. 1].

En la misma línea se encuentran los trabajos de Boley [6] y Beitia et al. [5]. Concretamente, el primero de ellos caracteriza los haces que pueden obtenerse a partir de otro al añadir una fila o columna. Al igual que en los artículos de Thompson [75, 76], la caracterización se obtiene a través de ciertas relaciones de dominancia, en este caso entre los índices minimales y la característica de Segre de uno y otro haz. En cambio, en el segundo de los trabajos, se caracterizan las matrices que pueden obtenerse a partir de otra por pequeñas perturbaciones (en norma) de una fila o columna. Nuevamente, la caracterización llega a través de una relación de dominancia entre determinadas series relacionadas con la forma de Jordan de la matriz. Ambos trabajos siguen la estela de los que anteriormente habían permitido caracterizar las matrices (respectivamente, los haces) que pueden obtenerse a partir de una matriz dada por perturbaciones arbitrarias de norma pequeña [49] (resp. [55]), y de los que hablaremos en el próximo capítulo.

6.9. Apéndice: Programas Matlab

Incluimos en este Apéndice los programas Matlab aludidos en la Sección 6.5.5. Hemos utilizado estos programas para comprobar empíricamente que las matrices mosaico Toeplitz $A_s(d_{min})$ y $A_s(d_{min} - 1)$ que aparecen en el enunciado del Teorema 6.5.10 tienen, genéricamente, rango completo. En la Sección 6.9.2 incluimos los experimentos realizados con estos programas.

6.9.1. Programas Matlab

En el programa Matlab **toecomp.m** se presupone un haz no perturbado en forma de Kronecker, de manera que una ROMB de dicho haz está formada por vectores polinómicos canónicos de la forma $[0 \dots 0 \ 1 \ \lambda \dots \lambda^i \ 0 \dots 0]^T$. El programa tiene como entradas:

- Los índices minimales derechos del haz no perturbado (tamaños de los bloques singulares derechos) (D),
- el rango de la perturbación (r),
- la suma de los índices minimales derechos de la perturbación (e),

y construye una perturbación aleatoria M con rango r y orden minimal derecho e . Más concretamente, construye una descomposición derecha (6.3) de una perturbación M . Para ello se usa el comando 'rand([e+r,n])', que proporciona una matriz $(e+r) \times n$, que hemos denotado con la misma letra M , cuyas $e+r$ filas son los vectores fila de una descomposición derecha como en (6.3) (nótese que una tal descomposición contiene e vectores polinómicos fila de grado 1 y otros $r-e$ de grado 0, lo que suma un total de $e+r$ vectores constantes). A partir de esta descomposición, obtiene un conjunto completo de polinomios conexión derechos en la forma que se describe en la Sección 6.5.1, es decir, multiplicando los vectores de la matriz M anterior por cada uno de los vectores polinómicos canónicos. Posteriormente se construyen las matrices mosaico Toeplitz asociadas $A_s(d_{mín})$ y $A_s(d_{mín} - 1)$ que aparecen en el Teorema 6.5.10. A continuación, calcula el rango de dichas matrices, utilizando el comando Matlab 'rank', y ofrece como salida un valor igual a 1 si las matrices son de rango completo e igual a 0 en caso contrario.

Por otro lado, el programa **testcomp.m** ejecuta el programa toecomp.m en un número de casos prefijado. Este número de casos es la primera entrada del programa. Las restantes entradas son las del programa toecomp.m. La salida de testcomp.m es un número que indica el número de casos en que se cumple la conjetura, es decir, el número de casos en que las matrices mosaico Toeplitz asociadas tienen rango máximo.

La siguiente tabla muestra la equivalencia entre la notación empleada por los programas y la empleada en el resto de la memoria.

<i>Programa</i>	<i>Memoria</i>
e	$\tilde{\varepsilon}$
r	ρ
$D = [d_1, d_2, \dots]$	$[\varepsilon_1, \varepsilon_2, \dots]$

A continuación, incluimos los programas toecomp.m y testcomp.m. Los comentarios, siguiendo la sintaxis Matlab, van precedidos del símbolo %.

```

function[z]=toecomp(e,r,D)

%D es un vector que contiene las dimensiones de los bloques
D=sort(D) %los ordena en orden no decreciente
%e es la suma de '\{i}ndices minimales de la perturbaci'\{o}n
%r es el rango de la perturbaci'\{o}n. Ha de ser mayor o igual que e
%y menor que n

if (e>r)
    disp('rango menor que e!!')
    return
end

[q p]=size(D);
%p es el n'\{u}mero de bloques del haz no perturbado

n=sum(D)+p;
%n es el n'\{u}mero de columnas

if (r>p)
    disp('rango mayor que no. de bloques!!')
    return
end

%calculamos la suma de los '\{i}ndices minimales:
for i=1:p
    S(i)=sum(D(1:i));
end

D=[D 0];

%calculamos los d_i's y el entero s (en la notaci'\{o}n de la memoria):
for i=r+1:p
    s=i;
    H=[D(1:i)];
    F(i-r)=floor((sum(H)+e)/(i-r));
    if D(i+1)>=F(i-r)
        break
        s=i+1;
    end
end

s;
d=min(F);
%d es el grado

for i=1:s
    L(i)=d-D(i);

```

```

end

L;
%L(i) es el grado de las inc\'}gnitas

M=rand([e+r,n]);
%M contiene por filas los vectores fila de una perturbaci\'}n gen\'}rica
%de rango r y suma de \'}ndices minimales e expresada a trav\'}s de una
%descomposici\'}n derecha. Hay e+r porque los primeros e vectores son de grado 1

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%CONSTRUCCION DE LOS POLINOMIOS CONEXION %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if (e==0)
    A=[];
else A=[[M(1,1:D(1)+1) 0]+[0 M(e+1,1:D(1)+1)] zeros(1,d-D(1))];

for i=2:s
    A=[A; [[M(1,S(i-1)+i:S(i)+i) 0]+[0 M(e+1,S(i-1)+i:S(i)+i)] zeros(1,d-D(i))]];
end
%hemos constru\'}do el primer bloque de columnas (que luego ser\'}{a}n filas)
%de los polinomios de grado e_i+1

for i=2:e
    B=[[M(i,1:D(1)+1) 0]+[0 M(e+i,1:D(1)+1)] zeros(1,d-D(1))];
for j=2:s
    B=[B; [[M(i,S(j-1)+j:S(j)+j) 0]+[0 M(e+i,S(j-1)+j:S(j)+j)] zeros(1,d-D(j))]];
end
A=[A B];
end
end

if (e<r) C=[[M(2*e+1,1:D(1)+1)] zeros(1,d-D(1))];
else
    C=[];
    E=[];
end

if (e<r)
for i=2:s
    C=[C; [M(2*e+1,S(i-1)+i:S(i)+i)] zeros(1,d-D(i))];
end
%hemos constru\'}do el primer bloque de columnas (que luego ser\'}{a}n filas)
%de los polinomios de grado e_i
end

if (e+1<r)

```

```

for i=2:(r-e)
    E=[[M(2*e+i,1:D(1)+1) zeros(1,d-D(1))]];
for j=2:s
    E=[E;[[M(2*e+i,S(j-1)+j:S(j)+j) zeros(1,d-D(j))]];
end
C=[C E];
end
end
A=A.'; C=C.';

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% CONSTRUCCION DE LAS MATRICES MOSAICO TOEPLITZ %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if (e==0)
    T0=[];
else T0=toeplitz(A(:,1),[A(1,1) zeros(1,d-D(1))]);
for i=2:s
    T2=toeplitz(A(:,i),[A(1,i) zeros(1,d-D(i))]);
    T0=[T0 T2];
end
end

if (e<r) T1=toeplitz(C(:,1),[C(1,1) zeros(1,d-D(1))]);
for i=2:s
    T2=toeplitz(C(:,i),[C(1,i) zeros(1,d-D(i))]);
    T1=[T1 T2];
end
T=[T0;T1];
else
    T=T0;
end
T;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% CONSTRUCCION DE LAS MATRICES MOSAICO TOEPLITZ %
% QUE SE OBTIENEN QUITANDO LA ULTIMA FILA Y %
% COLUMNA DE CADA BLOQUE %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

L1=T(:,1:L(1)); h=L(1)+2; for i=2:s
    L1=[L1 T(:,h:h+L(i)-1)];
    h=h+L(i)+1;
end
j=d+3;
T1=L1(1:d+1,:);

for i=2:e

```

```

    T1=[T1;L1(j:j+d,:)];
    j=j+d+2;
end
j=e*(d+2)+1;

for i=e+1:r
    T1=[T1;L1(j:j+d-1,:)];
    j=j+d+1;
end
T1;

%calculamos los rangos y comprobamos si las matrices mosaico
%Toeplitz son de rango m\{a}ximo

[a b]=size(T); r1=rank(T); c1=a-r1;

[m,n]=size(T1); r2=rank(T1); c2=n-r2;

if (c1==0) & (c2==0)
    z=1;
    %disp('El caso es gen\{e}rico')
else
    z=0;
    %disp('El caso es no gen\{e}rico')
end

function[suma,Y]=testcomp(n,e,r,D)
Y=zeros([1,n]);
for i=1:n
    X(i)=toecomp(e,r,D);
end
suma=sum(X)

if suma==n
    disp 'se cumple la conjetura'
else
    disp 'no se cumple la conjetura'
end

```

6.9.2. Experimentos

En esta sección incluimos los experimentos que hemos llevado a cabo con el programa testcomp.m. En la siguiente tabla mostramos los resultados y las características del conjunto de entradas sobre el que se ha ejecutado el programa. En todos los ejemplos, el número de casos que se ha ejecutado (primera entrada) es $n = 10000$.

e	r	D	salida(z)
0	3	[0, 1, 1, 1, 2]	10000
3	3	[0, 1, 6, 10]	10000
1	4	[0, 1, 2, 3, 4, 4, 5, 6]	10000
3	5	[1, 1, 2, 3, 4, 5, 5, 6, 7, 9]	10000
4	5	[0, 0, 1, 1, 5, 6, 6]	10000
4	8	[1, 1, 1, 3, 3, 4, 5, 6, 7, 7, 7, 8, 8]	10000

Como puede observarse en la última columna, en todos los casos (un total de 60000) las matrices mosaico Toeplitz asociadas tienen rango máximo.

