# Forma canónica de Kronecker

Juan-Miguel Gracia Melero
Lourdes Ortiz de Elguea Ugartondo

ARGITALPEN ZERBITZUA
SERVICIO EDITORIAL

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

www.argitalpenak.ehu.es
ISBN: 978-84-694-1235-0

# Forma canónica de Kronecker

Juan-Miguel Gracia Melero (Catedrático de Matemática Aplicada)

Lourdes Ortiz de Elguea Ugartondo (Profesora Titular de Álgebra)

23 de junio de 2010

# Índice general

# Prólogo

Este libro trata de explicar con claridad y sencillez la forma canónica de Kronecker de haces de matrices para la relación de equivalencia estricta. El tema es importante para los ingenieros, físicos, químicos, economistas, y otros científicos que estudian sistemas lineales con control, por lo que una introducción asequible y rigurosa se echa de menos. También esperamos que el libro será de utilidad para los matemáticos en un segundo curso de álgebra lineal como complemento natural del estudio de la forma canónica de Jordan. La forma canónica de Kronecker es llamada igualmente de Weierstrass-Kronecker, ya que Weierstrass desarrolló la teoría de los divisores elementales y Kronecker la de los índices minimales. Desde un punto de vista epistemológico e histórico deben relacionarse estas teorías con el estudio geométrico de los haces de cónicas y cuádricas para la formación del estudiante de matemáticas. Este libro no intenta establecer estas conexiones. Al lector que desee proseguir en los precedentes históricos le recomendamos el libro sobre historia de las matemáticas de Bourbaki y también artículos de Robert Thompson, Frank Uhlig y otros en la revista Linear Algebra and Its Applications en los años 1980. Aunque este librito no vuelve a mencionar la historia de las matemáticas, nos parece su estudio muy útil para poder comprender y contextualizar cualquier tema: por ejemplo, intente el lector comprender lo que sobre las formas canónicas de matrices escribieron Turnbull–Aitken [22] en su libro de 1932, y Wedderburn [23] en su libro de 1934.

La forma de Jordan de una matriz cuadrada $A$ encuentra una aplicación natural en el análisis de sistemas de ecuaciones diferenciales lineales de la forma $\dot{x} = Ax + b(t)$, donde $b(t)$ es un vector columna de funciones. De igual modo, la forma de Kronecker permite un estudio completo de la compatibilidad de sistemas $B\dot{x} = Ax + b(t)$, donde $A$ y $B$ son matrices rectangulares de las mismas dimensiones.

El desarrollo que hemos hecho ha estado basado en la presentación magistral del Capítulo 12, tomo 2, del libro de Gantmacher [6]; pero intentando explicar sus puntos oscuros. También somos deudores de publicaciones previas de Friedland [5], Forney [4] y Kailath [12]. El primer autor agradece a Ion Zaballa por haberle introducido en el estudio de los índices minimales de haces de matrices hacia 1983.

# Capítulo 1

# Introducción

## 1.1. Haces de matrices

Sea $\mathbb{F}$ un cuerpo infinito (habitualmente pensaremos que $\mathbb{F}$ es $\mathbb{C}$ cuando convenga, lo que se dirá explícitamente). Los elementos de $\mathbb{F}$ se llamarán constantes o escalares cuando convenga. Sea $\mathbb{F}^{m\times n}$ el espacio vectorial de las matrices $m \times n$ sobre $\mathbb{F}$ y sea $\mathrm{GL}_n(\mathbb{F})$ el conjunto de las matrices cuadradas $n \times n$ no singulares o invertibles. Por $\mathbb{F}[\lambda]$ designamos el anillo de polinomios en la indeterminada $\lambda$, y por $\mathbb{F}(\lambda)$ el cuerpo de fracciones (o funciones) racionales en $\lambda$. Por $\mathbb{F}[\lambda]^{m\times n}$ el conjunto de matrices $m \times n$ con elementos en el anillo $\mathbb{F}[\lambda]$ o *matrices polinómicas*, o $\lambda$–matrices. Y por $\mathbb{F}(\lambda)^{m\times n}$ el espacio de matrices $m \times n$ con elementos en el cuerpo $\mathbb{F}(\lambda)$, o *matrices racionales*. Por $\mathbb{F}[\lambda, \mu]$ designaremos el anillo de polinomios en dos indeterminadas $\lambda$ y $\mu$ con coeficientes en $\mathbb{F}$, y por $\mathbb{F}(\lambda, \mu)$ el cuerpo de fracciones racionales en $\lambda$ y $\mu$ con coeficientes en $\mathbb{F}$. Una matriz polinómica $H(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ puede identificarse con un polinomio matricial

$$H(\lambda) := A_0 + \lambda A_1 + \lambda^2 A_2 + \cdots + \lambda^p A_p \in \mathbb{F}^{m\times n}[\lambda],$$

con $A_0, A_1, \ldots, A_p \in \mathbb{F}^{m\times n}$ y $A_p \neq 0$, siendo $p$ el mayor de los grados de los polinomios $h_{ij}(\lambda) \in \mathbb{F}[\lambda]$ tales que $H(\lambda) = (h_{ij}(\lambda))$. Como $A_p \neq 0$, diremos que el polinomio matricial $H(\lambda)$ tiene grado $p$. Si $A_1 = A_2 = \cdots = A_p = 0$ y $A_o \neq 0$ entonces se dice que $H(\lambda)$ tiene grado 0; si además $A_0 = 0$, se dice que el grado de $H(\lambda)$ es $-\infty$. Un polinomio matricial $H(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ es un *haz de matrices* si tiene grado $\leq 1$, es decir si

$$H(\lambda) := \lambda B - A,$$

donde $A, B \in \mathbb{F}^{m\times n}$.

**Definición 1.1.1.** Sean $H(\lambda) = \lambda B - A$ y $H_1(\lambda) = \lambda B_1 - A_1$ dos haces de $\mathbb{F}^{m\times n}[\lambda]$.

(i) Se dice que son *equivalentes* como matrices polinómicas, lo que se denotará $\lambda B - A \sim \lambda B_1 - A_1$, si existen matrices $U(\lambda) \in \mathbb{F}[\lambda]^{m\times m}$, y $V(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ con determinantes constantes no nulas, tales que

$$\lambda B_1 - A_1 = U(\lambda)(\lambda B - A)V(\lambda).$$

(ii) Se dice que son *estrictamente equivalentes* si existen matrices $P \in \mathrm{GL}_m(\mathbb{F})$ y $Q \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$H_1(\lambda) = PH(\lambda)Q,$$

o equivalentemente $A_1 = PAQ$ y $B_1 = PBQ$; lo que escribiremos

$$H(\lambda) \overset{s}{\sim} H_1(\lambda) \quad \text{o} \quad \lambda B - A \overset{s}{\sim} \lambda B_1 - A_1.$$

**Definición 1.1.2.** Llamaremos *rango normal* del haz $H(\lambda) = \lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$ al orden del mayor menor de $\lambda B - A$ que sea diferente del polinomio cero. Lo denotaremos por $\operatorname{rgn} H(\lambda)$.

El rango normal, así definido, coincide con el rango de $\lambda B - A$ como matriz cuyas componentes son elementos del cuerpo $\mathbb{F}(\lambda)$, cuerpo de fracciones de $\mathbb{F}[\lambda]$. Una consecuencia inmediata del Lema 1.2.20, que veremos en la sección siguiente, es que dos haces equivalentes tienen el mismo rango normal.

Es conocido que para la equivalencia de $\lambda$–matrices, existe la forma normal de Smith, a saber, si $H(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ es una matriz polinómica de rango $r$, entonces existen $U(\lambda) \in \mathbb{F}[\lambda]^{m\times m}$ y $V(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ matrices unimodulares (es decir, matrices cuyo determinante es una constante no nula) tales que $U(\lambda)H(\lambda)V(\lambda) = D(\lambda)$ es la forma normal de Smith de $H(\lambda)$, esto es,

$$U(\lambda)H(\lambda)V(\lambda) = \left[\begin{array}{ccc|c} d_1(\lambda) & & & \\ & \ddots & & 0 \\ & & d_r(\lambda) & \\ \hline & 0 & & 0 \end{array}\right]$$

donde cada polinomio $d_i(\lambda)\,(i = 1, \ldots, r-1)$ divide al siguiente, $d_r(\lambda) \neq 0$, ya que $\operatorname{rg} D(\lambda) = \operatorname{rg} H(\lambda) = r$, y todos los $d_i(\lambda)$ son mónicos (es decir, de coeficiente principal 1) y se llaman *los factores invariantes* de $H(\lambda)$. La demostración puede encontrarse en Gantmacher, vol. I, p. 142, Théoréme 3.

Obviamente, la equivalencia estricta de haces implica su equivalencia como matrices polinómicas. Pero la equivalencia estricta es más exigente, esto es, requiere que $U(\lambda)$ y $V(\lambda)$ sean constantes o independientes de $\lambda$. En el capítulo siguiente (v. Ejemplo 2.1.2) probaremos que la equivalencia de haces no implica la equivalencia estricta, en general.

*Observación* 1.1.3. Es conocido que si $A, A_1 \in \mathbb{F}^{n\times n}$, se dice que $A$ y $A_1$ son semejantes, y se denota $A \approx A_1$, si existe una matriz $P \in \mathrm{GL}_n(\mathbb{F})$ tal que $A_1 = P^{-1}AP$. Es evidente que

$$A \approx A_1 \Longleftrightarrow \lambda I_n - A \overset{s}{\sim} \lambda I_n - A_1.$$

En efecto, si $A_1 = P^{-1}AP$, entonces

$$\lambda I_n - A_1 = \lambda P^{-1}I_nP - P^{-1}AP = P^{-1}(\lambda I_n - A)P.$$

Recíprocamente, si existen $P, Q \in \mathrm{GL}_n(\mathbb{F})$ tales que $\lambda I_n - A_1 = P(\lambda I_n - A)Q$, entonces $I_n = PQ$, $A_1 = PAQ$; luego $P = Q^{-1}$; de donde $A_1 = Q^{-1}AQ$; es decir que $A \approx A_1$.

En este libro, tratamos de establecer un criterio de equivalencia estricta de dos haces de matrices. Vamos a determinar un sistema completo de invariantes, lo que permite determinar, para cada haz, una forma canónica que le es estrictamente equivalente.

Al final de este libro quedará probado que si $\mathbb{F} = \mathbb{C}$, un sistema completo de invariantes está dado por:

1) Una sucesión creciente de enteros $\geq 0$:

$$\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p,$$

llamados *índices minimales (o de Kronecker) por columnas*.

2) Una sucesión creciente de enteros $\geq 0$:

$$\eta_1 \leq \eta_2 \leq \cdots \leq \eta_q,$$

llamados *índices minimales (o de Kronecker) por filas* .

3) Un sistema de divisores elementales de la forma $(\lambda - \lambda_0)^k$, con $\lambda_0 \in \mathbb{C}$.

4) Un sistema de polinomios en $\mu$ de la forma $\mu^j$, llamados *divisores elementales infinitos*.

Para un haz concreto dado pueden faltar uno o varios (hasta tres) de estos sistemas de invariantes. A este sistema completo de invariantes viene asociada una forma canónica llamada de Kronecker. La introducimos mediante un ejemplo.

**Ejemplo 1.1.4.** Sea $\lambda B - A \in \mathbb{C}[\lambda]^{17\times 16}$ y supongamos que este haz tiene los invariantes siguientes:

$$\begin{array}{llll}
\varepsilon_1 = 0, & \varepsilon_2 = 1, & \varepsilon_3 = 2, & \\
\eta_1 = 0, & \eta_2 = 0, & \eta_3 = 1, & \eta_4 = 2, \\
\quad\lambda^2, & (\lambda+2)^2, & & \\
\quad\mu^3. & & &
\end{array}$$

Su forma canónica será:

$$\begin{pmatrix}
0 &&&&&&&&&&&&&&& \\
0 &&&&&&&&&&&&&&& \\
& \lambda & -1 &&&&&&&&&&&&& \\
&&& \lambda & -1 & 0 &&&&&&&&&& \\
&&& 0 & \lambda & -1 &&&&&&&&&& \\
&&&&&& \lambda &&&&&&&&& \\
&&&&&& -1 &&&&&&&&& \\
&&&&&&& \lambda & 0 &&&&&&& \\
&&&&&&& -1 & \lambda &&&&&&& \\
&&&&&&& 0 & -1 &&&&&&& \\
&&&&&&&&& -1 & \lambda & 0 &&&& \\
&&&&&&&&& 0 & -1 & \lambda &&&& \\
&&&&&&&&& 0 & 0 & -1 &&&& \\
&&&&&&&&&&&& \lambda & -1 && \\
&&&&&&&&&&&& 0 & \lambda && \\
&&&&&&&&&&&&&& \lambda+2 & -1 \\
&&&&&&&&&&&&&& 0 & \lambda+2
\end{pmatrix}$$

donde $\left(\begin{array}{cc|cc} \lambda & -1 & & \\ 0 & \lambda & & \\ \hline & & \lambda+2 & -1 \\ & & 0 & \lambda+2 \end{array}\right)$ se llama la *parte de Jordan* y

$$\left(\begin{array}{ccc|cc|cc} -1 & \lambda & 0 & & & & \\ 0 & -1 & \lambda & & & & \\ 0 & 0 & -1 & & & & \\ \hline & & & \lambda & -1 & & \\ & & & 0 & \lambda & & \\ \hline & & & & & \lambda+2 & -1 \\ & & & & & 0 & \lambda+2 \end{array}\right)$$

se llama la *parte de Weierstrass.*

Nuestro problema encuentra una interpretación geométrica natural como sigue. Sean $V$ y $W$ espacios vectoriales sobre el cuerpo $\mathbb{F}$, $\boldsymbol{\mathcal{B}} = \{v_1, \ldots, v_n\}$ y $\boldsymbol{\mathcal{B}}' = \{w_1, \ldots, w_m\}$ sendas bases de $V$ y $W$. Consideremos una aplicación lineal

$$\mathbb{A} : V \longrightarrow W$$

de $V$ en $W$. Recordemos la bien conocida definición siguiente.

**Definición 1.1.5.** Si $\mathbb{A}(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad (j = 1, 2, \ldots, n)$, llamamos *matriz asociada a* $\mathbb{A}$ *respecto a las bases* $\boldsymbol{\mathcal{B}}$ *y* $\boldsymbol{\mathcal{B}}'$ a la matriz

$$M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{A}) := \begin{pmatrix} * & a_{1j} & * \\ * & a_{2j} & * \\ \vdots & \vdots & \vdots \\ * & a_{mj} & * \end{pmatrix} \in \mathbb{F}^{m \times n}.$$

Es bien conocido cómo afecta un cambio de bases a la matriz asociada a una aplicación lineal. Si $\boldsymbol{\mathcal{B}}_1$ y $\boldsymbol{\mathcal{B}}_1'$ son nuevas bases de $V$ y $W$ respectivamente, sabemos que existen matrices $P \in \mathrm{GL}_m(\mathbb{F})$ y $Q \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$A_1 = PAQ, \tag{1.1}$$

siendo $A := M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{A})$ y $A_1 := M_{\boldsymbol{\mathcal{B}}_1'}^{\boldsymbol{\mathcal{B}}_1}(\mathbb{A})$.

**Proposición 1.1.6.** *Sea* $\mathbb{A} : V \longrightarrow W$ *una aplicación lineal sobre el cuerpo* $\mathbb{F}$. *Entonces existen bases* $\mathcal{B}$ *en* $V$ *y* $\mathcal{B}'$ *en* $W$ *tales que*

$$M_{\mathcal{B}'}^{\mathcal{B}}(\mathbb{A}) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

*donde* $r$ *es la dimensión del subespacio imagen de* $\mathbb{A}$ *(abreviadamente* $\mathrm{Im}\,\mathbb{A}$*), es decir, el rango de* $\mathbb{A}$.

Demostración. Sea $\{u_1, \ldots, u_p\}$ una base de $\mathrm{Ker}\,\mathbb{A}$. Ampliamos esta base hasta obtener una base de $V$, digamos,

$$\mathcal{B} = \{v_1, \ldots, v_r, u_1, \ldots, u_p\}.$$

Entonces $\mathbb{A}(v_i) \neq 0$ para $i = 1, \ldots, r$. Además, el conjunto $\{\mathbb{A}(v_1), \ldots, \mathbb{A}(v_r)\}$ es linealmente independiente. En efecto, si se da la relación

$$\alpha_1 \mathbb{A}(v_1) + \cdots + \alpha_r \mathbb{A}(v_r) = 0$$

con $\alpha_i \in \mathbb{F}$ $(i = 1, \ldots, r)$, entonces

$$\mathbb{A}(\alpha_1 v_1 + \cdots + \alpha_r v_r) = 0.$$

Por lo tanto, existen $\beta_1, \ldots, \beta_p \in \mathbb{F}$ tales que

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = \beta_1 u_1 + \cdots + \beta_p u_p.$$

Por consiguiente

$$\alpha_1 v_1 + \cdots + \alpha_r v_r - \beta_1 u_1 - \cdots - \beta_p u_p = 0.$$

Y por ser $\mathcal{B}$ base se sigue que $\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_p = 0$. Además, como $\{\mathbb{A}(v_1), \ldots, \mathbb{A}(v_r)\}$ genera el subespacio $\mathbb{A}(V)$, tenemos que $\{\mathbb{A}(v_1), \ldots, \mathbb{A}(v_r)\}$ es base de $\mathbb{A}(V)$, así que $\operatorname{rg} \mathbb{A} = \dim \mathbb{A}(V) = r$. Ampliamos ahora el conjunto $\{\mathbb{A}(v_1), \ldots, \mathbb{A}(v_r)\}$ hasta conseguir una base

$$\mathcal{B}' = \{\mathbb{A}(v_1), \ldots, \mathbb{A}(v_r), w_1, \ldots, w_s\}$$

de $W$. Con lo cual, es obvio que

$$M_{\mathcal{B}'}^{\mathcal{B}}(\mathbb{A}) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

$\square$

Si $\mathbb{B} : V \longrightarrow W$ es ahora otra aplicación lineal de $V$ en $W$, llamando $B := M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{B})$ y $B_1 := M_{\boldsymbol{\mathcal{B}}_1'}^{\boldsymbol{\mathcal{B}}_1}(\mathbb{B})$ nos queda la fórmula análoga a (1.1)

$$B_1 = PBQ. \tag{1.2}$$

De modo que todos los haces de la clase de equivalencia estricta del haz $\lambda B - A$ nos proporcionan todas las matrices asociadas al par de aplicaciones $(\mathbb{A}, \mathbb{B})$ al efectuar cambios de bases en $V$ y $W$. De aquí que para obtener una forma canónica de un haz, es necesario encontrar bases, digamos $\boldsymbol{\mathcal{B}}_c$ en $V$ y $\boldsymbol{\mathcal{B}}_c'$ en $W$, respecto de las cuales el par de aplicaciones lineales $(\mathbb{A}, \mathbb{B})$ tenga matrices asociadas

$$M_{\boldsymbol{\mathcal{B}}_c'}^{\boldsymbol{\mathcal{B}}_c}(\mathbb{A}) \quad \text{y} \quad M_{\boldsymbol{\mathcal{B}}_c'}^{\boldsymbol{\mathcal{B}}_c}(\mathbb{B})$$

de la forma más "simple" posible, simultáneamente.

El problema principal a tratar en este libro es cómo elegir dichas bases $\boldsymbol{\mathcal{B}}_c$ en $V$ y $\boldsymbol{\mathcal{B}}_c'$ en $W$. Curiosamente la solución aportada no es de naturaleza geométrica, como se verá.

Obsérvese que si $g = \dim(\operatorname{Ker} \mathbb{A} \cap \operatorname{Ker} \mathbb{B})$, podemos elegir una base $\{e_1, \ldots, e_g\}$ de $\operatorname{Ker} \mathbb{A} \cap \operatorname{Ker} \mathbb{B}$, que se puede ampliar hasta una base

$$\boldsymbol{\mathcal{B}}_1 = \{*, \ldots, *, e_1, \ldots, e_g\}$$

de $V$. De modo que, para toda base $\mathcal{B}'_1$ de $W$, tenemos

$$M_{\mathcal{B}'_1}^{\mathcal{B}_1}(\mathbb{A}) = \begin{pmatrix} * & \dots & * & \overbrace{0 \dots 0}^{g} \\ \vdots & & \vdots & \vdots \quad \vdots \\ * & \dots & * & 0 \dots 0 \end{pmatrix} ; M_{\mathcal{B}'_1}^{\mathcal{B}_1}(\mathbb{B}) = \begin{pmatrix} * & \dots & * & \overbrace{0 \dots 0}^{g} \\ \vdots & & \vdots & \vdots \quad \vdots \\ * & \dots & * & 0 \dots 0 \end{pmatrix}.$$

La dificultad del problema estriba, pues, en la elección de los $n - g$ vectores primeros de $\mathcal{B}_1$ y de la base $\mathcal{B}'_1$ para que las $n - g$ columnas primeras de estas matrices tengan el mayor número posible de ceros.

Todos los haces de matrices $\lambda B - A$ de dimensión $m \times n$ son de dos tipos fundamentales: los haces *regulares* y los haces *singulares*.

**Definición 1.1.7.** Un haz de matrices $\lambda B - A$ se dice *regular* si

(i) $A$ y $B$ son matrices cuadradas $n \times n$, y

(ii) el determinante $|\lambda B - A| \neq 0$ .

En todos los demás casos ($m \neq n$ o $m = n$ pero $|\lambda B - A| = 0$ ) el haz se llama *singular*.

Un criterio de equivalencia estricta de haces regulares de matrices y, también, una forma canónica para tales haces fueron establecidos por Weierstrass en 1867 sobre la base de su teoría de los divisores elementales. Los problemas análogos para los haces singulares fueron resueltos más tarde, en 1890, por Kronecker. Los resultados de Kronecker forman el contenido fundamental de este libro.

## 1.2. Algunos preliminares algebraicos

Un problema algebraico que aparece de forma natural al estudiar los haces $\lambda B - A$ es que no bastan los polinomios en una variable. Para el estudio de la relación de semejanza de una matriz cuadrada $A$, es obvio que los menores de su matriz característica $\lambda I - A$ son polinomios en $\lambda$ cuyos coeficientes dependen de $A$. Pero, en un haz $\lambda B - A$ el papel desempeñado por las matrices $A$ y $B$ en la $\lambda$–matriz $\lambda B - A$ es asimétrico, y no da cuenta de la realidad, pues en un haz $\lambda B - A$ las matrices $A$ y $B$ están en pie de igualdad. Para poner esto de manifiesto algunos autores prefieren hablar de un par de matrices $(A, B) \in \mathbb{F}^{m \times n} \times \mathbb{F}^{m \times n}$ en vez de un haz $\lambda B - A$. Así pues, debemos introducir apropiadamente el haz homogéneo

$$\lambda B - \mu A;$$

los menores de esta matriz son polinomios en $\lambda$ y $\mu$. de esta manera los papeles de $A$ y $B$ en $\lambda B - \mu A$ son simétricos. Por consiguiente, se impone trabajar en el anillo de polinomios en dos variables $\mathbb{F}[\lambda, \mu]$ en vez de sólo en el anillo $\mathbb{F}[\lambda]$.

Sin embargo hay un problema, y es que mientras que el anillo $\mathbb{F}[\lambda]$ es euclídeo, el anillo $\mathbb{F}[\lambda, \mu]$ no lo es; ni siquiera es un dominio de ideales principales. Por consiguiente, como tenemos que trabajar con máximos comunes divisores de los menores de la matriz

$$\lambda B - \mu A,$$

necesitamos establecer el contexto en el que debemos situar el anillo $\mathbb{F}[\lambda, \mu]$. Este anillo es un dominio de factorización única. Por fortuna, la definición de máximo común divisor de varios elementos puede extenderse de los dominios de ideales principales a los dominios de factorización única. A continuación daremos un resumen de estos conceptos.

### 1.2.1. Elementos notables en un anillo

Por anillo entenderemos siempre anillo conmutativo con elemento identidad 1, igualmente todos los cuerpos serán conmutativos. Esta salvedad tiene su origen en la terminología utilizada en los años 1960, sobre todo en libros franceses, que contemplaban la existencia de anillos y cuerpos no conmutativos. Aunque hoy en día prevalece la terminología anglosajona en la que la conmutatividad debe darse por supuesta.

Si $A$ es un anillo y $a, b \in A$ son tales que $ab = 0$, $a \neq 0$, $b \neq 0$, entonces $a$ y $b$ son llamados *divisores de* 0. Un *dominio de integridad* es un anillo sin divisores de 0. Por ejemplo, el anillo de matrices $\mathbb{C}^{2\times 2}$ tiene divisores de cero pues existen matrices $A \neq 0$, $B \neq 0$, tales que

$$AB = 0;$$

basta tomar el caso

$$\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

En un anillo $A$ se dice que un elemento $u \in A$ es una *unidad* (o elemento invertible) si existe $v \in A$ tal que

$$uv = 1;$$

el propio elemento 1 de $A$ es una unidad de $A$, pero puede haber otras unidades en $A$. Por ejemplo, en el anillo $\mathbb{Z}$ de lo números enteros las unidades son 1 y $-1$. En el anillo de polinomios $\mathbb{C}[\lambda]$ las unidades son todos los números complejos salvo cero.

Se dice que $a \in A$ es un *divisor* de $b \in A$, si existe $x \in A$ tal que

$$ax = b;$$

lo que se denota por $a \mid b$ ( y se lee "$a$ divide a $b$").

Supongamos de ahora en adelante que $D$ es un dominio de integridad. Si $a, b \in D$ son tales que $a \mid b$ y $b \mid a$, entonces diremos que $a$ y $b$ son elementos *asociados* y escribiremos $a \sim b$. Es claro que $0 \sim 0$. Supongamos ahora que $a \neq 0$ y $a \sim b$, entonces existen $u, v \in D$ tales que $au = b$, $bv = a$. De donde,

$$a = bv = auv$$

y como $a \neq 0$, de aquí se sigue que $1 = uv$; es decir, $u$ y $v$ son unidades; además $b \neq 0$. Por tanto, $a \sim b$ si y sólo si existe una unidad $u$ tal que $b = au$.

Diremos que un elemento $q \in D$, $q \neq 0$, es un elemento *irreducible* si $q$ no es una unidad y sus únicos divisores son las unidades $u \in D$ y los elementos de la forma $uq$, con $u$ unidad. Diremos que $p \in D$, $p \neq 0$, es un elemento *primo* si $p$ no es una unidad y si para todos $a, b \in D$ tales que

$$p \mid ab$$

se tiene necesariamente que

$$p \mid a \ \text{ o } \ p \mid b.$$

**Proposición 1.2.1.** *Sea $D$ un dominio de integridad. Todo elemento primo en $D$ es irreducible.*

DEMOSTRACIÓN. Supongamos que $p \in D$ es primo y que $r \mid p$. Luego $p = rs$ para algún $s \in D$, y ya que $p$ es primo, tenemos que $p \mid r$ o $p \mid s$. Supongamos que $p \mid s$ de tal manera que $s = tp$ para algún $t \in D$. Entonces se tiene que $p = rs = rtp$, y de aquí que $rt = 1$, i.e. $r$ es una unidad. Por otra parte, si $p \mid r$, como ya $r \mid p$, entonces se sigue que $p$ y $r$ son asociados. Por lo tanto, los únicos divisores de $p$ son las unidades y los asociados con $p$. Así pues, $p$ es irreducible.

□

Pero no todo elemento irreducible en $D$ es necesariamente primo, como se puede ver por medio del ejemplo siguiente.

**Ejemplo 1.2.2.** Sea

$$D := \mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$$

siendo $\mathbb{Z}$ el anillo de los números enteros. Es obvio que $D$ es un subanillo de $\mathbb{C}$. Vamos a demostrar que

$$\alpha := 2 + \sqrt{-5} \in D$$

es irreducible pero no es primo.

Averigüemos la forma de las unidades $a+b\sqrt{-5}$ de $D$. Si existe $c+d\sqrt{-5} \in D$ tal que

$$(a + b\sqrt{-5})(c + d\sqrt{-5}) = 1,$$

multiplicando ambos miembros por $a - b\sqrt{-5}$ obtenemos

$$(a^2 + 5b^2)(c + d\sqrt{-5}) = a - b\sqrt{-5};$$

y como

$$a + b\sqrt{-5} = 0 \quad \Longleftrightarrow \quad a = 0,\ b = 0,$$

deducimos que

$$\begin{aligned} (a^2 + 5b^2)\, c &= a \quad (1) \\ (a^2 + 5b^2)\, d &= -b \quad (2). \end{aligned}$$

Si fuera $c = 0$, se tendría que $a = 0$ y $5b^2 d = -b$. Ya que suponemos que $a + b\sqrt{-5}$ es una unidad, i.e. que $b\sqrt{-5}$ es una unidad, sería $b \neq 0$. Por lo que seguiría que $5bd = -1$ con $b, d$ enteros, lo que es imposible. Así pues, es seguro que $c \neq 0$ y por ser $a$ ó $b$ no nulos, (1) implica que $a \neq 0$. Entonces ha de ser $b = 0$, pues si fuera $b \neq 0$ tendríamos que

$$|(a^2 + 5b^2)\, c| > |a|$$

en contradicción con (1). De modo que $a^2 c = a$; de donde $ac = 1$ y por tanto $a$ es una unidad en $\mathbb{Z}$; luego $a = \pm 1$. En resumidas cuentas, hemos probado que las únicas unidades de $D$ son

$$\begin{aligned} 1 + 0\sqrt{-5} &= 1 \\ -1 + 0\sqrt{-5} &= -1. \end{aligned}$$

En consecuencia, $\alpha$ no es unidad de $D$.

A continuación, veremos que $\alpha = 2 + \sqrt{-5}$ no es primo, observando que

$$(2 + \sqrt{-5})(2 - \sqrt{-5}) = 9 = 3{,}3,$$

luego se tiene que

$$(2 + \sqrt{-5}) \mid 3{,}3,$$

y sin embargo $2 + \sqrt{-5}$ no es divisor de 3. En efecto, si existiera $a + b\sqrt{-5} \in D$ tal que

$$(2 + \sqrt{-5})(a + b\sqrt{-5}) = 3,$$

multiplicando ambos miembros por $2 - \sqrt{-5}$ se deduciría que

$$9(a + b\sqrt{-5}) = 6 - 3\sqrt{-5},$$

de donde $9a = 6$, lo que es imposible con $a$ entero. Por consiguiente $\alpha$ no es primo.

Sólo falta demostrar que $\alpha$ es irreducible. Supongamos que se verifica

$$(a + b\sqrt{-5})(c + d\sqrt{-5}) = 2 + \sqrt{-5}, \tag{1.3}$$

siendo $a + b\sqrt{-5}, c + d\sqrt{-5} \in D$. Si $z$ es el número complejo $z := u + iv$, con $u, v \in \mathbb{R}$, denotamos por $\bar{z} := u - iv$ el conjugado de $z$. Si ahora observamos que

$$a + b\sqrt{-5} = a + b\sqrt{5}i$$

tenemos que

$$\overline{a + b\sqrt{-5}} = a - b\sqrt{5}i = a - b\sqrt{-5};$$

así pues, tomando conjugados complejos en (1.3), tenemos también

$$(a - b\sqrt{-5})(c - d\sqrt{-5}) = 2 - \sqrt{-5}, \tag{1.4}$$

y multiplicando después (1.3) por (1.4) miembro a miembro, se sigue que

$$(a^2 + 5b^2)(c^2 + 5d^2) = 9.$$

De manera que $a^2 + 5b^2$ divide a 9 en $\mathbb{Z}$; consecuentemente $a^2 + 5b^2$ es $1, 3$ ó 9. Si fuera $a^2 + 5b^2 = 1$, se tendría que $b = 0$ y $a = \pm 1$, en cuyo caso $a + b\sqrt{-5}$ sería una unidad de $D$. Si fuera $a^2 + 5b^2 = 3$, también se seguiría que $b = 0$ y $a^2 = 3$ lo que es imposible pues $a$ es entero. Por último, si $a^2 + 5b^2 = 9$, entonces $c^2 + 5d^2 = 1$, lo que implica que $d = 0$ y $c = \pm 1$. En otras palabras, (1.3) se convierte en

$$(a + b\sqrt{-5}) = \pm(2 + \sqrt{-5});$$

esto significa que $a + b\sqrt{-5}$ y $\alpha$ son asociados. La conclusión obtenida es que si $a + b\sqrt{-5}$ es un divisor de $\alpha$, $a + b\sqrt{-5}$ es una unidad de $D$ o es un asociado de $\alpha$; es decir que $\alpha$ es irreducible.

### 1.2.2. Máximo común divisor

**Definición 1.2.3.** Sea $a_1, a_2, \ldots, a_n$ una sucesión de $n$ elementos de un dominio de integridad $D$. Si

$$a_1 = a_2 = \cdots = a_n = 0,$$

decimos que el máximo común divisor de $a_1, a_2, \ldots, a_n$ es 0 , por definición.

Si para algún $i \in \{1, 2, \ldots, n\}$ es $a_i \neq 0$, decimos que un elemento $d \in D$ es *un máximo común divisor* de la secuencia $a_1, a_2, \ldots, a_n$ si

(1) $d \mid a_i$ para $i = 1, \ldots, n$,

(2) para cualquier $c \in D$ tal que $c \mid a_i$ para $i = 1, \ldots, n$, se tiene que $c \mid d$.

En dominios de integridad podemos hablar de los máximos comunes divisores de una sucesión finita de elementos; la principal dificultad es que éstos, en general, tal vez no existan. Pero en los dominios euclídeos y en los dominios de factorización única, su existencia es segura. Antes de nada, si $d_1$ y $d_2$ son máximos comunes divisores de $a_1, a_2, \ldots, a_n$, siendo algún $a_i$ distinto de 0, por la parte (2) de la definición se tiene que

$$d_1 \mid d_2 \text{ y } d_2 \mid d_1.$$

Por tanto $d_1 \sim d_2$ (son asociados); en consecuencia, los máximos comunes divisores de una secuencia $a_1, a_2, \ldots, a_n$ están determinados salvo asociados. Es decir, si $d$ es un máximo común divisor (mcd) de $a_1, a_2, \ldots, a_n$, entonces los restantes máximos comunes divisores de $a_1, a_2, \ldots, a_n$ son los elementos de la forma

$$ud$$

donde $u$ recorre las unidades de $D$.

**Definición 1.2.4.** Sea $D$ un dominio de integridad. Sea $\mathbb{N} = \{0, 1, 2, \ldots\}$ el conjunto de los números naturales. Se dice que $D$ es un *dominio euclídeo*, si es posible definir una aplicación

$$\delta : D - \{0\} \longrightarrow \mathbb{N}$$

que satisfaga las dos condiciones siguientes:

(1) $\delta(c) \leq \delta(cd)$ si $cd \neq 0$.

(2) Si $b \neq 0$ y $a \in D$, entonces existen $q$ y $r$ en $D$ tales que

$$a = bq + r$$

siendo $r = 0$ ó $\delta(r) < \delta(b)$.

*Observación* 1.2.5. Intuitivamente (2) nos dice que $D$ es un dominio euclídeo si podemos definir en él una "división euclídea".

**Ejemplo 1.2.6.** (1) El anillo $\mathbb{Z}$ de los enteros es un dominio euclídeo con la aplicación $\delta(m) = |m|$ para todo $m \in \mathbb{Z} - \{0\}$.

(2) Si $\mathbb{F}$ es un cuerpo, el anillo $\mathbb{F}[x]$ de polinomios en una variable $x$ sobre $\mathbb{F}$, es un dominio euclídeo con la aplicación

$$\delta(f(x)) = \text{grado} f(x), \;\; f(x) \neq 0.$$

Mediante el algoritmo de Euclides se puede demostrar que en todo dominio euclídeo existen máximos comunes divisores de dos elementos. Como el mcd es una operación asociativa, i.e. para todo $a, b, c \in D$

$$\mathrm{mcd}(\mathrm{mcd}(a, b), c) = \mathrm{mcd}(a, \mathrm{mcd}(b, c)),$$

esto permite probar que existen mcd de $n$ elementos de $D$ ($n \geq 2$). Por lo tanto, dados $n$ polinomios $f_1(x), \ldots, f_n(x) \in \mathbb{F}[x]$, con $\mathbb{F}$ un cuerpo, tiene sentido hablar de los mcd de $f_1(x), \ldots, f_n(x)$, los cuales son asociados entre sí, luego se diferencian sólo en un factor $c \in \mathbb{F} - \{0\}$, ya que las unidades de $\mathbb{F}[x]$ son las constantes no nulas. Llamamos *mónico* a un polinomio no nulo cuyo coeficiente principal es 1; así un polinomio mónico de grado $m$ tiene la forma

$$x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0.$$

Si entre los mcd de $f_1(x), \ldots, f_n(x)$ elegimos el que es mónico, digamos $d(x)$, este mcd queda unívocamente determinado y podemos llamarle el mcd de

$$f_1(x), \ldots, f_n(x)$$

y denotarlo

$$\mathrm{mcd}(f_1(x), \ldots, f_n(x)).$$

Nosotros necesitaremos el concepto de máximo común divisor de varios polinomios $f_1(x, y), \ldots, f_n(x, y) \in \mathbb{F}[x, y]$. Como el dominio $\mathbb{F}[x, y]$ no es un dominio euclídeo, no podemos echar mano del algoritmo de Euclides para probar la existencia de este máximo común divisor. Hay una clase de anillos más general que la clase de los dominios euclídeos, en la que puede hablarse de máximo común divisor. Esta clase es la de los dominios de ideales principales.

**Definición 1.2.7.** Un *ideal* $I$ de un dominio de integridad $D$ es un subconjunto no vacío de $D$ tal que $I$ es un subgrupo aditivo de $D$ y para cualesquiera $d \in D$, $a \in I$, se verifica que $da \in I$. El conjunto

$$(a) = \{da \mid d \in D\},$$

es un ideal de $D$, para todo $a \in D$; los ideales de este tipo se llaman *principales*. Se dice que un dominio de integridad $D$ es un *dominio de ideales principales* si todos sus ideales son principales.

*Observación* 1.2.8. Si $\mathbb{F}$ es cuerpo, entonces $\mathbb{F}[x]$ es un dominio de ideales principales (véase p. 126, Theorem 2.15 de [11]). Lamentablemente, el anillo $\mathbb{F}[x, y]$ no es un dominio de ideales principales (véase Ejercicio 1.1). Por eso, no centraremos nuestra atención en esta estructura.

Sabemos que todo entero $a \in \mathbb{Z}$ es producto de números primos, único salvo el orden y el signo $\pm 1$,

$$a = p_1 p_2 \cdots p_m \quad \text{ó} \quad a = -p_1 p_2 \cdots p_m$$

con $p_1, p_2, \ldots, p_m$ números primos positivos. Análogamente, todo polinomio $f(x) \in \mathbb{C}[x]$ puede factorizarse de forma única (salvo el orden de los factores) en producto de una unidad y de monomios de primer grado, $x - \alpha$, con $\alpha \in \mathbb{C}$. Es decir,

$$f(x) = c(x - \alpha_1) \cdots (x - \alpha_n)$$

con $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ las raíces de $f(x)$ y $c$ una constante no nula. Una propiedad análoga a esta factorización única la tienen los anillos $\mathbb{F}[x]$ y $\mathbb{F}[x, y]$, con $\mathbb{F}$ un cuerpo cualquiera.

**Definición 1.2.9.** Un dominio de integridad $D$ se llama un *dominio de factorización única* (o un anillo factorial) si todo elemento no nulo y no unidad $a \in D$ es expresable como producto de un número finito de elementos irreducibles y esta factorización es única salvo el orden y asociados. Así, pues, si

$$a = p_1 p_2 \cdots p_m = q_1 q_2 \cdots q_n$$

y $p_i$ $(i = 1, 2, \ldots, m)$ y $q_j$ $(j = 1, 2, \ldots, n)$ son elementos irreducibles de $D$, entonces $m = n$ y existirá una permutación $\sigma$ de $\{1, 2, \ldots, n\}$ tal que para todo $i \in \{1, 2, \ldots, n\}$ los elementos $p_i$ y $q_{\sigma(i)}$ son asociados.

*Observación* 1.2.10. Se puede demostrar que si $D$ es un dominio de ideales principales, entonces $D$ es un dominio de factorización única (véase p. 185, Theorem 3.6 de [17]).

**Proposición 1.2.11.** *En un dominio de factorización única $D$, el conjunto de los elementos primos coincide con el conjunto de elementos irreducibles.*

Demostración.Por la Proposición 1.2.1 ya sabemos que todo elemento primo de $D$ es irreducible. Supongamos que $a \in D$ es irreducible. Ya que $D$ es un dominio de factorización única, existen elementos irreducibles $p_1, \ldots, p_n \in D$ tal que

$$a = p_1 \cdots p_n.$$

Ahora bien, $a$ es irreducible, $p_1 \mid a$ y $p_1$ no es una unidad; por tanto, $p_1$ debe ser un asociado de $a$. Entonces $a$ es un asociado de un primo y, de aquí, también $a$ debe ser primo.

□

*Observación* 1.2.12. Entonces, del Ejemplo 1.2.2 se sigue, en virtud de la proposición anterior, que $\mathbb{Z}[\sqrt{-5}]$ no es un dominio de factorización única.

**Teorema 1.2.13.** *Si $D$ es un dominio de factorización única y $a_1, \ldots, a_n$ son elementos de $D$, con alguno no nulo, entonces existe un* mcd *de $a_1, \ldots, a_n$, que es único salvo asociados.*

Demostración. Sean $b_1, \ldots, b_m$ los elementos no nulos que hay en $a_1, \ldots, a_n$ (con $m \leq n$). Ya que $D$ es un dominio de factorización única existen elementos irreducibles (no asociados) $p_1, \ldots, p_r \in D$ y unidades $u_1, \ldots, u_m \in D$ tales que

$$b_k = u_k p_1^{e_{1k}} \cdots p_r^{e_{rk}}, \quad k = 1, \ldots, m.$$

donde los $e_{jk}$ son enteros no negativos (por supuesto, si $b_k$ es una unidad, tenemos que $e_{jk} = 0,\ j = 1, \ldots, r$). Sean

$$e_1 := \min_{1 \leq k \leq m} \{e_{1k}\}, \ldots, e_r := \min_{1 \leq k \leq m} \{e_{rk}\}$$

y sea

$$d := p_1^{e_1} \cdots p_r^{e_r}.$$

Como en un dominio de factorización única, dados dos elementos cualesquiera $a, b \in D$ pueden escribirse factorizados así

$$a = uq_1^{m_1} \cdots q_s^{m_s}, \quad b = vq_1^{n_1} \cdots q_s^{n_s},$$

con $u, v$ unidades $q_1, \ldots, q_s$ elementos irreducibles y exponentes $m_1, \ldots, m_s$, $n_1, \ldots, n_s$ enteros no negativos, se tiene que

$$a \mid b \quad \Longleftrightarrow \quad m_i \leq n_i, \ (i = 1, \ldots, s). \tag{1.5}$$

Por tanto, $d \mid b_k$ para $k = 1, \ldots, m$. Supongamos ahora que $c \mid b_k$ para $k = 1, \ldots, m$. Puesto que $D$ es un dominio de factorización única, se tiene que en la factorización de $c$ en elementos irreducibles no pueden aparecer irreducibles distintos de asociados de $p_1, \ldots, p_r$; además, después de multiplicar las unidades apropiadas, vemos que el exponente de cada $p_j$ es menor o igual que $e_j$, por (1.5). Dicho abreviadamente

$$c = wp_1^{f_1} \cdots p_r^{f_r}$$

donde $f_j \leq e_j$ $(j = 1, \ldots, r)$ y $w$ es una unidad. Por tanto, nuevamente en virtud de (1.5), se sigue que $c \mid d$, lo que prueba que $d$ es un mcd de $b_1, \ldots, b_m$ y, por ello, de $a_1, \ldots, a_n$. La unicidad de $d$ salvo asociados es consecuencia de la definición de mcd.

□

Por último, se puede demostrar que si $D$ es un dominio de factorización única entonces el anillo de polinomios $D[x]$ también es un dominio de factorización única (véase p. 147, Theorem 2.25 de [11]). De aquí que, si $\mathbb{F}$ es un cuerpo, $\mathbb{F}[x, y] = \mathbb{F}[x][y]$ es un dominio de factorización única, ya que $\mathbb{F}[x]$ lo es.

### 1.2.3. Invariantes de la equivalencia en un dominio de factorización única

Introducimos la notación siguiente

$$Q_{k,n} := \{\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{N}^k \mid 1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq n\},$$

esto es, $Q_{k,n}$ denota el conjunto de todas las sucesiones finitas estrictamente crecientes de $k$ enteros elegidos de $1, \ldots, n$.

Sea $D$ un anillo, $D^{m \times n}$ denota el conjunto de todas las matrices con elementos en $D$ de $m$ filas y $n$ columnas

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Si $A \in D^{m \times n}$, $\alpha \in Q_{k,m}$ y $\beta \in Q_{k,n}$, entonces $A[\alpha \mid \beta]$ es la submatriz $k \times k$ de $A$ formada por las filas numeradas por $\alpha$ y por las columnas numeradas por $\beta$.

**Definición 1.2.14.** Sea $D$ un anillo conmutativo con unidad. La matriz $A \in D^{n \times n}$ se llama *invertible (o unimodular)* si existe una matriz $B \in D^{n \times n}$ tal que

$$AB = I_n = BA,$$

donde $I_n$ es la matriz identidad (o unidad) de orden $n$.

**Proposición 1.2.15.** *Sea $D$ un anillo conmutativo con unidad y $A \in D^{n\times n}$. Entonces $A$ es invertible o unimodular si y sólo si* $\det A$ *es una unidad (o elemento invertible) de $D$.*

DEMOSTRACIÓN. (c.f. el libro de Jacobson [11] p. 94, Theorem 2.1.)

□

En adelante, denotaremos mediante $\mathrm{GL}_n(D)$ al grupo multiplicativo de todas las matrices $n \times n$ invertibles con elementos en $D$.

**Definición 1.2.16.** Sea $D$ un dominio de integridad. Sean $A, B \in D^{m\times n}$. Se dice que la matriz $A$ es *equivalente* a la matriz $B$, abreviadamente $A \sim B$, si existen matrices invertibles $P \in D^{m\times m}$ y $Q \in D^{n\times n}$ tales

$$B = PAQ.$$

Es fácil probar que $\sim$ es una relación de equivalencia en $D^{m\times n}$.

El problema de seleccionar, entre las matrices equivalentes a una matriz dada $A$, una que tenga una forma "normal" particularmente simple lo proporciona el siguiente resultado, cuya demostración puede encontrarse en el libro de Jacobson [11], p. 176, Theorem 3.8:

**Teorema 1.2.17.** *Si $D$ es un dominio de ideales principales y $A \in D^{m\times n}$, entonces $A$ es equivalente a una matriz que tiene la forma diagonal*

$$\mathrm{diag}\{d_1, d_2, \ldots, d_r, 0. \ldots, 0\},$$

*donde $d_i \neq 0$, para cada $i = 1, \ldots, r$ y $d_1 \mid d_2 \mid \cdots \mid d_r$ .*

**Definición 1.2.18.** Sea $D$ un anillo conmutativo con unidad. Dada $A \in D^{n\times n}$, se llama *rango* de $A$, y se denota $\operatorname{rg} A$, al máximo índice $k$, $k \in \{1, 2, \ldots, \text{mín}\{m, n\}\}$ para el que existen $\alpha \in Q_{k,m}$ y $\beta \in Q_{k,n}$ tales que

$$\det A[\alpha \mid \beta] \neq 0.$$

Es decir el rango $r$ de $A$ es el orden del mayor menor no nulo de $A$.

**Definición 1.2.19.** Sea $D$ un dominio de factorización única y sea $A \in D^{m\times n}$. Entonces definamos, para $k = 1, 2, \ldots, \text{mín}\{m, n\}$,

$$D_k(A) := \mathrm{mcd}\{\det A[\alpha \mid \beta] \mid \alpha \in Q_{k,m},\ \beta \in Q_{k,n}\}$$

$D_k(A)$ se llamará el *divisor determinantal*divisor determinantal $k$–ésimo de $A$.

**Lema 1.2.20.** *Sea $D$ un dominio de factorización única y sean $A, B \in D^{m\times n}$. Entonces se verifica*

$$A \sim B \Rightarrow D_k(A) = D_k(B) \ \ \textit{para todo } k \in \{1, 2, \ldots, \text{mín}\{m, n\}\}.$$

*Más aun, si* $\operatorname{rg} A = r$*, tomando $D_0(A) := 1$, entonces se tiene*

$$D_0(A) \mid D_1(A) \mid \cdots \mid D_r(A)$$

*y por definición* $D_{r+1}(A) = \ldots = D_{\text{mín}(m,n)}(A) = 0$.

*De este modo, los elementos siguientes*

$$i_k(A) = \frac{D_k(A)}{D_{k-1}(A)}, \ (k = 1, 2, \ldots, r),$$

*pertenecen a* $D$ *y son llamados los* factores invariantes *de* $A$.

*En consecuencia, si* $A \sim B$*, entonces* $\operatorname{rg} A = \operatorname{rg} B$ *e*

$$i_k(A) = i_k(B) \ \ \textit{para todo } k \in \{1, 2, \ldots, \text{mín}\{m, n\}\}.$$

*Además, si* $D$ *es un dominio de ideales principales, como es el caso de* $\mathbb{F}[\lambda]$*, de acuerdo con el teorema anterior se tiene que*

$$i_1(A) = d_1 \mid i_2(A) = d_2 \mid \cdots \mid i_r(A) = d_r \ \ \textit{para todo } k \in \{1, 2, \ldots, r\}.$$

Demostración. Como $A \sim B$, existen $P \in \mathrm{GL}_m(D)$ y $Q \in \mathrm{GL}_n(D)$ tales que $B = PAQ$. Sea $k \in \{1, 2, \ldots, \text{mín}\{m, n\}\}$. Entonces, aplicando la fórmula de Binet-Cauchy (c.f. Gantmacher Vol. I, p.12), se tiene que para cualesquiera $\alpha' \in Q_{k,m}$ y $\beta' \in Q_{k,n}$ se verifica

$$\det B[\alpha' \mid \beta'] = \sum_{\alpha \in Q_{k,m}, \beta \in Q_{k,m}} \det P[\alpha' \mid \alpha] \det A[\alpha \mid \beta] \det Q[\beta \mid \beta'],$$

de donde se sigue que $D_k(A)$ divide a todos los menores de orden $k$ de $B$ y por tanto $D_k(A) \mid D_k(B)$. Análogamente, del hecho de que $A = P^{-1}BQ^{-1}$, se obtiene $D_k(B) \mid D_k(A)$ y consecuentemente $D_k(A) = D_k(B)$. Las afirmaciones restantes son evidentes.

□

**Definición 1.2.21.** Siguiendo con la notación del lema anterior, definamos

$$i_k(A) := \frac{D_k(A)}{D_{k-1}(A)},$$

para $k = 1, 2, \ldots, r$ ($r = \operatorname{rg} A$). Los elementos $i_1(A), i_2(A), \ldots, i_r(A)$ son llamados los *factores invariantes* de $A$.

*Observación* 1.2.22. Del lema anterior sigue directamente que si $A$ y $B$ son matrices equivalentes, entonces tienen los mismos factores invariantes.

## 1.3. Polinomios en varias indeterminadas

**Definición 1.3.1.** Se llama *monomio* de $\mathbb{F}[x_1, \ldots, x_n]$ a todo polinomio de la forma $cx_1^{i_1}x_2^{i_2}\ldots x_n^{i_n}$ con $c$ constante no nula. El *grado* de este monomio se define como la suma $i_1 + i_2 + \cdots + i_n$, y se denotará así

$$\partial(cx_1^{i_1}x_2^{i_2}\ldots x_n^{i_n}) := i_1 + i_2 + \cdots + i_n.$$

Se llama grado (total) de un polinomio

$$f(x_1, \ldots, x_n) := \sum_{i_1, i_2, \ldots, i_n} a_{i_1, i_2 \ldots, i_n} x_1^{i_1}x_2^{i_2}\ldots x_n^{i_n} \in \mathbb{F}[x_1, \ldots, x_n]$$

al mayor de los grados de sus monomios y se denota por $\partial(f)$.

*Observación* 1.3.2. Propiedades evidentes del grado son :

1) $\partial(fg) = \partial(f) + \partial(g)$.

2) $\partial(f + g) \leq \max\{\partial(f), \partial(g)\}$.

Nótese que en el caso particular $n = 1$, la definición anterior coincide con la bien conocida del grado de un polinomio en una indeterminada.

**Ejemplo 1.3.3.** Por ejemplo, si

$$f(x_1, x_2, x_3, x_4) = 4x_1^2x_3 + 7x_1x_2x_3^2x_4 \in \mathbb{Q}[x_1, x_2, x_3, x_4],$$

se tiene que $\partial(f) = 5$, pues $\partial(4x_1^2x_3) = 3$ y $\partial(7x_1x_2x_3^2x_4) = 5$.

**Proposición 1.3.4.** *Sea $\mathbb{F}$ un cuerpo y sea $\mathbb{F}[x_1, \ldots, x_n]$ el anillo de polinomios en $n$ indeterminadas con coeficientes en $\mathbb{F}$. Entonces las unidades de $\mathbb{F}[x_1, \ldots, x_n]$ son los elementos no nulos de $\mathbb{F}$.*

DEMOSTRACIÓN. Supongamos que $f \in \mathbb{F}[x_1, \ldots, x_n]$ es una unidad, es decir, un elemento invertible del anillo $\mathbb{F}[x_1, \ldots, x_n]$. Entonces existirá otro $g \in \mathbb{F}[x_1, \ldots, x_n]$ tal que

$$fg = 1. \tag{1.6}$$

Sea $\partial_i(f)$ (resp. $\partial_i(g)$) el grado de $f$ (resp. de $g$) en la indeterminada $x_i$, para $i = 1, \ldots, n$. Entonces, por la propiedad del grado, de (1.6) sigue

$$\partial_i(f) + \partial_i(g) = \partial_i(1) = 0$$

para cada $i = 1, \ldots, n$; lo que obliga a que sea

$$\partial_i(f) = 0 = \partial_i(g)$$

para cada $i = 1, \ldots, n$, y por tanto $f$ es un polinomio constante no nulo.

□

**Definición 1.3.5.** Sea $\mathbb{F}$ un cuerpo y sea $\mathbb{F}[x_1, \ldots, x_n]$ el anillo de polinomios en $n$ indeterminadas con coeficientes en $\mathbb{F}$. Un polinomio $f \in \mathbb{F}[x_1, \ldots, x_n]$ no constante, se dice que es *irreducible* en $\mathbb{F}[x_1, \ldots, x_n]$, si siempre que $f$ es producto de otros dos polinomios, digamos

$$f = g\,h$$

con $g, h \in \mathbb{F}[x_1, \ldots, x_n]$, entonces alguno de los dos, $g$ o $h$, es una constante no nula.

*Observación* 1.3.6. Es fácil ver, que todo polinomio lineal (i.e. de grado 1) en $\mathbb{F}[x_1, \ldots, x_n]$ es irreducible.

**Teorema 1.3.7.** *Sea $\mathbb{F}$ un cuerpo y sea $\mathbb{F}[x_1, \ldots, x_n]$ el anillo de polinomios en $n$ indeterminadas con coeficientes en $\mathbb{F}$. Entonces $\mathbb{F}[x_1, \ldots, x_n]$ es un dominio de factorización única.*

DEMOSTRACIÓN. (c.f. el libro de Jacobson [11], p. 147 Theorem 2.25 y también en [9], p. 150 Corolario 2.)

□

Si $f(x) \in \mathbb{F}[x]$, es bien conocido que el número de raíces o ceros del polinomio $f$ en el cuerpo $\mathbb{F}$ es a lo sumo igual al grado de $f$. Por otra parte, si $\mathbb{F}$ es un cuerpo finito, entonces todos los elementos de $\mathbb{F}$ son raíces del polinomio $x^{|\mathbb{F}|} - x$. En consecuencia, si $|\mathbb{F}| > \partial(f)$, podemos asegurar que existe $a \in \mathbb{F}$ tal que $f(a) \neq 0$. Extendemos este resultado para $n$ indeterminadas en el teorema siguiente.

**Teorema 1.3.8.** *Sean $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ polinomios no nulos y sea $\mathbb{F}$ un cuerpo con más de $\partial(f) + \partial(g)$ elementos. Entonces existen elementos $a_1, a_2, \ldots, a_n$ en $\mathbb{F}$ tales que*

$$f(a_1, \ldots, a_n) \neq 0 \quad y \quad g(a_1, \ldots, a_n) \neq 0.$$

Demostración. Probaremos primero, por inducción sobre $n$, que para cualquier $f \in \mathbb{F}[x_1, \ldots, x_n]$ no nulo existen $a_1, a_2, \ldots, a_n \in \mathbb{F}$ tales que $f(a_1, \ldots, a_n) \neq 0$. Luego el resultado sigue sin más que aplicarlo al polinomio producto $f\,g$.

Para $n = 1$ el resultado es conocido. Lo suponemos cierto para $n-1$ y vamos a demostrarlo para $n$. Consideramos $f \in \mathbb{F}[x_1, \ldots, x_{n-1}][x_n]$, esto es,

$$f(x_1, \ldots, x_n) = h_0 + h_1 x_n + h_2 x_n^2 + \cdots + h_r x_n^r$$

donde $h_i \in \mathbb{F}[x_1, \ldots, x_{n-1}]$ y podemos suponer $h_r = h_r(x_1, \ldots, x_{n-1}) \neq 0$. Entonces, por la hipótesis inductiva, sabemos que existen $a_i \in \mathbb{F}$, $i = 1, \ldots, n-1$, tales que $h_r(a_1, \ldots, a_{n-1}) \neq 0$. Entonces el polinomio $f(a_1, \ldots, a_{n-1}, x_n) \in \mathbb{F}[x_n]$ es no nulo, y como $\mathbb{F}$ tiene más de $\partial(f)$ elementos, podemos elegir $a_n \in \mathbb{F}$ tal que $f(a_1, \ldots, a_n) \neq 0$.

□

## 1.4. Polinomios homogéneos

**Definición 1.4.1.** Un polinomio

$$f(x_1, \ldots, x_n) := \sum_{i_1, i_2, \ldots, i_n} a_{i_1, i_2 \ldots, i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \in \mathbb{F}[x_1, \ldots, x_n]$$

se dice *homogéneo de grado $k$* si todos los monomios $a_{i_1, i_2 \ldots, i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ (no nulos) tienen el mismo grado $k$.

**Ejemplo 1.4.2.** El polinomio de coeficientes racionales

$$3x_1^2 x_2^2 - 7x_1 x_2 x_3^2 + 5x_3^4$$

es homogéneo de grado 4.

**Teorema 1.4.3.** *Sea $\mathbb{F}$ un cuerpo. Entonces cualquier factor de un polinomio homogéneo en $\mathbb{F}[x_1, \ldots, x_n]$ debe ser homogéneo también.*

Demostración. Es suficiente probar que si $f \in \mathbb{F}[x_1, \ldots, x_n]$ es un polinomio no homogéneo, entonces, para cualquier polinomio $g = g(x_1, \ldots, x_n) \neq 0$, $fg$ tampoco es homogéneo. Podemos escribir $f$ en la forma siguiente

$$f = \sum_{i=1}^{r} f_{n_i}$$

siendo cada $f_{n_i} \neq 0$ homogéneo de grado $n_i$, con $0 \leq n_1 < \cdots < n_r$ y $r \geq 2$, puesto que $f$ no es homogéneo. Análogamente, escribimos

$$g = \sum_{i=1}^{s} g_{m_i},$$

siendo $g_{m_i} \neq 0$ homogéneo de grado $m_i$, con $0 \leq m_1 < \cdots < m_s$ y $s \geq 1$. Entonces

$$fg = f_{n_1} g_{m_1} + \cdots + f_{n_r} g_{m_s}, \tag{1.7}$$

donde cada uno de los términos de esta suma es homogéneo. Claramente

$$\partial(f_{n_1} g_{m_1}) = n_1 + m_1 < n_r + m_s = \partial(f_{n_r} g_{m_s})$$

y los sumandos intermedios en (1.7) deben tener grados estrictamente comprendidos entre $n_1 + m_1$ y $n_r + m_s$. Pero entonces $fg$ no es una suma de monomios en $x_1, \ldots, x_n$ todos del mismo grado, así que $fg$ es no homogéneo.

□

*Observación* 1.4.4. Si $f(\lambda) \in \mathbb{F}[\lambda]$ es el polinomio mónico de grado $k$

$$f(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + a_{k-2}\lambda^{k-2} + \cdots + a_1\lambda + a_0,$$

entonces $\mu^k f(\frac{\lambda}{\mu})$ resulta ser un polinomio de $\mathbb{F}[\lambda, \mu]$ homogéneo de grado $k$, con un monomio donde no aparece $\mu$, a saber $\lambda^k$, y otro donde no aparece $\lambda$, a saber $a_0\mu^k$.

Recíprocamente, todo polinomio homogéneo $g(\lambda, \mu) \in \mathbb{F}[\lambda, \mu]$ de grado (total) $m + k$ y de grado $k$ respecto de $\lambda$ se puede escribir en la forma

$$g(\lambda, \mu) = \mu^m \mu^k g(\frac{\lambda}{\mu}, 1), \tag{1.8}$$

donde $\partial(g(\lambda, 1))$ es $k$ y el polinomio $\mu^k g(\frac{\lambda}{\mu}, 1)$ es homogéneo de grado $k$ con un monomio de la forma $a\lambda^k$ siendo $a \in \mathbb{F}$.

**Ejemplo 1.4.5.** Sea $g(\lambda, \mu) \in \mathbb{C}[\lambda, \mu]$ dado por

$$g(\lambda, \mu) := 2\lambda^5\mu^3 - 7\lambda^4\mu^4 + 11\lambda^3\mu^5 + 17\lambda^2\mu^6 + 8\lambda\mu^7 - 4\mu^8.$$

Está claro que $g(\lambda, \mu)$ es un polinomio homogéneo de grado 8 y de grado 5 respecto de $\lambda$. Obviamente,

$$g(\lambda, 1) = 2\lambda^5 - 7\lambda^4 + 11\lambda^3 + 17\lambda^2 + 8\lambda - 4,$$

luego $g(\frac{\lambda}{\mu}, 1) = 2\frac{\lambda^5}{\mu^5} - 7\frac{\lambda^4}{\mu^4} + 11\frac{\lambda^3}{\mu^3} + 17\frac{\lambda^2}{\mu^2} + 8\frac{\lambda}{\mu} - 4$. Por lo tanto

$$\mu^5 g(\frac{\lambda}{\mu}, 1) = 2\lambda^5 - 7\lambda^4\mu + 11\lambda^3\mu^2 + 17\lambda^2\mu^3 + 8\lambda\mu^4 - 4\mu^5$$

y

$$\mu^3\mu^5 g(\frac{\lambda}{\mu}, 1) = 2\lambda^5\mu^3 - 7\lambda^4\mu^4 + 11\lambda^3\mu^5 + 17\lambda^2\mu^6 + 8\lambda\mu^7 - 4\mu^8.$$

Con lo cual tenemos $g(\lambda, \mu) = \mu^3\mu^5 g(\frac{\lambda}{\mu}, 1)$.

**Teorema 1.4.6.** *Sea $\mathbb{F}$ un cuerpo.*

*1) Si $p = p(\lambda,\mu) \in \mathbb{F}[\lambda,\mu]$ es un polinomio irreducible homogéneo, distinto de $\mu$, con grado $d := \partial(p(\lambda,\mu))$, entonces el polinomio $p(\lambda,1) \in \mathbb{F}[\lambda]$ es irreducible en $\mathbb{F}[\lambda]$, de grado $d$.*

*2) Recíprocamente, si $p(\lambda) \in \mathbb{F}[\lambda]$ es un polinomio irreducible de grado $d$, entonces $f(\lambda,\mu) := \mu^d p(\frac{\lambda}{\mu})$ es un polinomio homogéneo irreducible en $\mathbb{F}[\lambda,\mu]$, de grado (total) $d$.*

DEMOSTRACIÓN. 1) En efecto, si $k$ es el grado de $p(\lambda,1)$, por la Observación 1.4.4 tenemos

$$p(\lambda,\mu) = \mu^{d-k}\mu^k p(\frac{\lambda}{\mu},1),$$

siendo $\mu^k p(\frac{\lambda}{\mu},1) \in \mathbb{F}[\lambda,\mu]$; luego, de la irreducibilidad de $p$ sigue $d = k$. Para probar que $p(\lambda,1)$ es irreducible en $\mathbb{F}[\lambda]$, supongamos que existen $f_i(\lambda) \in \mathbb{F}[\lambda]$, con $d_i := \partial(f_i(\lambda))$ $(i = 1,2)$, tales que

$$p(\lambda,1) = f_1(\lambda)f_2(\lambda), \tag{1.9}$$

luego ha de ser $d = d_1 + d_2$. Ahora, por la Observación 1.4.4, de (1.9) obtenemos la factorización siguiente en $\mathbb{F}[\lambda,\mu]$

$$p(\lambda,\mu) = \mu^{d_1+d_2}p(\frac{\lambda}{\mu},1) = [\mu^{d_1}f_1(\frac{\lambda}{\mu})][\mu^{d_2}f_2(\frac{\lambda}{\mu})].$$

Ahora, de la irreducibilidad de $p(\lambda,\mu)$ se sigue que alguno de los dos factores $[\mu^{d_i}f_i(\frac{\lambda}{\mu})] \in \mathbb{F}[\lambda,\mu]$ $(i = 1,2)$ es una constante no nula; luego, $d_1 = 0$ o $d_2 = 0$. Así pues, queda probado que $f_1(\lambda)$ o $f_2(\lambda)$ es una constante no nula.

Para probar 2), supongamos que existen polinomios $g_i(\lambda,\mu) \in \mathbb{F}[\lambda,\mu]$, con $d_i := \partial(g_i(\lambda,\mu))$ $(i = 1,2)$, tales que

$$f(\lambda,\mu) = g_1(\lambda,\mu)g_2(\lambda,\mu). \tag{1.10}$$

Por la Observación 1.4.4, se tiene que $f(\lambda,\mu)$ es homogéneo y tiene grado $d$; luego $d = d_1 + d_2$. Ahora haciendo $\mu = 1$ en (1.10) tenemos

$$f(\lambda,1) = p(\lambda) = g_1(\lambda,1)g_2(\lambda,1),$$

y de ahí

$$d = \partial(p(\lambda)) = \partial(g_1(\lambda,1)) + \partial(g_2(\lambda,1)).$$

Como $p(\lambda)$ es irreducible en $\mathbb{F}[\lambda]$, algún de los dos divisores de $p(\lambda)$ es una constante no nula, digamos $g_2(\lambda,1)$, de donde sigue

$$d = \partial(g_1(\lambda,1)) \le d_1 \le d;$$

y por consiguiente, $d = d_1$ y $d_2 = 0$, lo que significa que $g_2(\lambda,\mu)$ es una constante no nula. Así pues, queda probado que $f(\lambda,\mu)$ es irreducible.

□

### 1.4.1. Sustitución lineal de indeterminadas

Sea $\mathbb{F}$ un cuerpo. Sean $\alpha,\beta,\gamma,\delta \in \mathbb{F}$ tales que la matriz $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ es no singular. Dado un polinomio $f(\lambda,\mu) \in \mathbb{F}[\lambda,\mu]$ hagamos en él la sustitución de $\lambda$ y $\mu$ por los polinomios $\alpha\tilde{\lambda}+\beta\tilde{\mu}, \gamma\tilde{\lambda}+\delta\tilde{\mu}$ pertenecientes a $\mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ en dos nuevas indeterminadas:

$$\left\{\begin{array}{rcl}\lambda & = & \alpha\tilde{\lambda}+\beta\tilde{\mu}\\ \mu & = & \gamma\tilde{\lambda}+\delta\tilde{\mu}\end{array}\right. . \tag{1.11}$$

Resulta así

$$\tilde{f}(\tilde{\lambda},\tilde{\mu}) := f(\alpha\tilde{\lambda}+\beta\tilde{\mu}, \gamma\tilde{\lambda}+\delta\tilde{\mu}) \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}].$$

Sea ahora

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix} := \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{-1}.$$

Dado un polinomio $\tilde{g}(\tilde{\lambda},\tilde{\mu}) \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ hagamos en él la sustitución de las indeterminadas $\tilde{\lambda}$ y $\tilde{\mu}$ por los polinomios $a\lambda+b\mu, c\lambda+d\mu \in \mathbb{F}[\lambda,\mu]$:

$$\left\{\begin{array}{rcl}\tilde{\lambda} & = & a\lambda+b\mu\\ \tilde{\mu} & = & c\lambda+d\mu\end{array}\right. \tag{1.12}$$

resultando

$$g(\lambda,\mu) := \tilde{g}(a\lambda+b\mu, c\lambda+d\mu) \in \mathbb{F}[\lambda,\mu].$$

Estas sustituciones son involutivas como se ve en la proposición siguiente.

**Proposición 1.4.7.** *Siguiendo con la notación anterior.*

*(1) Dado $f(\lambda,\mu) \in \mathbb{F}[\lambda,\mu]$. Si $\tilde{f}(\tilde{\lambda},\tilde{\mu}) := f(\alpha\tilde{\lambda}+\beta\tilde{\mu}, \gamma\tilde{\lambda}+\delta\tilde{\mu})$, entonces*

$$\tilde{f}(a\lambda+b\mu, c\lambda+d\mu) = f(\lambda,\mu).$$

*(2) Recíprocamente, dado $\tilde{g}(\tilde{\lambda},\tilde{\mu}) \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$. Si $g(\lambda,\mu) := \tilde{g}(a\lambda+b\mu, c\lambda+d\mu)$, entonces*

$$g(\alpha\tilde{\lambda}+\beta\tilde{\mu}, \gamma\tilde{\lambda}+\delta\tilde{\mu}) = \tilde{g}(\tilde{\lambda},\tilde{\mu}).$$

Demostración. (1) Como

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = I_2,$$

se verifica

$$\tilde{f}(a\lambda+b\mu, c\lambda+d\mu) = f(\alpha[a\lambda+b\mu]+\beta[c\lambda+d\mu], \gamma[a\lambda+b\mu]+\delta[c\lambda+d\mu])$$

$$= f((\alpha a+\beta c)\lambda+(\alpha b+\beta d)\mu, (\gamma a+\delta c)\lambda+(\gamma b+\delta d)\mu) = f(\lambda,\mu).$$

(2) Análogamente, como

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} = I_2,$$

resulta que

$$g(\alpha\tilde{\lambda}+\beta\tilde{\mu},\gamma\tilde{\lambda}+\delta\tilde{\mu}) = \tilde{g}(a[\alpha\tilde{\lambda}+\beta\tilde{\mu}]+b[\gamma\tilde{\lambda}+\delta\tilde{\mu}], c[\alpha\tilde{\lambda}+\beta\tilde{\mu}]+d[\gamma\tilde{\lambda}+\delta\tilde{\mu}])$$

$$= \tilde{g}((a\alpha+b\gamma)\tilde{\lambda}+(a\beta+b\delta)\tilde{\mu}, (c\alpha+d\gamma)\tilde{\lambda}+(c\beta+d\delta)\tilde{\mu}) = \tilde{g}(\tilde{\lambda},\tilde{\mu}),$$

como era requerido.

□

De este modo establecemos dos aplicaciones biyectivas

$$\begin{array}{rccl} \Phi : & \mathbb{F}[\lambda,\mu] & \longrightarrow & \mathbb{F}[\tilde{\lambda},\tilde{\mu}] \\ & f & \mapsto & \Phi(f) := \tilde{f} \end{array}$$

$$\begin{array}{rccl} \Psi : & \mathbb{F}[\tilde{\lambda},\tilde{\mu}] & \longrightarrow & \mathbb{F}[\lambda,\mu] \\ & \tilde{g} & \mapsto & \Psi(\tilde{g}) := g \end{array}$$

para las que es cierta la proposición siguiente.

**Proposición 1.4.8.** *En las condiciones anteriores se verifica:*

*1)* $\Phi$ *y* $\Psi$ *son isomorfismos de anillos.*

*2)* $\Phi^{-1} = \Psi$.

*3) Para todo* $f \in \mathbb{F}[\lambda,\mu]$, $\partial(\Phi(f)) = \partial(f)$, *donde* $\partial(f)$ *denota el grado de* $f$, *y para todo* $\tilde{f} \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$, $\partial(\Psi(\tilde{f})) = \partial(\tilde{f})$.

*4)* $p \in \mathbb{F}[\lambda,\mu]$ *es irreducible si y sólo si* $\Phi(p) := \tilde{p} \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ *es irreducible.*

*5)* $p \in \mathbb{F}[\lambda,\mu]$ *es homogéneo si y sólo si* $\Phi(p) := \tilde{p} \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ *es homogéneo.*

DEMOSTRACIÓN. 1) El hecho de que $\Phi$ (y $\Psi$) sea homomorfismo es debido a las propiedades del cálculo del valor de la suma $f_1 + f_2$ y del producto $f_1 f_2$ de dos polinomios $f_1, f_2 \in \mathbb{F}[\lambda,\mu]$ en el punto $(\alpha\tilde{\lambda}+\beta\tilde{\mu},\gamma\tilde{\lambda}+\delta\tilde{\mu}) \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]^2$. Además $\Phi(1) = 1$.

2) y 5) son evidentes. Vamos a probar 3). En efecto, por la definición de $\Phi(f)$ es claro que

$$\partial(\Phi(f)) \leq \partial(f). \tag{1.13}$$

Análogamente se verifica

$$\partial(\Psi(\tilde{g})) \leq \partial(\tilde{g}). \tag{1.14}$$

Como se tiene que $\Psi(\Phi(f)) = f$, basta aplicar (1.14) y (1.13) y sigue

$$\partial(f) \leq \partial(\Phi(f)) \leq \partial(f),$$

así pues $\partial(f) = \partial(\Phi(f))$, y del mismo modo se obtiene también $\partial(\tilde{f}) = \partial(\Psi(\tilde{f}))$.

Para probar 4), supongamos que $p$ es irreducible y que $\tilde{p}$ no lo fuera. Entonces habría polinomios $\tilde{f}, \tilde{g} \in \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ de grado $\geq 1$ tales que $\tilde{p} = \tilde{f}\tilde{g}$, de donde sigue

$$p = \Psi(\tilde{p}) = \Psi(\tilde{f})\Psi(\tilde{g}),$$

con $\partial(\Psi(\tilde{f})) \geq 1$ y $\partial(\Psi(\tilde{g})) \geq 1$; lo que es imposible, pues $p$ es irreducible.

Por un razonamiento simétrico es inmediato que si $\tilde{p}$ es irreducible, entonces $p = \Psi(\tilde{p}))$ también lo es.

□

## 1.5. Matrices polinómicas

En lo que sigue, $\mathbb{F}$ denotará un cuerpo cualquiera, $\mathbb{F}[\lambda]$ el anillo de polinomios con coeficientes en $\mathbb{F}$ y $\mathbb{F}(\lambda)$ su cuerpo de fracciones, esto es, $\mathbb{F}(\lambda) := \{\frac{p(\lambda)}{q(\lambda)} \mid p(\lambda), q(\lambda) \in \mathbb{F}[\lambda], q(\lambda) \neq 0\}$. Además, llamaremos espacio vectorial racional a cualquier $\mathbb{F}(\lambda)$–subespacio vectorial de $\mathbb{F}(\lambda)^m$.

**Definición 1.5.1.** Se llama *grado de un vector polinómico*

$$x(\lambda) = \begin{pmatrix} p_1(\lambda) \\ \vdots \\ p_m(\lambda) \end{pmatrix} \in \mathbb{F}[\lambda]^{m\times 1},$$

que denotaremos por $\partial(x(\lambda))$, al máximo de los grados de todas las componentes del vector, es decir,

$$\partial(x(\lambda)) := \text{máx}\{\partial(p_i(\lambda)) \mid i = 1, \dots, m\}.$$

(Consideramos que el polinomio nulo tiene grado $-\infty$.)