Capítulo 7

Órbitas de haces matriciales de rango fijo

7.1. Introducción

Como ya hemos mencionado en la Introducción del capítulo anterior, resulta llamativo el hecho de que el cambio genérico que experimenta la FCK de haces matriciales singulares tras perturbaciones de rango bajo dependa de los órdenes singulares de la perturbación. Este hecho, en cambio, tiene una explicación intrínseca al conjunto de perturbaciones que estamos considerando, es decir, al conjunto de perturbaciones de rango (menor o igual que) r (nótese que, por comodidad, hemos preferido emplear esta letra en lugar de ρ , utilizada para el rango de la perturbación en los dos capítulos previos). El objetivo del presente capítulo es dar una descripción de dicho conjunto en la línea del trabajo clásico de Arnold [2]. Según los resultados que contiene dicho trabajo, el espacio de matrices $n \times n$ (identificado de manera natural con el espacio \mathbb{C}^{n^2}) se descompone en una unión de subvariedades (*órbitas*) cuyos puntos son matrices con la misma forma de Jordan. La (co)dimensión de dichas subvariedades se calcula en términos de los tamaños de los bloques que aparecen en la forma canónica. Esta descripción se aplica, en particular, a la subvariedad de matrices $m \times n$ (es decir: no necesariamente cuadradas) de rango menor o igual que r . La generalización de las ideas de Arnold a haces matriciales, llevada a cabo esencialmente por Edelman, Elmroth y Kågström [21, 22], dejaba abierto el problema de dar una descripción análoga para el conjunto de haces matriciales de rango menor o igual que r . Nuestro trabajo, por lo tanto, se debe en gran parte a los de estos tres autores, especialmente [22]. En dicho artículo, a partir de la identificación natural del conjunto de haces $m \times n$ con el espacio \mathbb{C}^{2mn} , se describe a este conjunto como una unión finita de cierres de órbitas, cada una de los cuales está formado por haces estrictamente equivalentes que, por tanto, tienen la misma forma de Kronecker. A partir de la relación de inclusión entre los cierres obtienen un orden parcial en el conjunto de órbitas de haces matriciales de tamaño $m \times n$ y aportan condiciones necesarias y suficientes para que dos haces estén relacionados. El diagrama de Hasse asociado a este orden es de gran utilidad en la interpretación de los algoritmos que se emplean para obtener las formas canónicas que estudiamos en la presente memoria porque permite conocer las estructuras cercanas a un haz determinado del que se conoce su forma canónica total o parcialmente. Si, en particular, el haz se halla en una órbita situada en el cierre de otra, una pequeña perturbación (debida, acaso, a errores de redondeo) situará posiblemente el haz en esa órbita superior, cuya forma canónica puede ser muy distinta, lo que implicará un fallo en

el algoritmo que se esté utilizando.

Las razones expuestas en el párrafo anterior nos muestran la utilidad práctica que tiene conocer, si es que existe, la forma de Kronecker (es decir: la órbita) genérica de un haz $m \times n$ de rango r , donde r es un número entero entre 0 y $\min\{m, n\} - 1$. En el presente capítulo veremos que no existe una órbita genérica, en el sentido de que su cierre contenga a todas las demás, sino que existen $r + 1$ órbitas, que pueden llamarse *maximales*, cuyos cierres contienen a las restantes. La teoría de perturbaciones que hemos expuesto en el Capítulo 6 refleja este hecho y, en consecuencia, los resultados que hemos obtenido dependen de estas $r + 1$ formas maximales, que se parametrizan por los valores de los órdenes singulares compatibles con dicho rango. Esta es la razón intrínseca de la relevancia que tienen los órdenes singulares de la perturbación en el comportamiento genérico descrito en el Capítulo 6.

7.2. Preliminares y resultados previos

En esta sección resumimos brevemente los resultados que necesitaremos en el capítulo. Simultáneamente, introducimos algo de notación.

Recordemos que el *rango* de un haz $A + \lambda B$ es el orden del mayor menor no idénticamente nulo como función de λ .

7.2.1. Órbitas y la forma canónica de Kronecker

Vamos a usar la misma notación que en [22]. La **órbita**, $\mathcal{O}(H)$, de un haz matricial $H(\lambda) = A + \lambda B$ de tamaño $m \times n$, es el conjunto de haces matriciales estrictamente equivalentes a $H(\lambda)$:

$$\mathcal{O}(H) = \{PH(\lambda)Q : P \in \mathbb{C}^{m \times m}, Q \in \mathbb{C}^{n \times n}, P, Q \text{ invertibles}\}.$$

Estas órbitas son variedades en el espacio vectorial \mathbb{C}^{2mn} y, por tanto, nos referiremos a la **codimensión** de $\mathcal{O}(H)$ como la codimensión en este espacio. Denotaremos por $\overline{\mathcal{O}}(H)$ al cierre de esta órbita.

El elemento más significativo de la órbita $\mathcal{O}(H)$ es la *forma canónica de Kronecker* (o FCK por brevedad) de $H(\lambda)$ (véase la sección 2.3). La FCK de $H(\lambda)$ determina unívocamente la órbita $\mathcal{O}(H)$, y, en particular, determina completamente la codimensión de $\mathcal{O}(H)$ [11, Th. 2.2].

7.2.2. Particiones enteras

El **orden de dominancia** en el conjunto de sucesiones de enteros no negativos especifica que $(a_1, a_2, \dots) \geq (b_1, b_2, \dots)$ si $a_1 + \dots + a_i \geq b_1 + \dots + b_i$ para $i = 1, 2, \dots$. Diremos que $(a_1, a_2, \dots) > (b_1, b_2, \dots)$ si $(a_1, a_2, \dots) \geq (b_1, b_2, \dots)$ y $(a_1, a_2, \dots) \neq (b_1, b_2, \dots)$ [22, Section 2.1]. Una **partición**, \mathbf{n} , de un entero n es una sucesión de enteros $\mathbf{n} = (n_1, n_2, \dots, n_c)$ con $n_1 \geq n_2 \geq \dots \geq n_c \geq 0$ y tal que $n_1 + n_2 + \dots + n_c = n$. Eventualmente, algunas de las entradas n_i pueden ser cero. El número c se conoce como la **longitud** de la partición. La **partición conjugada** de $\mathbf{n} = (n_1, n_2, \dots)$ es la sucesión $\mathbf{n}^* = (n_0^*, n_1^*, \dots)$, donde n_i^* es el número de entradas en \mathbf{n} que son mayores o iguales que i , para $i = 0, 1, \dots$. Nótese que \mathbf{n}^* es una partición de $n + c$, donde c es la longitud de \mathbf{n} . Es bien sabido que, si \mathbf{a} y

\mathbf{b} son particiones del mismo entero n con la misma longitud, entonces $\mathbf{a} \geq \mathbf{b}$ si y solo si $\mathbf{a}^* \leq \mathbf{b}^*$.

En las secciones restantes de este capítulo usaremos los siguientes dos lemas. Como es habitual, $\lfloor x \rfloor$ denotará la *parte entera* de x , es decir: el mayor entero que es menor o igual que x .

Lema 7.2.1 Sean n y c enteros no negativos y sea $n = c\alpha + t$, con $0 \leq t < c$, la división entera de n entre c , es decir, $\alpha = \lfloor n/c \rfloor$ y $t = n \bmod c$. Definamos $\mathbf{n}_c := (\alpha + 1, \dots, \alpha + 1, \alpha, \dots, \alpha)$, que es una partición de n de longitud c . Entonces

$$\mathbf{n}_c^* = (\underbrace{c, \dots, c}_{\alpha+1}, t, 0, \dots),$$

y, para toda partición \mathbf{n} de n de longitud c , $\mathbf{n} \geq \mathbf{n}_c$ o, equivalentemente, $\mathbf{n}^* \leq \mathbf{n}_c^*$.

Demostración. Es trivial que $\mathbf{n}_c^* = (c, \dots, c, t, 0, \dots)$. Por otro lado, dada una partición \mathbf{n} de n de longitud c , tenemos que $\mathbf{n}^* = (c, \dots, c, t', \dots)$, con $0 \leq t' < c$ y $\alpha' \geq 0$. Entonces, $n + c \geq (\alpha' + 1)c + t'$, lo que implica que $\alpha \geq \alpha'$. El caso $\alpha = \alpha'$ implica que $t \geq t'$, luego, en cualquier caso, $\mathbf{n}_c^* \geq \mathbf{n}^*$. ■

Lema 7.2.2 Si $n > m$, entonces $\mathbf{n}_c^* > \mathbf{m}_c^*$ y $\mathbf{n}_c > \mathbf{m}_c$, para cualquier entero no negativo c .

La demostración de este Lema es inmediata.

7.2.3. Relaciones de inclusión entre cierres de órbitas

Para cada haz matricial, $H(\lambda)$, de rango r , consideramos las siguientes tres sucesiones definidas en [22]:

$$\mathcal{R}(H) + r = (r_0 + r, r_1 + r, r_2 + r, \dots),$$

donde r_i es el número de bloques singulares derechos L_j en la FCK de $H(\lambda)$ con $j \geq i$;

$$\mathcal{L}(H) + r = (l_0 + r, l_1 + r, l_2 + r, \dots),$$

donde l_i es el número de bloques singulares izquierdos L_j^T en la FCK de $H(\lambda)$ con $j \geq i$; y, para cada $\mu \in \mathbb{C} \cup \{\infty\}$,

$$\mathcal{J}_\mu(H) + p = (w_1(\mu) + p, w_2(\mu) + p, \dots),$$

donde $w_i(\mu)$ es el número de bloques de Jordan asociados con el autovalor μ de dimensión mayor o igual que i en la estructura regular de la FCK de $H(\lambda)$, y p es el número de bloques singulares derechos en la FCK de $H(\lambda)$. Estas sucesiones nos permiten obtener relaciones de inclusión entre los cierres de las órbitas de dos haces matriciales diferentes. Esto es lo que se presenta en el Teorema 7.2.3 siguiente, obtenido en [55], y reformulado más tarde en [8] y [22]. Nosotros enunciamos el teorema tal y como aparece en [22].