**Proposición 1.5.2.** *Sea $x(\lambda) \in \mathbb{F}[\lambda]^{m\times 1}$ y sea $U \in \mathbb{F}^{m\times m}$ una matriz constante y no singular. Entonces $\partial(x(\lambda)) = \partial(Ux(\lambda))$.*

Demostración. Llamamos $z(\lambda) := Ux(\lambda)$. Por las propiedades del grado de un polinomio, es fácil ver que el grado de cada componente de $z(\lambda)$ es menor o igual que $\partial(x(\lambda))$, y de ahí se sigue $\partial(z(\lambda)) \leq \partial(x(\lambda))$. Análogamente, como también $x(\lambda) := U^{-1}z(\lambda)$, por la misma razón se tiene $\partial(x(\lambda)) \leq \partial(z(\lambda))$, y por tanto la igualdad.

□

**Definición 1.5.3.** Dada la matriz polinómica $D(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$, donde $k_i$ es el grado de la columna $i$–ésima para $i = 1, \dots, n$, se define $D_c \in \mathbb{F}^{m\times n}$ como la matriz cuya columna $i$–ésima está formada por los coeficientes de $\lambda^{k_i}$ en la misma columna de $D(\lambda)$. Llamaremos a $D_c$ la matriz de los coeficientes principales por columnas de $D(\lambda)$. Además, se dice que $D(\lambda)$ es *reducida por columnas o propia por columnas* si rg $D(\lambda) = n$ sobre el cuerpo $\mathbb{F}(\lambda)$ y el rango de $D_c$ por columnas es completo, es decir, si rg $D_c = n$.

*Observación* 1.5.4. De la definición anterior se deduce la descomposición siguiente:

$$D(\lambda) = D_c \begin{pmatrix} \lambda^{k_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda^{k_n} \end{pmatrix} + L(\lambda) \tag{1.15}$$

donde la matriz $L(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ verifica que el grado de su columna $i$–ésima es menor que $k_i$, para cada $i = 1, \dots, n$.

**Ejemplo 1.5.5.** Sea

$$D(\lambda) := \begin{pmatrix} 5\lambda^4 + \lambda^3 & -2\lambda^3 + \lambda \\ \lambda^3 + \lambda^2 & \lambda^3 - \lambda \\ 2\lambda^4 - \lambda^2 & \lambda^2 \end{pmatrix}.$$

Usando la notación anterior, tenemos $k_1 = 4$, $k_2 = 3$ y

$$D_c = \begin{pmatrix} 5 & -2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}.$$

Y como rg $D_c = 2$, se sigue que $D(\lambda)$ es reducida por columnas. Además, comprobamos la correspondiente descomposición (1.15) en este caso:

$$\begin{pmatrix} 5\lambda^4 + \lambda^3 & -2\lambda^3 + \lambda \\ \lambda^3 + \lambda^2 & \lambda^3 - \lambda \\ 2\lambda^4 - \lambda^2 & \lambda^2 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} \lambda^4 & 0 \\ 0 & \lambda^3 \end{pmatrix} + \begin{pmatrix} \lambda^3 & \lambda \\ \lambda^3 + \lambda^2 & -\lambda \\ -\lambda^2 & \lambda^2 \end{pmatrix}.$$

Dada una matriz polinómica de rango completo por columnas, pero no reducida por columnas, mediante operaciones elementales sucesivas sobre las columnas en el anillo $\mathbb{F}[\lambda]$, o equivalentemente multiplicando a la derecha por matrices unimodulares, que rebajan el grado de algunas columnas, se llega a obtener otra matriz polinómica de rango completo por columnas y reducida por columnas. Para que se pueda entender mejor en qué consiste este procedimiento, damos el ejemplo siguiente.

**Ejemplo 1.5.6.** Consideramos la matriz siguiente

$$D(\lambda) := \begin{pmatrix} \lambda^4 + 2\lambda & -\lambda^3 + \lambda^2 \\ \lambda^2 & \lambda + 2 \\ 0 & 1 \end{pmatrix},$$

que tiene rango 2, pero no es reducida por columnas, porque su matriz de coeficientes principales

$$D_c = \begin{pmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

tiene rango menor que 2. Ahora, sumando a la primera columna de $D(\lambda)$ la segunda multiplicada por $\lambda$, el grado de la primera columna baja de 4 a 3, como se puede ver

$$D(\lambda) \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} = \begin{pmatrix} \lambda^3 + 2\lambda & -\lambda^3 + \lambda^2 \\ 2\lambda^2 + 2\lambda & \lambda + 2 \\ \lambda & 1 \end{pmatrix}.$$

Como la matriz resultante tampoco es reducida por columnas, sumamos a la primera columna la segunda, con lo que su grado baja a 2, y queda

$$D(\lambda) \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \lambda^2 + 2\lambda & -\lambda^3 + \lambda^2 \\ 2\lambda^2 + 3\lambda + 2 & \lambda + 2 \\ \lambda + 1 & 1 \end{pmatrix},$$

siendo la correspondiente matriz de los coeficientes principales

$$\begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 0 \end{pmatrix}$$

de rango 2, como era requerido.

El resultado siguiente es conocido por *la propiedad del grado predecible* .

**Proposición 1.5.7.** *Sea $D(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ con rango por columnas completo, es decir $n$, y cuyos grados de las columnas son respectivamente $k_1, \ldots, k_n$ (por tanto, cada $k_i \geq 0$). Entonces, $D(\lambda)$ es reducida por columnas si y sólo si para todo*

$$p(\lambda) := \begin{pmatrix} p_1(\lambda) \\ \vdots \\ p_n(\lambda) \end{pmatrix} \in \mathbb{F}[\lambda]^{n\times 1}$$

*se verifica*

$$\partial(D(\lambda)p(\lambda)) = \text{máx}\{\partial(\lambda^{k_i} p_i(\lambda)) \mid i = 1, \ldots, n\}.$$

Demostración. Primero consideramos $D(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ cualquiera y $p(\lambda)$ cualquier vector polinómico no nulo, pues si no el resultado es trivial. Llamamos $q(\lambda) := D(\lambda)p(\lambda)$ y $d := \text{máx}\{\partial(\lambda^{k_i} p_i(\lambda)) \mid i = 1, \ldots, n\}$. Denotamos por $D(\lambda)^i$ la $i$–ésima columna de $D(\lambda)$ para $i = 1, \ldots, n$. Entonces, se tiene

$$q(\lambda) = p_1(\lambda)D(\lambda)^1 + \cdots + p_n(\lambda)D(\lambda)^n,$$

de donde sigue

$$\begin{array}{rcl} \partial(q(\lambda)) & \leq & \text{máx}\{\partial(p_i(\lambda)D(\lambda)^i) \mid i = 1, \ldots, n\} \\ & = & \text{máx}\{\partial(p_i(\lambda)) + k_i \mid i = 1, \ldots, n\} = d. \end{array}$$

Ahora nos queda probar que se da la igualdad, supuesto que $D(\lambda)$ es de rango completo por columnas y reducida por columnas. Por la igualdad (1.15), se tiene

$$q(\lambda) = D_c \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} + L(\lambda)p(\lambda). \tag{1.16}$$

Como la $i$–ésima columna de $L(\lambda)$ tiene grado menor que $k_i$ ($i = 1, \ldots, n$), es claro que $\partial(L(\lambda)p(\lambda)) < d$. Por tanto, de (1.16) se sigue que $\partial(q(\lambda)) = d$ si y sólo si el vector polinómico

$$D_c \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix}$$

tiene grado exactamente $d$. Como $\text{rg}(D_c) = n$, es conocido que existen matrices $U \in \mathbb{F}^{m\times m}$ y $V \in \mathbb{F}^{n\times n}$ no singulares tal que $D_c = U \begin{bmatrix} I_n \\ 0 \end{bmatrix} V$, donde $\begin{bmatrix} I_n \\ 0 \end{bmatrix} \in \mathbb{F}^{m\times n}$. Luego,

$$D_c \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} = U \begin{bmatrix} I_n \\ 0 \end{bmatrix} V \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} = U \begin{bmatrix} V \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} \\ 0 \end{bmatrix}$$

que, según Proposición 1.5.2, tiene el mismo grado que

$$\begin{bmatrix} V \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} \\ 0 \end{bmatrix},$$

el cual, por la misma razón, coincide con el grado de

$$\begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix},$$

que es claramente $d$.

Para probar el recíproco, lo hacemos por reducción al absurdo. Supongamos lo contrario, que $D(\lambda)$ no es reducida por columnas y vamos a llegar a contradicción con la hipótesis comprobando que existe un vector polinómico $p(\lambda)$ que no la cumple. Denotamos por $D_c^1, \ldots, D_c^n$ las columnas de $D_c$, la matriz de los coeficientes principales de $D(\lambda)$. Como rg $D_c < n$, existe una columna que es $\mathbb{F}$–combinación lineal de las otras, digamos $D_c^j = \sum_{i \neq j} a_i D_c^i$, para ciertos $a_i \in \mathbb{F}$, $i = 1, \ldots, n$. Supuesto que $k_l := \text{máx}\{k_1, \ldots, k_n\}$, elegimos $p_i(\lambda) := a_i \lambda^{k_l - k_i}$ para cada $i = 1, \ldots, n$, con $i \neq j$, y $p_j(\lambda) := -\lambda^{k_l - k_j}$. Entonces, usando la igualdad (1.15), para este vector $p(\lambda) \neq 0$ tenemos

$$D(\lambda)p(\lambda) = D_c \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} + L(\lambda)p(\lambda) = L(\lambda)p(\lambda),$$

ya que

$$\begin{aligned} D_c \begin{pmatrix} \lambda^{k_1} p_1(\lambda) \\ \vdots \\ \lambda^{k_n} p_n(\lambda) \end{pmatrix} &= \lambda^{k_1} p_1(\lambda) D_c^1 + \cdots + \lambda^{k_n} p_n(\lambda) D_c^n \\ &= \lambda^{k_l} (\sum_{i \neq j} a_i D_c^i) - \lambda^{k_l} D_c^j = 0. \end{aligned}$$

Así pues, se concluye

$$\partial(D(\lambda)p(\lambda)) = \partial(L(\lambda)p(\lambda)) < d := \text{máx}\{\partial(\lambda^{k_i} p_i(\lambda)) \mid i = 1, \ldots, n\},$$

aplicando la primera parte de la demostración a $L(\lambda)p(\lambda)$ y por ser la columna $i$–ésima de $L(\lambda)$ de grado menor que $k_i$ para $i = 1, \ldots, n$.

□

**Proposición 1.5.8.** *Sea $F(\lambda) \in \mathbb{F}[\lambda]^{m \times r}$ una matriz polinómica de rango completo por columnas ($r \leq m$). El sistema lineal $F(\lambda)x(\lambda) = p(\lambda)$ tiene una solución polinómica $x(\lambda) \in \mathbb{F}[\lambda]^{m \times 1}$ para cada $p(\lambda) \in \mathbb{F}[\lambda]^{m \times 1}$ perteneciente al espacio vectorial racional generado por las columnas de $F(\lambda)$ si y sólo si todos los factores invariantes de $F(\lambda)$ son triviales (i.e., iguales a 1).*

Demostración. Sean $U(\lambda) \in \mathbb{F}[\lambda]^{m\times m}$ y $V(\lambda) \in \mathbb{F}[\lambda]^{r\times r}$ matrices unimodulares (es decir, matrices cuyo determinante es una constante no nula) tales que $U(\lambda)F(\lambda)V(\lambda) = D(\lambda)$ es la forma normal de Smith de $F(\lambda)$, esto es,

$$U(\lambda)F(\lambda)V(\lambda) = \begin{bmatrix}\operatorname{diag}(d_1(\lambda),\dots,d_r(\lambda))\\ 0\end{bmatrix}$$

donde $d_1(\lambda) \mid \cdots \mid d_r(\lambda)$ son polinomios mónicos y además $d_r(\lambda) \neq 0$, ya que $\operatorname{rg} D(\lambda) = \operatorname{rg} F(\lambda) = r$.

Dado que $U(\lambda)$ y $V(\lambda)$ son matrices unimodulares, es claro que

$$F(\lambda)x(\lambda) = p(\lambda)$$

tiene una solución polinómica o racional si y sólo si el sistema

$$D(\lambda)y(\lambda) = q(\lambda), \tag{1.17}$$

con $q(\lambda) := U(\lambda)p(\lambda)$, tiene una solución polinómica o racional, respectivamente. De hecho, las soluciones de ambos sistemas están relacionadas mediante $x(\lambda) = V(\lambda)y(\lambda)$. Así pues, el sistema $F(\lambda)x(\lambda) = p(\lambda)$ tiene una solución polinómica para cada $p(\lambda) \in \mathbb{F}[\lambda]^{m\times 1} \cap \operatorname{Im} F(\lambda)$ si y sólo si el sistema $D(\lambda)y(\lambda) = q(\lambda)$ tiene una solución polinómica para cada $q(\lambda) \in \mathbb{F}[\lambda]^{m\times 1} \cap \operatorname{Im} D(\lambda)$.

Por ser cero las $m-r$ filas últimas de $D(\lambda)$ para que (1.17) tenga solución (en $\mathbb{F}(\lambda)^{m\times 1}$) es necesario que

$$q(\lambda) = \big(q_1(\lambda),\dots,q_r(\lambda),0,\dots,0\big)^{\mathrm{T}}; \tag{1.18}$$

por ser $D(\lambda)$ cuasi-diagonal la condición (1.18) es también suficiente. Por consiguiente,

$$\operatorname{Im} D(\lambda) = \left\{ \begin{pmatrix} s_1(\lambda)\\ \vdots \\ s_r(\lambda) \\ 0 \\ \vdots \\ 0\end{pmatrix} \in \mathbb{F}(\lambda)^{m\times 1} \right\};$$

luego,

$$\mathbb{F}[\lambda]^{m\times 1} \cap \operatorname{Im} D(\lambda) = \left\{ \begin{pmatrix} q_1(\lambda)\\ \vdots \\ q_r(\lambda) \\ 0 \\ \vdots \\ 0\end{pmatrix} \in \mathbb{F}[\lambda]^{m\times 1} \right\}.$$

Si $d_i(\lambda) = 1$ para $i = 1,\dots,r$, es claro que para todo

$$q(\lambda) = \big(q_1(\lambda),\dots,q_r(\lambda),0,\dots,0\big)^{\mathrm{T}} \in \mathbb{F}[\lambda]^{m\times 1},$$

el sistema

$$D(\lambda)y(\lambda) = q(\lambda)$$

admite la solución polinómica $y(\lambda) = \big(q_1(\lambda), \dots, q_r(\lambda)\big)$. Recíprocamente, si para todos $q_1(\lambda), \dots, q_r(\lambda) \in \mathbb{F}[\lambda]$ el sistema

$$D(\lambda) \begin{pmatrix} y_1(\lambda) \\ \vdots \\ y_r(\lambda) \end{pmatrix} = \begin{pmatrix} q_1(\lambda) \\ \vdots \\ q_r(\lambda) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

admite una solución polinómica, necesariamente deberá ser

$$d_i(\lambda) y_i(\lambda) = q_i(\lambda), \quad \forall i = 1, \dots, r.$$

Por lo tanto, $d_i(\lambda) | q_i(\lambda)$; y esto para cualquier polinomio $q_i(\lambda) \in \mathbb{F}[\lambda]$. Por consiguiente, $d_i(\lambda) = 1$ para $i = 1, \dots, r$.

$\square$

**Corolario 1.5.9.** *Sea $\mathbb{F}$ un cuerpo algebraicamente cerrado y sea $F(\lambda) \in \mathbb{F}[\lambda]^{m \times r}$ una matriz polinómica de rango completo por columnas ($r \leq m$). Entonces, todos los factores invariantes de $F(\lambda)$ son triviales si y sólo si* $\operatorname{rg} F(a) = r$ *para todo $a \in \mathbb{F}$.*

Demostración. Sea $U(\lambda)F(\lambda)V(\lambda) = D(\lambda) = \begin{bmatrix} \operatorname{diag}(d_1(\lambda), \dots, d_r(\lambda)) \\ 0 \end{bmatrix}$ la forma de Smith de $F(\lambda)$, como en la demostración de la proposición anterior. Como $\det U(a) \neq 0$ y $\det V(a) \neq 0$, para todo $a \in \mathbb{F}$, se tiene que $\operatorname{rg} F(a) = \operatorname{rg} D(a)$ para todo $a \in \mathbb{F}$. Finalmente, puesto que $\mathbb{F}$ es algebraicamente cerrado, es claro que $\operatorname{rg} D(a) = r$ para todo $a \in \mathbb{F}$ si y sólo si $d_1(\lambda), \dots, d_r(\lambda)$ no tienen raíces en $\mathbb{F}$; lo que equivale a $d_i(\lambda) = 1$ para $i = 1, \dots, r$.

$\square$

**Definición 1.5.10.** Si $\mathbb{F}$ es un cuerpo algebraicamente cerrado y $r \leq m$, se dice que una matriz $F(\lambda) \in \mathbb{F}[\lambda]^{m \times r}$ es *irreducible* si para todo $a \in \mathbb{F}$, $\operatorname{rg} F(a) = r$.

## 1.6. Espacios vectoriales racionales

En lo que sigue, $\mathbb{F}$ denotará un cuerpo cualquiera, $\mathbb{F}[\lambda]$ el anillo de polinomios con coeficientes en $\mathbb{F}$ y $\mathbb{F}(\lambda)$ su cuerpo de fracciones, esto es, $\mathbb{F}(\lambda) := \{\frac{p(\lambda)}{q(\lambda)} \mid p(\lambda), q(\lambda) \in \mathbb{F}[\lambda], q(\lambda) \neq 0\}$. Además, llamaremos espacio vectorial racional a cualquier $\mathbb{F}(\lambda)$–subespacio de $\mathbb{F}(\lambda)^m$.

**Proposición 1.6.1.** *Los vectores $x_1(\lambda), \dots, x_r(\lambda) \in \mathbb{F}[\lambda]^{n \times 1}$ son $\mathbb{F}[\lambda]$–linealmente dependientes si y sólo si son $\mathbb{F}(\lambda)$–linealmente dependientes.*
*Equivalentemente, se tiene que $x_1(\lambda), \dots, x_r(\lambda) \in \mathbb{F}[\lambda]^{n \times 1}$ son $\mathbb{F}[\lambda]$–linealmente independientes si y sólo si son $\mathbb{F}(\lambda)$–linealmente independientes.*

Demostración. Como $\mathbb{F}[\lambda] \subseteq \mathbb{F}(\lambda)$, está claro que si $x_1(\lambda), \dots, x_r(\lambda) \in \mathbb{F}[\lambda]^{n \times 1}$ son $\mathbb{F}[\lambda]$–linealmente dependientes, entonces son $\mathbb{F}(\lambda)$–linealmente dependientes.

Recíprocamente, supongamos que existen fracciones $\frac{p_i(\lambda)}{q_i(\lambda)} \in \mathbb{F}(\lambda)(i = 1, \ldots, r)$, no todas nulas, por ejemplo $p_j(\lambda) \neq 0$, tales que

$$\frac{p_1(\lambda)}{q_1(\lambda)} x_1(\lambda) + \cdots + \frac{p_r(\lambda)}{q_r(\lambda)} x_r(\lambda) = 0. \tag{1.19}$$

Sea $q(\lambda) := \text{m.c.m.}\{q_1(\lambda), \ldots, q_r(\lambda)\}$. Entonces para cada $i = 1, \ldots, r$ existen polinomios $\tilde{q_i}(\lambda)$, tales que $q_i(\lambda)\tilde{q_i}(\lambda) = q(\lambda)$. Ahora, multiplicando (1.19) por $q(\lambda)$ queda

$$p_1(\lambda)\tilde{q_1}(\lambda)x_1(\lambda) + \cdots + p_r(\lambda)\tilde{q_r}(\lambda)x_r(\lambda) = 0$$

siendo $p_j(\lambda)\tilde{q_j}(\lambda) \neq 0$. Así pues, $x_1(\lambda), \ldots, x_r(\lambda)$ son $\mathbb{F}[\lambda]$-linealmente dependientes.

□

### 1.6.1. Bases polinómicas

Deseamos esclarecer el proceso de definición de los *índices minimales* de una matriz $H(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ con el fin de aplicarlo después a $H(\lambda) = \lambda B - A$ un haz singular. Para ello, consideremos el sistema de ecuaciones lineales homogéneo

$$H(\lambda)\, x(\lambda) = 0, \quad x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1}. \tag{1.20}$$

Llamemos $S$ al $\mathbb{F}(\lambda)$–subespacio vectorial de $\mathbb{F}(\lambda)^{n\times 1}$ formado por las soluciones de (1.20), también denominado el núcleo por la derecha de $H(\lambda)$, y denotado alternativamente por $\operatorname{Ker} H(\lambda)$. Es conocido que $\dim_{\mathbb{F}(\lambda)} S = n - r$, donde $r$ denota el rango de $H(\lambda)$ sobre el cuerpo $\mathbb{F}(\lambda)$. En adelante será $\alpha := n - r$ y suponemos que $r < n$.

**Proposición 1.6.2.** *En las condiciones anteriores, existen bases polinómicas de $S$.*

DEMOSTRACIÓN. Sea $\{x_1(\lambda), \ldots, x_\alpha(\lambda)\} \subset \mathbb{F}(\lambda)^{n\times 1}$ una base de $S$, donde

$$x_i(\lambda) = \begin{pmatrix} \dfrac{p_{1i}(\lambda)}{q_{1i}(\lambda)} \\ \vdots \\ \dfrac{p_{ni}(\lambda)}{q_{ni}(\lambda)} \end{pmatrix},$$

para $i = 1, \ldots, \alpha$. Entonces, considerando los polinomios no nulos $q_i(\lambda) := \text{m.c.m.}\{q_{1i}(\lambda), \ldots, q_{ni}(\lambda)\}$ $(i = 1, \ldots, \alpha)$, se verifica que

$$\{q_1(\lambda)x_1(\lambda), \cdots, q_\alpha(\lambda)x_\alpha(\lambda)\}$$

constituye una base de $S$ formada por vectores polinómicos.

□

Construiremos ahora una base polinómica de $S$ de una determinada manera iterativamente:

- Entre todas las soluciones polinómicas no nulas de (1.20) elegimos una, $x_1(\lambda)$, de grado mínimo $\varepsilon_1$.
- Consideramos ahora las soluciones polinómicas de (1.20), que no son $\mathbb{F}(\lambda)$–combinación lineal de $x_1(\lambda)$. De entre todas ellas, tomamos una de grado mínimo $\varepsilon_2$. Obviamente, se verifica $\varepsilon_1 \leq \varepsilon_2$.
- Continuando de este modo, en el paso $i$-ésimo ($i < \alpha$) tendremos elegidas las soluciones polinómicas linealmente independientes $x_1(\lambda), \ldots, x_i(\lambda)$ con grados respectivos $\varepsilon_1 \leq \cdots \leq \varepsilon_i$. Como existen soluciones polinómicas $x(\lambda) \notin \langle x_1(\lambda), \ldots, x_i(\lambda)\rangle_{\mathbb{F}(\lambda)}$, elegimos una, $x_{i+1}(\lambda)$, cuyo grado $\varepsilon_{i+1} := \partial(x_{i+1}(\lambda))$ sea el mínimo entre todos esos grados $\partial(x(\lambda))$.
- Evidentemente, este proceso finaliza al llegar a elegir una base polinómica de $S$, a saber,
$$\boldsymbol{\mathcal{B}} := \{x_1(\lambda),\, x_2(\lambda), \ldots,\, x_\alpha(\lambda)\},$$
donde $\varepsilon_i := \partial(x_i(\lambda))$ es el menor grado de todas las soluciones polinómicas de (1.20), $x(\lambda)$, que verifican $x(\lambda) \notin \langle x_1(\lambda), x_2(\lambda), \ldots, x_{i-1}(\lambda)\rangle_{\mathbb{F}(\lambda)}$, para $i = 1, \ldots, \alpha$, y de ahí se sigue
$$0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_\alpha.$$

### 1.6.2. Bases minimales

**Definición 1.6.3.** Una base polinómica $\boldsymbol{\mathcal{B}} := \{x_1(\lambda),\, x_2(\lambda), \ldots,\, x_\alpha(\lambda)\}$ de $S$, elegida según el proceso anterior, se llama una *base polinómica minimal* de $S$.

Naturalmente, procediendo de este modo se pueden encontrar muchas bases polinómicas diferentes, pero todas ellas deben tener algo en común, como lo prueba la proposición siguiente.

**Proposición 1.6.4.** *Siguiendo con la notación anterior, los números $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_\alpha$ son independientes de la base polinómica minimal de $S$ a la que están asociados.*

DEMOSTRACIÓN. Supongamos que, siguiendo el proceso anterior, hemos obtenido otra base polinómica de $S$, $\boldsymbol{\mathcal{B}}' := \{z_1(\lambda),\, z_2(\lambda), \ldots,\, z_\alpha(\lambda)\}$, siendo $\mu_i := \partial(z_i(\lambda))$ para $i = 1, \ldots, \alpha$, y $0 \leq \mu_1 \leq \mu_2 \leq \cdots \leq \mu_\alpha$. Precisemos los lugares donde las sucesiones $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_\alpha$ y $\mu_1, \mu_2, \ldots, \mu_\alpha$ son estrictamente crecientes. Supongamos que
$$\varepsilon_1 = \cdots = \varepsilon_{n_1} < \varepsilon_{n_1+1} = \cdots = \varepsilon_{n_2} < \cdots$$
y análogamente
$$\mu_1 = \cdots = \mu_{m_1} < \mu_{m_1+1} = \cdots = \mu_{m_2} < \cdots$$
Es evidente que $\varepsilon_1 = \mu_1$, por coincidir en ambos casos con el menor grado de todas las soluciones polinómicas no nulas de (1.20) y nos proponemos probar que $n_1 = m_1$ y $\varepsilon_{n_1+1} = \mu_{m_1+1}$. Sabemos que cada $z_i(\lambda)$ es una $\mathbb{F}(\lambda)$–combinación lineal de $\{x_1(\lambda), x_2(\lambda), \ldots, x_\alpha(\lambda)\}$, pero para $i = 1, \ldots, m_1$ podemos precisar aun más, a saber,
$$z_i(\lambda) \in \langle x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda)\rangle_{\mathbb{F}(\lambda)}.$$

En efecto, si para algún $i \in \{1, \ldots, m_1\}$ se tiene que

$$z_i(\lambda) \notin \langle x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda)\rangle_{\mathbb{F}(\lambda)},$$

entonces $\{x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda), z_i(\lambda)\}$ sería $\mathbb{F}(\lambda)$–linealmente independiente y podríamos sustituir $x_{n_1+1}(\lambda)$ por $z_i(\lambda)$ por tener éste grado $\varepsilon_1$ menor que $\varepsilon_{n_1+1} = \partial(x_{n_1+1}(\lambda))$, lo que es contrario al modo de elegir $x_{n_1+1}(\lambda)$. En consecuencia, tenemos

$$\langle z_1(\lambda), z_2(\lambda), \ldots, z_{m_1}(\lambda)\rangle_{\mathbb{F}(\lambda)} \subseteq \langle x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda)\rangle_{\mathbb{F}(\lambda)},$$

y tomando las respectivas dimensiones, se sigue $m_1 \leq n_1$. Por otro lado, invirtiendo los papeles, cada $x_i(\lambda)$ $(i = 1, \ldots, n_1)$ pertenece al subespacio

$$\langle z_1(\lambda),\, z_2(\lambda), \ldots,\, z_{m_1}(\lambda)\rangle_{\mathbb{F}(\lambda)},$$

pues en caso contrario podríamos reemplazar $z_{m_1+1}(\lambda)$ por $x_i(\lambda)$ de grado inferior $\mu_1 = \varepsilon_1$, una contradicción. Así pues, como antes se obtiene $n_1 \leq m_1$ y por tanto la igualdad $n_1 = m_1$. Además, al ser

$$S_{n_1} := \langle x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda)\rangle_{\mathbb{F}(\lambda)} = \langle z_1(\lambda), z_2(\lambda), \ldots, z_{m_1}(\lambda)\rangle_{\mathbb{F}(\lambda)},$$

por la forma de elegir $x_{n_1+1}(\lambda)$ y $z_{m_1+1}(\lambda)$, ambos han de tener el mismo grado, a saber, el mínimo de los grados de los vectores polinómicos de $S$ que no pertenecen a $S_{n_1}$.

De este modo, repitiendo el mismo argumento sucesivamente, se llega a probar que $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_\alpha) = (\mu_1, \mu_2, \ldots, \mu_\alpha)$.

□

**Definición 1.6.5.** Se llaman los *índices minimales de Kronecker por la derecha (o por columnas)* de la matriz racional $H(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ a los términos de la sucesión creciente de enteros $(\varepsilon_1, \ldots, \varepsilon_\alpha)$. Se llaman los *índices minimales de Kronecker por la izquierda (o por filas)* de la matriz racional $H(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ a los índices minimales por la derecha de la matriz transpuesta $H(\lambda)^{\mathrm{T}}$.

**Definición 1.6.6.** Sea $S$ el subespacio definido anteriormente. Entonces, para cualquier base polinómica de $S$, digamos $\boldsymbol{\mathcal{B}}_1 := \{z_1(\lambda),\, z_2(\lambda), \ldots,\, z_\alpha(\lambda)\}$ con $\mu_i := \partial(z_i(\lambda))$ para $i = 1, \ldots, \alpha$, se define el *orden* de $\boldsymbol{\mathcal{B}}_1$, como la suma de los grados de sus vectores, esto es,

$$\partial(\boldsymbol{\mathcal{B}}_1) := \mu_1 + \cdots + \mu_\alpha.$$

Es claro que cualquier base polinómica de $S$, siempre se puede reordenar para que los grados de sus vectores estén en forma creciente. Pues bien, en el Teorema 1.6.11 obtendremos una caracterización muy sencilla de las bases polinómicas minimales cuando el cuerpo $\mathbb{F}$ es algebraicamente cerrado, a saber, “ una base polinómica es minimal (tras una reordenación, en caso necesario) si y sólo si tiene orden minimal”, es decir, si su orden toma el mínimo valor posible al considerar todas las bases polinómicas de $S$.

**Definición 1.6.7.** Diremos que $F(\lambda) \in \mathbb{F}(\lambda)^{n\times\alpha}$ es una *matriz-base* de un subespacio vectorial $W$ de $\mathbb{F}(\lambda)^{n\times 1}$ si las columnas de $F(\lambda)$ forman una base de $W$.

Si $F(\lambda)$ es una matriz polinómica, diremos que es una *matriz-base polinómica* de $W$; sea en este caso $F(\lambda)^i$, la $i$–ésima columna de $F(\lambda)$ para $i = 1, \ldots, \alpha$; si $\{F(\lambda)^1, \ldots, F(\lambda)^\alpha\}$ constituye una base polinómica minimal de $W$, diremos que $F(\lambda)$ es una *matriz-base minimal* de $W$.

Evidentemente, dar una base $\boldsymbol{\mathcal{B}} := \{x_1(\lambda), x_2(\lambda), \ldots, x_\alpha(\lambda)\}$ de $S$, el núcleo por la derecha de $H(\lambda)$, con $\varepsilon_i := \partial(x_i(\lambda))$ $(i = 1, \ldots, \alpha)$, es equivalente a dar una matriz por columnas

$$F(\lambda) := \begin{bmatrix} x_1(\lambda) & x_2(\lambda) & \cdots & x_\alpha(\lambda) \end{bmatrix} \in \mathbb{F}[\lambda]^{n\times\alpha}$$

cuyos grados de las columnas sean $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_\alpha)$, que tenga rango $\alpha$ sobre el cuerpo $\mathbb{F}(\lambda)$, y que verifique $H(\lambda)F(\lambda) = 0$, según (1.20). Ahora consideramos el caso especial donde $H(\lambda)$ es un haz lineal $H(\lambda) := \lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$.

**Proposición 1.6.8.** *Dos haces lineales estrictamente equivalentes tienen la misma secuencia de índices minimales por la derecha y por la izquierda.*

Demostración. Sean $\lambda B - A$ y $U(\lambda B - A)V$ dos haces estrictamente equivalentes en $\mathbb{F}[\lambda]^{m\times n}$, donde $U$ y $V$ son dos matrices no singulares y constantes. Entonces es fácil ver que, si las columnas de la matriz

$$F(\lambda) = \begin{bmatrix} x_1(\lambda) & x_2(\lambda) & \cdots & x_\alpha(\lambda) \end{bmatrix} \in \mathbb{F}[\lambda]^{n\times\alpha},$$

con $\varepsilon_i := \partial(x_i(\lambda))$ $(i = 1, \ldots, \alpha)$, constituyen una base minimal por la derecha de $\lambda B - A$, entonces las columnas de la matriz

$$G(\lambda) = V^{-1}F(\lambda) = \begin{bmatrix} V^{-1}x_1(\lambda) & V^{-1}x_2(\lambda) & \cdots & V^{-1}x_\alpha(\lambda) \end{bmatrix},$$

forman una base minimal por la derecha de $U(\lambda B - A)V$. Además, por la Proposición 1.5.2, se tiene que $\partial(V^{-1}x_i(\lambda)) = \partial(x_i(\lambda))$, $(i = 1, \ldots, \alpha)$. Por tanto, queda probado que ambos haces tienen la misma secuencia de índices minimales por la derecha.

Haciendo esta misma demostración para los haces $\lambda B^{\mathrm{T}} - A^{\mathrm{T}}$ y $V^{\mathrm{T}}(\lambda B^{\mathrm{T}} - A^{\mathrm{T}})U^{\mathrm{T}}$, se deduce que $\lambda B - A$ y $U(\lambda B - A)V$ tienen la misma secuencia de índices minimales por la izquierda.

□

Como consecuencia de la proposición anterior, las secuencias de índices minimales por la derecha y por la izquierda de un haz $\lambda B - A$ coincide con las secuencias de índices minimales por la derecha y por la izquierda de su correspondiente forma canónica de Kronecker.

**Ejemplo 1.6.9.** Sea

$$L_\mu := \begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \end{pmatrix} \in \mathbb{F}[\lambda]^{\mu\times(\mu+1)},$$

que tiene rango $\mu$ sobre el cuerpo $\mathbb{F}(\lambda)$. Observemos que

$$L_\mu \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

De modo que el vector $f(\lambda) := \begin{pmatrix} 1 & \lambda & \cdots & \lambda^{\mu} \end{pmatrix}^{\mathrm{T}}$ constituye una base polinómica de $\operatorname{Ker} L_{\mu}$, que tiene dimensión 1. Veamos que también es una base minimal, esto es, que $f(\lambda)$ es el de menor grado entre los vectores polinómicos no nulos $g(\lambda) \in \mathbb{F}[\lambda]^{(\mu+1)\times 1}$ tales que $L_{\mu} g(\lambda) = 0$. En efecto, sea $g(\lambda) = \begin{pmatrix} g_1(\lambda) & \cdots & g_{\mu+1}(\lambda) \end{pmatrix}^{\mathrm{T}}$ uno de estos vectores polinómicos. Entonces existe un escalar $\frac{p(\lambda)}{q(\lambda)} \in \mathbb{F}(\lambda)$ que satisface la relación siguiente

$$\begin{pmatrix} g_1(\lambda) \\ \vdots \\ g_{\mu+1}(\lambda) \end{pmatrix} = \frac{p(\lambda)}{q(\lambda)} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{\mu} \end{pmatrix}.$$

De ahí sigue directamente que $g_1(\lambda) = \frac{p(\lambda)}{q(\lambda)}$ y no nulo; pues, si no, sería $g(\lambda) = 0$. De este modo tenemos $\partial(g_i(\lambda)) = \partial(g_1(\lambda)) + \partial(\lambda^{i-1})$ para cada $i = 1, \ldots, \mu+1$, y de ahí $\partial(g(\lambda)) \geq \partial(f(\lambda))$, como era requerido.

**Ejemplo 1.6.10.** Sea ahora $\lambda B - A \in \mathbb{C}^{14\times 16}$ la forma canónica de Kronecker siguiente

$$\begin{pmatrix}
0&0&\lambda&-1&&&&&&&&&&&&\\
&&&&\lambda&-1&0&&&&&&&&&\\
&&&&0&\lambda&-1&&&&&&&&&\\
&&&&&&0&0&&&&&&&&\\
&&&&&&&\lambda&0&0&&&&&&\\
&&&&&&&-1&\lambda&0&&&&&&\\
&&&&&&&0&-1&\lambda&&&&&&\\
&&&&&&&0&0&-1&&&&&&\\
&&&&&&&&&&-1&&&&&\\
&&&&&&&&&&&-1&\lambda&&&\\
&&&&&&&&&&&0&-1&&&\\
&&&&&&&&&&&&&\lambda-2&&\\
&&&&&&&&&&&&&&\lambda-3&-1\\
&&&&&&&&&&&&&&0&\lambda-3
\end{pmatrix},$$

o escrito de otro modo,

$$\lambda B - A = \operatorname{diag}(L_0, L_0, L_1, L_2.L_0^{\mathrm{T}}, L_3^{\mathrm{T}}, \lambda N - I_3, \lambda I_3 - J)$$

para $N := \left(\begin{array}{c|cc} 0 & & \\ \hline & 0 & 1 \\ & 0 & 0 \end{array}\right)$, $J := \left(\begin{array}{c|cc} 2 & & \\ \hline & 3 & 1 \\ & 0 & 3 \end{array}\right)$, los índices de Kronecker por columnas son $\mu_1 = 0, \mu_2 = 0, \mu_3 = 1, \mu_4 = 2$, y los de filas son $\nu_1 = 0, \nu_2 = 3$. Como este haz tiene rango 12 sobre el cuerpo $\mathbb{C}(\lambda)$, la dimensión del subespacio $S = \operatorname{Ker}(\lambda B - A)$ es $16 - 12 = 4$. Entonces, una vez visto el ejemplo anterior para $L_{\mu}$, se tiene que una base polinómica para $S$ es la formada por las columnas de la matriz siguiente $F(\lambda)$

$$F(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda \\ 0 & 0 & 0 & \lambda^2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}[\lambda]^{16\times 4},$$

cuyos respectivos grados son precísamente $\mu_1 = 0, \mu_2 = 0, \mu_3 = 1, \mu_4 = 2$. Es fácil ver que $F(\lambda)$ tiene rango completo por columnas, es reducida por columnas y es irreducible (es decir, rg $F(z) = 4$ para todo $z \in \mathbb{C}$), lo que, según probaremos en el teorema siguiente, equivale a decir que la base anterior es minimal y por tanto los índices minimales del núcleo del haz $\lambda B - A$ coinciden con los índices minimales de Kronecker por columnas.

**Teorema 1.6.11.** *Sea $\mathbb{F}$ un cuerpo algebraicamente cerrado. Consideramos una matriz polinómica de rango completo por columnas*

$$F(\lambda) := \begin{bmatrix} f_1(\lambda) & \cdots & f_\alpha(\lambda) \end{bmatrix} \in \mathbb{F}[\lambda]^{n\times\alpha},$$

*con grados de las columnas $\mu_1 \leq \cdots \leq \mu_\alpha$. Entonces las afirmaciones siguientes son equivalentes:*

(i) *Las columnas de $F(\lambda)$ constituyen una base minimal para el espacio vectorial racional generado por ellas.*

(ii) *$F(\lambda)$ es reducida por columnas e irreducible.*

(iii) *Las columnas de $F(\lambda)$ constituyen una base de orden minimal para el espacio vectorial racional generado por ellas.*

Demostración. Demostraremos que (i) $\Rightarrow$ (ii) y (iii) $\Rightarrow$ (ii) a la vez, por reducción al absurdo. Supongamos pues, que $F(\lambda)$ no es reducida por columnas o no es irreducible. Entonces vamos a probar que existe otra base polinómica $\{\overline{f}_1(\lambda), \ldots, \overline{f}_\alpha(\lambda)\} \subset \mathbb{F}[\lambda]^{n\times 1}$ para el mismo espacio, cuyos respectivos grados $\overline{\mu}_i$ $(i = 1, \ldots, \alpha)$ coinciden con los $\mu_i$ para todos los índices $i$ salvo uno, digamos el $j$–ésimo, esto es

$$\overline{\mu}_i = \mu_i \quad \forall\, i \neq j \quad \text{tal que } \ 1 \leq i \leq \alpha \quad \text{y} \quad \overline{\mu}_j < \mu_j,$$

lo que contradice la hipótesis (i) y la (iii).

Si $F(\lambda)$ no es reducida por columnas, entonces haciendo una operación elemental en una sola columna (véase el método en Ejemplo 1.5.6), digamos la $j$–ésima, podemos obtener una nueva matriz $\overline{F}(\lambda) := \begin{bmatrix} \overline{f}_1(\lambda) & \ldots & \overline{f}_\alpha(\lambda) \end{bmatrix}$ donde todas las columnas siguen igual, salvo la $j$–ésima que baja de grado, contradicción con (i) y (iii).

Si $F(\lambda)$ no es irreducible, entonces rg $F(a) < \alpha$ para algún $a \in \mathbb{F}$, y por tanto existen constantes $\{c_i\}_{i=1}^{\alpha}$, no todas nulas, tal que

$$\sum_{i=1}^{\alpha} c_i f_i(a) = 0.$$

De ahí sigue que $\lambda - a$ divide a cada polinomio componente de $\sum_{i=1}^{\alpha} c_i f_i(\lambda) \in \mathbb{F}[\lambda]^{n\times 1}$, y por ello tenemos que

$$\overline{f}(\lambda) := \sum_{i=1}^{\alpha} \frac{c_i f_i(\lambda)}{\lambda - a}$$

es un vector polinómico. Ahora, supongamos que $f_j(\lambda)$ es un vector columna de $F(\lambda)$, que tiene grado máximo entre todos los $f_i(\lambda)$ para los cuales $c_i \neq 0$. Entonces, es claro que $\partial(\overline{f}(\lambda)) < \partial(f_j(\lambda))$. Luego, reemplazando $f_j(\lambda)$ por $\overline{f}(\lambda)$ en la base original formada por las columnas de $F(\lambda)$, resulta una nueva base polinómica de orden menor para el mismo espacio, lo que es una contradicción con (iii) y también con (i).

La demostración de los recíprocos (ii) $\Rightarrow$ (i) y (ii) $\Rightarrow$ (iii) se hará también conjuntamente, por reducción al absurdo y aplicaremos la propiedad del grado predecible (véanse las Proposiciones 1.5.7 y 1.5.8). Supongamos que $F(\lambda)$ es reducida por columnas e irreducible, pero no se verifica (i) o no se verifica (iii). Sea $S$ el subespacio racional generado por las columnas de $F(\lambda)$ y sea

$$G(\lambda) := \begin{bmatrix} g_1(\lambda) & \cdots & g_\alpha(\lambda) \end{bmatrix} \in \mathbb{F}[\lambda]^{n\times\alpha}$$

tal que sus columnas constituyen una base polinómica minimal de $S$ o una base polinómica de orden minimal, cuyos grados de las columnas son respectivamente $\varepsilon_1 \leq \cdots \leq \varepsilon_\alpha$. En ambos casos, no (i) o no (iii), se sigue que no puede ser $\mu_i \leq \varepsilon_i$ para todo $i = 1, \cdots, \alpha$, entonces debe haber algún índice $j$ tal que

$$\mu_i \leq \varepsilon_i \quad \forall\, i \in \{1, \ldots, j-1\}; \qquad \text{y} \qquad \mu_j > \varepsilon_j \tag{1.21}$$

Ahora, como $\{f_1(\lambda), \ldots, f_\alpha(\lambda)\}$ es una base del espacio $S$, debemos tener

$$\begin{pmatrix} g_{1,k}(\lambda) \\ \vdots \\ g_{n,k}(\lambda) \end{pmatrix} = g_k(\lambda) = \sum_{i=1}^{\alpha} f_i(\lambda) p_{i,k}(\lambda) \quad (k = 1, \ldots, j), \tag{1.22}$$

donde los coeficientes $p_{i,k}(\lambda) \in \mathbb{F}(\lambda)$, para todo $k = 1, \ldots, j$ e $i = 1, \ldots, \alpha$. Eso es equivalente a decir que $p_k(\lambda) := \begin{pmatrix} p_{1,k}(\lambda) & \cdots & p_{n,k}(\lambda) \end{pmatrix}^{\mathrm{T}} \in \mathbb{F}(\lambda)^{n\times 1}$ verifica

$$F(\lambda) p_k(\lambda) = g_k(\lambda).$$

Luego, por la irreducibilidad de $F(\lambda)$ y aplicando la Proposición 1.5.8 y el Corolario 1.5.9, podemos suponer sin pérdida de generalidad que cada $p_k(\lambda)$ es un vector polinómico. Ahora, aplicando la propiedad del grado predecible (cf. Proposición 1.5.7), resulta que en el sumatorio (1.22), $p_{i,k}(\lambda) = 0$ para todo $i \geq j$ y todo $k = 1, \ldots, j$. En efecto, supongamos lo contrario, si existe un $l \geq j$ y un $k \leq j$ tal que $p_{l,k}(\lambda) \neq 0$, entonces $\partial(g_k(\lambda)) = \varepsilon_k \geq \partial(f_l(\lambda)) = \mu_l \geq \mu_j$, contradicción con las desigualdades (1.21). Así pues, hemos obtenido la contradicción siguiente con la independencia de los $\{g_k(\lambda)\}_{k=1}^{j}$, a saber,

$$\langle g_1(\lambda), \ldots, g_j(\lambda) \rangle \subseteq \langle f_1(\lambda), \ldots, f_{j-1}(\lambda) \rangle,$$

contradicción final que prueba las afirmaciones (i) y (iii).

□

**Definición 1.6.12.** Dadas dos sucesiones finitas crecientes de números naturales $a = (a_1, \dots, a_p)$ y $b = (b_1, \dots, b_q)$ llamaremos *unión* a la sucesión $c = (c_1, \dots, c_{p+q})$ formada por todos los términos de $a$ y $b$ reordenados en sentido creciente. Denotaremos esta unión por $c = a \cup b$, que no se debe confundir con la unión conjuntista.

**Ejemplo 1.6.13.** Sean $a = (0, 0, 0, 1, 3, 3)$ y $b = (0, 1, 1, 2, 2, 5)$. Entonces,

$$a \cup b = (0, 0, 0, 0, 1, 1, 1, 2, 2, 3, 3, 5).$$

Es evidente que la unión es asociativa y conmutativa.

**Corolario 1.6.14.** *Sea $\mathbb{F}$ un cuerpo algebraicamente cerrado. Sea*

$$H(\lambda) := \begin{bmatrix} H_1(\lambda) & 0 & \cdots & 0 \\ 0 & H_2(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_q(\lambda) \end{bmatrix}$$

*donde $H_i(\lambda) \in \mathbb{F}^{m_i \times n_i}$. Si $F_i(\lambda) \in \mathbb{F}[\lambda]^{n_i \times \alpha_i}$, con $\alpha_i = n_i - \operatorname{rg} H_i(\lambda)$, es una matriz-base minimal del núcleo por la derecha de $H_i(\lambda)$, con grados de las columnas $\mu_1^{(i)} \le \mu_2^{(i)} \le \cdots \le \mu_{\alpha_i}^{(i)}$ para $i = 1, \dots, q$, entonces las columnas de*

$$F(\lambda) := \begin{bmatrix} F_1(\lambda) & 0 & \cdots & 0 \\ 0 & F_2(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_q(\lambda) \end{bmatrix}$$

*constituyen una base minimal del núcleo por la derecha de $H(\lambda)$, cuya secuencia de índices minimales por la derecha es $\cup_{i=1}^{q} \big(\mu_1^{(i)}, \mu_2^{(i)}, \dots, \mu_{\alpha_i}^{(i)}\big)$.*

Demostración. Usando un argumento inductivo, es suficiente probarlo para $q = 2$. Sean pues

$$H(\lambda) := \begin{bmatrix} H_1(\lambda) & 0 \\ 0 & H_2(\lambda) \end{bmatrix} \quad \text{y} \quad F(\lambda) := \begin{bmatrix} F_1(\lambda) & 0 \\ 0 & F_2(\lambda) \end{bmatrix},$$

como en el enunciado, y $S$ denota el subespacio de las soluciones de la ecuación $H(\lambda)\,x = 0$. Está claro que $H(\lambda)F(\lambda) = 0$, dim $S = n_1 + n_2 - (\operatorname{rg} H_1(\lambda) + \operatorname{rg} H_2(\lambda)) = \alpha_1 + \alpha_2$ y que $\operatorname{rg} F_1(\lambda) + \operatorname{rg} F_2(\lambda) = \operatorname{rg} F(\lambda)$. Luego $F(\lambda)$ es una matriz-base polinómica de $S$ y sólo queda ver que es minimal. Para ello usaremos la caracterización del Teorema 1.6.11, según la cual las columnas de $F(\lambda)$ forman una base minimal de $S$ si y sólo si $F(\lambda)$ es reducida por columnas e irreducible. Es inmediato ver que la matriz de coeficientes principales de $F(\lambda)$, digamos $F_c$, es suma diagonal por bloques de las correspondientes matrices $(F_1)_c$ y $(F_2)_c$, en $F_1(\lambda)$ y $F_2(\lambda)$ respectivamente, cuyos rangos son $\alpha_1$ y $\alpha_2$, por ser ambas reducida por columnas por hipótesis. Así pues, el rango de $F_c$ es completo por columnas y por tanto $F(\lambda)$ es reducida por columnas. Asímismo, es inmediato que $F(\lambda)$ es irreducible por serlo cada $F_i(\lambda)$ $(i = 1, 2)$, usando la Proposición 1.5.8 y el Corolario 1.5.9.

$\square$

**Ejemplo 1.6.15.** Sea

$$H(\lambda) := \begin{pmatrix} \lambda^{-1} & 0 & \lambda^{-1} & \lambda \\ 0 & (\lambda+1)^2 & (\lambda+1)^2 & 0 \\ -1 & (\lambda+1)^2 & \lambda^2+2\lambda & -\lambda^2 \end{pmatrix} \in \mathbb{C}(\lambda)^{3\times 4}.$$

Entonces haciendo operaciones elementales por filas y columnas en $\mathbb{C}[\lambda]$ se obtiene

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda & -1 & 1 \end{pmatrix} H(\lambda) \begin{pmatrix} 1 & 0 & -1 & -\lambda^2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} \lambda^{-1} & 0 & 0 & 0 \\ 0 & (\lambda+1)^2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right).$$

Lo que significa que las columnas de la matriz $F_d(\lambda) = \begin{pmatrix} -1 & -\lambda^2 \\ -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{C}[\lambda]^{4\times 2}$, forman una base polinómica del núcleo por la derecha de $H(\lambda)$, ya que se verifica $H(\lambda)F_d(\lambda) = 0$. Además, por ser la matriz $F_d(\lambda)$ reducida por columnas e irreducible, del Teorema 1.6.11 se sigue que dicha base también es minimal, y por tanto los grados de sus columnas $\{0, 2\}$, son los índices minimales por columnas de la matriz $H(\lambda)$.

Análogamente a lo hecho por columnas se hace por filas, y se obtiene que la fila de la matriz $F_i(\lambda) = \begin{pmatrix} \lambda & -1 & 1 \end{pmatrix} \in \mathbb{C}[\lambda]^{1\times 3}$, es una base polinómica del núcleo por la izquierda de $H(\lambda)$, ya que se verifica $F_i(\lambda)H(\lambda) = 0$. Además, por ser la matriz $F_i(\lambda)$ reducida por filas e irreducible, del Teorema 1.6.11 se sigue que dicha base también es minimal, y por tanto el grado de su fila $\{1\}$, es el único índice minimal por filas de la matriz $H(\lambda)$.

## 1.7. Ejercicios

**1.1.** En el anillo $\mathbb{F}[x_1, \dots, x_n]$ con $n > 1$, sea $I$ el conjunto de los polinomios con término constante igual a 0, es decir los polinomios que tienen la forma

$$f(x_1, \dots, x_n) := \sum_{i_1, i_2, \dots, i_n} a_{i_1, i_2 \dots, i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \in \mathbb{F}[x_1, \dots, x_n]$$

con $a_{0,\dots,0} = 0$. Probar que $I$ es un ideal que no es principal.

**1.2.** Sea $\tilde{f}(\tilde{\lambda}, \tilde{\mu}) = \tilde{\lambda}^2\tilde{\mu}^3 + 2\tilde{\lambda}\tilde{\mu}^4 \in \mathbb{C}[\tilde{\lambda}, \tilde{\mu}]$. Hagamos la sustitución lineal de variables

$$\tilde{\lambda} = a\lambda + b\mu, \quad \tilde{\mu} = c\lambda + d\mu,$$

donde $a, b, c, d \in \mathbb{C}$ son tales que $ad - bc \neq 0$. Sin utilizar la Proposición 1.4.8, demostrar directamente que el grado del polinomio

$$f(\lambda, \mu) := \tilde{f}(a\lambda + b\mu, c\lambda + d\mu)$$

es igual a 5.

¿Es cierto el apartado 3) de la Proposición 1.4.8 si $\alpha\delta - \beta\gamma = 0$?

**1.3.** Sea $D(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ donde $\mathbb{F}$ es un cuerpo. Sea $D_c$ la matriz de los coeficientes principales por columnas de $D(\lambda)$. Demostrar que

$$\mathrm{rg}\ (D_c) \leq \mathrm{rgn}\ (D(\lambda)).$$

**1.4.** Dada una matriz compleja $X$ denotamos su traspuesta conjugada por $X^*$. Sean las matrices $A \in \mathbb{C}^{p\times n}$ y $B \in \mathbb{C}^{q\times n}$. Demostrar que si $\{0\} \neq \mathrm{Ker}\, B \cap \mathrm{Ker}\, A$ entonces el haz $\lambda B^*B - A^*A$ es singular. Encontrar un contraejemplo que pruebe que el recíproco es falso.

**1.5.** Sean $L$ un cuerpo y $K$ un subcuerpo de $L$ ( i.e. $K$ es un subconjunto de $L$ que contiene a 0 y 1 y que es un cuerpo con la restricción de las operaciones de $L$ a $K$). Sean $A_1, A_2, B_1, B_2 \in K^{m\times n}$ y supongamos que $K$ tiene más de $m+n$ elementos. Demostrar que si los haces $\lambda B_1 - A_1$ y $\lambda B_2 - A_2$ son estrictamente equivalentes sobre el cuerpo $L$, también lo son sobre el cuerpo $K$; dicho de otro modo, si existen matrices $P \in \mathrm{GL}_m(L)$ y $Q \in \mathrm{GL}_n(L)$ tales que

$$P(\lambda B_1 - A_1)Q = \lambda B_2 - A_2,$$

demostrar que entonces existen matrices $R \in \mathrm{GL}_m(K)$ y $S \in \mathrm{GL}_n(K)$ tales que

$$R(\lambda B_1 - A_1)S = \lambda B_2 - A_2.$$

Proceder del siguiente modo:

(i) Probar que los conjuntos

$$\mathcal{S} := \{(X,Y) \in K^{m\times m} \times K^{n\times n} \mid XB_1 = B_2Y,\ XA_1 = A_2Y\}$$

y

$$\mathcal{S}^L := \{(X,Y) \in L^{m\times m} \times L^{n\times n} \mid XB_1 = B_2Y,\ XA_1 = A_2Y\},$$

son subespacios vectoriales de $K^{m\times m} \times K^{n\times n}$ y $L^{m\times m} \times L^{n\times n}$, respectivamente.

(ii) Sea $\mathcal{B} = \{(X_1,Y_1),\ldots,(X_t,Y_t)\}$ una $K$–base de $\mathcal{S}$. Probar que se verifica

$$\mathcal{S}^L = \{\sum_{i=1}^{t} \alpha_i(X_i,Y_i) \mid \alpha_i \in L, i = 1,\ldots,t\}.$$

(iii) Sean $\lambda_1,\ldots,\lambda_t$ indeterminadas independientes sobre el cuerpo $K$ y sean

$$f(\lambda_1,\ldots,\lambda_t) = \det(\lambda_1 X_1 + \ldots + \lambda_t X_t) \in K[\lambda_1,\ldots,\lambda_t],$$

y

$$g(\lambda_1,\ldots,\lambda_t) = \det(\lambda_1 Y_1 + \ldots + \lambda_t Y_t) \in K[\lambda_1,\ldots,\lambda_t].$$

Probar que $f$ y $g$ son polinomios homogéneos no nulos, de grados $m$ y $n$, respectivamente. (Ayuda: Notar que $(P, Q^{-1}) \in \mathcal{S}^L$ y usar el apartado $(ii)$ para probar que $f$ y $g$ son no nulos.)

(iv) Probar que existen $\beta_1,\ldots,\beta_t \in K$ tales que

$$f(\beta_1,\ldots,\beta_t) \neq 0 \ \text{ y } \ g(\beta_1,\ldots,\beta_t) \neq 0.$$

(Ayuda: Aplicar el Teorema 1.3.8 a los polinomios $f$ y $g$.)

(v) Concluir que existen matrices $R \in \mathrm{GL}_m(K)$ y $S \in \mathrm{GL}_n(K)$ tales que

$$R(\lambda B_1 - A_1)S = \lambda B_2 - A_2.$$

## 1.8. Notas al capítulo

No habiendo podido mejorar algunas presentaciones, hemos seguido en algunas secciones a libros excelentes. La Sección 1.2 y los Teoremas 1.3.8 y 1.4.3 siguen a Marcus [17]. La Proposición 1.5.7 y el Teorema 1.6.11 están tomados del libro de Kailath [12]. Nuestro objetivo ha facilitar la lectura del texto en su conjunto, sin obligar al lector a consultar otros libros para poder seguir la línea argumental.

# Capítulo 2

# Haces regulares

## 2.1. Planteamiento del problema

Observemos que el determinante del haz $\lambda B - A \in \mathbb{F}[\lambda]^{n\times n}$ tiene la forma

$$|\lambda B - A| = |B|\lambda^n + \cdots + (-1)^n|A| \tag{2.1}$$

siendo los monomios que faltan de la forma $a_k\lambda^k$ con $a_k \in \mathbb{F}$ y $1 \leq k \leq n-1$. En efecto,

$$|\lambda B - A| = \begin{vmatrix} \lambda b_{11} - a_{11} & \lambda b_{12} - a_{12} & \ldots & \lambda b_{1n} - a_{1n} \\ \lambda b_{21} - a_{21} & \lambda b_{22} - a_{22} & \ldots & \lambda b_{2n} - a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda b_{n1} - a_{n1} & \lambda b_{n2} - a_{n2} & \ldots & \lambda b_{nn} - a_{nn} \end{vmatrix}$$

$$= \sum_{\sigma\in S_n} \varepsilon(\sigma)(\lambda b_{\sigma(1),1} - a_{\sigma(1),1})(\lambda b_{\sigma(2),2} - a_{\sigma(2),2})\cdots(\lambda b_{\sigma(n),n} - a_{\sigma(n),n}) \tag{2.2}$$

donde $S_n$ es el grupo simétrico de grado $n$, y $\varepsilon(\sigma)$ denota la signatura de la permutación $\sigma$.

Cada sumando de (2.2) se forma eligiendo $k$ $b$es y $n-k$ $a$es; de ahí sale la aportación al coeficiente de $\lambda^k$ en el polinomio $|\lambda B - A|$. Luego variamos $k$ desde $n, n-1, \ldots$ hasta 0. En particular, así se ve que el coeficiente de $\lambda^n$ es

$$\sum_{\sigma\in S_n} \varepsilon(\sigma) b_{\sigma(1),1} b_{\sigma(2),2}\cdots b_{\sigma(n),n} = |B|,$$

y el término independiente es

$$\sum_{\sigma\in S_n} \varepsilon(\sigma)(-1)^n a_{\sigma(1),1} a_{\sigma(2),2}\cdots a_{\sigma(n),n} = (-1)^n|A|.$$

En consecuencia, si $|A| \neq 0$ o si $|B| \neq 0$, el haz $\lambda B - A$ es necesariamente regular. Pero existen haces regulares con $|B| = 0$, e incluso con $|A| = |B| = 0$; por ejemplo, los haces

$$\begin{pmatrix} \lambda+2 & \lambda+1 & 2\lambda+3 \\ \lambda+3 & \lambda+2 & 2\lambda+5 \\ \lambda+3 & \lambda+2 & 3\lambda+6 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \lambda+1 & 3\lambda+1 \\ \lambda+2 & 3\lambda+2 \end{pmatrix}$$

son regulares, pues su determinante es distinto del polinomio cero.

Recordemos que la equivalencia estricta de haces implica su equivalencia como matrices polinómicas. El recíproco no es cierto en general como se verá en el Ejemplo 2.1.2. Sin embargo, en el caso especial de haces regulares $\lambda B - A$ y $\lambda B_1 - A_1$ en los que $|B| \neq 0$ y $|B_1| \neq 0$, los dos conceptos de equivalencia y de equivalencia estricta coinciden, como se observa en el siguiente

**Teorema 2.1.1.** *Dos haces $\lambda B - A$ y $\lambda B_1 - A_1$, pertenecientes a $\mathbb{F}[\lambda]^{n\times n}$, con $|B| \neq 0$ y $|B_1| \neq 0$, son estrictamente equivalentes si y sólo si tienen los mismos factores invariantes (o, equivalentemente, los mismos divisores elementales) en $\mathbb{F}[\lambda]$.*

No damos la demostración de este teorema que puede verse en [Gantmacher, vol. I, p. 147, Théorème 6].

Para comprobar si el Teorema 2.1.1 es cierto para haces regulares cualesquiera, consideremos el ejemplo siguiente:

**Ejemplo 2.1.2.** Sean

$$\lambda B - A = \begin{pmatrix} \lambda+2 & \lambda+1 & 2\lambda+3 \\ \lambda+3 & \lambda+2 & 2\lambda+5 \\ \lambda+3 & \lambda+2 & 3\lambda+6 \end{pmatrix}, \quad \lambda B_1 - A_1 = \begin{pmatrix} \lambda+2 & \lambda+1 & \lambda+1 \\ \lambda+1 & \lambda+2 & \lambda+1 \\ \lambda+1 & \lambda+1 & \lambda+1 \end{pmatrix}.$$

Es fácil probar, mediante operaciones elementales por filas y columnas en $\mathbb{F}[\lambda]$, que ambos haces son equivalentes a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda+1 \end{pmatrix},$$

de ahí que cada uno de dichos haces sólo tiene un divisor elemental, que es $\lambda+1$. Sin embargo, estos haces no son estrictamente equivalentes, pues las matrices

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{y} \quad B_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

tienen rangos 2 y 1 respectivamente; lo que hace imposible que existan matrices invertibles $P$ y $Q$ tales que $B_1 = PBQ$. No obstante, estos haces son regulares pues

$$|\lambda B - A| = |\lambda B_1 - A_1| = \lambda + 1.$$

Por lo tanto, el Teorema 2.1.1 no es cierto para haces regulares cualesquiera. Antes de extender este resultado a la clase de todos los haces regulares, necesitamos algunos conceptos y resultados previos.

## 2.2. Factores invariantes homogéneos de un haz rectangular

Necesitamos introducir el concepto de factores invariantes homogéneos. En la Definición 1.1.1, la condición (ii) de equivalencia estricta de haces es equivalente a

$$A_1 = PAQ, \qquad B_1 = PBQ.$$

Así los papeles de $A$ y $B$ pueden ser intercambiados sin afectar a la equivalencia estricta. Por ello es natural considerar haces *homogéneos*

$$H(\lambda,\mu) = \lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{m\times n}.$$

Definiremos el rango normal de $H(\lambda,\mu)$ como el orden del mayor menor de $H(\lambda,\mu)$ distinto del polinomio cero de $\mathbb{F}[\lambda,\mu]$, o como el rango de $H(\lambda,\mu)$ mirada como matriz con elementos en el cuerpo de fracciones $\mathbb{F}(\lambda,\mu)$, indistintamente. Lo denotaremos por rgn $H(\lambda,\mu)$. Es fácil ver que

$$\operatorname{rgn} H(\lambda,\mu) = \operatorname{rgn} H(\lambda).$$

Lo cual es consecuencia de la siguiente proposición, que es útil también para otros menesteres.

**Proposición 2.2.1.** *Sean $C, D \in \mathbb{F}^{k\times k}$. Llamaremos $G(\lambda) := \lambda D - C \in \mathbb{F}[\lambda]^{k\times k}$ y $G(\lambda,\mu) := \lambda D - \mu C \in \mathbb{F}[\lambda,\mu]^{k\times k}$. Entonces*

$$|G(\lambda,\mu)| = \mu^k |G(\frac{\lambda}{\mu})|$$

*como igualdad en $\mathbb{F}(\lambda,\mu)$, o en $\mathbb{F}[\lambda,\mu]$ al ser el primer miembro de la identidad un polinomio en $\lambda,\mu$.*

*Además $|G(\lambda,\mu)|$ es un polinomio homogéneo de grado $k$ en $\lambda,\mu$, o el polinomio nulo.*

Demostración. Sean $C = (c_{ij})$ y $D = (d_{ij})$. Entonces

$$|G(\lambda,\mu)| = \begin{vmatrix} \lambda d_{11} - \mu c_{11} & \lambda d_{12} - \mu c_{12} & \dots & \lambda d_{1k} - \mu c_{1k} \\ \lambda d_{21} - \mu c_{21} & \lambda d_{22} - \mu c_{22} & \dots & \lambda d_{2k} - \mu c_{2k} \\ \vdots & \vdots & & \vdots \\ \lambda d_{k1} - \mu c_{k1} & \lambda d_{k2} - \mu c_{k2} & \dots & \lambda d_{kk} - \mu c_{kk} \end{vmatrix} =$$

$$\sum_{\sigma\in S_k} \varepsilon(\sigma)(\lambda d_{\sigma(1),1} - \mu c_{\sigma(1),1})(\lambda d_{\sigma(2),2} - \mu c_{\sigma(2),2})\cdots(\lambda d_{\sigma(k),k} - \mu c_{\sigma(k),k}). \quad (2.3)$$

Por otro lado, tenemos

$$|G(\lambda)| = |\lambda D - C| = \sum_{\sigma\in S_k} \varepsilon(\sigma)(\lambda d_{\sigma(1),1} - c_{\sigma(1),1})(\lambda d_{\sigma(2),2} - c_{\sigma(2),2})\cdots$$
$$\cdots(\lambda d_{\sigma(k),k} - c_{\sigma(k),k}). \quad (2.4)$$

Sustituyendo $\lambda$ por la fracción racional $\frac{\lambda}{\mu}$ en (2.4), y multiplicando por $\mu^k$, se sigue

$$\mu^k|G(\frac{\lambda}{\mu})| = \mu^k|(\frac{\lambda}{\mu})D - C| = \sum_{\sigma\in S_k} \varepsilon(\sigma)\mu^k[(\frac{\lambda}{\mu})d_{\sigma(1),1} - c_{\sigma(1),1}][(\frac{\lambda}{\mu})d_{\sigma(2),2} - c_{\sigma(2),2}]$$
$$\cdots[(\frac{\lambda}{\mu})d_{\sigma(k),k} - c_{\sigma(k),k}] = |G(\lambda,\mu)|.$$

De la segunda parte de (2.3) se sigue que

$$|G(\lambda,\mu)| = \mu^k(-1)^k \sum_{\sigma\in S_k} \varepsilon(\sigma)c_{\sigma(1),1}c_{\sigma(2),2}\cdots c_{\sigma(k),k} + \sum_{l=1}^{k} \lambda^l\mu^{k-l}(-1)^{k-l} \sum_{\sigma\in S_k} \varepsilon(\sigma)$$
$$\sum_{1\le j_1<j_2<\cdots<j_l\le k} d_{\sigma(j_1),1}\cdots d_{\sigma(j_l),j_l}c_{\sigma(j_1'),j_1'}\cdots c_{\sigma(j_{k-l}'),j_{k-l}'}$$

donde la última suma está guiada por el índice $(j_1,\ldots,j_l)$ que recorre $Q_{l,k}$ y

$$\{j_1,\ldots,j_l\}\cup\{j_1',\ldots,j_{k-l}'\} = \{1,2,\ldots,k\}.$$

Luego

$$|G(\lambda,\mu)| = \mu^k(-1)^k|C| + \sum_{l=1}^{k} \lambda^l\mu^{k-l}(-1)^{k-l}$$
$$\sum_{\sigma\in S_k} \varepsilon(\sigma) \sum_{1\le j_1<j_2<\cdots<j_l\le k} d_{\sigma(j_1),1}\cdots d_{\sigma(j_l),j_l}c_{\sigma(j_1'),j_1'}\cdots c_{\sigma(j_{k-l}'),j_{k-l}'}$$
$$= \mu^k(-1)^k|C| + \sum_{l=1}^{k} \lambda^l\mu^{k-l}(-1)^{k-l} \sum_{1\le j_1<j_2<\cdots<j_l\le k} |(C|D)(j_1,\ldots,j_l)|$$

donde para $1 \le j_1 < j_2 < \cdots < j_l \le k$, $(C|D)(j_1,\ldots,j_l)$ es la matriz $k\times k$ obtenida al cambiar en la matriz $C$ las columnas $j_1, j_2, \ldots, j_l$ por las correspondientes columnas de la matriz $D$.

Así pues, queda probado que $|G(\lambda,\mu)|$ es un polinomio homogéneo de grado $k$ en $\lambda,\mu$, o el polinomio cero.

□

Recordemos que $\mathbb{F}[\lambda,\mu]$ es un dominio de factorización única (abreviadamente DFU), y es conocida la existencia de máximo común divisor en un DFU. Es decir, dados dos polinomios arbitrarios no nulos de $\mathbb{F}[\lambda,\mu]$ siempre existe un polinomio que es máximo común divisor de los dos polinomios dados. Este polinomio está determinado salvo multiplicación por una constante no nula de $\mathbb{F}$. Por ello, para que sea único, convenimos en elegirlo de forma que sea *mónico respecto de* $\lambda$, es decir, con el coeficiente del monomio donde $\lambda$ aparece elevado al exponente más alto, igual a 1. Por lo tanto, dado un haz homogéneo

$$H(\lambda,\mu) = \lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{m\times n},$$

podemos definir el *divisor determinantal de orden* $k$ como el máximo común divisor de los menores de orden $k$ de $H(\lambda,\mu)$, que denotamos por $D_k(\lambda,\mu)$ para

$k = 1, 2, \ldots, \operatorname{rgn} H(\lambda, \mu)$. De las propiedades del máximo común divisor, se sigue que

$$D_1(\lambda,\mu) \bigm| D_2(\lambda,\mu) \bigm| \cdots \bigm| D_r(\lambda,\mu), \quad r = \operatorname{rgn} H(\lambda,\mu), \tag{2.5}$$

Y los *factores invariantes* de $H(\lambda, \mu)$ vienen definidos por las fórmulas

$$i_k(\lambda,\mu) = \frac{D_k(\lambda,\mu)}{D_{k-1}(\lambda,\mu)}, \ (k = 1, 2, \ldots, r); \ D_0(\lambda,\mu) = 1.$$

También es cierto que

$$i_1(\lambda,\mu) \bigm| i_2(\lambda,\mu) \bigm| \cdots \bigm| i_r(\lambda,\mu), \tag{2.6}$$

pero esto no resulta inmediatamente de (2.5) y será probado más adelante.

A continuación demostraremos que los polinomios $D_k(\lambda, \mu)$ e $i_k(\lambda, \mu)$, $k = 1, 2, \ldots, r$, son polinomios homogéneos en $\mathbb{F}[\lambda, \mu]$. Son llamados, respectivamente, los *divisores determinantales homogéneos* y los *factores invariantes homogéneos* del haz $H(\lambda) = \lambda B - A$, o del haz $H(\lambda, \mu)$.

**Lema 2.2.2.** *Sean $A, B \in \mathbb{F}^{m\times n}$. Entonces los divisores determinantales $D_k(\lambda, \mu)$ y los factores invariantes $i_k(\lambda, \mu)$, $k = 1, 2, \ldots, r$, del haz homogéneo*

$$H(\lambda,\mu) = \lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{m\times n}; \ r := \operatorname{rgn} H(\lambda,\mu),$$

*son polinomios homogéneos.*

*Además, si $D_k(\lambda)$ e $i_k(\lambda)$ son los divisores determinantales y los factores invariantes del haz*

$$H(\lambda) = \lambda B - A \in \mathbb{F}[\lambda]^{m\times n}; \ k = 1, \ldots, \operatorname{rgn} H(\lambda),$$

*entonces se verifica*

$$D_k(\lambda) = D_k(\lambda, 1), \ i_k(\lambda) = i_k(\lambda, 1), \ k = 1, \ldots, \operatorname{rgn} H(\lambda)(= r). \tag{2.7}$$

Demostración. Por la Proposición 2.2.1 cualquier menor de orden $k$ de $H(\lambda, \mu)$ es o cero o un polinomio homogéneo de grado $k$. En el Teorema 1.4.3 hemos visto que todo divisor de un polinomio homogéneo de $\mathbb{F}[\lambda, \mu]$ es homogéneo también. Por lo tanto el máximo común divisor de todos los menores de orden $k$ es un polinomio homogéneo, $D_k(\lambda, \mu)$. Como $i_k(\lambda, \mu)$ es factor de $D_k(\lambda, \mu)$, $i_k(\lambda, \mu)$ también será homogéneo.

En lo que sigue se hará implícitamente uso de resultados fáciles sobre sustitución en polinomios de $\mathbb{F}[\lambda, \mu]$ y en fracciones racionales de $\mathbb{F}(\lambda, \mu)$ de elementos de superanillos y de supercuerpos de $\mathbb{F}$ en lugar de las indeterminadas $\lambda$ y $\mu$.

Tenemos

$$H(\lambda) = H(\lambda, 1).$$

Como $D_k(\lambda, \mu)$ divide a todos los menores $k\times k$ de $H(\lambda, \mu)$, tenemos que $D_k(\lambda, 1)$ divide a todos los menores $k \times k$ de $H(\lambda, 1) = H(\lambda)$. Luego $D_k(\lambda)$, el máximo común divisor de todos los menores $k \times k$ de $H(\lambda)$, es divisible por $D_k(\lambda, 1)$:

$$D_k(\lambda, 1) \bigm| D_k(\lambda). \tag{2.8}$$

En virtud de la Proposición 2.2.1, tenemos la relación siguiente entre los menores $k \times k$ de $H(\lambda, \mu)$ y $H(\lambda)$: Para todos $\alpha \in Q_{k,m}, \beta \in Q_{k,n}$,

$$|H(\lambda,\mu)[\alpha \mid \beta]| = \mu^k |H(\frac{\lambda}{\mu})[\alpha \mid \beta]|.$$

Sea $\rho_k := \partial(D_k(\lambda))$. Como $D_k(\lambda)$ divide al menor $|H(\lambda)[\alpha \mid \beta]|$, se sigue que existe un polinomio $q(\lambda) \in \mathbb{F}[\lambda]$ tal que

$$D_k(\lambda)q(\lambda) = |H(\lambda)[\alpha \mid \beta]|. \tag{2.9}$$

Sustituyendo la fracción racional $\frac{\lambda}{\mu} \in \mathbb{F}(\lambda, \mu)$ en vez de $\lambda$ en (2.9), tenemos

$$D_k(\frac{\lambda}{\mu})q(\frac{\lambda}{\mu}) = |H(\frac{\lambda}{\mu})[\alpha \mid \beta]|. \tag{2.10}$$

Multiplicando en (2.10) por $\mu^k$ queda

$$\mu^k D_k(\frac{\lambda}{\mu})q(\frac{\lambda}{\mu}) = \mu^k|H(\frac{\lambda}{\mu})[\alpha \mid \beta]| = |H(\lambda, \mu)[\alpha \mid \beta]|. \tag{2.11}$$

De (2.9) deducimos que

$$\partial(D_k(\lambda)) + \partial(q(\lambda)) = \partial(|H(\lambda)[\alpha \mid \beta]|),$$

lo que, según (2.1), se traduce por

$$\rho_k + \partial(q(\lambda)) = j \quad (j \leq k);$$

y esto implica que $\partial(q(\lambda)) = j - \rho_k$. Reescribamos (2.11) en la forma

$$\mu^{\rho_k} D_k(\frac{\lambda}{\mu})\mu^{k-j}\mu^{j-\rho_k}q(\frac{\lambda}{\mu}) = |H(\lambda, \mu)[\alpha \mid \beta]|.$$

Por la Observación 1.4.4, es claro que

$$\mu^{\rho_k} D_k(\frac{\lambda}{\mu}) \quad \text{y} \quad \mu^{j-\rho_k}q(\frac{\lambda}{\mu})$$

pertenecen a $\mathbb{F}[\lambda, \mu]$; por consiguiente,

$$\mu^{\rho_k} D_k(\frac{\lambda}{\mu}) \,\Big|\, |H(\lambda, \mu)[\alpha \mid \beta]| \quad (\text{ en } \ \mathbb{F}[\lambda, \mu]). \tag{2.12}$$

Como (2.12) es cierto para todos $\alpha \in Q_{k,m}$ y $\beta \in Q_{k,n}$, deducimos que

$$\mu^{\rho_k} D_k(\frac{\lambda}{\mu}) \,\Big|\, D_k(\lambda, \mu) \quad (\textit{en } \mathbb{F}[\lambda, \mu]).$$

Por lo tanto, existe un polinomio $p(\lambda, \mu) \in \mathbb{F}[\lambda, \mu]$ tal que

$$p(\lambda, \mu)\mu^{\rho_k} D_k(\frac{\lambda}{\mu}) = D_k(\lambda, \mu). \tag{2.13}$$

Haciendo $\mu = 1$ en (2.13), se tiene que

$$p(\lambda, 1)D_k(\lambda) = D_k(\lambda, 1);$$

lo que prueba que

$$D_k(\lambda) \,\Big|\, D_k(\lambda, 1) \quad (\textit{en } \mathbb{F}[\lambda]). \tag{2.14}$$

De (2.8) y (2.14) se sigue que

$$D_k(\lambda) = D_k(\lambda, 1), \tag{2.15}$$

como queríamos demostrar.