Teorema 7.2.3 [22, Th. 3.1] Sean H_1 y H_2 dos haces matriciales complejos $m \times n$ con, respectivamente, $p(H_1)$ y $p(H_2)$ bloques singulares derechos en su FCK. Entonces $\overline{\mathcal{O}}(H_1) \supseteq \overline{\mathcal{O}}(H_2)$ si y sólo si se dan las siguientes relaciones:

$$(i) \quad \mathcal{R}(H_1) + \text{rg}(H_1) \geq \mathcal{R}(H_2) + \text{rg}(H_2),$$

$$(ii) \quad \mathcal{L}(H_1) + \text{rg}(H_1) \geq \mathcal{L}(H_2) + \text{rg}(H_2),$$

$$(iii) \quad \mathcal{J}_\mu(H_1) + p(H_1) \leq \mathcal{J}_\mu(H_2) + p(H_2),$$

para todo $\mu \in \mathbb{C} \cup \{\infty\}$.

7.3. El conjunto de haces singulares de rango menor o igual que r

Si nos restringimos al conjunto de haces de matrices $m \times n$ con rango fijo igual a r , las condiciones del Teorema 7.2.3 se simplifican significativamente. En este caso, $\overline{\mathcal{O}}(H_1) \supseteq \overline{\mathcal{O}}(H_2)$ si y sólo si se dan las siguientes condiciones: (i) $\mathcal{R}(H_1) \geq \mathcal{R}(H_2)$; (ii) $\mathcal{L}(H_1) \geq \mathcal{L}(H_2)$, y; (iii) $\mathcal{J}_\mu(H_1) \leq \mathcal{J}_\mu(H_2)$, para todo $\mu \in \mathbb{C} \cup \{\infty\}$ (ya que $p(H) = n - \text{rg}(H)$). Nuestro primer resultado caracteriza las órbitas maximales de haces de rango r .

Lema 7.3.1 *Sea r un número entero tal que $1 \leq r \leq \min\{m, n\} - 1$. Definamos, en el conjunto de haces matriciales complejos $m \times n$ de rango r , las siguientes $r + 1$ formas canónicas de Kronecker:*

$$\mathcal{K}_a(\lambda) = \text{diag} \left(\underbrace{L_{\alpha+1}, \dots, L_{\alpha+1}}_s, \underbrace{L_\alpha, \dots, L_\alpha}_{n-r-s}, \underbrace{L_{\beta+1}^T, \dots, L_{\beta+1}^T}_t, \underbrace{L_\beta^T, \dots, L_\beta^T}_{m-r-t} \right), \quad (7.1)$$

para $a = 0, 1, \dots, r$, donde $\alpha = \lfloor a/(n-r) \rfloor$, $s = a \bmod (n-r)$, $\beta = \lfloor (r-a)/(m-r) \rfloor$ y $t = (r-a) \bmod (m-r)$. Entonces, para todo haz $m \times n$, $H(\lambda)$, de rango r existe un número “ a ” tal que $\overline{\mathcal{O}}(\mathcal{K}_a) \supseteq \overline{\mathcal{O}}(H)$, y $\overline{\mathcal{O}}(\mathcal{K}_a) \not\supseteq \overline{\mathcal{O}}(\mathcal{K}_{a'})$ siempre que $a \neq a'$.

Demostración. Debemos probar que, para cada FCK de tamaño $m \times n$ y rango r , $\tilde{\mathcal{K}}(\lambda)$, existe un número “ a ” tal que $\overline{\mathcal{O}}(\mathcal{K}_a) \supseteq \overline{\mathcal{O}}(\tilde{\mathcal{K}})$, y también que $\overline{\mathcal{O}}(\mathcal{K}_a) \not\supseteq \overline{\mathcal{O}}(\mathcal{K}_{a'})$ siempre que $a \neq a'$.

Cualquier estructura de Kronecker de tamaño $m \times n$ y rango r es de la forma

$$\tilde{\mathcal{K}}(\lambda) = \text{diag} (L_{\alpha_1}, \dots, L_{\alpha_{n-r}}, L_{\beta_1}^T, \dots, L_{\beta_{m-r}}^T, J),$$

donde $\alpha_1 \geq \dots \geq \alpha_{n-r} \geq 0$, $\beta_1 \geq \dots \geq \beta_{m-r} \geq 0$, y J es la estructura regular, que suponemos de tamaño $k \times k$. Sea $R = \alpha_1 + \dots + \alpha_{n-r}$ el número de filas ocupadas por los bloques singulares derechos $L_{\alpha_1}, \dots, L_{\alpha_{n-r}}$, y sea $a = R + k$. Nótese que $0 \leq a \leq r$, luego podemos considerar, para este valor de a , la estructura de Kronecker $\mathcal{K}_a(\lambda)$ dada en (7.1). Demostraremos que $\mathcal{R}(\mathcal{K}_a) \geq \mathcal{R}(\tilde{\mathcal{K}})$, $\mathcal{L}(\mathcal{K}_a) \geq \mathcal{L}(\tilde{\mathcal{K}})$, y $\mathcal{J}_\mu(\mathcal{K}_a) \leq \mathcal{J}_\mu(\tilde{\mathcal{K}})$, para todo $\mu \in \mathbb{C} \cup \{\infty\}$. De acuerdo con el Teorema 7.2.3, esto implicará que $\overline{\mathcal{O}}(\mathcal{K}_a) \supseteq \overline{\mathcal{O}}(\tilde{\mathcal{K}})$, para $a = R + k$.

Teniendo en cuenta que $\mathcal{J}_\mu(\mathcal{K}_a) = (0, 0, \dots)$, para todos μ y a , la relación $\mathcal{J}_\mu(\mathcal{K}_a) \leq \mathcal{J}_\mu(\tilde{\mathcal{K}})$ se obtiene trivialmente.

Ahora, para demostrar $\mathcal{R}(\mathcal{K}_a) \geq \mathcal{R}(\tilde{\mathcal{K}})$, nótese que $\mathcal{R}(\tilde{\mathcal{K}})$ es la partición conjugada de $(\alpha_1, \dots, \alpha_{n-r})$, que es una partición de R de longitud $n-r$. Por lo tanto, si $(n-r)\tilde{\alpha} + \tilde{s} = R$ es la división entera de R entre $n-r$, el Lema 7.2.1 implica que

$$(n-r, \tilde{\alpha}^{+1}, n-r, \tilde{s}, 0, \dots) \geq \mathcal{R}(\tilde{\mathcal{K}}). \quad (7.2)$$

Por otro lado, $\mathcal{R}(\mathcal{K}_a) = (n-r, \alpha^{+1}, n-r, s, 0, \dots)$, es la partición conjugada de $(\alpha+1, \dots, \alpha+1, \alpha, \dots, \alpha, n-r-s, \alpha)$, que es una partición de $a = R + k$ de longitud $n-r$. Finalmente, el Lema 7.2.2, junto con (7.2), implica

$$\mathcal{R}(\mathcal{K}_a) \geq (n-r, \tilde{\alpha}^{+1}, n-r, \tilde{s}, 0, \dots) \geq \mathcal{R}(\tilde{\mathcal{K}}).$$

Análogamente, para demostrar $\mathcal{L}(\mathcal{K}_a) \geq \mathcal{L}(\tilde{\mathcal{K}})$, nótese que $\mathcal{L}(\tilde{\mathcal{K}})$ es la partición conjugada de $(\beta_1, \dots, \beta_{m-r})$, que es una partición de $r-a$ de longitud $m-r$, donde $a = R + k$.

También, $\mathcal{L}(\mathcal{K}_a)$ es la partición conjugada de $(\beta + 1, \cdot^t, \beta + 1, \beta, \overset{m-r-t}{\cdot}, \beta)$, que es una partición de $r - a$ de longitud $m - r$. Así, la relación $\mathcal{L}(\mathcal{K}_a) \geq \mathcal{L}(\tilde{\mathcal{K}})$ es una consecuencia inmediata del Lema 7.2.1.

Finalmente, hemos de probar que $\overline{\mathcal{O}}(\mathcal{K}_a) \not\supseteq \overline{\mathcal{O}}(\mathcal{K}_{a'})$ siempre que $a \neq a'$. Para esto, es suficiente comprobar que para dos valores $a, a' \in \{0, 1, \dots, r\}$ distintos la versión simplificada de las tres condiciones (i), (ii), y (iii) del Teorema 7.2.3 no se satisfacen simultáneamente. Este hecho es inmediato a partir del Lema 7.2.2, teniendo en cuenta que $\mathcal{R}(\mathcal{K}_a) = \mathbf{a}_{n-r}^*$ y $\mathcal{L}(\mathcal{K}_a) = (\mathbf{r} - \mathbf{a})_{m-r}^*$. Así, $a > a'$ implica $\mathcal{R}(\mathcal{K}_a) > \mathcal{R}(\mathcal{K}_{a'})$ y $\mathcal{L}(\mathcal{K}_a) < \mathcal{L}(\mathcal{K}_{a'})$. ■

Ejemplo 7.3.2 Consideremos el caso cuadrado $m = n = 6$ y $r = 4$. En la notación del teorema anterior, tenemos las siguientes cinco formas canónicas de Kronecker maximales

$$\begin{aligned} a = 0 &\rightarrow \mathcal{K}_0 = L_0 \oplus L_0 \oplus L_2^T \oplus L_2^T \\ a = 1 &\rightarrow \mathcal{K}_1 = L_0 \oplus L_1 \oplus L_1^T \oplus L_2^T \\ a = 2 &\rightarrow \mathcal{K}_2 = L_1 \oplus L_1 \oplus L_1^T \oplus L_1^T \\ a = 3 &\rightarrow \mathcal{K}_3 = L_1 \oplus L_2 \oplus L_0^T \oplus L_1^T \\ a = 4 &\rightarrow \mathcal{K}_4 = L_2 \oplus L_2 \oplus L_0^T \oplus L_0^T. \end{aligned}$$

□

El Lema 7.3.3 relaciona las formas canónicas de Kronecker que aparecen en (7.1) para rangos diferentes.

Lema 7.3.3 Sean $1 \leq r' < r \leq \min\{m, n\} - 1$ y $0 \leq a \leq r'$. Sean $\mathcal{K}_a(\lambda)$ y $\mathcal{K}'_a(\lambda)$ respectivamente, las formas canónicas de Kronecker dadas en (7.1) para los rangos r y r' . Entonces $\overline{\mathcal{O}}(\mathcal{K}_a) \supseteq \overline{\mathcal{O}}(\mathcal{K}'_a)$.