Por (2.15), el grado de $D_k(\lambda, \mu)$ respecto de $\lambda$ es igual al grado de $D_k(\lambda)$, $\rho_k$. Como el grado (total) de $D_k(\lambda, \mu)$ es mayor o igual que el grado de $D_k(\lambda, \mu)$ respecto de $\lambda$, se tiene que

$$\phi_k := \partial(D_k(\lambda, \mu)) - \rho_k \geq 0.$$

Por (1.8) tenemos que

$$D_k(\lambda, \mu) = \mu^{\phi_k} \mu^{\rho_k} D_k(\frac{\lambda}{\mu}, 1);$$

de aquí y de (2.15) deducimos

$$D_k(\lambda, \mu) = \mu^{\phi_k} \mu^{\rho_k} D_k(\frac{\lambda}{\mu}), \tag{2.16}$$
$$\rho_k := \partial(D_k(\lambda));\ \phi_k \geq 0;\ k = 1, 2, \ldots,\ r := \operatorname{rgn} H(\lambda).$$

Ahora de la fórmula

$$i_k(\lambda, \mu) := \frac{D_k(\lambda, \mu)}{D_{k-1}(\lambda, \mu)},$$

que define los factores invariantes homogéneos, deducimos

$$i_k(\lambda, \mu) D_{k-1}(\lambda, \mu) = D_k(\lambda, \mu),$$

y evaluando en $\mu = 1$ resulta

$$i_k(\lambda, 1) D_{k-1}(\lambda, 1) = D_k(\lambda, 1),$$

que por (2.15) llega a ser

$$i_k(\lambda, 1) D_{k-1}(\lambda) = D_k(\lambda).$$

Así pues, se verifica

$$i_k(\lambda, 1) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} = i_k(\lambda). \tag{2.17}$$

Entonces, de acuerdo con (1.8) y (2.17), obtenemos

$$i_k(\lambda, \mu) = \mu^{\psi_k} \mu^{\sigma_k} i_k(\frac{\lambda}{\mu}), \tag{2.18}$$

donde $\sigma_k := \partial(i_k(\lambda));\ \psi_k := \partial(i_k(\lambda, \mu)) - \sigma_k \geq 0;\ k = 1, 2, \ldots,\ r := \operatorname{rgn} H(\lambda)$.

Finalmente, de la Proposición 2.2.1 sigue inmediatamente la igualdad

$$\operatorname{rgn} H(\lambda, \mu) = \operatorname{rgn} H(\lambda).$$

□

**Lema 2.2.3.** *Sean $A, B \in \mathbb{F}^{m\times n}$. Sean $D_k(\lambda,\mu)$ e $i_k(\lambda,\mu)$ los divisores determinantales homogéneos y los factores invariantes homogéneos del haz homogéneo*

$$H(\lambda,\mu) = \lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{m\times n},$$

*para $k = 1, 2, \ldots, r := \operatorname{rgn} H(\lambda,\mu)$.*

*Si $D_k(\mu)$ e $i_k(\mu)$ son los divisores determinantales y los factores invariantes de la $\mu$–matriz*

$$B - \mu A \in \mathbb{F}[\mu]^{m\times n}; \;\; k = 1, \ldots, \operatorname{rgn}\,(B - \mu A),$$

*entonces se tiene que*

$$D_k(\mu) = D_k(1,\mu), \quad i_k(\mu) = i_k(1,\mu), \tag{2.19}$$

*para $k = 1, 2, \ldots, \operatorname{rgn}\,(B - \mu A)(= r)$.*

La demostración de este lema es completamente análoga a la del lema anterior. Con las notaciones de la Proposición 2.2.1 puede demostrarse análogamente que

$$|G(\lambda,\mu)| = \lambda^k |G_1(\frac{\mu}{\lambda})|,$$

siendo $G_1(\mu) := D - \mu C \in \mathbb{F}[\mu]^{k\times k}$. De aquí se deduce la igualdad

$$\operatorname{rgn}\,(\lambda B - \mu A) = \operatorname{rgn}\,(B - \mu A).$$

**Lema 2.2.4.** *Sean $i_k(\lambda,\mu)$ los factores invariantes homogéneos del haz $H(\lambda) = \lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$, para $k = 1, 2, \ldots, r := \operatorname{rgn} H(\lambda)$. Entonces se verifica*

$$i_k(\lambda,\mu) \,\big|\, i_{k+1}(\lambda,\mu), \quad k = 1, 2, \ldots, r-1. \tag{2.20}$$

Demostración. De (2.18) se sigue que

$$i_k(\lambda,\mu) = \mu^{\psi_k}\mu^{\sigma_k} i_k(\frac{\lambda}{\mu}), \quad \sigma_k := \partial(i_k(\lambda)), \;\psi_k \geq 0,$$

$$i_{k+1}(\lambda,\mu) = \mu^{\psi_{k+1}}\mu^{\sigma_{k+1}} i_{k+1}(\frac{\lambda}{\mu}), \quad \sigma_{k+1} := \partial(i_{k+1}(\lambda)), \;\psi_{k+1} \geq 0. \tag{2.21}$$

Haciendo $\lambda = 1$ en estas identidades queda

$$i_k(1,\mu) = \mu^{\psi_k}\mu^{\sigma_k} i_k(\frac{1}{\mu}), \tag{2.22}$$

$$i_{k+1}(1,\mu) = \mu^{\psi_{k+1}}\mu^{\sigma_{k+1}} i_k(\frac{1}{\mu}).$$

Obviamente $\mu^{\sigma_k} i_k(\frac{1}{\mu})$ es un polinomio de la forma

$$1 + a_1\mu + \cdots + a_{\sigma_k}\mu^{\sigma_k} \in \mathbb{F}[\mu].$$

Del mismo modo

$$\mu^{\sigma_{k+1}} i_{k+1}(\frac{1}{\mu}) = 1 + b_1\mu + \cdots + b_{\sigma_{k+1}}\mu^{\sigma_{k+1}} \in \mathbb{F}[\mu].$$

Como $i_k(1,\mu)$ e $i_{k+1}(1,\mu)$ son factores invariantes sucesivos de la $\mu$–matriz $B-\mu A$ por (2.19), se tiene que

$$i_k(1,\mu) = i_k(\mu) \,\big|\, i_{k+1}(1,\mu) = i_{k+1}(\mu);$$

de aquí que por (2.22) se verifica

$$\mu^{\psi_k} \,\big|\, \mu^{\psi_{k+1}}[1 + b_1\mu + \cdots + b_{\sigma_{k+1}}\mu^{\sigma_{k+1}}];$$

y como estamos en $\mathbb{F}[\mu]$ (DFU) y

$$\mathrm{mcd}(\mu^{\psi_k}, 1 + b_1\mu + \cdots + b_{\sigma_{k+1}}\mu^{\sigma_{k+1}}) = 1,$$

se sigue que $\mu^{\psi_k} \,\big|\, \mu^{\psi_{k+1}}$, de donde

$$\psi_k \leq \psi_{k+1}. \tag{2.23}$$

Por otro lado, por ser factores invariantes sucesivos de la $\lambda$–matriz $\lambda B - A$, tenemos

$$i_k(\lambda) \,\big|\, i_{k+1}(\lambda);$$

luego existe $q_k(\lambda) \in \mathbb{F}[\lambda]$ tal que

$$i_k(\lambda) q_k(\lambda) = i_{k+1}(\lambda),$$

con $\partial(q_k(\lambda)) = \sigma_{k+1} - \sigma_k$; y sustituyendo $\lambda$ por la fracción racional $\frac{\lambda}{\mu}$ en esta expresión resulta

$$i_k(\frac{\lambda}{\mu}) q_k(\frac{\lambda}{\mu}) = i_{k+1}(\frac{\lambda}{\mu}).$$

Multiplicando ambos miembros de esta igualdad por $\mu^{\psi_{k+1}}\mu^{\sigma_{k+1}}$ resulta

$$\mu^{\psi_k}\mu^{\sigma_k} i_k(\frac{\lambda}{\mu}) \mu^{\psi_{k+1}-\psi_k}\mu^{\sigma_{k+1}-\sigma_k} q_k(\frac{\lambda}{\mu}) = \mu^{\psi_{k+1}}\mu^{\sigma_{k+1}} i_{k+1}(\frac{\lambda}{\mu}).$$

Ahora, por (2.21), se obtiene

$$i_k(\lambda,\mu)\mu^{\psi_{k+1}-\psi_k}\mu^{\sigma_{k+1}-\sigma_k} q_k(\frac{\lambda}{\mu}) = i_{k+1}(\lambda,\mu). \tag{2.24}$$

De acuerdo con la Observación 1.4.4, $\mu^{\sigma_{k+1}-\sigma_k} q_k(\frac{\lambda}{\mu})$ es un polinomio homogéneo de $\mathbb{F}[\lambda,\mu]$ de grado $\sigma_{k+1} - \sigma_k$, y como, por (2.23), $\psi_{k+1} - \psi_k \geq 0$

$$\mu^{\psi_{k+1}-\psi_k}\mu^{\sigma_{k+1}-\sigma_k} q_k(\frac{\lambda}{\mu})$$

es también un polinomio homogéneo de $\mathbb{F}[\lambda,\mu]$. Así pues, (2.24) implica el resultado buscado

$$i_k(\lambda,\mu) \,\big|\, i_{k+1}(\lambda,\mu).$$

□

**Lema 2.2.5.** *Sean $H(\lambda) = \lambda B - A$ y $H_1(\lambda) = \lambda B_1 - A_1$ dos haces estrictamente equivalentes en $\mathbb{F}[\lambda]^{m\times n}$. Entonces los los factores invariantes homogéneos de $H(\lambda)$ y $H_1(\lambda)$ son los mismos.*

Demostración. Por hipótesis existen $P \in \mathrm{GL}_m(\mathbb{F})$ y $Q \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$H_1(\lambda) = PH(\lambda)Q.$$

Por lo tanto,

$$\lambda B_1 - \mu A_1 = P(\lambda B - \mu A)Q.$$

Como $P$ y $Q$ son matrices invertibles sobre el cuerpo $\mathbb{F}$, también lo son sobre el anillo $\mathbb{F}[\lambda, \mu]$, de modo que $\lambda B_1 - \mu A_1$ es quivalente a $\lambda B - \mu A$ en $\mathbb{F}[\lambda, \mu]$. Por ser este anillo un dominio de factorización única y por el Lema 1.2.20 se sigue que los factores invariantes de $\lambda B_1 - \mu A_1$ y $\lambda B - \mu A$ son los mismos.

□

## 2.3. Caracterización de la equivalencia estricta de haces regulares

Ahora ya estamos en condiciones de hacer la siguiente generalización del Teorema 2.1.1 a haces regulares cualesquiera.

**Teorema 2.3.1.** *Sean $A, B, A_1, B_1 \in \mathbb{F}^{n\times n}, H(\lambda) = \lambda B - A$, un haz regular, y $H_1(\lambda) = \lambda B_1 - A_1 \in \mathbb{F}[\lambda]^{n\times n}$. Entonces $H(\lambda) \overset{s}{\sim} H_1(\lambda)$ si y sólo si $H(\lambda)$ y $H_1(\lambda)$ tienen los mismos factores invariantes homogéneos. En cuyo caso, el haz $H_1(\lambda)$ también es regular*

Demostración. Por el Lema 2.2.5, si $H(\lambda) \overset{s}{\sim} H_1(\lambda)$ entonces los factores invariantes homogéneos de $H(\lambda)$ y $H_1(\lambda)$ son iguales.

Recíprocamente, supongamos ahora que $H(\lambda)$ y $H_1(\lambda)$ tienen los mismos factores invariantes homogéneos. De acuerdo con (2.7) del Lema 2.2.2, los haces $H(\lambda)$ y $H_1(\lambda)$ tienen los mismos factores invariantes en $\mathbb{F}[\lambda]$; consecuentemente, $H(\lambda) \sim H_1(\lambda)$ sobre $\mathbb{F}[\lambda]$. Pueden ahora ocurrir dos casos:

(i) $|B| \neq 0$ y $|B_1| \neq 0$; o bien

(ii) $|B| = 0$ o $|B_1| = 0$.

Si ocurre el caso $(i)$, entonces, por el Teorema 2.1.1, $H(\lambda) \overset{s}{\sim} H_1(\lambda)$ y habríamos terminado la demostración.

Si, por el contrario, $|B| = 0$ o $|B_1| = 0$, vamos a demostrar el Teorema reduciendo la prueba al caso $(i)$.

Definimos $H(\lambda, \mu) := \lambda B - \mu A$, $H_1(\lambda, \mu) := \lambda B_1 - \mu A_1$ pertenecientes a $\mathbb{F}[\lambda, \mu]^{n\times n}$. Por la Proposición 2.2.1 tenemos

$$|H(\lambda, \mu)| = \mu^n |H(\frac{\lambda}{\mu})| \neq 0,$$

donde la última desigualdad es por ser $H(\lambda)$ un haz regular. Por hipótesis $H_1(\lambda, \mu)$ tiene los mismos factores invariantes homogéneos que $H(\lambda, \mu)$; luego se sigue $|H_1(\lambda, \mu)| = |H(\lambda, \mu)| \neq 0$, pues $|H_1(\lambda, \mu)|$ es igual al producto de todos esos factores invariantes homogéneos.

Teniendo en cuenta el Teorema 1.3.8, se obtiene que por ser $\mathbb{F}$ infinito existen $\alpha, \gamma \in \mathbb{F}$ tales que

$$|\alpha B - \gamma A| = |\alpha B_1 - \gamma A_1| \neq 0. \tag{2.25}$$

Ahora, para estos $\alpha, \gamma$, elegimos $\beta, \delta \in \mathbb{F}$ de manera que la matriz

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

sea invertible.

Hagamos a continuación la sustitución de indeterminadas $\lambda$ y $\mu$ por los polinomios $\alpha\tilde{\lambda} + \beta\tilde{\mu}, \gamma\tilde{\lambda} + \delta\tilde{\mu} \in \mathbb{F}[\tilde{\lambda}, \tilde{\mu}]$

$$\left\{ \begin{array}{rcl} \lambda & = & \alpha\tilde{\lambda} + \beta\tilde{\mu} \\ \mu & = & \gamma\tilde{\lambda} + \delta\tilde{\mu} \end{array} \right. \tag{2.26}$$

en los haces $H(\lambda, \mu) = \lambda B - \mu A$, $H_1(\lambda, \mu) = \lambda B_1 - \mu A_1 \in \mathbb{F}[\lambda, \mu]^{n \times n}$, obtenemos los nuevos haces

$$\tilde{H}(\tilde{\lambda}, \tilde{\mu}) := \tilde{\lambda}\tilde{B} - \tilde{\mu}\tilde{A}, \ \tilde{H}_1(\tilde{\lambda}, \tilde{\mu}) := \tilde{\lambda}\tilde{B}_1 - \tilde{\mu}\tilde{A}_1 \in \mathbb{F}[\tilde{\lambda}, \tilde{\mu}]^{n \times n}$$

donde la expresión de $\tilde{A}, \tilde{B}, \tilde{A}_1, \tilde{B}_1$ en términos de $A, B, A_1, B_1$ se obtiene de la siguiente manera:

$$\begin{array}{rcl} \lambda B - \mu A & = & (\alpha\tilde{\lambda} + \beta\tilde{\mu})B - (\gamma\tilde{\lambda} + \delta\tilde{\mu})A \\ & = & \tilde{\lambda}\,(\alpha B - \gamma A) - \tilde{\mu}\,(\delta A - \beta B); \end{array}$$

análogamente resulta

$$\lambda B_1 - \mu A_1 = \tilde{\lambda}\,(\alpha B_1 - \gamma A_1) - \tilde{\mu}\,(\delta A_1 - \beta B_1);$$

así sigue que

$$\begin{array}{rcl} \tilde{A} & = & \delta A - \beta B, \qquad \tilde{B} = \alpha B - \gamma A \\ \tilde{A}_1 & = & \delta A_1 - \beta B_1, \quad \tilde{B}_1 = \alpha B_1 - \gamma A_1. \end{array} \tag{2.27}$$

Esto en notación matricial por bloques, se escribe

$$\begin{bmatrix} \tilde{B} \\ -\tilde{A} \end{bmatrix} = \begin{bmatrix} \alpha I_n & \gamma I_n \\ \beta I_n & \delta I_n \end{bmatrix} \begin{bmatrix} B \\ -A \end{bmatrix} \ y \ \begin{bmatrix} \tilde{B}_1 \\ -\tilde{A}_1 \end{bmatrix} = \begin{bmatrix} \alpha I_n & \gamma I_n \\ \beta I_n & \delta I_n \end{bmatrix} \begin{bmatrix} B_1 \\ -A_1 \end{bmatrix}.$$

Se puede comprobar, que si

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}^{-1} = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

entonces se verifica

$$\begin{bmatrix} B \\ -A \end{bmatrix} = \begin{bmatrix} aI_n & cI_n \\ bI_n & dI_n \end{bmatrix} \begin{bmatrix} \tilde{B} \\ -\tilde{A} \end{bmatrix} \ y \ \begin{bmatrix} B_1 \\ -A_1 \end{bmatrix} = \begin{bmatrix} aI_n & cI_n \\ bI_n & dI_n \end{bmatrix} \begin{bmatrix} \tilde{B}_1 \\ -\tilde{A}_1 \end{bmatrix}.$$

Así obtenemos el cambio inverso de (2.27), a saber,

$$\begin{array}{rcl} A & = & d\tilde{A} - b\tilde{B}, \qquad B = a\tilde{B} - c\tilde{A} \\ A_1 & = & d\tilde{A}_1 - b\tilde{B}_1, \quad B_1 = a\tilde{B}_1 - c\tilde{A}_1. \end{array} \tag{2.28}$$

Además, según (2.25), tenemos ahora

$$|\tilde{B}| \neq 0 \quad y \quad |\tilde{B}_1| \neq 0. \tag{2.29}$$

Luego, para poder aplicar el Teorema 2.1.1 a los haces $\tilde{H}(\tilde{\lambda},\tilde{\mu}), \tilde{H}_1(\tilde{\lambda},\tilde{\mu})$, tenemos antes que probar que estos haces tienen los mismos factores invariantes homogéneos en el anillo $\mathbb{F}[\tilde{\lambda},\tilde{\mu}]$.

En efecto, por hipótesis $H(\lambda,\mu)$ y $H_1(\lambda,\mu)$ tienen los mismos factores invariantes homogéneos; lo cual equivale a que $H(\lambda,\mu)$ y $H_1(\lambda,\mu)$ tienen los mismos divisores determinantales homogéneos

$$D_k(\lambda,\mu), \quad k = 1,\ldots,n = \operatorname{rgn} H(\lambda,\mu) = \operatorname{rgn} H_1(\lambda,\mu),$$

donde, recordemos que

$$\begin{aligned} D_k(\lambda,\mu) &= \operatorname{mcd}\{|H(\lambda,\mu)[\alpha',\beta']| \mid \alpha',\beta' \in Q_{k,n}\} \\ &= \operatorname{mcd}\{|H_1(\lambda,\mu)[\alpha',\beta']| \mid \alpha',\beta' \in Q_{k,n}\}. \end{aligned} \tag{2.30}$$

Sean $p_1,\ldots,p_t \in \mathbb{F}[\lambda,\mu]$ todos los polinomios irreducibles distintos que aparecen al factorizar en factores irreducibles todos los menores de orden $k$ de $H(\lambda,\mu)$ y $H_1(\lambda,\mu)$. Entonces

$$D_k(\lambda,\mu) = p_1^{s_1}\cdots p_t^{s_t} \quad \textit{con } s_i \in \mathbb{N}, \textit{ para } i = 1,\ldots,t. \tag{2.31}$$

Si $\Phi : \mathbb{F}[\lambda,\mu] \longrightarrow \mathbb{F}[\tilde{\lambda},\tilde{\mu}]$ es el isomorfismo de anillos definido en las Proposiciones 1.4.7 y 1.4.8, se tiene que para todo $(\alpha',\beta') \in Q_{k,n}^2$,

$$\Phi(|H(\lambda,\mu)[\alpha',\beta']|) = |\tilde{H}(\tilde{\lambda},\tilde{\mu})[\alpha',\beta']|. \tag{2.32}$$

De modo que, si

$$|H(\lambda,\mu)[\alpha',\beta']| = p_1^{m_1(\alpha',\beta')}\cdots p_t^{m_t(\alpha',\beta')}$$

para unos $m_i(\alpha',\beta') \in \mathbb{N}$, por (2.32) se sigue

$$|\tilde{H}(\tilde{\lambda},\tilde{\mu})[\alpha',\beta']| = \tilde{p}_1^{m_1(\alpha',\beta')}\cdots \tilde{p}_t^{m_t(\alpha',\beta')},$$

donde $\tilde{p}_i := \Phi(p_i)$ para $i = 1,\ldots,t$. Por otro lado, si

$$|H_1(\lambda,\mu)[\alpha',\beta']| = p_1^{n_1(\alpha',\beta')}\cdots p_t^{n_t(\alpha',\beta')}$$

para unos $n_i(\alpha',\beta') \in \mathbb{N}$ $(i = 1,\ldots,t)$, análogamente se obtiene

$$|\tilde{H}_1(\tilde{\lambda},\tilde{\mu})[\alpha',\beta']| = \tilde{p}_1^{n_1(\alpha',\beta')}\cdots \tilde{p}_t^{n_t(\alpha',\beta')}.$$

En un dominio de factorización única, como es $\mathbb{F}[\lambda,\mu]$, son válidas las reglas elementales de la aritmética para calcular el máximo común divisor. Entonces, teniendo encuenta (2.30) y (2.31) tenemos

$$\min\{m_i(\alpha',\beta') \mid (\alpha',\beta') \in Q_{k,n}^2\} = s_i = \min\{n_i(\alpha',\beta') \mid (\alpha',\beta') \in Q_{k,n}^2\}. \tag{2.33}$$

Por la Proposición 1.4.8, los polinomios $\tilde{p}_1,\ldots,\tilde{p}_t$ son elementos irreducibles de $\mathbb{F}[\tilde{\lambda},\tilde{\mu}]$, que también es un DFU. De ahí que, por (2.33), el mcd de todos los menores de orden $k$, tanto de $\tilde{H}(\tilde{\lambda},\tilde{\mu})$ como de $\tilde{H}_1(\tilde{\lambda},\tilde{\mu})$, es

$$D_k(\tilde{\lambda},\tilde{\mu}) = \tilde{p}_1^{s_1}\cdots \tilde{p}_t^{s_t}.$$

Así pues, los factores invariantes homogéneos de $\tilde{H}(\tilde{\lambda},\tilde{\mu})$ y $\tilde{H}_1(\tilde{\lambda},\tilde{\mu})$ son iguales. Por consiguiente, del Teorema 2.1.1 se deduce que

$$\tilde{H}(\tilde{\lambda},\tilde{\mu}) \overset{s}{\sim} \tilde{H}_1(\tilde{\lambda},\tilde{\mu}). \tag{2.34}$$

Para terminar la demostración, veamos que $H(\tilde{\lambda},\tilde{\mu}) \overset{s}{\sim} \tilde{H}_1(\tilde{\lambda},\tilde{\mu})$ es equivalente a $H(\lambda,\mu) \overset{s}{\sim} H_1(\lambda,\mu)$.

Supongamos que $H(\lambda,\mu) \overset{s}{\sim} H_1(\lambda,\mu)$. Entonces existen matrices $P, Q \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$PAQ = A_1 \quad y \quad PBQ = B_1.$$

De aquí, aplicando (2.27), resulta

$$P\tilde{A}Q = P(\delta A - \beta B)Q = \delta PAQ - \beta PBQ = \delta A_1 - \beta B_1 = \tilde{A}_1,$$

y análogamente

$$P\tilde{B}Q = P(\alpha B - \gamma A)Q = \alpha PBQ - \gamma PAQ = \alpha B_1 - \gamma A_1 = \tilde{B}_1.$$

Recíprocamente, supongamos ahora que $\tilde{H}(\tilde{\lambda},\tilde{\mu}) \overset{s}{\sim} \tilde{H}_1(\tilde{\lambda},\tilde{\mu})$. Esto quiere decir que existen matrices $\tilde{P}, \tilde{Q} \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$\tilde{P}\tilde{A}\tilde{Q} = \tilde{A}_1 \quad y \quad \tilde{P}\tilde{B}\tilde{Q} = \tilde{B}_1.$$

Entonces, haciendo uso de (2.28), se obtiene

$$\tilde{P}A\tilde{Q} = \tilde{P}(d\tilde{A} - b\tilde{B})\tilde{Q} = d\tilde{P}\tilde{A}\tilde{Q} - b\tilde{P}\tilde{B}\tilde{Q} = d\tilde{A}_1 - b\tilde{B}_1 = A_1,$$

y análogamente

$$\tilde{P}B\tilde{Q} = \tilde{P}(a\tilde{B} - c\tilde{A})\tilde{Q} = a\tilde{P}\tilde{B}\tilde{Q} - c\tilde{P}\tilde{A}\tilde{Q} = a\tilde{B}_1 - c\tilde{A}_1 = B_1.$$

□

## 2.4. Divisores elementales homogéneos de un haz rectangular

Dado un haz homogéneo $\lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{n\times m}$, sean $i_k(\lambda,\mu)$ para $k = 1,\ldots,r = \mathrm{rgn}(\lambda B - \mu A)$, sus factores invariantes homogéneos. Factorizando estos polinomios homogéneos en potencias de polinomios irreducibles de $\mathbb{F}[\lambda,\mu]$, obtenemos los divisores elementales $e_k(\lambda,\mu)$ del haz. El Teorema 1.4.3, nos permite afirmar que los divisores elementales son polinomios homogéneos, así como también los polinomios irreducibles mencionados antes.

**Definición 2.4.1.** Los divisores elementales del haz $\lambda B - \mu A \in \mathbb{F}[\lambda,\mu]^{m\times n}$ de la forma $\mu^q$, $q$ entero positivo, se llaman *divisores elementales infinitos* del haz $\lambda B - \mu A$ ( o del haz $\lambda B - A$).

A los divisores elementales $e_k(\lambda,\mu)$ del haz $\lambda B - \mu A$ los llamaremos también divisores elementales homogéneos del haz $\lambda B - A$.

A continuación nos proponemos determinar la relación existente entre los divisores elementales homogéneos, no infinitos, del haz $\lambda B - \mu A$ y los divisores elementales de la matriz polinómica $\lambda B - A$.

Sea $e_j(\lambda, \mu)$ un divisor elemental homogéneo, no infinito, que aparece en la factorización del factor invariante $i_k(\lambda, \mu)$ en producto de potencias de polinomios irreducibles distintos. Entonces tenemos $e_j(\lambda, \mu) = p_j(\lambda, \mu)^{m_j}$, siendo $p_j(\lambda, \mu)$ un polinomio irreducible (homogéneo) y $m_j > 0$ un entero y se verifica

$$i_k(\lambda, \mu) = \cdots p_j(\lambda, \mu)^{m_j} \cdots = \cdots e_j(\lambda, \mu) \cdots . \tag{2.35}$$

Recordemos que, según (2.7) del Lema 2.2.2, $i_k(\lambda, 1) = i_k(\lambda)$ factor invariante de $\lambda B - A$. Hacemos pues $\mu = 1$ en (2.35) y queda

$$i_k(\lambda) = i_k(\lambda, 1) = \cdots p_j(\lambda, 1)^{m_j} \cdots = \cdots e_j(\lambda, 1) \cdots .$$

De acuerdo con Teorema 1.4.6 apartado 1), si $\partial(p_j(\lambda, \mu)) = d_j$, entonces el polinomio $p_j(\lambda, 1)$ tiene grado $d_j$ y es irreducible en $\mathbb{F}[\lambda]$.

Así pues, se concluye que $e_j(\lambda, 1) = p_j(\lambda, 1)^{m_j}$ es un divisor elemental de $\lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$.

Recíprocamente, partiendo de un divisor elemental $e_j(\lambda)$ de grado $d_j$ de $\lambda B - A$, definimos

$$e_j(\lambda, \mu) := \mu^{d_j} e_j(\frac{\lambda}{\mu}),$$

y aplicando el Teorema 1.4.6 apartado 2), tenemos que $e_j(\lambda, \mu)$ es un divisor elemental homogéneo del haz $\lambda B - \mu A$. Podemos obtener así todos los divisores elementales homogéneos del haz $\lambda B - \mu A$, excepto los de la forma $\mu^q$.

Al respecto de estos divisores elementales infinitos (los de la forma $\mu^q$) podemos formular tres proposiciones, a continuación, la primera de las cuales es una conscuencia casi inmediata de lo anterior.

**Proposición 2.4.2.** *Dado el haz $\lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$, el factor invariante homogéneo $i_k(\lambda, \mu)$ aporta el divisor elemental infinito $\mu^{\psi_k}$, si en la fórmula* (2.18)

$$i_k(\lambda, \mu) = \mu^{\psi_k} \mu^{\sigma_k} i_k(\frac{\lambda}{\mu}),$$

*donde $i_k(\lambda)$ es el factor invariante $k$–ésimo de $\lambda B - A$ y $\sigma_k = \partial(i_k(\lambda))$, se verifica que $\psi_k := \partial(i_k(\lambda, \mu)) - \sigma_k \geq 1$.*

*Además, estos polinomios $\mu^{\psi_k}$ son los únicos divisores elementales infinitos (a lo más hay uno por cada factor invariante homogéneo $i_k(\lambda, \mu)$).*

Demostración. Está claro, ya que la Observación 1.4.4 nos asegura que si

$$i_k(\lambda) = \lambda^{\sigma_k} + a_{\sigma_k - 1}\lambda^{\sigma_k - 1} + a_{\sigma_k - 2}\lambda^{\sigma_k - 2} + \cdots + a_1\lambda + a_0,$$

entonces $\mu^{\sigma_k} i_k(\frac{\lambda}{\mu}) \in \mathbb{F}[\lambda, \mu]$ tiene un monomio donde no aparece $\mu$, a saber $\lambda^{\sigma_k}$, por lo que no es múltiplo de $\mu$.

□

**Proposición 2.4.3.** *Los divisores elementales infinitos del haz $\lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$ pueden hallarse buscando todos los divisores elementales de la forma $\mu^q$ (si los hay) de la matriz polinómica $B - \mu A$.*

DEMOSTRACIÓN. Sea $i_k(\mu)$ el $k$–ésimo factor invariante de $B - \mu A$. Por el Lema 2.2.3 (2.19), se tiene que

$$i_k(\mu) = i_k(1, \mu),$$

siendo $i_k(\lambda, \mu)$ el $k$–ésimo factor invariante homogéneo. Buscamos los factores de la forma $\mu^q$ que hay en el polinomio $i_k(\mu)$ ( a lo más habrá uno de estos $\mu^q$ en la factorización de $i_k(\mu)$ en polinomios irreducibles de $\mathbb{F}[\mu]$). Por la fórmula (2.22) y las consideraciones que siguen

$$i_k(1, \mu) = \mu^{\psi_k} \mu^{\sigma_k} i_k(\frac{1}{\mu}),$$

siendo

$$i_k(\lambda, \mu) = \mu^{\psi_k} \mu^{\sigma_k} i_k(\frac{\lambda}{\mu}),$$

con $\psi_k \geq 0$ y $\sigma_k = \partial(i_k(\lambda))$ y el polinomio $\mu^{\sigma_k} i_k(\frac{1}{\mu})$ no es múltiplo de $\mu$, por ser su término constante igual a 1.

De todo lo anterior se deduce que, si $\psi_k \geq 1$, entonces $\mu^{\psi_k}$ es un divisor elemental de $B - \mu A$, que según la Proposición 2.4.2 también es el único divisor elemental infinito del haz $\lambda B - A$ aportado por $i_k(\lambda, \mu)$, cuando $\psi_k \geq 1$.

□

**Proposición 2.4.4.** *El haz regular* $\lambda B - A \in \mathbb{F}[\lambda]^{n \times n}$ *tiene divisores elementales infinitos si y sólo si* $|B| = 0$.

DEMOSTRACIÓN. Sea $D_n(\mu) = |B - \mu A| \neq 0$. Entonces

$$\mu \mid D_n(\mu) \iff D_n(0) = 0,$$

lo que, según (2.1), equivale a $|B| = 0$.

□

*Observación* 2.4.5. La Proposición 2.4.4 es falsa para haces singulares. Por ejemplo, consideremos el haz

$$\lambda B - A := \left( \begin{array}{c|cc} 0 & & \\ \hline & \lambda + 2 & 1 \\ & 0 & \lambda + 2 \end{array} \right),$$

Este haz es singular, pues $|\lambda B - A| = 0$. Sus divisores determinantales homogéneos son claramente

$$1, (\lambda + 2\mu)^2;$$

de ahí que $i_1(\lambda, \mu) = (\lambda + 2\mu)^2$ (nótese que $\psi_1 = 0$), es el único factor invariante homogéneo del haz. De donde sigue que $i_1(\lambda) = (\lambda + 2)^2$ es el único factor invariante de $\lambda B - A$. Como $\lambda + 2$ es un polinomio irreducible en $\mathbb{F}[\lambda]$, tenemos $i_1(\lambda) = e_1(\lambda) = (\lambda + 2)^2$; consecuentemente, $e_1(\lambda, \mu) = (\lambda + 2\mu)^2$ es el único divisor elemental homogéneo del haz . Por lo tanto $\lambda B - A$ no tiene divisores elementales infinitos, a pesar de ser $|B| = 0$.

El Teorema 2.3.1 puede enunciarse de manera equivalente:

**Teorema 2.4.6.** *Dos haces regulares $\lambda B - A, \lambda B_1 - A_1 \in \mathbb{F}[\lambda]^{n\times n}$ son estrictamente equivalentes si y sólo si tienen los mismos divisores elementales finitos (i.e. como $\lambda$–matrices) y los mismos divisores elementales infinitos.*

Para acabar esta sección, damos un ejemplo de dos haces regulares que son equivalentes, pero no estrictamente equivalentes, como anunciábamos antes de Teorema 2.1.1.

**Ejemplo 2.4.7.** Consideremos los haces $H(\lambda) = \lambda B - A$ y $H_1(\lambda) = \lambda B_1 - A_1$, dados como sigue

$$H(\lambda) := \begin{pmatrix} \lambda+2 & \lambda+1 & 2\lambda+3 \\ \lambda+3 & \lambda+2 & 2\lambda+5 \\ \lambda+3 & \lambda+2 & 3\lambda+6 \end{pmatrix}, \quad H_1(\lambda) := \begin{pmatrix} \lambda+2 & \lambda+1 & \lambda+1 \\ \lambda+1 & \lambda+2 & \lambda+1 \\ \lambda+1 & \lambda+1 & \lambda+1 \end{pmatrix}.$$

Mediante operaciones elementales por filas y columnas en $\mathbb{F}[\lambda, \mu]$, efctuadas en los haces homogéneos $H(\lambda, \mu)$ y $H_1(\lambda, \mu)$ obtenemos sendos haces $G(\lambda, \mu)$ y $G_1(\lambda, \mu)$, tales que

$$H(\lambda, \mu) \overset{s}{\sim} G(\lambda, \mu) \quad y \quad H_1(\lambda, \mu) \overset{s}{\sim} G_1(\lambda, \mu),$$

donde sea más fácil calcular los factores invariantes homogéneos, a saber,

$$G(\lambda, \mu) := \begin{pmatrix} \mu & \lambda & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \lambda+\mu \end{pmatrix}, \quad G_1(\lambda, \mu) := \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \lambda+\mu \end{pmatrix};$$

de hecho

$$G(\lambda, \mu) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} H(\lambda, \mu) \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$G_1(\lambda, \mu) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} H_1(\lambda, \mu) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

Ahora es fácil calcular los factores invariantes de $G(\lambda, \mu)$ y $G_1(\lambda, \mu)$, que son respectivamente los de $H(\lambda, \mu)$ y $H_1(\lambda, \mu)$.

En efecto, los de $H(\lambda, \mu)$ son:

$$i_1(\lambda, \mu) = 1,\ i_2(\lambda, \mu) = 1,\ i_3(\lambda, \mu) = \mu^2(\lambda+\mu);$$

por lo tanto sus divisores elementales homogéneos son:

$$e_1(\lambda, \mu) = \mu^2,\ e_2(\lambda, \mu) = (\lambda+\mu),$$

asociados a los polinomios irreducibles $\mu$ y $\lambda + \mu$. Así pues el único divisor elemental finito de $H(\lambda)$ es $\lambda + 1$; y el único divisor elemental infinito es $\mu^2$. Además, $|H(\lambda)| = \lambda + 1$, luego $H(\lambda)$ es un haz regular.

Análogamente, los de $H_1(\lambda, \mu)$ son:

$$i_1(\lambda, \mu) = 1,\ i_2(\lambda, \mu) = \mu,\ i_3(\lambda, \mu) = \mu(\lambda+\mu);$$

por lo tanto sus divisores elementales homogéneos son:

$$e_1(\lambda,\mu) = \mu,\ e_2(\lambda,\mu) = \mu,\ e_3(\lambda,\mu) = (\lambda + \mu),$$

asociados a los polinomios irreducibles $\mu$ y $\lambda + \mu$. Así pues el único divisor elemental finito de $H(\lambda)$ es $\lambda + 1$; y los divisores elementales infinitos son $\mu$, $\mu$. Además, $|H_1(\lambda)| = \lambda + 1$, luego $H_1(\lambda)$ es un haz regular.

De modo que, podemos afirmar que $H(\lambda) \sim H_1(\lambda)$, puesto que tienen los mismos divisores elementales finitos. Sin embargo, como no tienen los mismos divisores elementales infinitos, de acuerdo con los Teoremas 2.3.1 o 2.4.6, se concluye que $H(\lambda) \overset{s}{\not\sim} H_1(\lambda)$; cosa que se sabía de antemano ya que $\mathrm{rg}(B) = 2$ y $\mathrm{rg}(B_1) = 1$, por lo que no pueden existir matrices invertibles $P$ y $Q$ tales que $B_1 = PBQ$.

## 2.5. Forma canónica de Weierstrass

En esta sección emplearemos la siguiente notación: Si $M_1, M_2, \dots, M_p$ son matrices cuadradas de diferentes o iguales órdenes, denotaremos

$$\bigoplus_{i=1}^{p} M_i := M_1 \oplus M_2 \oplus \cdots \oplus M_p := \mathrm{diag}, \{M_1, M_2, \dots, M_p\} :=$$

$$\begin{bmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_p \end{bmatrix}.$$

El Teorema 2.3.1 (o 2.4.6) nos va a permitir obtener una forma canónica para la equivalencia estricta de haces regulares, como sigue:

**Teorema 2.5.1.** *Sea $\lambda B - A \in \mathbb{F}[\lambda]^{n\times n}$ un haz regular. Sea*

$$\mu^{l_i}, \quad i = 1, \dots, r,$$

*el sistema de sus divisores elementales infinitos (donde puede ocurrir que $l_{i_1} = l_{i_2}$, aunque sea $i_1 \neq i_2$); y sea*

$$(\lambda - \lambda_k)^{m_k}, \quad \lambda_k \in \mathbb{F}, \quad k = 1, \dots, s,$$

*el sistema de sus divisores elementales finitos ( donde puede ocurrir también que $\lambda_{k_1} = \lambda_{k_2}$ e incluso $m_{k_1} = m_{k_2}$, aunque sea $k_1 \neq k_2$).*

*Llamando $l := l_1 + \cdots + l_r$ y $m = m_1 + \cdots + m_s$, se tiene que $l + m = n$.*

*Definamos las matrices constantes $N$ y $J$:*

$$N := \bigoplus_{i=1}^{r} N_i, \quad con\ N_i := \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{F}^{l_i\times l_i}\ (i = 1, \dots, r);$$

$$J := \bigoplus_{k=1}^{s} J_k(\lambda_k), \quad \textit{con } J_k(\lambda_k) := \begin{pmatrix} \lambda_k & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix} \in \mathbb{F}^{m_k \times m_k} \ (k = 1, \dots, s).$$

*Entonces el haz*

$$W(\lambda) := \begin{bmatrix} \lambda N - I_l & 0 \\ 0 & \lambda I_m - J \end{bmatrix} \tag{2.36}$$

*es estrictamente equivalente al haz $\lambda B - A$.*

**Definición 2.5.2.** El haz $W(\lambda)$, dado en el teorema anterior, se llama la *forma canónica de Weierstrass* del haz regular $\lambda B - A$. Así pues,

$$W(\lambda) = \text{diag}, \{\lambda N_{l_1} - I_{l_1}, \dots, \lambda N_{l_r} - I_{l_r}, \lambda I_{m_1} - J_1(\lambda_1), \dots, \lambda I_{m_s} - J_s(\lambda_s)\},$$

es decir, $W(\lambda)$ es una matriz diagonal por bloques con $r + s$ bloques en su diagonal, y está unívocamente determinada salvo permutación de estos $r + s$ bloques.

Demostración. **(del Teorema 2.5.1)** Los divisores elementales finitos de la $\lambda$–matriz

$$W(\lambda) = \begin{bmatrix} \lambda N - I_l & 0 \\ 0 & \lambda I_m - J \end{bmatrix}$$

están formados por la unión de los divisores elementales de $\lambda N - I_l$ y de $\lambda I_m - J$.

Los divisores elementales de $\lambda N - I_l$ son la unión de los divisores elementales de cada una de los haces

$$N_i(\lambda) := \lambda N_i - I_{l_i} = \begin{pmatrix} -1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda \\ 0 & 0 & \cdots & -1 \end{pmatrix} \ (i = 1, \dots, r);$$

pero como su divisor determinantal $l_i$–ésimo es $|N_i(\lambda)| = (-1)^{l_i}$, se sigue que los haces $N_i(\lambda)$ no tienen divisores elementales finitos.

En cuanto a los divisores elementales del haz $\lambda I_m - J$ son la unión de los divisores elementales de cada una de las matrices

$$\lambda I_{m_k} - J_k(\lambda_k) = \begin{pmatrix} \lambda - \lambda_k & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \lambda - \lambda_k \end{pmatrix} \ (k = 1, \dots, s),$$

cuyo único divisor elemental es $(\lambda - \lambda_k)^{m_k}$, lo que vale su menor de orden $m_k$, puesto que hay un menor de orden $m_{k-1}$ igual a 1. Así pues, los divisores elementales finitos del haz $W(\lambda)$ son

$$(\lambda - \lambda_k)^{m_1}, \dots, (\lambda - \lambda_k)^{m_s}.$$

Y como

$$|W(\lambda)| = |\lambda N - I_l||\lambda I_m - J| = (-1)^l(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_s)^{m_s},$$

entonces $W(\lambda)$ es regular.

Por otro lado, los divisores elementales infinitos del haz $W(\lambda)$ son los divisores elementales de la forma $\mu^q$ de la $\mu$–matriz

$$\begin{bmatrix} N - \mu I_l & 0 \\ 0 & I_m - \mu J \end{bmatrix}; \tag{2.37}$$

y se obtienen uniendo los de la forma $\mu^q$ de $N - \mu I_l$ y de $I_m - \mu J$.

Se tiene que

$$I_m - \mu J = \bigoplus_{k=1}^{s}(I_{m_k} - \mu J_k(\lambda_k)) = \bigoplus_{k=1}^{s} \begin{pmatrix} 1 - \mu\lambda_k & -\mu & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\mu \\ 0 & 0 & \cdots & 1 - \mu\lambda_k \end{pmatrix}.$$

Como el menor de orden $m_k$ de $I_{m_k} - \mu J_k(\lambda_k)$ es $(1-\mu\lambda_k)^{m_k}$ y $\mu \not| (1-\mu\lambda_k)^{m_k}$, podemos concluir que el haz $I_m - \mu J$ no posee divisores elementales de la forma $\mu^q$. De modo que, los divisores elementales de la forma $\mu^q$ de (2.37) serán los que de esa forma aporten los bloques diagonales de $N - \mu I_l$:

$$N - \mu I_l = \bigoplus_{i=1}^{r}(-\mu I_{l_j} + N_{l_j}) = \bigoplus_{i=1}^{r} \begin{pmatrix} -\mu & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & -\mu \end{pmatrix}.$$

Ahora bien, $\mu^{l_j}$ es el único divisor elemental de $-\mu I_{l_j} + N_{l_j}$. En consecuencia,

$$\mu^{l_1}, \ldots, \mu^{l_r}$$

son los divisores elementales infinitos del haz $W(\lambda)$.

En resumidas cuentas, por el Teorema 2.4.6 se concluye que $W(\lambda) \overset{s}{\sim} \lambda B - A$ (de aquí también se deduce que $W(\lambda)$ es regular) y que $\; l + m = n$ .

□

### 2.5.1. Reducción a la forma canónica de Weierstrass de un haz regular

Para lo que sigue, se supone que $\mathbb{F}$ es algebraicamente cerrado (se necesita para la existencia de la forma canónica de Jordan, de la que se hace uso).

Dado el haz regular $\lambda B - A \in \mathbb{F}[\lambda]^{n\times n}$, nos interesa concretar un proceso de reducción que nos permita obtener matrices de paso $P, Q \in \mathrm{GL}_n(\mathbb{F})$ a su forma canónica de Weierstrass:

$$P(\lambda B - A)Q = W(\lambda).$$

Como el determinante de $\lambda B - A$ es distinto del polinomio nulo, y $\mathbb{F}$ es infinito, existe un $c \in \mathbb{F}$ tal que $|cB - A| \neq 0$. Escribamos

$$\lambda B - A = (\lambda - c)B - (A - cB) = (\lambda - c)B - A_1$$

con $A_1 := A - cB$; así $|A_1| \neq 0$. Multipliquemos a la izquierda el haz $\lambda B - A$ por $A_1^{-1}$

$$A_1^{-1}(\lambda B - A)I_n = (\lambda - c)A_1^{-1}B - I_n.$$

Luego $\lambda B - A \overset{s}{\sim} (\lambda - c)A_1^{-1}B - I_n$. Sea ahora $\bar{P} \in \mathrm{GL}_n(\mathbb{F})$ una matriz que lleva $A_1^{-1}B$ a su forma canónica de Jordan:

$$\bar{P}^{-1}(A_1^{-1}B)\bar{P} = \mathrm{diag}\,\{J_0, J_1\},$$

donde en $J_0$ recogemos todos los bloques de Jordan asociados al valor propio 0 de $A_1^{-1}B$ (observar que la matriz $A_1^{-1}B$ tiene a 0 como valor propio si y sólo si $|B| = 0$), y en $J_1$ recogemos los restantes bloques de Jordan. De modo que, $|J_1| \neq 0$ y $J_0$ es una matriz nilpotente, por ser suma diagonal de bloques del tipo

$$\begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Dado que

$$\bar{P}^{-1}((\lambda - c)A_1^{-1}B - I_n)\bar{P} = (\lambda - c)\bar{P}^{-1}A_1^{-1}B\bar{P} - I_n,$$

tenemos

$$\lambda B - A \overset{s}{\sim} (\lambda - c)\mathrm{diag}\,\{J_0, J_1\} - I_n = \mathrm{diag}\,\{(\lambda - c)J_0 - I_l, (\lambda - c)J_1 - I_m\}$$

siendo $I_l$ la matriz unidad del mismo orden que $J_0$ e $I_m$ la matriz unidad del mismo orden que $J_1$. Por lo que

$$\lambda B - A \overset{s}{\sim} \mathrm{diag}\,\{\lambda J_0 - (I_l + cJ_0), \lambda J_1 - (I_m + cJ_1)\}. \tag{2.38}$$

Como $I_l + cJ_0$ es una matriz triangular superior con todos los elementos de la diagonal principal iguales a 1, existe $(I_l + cJ_0)^{-1}$. Multipliquemos por la izquierda el primer bloque diagonal del segundo miembro de (2.38) por esta matriz inversa. Obtenemos

$$\lambda(I_l + cJ_0)^{-1}J_0 - I_l,$$

donde $(I_l + cJ_0)^{-1}J_0$ también es una matriz nilpotente, por serlo $J_0$; es decir, como existe un entero $j > 0$ (por ejemplo $j = l$) tal que $J_0^j = 0$, entonces la matriz inversa de $(I_l + cJ_0)$ es un polinomio en $J_0$, luego conmuta con $J_0$, a saber,

$$(I_l + cJ_0)^{-1} = I_l - cJ_0 + c^2J_0^2 + \cdots + (-1)^{j-1}c^{j-1}J_0^{j-1},$$

y por ello se verifica

$$[(I_l + cJ_0)^{-1}J_0]^j = [(I_l + cJ_0)^{-1}]^j J_0^j = 0.$$

Por consiguiente, existe una matriz $P_0 \in \mathrm{GL}_l(\mathbb{F})$ que lleva $(I_l + cJ_0)^{-1}J_0$ a su forma canónica de *Jordan*, $N$:

$$P_0^{-1}[(I_l + cJ_0)^{-1}J_0]P_0 = N = \mathrm{diag}\,\{N_{l_1}, \dots, N_{l_r}\},$$

siendo las matrices $N_{l_i}$ como las dadas en el Teorema 2.5.1, y en consecuencia

$$\lambda[(I_l + cJ_0)^{-1}J_0] - I_l \overset{s}{\sim} \mathrm{diag}\,\{\lambda N_{l_1} - I_{l_1}, \dots, \lambda N_{l_r} - I_{l_r}\} = \lambda N - I_l.$$

Ahora, multipliquemos por la derecha el segundo bloque diagonal del segundo miembro de (2.38) por $J_1^{-1}$:

$$[\lambda J_1 - (I_m + cJ_1)]J_1^{-1} = \lambda I_m - (cI_m + J_1^{-1});$$

y a través de una matriz de paso $P_1 \in \mathrm{GL}_m(\mathbb{F})$ reducimos la matriz $cI_m + J_1^{-1}$ a su forma canónica de Jordan, $J$, entre cuyos valores propios puede estar también el 0. Esto es,

$$P_1^{-1}(cI_m + J_1^{-1})P_1 = J.$$

Lo que implica que

$$P_1^{-1}(\lambda I_m - (cI_m + J_1^{-1}))P_1 = \lambda I_m - J.$$

Por este procedimiento, hemos llegado a establecer que

$$\lambda B - A \overset{s}{\sim} \mathrm{diag}\,\{\lambda N - I_l, \lambda I_m - J\}$$

donde la $\lambda$–matriz de la derecha (diagonal por bloques) está en la forma canónica de Weierstrass, que en virtud del Teorema 2.5.1 es única salvo el orden de sus bloques.

En resumidas cuentas, llamando

$$P := \begin{bmatrix} P_0^{-1} & 0 \\ 0 & P_1^{-1} \end{bmatrix} \begin{bmatrix} (I_l + cJ_0)^{-1} & 0 \\ 0 & I_m \end{bmatrix} \bar{P}^{-1} A_1^{-1}$$

y

$$Q := \bar{P} \begin{bmatrix} P_0 & 0 \\ 0 & J_1^{-1}P_1 \end{bmatrix}$$

obtenemos dos matrices $n \times n$ invertibles tales que

$$P(\lambda B - A)Q = \mathrm{diag}\,\{\lambda N - I_l, \lambda I_m - J\}.$$

## 2.6. Ejercicios

En lo que sigue y mientras no se diga otra cosa, $\mathbb{F}$ es un cuerpo cualquiera y $A, B \in \mathbb{F}^{n\times n}$. Un subespacio $\mathcal{N}$ de $\mathbb{F}^{n\times 1}$ se dice que es $(\lambda B - A)$–*invariante* si se verifica

$$\dim(A\mathcal{N} + B\mathcal{N}) \leq \dim(\mathcal{N}).$$

(Esta definición fue introducida, para haces regulares, por Stewart en 1972.)

**2.1.** Sea $\mathcal{N}$ un subespacio de $\mathbb{F}^{n\times 1}$. Demostrar:

(i) $\mathcal{N}$ es $(\lambda I - A)$–invariante si y sólo si $\mathcal{N}$ es $A$–invariante.

(ii) $\mathcal{N}$ es $(\lambda B - I)$–invariante si y sólo si $\mathcal{N}$ es $B$–invariante.

(iii) Si $A\mathcal{N} \subseteq B\mathcal{N}$, entonces $\mathcal{N}$ es $(\lambda B - A)$–invariante.

(iv) Si $B\mathcal{N} \subseteq A\mathcal{N}$, entonces $\mathcal{N}$ es $(\lambda B - A)$–invariante.

(v) Si $P, Q \in \mathrm{GL}_n(\mathbb{F})$, satisfacen

$$P(\lambda B - A)Q = \lambda B_1 - A_1,$$

entonces el subespacio $\mathcal{N}$ es $(\lambda B - A)$–invariante si y sólo si $Q^{-1}\mathcal{N}$ es $(\lambda B_1 - A_1)$–invariante.

(vi) Sea $\det B \neq 0$, entonces $\mathcal{N}$ es $(\lambda B - A)$–invariante si y sólo si $A\mathcal{N} \subseteq B\mathcal{N}$. (*Ayuda*: Utilizar apartados (v) e (i).)

(vii) Sea $\det A \neq 0$, entonces $\mathcal{N}$ es $(\lambda B - A)$–invariante si y sólo si $B\mathcal{N} \subseteq A\mathcal{N}$. (*Ayuda*: Utilizar apartados (v) y (ii).)

**2.2.** Sea $\mathcal{N}$ un subespacio de $\mathbb{F}^{n\times 1}$. Demostrar:

(i) Si $X \in \mathbb{F}^{n\times r}$, para $r = \dim(\mathcal{N})$, denota una matriz base de $\mathcal{N}$, entonces

$$\mathcal{N}\ \textit{es}\ (\lambda B - A) - \textit{invariante} \quad \Longleftrightarrow \quad \mathrm{rg}[AX, BX] \leq \dim(\mathcal{N}).$$

(*Ayuda*: Observar que las columnas de la matriz, dada por bloques, $[AX, BX]$, son un sistema generador del espacio $A\mathcal{N} + B\mathcal{N}$.)

(ii) Si $\mathcal{N}$ y $\mathcal{M}$ son dos subespacios $(\lambda B - A)-$invariantes de $\mathbb{F}^{n\times 1}$ tales que $\mathcal{N} \cap \mathcal{M} = \{0\}$, entonces la suma directa $\mathcal{N} \dotplus \mathcal{M}$ es $(\lambda B - A)$–invariante. (*Ayuda*: Considerar $X, Y$ sendas matrices base de $\mathcal{N}$ y $\mathcal{M}$, respectivamente, y aplicar el apartado anterior.)

**2.3.** Sea $\mathbb{F}$ un cuerpo infinito, $(\lambda B - A)$ un haz regular y $\mathcal{N}$ un subespacio de $\mathbb{F}^{n\times 1}$. Demostrar que

$$\mathcal{N} \text{ es } (\lambda B - A)\text{–invariante} \quad \Longleftrightarrow \quad \dim(A\mathcal{N} + B\mathcal{N}) = \dim(\mathcal{N}).$$

(*Ayuda*: Considerar $\lambda_0 \in \mathbb{F}$ tal que $\det(\lambda_0 B - A) \neq 0$ y observar que

$$(\lambda_0 B - A)\mathcal{N} + B\mathcal{N} = A\mathcal{N} + B\mathcal{N};$$

deducir de ahí que $\dim(A\mathcal{N} + B\mathcal{N}) \geq \dim(\mathcal{N})$. )

# Capítulo 3

# Haces singulares

## 3.1. Teorema de reducción

Vamos a considerar ahora un haz singular de matrices $\lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$. Designamos por $r$ el rango del haz, es decir, el orden del mayor menor no nulo. Por la singularidad del haz, debe ser

$$r < m \ \text{ o } \ r < n,$$

y supongamos para concretar que $r < n$. Entonces las columnas de la $\lambda$–matriz $\lambda B - A$ son linealmente dependientes, tanto sobre $\mathbb{F}[\lambda]$ como sobre $\mathbb{F}(\lambda)$, según la Proposición 1.6.1. Es decir, la ecuación

$$(\lambda B - A)\, x(\lambda) = 0, \quad x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} \tag{3.1}$$

donde $x(\lambda)$ es la columna de las incógnitas, tiene una solución no nula. Toda solución no nula de esta ecuación determina una dependencia lineal entre las columnas de $\lambda B - A$. Nos limitaremos a examinar las soluciones no nulas $x(\lambda)$ de (3.1) que son polinomios en $\lambda$, como vimos en la Subsección 1.6.1; y entre estas soluciones elegiremos las solución polinómica vectorial que tenga el menor grado posible $\varepsilon$, a saber,

$$x(\lambda) := x_0 + \lambda x_1 + \lambda^2 x_2 + \cdots + \lambda^{\varepsilon} x_{\varepsilon} \ \ (x_{\varepsilon} \neq 0),$$

donde cada $x_i \in \mathbb{F}^{n\times 1}$. Obsérvese, que ha de ser $x_0 \neq 0$; si no, $\frac{x(\lambda)}{\lambda}$ sería una solución de (3.1) de grado $\varepsilon - 1$, lo que es imposible.

Sustituyendo esta solución en (3.1) queda

$$(\lambda B - A)\, x(\lambda) = Ax_0 + \sum_{i=0}^{\varepsilon-1}(Bx_i - Ax_{i+1})\lambda^{i+1} + Bx_{\varepsilon}\lambda^{\varepsilon+1} = 0,$$

e igualando a cero los coeficientes de las potencias de $\lambda$, obtenemos

$$Ax_0 = 0,\ Bx_i - Ax_{i+1} = 0\ (i = 0, \ldots, \varepsilon - 1),\ Bx_{\varepsilon} = 0, \tag{3.2}$$

o equivalentemente,

$$Ax_0 = 0, \quad Bx_i = Ax_{i+1} \quad (i = 0, \ldots, \varepsilon - 1), \quad Bx_{\varepsilon} = 0. \tag{3.3}$$

Considerando esto como un sistema de ecuaciones homogéneas lineales, que satisfacen los elementos de las columnas $x_0, x_1, \ldots, x_\varepsilon$, se deduce que la matriz compuesta por $(\varepsilon+2) \times (\varepsilon+1)$ bloques, como sigue

$$M_\varepsilon := M_\varepsilon[\lambda B - A] := \begin{bmatrix} -A & 0 & \cdots & 0 \\ B & -A & \cdots & \vdots \\ 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & -A \\ 0 & 0 & \cdots & B \end{bmatrix}$$

satisface

$$M_\varepsilon \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_\varepsilon \end{bmatrix} = 0;$$

de ahí que

$$\operatorname{rg} M_\varepsilon = \rho_\varepsilon < (\varepsilon+1)n.$$

En virtud de la propiedad minimal de $\varepsilon$, los rangos $\rho_0, \ldots, \rho_{\varepsilon-1}$ de las matrices

$$M_0 = \begin{bmatrix} -A \\ B \end{bmatrix}, M_1 = \begin{bmatrix} -A & 0 \\ B & -A \\ 0 & B \end{bmatrix}, \ldots, M_{\varepsilon-1} = \begin{bmatrix} -A & 0 & \cdots & 0 \\ B & -A & \cdots & \vdots \\ 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & -A \\ 0 & 0 & \cdots & B \end{bmatrix} \tag{3.4}$$

son, respectivamente,

$$\rho_0 = n, \rho_1 = 2n \ldots, \rho_{\varepsilon-1} = \varepsilon n.$$

Así pues, el número $\varepsilon$ es el menor índice $k$ tal que

$$\rho_k < (k+1)n.$$

**Lema 3.1.1.** *En las condiciones anteriores y siguiendo con la misma notación, se verifica:*

*(i)* $Ax_1, \ldots, Ax_\varepsilon \in \mathbb{F}^{m\times 1}$ *son linealmente independientes.*

*(ii)* $x_0, x_1, \ldots, x_\varepsilon \in \mathbb{F}^{n\times 1}$ *son linealmente independientes.*

Demostración. En primer lugar, tiene que ser $Ax_1 \neq 0$; pues si no, por (3.2) sería

$$x_1 + \lambda x_2 + \lambda^2 x_3 + \cdots + \lambda^{\varepsilon-1} x_\varepsilon \quad (x_\varepsilon \neq 0),$$

una solución de (3.2) de grado $< \varepsilon$, lo que es imposible.

Ahora para probar la afirmación $(i)$, supongamos lo contrario, esto es, que $Ax_1, \ldots, Ax_\varepsilon$ son linealmente dependientes. Entonces, como $Ax_1 \neq 0$, existe un vector $Ax_h$ $(1 < h \leq \varepsilon)$, que es combinación lineal de los que le preceden, digamos

$$Ax_h = \alpha_1 Ax_1 + \cdots + \alpha_{h-1} Ax_{h-1},$$

para ciertos elementos $\alpha_i \in \mathbb{F}$. De (3.3) sigue

$$Bx_{h-1} = \alpha_1 Bx_0 + \cdots + \alpha_{h-1} Bx_{h-2},$$

donde $x_{h-i} = 0$, si $h - i < 0$. Luego, para

$$x^*_{h-1} := \alpha_1 x_0 + \cdots + \alpha_{h-1} x_{h-2} - x_{h-1},$$

se tiene

$$Bx^*_{h-1} = 0.$$

Además, otra vez por (3.3) resulta

$$\begin{aligned} Ax^*_{h-1} &= \alpha_1 Ax_0 + \alpha_2 Ax_1 \cdots + \alpha_{h-1} Ax_{h-2} - Ax_{h-1} \\ &= \alpha_2 Bx_0 + \cdots + \alpha_{h-1} Bx_{h-3} - Bx_{h-2}, \end{aligned}$$

donde $x_{h-i} = 0$, si $h - i < 0$. Así para

$$x^*_{h-2} := \alpha_2 x_0 + \cdots + \alpha_{h-1} x_{h-3} - x_{h-2},$$

se verifica

$$Ax^*_{h-1} = Bx^*_{h-2}.$$

Continuando este proceso e introduciendo los vectores

$$x^*_{h-3} := \alpha_3 x_0 + \cdots + \alpha_{h-1} x_{h-4} - x_{h-3},$$

y siguientes, hasta $x^*_0 = x_0$, obtenemos una cadena de ecuaciones

$$Bx^*_{h-1} = 0,\ Ax^*_{h-1} = Bx^*_{h-2}, \ldots, Ax^*_1 = Bx^*_0,\ Ax^*_0 = 0.$$

Esto significa que

$$x^*(\lambda) := x^*_0 + \lambda x^*_1 + \lambda^2 x^*_2 + \cdots + \lambda^{h-1} x^*_{h-1} \quad (x^*_0 = x_0 \neq 0),$$

es una solución no nula de (3.1) de grado $\leq (h-1) < \varepsilon$, lo que contradice la minimalidad de $\varepsilon$. Así pues, queda probado $(i)$.

Ahora es fácil probar el enunciado $(ii)$. En efecto, supongamos

$$k_0 x_0 + k_1 x_1 + \cdots + k_\varepsilon x_\varepsilon = 0, \tag{3.5}$$

con $k_i \in \mathbb{F}$ para $i = 0, 1, \ldots, \varepsilon$. Entonces, multiplicando por $A$ a la izquierda en (3.5) y teniendo en cuenta que $Ax_0 = 0$, sigue

$$k_1 Ax_1 + k_2 Ax_2 + \cdots + k_\varepsilon Ax_\varepsilon = 0;$$

luego, por el apartado $(i)$, obtenemos

$$k_1 = k_2 = \cdots = k_\varepsilon = 0.$$

Ahora, la ecuación (3.5) queda así

$$k_0 x_0 = 0,$$

de donde, por ser $x_0 \neq 0$ se concluye que también $k_0 = 0$. Por consiguiente, hemos probado que $x_0, x_1, \ldots, x_\varepsilon \in \mathbb{F}^{n\times 1}$ son linealmente independientes.

$\square$

Podemos ahora enunciar y demostrar el teorema fundamental siguiente:

**Teorema 3.1.2.** *Sea $\lambda B - A \in \mathbb{F}[\lambda]^{m\times n}$, con* $\operatorname{rg}(\lambda B - A) := r < n$*. Si la solución polinómica vectorial, no nula, de grado minimal, de la ecuación* (3.1)*, que sabemos que existe, tiene grado $\varepsilon > 0$, el haz dado $\lambda B - A$ es estrictamente equivalente a un haz de la forma*

$$\begin{bmatrix} L_\varepsilon & 0 \\ 0 & \lambda\widehat{B} - \widehat{A} \end{bmatrix}, \tag{3.6}$$

*donde*

$$L_\varepsilon = \begin{pmatrix} \lambda & -1 & \cdots & 0 & 0 \\ 0 & \lambda & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \end{pmatrix}$$

*y $\lambda\widehat{B} - \widehat{A}$ es un haz de matrices para el cual la ecuación análoga a* (3.1) *no tiene una solución polinómica vectorial de grado inferior a $\varepsilon$.*

Demostración. Llevaremos a cabo la demostración en tres etapas. Primeramente, vamos a probar que el haz $\lambda B - A$ es estrictamente equivalente a uno de la forma

$$\begin{bmatrix} L_\varepsilon & \lambda D - C \\ 0 & \lambda\widehat{B} - \widehat{A} \end{bmatrix}, \tag{3.7}$$

donde $\widehat{A}, \widehat{B}, C, D$ son matrices rectangulares constantes de dimensiones apropiadas. Luego, mostraremos que la ecuación

$$(\lambda\widehat{B} - \widehat{A})\,\widehat{x} = 0,$$

no tiene soluciones polinómicas vectoriales $\widehat{x}(\lambda)$ no nulas de grado menor que $\varepsilon$. Por último, vamos a probar que mediante otras transformaciones, el haz (3.7) se puede pasar a la forma casidiagonal (3.6).

1.- La primera parte de la demostración será dada bajo la forma geométrica. En efecto, consideremos las aplicaciones lineales

$$\begin{array}{rccc} \mathbb{A}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{m\times 1}, \\ & u & \mapsto & Au \end{array} \qquad \begin{array}{rccc} \mathbb{B}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{m\times 1} \\ & u & \mapsto & Bu \end{array}$$

cuyas matrices asociadas, respecto a las bases canónicas $\boldsymbol{\mathcal{B}}$ de $\mathbb{F}^{n\times 1}$ y $\boldsymbol{\mathcal{B}}'$ de $\mathbb{F}^{m\times 1}$, son respectivamente

$$M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{A}) = A \quad y \quad M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{B}) = B.$$

Por el Lema 3.1.1, tanto $\{x_0, x_1, \ldots, x_\varepsilon\}$, en $\mathbb{F}^{n\times 1}$, como $\{Ax_1, \ldots, Ax_\varepsilon\}$, en $\mathbb{F}^{m\times 1}$, son linealmente independientes, entonces ambos conjuntos se pueden completar hasta sendas bases de $\mathbb{F}^{n\times 1}$ y $\mathbb{F}^{m\times 1}$, respectivamente, digamos, $\boldsymbol{\mathcal{B}}_1 := \{x_0, x_1, \ldots, x_\varepsilon, \ldots\}$ y $\boldsymbol{\mathcal{B}}_1' := \{Ax_1, \ldots, Ax_\varepsilon, \ldots\}$. Ahora, de acuerdo con (3.3), es fácil ver que

$$\bar{A} := M_{\mathcal{B}'_1}^{\mathcal{B}_1}(\mathbb{A}) = \left[\begin{array}{ccccc|ccc} 0 & 1 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & 0 & \ddots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ \hline & & 0 & & & & \widehat{A} & \end{array}\right],$$

donde el primer bloque es de orden $\varepsilon \times (\varepsilon + 1)$ y $\widehat{A}$ designa una matriz de orden $(m - \varepsilon) \times (n - \varepsilon - 1)$, y

$$\bar{B} := M_{\mathcal{B}'_1}^{\mathcal{B}_1}(\mathbb{B}) = \left[\begin{array}{ccccc|ccc} 1 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & 1 & \ddots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & * & \cdots & * \\ \hline & & 0 & & & & \widehat{B} & \end{array}\right],$$

siendo también aquí, el primer bloque, de orden $\varepsilon \times (\varepsilon + 1)$ y $\widehat{B}$ representa una matriz de orden $(m - \varepsilon) \times (n - \varepsilon - 1)$. Entonces existen matrices $P \in \mathrm{GL}_m(\mathbb{F})$ y $Q \in \mathrm{GL}_n(\mathbb{F})$ tales que

$$\bar{A} = PAQ, \ \bar{B} = PBQ,$$

luego

$$P(\lambda B - A)Q = \lambda\bar{B} - \bar{A} = \begin{bmatrix} L_\varepsilon & \lambda D - C \\ 0 & \lambda\widehat{B} - \widehat{A} \end{bmatrix}.$$

como era requerido.

Además, si $x(\lambda) \in \mathbb{F}[\lambda]^{n\times 1}$ es una solución de la ecuación (3.1), entonces $Q^{-1}x(\lambda)$ es solución de la ecuación correspondiente al nuevo haz:

$$(\lambda\bar{B} - \bar{A})\,\bar{x} = 0; \tag{3.8}$$

y recíprocamente, si $y(\lambda)$ es solución de (3.8), entonces $Qy(\lambda)$, lo es de (3.1). Por lo tanto, de la Proposición 1.5.2, se sigue que el mínimo grado de las soluciones polinómicas, no nulas, para ambas ecuaciones (3.1) y (3.8) coincide y, según la hipótesis, su valor es $\varepsilon$.

2.- Vamos a probar ahora que la ecuación

$$(\lambda\widehat{B} - \widehat{A})\,\widehat{x} = 0$$

no tiene una solución no nula de grado menor que $\varepsilon$. Recordemos que, según vimos en el Ejemplo 1.6.9, la ecuación $L_\varepsilon y = 0$ tiene una solución polinómica minimal no nula de grado $\varepsilon$, a saber,

$$\begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^\varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Por otra parte, si el haz tiene la forma triangular (3.7), los haces de matrices correspondientes a ese haz, $M_k$ ($k = 0, 1, \ldots$), tras una permutación adecuada de filas y columnas pueden ponerse en la forma triangular

$$T_k := \begin{bmatrix} M_k[L_\varepsilon] & M_k[\lambda D - C] \\ 0 & M_k[\lambda\widehat{B} - \widehat{A}] \end{bmatrix}.$$

Ahora, por el argumento dado en torno a (3.4), antes del Lema 3.1.1, tanto las columnas de $M_{\varepsilon-1}[L_\varepsilon]$, como las de $M_{\varepsilon-1}[\lambda\bar{B} - \bar{A}]$ y $T_{\varepsilon-1}$, son linealmente independientes. Pero, como $M_{\varepsilon-1}[L_\varepsilon]$ es una matriz cuadrada de orden $\varepsilon(\varepsilon+1)$, se infiere que las columnas de $M_{\varepsilon-1}[\lambda D - C]$ son $\mathbb{F}$–combinación lineal de las columnas de $M_{\varepsilon-1}[L_\varepsilon]$. De este modo, mediante las correspondientes operaciones elementales por columnas, podemos encontrar una matriz $S \in \mathbb{F}^{\varepsilon(\varepsilon+1)\times t}$ tal que

$$\begin{bmatrix} M_{\varepsilon-1}[L_\varepsilon] & M_{\varepsilon-1}[\lambda D - C] \end{bmatrix} \begin{bmatrix} I_{\varepsilon(\varepsilon+1)} & S \\ 0 & I_t \end{bmatrix} = \begin{bmatrix} M_k[L_\varepsilon] & 0 \end{bmatrix},$$

donde $t = \varepsilon n - \varepsilon(\varepsilon+1)$; y obviamente también

$$\begin{bmatrix} M_{\varepsilon-1}[L_\varepsilon] & M_{\varepsilon-1}[\lambda D - C] \\ 0 & M_{\varepsilon-1}[\lambda\widehat{B} - \widehat{A}] \end{bmatrix} \begin{bmatrix} I_{\varepsilon(\varepsilon+1)} & S \\ 0 & I_t \end{bmatrix} = \begin{bmatrix} M_{\varepsilon-1}[L_\varepsilon] & 0 \\ 0 & M_{\varepsilon-1}[\lambda\widehat{B} - \widehat{A}] \end{bmatrix}.$$

Así pues, hemos obtenido que también son linealmente independientes todas las columnas de $M_{\varepsilon-1}[\lambda\widehat{B} - \widehat{A}]$, y eso significa, según se ha explicado justo antes del Lema 3.1.1, que la ecuación

$$(\lambda\widehat{B} - \widehat{A})\,\widehat{x} = 0,$$

no tiene soluciones polinómicas no nulas de grado menor que $\varepsilon$, como teníamos que probar.