Demostración. Sean $(n - r')\alpha' + s' = a$ y $(m - r')\beta' + t' = r' - a$ las divisiones enteras de, respectivamente, a entre $n - r'$ y $r' - a$ entre $m - r'$. Por tanto, tenemos

$$\begin{aligned} \mathcal{R}(\mathcal{K}'_a) + \text{rg}(\mathcal{K}'_a) &= (n, \overset{\alpha'+1}{\cdot}, n, s' + r', 0, 0, \dots), \\ \mathcal{L}(\mathcal{K}'_a) + \text{rg}(\mathcal{K}'_a) &= (m, \overset{\beta'+1}{\cdot}, m, t' + r', 0, 0, \dots), \\ \mathcal{J}_\mu(\mathcal{K}'_a) + p(\mathcal{K}'_a) &= (n - r', n - r', \dots). \end{aligned}$$

Análogamente, para $\mathcal{K}_a(\lambda)$ y con la notación de (7.1), se tiene

$$\begin{aligned} \mathcal{R}(\mathcal{K}_a) + \text{rg}(\mathcal{K}_a) &= (n, \overset{\alpha+1}{\cdot}, n, s + r, 0, 0, \dots), \\ \mathcal{L}(\mathcal{K}_a) + \text{rg}(\mathcal{K}_a) &= (m, \overset{\beta+1}{\cdot}, m, t + r, 0, 0, \dots), \\ \mathcal{J}_\mu(\mathcal{K}_a) + p(\mathcal{K}_a) &= (n - r, n - r, \dots). \end{aligned}$$

Es obvio que la relación $\mathcal{J}_\mu(\mathcal{K}_a) + p(\mathcal{K}_a) \leq \mathcal{J}_\mu(\mathcal{K}'_a) + p(\mathcal{K}'_a)$ se da para todo $\mu \in \mathbb{C} \cup \{\infty\}$. Por otro lado, tenemos que

$$(n - r)\alpha' \leq (n - r')\alpha' \leq (n - r')\alpha' + s' = (n - r)\alpha + s < (n - r)(\alpha + 1),$$

y

$$(m - r')\beta' \leq (m - r')\beta' + t' < (m - r)\beta + t < (m - r)(\beta + 1) < (m - r')(\beta + 1),$$

luego $\alpha \geq \alpha'$ y $\beta \geq \beta'$, y las igualdades implican, respectivamente, $s \geq s'$ y $t \geq t'$. De aquí se sigue que $\mathcal{R}(\mathcal{K}_a) + \text{rg}(\mathcal{K}_a) \geq \mathcal{R}(\mathcal{K}'_a) + \text{rg}(\mathcal{K}'_a)$ y $\mathcal{L}(\mathcal{K}_a) + \text{rg}(\mathcal{K}_a) \geq \mathcal{L}(\mathcal{K}'_a) + \text{rg}(\mathcal{K}'_a)$.

■

Nótese que, aunque el rango de las formas canónicas de Kronecker dadas en (7.1) es exactamente r , los cierres de sus órbitas incluyen al conjunto de todos los haces con rango menor o igual que r , por el Lema 7.3.3. El Teorema 7.3.4 establece este hecho, y este es el resultado principal de este Capítulo. Nos permite decir con propiedad que los haces genéricos de rango menor o igual que r tienen estructuras de Kronecker dadas por (7.1).

Teorema 7.3.4 *Sea r un número entero tal que $1 \leq r \leq \min\{m, n\} - 1$, y sean $\mathcal{K}_a(\lambda)$, para $a = 0, 1, \dots, r$, las $r + 1$ formas canónicas de Kronecker definidas en (7.1). Entonces*

- (i) *El conjunto de haces matriciales complejos $m \times n$ con rango menor o igual que r es un conjunto cerrado igual a $\bigcup_{0 \leq a \leq r} \overline{\mathcal{O}(\mathcal{K}_a)}$.*
- (ii) *Para cada haz de tamaño $m \times n$, $H(\lambda)$, con rango menor o igual que r existe un número “ a ” tal que $\overline{\mathcal{O}(\mathcal{K}_a)} \supseteq \overline{\mathcal{O}(H)}$.*
- (iii) *$\overline{\mathcal{O}(\mathcal{K}_a)} \not\supseteq \overline{\mathcal{O}(\mathcal{K}_{a'})}$ siempre que $a \neq a'$.*

Demostración. Se sigue directamente de los Lemas 7.3.1 y 7.3.3. ■

A continuación, fijamos nuestra atención en la codimensión de las órbitas $\mathcal{O}(\mathcal{K}_a)$ de las formas de Kronecker genéricas con rango menor o igual que r . Veremos que estas codimensiones son distintas si $m \neq n$. En este caso, la codimensión (resp. dimensión) del conjunto de haces matriciales de rango menor o igual que r se define, de acuerdo con [91], como la menor (resp. mayor) de las codimensiones (resp. dimensiones) de $\mathcal{O}(\mathcal{K}_a)$, para $a = 0, 1, \dots, r$.

Teorema 7.3.5 *Sea r un número entero tal que $1 \leq r \leq \min\{m, n\} - 1$, y sean $\mathcal{K}_a(\lambda)$, para $a = 0, 1, \dots, r$, las $r + 1$ formas canónicas de Kronecker definidas en (7.1). Entonces*

1. *La codimensión de $\mathcal{O}(\mathcal{K}_a)$ es $(n - r)(2m - r) + a(m - n)$.*
2. *La codimensión del conjunto de haces matriciales complejos de tamaño $m \times n$ con rango menor o igual que r es igual a*
 - (i) $(n - r)(2m - r)$ si $m \geq n$,
 - (ii) $(m - r)(2n - r)$ si $m \leq n$.

Demostración. El primer apartado es una consecuencia directa de [11, Teorema 2.2]. El segundo apartado se sigue de calcular $\min_a \{(n - r)(2m - r) + a(m - n)\}$. ■

Ejemplo 7.3.6 En [23] se estudia en profundidad la estratificación del conjunto de órbitas de haces singulares de tamaño 2×3 . En este ejemplo, analizamos la jerarquía de los cierres de las órbitas del conjunto de haces singulares 3×3 . Usando el Teorema 7.2.3, podemos obtener el diagrama de Hasse completo asociado a esta jerarquía:

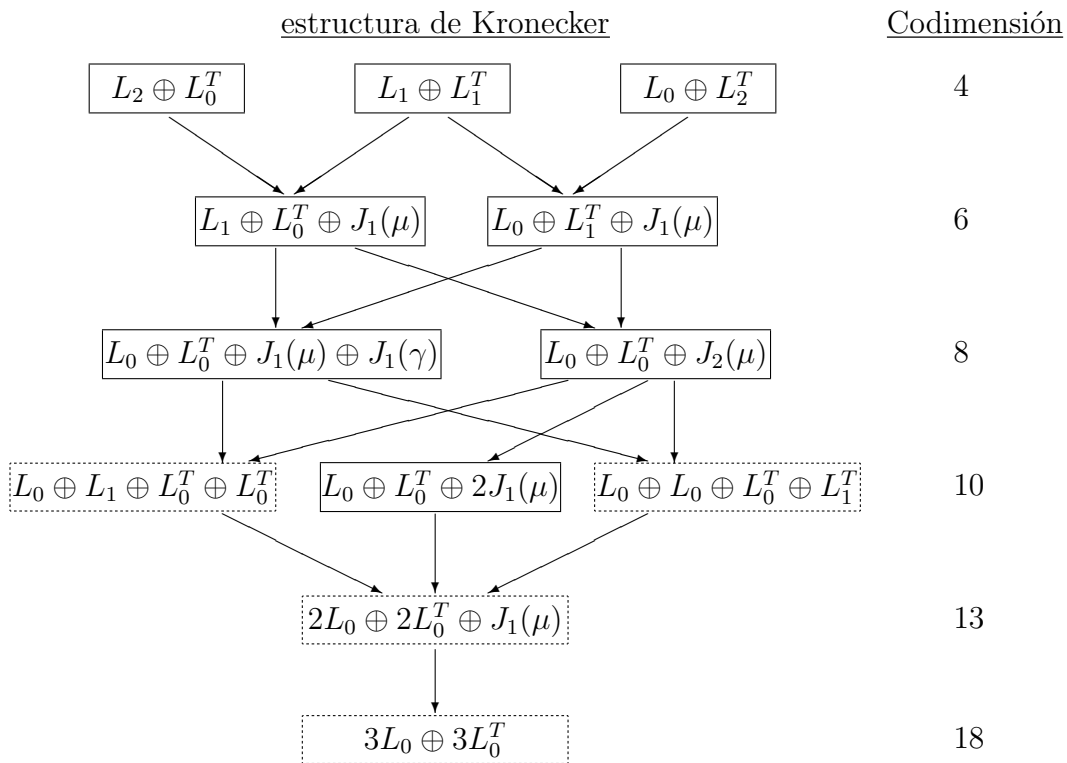


FIG. 1. Jerarquía de los cierres de órbitas del conjunto de haces singulares 3×3

Como se observa, hay tres estructuras de Kronecker maximales de codimensión 4, cuyos cierres contienen a las restantes estructuras.

Por otro lado, los recuadros en trazo discontinuo de la Fig. 1 corresponden a las estructuras de Kronecker de rango menor o igual que 1. Esto significa que la jerarquía de los cierres en el subconjunto de los haces 3×3 con rango menor o igual que 1 se reduce a tres estructuras de Kronecker no nulas, dos de las cuales tienen codimensión 10 (maximales) y una tiene codimensión 13. \square

Actualmente existe un software, bautizado por sus creadores como *Stratigraph*, que permite obtener el diagrama completo de la estratificación de los cierres de las órbitas de cualquier conjunto de haces matriciales (también de matrices) con dimensiones relativamente bajas [70]. El ejemplo anterior puede obtenerse con dicho programa.