3.- Reemplacemos el haz (3.7) por el haz estrictamente equivalente

$$\begin{aligned} &\begin{bmatrix} I_\varepsilon & Y \\ 0 & \lambda I_{m-\varepsilon} \end{bmatrix} \begin{bmatrix} L_\varepsilon & \lambda D - C \\ 0 & \lambda\widehat{B} - \widehat{A} \end{bmatrix} \begin{bmatrix} I_{\varepsilon+1} & -X \\ 0 & \lambda I_{n-\varepsilon-1} \end{bmatrix} \\ &= \begin{bmatrix} L_\varepsilon & -L_\varepsilon X + \lambda D - C + Y(\lambda\widehat{B} - \widehat{A}) \\ 0 & \lambda\widehat{B} - \widehat{A} \end{bmatrix}, \end{aligned} \tag{3.9}$$

donde $X$ e $Y$ son matrices rectangulares constantes, arbitrarias, de dimensiones apropiadas. El teorema quedará completamente demostrado cuando probemos que podemos elegir $X$ e $Y$, de forma que se satisfaga la ecuación

$$L_\varepsilon X = \lambda D - C + Y(\lambda\widehat{B} - \widehat{A}). \tag{3.10}$$

Emplearemos la notación siguiente:

$$\begin{gathered} D = (d_{ik}), \quad C = (c_{ik}), \quad X = (x_{jk}) \\ (i = 1, 2, \ldots, \varepsilon;\ k = 1, 2, \ldots, n - \varepsilon - 1; j = 1, 2, \ldots, \varepsilon + 1), \end{gathered}$$

y para las filas de $Y$ y las columnas de $\widehat{A}$ y $\widehat{B}$

$$Y=\begin{bmatrix} y_1\\ y_2\\ \vdots\\ y_\varepsilon \end{bmatrix}, \widehat{A}=\begin{bmatrix} a_1 & a_2 & \cdots & a_{n-\varepsilon-1}\end{bmatrix}, \widehat{B}=\begin{bmatrix} b_1 & b_2 & \cdots & b_{n-\varepsilon-1}\end{bmatrix}.$$

Entonces, la ecuación matricial (3.10) es equivalente a un sistema de ecuaciones escalares que resultan de igualar los elementos de la $k$–ésima columna de las matrices que están a ambos lados de la igualdad en (3.10), para cada $k = 1, 2, \dots, n-\varepsilon-1$:

$$\begin{aligned} \lambda x_{1k} - x_{2k} &= \lambda d_{1k} - c_{1k} + \lambda y_1 b_k - y_1 a_k,\\ \lambda x_{2k} - x_{3k} &= \lambda d_{2k} - c_{2k} + \lambda y_2 b_k - y_2 a_k,\\ \cdots\cdots&\cdots\cdots\cdots\cdots\\ \lambda x_{\varepsilon k} - x_{(\varepsilon+1)k} &= \lambda d_{\varepsilon k} - c_{\varepsilon k} + \lambda y_\varepsilon b_k - y_\varepsilon a_k,\\ &\qquad\qquad (k = 1, 2, \dots, n-\varepsilon-1). \end{aligned} \tag{3.11}$$

Identificando los coeficientes de $\lambda$ y los términos constantes a ambos lados de la igualdad en cada ecuación de (3.11) resulta

$$\begin{array}{rcccl} & & x_{1k} & = & d_{1k} + y_1 b_k,\\ c_{1k} + y_1 a_k & = & x_{2k} & = & d_{2k} + y_2 b_k,\\ c_{2k} + y_2 a_k & = & x_{3k} & = & d_{3k} + y_3 b_k,\\ \cdots\cdots & & \cdots & & \cdots\cdots\\ c_{\varepsilon k} + y_\varepsilon a_k & = & x_{(\varepsilon+1)k}, & & \\ & & & & (k = 1, 2, \dots, n-\varepsilon-1). \end{array}$$

Por tanto, los elementos de $X$ pueden ser determinados siempre que encontremos las filas $y_k$ de $Y$ que satisfagan las siguientes ecuaciones

$$\begin{aligned} y_2 b_k - y_1 a_k &= c_{1k} - d_{2k},\\ y_3 b_k - y_2 a_k &= c_{2k} - d_{3k},\\ \cdots\cdots & \quad \cdots\cdots\\ y_\varepsilon b_k - y_{\varepsilon-1} a_k &= c_{(\varepsilon-1)k} - d_{\varepsilon k},\\ &\quad (k = 1, 2, \dots, n-\varepsilon-1). \end{aligned} \tag{3.12}$$

Ahora bien, el sistema (3.12), que puede esciribirse más claramente así

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_\varepsilon \end{bmatrix} \begin{bmatrix} -\widehat{A} & 0 & \cdots & 0\\ \widehat{B} & -\widehat{A} & \cdots & \vdots\\ 0 & \widehat{B} & \ddots & \vdots\\ \vdots & \vdots & \ddots & -\widehat{A}\\ 0 & 0 & \cdots & \widehat{B} \end{bmatrix} = (c_{11}-d_{21}, \cdots, c_{(\varepsilon-1)(n-\varepsilon-1)}-d_{\varepsilon(n-\varepsilon-1)})$$

(sistema con $\varepsilon(m-\varepsilon)$ incógnitas y $(\varepsilon-1)(n-\varepsilon-1)$ ecuaciones), siempre tiene solución, cualesquiera que sean

$$c_{ik},\ d_{ik} \quad (i = 1, 2, \dots, \varepsilon; \quad k = 1, 2, \dots, n-\varepsilon-1),$$

porque su matriz de coeficientes es $M_{\varepsilon-2}[\lambda\widehat{B}-\widehat{A}]$; la cual, según se ha visto en la etapa anterior, tiene todas sus columnas linealmente independientes; por tanto, su rango coincide con el número de ecuaciones del sistema, y tales sistemas siempre tienen solución.

$\square$

## 3.2. Forma canónica de un haz singular

Sea $\lambda B - A$ un haz singular arbitrario de matrices m $\times n$. En esta sección obtendremos la forma canónica para este haz, en cada uno de los dos casos posibles, a saber:

1) las filas y las columnas del haz son $\mathbb{F}$–linealmente independientes,

2) las filas o las columnas del haz son $\mathbb{F}$–linealmente dependientes.

Caso 1): Supongamos que todas las filas y las columnas del haz $\lambda B - A$ son $\mathbb{F}$–linealmente independientes ( es claro que, cualquier haz estictamente equivalente a él, también lo verifica). Sea $r < n$, donde $r$ es el rango del haz, de manera que las columnas del haz son $\mathbb{F}[\lambda]$–linealmente dependientes. Entonces, la ecuación $(\lambda B - A)\, x = 0$ tiene una solución polinómica no nula de grado minimal, $\varepsilon_1$, siendo además $\varepsilon_1 > 0$, por la restricción del caso en que estamos. Así, en virtud del Teorema 3.1.2, el haz dado es estrictamente equivalente a

$$\begin{bmatrix} L_{\varepsilon_1} & 0 \\ 0 & \lambda B_1 - A_1 \end{bmatrix},$$

donde la ecuación $(\lambda B_1 - A_1)\, x^{(1)} = 0$ no tiene soluciones polinómicas, no nulas, $x^{(1)}$ de grado menor que $\varepsilon_1$.

Si esta ecuación tiene una solución no nula de grado minimal $\varepsilon_2$ ( necesariamente, $\varepsilon_1 \le \varepsilon_2$), aplicando el Teorema 3.1.2 al haz $\lambda B_1 - A_1$ podemos pasar al haz estrictamente equivalente

$$\begin{bmatrix} L_{\varepsilon_1} & 0 & 0 \\ 0 & L_{\varepsilon_2} & 0 \\ 0 & 0 & \lambda B_2 - A_2 \end{bmatrix}.$$

Continuando con este proceso, se obtiene que el haz $\lambda B - A$ es estrictamente equivalente al siguiente, dado en forma casi diagonal

$$\begin{bmatrix} L_{\varepsilon_1} & & & & 0 \\ & L_{\varepsilon_2} & & & \\ & & \ddots & & \\ & & & L_{\varepsilon_p} & \\ 0 & & & & \lambda B_p - A_p \end{bmatrix}, \tag{3.13}$$

donde $0 < \varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_p$, y la ecuación

$$(\lambda B_p - A_p) x^{(p)} = 0$$

no tiene soluciones no nulas, de modo que todas sus columnas son $\mathbb{F}[\lambda]$–linealmente independientes. Obviamente, si $\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_p = m$, el bloque $\lambda B_p - A_p$ no aparece.

Si las filas de $\lambda B_p - A_p$ son $\mathbb{F}[\lambda]$–linealmente dependientes, el haz traspuesto $\lambda B_p^{\mathrm{T}} - A_p^{\mathrm{T}}$ es estrictamente equivalente a uno de la forma (3.13), donde, en lugar de los números $\varepsilon_1, \ldots, \varepsilon_p$, tenemos $0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$. En consecuencia, el haz dado $\lambda B - A$ es estrictamente equivalente a

$$\begin{bmatrix} L_{\varepsilon_1} & & & & & & & & 0 \\ & L_{\varepsilon_2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & L_{\varepsilon_p} & & & & & \\ & & & & L_{\eta_1}^{\mathrm{T}} & & & & \\ & & & & & L_{\eta_2}^{\mathrm{T}} & & & \\ & & & & & & \ddots & & \\ & & & & & & & L_{\eta_q}^{\mathrm{T}} & \\ 0 & & & & & & & & \lambda B_0 - A_0 \end{bmatrix}, \tag{3.14}$$

$$(0 < \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p,\ 0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q),$$

donde las columnas y las filas de $\lambda B_0 - A_0$ son $\mathbb{F}[\lambda]$–linealmente independientes (también $\mathbb{F}(\lambda)$–l.i., según Proposición 1.6.1), luego $\lambda B_0 - A_0$ es un haz regular.

Evidentemente, si en el haz dado, tenemos $r = n$, es decir si las columnas del haz $\lambda B - A$ son $\mathbb{F}[\lambda]$–linealmente independientes, los $p$ primeros bloques diagonales de (3.14) de la forma $L_\varepsilon$ no aparecen ($p = 0$). Análogamente, si $r = m$, es decir si las filas del haz dado son $\mathbb{F}[\lambda]$–linealmente independientes, entonces los bloques diagonales de la forma $L_\eta^{\mathrm{T}}$ no aparecen en (3.14) ($q = 0$).

Caso 2): Ahora, las columnas o las filas del haz $\lambda B - A$ están ligados por relaciones lineales de coeficientes constantes. Esto equivale a decir que que las ecuaciones

$$(\lambda B - A)\, x = 0 \quad o \quad (\lambda B^{\mathrm{T}} - A^{\mathrm{T}})\, y = 0,$$

respectivamente, tienen soluciones constantes no nulas. Sea $g$ el número máximo de soluciones constantes $\mathbb{F}[\lambda]$–linealmente independientes (equivalente a $\mathbb{F}$–l.i. en este caso) de la primera ecuación y $h$ lo mismo de la segunda. Nótese, que $g$ y $h$ son invariantes para una clase de equivalencia estricta de haces. Naturalmente, $g = 0$ (respectivamente, $h = 0$) cuando las columnas (respectivamente, filas) de $\lambda B - A$ son $\mathbb{F}$–linealmente independientes.

Consideremos, como en la demostración del Teorema 3.1.2, las aplicaciones lineales

$$\begin{array}{cccc} \mathbb{A}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{m\times 1}, \\ & u & \mapsto & Au \end{array} \qquad \begin{array}{cccc} \mathbb{B}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{m\times 1} \\ & u & \mapsto & Bu \end{array}$$

cuyas matrices asociadas, respecto a las bases canónicas $\boldsymbol{\mathcal{B}}$ de $\mathbb{F}^{n\times 1}$ y $\boldsymbol{\mathcal{B}}'$ de $\mathbb{F}^{m\times 1}$, son respectivamente

$$M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{A}) = A \quad y \quad M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}}(\mathbb{B}) = B.$$

Denotemos por $c_1, c_2, \ldots, c_g$ las soluciones constante $\mathbb{F}[\lambda]$–linealmente independientes (es decir, $\mathbb{F}$–l.i.) de $(\lambda B - A)\, x = 0$; de hecho, constituyen una base de

$\ker(\mathbb{A}) \cap \ker(\mathbb{B})$, y los cuales tomaremos como los $g$ primeros vectores de una nueva base $\boldsymbol{\mathcal{B}}_1$ de $\mathbb{F}^{n\times 1}$. Entonces, las $g$ primeras columnas de las dos matrices

$$M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}_1}(\mathbb{A}) = \tilde{A} \quad y \quad M_{\boldsymbol{\mathcal{B}}'}^{\boldsymbol{\mathcal{B}}_1}(\mathbb{B}) = \tilde{B}$$

están formadas por ceros; esto es

$$\lambda\tilde{B} - \tilde{A} = [\ \overbrace{0}^{g} \quad \lambda\tilde{B}_1 - \tilde{A}_1\,].$$

Como el haz $(\lambda\tilde{B}^{\mathrm{T}} - \tilde{A}^{\mathrm{T}})\,y = 0$ tiene $h$ soluciones constantes linealmente independientes, siendo

$$\lambda\tilde{B}^{\mathrm{T}} - \tilde{A}^{\mathrm{T}} = \begin{bmatrix} {}^{g}\{0 \\ \lambda\tilde{B}_1^{\mathrm{T}} - \tilde{A}_1^{\mathrm{T}} \end{bmatrix},$$

entonces la ecuación $(\lambda\tilde{B}_1^{\mathrm{T}} - \tilde{A}_1^{\mathrm{T}})\,z = 0$ tiene también $h$ soluciones constantes linealmente independientes; luego razonando análogamente, existirán $P \in \mathrm{GL}_{n-g}(\mathbb{F})$ y $Q \in \mathrm{GL}_m(\mathbb{F})$ tal que

$$P(\lambda\tilde{B}_1^{\mathrm{T}} - \tilde{A}_1^{\mathrm{T}})Q = [\ \overbrace{0}^{h} \quad \lambda\tilde{B}_2^{\mathrm{T}} - \tilde{A}_2^{\mathrm{T}}\,];$$

de modo que

$$\begin{bmatrix} I_g & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} {}^{g}\{0 \\ \lambda\tilde{B}_1^{\mathrm{T}} - \tilde{A}_1^{\mathrm{T}} \end{bmatrix} Q = \begin{bmatrix} {}^{g}\{0 \\ P(\lambda\tilde{B}_1^{\mathrm{T}} - \tilde{A}_1^{\mathrm{T}})Q \end{bmatrix} = \begin{bmatrix} {}^{g}\{\ \overbrace{0}^{h} & 0 \\ 0 & \lambda\tilde{B}_2^{\mathrm{T}} - \tilde{A}_2^{\mathrm{T}} \end{bmatrix}.$$

Luego, trasponiendo en esta identidad, se sigue que

$$\lambda\tilde{B} - \tilde{A} \overset{s}{\sim} \begin{bmatrix} {}^{h}\{\ \overbrace{0}^{g} & 0 \\ 0 & \lambda\tilde{B}_2 - \tilde{A}_2 \end{bmatrix},$$

donde ya no hay dependencia lineal, con coeficientes constantes, entre las filas o las columnas del haz $\lambda\tilde{B}_2 - \tilde{A}_2$. Por tanto, éste es estrictamente equivalente a uno del tipo (3.14).

Así pues, en el caso general, el haz $\lambda B - A$ siempre puede ser llevado, mediante equivalencia estricta, a la forma canónica casi diagonal

$$\operatorname{diag}\{\ {}^{h}\{\overbrace{\ 0\ }^{g}, L_{\varepsilon_{g+1}}, \ldots, L_{\varepsilon_p}, L_{\eta_{h+1}}^{\mathrm{T}}, \ldots, L_{\eta_q}^{\mathrm{T}}, \lambda B_0 - A_0\}. \tag{3.15}$$

La elección de los índices de $\varepsilon$ y $\eta$ está asociado al hecho de que conviene tomar

$$\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_g = 0 \quad y \quad \eta_1 = \eta_2 = \ldots = \eta_h = 0.$$

Reemplazando, en (3.15), el haz regular $\lambda B_0 - A_0$ por su forma canónica de Weierstrass dada en Teorema 2.5.1 y subsección 2.5.1 (recordar que para que exista, en general, se precisa que $\mathbb{F}$ sea algebraicamente cerrado), obtenemos finalmente la matriz casi diagonal siguiente

$$\operatorname{diag}\{\ {}^{h}\{\overbrace{\ 0\ }^{g}, L_{\varepsilon_{g+1}}, \ldots, L_{\varepsilon_p}, L_{\eta_{h+1}}^{\mathrm{T}}, \ldots, L_{\eta_q}^{\mathrm{T}}, \lambda N - I_{l_0}, \lambda I_{m_0} - J\}. \tag{3.16}$$

Para determinar inmediatamente la forma canónica (3.16) de un haz dado sin pasar por el proceso de reducciones sucesivas, vamos a introducir en la sección siguiente, el concepto de índices minimales de Kronecker de un haz.

## 3.3. Indices minimales de Kronecker

Sea $\mathbb{F}$ un cuerpo algebraicamente cerrado y sea $\lambda B - A$ un haz singular arbitrario de matrices rectangulares $m \times n$. Las $k$ columnas polinómicas $x_1(\lambda), x_2(\lambda), \dots, x_k(\lambda)$ que son las soluciones de la ecuación

$$(\lambda B - A)\, x = 0, \tag{3.17}$$

son linealmente dependientes en $\mathbb{F}[\lambda]$ o en $\mathbb{F}(\lambda)$, según la Proposición 1.6.1, si el rango de la matriz polinómica,

$$X = \begin{bmatrix} x_1(\lambda) & x_2(\lambda) & \cdots & x_k(\lambda) \end{bmatrix},$$

formada a partir de sus columnas es menor que $k$. En ese caso, existen $k$ polinomios $p_1(\lambda), p_2(\lambda), \dots, p_k(\lambda)$, de los cuales no todos son idénticamente nulos, tales que

$$p_1(\lambda)x_1(\lambda) + p_2(\lambda)x_2(\lambda) + \ldots + p_k(\lambda)x_k(\lambda) = 0.$$

Pero si el rango de $X$ es $k$, una tal dependencia no existe y las soluciones $x_1(\lambda), x_2(\lambda), \dots, x_k(\lambda)$ son linealmente independientes.

Entre todas las soluciones de (3.17), elegimos una solución no nula $x_1(\lambda)$ de grado minimal $\varepsilon_1$. Entre todas las soluciones de la misma ecuación linealmente independientes de $x_1(\lambda)$, elegimos una solución $x_2(\lambda)$ de grado minimal $\varepsilon_2$. Es evidente que $\varepsilon_1 \leq \varepsilon_2$. Continuamos el proceso eligiendo, entre las soluciones que son linealmente independientes de $x_1(\lambda)$ y $x_2(\lambda)$, una solución $x_3(\lambda)$ de grado minimal $\varepsilon_3$, etc. Puesto que el número de soluciones linealmente independientes de (3.17) es a lo sumo $n$, el proceso es finito. Obtenemos así una serie fundamental de soluciones de (3.17) ( también llamada una base minimal del subespacio de las soluciones de (3.17))

$$x_1(\lambda), x_2(\lambda), \dots, x_p(\lambda)$$

de grado

$$\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p.$$

En general, una serie fundamental de soluciones no está determinada de manera única (salvo factores escalares) por el haz $\lambda B - A$. Sin embargo, dos series fundamentales de soluciones distintas tienen siempre una única serie de grados $\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$, tal y como se demostró en la Proposición 1.6.4, que son llamados los *índices minimales para las columnas* (o también, índices minimales a derecha) del haz $\lambda B - A$.

Los índices minimales $\eta_1, \eta_2, \dots, \eta_q$ para las filas (o índices minimales a izquierda) del haz $\lambda B - A$ son introducidos de la misma manera, sustituyendo la ecuación (3.17)) por $(\lambda B^{\mathrm{T}} - A^{\mathrm{T}})\, y = 0$, y definiendo $\eta_1, \eta_2, \dots, \eta_q$ como los índices minimales para las columnas del haz traspuesto $\lambda B^{\mathrm{T}} - A^{\mathrm{T}}$.

Además, haces estrictamente equivalentes tienen los mismos índices minimales, según se ha probado ya en la Proposición 1.6.8.

Calculemos los índices minimales para la matriz casi diagonal canónica

$$\operatorname{diag}\{\, {}^{h}\{\overbrace{\;0\;}^{g}, L_{\varepsilon_{g+1}}, \dots, L_{\varepsilon_p}, L^{\mathrm{T}}_{\eta_{h+1}}, \dots, L^{\mathrm{T}}_{\eta_q}, \lambda B_0 - A_0\}, \tag{3.18}$$

donde $\lambda B_0 - A_0$ es un haz regular teniendo la forma canónica de Wierstrass dada en el Teorema 2.5.1.

Notemos en primer lugar, que el sistema completo de los índices para las columnas (filas) de una matriz casi diagonal se obtiene de la unión de los sistemas correspondientes de índices minimales de los bloques diagonales individuales, resultado probado en el Corolario 1.6.14.

Se ha visto, en el Ejemplo 1.6.9, que la matriz $L_\varepsilon$ no tiene más que un solo índice $\varepsilon$ para las columnas, y todas sus filas son linealmente independientes. Análogamente, la matriz $L_\eta^{\mathrm{T}}$ no tiene más que un solo índice $\eta$ para las filas, y todas sus columnas son linealmente independientes. El haz regular $\lambda B_0 - A_0$ no tiene índices minimales. De ahí sigue que la matriz (3.18) tiene como índices minimales para las columnas

$$\varepsilon_1 = \cdots = \varepsilon_g = 0,\ \varepsilon_{g+1}, \ldots, \varepsilon_p,$$

y para las filas

$$\eta_1 = \cdots = \eta_h = 0,\ \eta_{h+1}, \ldots, \eta_q.$$

Por otro lado, la matriz $L_\varepsilon$ no tiene divisores elementales ya que, entre sus menores de orden máximo, hay uno igual a 1 y otro igual a $\lambda^\varepsilon$, y lo mismo sucede con la matriz traspuesta $L_\varepsilon^{\mathrm{T}}$. Puesto que los divisores elementales de una matriz casi diagonal por bloques se obtienen uniendo los de los bloques diagonales individuales (v. Gantmacher vol. 1, chap. 6, p. 143), los divisores elementales de la $\lambda$–matriz (3.18) coinciden con los de su parte regular $\lambda B_0 - A_0$.

Así pues, la forma canónica del haz (3.18) está completamente determinada por los índices minimales $\varepsilon_1, \ldots, \varepsilon_p, \eta_1, \ldots, \eta_q$ y los divisores elementales de este haz o, lo que es lo mismo, del haz estrictamente equivalente $\lambda B - A$. Como dos haces que tienen la misma forma canónica son estrictamente equivalentes, hemos probado el teorema siguiente.

**Teorema 3.3.1** (Kronecker)**.** *Sea $\mathbb{F}$ un cuerpo algebraicamente cerrado. Entonces, dos haces arbitrarios $\lambda B - A$ y $\lambda B_1 - A_1$ de matrices rectangulares sobre el cuerpo $\mathbb{F}$, de las mismas dimensiones $m \times n$, son estrictamente equivalentes si y sólo si tienen los mismos índices minimales y los mismos divisores elementales (finitos e infinitos).*

A modo de ilustración, conviene releer el Ejemplo 1.1.4 dado en la sección 1.1.

**Ejemplo 3.3.2.** Consideramos el haz siguiente

$$\lambda B - A := \begin{pmatrix} \lambda & 0 & \lambda & 1 \\ 0 & \lambda+1 & \lambda+1 & 0 \\ -1 & \lambda+1 & \lambda & -\lambda \end{pmatrix},$$

que tiene rango 3 y cuatro columnas, luego existe un índice minimal por columnas $\epsilon$. Procediendo como en el Ejemplo 1.5.6 mediante operaciones elementales por filas y columnas obtenemos la forma normal de Smith, a saber

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda & -\lambda & 1 \end{pmatrix} (\lambda B - A) \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 1 & -\lambda & \lambda & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & 1-\lambda^2 & 0 \end{pmatrix},$$

de modo que los invariantes del haz singular $\lambda B - A$ son $\varepsilon = 0$ ya que

$$(\lambda B - A) \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

y los divisores elementales finitos $\lambda+1, \lambda+1$ y $\lambda-1$. Luego la forma canónica de Kronecker de este haz es

$$\lambda B_c - A_c = \begin{pmatrix} 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & \lambda+1 & 0 \\ 0 & 0 & 0 & \lambda-1 \end{pmatrix}.$$

Ahora, para obtener las matrices escalares invertibles $P$ y $Q$ tales que

$$P(\lambda B - A)Q = \lambda B_c - A_c$$

se puede proceder en varios pasos. Primero, procediendo como se indica en el Caso 2) de la Sección 3.2, tenemos

$$(\lambda B - A)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \lambda & 1 \\ 0 & \lambda+1 & \lambda+1 & 0 \\ 0 & \lambda+1 & \lambda & -\lambda \end{pmatrix}.$$

A continuación, se puede aplicar lo visto en la Sección 2.5 al haz regular

$$\lambda B_0 - A_0 = \begin{pmatrix} 0 & \lambda & 1 \\ \lambda+1 & \lambda+1 & 0 \\ \lambda+1 & \lambda & -\lambda \end{pmatrix}.$$

Otra posibilidad también es hacer operaciones elementales de filas y columnas para llegar a obtener el haz $\lambda B_c - A_c$, como sigue

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{pmatrix}(\lambda B - A)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & \lambda+1 & 1 \\ 0 & 0 & 0 & \lambda-1 \end{pmatrix}.$$

Ahora, en la esquina inferior derecha tenemos $\lambda I_2 - C$, siendo $C = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$, y hallando la forma canónica de Jordan de $C$ se tiene que

$$\begin{pmatrix} 1 & 1/2 \\ 0 & -1/2 \end{pmatrix} C \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Así pues, hemos obtenido que

$$P(\lambda B - A)Q = \lambda B_c - A_c$$

para las matrices invertibles siguientes

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \\ 0 & 0 & -1/2 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \end{pmatrix},$$

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

## 3.4. Caracterización de los subespacios invariantes de haces cuadrados

Las definiciones siguientes de subespacio invariante fueron avanzadas en la Sección 2.6, Ejercicios 2.1, 2.1 y 2.3. Las repetimos con generalidad mayor. Sean las matrices $A, B \in \mathbb{C}^{m\times n}$, $\boldsymbol{\mathcal{M}}$ un subespacio $r$–dimensional de $\mathbb{C}^{n\times 1}$, y $X \in \mathbb{C}^{n\times r}$ una matriz base de $\boldsymbol{\mathcal{M}}$.

**Definición 3.4.1.** El haz $\lambda BX - AX$ será llamado una restricción del haz $\lambda B - A$ al subespacio $\boldsymbol{\mathcal{M}}$.

Obsérvese que dado el subespacio $\boldsymbol{\mathcal{M}}$, el haz $\lambda BX - AX$ no queda determinado unívocamente, ya que la definición pasa por la elección particular de una matriz base de $\boldsymbol{\mathcal{M}}$; si $X'$ es otra matriz base de $\boldsymbol{\mathcal{M}}$, entonces los haces $\lambda BX' - AX'$ y $\lambda BX - AX$ son estrictamente equivalentes. Así pues, todas las restricciones de $\lambda B - A$ a $\boldsymbol{\mathcal{M}}$ tienen la misma forma canónica de Kronecker. Denotamos a esta clase de haces de $\mathbb{C}[\lambda]^{m\times r}$ estrictamente equivalentes, mediante $\lambda B - A\big|_{\boldsymbol{\mathcal{M}}}$.

**Proposición 3.4.2.** *Sean $A, B \in \mathbb{C}^{n\times n}$ matrices tales que el haz $\lambda B - A$ es regular, y sea $\boldsymbol{\mathcal{M}}$ un subespacio de $\mathbb{C}^{n\times 1}$. Entonces la clase $\lambda B - A\big|_{\boldsymbol{\mathcal{M}}}$ no tiene índices minimales por columnas.*

La demostración forma la parte (iii) del Ejercicio 3.2. Se habrá podido leer al comienzo de la Sección 2.6, que un subespacio $\boldsymbol{\mathcal{M}}$ de $\mathbb{C}^{n\times 1}$ se llama subespacio invariante de un haz cuadrado $\lambda B - A$ si

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) \leq \dim \boldsymbol{\mathcal{M}}. \tag{3.19}$$

Veamos que este concepto se llena de significado cuando el haz $\lambda B - A$ es regular.

**Teorema 3.4.3.** *Sean las matrices $A, B \in \mathbb{C}^{n\times n}$ tales que el haz $\lambda B - A$ es regular, y sea $\boldsymbol{\mathcal{N}}$ un subespacio $r$–dimensional de $\mathbb{C}^{n\times 1}$. Entonces $\boldsymbol{\mathcal{N}}$ es un subespacio invariante del haz $\lambda B - A$ si y sólo si la clase $\lambda B - A\big|_{\boldsymbol{\mathcal{N}}}$ no tiene índices minimales por columnas, ni índices minimales por filas mayores que cero, y caso de tener índices minimales por filas éstos consisten en $n - r$ ceros.*

*Observación* 3.4.4. La clase $\lambda B - A\big|_{\boldsymbol{\mathcal{N}}}$ puede tener divisores elementales.

La demostración del Teorema 3.4.3 es la parte (vi) del Ejercicio 3.2.

*Observación* 3.4.5. Si el haz $\lambda B - A \in \mathbb{C}[\lambda]^{n\times n}$ es regular y $\boldsymbol{\mathcal{M}}$ es un subespacio de $\mathbb{C}^{n\times 1}$, entonces

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) \geq \dim \boldsymbol{\mathcal{M}}. \tag{3.20}$$

Para demostrar (3.20) véase la ayuda del Ejercicio 2.3. La desigualdad (3.20) demuestra que en la definición de subespacio invariante ("deflating subspace") dada por Stewart se debería usar el signo de igualdad en vez de $\leq$, ya que no hay ningún subespacio $\boldsymbol{\mathcal{M}}$ de $\mathbb{C}^{n\times 1}$ para el que

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) < \dim \boldsymbol{\mathcal{M}}. \tag{3.21}$$

Sin embargo, si el haz $\lambda B - A \in \mathbb{C}[\lambda]^{n\times n}$ no es regular puede ocurrir la desigualdad estricta (3.21). Lo que es puesto de manifiesto por el teorema siguiente.

**Teorema 3.4.6.** *Sean $A, B \in \mathbb{C}^{n\times n}$ matrices cualesquiera y sea $\boldsymbol{\mathcal{M}}$ un subespacio de $\mathbb{C}^{n\times 1}$. Entonces*

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) < \dim \boldsymbol{\mathcal{M}}$$

*si y sólo si el número de índices minimales por filas no nulos es menor que el número de índices minimales por columnas de la clase $\lambda B - A|_{\boldsymbol{\mathcal{M}}}$.*

Demostración. Supongamos que $\dim \boldsymbol{\mathcal{M}} = r$ y sea $X \in \mathbb{C}^{n\times r}$ una matriz base de $\boldsymbol{\mathcal{M}}$; consideremos el haz $\lambda BX - AX$. Existen matrices invertibles $R \in \mathbb{C}^{n\times n}$ y $S \in \mathbb{C}^{r\times r}$ que transforman este haz a su forma de Kronecker:

$$\lambda RBXS - RAXS; \tag{3.22}$$

obviamente,

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) = \mathrm{rg}[BX, AX] = \mathrm{rg}[RBXS, RAXS].$$

La forma típica de (3.22) será

$$RBXS = \begin{pmatrix} 0_{h\times g} & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & [I_{\varepsilon_i}, 0] & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & \begin{bmatrix} I_{\eta_j} \\ 0 \end{bmatrix} & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & I_k & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & N_\ell & \\ & & & & & & & & & \ddots \end{pmatrix},$$

$$RAXS = \begin{pmatrix} 0_{h\times g} & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & [0, I_{\varepsilon_i}] & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & \begin{bmatrix} 0 \\ I_{\eta_j} \end{bmatrix} & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & J_k(\alpha) & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & I_\ell & \\ & & & & & & & & & \ddots \end{pmatrix}.$$

Por tanto,

$$\begin{aligned}
&\operatorname{rg}[RBXS, RAXS] = \\
&\sum_{i=g+1}^{p} \varepsilon_i + \sum_{j=h+1}^{q} (\eta_j + 1) + \sum_{i=1}^{t} k_i + \sum_{i=1}^{s} \ell_i \\
&= q - h + \sum_{i=g+1}^{p} \varepsilon_i + \sum_{j=h+1}^{q} \eta_j + \sum_{i=1}^{t} k_i + \sum_{i=1}^{s} \ell_i \\
&= q - h + \nu,
\end{aligned}$$

llamando

$$\nu := \sum_{i=g+1}^{p} \varepsilon_i + \sum_{j=h+1}^{q} \eta_j + \sum_{i=1}^{t} k_i + \sum_{i=1}^{s} \ell_i.$$

Este número es la suma de todos los invariantes enteros del sistema completo de invariantes de $\lambda BX - AX$. En resumidas cuentas,

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) = q - h + \nu;$$

notemos que $q - h$ es el número de índices minimales por filas no nulos.

Observando la forma canónica de Kronecker de un haz rectangular cualquiera $\lambda F - G$ se ve que el número de columnas del haz debe ser igual a la suma de todos sus invariantes enteros más el número total de índices minimales por columnas. Aplicando esta igualdad al haz $\lambda BX - AX$ se tiene que $r = p + \nu$, siendo $p$ el número total de índices minimales por columnas. De todo esto se deduce que si $q - h < p$, se seguirá la desigualdad

$$\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) = q - h + \nu < p + \nu = r = \dim \boldsymbol{\mathcal{M}}.$$

Recíprocamente, si $\dim(A\boldsymbol{\mathcal{M}} + B\boldsymbol{\mathcal{M}}) < \dim \boldsymbol{\mathcal{M}}$, entonces $q - h + \nu < p + \nu$; de donde $q - h < p$. $\square$

## 3.5. Ejercicios

**3.1.** Sea

$$\lambda B - A := \begin{pmatrix} 2\lambda - 2 & \lambda & 0 & 0 & -1 \\ \lambda + 1 & 2\lambda - 1 & -1 & \lambda - 2 & 0 \\ 3\lambda - 1 & 4\lambda - 2 & -1 & 2\lambda - 3 & 0 \\ 1 & \lambda - 1 & 0 & \lambda - 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{C}[\lambda]^{5\times 5}.$$

(i) Proceder como en el Ejemplo 1.5.6 para obtener la forma normal de Smith del haz $\lambda B - A$. Deducir que $\lambda B - A$ tiene un índice minimal por columnas $\varepsilon = 3$, tiene un índice minimal por filas $\eta = 1$ y no tiene divisores elementales finitos.

(ii) Aplicar el método de la demostración del Teorema 3.1.2 para encontrar matrices $P, Q \in \mathrm{GL}_5(\mathbb{C})$ tales que

$$P(\lambda B - A)Q = \lambda B_c - A_c,$$

donde $\lambda B_c - A_c$ denota la forma canónica de Kronecker del haz $\lambda B - A$.

**3.2.** Sea $\boldsymbol{\mathcal{N}}$ un subespacio $(\lambda B - A)$-invariante (véase la Sección 2.6), con $\dim(\boldsymbol{\mathcal{N}}) = r$, y sea $\boldsymbol{\mathcal{V}} := A\boldsymbol{\mathcal{N}} + B\boldsymbol{\mathcal{N}}$ con $\dim(\boldsymbol{\mathcal{V}}) = s$, luego $s \leq r$. Sea $X$ una matriz base cualquiera de $\boldsymbol{\mathcal{N}}$, es decir las columnas de $X = \begin{bmatrix} x_1 & \cdots & x_r \end{bmatrix}$, constituyen una base de $\boldsymbol{\mathcal{N}}$, y sea $Y$ una matriz base cualquiera de $\boldsymbol{\mathcal{V}}$, es decir las columnas de $Y = \begin{bmatrix} y_1 & \cdots & y_s \end{bmatrix}$, constituyen una base de $\boldsymbol{\mathcal{V}}$. Consideramos

$$\begin{array}{cccc} \mathbb{A}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{n\times 1}, \\ & u & \mapsto & Au \end{array} \qquad \begin{array}{cccc} \mathbb{B}: & \mathbb{F}^{n\times 1} & \longrightarrow & \mathbb{F}^{n\times 1} \\ & u & \mapsto & Bu \end{array}$$

cuyas matrices asociadas, respecto a la base canónica $\boldsymbol{\mathcal{B}}$ de $\mathbb{F}^{n\times 1}$ , son respectivamente

$$M_{\boldsymbol{\mathcal{B}}}^{\boldsymbol{\mathcal{B}}}(\mathbb{A}) = A \quad y \quad M_{\boldsymbol{\mathcal{B}}}^{\boldsymbol{\mathcal{B}}}(\mathbb{B}) = B.$$

(i) Considerar las restricciones de $\mathbb{A}$ y $\mathbb{B}$ a $\boldsymbol{\mathcal{N}}$, como sigue

$$\begin{array}{cccc} \mathbb{A}|_{\boldsymbol{\mathcal{N}}}: & \boldsymbol{\mathcal{N}} & \longrightarrow & \mathbb{F}^{n\times 1}, \\ & u & \mapsto & Au \end{array} \qquad \begin{array}{cccc} \mathbb{B}|_{\boldsymbol{\mathcal{N}}}: & \boldsymbol{\mathcal{N}} & \longrightarrow & \mathbb{F}^{n\times 1} \\ & u & \mapsto & Bu \end{array}$$

y probar que las matrices asociadas, respecto a las bases $X$ de $\boldsymbol{\mathcal{N}}$ y la canónica $\boldsymbol{\mathcal{B}}$ de $\mathbb{F}^{n\times 1}$ , son respectivamente

$$M_{\boldsymbol{\mathcal{B}}}^{X}(\mathbb{A}|_{\boldsymbol{\mathcal{N}}}) = AX \quad y \quad M_{\boldsymbol{\mathcal{B}}}^{X}(\mathbb{B}|_{\boldsymbol{\mathcal{N}}}) = BX.$$

El haz $\lambda BX - AX \in \mathbb{F}[\lambda]^{n\times r}$ se llama *la restricción del haz $\lambda B - A$ al subespacio invariante $\boldsymbol{\mathcal{N}}$*.

(ii) Demostrar que si $X'$ es otra base de $\boldsymbol{\mathcal{N}}$, entonces

$$\lambda BX - AX \overset{s}{\sim} \lambda BX' - AX'.$$

(iii) Demostrar que si existe $v(\lambda) \in \mathbb{F}[\lambda]^{r\times 1}, v(\lambda) \neq 0$, tal que

$$(\lambda BX - AX)\, v(\lambda) = 0$$

entonces $(\lambda B - A)\, Xv(\lambda) = 0$ con $Xv(\lambda) \neq 0$. Deducir de ahí que si una restricción del haz $\lambda B - A$ al subespacio $\boldsymbol{\mathcal{N}}$ tiene índices minimales por columnas, entonces el haz $\lambda B - A$ también tiene, y por tanto éste ha de ser un haz singular.