7.3.1. Irreducibilidad en la topología de Zariski

Todas las ideas topológicas utilizadas hasta el momento se refieren a la topología usual en \mathbb{C}^{2mn} . La topología de Zariski ha sido utilizada en [91] para demostrar que el conjunto de haces matriciales singulares $n \times n$ con entradas en un cuerpo infinito arbitrario tiene exactamente n componentes irreducibles, cada una de codimensión $n + 1$. En esta subsección demostraremos que los cierres $\mathcal{O}(\mathcal{K}_a(\lambda))$ de las formas canónicas de Kronecker que aparecen en (7.1) son irreducibles en la topología de Zariski. Un resumen claro y conciso acerca de la topología de Zariski se encuentra en la Introducción de [91]. Aquí simplemente recordamos las siguientes ideas: (i) un subconjunto de \mathbb{C}^q es cerrado en la topología de Zariski si es el conjunto de ceros comunes a uno o varios polinomios; (ii) los cerrados de Zariski son cerrados en el sentido usual pero el recíproco no es cierto; (iii)

un subconjunto de \mathbb{C}^q es irreducible si no es la unión de dos subconjuntos propios cerrados en la topología de Zariski; (iv) todo cerrado en la topología de Zariski es una unión finita de subconjuntos irreducibles que se conocen como sus *componentes irreducibles*.

Un resultado importante en este contexto es que los cierres de una órbita de haces estrictamente equivalentes son los mismos tanto en la topología de Zariski como en la topología usual de \mathbb{C}^{2mn} [33]. Por lo tanto, no hay ambigüedad en el hecho de usar el símbolo $\overline{\mathcal{O}(\mathcal{K}_a)}$ en esta subsección porque se refiere exactamente al mismo conjunto que en el resto de la memoria. El resultado principal en este apartado establece que las órbitas de haces matriciales son irreducibles.

Lema 7.3.7 *El cierre $\overline{\mathcal{O}(H)}$ de la órbita $\mathcal{O}(H)$ de un haz matricial complejo $H(\lambda)$ de tamaño $m \times n$ es una variedad irreducible en la topología de Zariski.*

Demostración. Identifiquemos el conjunto de haces matriciales de tamaño $m \times n$ con \mathbb{C}^{2mn} , donde el haz $H(\lambda) = A + \lambda B$ se identifica con el par (A, B) . Sea U el conjunto de pares (P, Q) con $P \in \mathbb{C}^{m \times m}$ y $Q \in \mathbb{C}^{n \times n}$ invertibles. Se trata de un conjunto denso de $\mathbb{C}^{m^2+n^2}$ (esto es: $\overline{U} = \mathbb{C}^{m^2+n^2}$). Dado un haz matricial (A, B) , podemos considerar la función continua (polinómica), φ_H , de $\mathbb{C}^{m^2+n^2}$ a \mathbb{C}^{2mn} que envía el par (P, Q) al par (PAQ, PBQ) . Tenemos que $\mathcal{O}(H) = \varphi_H(U)$. Por [91, Section 1] sabemos que $\overline{\varphi_H(\mathbb{C}^{m^2+n^2})}$ es un conjunto irreducible. Por otro lado, para toda función continua φ , se tiene que $\overline{\varphi(W)} \subset \overline{\varphi(W)}$, donde W es un conjunto arbitrario, y esto implica que $\overline{\varphi(W)} = \overline{\varphi(W)}$. En nuestro caso, tenemos

$$\overline{\varphi_H(U)} = \overline{\varphi_H(\mathbb{C}^{m^2+n^2})},$$

y esto es igual a $\overline{\mathcal{O}(H)}$, lo que concluye la demostración. ■

Con este Lema, el Teorema 7.3.4 puede ser complementado como sigue.

Teorema 7.3.8 *Sea r un número entero tal que $1 \leq r \leq \min\{m, n\} - 1$, y sean $\mathcal{K}_a(\lambda)$, para $a = 0, 1, \dots, r$, las $r+1$ formas canónicas de Kronecker definidas en (7.1). Entonces, el conjunto de haces matriciales complejos $m \times n$ con rango menor o igual que r es un conjunto cerrado que tiene exactamente $r+1$ componentes irreducibles en la topología de Zariski. Estas componentes irreducibles son $\overline{\mathcal{O}(\mathcal{K}_a)}$, para $a = 0, 1, \dots, r$.*

Este Teorema incluye a [91, Th. 1] como un caso particular.

7.4. Notas históricas

Aunque los conceptos de órbita y fibrado proceden de otras áreas de la matemática (concretamente del álgebra y la geometría diferencial) su aplicación explícita en la teoría del análisis matricial fue introducida por Arnold en 1971 [2]. El término *fibrado*, que nosotros no hemos empleado en esta memoria, hace referencia al conjunto de matrices (haces matriciales) cuya forma de Jordan (de Weierstrass en el caso de haces regulares o de Kronecker en el de singulares) difiere únicamente en los autovalores. Tanto las órbitas como los fibrados son invariantes por un tipo de relación. En el caso de matrices esta relación es la relación de *semejanza*. En términos de teoría de grupos, se trata de las órbitas por la acción de conjugación del grupo lineal. En cambio, en el caso de haces matriciales las órbitas son invariantes por equivalencia estricta. Arnold, en el artículo citado [2] obtuvo la (co)dimensión tanto de las órbitas como de los fibrados de matrices.

Waterhouse [91], por otra parte, sin utilizar las ideas de Arnold, describió el conjunto de haces singulares de tamaño $n \times n$ como una variedad cerrada formada por exactamente n componentes irreducibles de codimensión $n + 1$, e identificó las formas de Kronecker correspondientes a las órbitas que se obtienen por equivalencia estricta. Un año antes, Markus y Parilis [49], sin utilizar tampoco la terminología de Arnold, dieron condiciones necesarias y suficientes para que la órbita (por semejanza) de una matriz se encuentre en el cierre de otra. Es decir: caracterizaron la relación de orden del cierre de órbitas de matrices. Esta caracterización es enunciada de forma muy elemental mediante la relación de dominancia de particiones enteras, y servirá también para caracterizar la relación de orden en el caso de haces matriciales. Dicha generalización (que corresponde al Teorema 7.2.3 anterior) fue obtenida por Pokrzywa [55] tres años más tarde, introduciendo ya la terminología de cierres de órbitas. En otras palabras, lo que ambos trabajos contienen es la caracterización de la relación de pertenencia de una matriz (resp. un haz) al cierre de la órbita de otra matriz (resp. haz) en términos de propiedades intrínsecas de dichas matrices (resp. haces). Previamente, se habían obtenido caracterizaciones no intrínsecas para el caso de operadores finitos (matrices cuadradas) [3] y operadores en espacios de Hilbert [1]. El trabajo de De Hoyos [8] puso de manifiesto la relación entre el resultado de Pokrzywa y los invariantes de los haces matriciales que caracterizan la forma canónica de Kronecker, esto es: la *característica de Segre* y de los *índices minimales por filas y columnas*. Esto permitió a su autora, además, enunciar los resultados de una forma más simplificada a través de las particiones de dichos invariantes. En el trabajo citado se estudia la continuidad de la forma de Kronecker como función de \mathbb{C}^{2mn} en \mathbb{C}^{2mn} y, como era de esperar (para aquellos pocos que conocieran por aquel entonces los trabajos que citaremos en el próximo párrafo), los únicos puntos en los que dicha función es continua corresponden a haces matriciales cuya forma de Kronecker es genérica [8, Teoremas 2.7, 2.8 y 2.9]. La introducción a la tesis de De Hoyos [9] contiene gran parte de la reseña histórica incluida en este párrafo y ampliada con información sobre algunos temas relacionados.

En 1995, Demmel y Edelman [11] calculan la (co)dimensión de las órbitas de haces $m \times n$ (incluyendo el caso cuadrado $m = n$). Como consecuencia, obtienen la forma de Kronecker genérica (entendiendo como tal aquella cuya codimensión es cero) para haces rectangulares, que ya era conocida, aunque no del todo bien difundida, desde hacía más de una década (véase [82, 3.55]).

Finalmente, los artículos de Edelman, Elmroth y Kågström [21] y [22] recogen todo el trabajo anteriormente descrito y lo unifican en una presentación clarividente, cuya primera parte, de carácter eminentemente geométrico, está muy en la línea del trabajo de Arnold. Gran parte de su interés se centra en la aplicación de esta teoría geométrica a la interpretación de los resultados que ofrecen los algoritmos del cálculo de la forma canónica de Kronecker y a la mejora de los fallos debidos a la presencia de estructuras dominantes próximas. En estos artículos se introduce y caracteriza la relación de *cubrimiento*, es decir, la que satisfacen dos estructuras entre las que no hay ninguna intermedia en el orden de los cierres de las órbitas (y los fibrados). Esto permite dibujar el diagrama de Hasse completo de dicha relación de orden. Sin duda, estos trabajos han permitido la difusión entre la comunidad matemática de todos los resultados que hemos mencionado y suponen una referencia inevitable para todo aquel que quiera entender la geometría subyacente en el problema del cálculo de las formas canónicas que tratamos en esta tesis. Recomendamos, además, su lectura a todos los interesados en ampliar esta breve introducción histórica.

En los últimos años, varios profesores de la Universidad de Umeå han desarrollado un software que permite visualizar la jerarquía completa de la estratificación de órbitas y

fibrados tanto de matrices como de haces matriciales para una gran variedad de dimensiones no demasiado elevadas. Esta herramienta, denominada *Stratigraph* está disponible en la dirección que aparece en [70], donde también puede encontrarse una guía para su utilización. Una información más completa acerca de ella se encuentra en [36].

Capítulo 8

Conclusiones, publicaciones y problemas abiertos

A modo de resumen, incluimos en este Capítulo los resultados fundamentales que se han presentado en esta memoria, organizados por Capítulos. Indicamos, asimismo, los trabajos a que han dado lugar.

Capítulo 4: Hemos desarrollado, por vez primera, una teoría de perturbación de primer orden para autovalores de haces matriciales singulares cuadrados en términos de la información espectral del haz de partida. Hemos demostrado que, genéricamente, en torno a un autovalor de un haz singular existe, como ocurre en el caso regular, un número de desarrollos igual a la multiplicidad algebraica de dicho autovalor, que se agrupan en “anillos”, uno por cada bloque de Jordan asociado al autovalor en la FCK del haz no perturbado. El exponente del término de primer orden en cada uno de estos anillos es igual a $1/n_j$, donde n_j es el tamaño del bloque de Jordan, y los coeficientes directores son las raíces n_j -ésimas de los autovalores de un haz que depende de la perturbación evaluada en el autovalor y de ciertos espacios reductores del haz no perturbado asociados a dicho autovalor. También hemos determinado el vector al que se aproximan, genéricamente, los desarrollos de los autovectores asociados a los anteriores desarrollos de los autovalores. El contenido de este capítulo es el trabajo “*First order spectral perturbation theory of square singular matrix pencils*” (con F. M. Dopico y Julio Moro), que ha sido enviado en 2007 y se encuentra en fase de revisión para su posible publicación en *Linear Algebra and its Applications*.

Capítulo 5: Hemos demostrado que una perturbación genérica, $B_0 + \lambda B_1$, de rango bajo de un haz matricial regular produce la siguiente alteración en la estructura de Jordan asociada a un autovalor cualquiera, λ_0 , del haz en su forma canónica de Weierstrass:

- 1) Destruye los ρ_0 bloques de mayor tamaño,
- 2) convierte en bloques de tamaño 1×1 los siguientes ρ_1 bloques de mayor tamaño,
y
- 3) mantiene inalterados los restantes bloques,

donde $\rho_0 = \text{rg}(B_0 + \lambda_0 B_1)$ y $\rho_1 = \text{rg}(B_1)$. Demostramos, asimismo, que el conjunto de perturbaciones para las que no se da dicho cambio es una subvariedad alge-

braica propia del conjunto de haces de rango ρ que está definida por una ecuación polinómica que involucra las entradas de la perturbación. Estos resultados se incluyen en el trabajo “*Low rank perturbation of Weierstrass structure*” (con F. M. Dopico y Julio Moro), que ha sido enviado en 2005 y está en fase de revisión para su posible publicación en SIAM Journal of Matrix Analysis and Applications.

Capítulo 6: Hemos descrito el cambio genérico que experimenta la forma canónica de Kronecker de haces matriciales de rango no completo por perturbaciones de rango bajo (ρ) que hacen que el haz perturbado no tenga rango completo. Dicho cambio genérico puede resumirse en:

- 1) La estructura regular del haz perturbado es la suma directa de las estructuras regulares del haz no perturbado y el haz perturbación,
- 2) la perturbación reduce en ρ el número de bloques singulares derechos, y los tamaños de los nuevos bloques se obtienen a partir de los tamaños de los bloques antiguos, del rango, ρ , y del orden singular derecho de la perturbación,
- 3) la perturbación reduce en ρ el número de bloques singulares izquierdos, y los tamaños de los nuevos bloques se obtienen a partir de los tamaños de los bloques antiguos, del rango, ρ , y del orden singular izquierdo de la perturbación.

También describimos el cambio en la estructura de Kronecker para el caso límite en que el rango de la perturbación y el del haz de partida suman exactamente el rango máximo admisible (la menor de las dimensiones del haz). La nueva forma de Kronecker consta de:

- 1) La estructura regular del haz no perturbado,
- 2) una nueva estructura regular que se crea a partir de la estructura singular del haz no perturbado,
- 3) $n - m$ bloques singulares derechos (si $m \leq n$) o $m - n$ bloques singulares izquierdos (si $n \leq m$) cuyos tamaños se obtienen de igual forma que en el caso anterior (es decir, aquel en que el haz perturbado no es de rango completo).

Por último, damos condiciones necesarias sobre el conjunto de perturbaciones para que dicho cambio se produzca. Los resultados que conciernen al caso en que el rango del haz perturbado no es completo han dado lugar al artículo “*Low rank perturbation of Kronecker structures without full rank*” (con F. M. Dopico), que ha sido publicado en SIAM Journal of Matrix Analysis and Applications, 29, no. 2 (2007), pp. 496–529.

Capítulo 7: Hemos demostrado que el conjunto de haces matriciales singulares de rango menor o igual que r es una variedad que consiste en la unión de los cierres topológicos de $r + 1$ órbitas que son irreducibles en la topología de Zariski. Hemos obtenido las formas de Kronecker canónicas correspondientes a estas órbitas y su dimensión. Estos resultados se encuentran en el trabajo “*A note on generic Kronecker orbits of matrix pencils with fixed rank*” (con F. M. Dopico), que ha sido enviado en 2006 y está en fase de revisión para su posible publicación en SIAM Journal of Matrix Analysis and Applications.

Comentamos, a continuación, algunos de los problemas abiertos que pueden abordarse de manera más inmediata a partir de lo ya realizado:

1. **Extender la teoría de perturbaciones de rango bajo a polinomios matriciales de grado superior.** Una manera natural de abordar este problema para polinomios regulares es hacerlo a través de alguna linealización adecuada que conserve la estructura de Jordan del polinomio, como puede ser el primer haz compañero. El principal inconveniente que plantea este procedimiento es que las condiciones naturales de rango bajo que hay que imponer al polinomio perturbador no se traducen en las condiciones de rango bajo que se han utilizado en esta memoria para el caso de haces matriciales cuando se trasladan al haz compañero. Concretamente, una condición natural de rango bajo para el autovalor λ_0 en el polinomio perturbador $M(\lambda) = B_0 + \lambda B_1 + \dots + \lambda^l B_l$ es

$$\text{rg } M(\lambda_0) < g, \quad (8.1)$$

donde g es la multiplicidad geométrica de λ_0 en el polinomio no perturbado. En cambio, al linealizar el problema mediante el primer haz compañero obtenemos el haz perturbador asociado

$$\begin{bmatrix} 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \\ B_0 & B_1 & \dots & B_{l-1} \end{bmatrix} + \lambda \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & B_l \end{bmatrix},$$

para el que la condición de rango bajo habría de ser

$$\text{rg} [B_0 \ B_1 \ \dots \ B_{l-1} + \lambda_0 B_l] < g. \quad (8.2)$$

No obstante, se tiene la siguiente relación

$$\text{rg} [B_0 \ B_1 \ \dots \ B_{l-1} + \lambda_0 B_l] \geq \text{rg } M(\lambda_0),$$

luego la condición (8.2) implica la condición (8.1).

2. **Caracterizar todas las posibles formas de Kronecker que se pueden obtener al perturbar un haz dado mediante perturbaciones de rango bajo.** Hemos citado a lo largo de esta memoria una serie de trabajos en los que se han obtenido caracterizaciones de todas las posibles estructuras canónicas que pueden alcanzarse al perturbar una matriz o un haz por diferentes tipos de perturbaciones: perturbaciones de norma pequeña de matrices [49], o de haces [55]; perturbaciones de una fila o una columna de norma pequeña de matrices [5], perturbaciones de rango 1 de matrices sobre un Dominio de Ideales Principales (lo que incluye tanto el caso de matrices como el de haces con entradas complejas) en [76] o adición de una fila o columna en matrices sobre un Dominio de Ideales Principales [75] o en haces matriciales [6]. En todos estos trabajos las caracterizaciones están dadas por unas condiciones de entrelazamiento o mayorización sobre algunas de las sucesiones asociadas a las formas canónicas que han sido el objeto de esta memoria (las características de Segre o de Weyr y las sucesiones de índices minimales derechos e izquierdos). Esto nos lleva a creer que este tipo de condiciones son también las que caracterizan las formas canónicas de Kronecker que pueden alcanzarse por perturbaciones de rango bajo de haces matriciales. De hecho, estas condiciones ya están

descritas en [76] para la parte regular (la característica de Segre) cuando las perturbaciones son de rango 1. Resta, por tanto, caracterizar el cambio en la parte singular o, en otras palabras, el cambio de los índices minimales y comprobar si ambos comportamientos, al estudiarse por separado, permiten describir toda la estructura del haz perturbado o si, por el contrario, pueden aparecer nuevos elementos (como ocurre en lo que hemos llamado el “caso límite”). Un último paso consistiría en generalizar los resultados de rango 1 a rango bajo arbitrario.

3. **Estudiar el cambio genérico de la forma canónica de Kronecker en perturbaciones de rango bajo de haces singulares de rango completo.** El caso en que los haces son cuadrados y el haz no perturbado tiene rango completo corresponde al caso regular y ha sido estudiado en el Capítulo 5. El caso en que el haz de partida, $H(\lambda)$, tiene rango completo pero es rectangular es un problema abierto. En este caso, la FCK de $H(\lambda)$ tiene únicamente un tipo de bloques singulares: bloques singulares derechos si $m < n$, y bloques singulares izquierdos si $m > n$. Genéricamente, lo mismo ocurre con el haz perturbado $(H + M)(\lambda)$, pero las dimensiones de los bloques singulares pueden variar. Este problema es de gran importancia pues incluye los haces que aparecen con más frecuencia en Teoría de Control, es decir, los que corresponden a sistemas controlables. Una primera tarea fundamental en este contexto es definir el significado preciso del concepto *perturbación de rango bajo*.
4. **Estudiar el cambio genérico de la forma canónica de Kronecker de haces sin rango completo pero en las que el haz perturbado puede tener rango completo tras una perturbación de rango bajo.** Un segundo problema relacionado con el del punto anterior consiste en considerar un haz no perturbado, $H(\lambda)$, sin rango completo, pero perturbaciones cuyo rango, ρ , no satisface la desigualdad (6.1). Por ejemplo, si $H(\lambda)$ es un haz de tamaño 100×200 con $\text{rg}(H) = 98$ y con el rango de las perturbaciones igual a $\rho \equiv \text{rg}(M) = 3$, entonces las perturbaciones $M(\lambda)$ son, intuitivamente, perturbaciones de rango bajo de $H(\lambda)$. La solución de este tipo de problemas está conectada de forma natural con los resultados que se han presentado en el Capítulo 6 y con el problema propuesto en el punto anterior. En nuestro ejemplo previo, la descomposición derecha de $M(\lambda)$ dada en (6.3) nos permite escribir $M(\lambda) = M_1(\lambda) + M_2(\lambda)$, donde $\text{rg}(M_1) = 2$ y $\text{rg}(M_2) = 1$. Por tanto, podemos descomponer el problema de perturbación original, $H + M = H + M_1 + M_2$, en dos problemas de perturbación de rango menor, $H + M_1$ y $(H + M_1) + M_2$. El primero de ellos es del tipo que hemos considerado en el Capítulo 6, y, en el segundo de ellos, el haz no perturbado $H + M_1$ tiene genéricamente rango completo.
5. **Obtener condiciones necesarias y suficientes para que una matriz mosaico Toeplitz tenga rango máximo.** Como ya hemos mencionado en la Sección 6.5.5, el hecho de que las matrices mosaico Toeplitz que aparecen en la sección 6.5.1 y siguientes son, genéricamente, de rango completo no está rigurosamente demostrado. Este es, por lo tanto, y, hasta donde alcanza nuestro conocimiento, un problema abierto que puede abordarse con carácter más general desde el estudio de las matrices mosaico Toeplitz arbitrarias.
6. **Extender a polinomios matriciales singulares (cuadrados) de grado arbitrario los resultados de la teoría de primer orden de perturbaciones desarrollada en el Capítulo 4.** El procedimiento de regularizar el problema singular que se ha seguido para obtener los desarrollos de autovalores en el caso de

perturbaciones de haces matriciales singulares continúa siendo válido si los haces se sustituyen por polinomios de grado arbitrario. Esto significa que existirán desarrollos asintóticos en torno a los autovalores de un polinomio singular. En cambio, al intentar ofrecer una descripción de los términos de primer orden de dichos desarrollos similar a la que se ha dado para haces nos encontramos con el problema de que los dos elementos esenciales que nos han permitido dar esa descripción no son fácilmente generalizables a polinomios de grado superior. Por un lado, no existe una forma canónica equivalente a la FCK para polinomios de grado superior, es decir, no existe una forma canónica invariante por equivalencia estricta. Por otra parte, el concepto de subespacio reductor no admite una generalización elemental. Por ejemplo, una definición inmediata de par de subespacios reductores podría ser un par $(\mathcal{X}, \mathcal{Y})$ de modo que $\mathcal{Y} = A_0\mathcal{X} + \dots + A_\ell\mathcal{X}$ y $\dim \mathcal{Y} = \dim \mathcal{X} - \dim \mathcal{N}(L)$, siendo $L(\lambda) = A_0 + \lambda A_1 + \dots + \lambda^\ell A_\ell$, pero esta definición enseguida se muestra inapropiada. A título ilustrativo, considérese el siguiente ejemplo: sea el polinomio singular de segundo grado:

$$L(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2 = \begin{bmatrix} \lambda^2 & \lambda \\ \lambda & 1 \end{bmatrix}.$$

Se tiene que $\dim \mathcal{N}(L) = 1$, pero es inmediato comprobar que no existe ningún par de subespacios \mathcal{X}, \mathcal{Y} , de manera que $\mathcal{Y} = A_0\mathcal{X} + A_1\mathcal{X} + A_2\mathcal{X}$ y $\dim \mathcal{Y} = \dim \mathcal{X} - 1$.

Quizá este problema pueda abordarse utilizando una linealización adecuada (por ejemplo, el primer haz compañero) que permita definir la forma canónica de un polinomio de grado superior a través de la forma de Kronecker de la linealización. Otra opción posible es usar la forma local de Smith, como se ha hecho en el Capítulo 4 siguiendo a Langer y Najman [46]. Para autovalores simples es de esperar que el uso de los núcleos del polinomio evaluado en el autovalor sea suficiente.

7. **Extender los resultados de teoría de primer orden a perturbaciones no genéricas de haces matriciales singulares como las que se utilizan en los algoritmos que calculan las formas canónicas.** Como ya hemos mencionado, una perturbación genérica de un haz singular cuadrado convierte este haz en un haz regular. En cambio, para el desarrollo de algoritmos estables que permitan obtener la información espectral de un haz, como la forma GUPTRI [14, 15] es importante considerar perturbaciones que mantengan el haz singular. En este contexto, puede ser de utilidad analizar si es factible desarrollar una teoría de primer orden de perturbaciones no genéricas que se muevan dentro de la variedad concreta en la que se quiere trabajar.

Bibliografía

- [1] C. APOSTOL, D. A. HERRERO Y D. VOICULESCU, *The closure of the similarity orbit of a Hilbert space operator*, Bull. A. M. S., 6, no. 3 (1982), pp. 421–426.
- [2] V. I. ARNOLD, *On matrices depending on parameters*, Russian Math. Surveys, 26 (1971), pp. 29–43.
- [3] J. BARRÍA Y D. A. HERRERO, *Closure of similarity orbits of nilpotent operators I. Finite rank operators*, J. Operator Theory, 1 (1979), 177–185.
- [4] H. BAUMGÄRTEL, *Analytic Perturbation Theory for Matrices and Operators*, Birkhäuser, Basel (1985).
- [5] M. A. BEITIA, I. DE HOYOS E I. ZABALLA, *The change of the Jordan structure under one row perturbations*, Linear Algebra Appl., 401 (2005), pp 119–134.
- [6] D. L. BOLEY, *The algebraic structure of pencils and block Toeplitz matrices*, Linear Algebra Appl., 279 (1998), pp. 255–279.
- [7] K. W. E. CHU, *On Multiple Eigenvalues of Matrices Depending on Several Parameters*, SIAM J. Numer. Anal., 27 (1990), pp. 1368–1385.
- [8] I. DE HOYOS, *Points of continuity of the Kronecker canonical form*, SIAM J. Matrix Anal. Appl., 11, no. 2 (1990), pp. 278–300.
- [9] I. DE HOYOS, *Perturbación de matrices rectangulares y haces de matrices*, Tesis Doctoral, Universidad del País Vasco, Bilbao (1990).
- [10] J. W. DEMMEL, *Applied Numerical Linear Algebra*, SIAM, Philadelphia (1997).
- [11] J. W. DEMMEL Y A. EDELMAN, *The Dimension of Matrices (Matrix Pencils) with Given Jordan (Kronecker) Canonical Forms*, Linear Algebra Appl., 230 (1995), pp. 61–87.
- [12] J. W. DEMMEL Y B. KÅGSTRÖM, *Computing Stable Eigendecompositions of Matrix Pencils*, Linear Algebra Appl., 88-89 (1987), pp. 139–186.
- [13] J. W. DEMMEL Y B. KÅGSTRÖM, *Accurate solutions of ill-posed problems in Control Theory*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 126–145.
- [14] J. W. DEMMEL Y B. KÅGSTRÖM, *The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part I: Theory and algorithms*, ACM Trans. Math. Software, 19 (1993), pp. 160–174.

- [15] J. W. DEMMEL Y B. KÅGSTRÖM, *The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part II: Software and applications*, ACM Trans. Math. Software, 19 (1993), pp. 175–201.
- [16] F. DE TERÁN Y F. M. DOPICO, *Low rank perturbation of Kronecker structures without full rank*, SIAM J. Matrix Anal. Appl., 29, no. 2 (2007), pp. 496–529 (electrónico).
- [17] F. DE TERÁN Y F. M. DOPICO, *A note on generic Kronecker orbits of matrix pencils with fixed rank*, enviado en 2006.
- [18] F. DE TERÁN, F. M. DOPICO Y J. MORO, *Low rank perturbation of Weierstrass structure*, enviado en 2005.
- [19] F. DE TERÁN, F. M. DOPICO Y J. MORO, *First order spectral perturbation theory of square singular matrix pencils*, enviado en 2007.
- [20] F. DE TERÁN, F. M. DOPICO Y D. S. MACKEY, *Linearizations of singular matrix polynomials*, en preparación.
- [21] A. EDELMAN, E. ELMROTH Y B. KÅGSTRÖM, *A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations*, SIAM J. Matrix Anal. Appl., 18, no. 3, (1999), pp. 653–692.
- [22] A. EDELMAN, E. ELMROTH Y B. KÅGSTRÖM, *A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 667–699.
- [23] E. ELMROTH Y B. KÅGSTRÖM, *The Set of 2×3 Matrix Pencils – Kronecker Structures and Their Transition Under Perturbations*, SIAM J. Matrix Anal. Appl., 17, no. 1 (1996), pp. 1–34.
- [24] G. D. FORNEY, *Minimal bases of rational vector spaces, with applications to multi-variable linear systems*, SIAM J. Control, 13 (1975), pp. 493–520.
- [25] M. I. FRISWELL, U. PRELLS, Y S. D. GARVEY, *Low-rank damping modifications and defective systems*, J. Sound Vibration, 279 (2005), pp. 757–774.
- [26] F. R. GANTMACHER, *The theory of matrices*, Chelsea, New York (1959).
- [27] I. GOHBERG, M. A. KAASHOEK Y P. LANCASTER, *General theory of regular matrix polynomials and band Toeplitz operators*, Int. Equat. Oper. Th., 11 (1988), pp. 776–882.
- [28] I. GOHBERG, P. LANCASTER Y L. RODMAN, *Matrix Polynomials*, Academic Press (1982).
- [29] I. GOHBERG, P. LANCASTER Y L. RODMAN, *Invariant Subspaces of Matrices with Applications*, Wiley-Interscience, New York (1986).
- [30] N. J. HIGHAM, D. S. MACKEY Y F. TISSEUR, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28, no. 4, (2007), pp. 1005–1028.

- [31] N. J. HIGHAM, R.-C. LI Y F. TISSEUR, *Backward error of Polynomial Eigenproblems Solved by Linearization*, MIMS EPrint: 2006.137, Manchester Institute for Mathematical Sciences. School of Mathematics, The University of Manchester, UK (2007).
- [32] E. HILLE, *Analytic Function Theory*, vol. II, Chelsea (2a. ed.) (1987).
- [33] D. HINRICHSSEN Y J. O'HALLORAN, *Orbit closures of singular matrix pencils*, J. Pure Appl. Algebra, 81 (1992), pp. 117–137.
- [34] L. HÖRMANDER Y A. MELIN, *A remark on perturbations of compact operators*, Mathematica Scandinavica, 75, no. 2 (1994), pp. 255–262.
- [35] R. A. HORN Y C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge (1985).
- [36] P. JOHANSSON, *Software Tools for matrix canonical computations and web-based software library environments*, Ph.D. Thesis, Departament of Computing Science, Umeå University, (2006).
- [37] N. KARCANIAS, *Minimal bases of matrix pencils: algebraic Toeplitz structure and geometric properties*, Linear Algebra Appl., 205/206 (1994), pp. 831–868.
- [38] T. KATO, *Perturbation Theory for Linear Operators*, Springer, Berlin (1980).
- [39] L. KRONECKER, *Algebraische Reduction der Schaaren bilinearer formen*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin (1890), pp. 1225–1237.
- [40] P. LANCASTER, *On Eigenvalues of Matrices Dependent on a Parameter*, Numer. Math., 6 (1964), pp. 377–387.
- [41] P. LANCASTER, *Lambda-matrices and vibrating systems*, Pergamon Press, Oxford, 1966 (reeditado por Dover en 2002).
- [42] P. LANCASTER, A. S. MARKUS Y F. ZHOU, *Perturbation theory for analytic matrix functions: the semisimple case*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 606–626.
- [43] P. LANCASTER Y L. MAROULAS, *Inverse eigenvalue problems for damped vibrating systems*, J. Math Analysis and Applications, 123, (1987), pp. 238–261.
- [44] P. LANCASTER Y U. PRELLS *Inverse problems for damped vibrating systems*, J. Sound Vibration, 283 (2005), pp. 891–914.
- [45] H. LANGER Y B. NAJMAN, *Remarks on the perturbation of analytic matrix functions II*, Integr. Equat. Oper. Th., 12 (1989), pp. 392–407.
- [46] H. LANGER Y B. NAJMAN, *Remarks on the perturbation of analytic matrix functions III*, Integr. Equat. Oper. Th., 15 (1992), pp. 796–806.
- [47] V. B. LIDSKII, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Math. and Math. Phys., 1 (1965), pp. 73–85.

- [48] D. S. MACKEY, N. MACKEY, C. MEHL Y V. MEHRMANN, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.
- [49] A. S. MARKUS Y E. E. PARILIS, *The change of the Jordan structure of a matrix under small perturbations*, Linear Algebra Appl., 54 (1983), pp. 139–152.
- [50] V. MEHRMANN Y C. SHI, *Transformation of high order linear differential-algebraic systems to first order*, Numer. Algorithms, 42, no. 3-4 (2006), pp. 281–307.
- [51] J. MORO, J. V. BURKE Y M. L. OVERTON, *On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.
- [52] J. MORO Y F. M. DOPICO, *Low rank perturbation of Jordan structure*, SIAM J. Matrix Anal. Appl. 25, no. 2 (2003), pp. 495–506.
- [53] J. MORO Y F. M. DOPICO, *First order eigenvalue perturbation theory and the Newton diagram*, en *Applied Mathematics and Scientific Computing*, edited by Z. Drmač et al. Proceedings of the Second Conference on Applied Mathematics and Scientific Computing, dada del 4 al 9 de Junio de 2001 en Dubrovnik, Croacia, pp. 143–175, Kluwer Academic Publishers (2002).
- [54] J. E. MOTTERSHEAD Y G. LALLEMENT, *Vibration nodes, and the cancellation of poles and zeros by unit-rank modifications to structures*, J. Sound and Vibration, 222(5) (1999), pp. 833–851.
- [55] A. POKRZYWA, *On perturbations and the equivalence orbit of a matrix pencil*, Linear Algebra Appl., 82 (1986), pp. 99–121.
- [56] U. PRELLS, J. E. MOTTERSHEAD, Y M. I. FRISWELL, *On pole-zero placement by unit-rank modification*, Mechanical Systems and Signal Processing, 17 (2003), pp. 611–633.
- [57] LORD RAYLEIGH, *The Theory of Sound*, London (1894) (reeditado por Dover (1945)).
- [58] F. RELICH, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach, New York (1964).
- [59] H. H. ROSENBRock, *State-Space and Multivariable Theory*, Wiley, New York (1970).
- [60] S. V. SAVCHENKO, *Typical changes in the spectral properties under perturbations by a rank-one operator* (en Ruso), Mat. Zametki 74, no. 4 (2004), pp. 590-602; trad. al inglés en Math. Notes 74, no. 4 (2003), pp. 557–568.
- [61] S. V. SAVCHENKO, *On the change in the spectral properties of a matrix under perturbations of sufficiently low rank* (en Ruso), Funkt. Anal. i Ego Pril. 38, no. 1 (2004), pp. 85–88; trad. al inglés en Funct. Anal. Appl. 38, no. 1 (2004), pp. 69–71.
- [62] S. V. SAVCHENKO, *Laurent expansion for the determinant of the matrix of scalar resolvents* (en Ruso), Mat. Sb. 196, no. 5 (2005), pp. 121–144, trad. al inglés en Sb. Math. 196 (2005), no. 5–6, pp. 743–764.

- [63] S. V. SAVCHENKO, *On the change of the Weyr characteristic under low rank perturbations*, preprint.
- [64] E. SCHRÖDINGER, *Quantisierung als Eigenwertproblem III. Störungstheorie, mit Anwendung auf der Starkeffekt der Balmen-Linien*, Ann. Phys. 80 (1926), pp. 437–490.
- [65] E. D. SONTAG, *Mathematical Control Theory*, Springer, New York, 2a edición (1998)
- [66] G. W. STEWART, *On the Sensitivity of the Eigenvalue Problem $Ax = \lambda Bx$* , SIAM J. Numer. Anal., 9, no. 4 (1972), pp. 669–686.
- [67] G. W. STEWART, *Perturbation Theory for the Generalized Eigenvalue Problem*, en *Recent Advances in Numerical Analysis*, de Boor y Golub (eds.), Academic Press, New York (1978)
- [68] G. W. STEWART, *Perturbation theory for rectangular matrix pencils*, Linear Algebra Appl., 208/209 (1994), pp. 297–301.
- [69] G. W. STEWART Y J. SUN, *Matrix perturbation theory*, Academic Press, New York (1990).
- [70] STRATIGRAPH, disponible en: www.cs.umu.se/pedher/research/stratigraph.
- [71] J. G. SUN, *Orthogonal projections and the perturbation of the eigenvalues of singular pencils*, J. Comp. Math., 1 (1983), pp. 63–74.
- [72] J. G. SUN, *Eigenvalues and eigenvectors for a matrix dependent on several parameters*, J. Comp. Math., 3 (1985), pp. 351–364.
- [73] J. G. SUN, *Gerschgorin type theorem and the perturbation of the eigenvalues of a singular pencil*, Chinese J. Num. Math. Appl., 10:1 (1988), pp. 1–13. (Math. Num. Sin., 7:3 (1985), pp. 253–264)
- [74] J. G. SUN, *Multiple Eigenvalue Sensitivity Analysis*, Linear Algebra Appl., 137/138 (1990), pp. 183–211.
- [75] R. C. THOMPSON, *Interlacing inequalities for invariant factors*, Linear Algebra Appl., 24 (1979), pp. 1–31.
- [76] R. C. THOMPSON, *Invariant factors under rank one perturbations*, Can. J. Math., XXXII, no. 1 (1980), pp. 240–245.
- [77] F. TISSEUR Y K. MEERBERGEN, *A survey of the quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 234–286.
- [78] M. M. VAINBERG Y V. A. TRENIGIN, *The methods of Lyapunov and Schmidt in the Theory of Non-Linear Equations and their further development*, P. Noordhoff, Leyden (1974).
- [79] P. VAN DOOREN, *The computation of Kronecker's Canonical Form of a Singular Pencil*, Linear Algebra Appl., 27 (1979), pp. 103–140.
- [80] P. VAN DOOREN, *The generalized eigenstructure problem in linear system theory*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 111–128.

- [81] P. VAN DOOREN, *Reducing subspaces: Definitions, Properties and Algorithms*, In: Matrix Pencils, Proceedings, Pite Havsbad, 1982, Lecture Notes in Mathematics, 973, Springer, Berlin (1983), pp. 58–73.
- [82] P. VAN DOOREN, *Eigenstructuur van polynome en rationale matrices toepassingen in de systeemtheorie*, Tesis Doctoral, Katholieke Univ. Leuven (1979).
- [83] P. VAN DOOREN Y P. DEWILDE, *The Eigenstructure of an Arbitrary Polynomial Matrix: Computational Aspects*, Linear Algebra Appl., 50 (1983), pp. 545–579.
- [84] A. I. G. VARDULAKIS Y G. FRAGULIS, *Infinite Elementary Divisors of Polynomial Matrices and Impulsive Solutions of Linear Homogeneous Matrix Differential Equations*, Circuits System Signal Process, 8, no. 3 (1989), pp. 357–373.
- [85] A. I. G. VARDULAKIS Y N. KARCANIAS, *Relations Between Equivalence Invariants and Structure at Infinity of Matrix Pencils*, IEEE Trans. on Aut. Control, 28, no. 4 (1983), pp. 514–516.
- [86] G. VERGHESE, P. VAN DOOREN Y T. KAILATH, *Properties of the system matrix of a generalized state-space system*, Internat. J. Control, 30 (1979), pp. 235–243.
- [87] G. VERGHESE, *Comments on "Properties of the system matrix of a generalized state-space system"*, Internat. J. Control, 26 (1980), pp. 1007–1009.
- [88] K. VESELIĆ, *On Linear Vibrational Systems with One Dimensional Damping*, Applicable Analysis, 29 (1988), pp. 1–18.
- [89] K. VESELIĆ, *On Linear Vibrational Systems with One Dimensional Damping II*, Integr. Equat. Oper. Theory, 13 (1990), pp. 883–897.
- [90] M. I. VISHIK Y L. A. LYUSTERNIK, *The solution of problems on perturbation in the case of matrices and selfconjugate and non-selfconjugate differential equations*, Uspeshki mat. Nauk., 15 (1960), pp. 3–80.
- [91] W. C. WATERHOUSE, *The codimension of singular matrix pairs*, Linear Algebra Appl., 57 (1984), pp. 227–245.
- [92] K. WEIERSTRASS, *Zur Theorie der bilinearen und quadratischen Formen*, Monatsh. Akad. Wiss. Berlin (1867), pp. 310–368.
- [93] J. H. WILKINSON, *Linear Differential Equations and Kronecker's Canonical Form*, en *Recent Advances in Numerical Analysis*, de Boor and Golub, eds., Academic Press, New York (1978).
- [94] J. H. WILKINSON, *Kronecker's canonical form and the QZ algorithm*, Linear Algebra Appl., 28 (1983), pp. 285–303.
- [95] W. M. WONHAM, *Linear Multivariable Control: A Geometric Approach*, Springer, New York (1979).
- [96] H. XIE Y H. DAI, *On the sensitivity of multiple eigenvalues of nonsymmetric matrix pencils*, Linear Alg. Appl., 374 (2003), pp. 143–158.