(iv) Considerar ahora las aplicaciones lineales siguientes

$$\begin{array}{cccc} \mathbb{A}_1: & \boldsymbol{\mathcal{N}} & \longrightarrow & \boldsymbol{\mathcal{V}}, \\ & u & \mapsto & Au \end{array} \qquad \begin{array}{cccc} \mathbb{B}_1: & \boldsymbol{\mathcal{N}} & \longrightarrow & \boldsymbol{\mathcal{V}} \\ & u & \mapsto & Bu \end{array}$$

cuyas matrices asociadas, respecto a las bases $X$ de $\boldsymbol{\mathcal{N}}$ e $Y$ de $\boldsymbol{\mathcal{V}}$ , son respectivamente

$$M_{Y}^{X}(\mathbb{A}_1) = A_1 \quad y \quad M_{Y}^{X}(\mathbb{B}_1) = B_1.$$

Demostrar que si completamos la matriz base de $\boldsymbol{\mathcal{N}}$, $X$, hasta una matriz base $P$ de $\mathbb{F}^{n\times 1}$, esto es, $P = \begin{bmatrix} X & x_{r+1} & \cdots & x_n \end{bmatrix}$, y completamos la matriz base de $\boldsymbol{\mathcal{V}}$, $Y$, hasta una matriz base $Q$ de $\mathbb{F}^{n\times 1}$, esto es, $Q = \begin{bmatrix} Y & y_{s+1} & \cdots & y_n \end{bmatrix}$, entonces

$$Q^{-1}(\lambda B - A)P = \begin{bmatrix} \lambda B_{11} - A_{11} & \lambda B_{12} - A_{12} \\ 0 & \lambda B_{22} - A_{22} \end{bmatrix}, \qquad (3.23)$$

donde $\lambda B_{11} - A_{11} \in \mathbb{F}[\lambda]^{r\times r}$, siendo $A_{11} := \begin{bmatrix} A_1 \\ 0 \end{bmatrix}$, $B_{11} := \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$, y

$$Q^{-1}(\lambda BX - AX) = \begin{bmatrix} \lambda B_1 - A_1 \\ 0 \end{bmatrix}. \qquad (3.24)$$

(v) Aplicando determinantes en (3.23), probar que el haz $\lambda B - A$ es regular si y sólo si los dos haces $\lambda B_{11} - A_{11}$ y $\lambda B_{22} - A_{22}$ son regulares. Deducir que $\lambda B_{11} - A_{11}$ es regular si y sólo si $s = r$, siendo $\lambda B_{11} - A_{11} = \lambda B_1 - A_1$, en este caso. Concluir que $\boldsymbol{\mathcal{N}}$ es un subespacio invariante de un haz regular $\lambda B - A$ si y sólo si

$$\dim (A\boldsymbol{\mathcal{N}} + B\boldsymbol{\mathcal{N}}) = \dim (\boldsymbol{\mathcal{N}}).$$

(vi) Deducir de (3.24) y del apartado anterior, que si $\lambda B - A$ es regular la forma canónica de Kronecker del haz $\lambda BX - AX$ sólo tiene divisores elementales, los del haz regular $\lambda B_1 - A_1$, y $n - r$ índices minimales por filas iguales a cero.

# Capítulo 4

# Invariantes por rangos

En este capítulo vamos a exponer un método que permite calcular los invariantes enteros, exponentes de los divisores elementales e índices minimales, de un haz mediante el cálculo de los rangos de determinadas matrices; estas matrices son denominadas de Gantmacher, pues él fue quien primero las consideró en su exposición sobre haces. Para el cálculo de los exponentes de los divisores elementales finitos supondremos que los valores propios del haz son conocidos de manera exacta. En lo que sigue consideramos $\mathbb{F} = \mathbb{C}$, el cuerpo de los números complejos.

## 4.1. Definiciones y notaciones

### 4.1.1. Particiones de enteros

Una *partición* es una sucesión finita o infinita de enteros no negativos

$$a = (a_1, a_2, \ldots)$$

ordenados en sentido decreciente y tal que sólo hay un número finito de términos diferentes de cero,

$$a_1 \geq a_2 \geq \cdots \geq a_{\ell(a)} > 0 = a_{\ell(a)+1} = \cdots .$$

Llamamos *longitud* de $a$, $\ell(a)$, el número de términos de $a$ diferentes de cero.

Si $a$ es una partición, definimos la *partición conjugada*, $\bar{a}$, como la partición cuya $i$–ésima componente es

$$\bar{a}_i = \mathrm{Card}\{j : a_j \geq i\}, \quad i = 1, 2, \ldots .$$

Si $a$ y $b$ son particiones y $m := \max\{\ell(a), \ell(b)\}$ decimos que $a$ está *mayorada* por $b$ y lo denotamos por $a \prec b$ si

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i, \quad k = 1, 2, \ldots, m-1,$$

y

$$\sum_{i=1}^{m} a_i = \sum_{i=1}^{m} b_i;$$

decimos que $a$ está *débilmente mayorada* por $b$ y lo denotamos por $a \prec\prec b$ si

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i, \quad k = 1, 2, \ldots.$$

Es bien conocido (cf. [16, Macdonald, pp. 5,6]), que

$$a \prec b \Leftrightarrow \bar{b} \prec \bar{a}.$$

La *suma* de $a$ y $b$ se denota por $a+b$ y es la partición cuya $i$–ésima componente es $a_i + b_i$. La *unión* de $a$ y $b$ se denota por $a \cup b$ y es la partición obtenida por reordenación de todas las componentes de $a$ y $b$ en orden no creciente. Es también conocido que

$$\overline{a \cup b} = \bar{a} + \bar{b}.$$

### 4.1.2. Haces de matrices

Un haz de matrices es una matriz polinómica de gtrado menor o igual que uno; lo denotamos por $\lambda B - A$ o abreviadamente por $\mathcal{H}$; the matrices $B$ y $A$ pertenecen a $\mathbb{C}^{m\times n}$ y, de este modo, $\lambda B - A \in \mathbb{C}[\lambda]^{m\times n}$.

El conjunto de $\mathbb{C}[\lambda]^{m\times n}$ formado por todas los haces de matrices de tamaño $m \times n$, será denotado por $\mathcal{P}_{m\times n}$.

Decimos que los haces de matrices $\lambda B_1 - A_1, \lambda B_2 - A_2 \in \mathcal{P}_{m\times n}$ son *estrictamente equivalentes* si existen matrices invertibles $P \in \mathbb{C}^{m\times m}$ and $Q \in \mathbb{C}^{n\times n}$ such that $P(\lambda B_1 - A_1)Q = \lambda B_2 - A_2$.

El *rango normal* del haz $\mathcal{H} = \lambda B - A \in \mathcal{P}_{m\times n}$ es el orden del mayor menor distinto del polinomio cero. Lo denotamos por $\operatorname{rgn}(\lambda B - A)$ o $\operatorname{rgn}(\mathcal{H})$.

El rango normal, así definido, coincide con el rango ordinario de $\lambda B - A$ como matriz cuyos elementos pertenecen a $\mathbb{C}(\lambda)$, el cuerpo de fracciones de $\mathbb{C}[\lambda]$.

Decimos que un haz $\mathcal{H}$ es *regular* si $\operatorname{rgn}(\mathcal{H}) = n = m$. Decimos que $\mathcal{H}$ es *regular a derecha* si $\operatorname{rgn}(\mathcal{H}) = n \leq m$. Decimos que $\mathcal{H}$ es *regular a izquierda* si $\operatorname{rgn}(\mathcal{H}) = m \leq n$. Consecuentemente, un haz es regular si y sólo si es regular a izquierda y a derecha. En los tres casos, $\operatorname{rgn}(\mathcal{H})$ es completo (i.e. igual a $\min\{m, n\}$) y, recíprocamente, si $\operatorname{rgn}(\mathcal{H})$ es completo el haz $\mathcal{H}$ debe pertenecer a uno (o varios) de lo tres casos citados.

Decimos que un número complejo $\alpha$ es valor propio del haz $\lambda B - A \in \mathcal{P}_{m\times n}$ si

$$\operatorname{rg}(\alpha B - A) < \operatorname{rgn}(\lambda B - A);$$

y decimos que $\infty$ es valor propio del haz $\lambda B - A \in \mathcal{P}_{m\times n}$ si 0 es un valor propio del haz $B - \mu A$, o equivalentemente si

$$\operatorname{rg} B < \operatorname{rgn}(\lambda B - A).$$

De acuerdo con los Teoremas 2.5.1 y 3.3.1, que provienen de [6, Gantmacher, Section XII.5], un sistema completo de invariantes para la equivalencia estricta de haces está formado por los siguientes tipos de invariantes, asociados a cada haz $\mathcal{H}$:

(1) *Indices minimales por columnas* denotados por

$$\varepsilon_1 \geq \cdots \geq \varepsilon_{r_1} > \varepsilon_{r_1+1} = \cdots = \varepsilon_{r_0} = 0.$$

Definimos para $i = 0, 1, 2, \ldots$

$$r_i := \mathrm{Card}\{j : \varepsilon_j \geq i\}.$$

Los números $r_0, r_1, r_2, \ldots$ se llamarán los *r–números* del haz $\mathcal{H}$. De la definición de los $r_i$ deducimos que las particiones $(\varepsilon_1, \ldots, \varepsilon_{r_1}, 0, \ldots)$ y $\mathrm{r}(\mathcal{H}) := (r_1, \ldots, r_{\varepsilon_1}, 0, \ldots)$ son conjugadas. Si $r_1 = 0, r_2 = 0, \ldots,$ ponemos $\mathrm{r}(\mathcal{H}) := 0$, donde $0 := (0, 0, \ldots)$ es la partición nula; en este caso $r_0$ puede ser cero o no. Del concepto de rango normal tenemos que $r_0 = n - \mathrm{rgn}(\mathcal{H})$. Denotaremos por $\mathrm{ci}(\mathcal{H})$ el número, $r_0$, de índices minimales por columnas de $\mathcal{H}$.

(2) *Indices minimales por filas* denotados por

$$\eta_1 \geq \cdots \geq \eta_{s_1} > \eta_{s_1+1} = \cdots = \eta_{s_0} = 0.$$

Definimos para $i = 0, 1, 2, \ldots$

$$s_i := \mathrm{Card}\{j : \eta_j \geq i\}.$$

Los números $s_0, s_1, s_2, \ldots$ se llamarán los *s–números* del haz $\mathcal{H}$. De la definición de los $s_i$ deducimos que las particiones $(\eta_1, \ldots, \eta_{s_1}, 0, \ldots)$ y $\mathrm{s}(\mathcal{H}) := (s_1, \ldots, s_{\eta_1}, 0, \ldots)$ son conjugadas. Si $s_1 = 0, s_2 = 0, \ldots,$ ponemos $\mathrm{s}(\mathcal{H}) := 0$, donde $0 := (0, 0, \ldots)$ es la partición nula; en este caso $s_0$ puede ser cero o no. Del concepto de rango normal tenemos que $s_0 = m - \mathrm{rgn}(\mathcal{H})$. Denotaremos por $\mathrm{ri}(\mathcal{H})$ el número, $s_0$, de índices minimales por filas de $\mathcal{H}$.

(3) *Divisores elementales infinitos* de la forma

$$\mu^{n_{\infty 1}}, \ldots, \mu^{n_{\infty \nu_\infty}}, \quad \text{with} \quad n_{\infty 1} \geq \cdots \geq n_{\infty \nu_\infty} \geq 1.$$

Diremos que

$$\mathrm{S}(\infty, \mathcal{H}) := (n_{\infty 1}, \ldots, n_{\infty \nu_\infty}, 0, \ldots)$$

es la partición de la característica de Segré del haz $\mathcal{H}$ para el valor propio infinito y su partición conjugada

$$\mathrm{W}(\infty, \mathcal{H}) := \overline{\mathrm{S}(\infty, \mathcal{H})} := (m_{\infty 1}, m_{\infty 2}, \ldots)$$

es la partición de la característica de Weyr del haz $\mathcal{H}$ para el valor propio infinito.

Por tanto $m_{\infty 1} = \nu_\infty$. Si $\infty$ no es un valor propio de $\mathcal{H}$, escribimos $\mathrm{S}(\infty, \mathcal{H}) := 0$ y $\mathrm{W}(\infty, \mathcal{H}) := 0$.

(4) *Divisores elementales finitos* de la forma

$$(\lambda - \lambda_1)^{n_{\lambda_1 1}}, \ldots, (\lambda - \lambda_1)^{n_{\lambda_1 \nu_1}}, \ldots, (\lambda - \lambda_u)^{n_{\lambda_u 1}}, \ldots, (\lambda - \lambda_u)^{n_{\lambda_u \nu_u}}$$

con $n_{\lambda_i 1} \geq \cdots \geq n_{\lambda_i \nu_i} \geq 1 \quad (i = 1, \ldots, u)$.

Diremos que

$$\mathrm{S}(\lambda_i, \mathcal{H}) := (n_{\lambda_i 1}, \ldots, n_{\lambda_i \nu_i}, 0, \ldots)$$

es la partición de la característica de Segré del haz $\mathcal{H}$ correspondiente al valor propio $\lambda_i$ $(i = 1, \ldots, u)$. Su conjugate partición conjugada

$$\mathrm{W}(\lambda_i, \mathcal{H}) := \overline{\mathrm{S}(\lambda_i, \mathcal{H})} := (m_{\lambda_i 1}, m_{\lambda_i 2}, \ldots)$$

será llamada la partición de la característica de Weyr del haz $\mathcal{H}$ correspondiente al valor propio $\lambda_i$ $(i = 1, \ldots, u)$. Consecuentemente, $m_{\lambda_i 1} = \nu_{\lambda_i}$ $(i = 1, \ldots, u)$.

Sea $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$. El subconjunto de $\overline{\mathbb{C}}$ formado por todos los valores propios de $\mathcal{H}$ será llamado el *espectro* del haz $\mathcal{H}$, y será denotado por $\sigma(\mathcal{H})$.

La *característica de Segré* de $\mathcal{H}$ es el sistema de particiones

$$\mathrm{S}(\mathcal{H}) := (\mathrm{S}(\alpha, \mathcal{H}))_{\alpha \in \sigma(\mathcal{H})}.$$

La *característica de Weyr* de $\mathcal{H}$ es el sistema de particiones

$$\mathrm{W}(\mathcal{H}) := (\mathrm{W}(\alpha, \mathcal{H}))_{\alpha \in \sigma(\mathcal{H})}.$$

Generalizamos las notaciones de (3) y (4): Si $z \in \overline{\mathbb{C}}$ definimos

$$\mathrm{S}(z, \mathcal{H}) := \begin{cases} \mathrm{S}(z, \mathcal{H}) & \text{si} \quad z \in \sigma(\mathcal{H}) \\ 0 \text{ (partición nula)} & \text{if} \quad z \notin \sigma(\mathcal{H}) \end{cases}$$

y análogamente

$$\mathrm{W}(z, \mathcal{H}) := \begin{cases} \mathrm{W}(z, \mathcal{H}) & \text{si} \quad z \in \sigma(\mathcal{H}) \\ 0 \text{ (partición nula)} & \text{if} \quad z \notin \sigma(\mathcal{H}) \end{cases}.$$

*Observación* 4.1.1. Un haz regular a derecha $\mathcal{H}$ no tiene índices minimales por columnas, porque $r_0 = n - \mathrm{rgn}(\mathcal{H})$ y $n = \mathrm{rgn}(\mathcal{H})$. Análogamente, un haz regular a izquierda no tiene índices minimales por filas. Así pues, los únicos invariantes de un haz regular son los divisores elementales (finitos e infinitos).

Dado un haz $\mathcal{H}$ con los invariantes descritos más arriba, podemos asociarles un haz en la *forma canónica de Kronecker* según se ha expuesto en el Capítulo 3:

(1) Si $\varepsilon_j$ es un índice minimal por columnas $> 0$, ponemos

$$L_{\varepsilon_j} := \begin{bmatrix} \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \lambda & -1 \end{bmatrix} \in \mathcal{P}_{\varepsilon_j \times (\varepsilon_j + 1)}.$$

(2) Si $\eta_j$ es un índice minimal por filas $> 0$, ponemos

$$\tilde{L}_{\eta_j} := L_{\eta_j}^T \in \mathcal{P}_{(\eta_j + 1) \times \eta_j},$$

donde $^T$ denota traspuesta.

(3) Si $\mu^{n_{\infty j}}$ es un divisor elemental infinito,

$$J_{n_{\infty j}}(\infty) := \begin{bmatrix} -1 & \lambda & & \\ & \ddots & \ddots & \\ & & -1 & \lambda \\ & & & -1 \end{bmatrix} \in \mathcal{P}_{n_{\infty j} \times n_{\infty j}};$$

esto es lo mismo que decir,

$$J_{n_{\infty j}}(\infty) = \lambda J_{n_{\infty j}}(0) - I_{n_{\infty j}},$$

siendo $J_{n_{\infty j}}(0) \in \mathbb{C}^{n_{\infty j} \times n_{\infty j}}$ un bloque de Jordan asociado al valor propio 0.

(4) Si $(\lambda - \alpha)^{n_{\alpha j}}$ es un divisor elemental finito,

$$JF_{n_{\alpha j}}(\alpha) := \begin{bmatrix} \lambda - \alpha & -1 & & \\ & \ddots & \ddots & \\ & & \lambda - \alpha & -1 \\ & & & \lambda - \alpha \end{bmatrix} \in \mathcal{P}_{n_{\alpha j} \times n_{\alpha j}};$$

dicho de otro modo,

$$JF_{n_{\alpha j}}(\alpha) = \lambda I_{n_{\alpha j}} - J_{n_{\alpha j}}(\alpha)$$

siendo $J_{n_{\alpha j}}(\alpha) \in \mathbb{C}^{n_{\alpha j} \times n_{\alpha j}}$ un bloque de Jordan asociado al valor propio $\alpha$.

La *forma canónica de Kronecker* del haz $\mathcal{H}$ es el haz

$$C_K(\mathcal{H}) := \operatorname{diag}(L, \tilde{L}, J(\infty), JF),$$

donde

$$L := \left[\operatorname{diag}(L_{\varepsilon_1}, \ldots, L_{\varepsilon_{r_1}}), 0_{(\varepsilon_1 + \cdots + \varepsilon_{r_1}) \times (r_0 - r_1)}\right],$$

$$\tilde{L} := \begin{bmatrix} \operatorname{diag}(L^{\mathrm{T}}_{\eta_1}, \ldots, L^{\mathrm{T}}_{\eta_{s_1}}) \\ 0_{(s_0 - s_1) \times (\eta_1 + \cdots + \eta_{s_1})} \end{bmatrix},$$

$$J(\infty) := \operatorname{diag}\left(J_{n_{\infty 1}}(\infty), \ldots, J_{n_{\infty \nu_\infty}}(\infty)\right)$$

y

$$JF := \operatorname{diag}\left(JF_{n_{\lambda_1 1}}(\lambda_1), \ldots, JF_{n_{\lambda_1 \nu_1}}(\lambda_1), \ldots, JF_{n_{\lambda_u 1}}(\lambda_u), \ldots, JF_{n_{\lambda_u \nu_u}}(\lambda_u)\right).$$

Denotamos a la matriz nula $p \times q$ por $0_{p \times q}$.

*Observación* 4.1.2. La relación entre el número de columnas (respectivamente, el número de filas) de un haz $\mathcal{H}$ $m \times n$ y sus invariantes enteros es la siguiente:

$$\sum_{\alpha \in \sigma(\mathcal{H})} \sum_{k=1}^{l(\mathrm{W}(\alpha,\mathcal{H}))} m_{\alpha k} + \sum_{i=1}^{l(\mathrm{r}(\mathcal{H}))} r_i + \sum_{i=1}^{l(\mathrm{s}(\mathcal{H}))} s_i + \mathrm{ci}(\mathcal{H}) = n,$$

respectivamente,

$$\sum_{\alpha \in \sigma(\mathcal{H})} \sum_{k=1}^{l(\mathrm{W}(\alpha,\mathcal{H}))} m_{\alpha k} + \sum_{i=1}^{l(\mathrm{r}(\mathcal{H}))} r_i + \sum_{i=1}^{l(\mathrm{s}(\mathcal{H}))} s_i + \mathrm{ri}(\mathcal{H}) = m.$$

**Ejemplo 4.1.3.** Según esta nueva notación, los invariantes del Ejemplo 1.1.4, quedarían de este modo:

$\varepsilon_1 = 2 > \varepsilon_2 = 1 > \varepsilon_3 = 0, \quad (\Longrightarrow r_1 = 2, r_0 = 3 = \mathrm{ci}(\mathcal{H}))$
$\eta_1 = 2 > \eta_2 = 1 > \eta_3 = 0 > \eta_4 = 0, \quad (\Longrightarrow s_1 = 2, s_0 = 4 = \mathrm{ri}(\mathcal{H}))$
$\mu^3,$
$\lambda^2, \ (\lambda + 2)^2.$

Por tanto, tenemos que

$$\mathrm{r}(\mathcal{H}) = (2, 1, 0, \ldots), \quad \mathrm{s}(\mathcal{H}) = (2, 1, 0, \ldots),$$

$$\mathrm{W}(\infty, \mathcal{H}) = (1, 1, 1, 0, \ldots), \mathrm{W}(0, \mathcal{H}) = (1, 1, 0, \ldots) = \mathrm{W}(-2, \mathcal{H}).$$

Así pues,

$$\sum_{\alpha \in \sigma(\mathcal{H})} \sum_{k=1}^{l(\mathrm{W}(\alpha,\mathcal{H}))} m_{\alpha k} + \sum_{i=1}^{l(\mathrm{r}(\mathcal{H}))} r_i + \sum_{i=1}^{l(\mathrm{s}(\mathcal{H}))} s_i + \mathrm{ci}(\mathcal{H}) = 7 + 3 + 3 + 3 = 16,$$

$$\sum_{\alpha \in \sigma(\mathcal{H})} \sum_{k=1}^{l(\mathrm{W}(\alpha,\mathcal{H}))} m_{\alpha k} + \sum_{i=1}^{l(\mathrm{r}(\mathcal{H}))} r_i + \sum_{i=1}^{l(\mathrm{s}(\mathcal{H}))} s_i + \mathrm{ri}(\mathcal{H}) = 7 + 3 + 3 + 4 = 17.$$

## 4.2. Caracterización de los invariantes enteros por rangos

Una caracterización de $\mathrm{W}(\alpha,\mathcal{H})$ para $\alpha\in\overline{\mathbb{C}}$ y $\mathcal{H}=\lambda B-A$ un haz regular (o haz regular a derecha o a izquierda) con $B$ y $A$ matrices reales, viene dado en [13, Karcanias, Kalogeropoulos]. Estos resultados pueden ser generalizados a cualquier haz de matrices complejas [18, Pokrzywa].

**Definición 4.2.1.** Si $A=(a_{ij})\in\mathbb{C}^{n\times m}$ y $B\in\mathbb{C}^{p\times q}$ llamamos *producto de Kronecker* de $A$ y $B$, y lo denotamos por $A\otimes B$, a la matriz por bloques siguiente

$$A\otimes B:=\begin{bmatrix} a_{11}B & \cdots & a_{1n}B\\ \vdots & \ddots & \vdots\\ a_{m1}B & \cdots & a_{mn}B\end{bmatrix}\in\mathbb{C}^{mp\times nq}.$$

*Observación* 4.2.2. Destacamos dos propiedades que verifica el producto de Kronecker de matrices:

(1) Siempre que se pueden realizar los propuctos $AC$ y $BD$, se verifica

$$(A\otimes B)(C\otimes D)=(AC\otimes BD). \tag{4.1}$$

(2) Si $A$ y $B$ son no singulares, entonces $A\otimes B$ también lo es y

$$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}.$$

Sea $\lambda B-A\in\mathcal{P}_{m\times n}$ y $\alpha$ cualquier número complejo. Definimos para $k=1,2,\dots$

$$P_\alpha^k(B,A):=\begin{bmatrix} \alpha B-A & 0 & \dots & \dots & 0\\ B & \alpha B-A & \ddots & & \vdots\\ 0 & B & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \alpha B-A & 0\\ 0 & \dots & 0 & B & \alpha B-A\end{bmatrix}\in\mathbb{C}^{km\times kn}.$$

Observemos que

$$P_\alpha^k(B,A)=I_k\otimes(\alpha B-A)+\begin{pmatrix} 0 & 0\\ I_{k-1} & 0\end{pmatrix}\otimes B, \tag{4.2}$$

donde $\otimes$ es el producto de Kronecker.

Si consideramos $\infty$ en lugar de $\alpha$ definimos para $k=1,2,\dots$

$$P_\infty^k(B,A):=\begin{bmatrix} B & 0 & \dots & \dots & 0\\ -A & B & \ddots & & \vdots\\ 0 & -A & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & B & 0\\ 0 & \dots & 0 & -A & B\end{bmatrix}=I_k\otimes B-\begin{pmatrix} 0 & 0\\ I_{k-1} & 0\end{pmatrix}\otimes A\in\mathbb{C}^{km\times kn}. \tag{4.3}$$

Para cualquier $\mu\in\overline{\mathbb{C}}$, escribimos también $P_\mu^k(\mathcal{H}):=P_\mu^k(B,A)$.

**Lema 4.2.3.** *Sean* $\lambda B - A, \lambda\bar{B} - \bar{A} \in \mathcal{P}_{m\times n}$ *tales que* $\lambda B - A \overset{s}{\sim} \lambda\bar{B} - \bar{A}$. *Entonces, para todo* $\alpha \in \overline{\mathbb{C}}$ *y para todo* $k = 1, 2, \dots$ *se verifica*

$$\mathrm{rg}(P_\alpha^k(B, A)) = \mathrm{rg}(P_\alpha^k(\bar{B}, \bar{A})).$$

Demostración. Sean $P$ y $Q$ invertibles tales que

$$\lambda\bar{B} - \bar{A} = P(\lambda B - A)Q,$$

es decir se cumple que

$$\bar{B} = PBQ, \quad \bar{A} = PAQ.$$

Entonces, considerando las matrices invertibles $I_k \otimes P$, $I_k \otimes Q$ obtenemos en cada caso lo siguiente:

(1) si $\alpha \in \mathbb{C}$, aplicando la propiedad (4.1) y la identidad (4.2), se sigue que

$$(I_k \otimes P)P_\alpha^k(B, A)(I_k \otimes Q) = (I_k \otimes P)(I_k \otimes (\alpha B - A) + \begin{pmatrix} 0 & 0 \\ I_{k-1} & 0 \end{pmatrix} \otimes B)(I_k \otimes Q)$$

$$= I_k \otimes P(\alpha B - A)Q + \begin{pmatrix} 0 & 0 \\ I_{k-1} & 0 \end{pmatrix} \otimes PBQ = P_\alpha^k(\bar{B}, \bar{A}).$$

(2) si $\alpha = \infty$, aplicando (4.1) y la identidad (4.3) sigue análogamente

$$(I_k \otimes P)P_\infty^k(B, A)(I_k \otimes Q) = P_\infty^k(\bar{B}, \bar{A}).$$

Luego, para todo $\alpha \in \overline{\mathbb{C}}$ y para todo $k = 1, 2, \dots$, las matrices $P_\alpha^k(B, A)$ y $P_\alpha^k(\bar{B}, \bar{A})$ son equivalentes y por tanto tienen el mismo rango.

□

**Teorema 4.2.4.** *Sea* $\lambda B - A \in \mathcal{P}_{m\times n}$. *Entonces, para todo* $\alpha \in \overline{\mathbb{C}}$ *y para todo* $k = 1, 2, \dots$ *se verifica*

$$\nu(P_\alpha^k(B, A)) = \sum_{i=1}^{k} m_{\alpha i} + kr_0,$$

*donde* $(m_{\alpha 1}, m_{\alpha 2}, \dots) := \mathrm{W}(\alpha, \lambda B - A)$, $r_0 := \mathrm{ci}(\lambda B - A)$ *y* $\nu(M) := \dim \mathrm{Ker}\, M = \dim\{X \in \mathbb{C}^{n\times 1} \mid MX = 0\}$ *denota la nulidad de cualquier matriz* $M \in \mathbb{C}^{m\times n}$.

Demostración. Si el haz es regular por la derecha (o equivalentemente, $r_0 = 0$) la demostración aparece en la sección 5 de [13, Karcanias,Kalogeropoulos], para el caso real, y en [18, Corolario 1, Pokrzywa], para el caso complejo.

Si $\lambda B - A$ es un haz cualquiera de $\mathcal{P}_{m\times n}$, por el Lema 4.2.3 podemos sustituirlo por un haz estrictamente equivalente a él, por ejemplo su forma canónica de Kronecker. Este haz en forma canónica es $\mathrm{diag}(L, \bar{H})$ siendo $\bar{H}$ un haz regular por la derecha y $L$ el haz que recoge los índices minimales por columnas de $\lambda B - A$.

Haciendo permutaciones de filas y columnas, para todo $\alpha \in \overline{\mathbb{C}}$ y para todo $k = 1, 2, \dots$ obtenemos

$$\nu(P_\alpha^k(\mathrm{diag}(L, \bar{H}))) = \nu(\mathrm{diag}(P_\alpha^k(L), P_\alpha^k(\bar{H}))) = \nu(P_\alpha^k(L)) + \nu(P_\alpha^k(\bar{H})).$$

Por ser $\bar{H}$ la parte regular por la derecha de $\lambda B - A$ se tiene que

$$\nu(P_\alpha^k(\bar{H})) = \sum_{i=1}^{k} m_{\alpha i}.$$

Por contener $L$ sólo índices minimales por columnas se comprueba fácilmente que

$$\nu(P_\alpha^k(L)) = kr_0.$$

Así pues, el resultado queda demostrado.

$\square$

**Notación.** Vamos a extender la notación $\alpha B - A$ para $\alpha = \infty$, definiendo

$$\infty B - A := B.$$

Teniendo en cuenta este convenio, veamos unas consecuencias importantes del teorema anterior.

**Corolario 4.2.5.** *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$ y $\alpha \in \overline{\mathbb{C}}$. Se verifica:*

*(i)* $m_{\alpha 1} = \operatorname{rgn}(\lambda B - A) - \operatorname{rg}(\alpha B - A)$,

*(ii) $\alpha$ es valor propio de $\lambda B - A$ si y sólo si*

$$\operatorname{rg}(\alpha B - A) < \operatorname{rgn}(\lambda B - A).$$

Demostración. Por el Teorema 4.2.4, para todo $\alpha \in \overline{\mathbb{C}}$ se tiene

$$m_{\alpha 1} = \nu(P_\alpha^1(B, A)) - r_0 = n - \operatorname{rg}(\alpha B - A) - r_0 = \operatorname{rgn}(\lambda B - A) - \operatorname{rg}(\alpha B - A).$$

Como $\alpha$ es valor propio de $\lambda B - A$ si y sólo si $m_{\alpha 1} > 0$, de la igualdad anterior se obtiene la equivalencia de $(ii)$.

$\square$

**Corolario 4.2.6.** *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$ . Para $k = 1, 2, \ldots$ se tiene que*

$$k \operatorname{rgn}(\lambda B - A) = \max_{\alpha \in \overline{\mathbb{C}}} \operatorname{rg}(P_\alpha^k(B, A)).$$

Demostración. Por el Teorema 4.2.4, para todo $\alpha \in \overline{\mathbb{C}}$ y para todo $k = 1, 2, \ldots$ se tiene

$$\operatorname{rg}(P_\alpha^k(B, A)) = kn - \nu(P_\alpha^k(B, A)) = kn - \sum_{i=1}^{k} m_{\alpha i} - kr_0 = k \operatorname{rgn}(\lambda B - A) - \sum_{i=1}^{k} m_{\alpha i}.$$

Por tanto, para todo $\alpha \in \overline{\mathbb{C}}$ y para todo $k = 1, 2, \ldots$

$$k \operatorname{rgn}(\lambda B - A) \geq \operatorname{rg}(P_\alpha^k(B, A)),$$

y se da la igualdad cuando $\sum_{i=1}^{k} m_{\alpha i} = 0$, es decir, en el caso en que $\alpha$ no sea valor propio de $\lambda B - A$.

□

*Observación* 4.2.7. En la demostración anterior se observa que si $\beta \notin \sigma(\lambda B - A)$ entonces

$$\mathrm{rg}(P_\beta^k(B,A)) = k\,\mathrm{rgn}(\lambda B - A) = \max_{\alpha \in \overline{\mathbb{C}}} \mathrm{rg}(P_\alpha^k(B,A)).$$

Si $k = 1$ la igualdad anterior se reduce a

$$\mathrm{rgn}(\lambda B - A) = \max_{\alpha \in \overline{\mathbb{C}}} \mathrm{rg}(\alpha B - A).$$

A continuación, damos sendas caracterizaciones de los $r$–números y de los $s$–números de un haz a partir de determinadas secuencias de nulidades que, como se demuestra en [14, Karcanias,Kalogeropoulos] para el caso real, tienen relación con los valores de los índices minimales por columnas del haz y el número de veces que aparecen cada uno de ellos.

Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Definimos para $k = 1, 2, \ldots$

$$T_k(B,A) := \begin{bmatrix} B & 0 & \ldots & \ldots & 0 \\ -A & B & \ddots & & \vdots \\ 0 & -A & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & B & 0 \\ \vdots & & \ddots & -A & B \\ 0 & \ldots & \ldots & 0 & -A \end{bmatrix} = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \otimes B - \begin{pmatrix} 0 \\ I_k \end{pmatrix} \otimes A \in \mathbb{C}^{(k+1)m\times kn}.$$

Si $H = \lambda B - A$ también denotaremos $T_k(B,A)$ por $T_k(H)$.

Obsérvese que, para $k = 1, 2, \ldots$, reordenando los $k+1$ bloques por filas de la matriz $T_k(B,A)$, en orden inverso al dado, y después, reordenando los $k$ bloques por columnas de la matriz resultante, también en orden inverso, se obtiene finalmente la matriz $M_{k-1} := M_{k-1}[\lambda B - A]$ dada en (3.4). Por tanto, para $k = 1, 2, \ldots$, se tiene

$$\mathrm{rg}(T_k(B,A)) = \mathrm{rg}(M_{k-1}[\lambda B - A]).$$

Mediante una demostración análoga a la del Lema 4.2.3 tenemos el resultado siguiente.

**Lema 4.2.8.** *Sean $\lambda B - A$, $\lambda\bar{B} - \bar{A} \in \mathcal{P}_{m\times n}$ tales que $\lambda B - A \overset{s}{\sim} \lambda\bar{B} - \bar{A}$. Entonces, para todo $k = 1, 2, \ldots$ se verifica*

$$\mathrm{rg}(T_k(B,A)) = \mathrm{rg}(T_k(\bar{B},\bar{A})).$$

**Teorema 4.2.9.** *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Entonces, para todo $k = 1, 2, \ldots$ se tiene*

$$\nu(T_k(B,A)) = kr_0 - \sum_{i=1}^{k} r_i,$$

*donde $r_0 := \mathrm{ci}(\lambda B - A)$ y $(r_1, r_2, \ldots) := \mathrm{r}(\lambda B - A)$.*

Demostración. Si denotamos por $c_j$ el número de índices minimales por columnas iguales a $j$ ($j = 0, 1, \ldots$) es fácil ver que según el Corolario 3.2 de [14, Karcanias, Kalogeropoulos], para $k = 1, 2, \ldots$, se verifica

$$\nu(T_k(H)) - \nu(T_{k-1}(H) = \sum_{j=1}^{k-1} c_j,$$

conviniendo que $\nu(T_0(H)) = 0$. Teniendo en cuenta este convenio se deduce fácilmente que para $k = 1, 2, \ldots$

$$\nu(T_k(H)) = \sum_{j=1}^{k-1}(k-j)c_j. \tag{4.4}$$

Como por la definición de $r$–números tenemos

$$c_j = r_j - r_{j+1} \quad (j = 1, 2, \ldots),$$

de (4.4), para $k = 1, 2, \ldots$, sigue que

$$\nu(T_k(H)) = \sum_{j=1}^{k-1}(k-j)(r_j - r_{j+1}) =$$

$$k(r_0 - r_1) + (k-1)(r_1 - r_2) + \cdots + 2(r_{k-2} - r_{k-1}) + (r_{k-1} - r_k) = kr_0 - \sum_{i=1}^{k} r_i.$$

□

**Teorema 4.2.10.** *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Entonces, para todo $k = 1, 2, \ldots$ se tiene*

$$\nu(T_k(B^T, A^T)) = ks_0 - \sum_{i=1}^{k} s_i,$$

*donde $s_0 := \mathrm{ri}(\lambda B - A)$ y $(s_1, s_2, \ldots) := \mathrm{s}(\lambda B - A)$.*

Demostración. Basta tener en cuenta que si $\lambda B^T - A^T$ es el haz traspuesto de $\lambda B - A$ los índices minimales por columnas (respectivamente, los $r$–números) de $\lambda B^T - A^T$ coinciden con los índices minimales por filas (respectivamente, los $s$–números) de $\lambda B - A$. Aplicando ahora el Teorema 4.2.9 se obtiene la igualdad.

□

Para acabar, definimos unas matrices parecidas a las $P_\alpha^k(B, A)$ y que contienen a éstas en alguno de sus bloques. En función de ellas, daremos unos resultados que arrojan alguna luz sobre el Teorema 4.2.4.

Sean $\lambda B - A \in \mathcal{P}_{m\times n}$ y $\alpha_1, \ldots, \alpha_q$ números complejos arbitrarios, con $\alpha_i \neq \alpha_j$ para $i \neq j$, $i, j \in \{1, \ldots, q\}$. Definimos para $k_1, \ldots, k_q \in \{1, 2, \ldots\}$

$$P_{\alpha_1,\ldots,\alpha_q}^{k_1,\ldots,k_q}(B, A) := \begin{bmatrix} P_{\alpha_1}^{k_1}(B, A) & 0 & \ldots & \ldots & 0 \\ Q_{k_2,k_1}(B) & P_{\alpha_2}^{k_2}(B, A) & \ddots & & \vdots \\ 0 & Q_{k_3,k_2}(B) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & P_{\alpha_{q-1}}^{k_{q-1}}(B, A) & 0 \\ 0 & \ldots & 0 & Q_{k_q,k_{q-1}}(B) & P_{\alpha_q}^{k_q}(B, A) \end{bmatrix}$$

donde

$$Q_{k_{j+1},k_j}(B) := \begin{bmatrix} 0 & \dots & 0 & B \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{C}^{(k_{j+1}m)\times(k_j n)}$$

para $j = 1, \dots, q-1$, y

$$P^{k_1,\dots,k_q}_{\alpha_1,\dots,\alpha_q}(\mathcal{H}) := P^{k_1,\dots,k_q}_{\alpha_1,\dots,\alpha_q}(B,A) \in \mathbb{C}^{[(k_1+\dots+k_q)m]\times[(k_1+\dots+k_q)n]},$$

si $\mathcal{H} := \lambda B - A$.

**Lema 4.2.11.** *[10, I. de Hoyos] Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Entonces, para cualesquiera $\alpha_1, \dots, \alpha_q \in \mathbb{C}$ distintos y para cualesquiera $k_1, \dots, k_q \in \{1, 2, \dots\}$ sigue que*

$$\nu(P^{k_1,\dots,k_q}_{\alpha_1,\dots,\alpha_q}(B,A)) = \nu(\mathrm{diag}(P^{k_1}_{\alpha_1}(B,A), \dots, P^{k_q}_{\alpha_q}(B,A))).$$

DEMOSTRACIÓN. El resultado sigue, si probamos que las dos matrices tienen el mismo rango. Se puede hacer tomando $\lambda B - A$ en forma canónica de Kronecker y pasando de una matriz a otra mediante transformaciones elementales, ya que éstas conservan el rango.

También se pueden hacer por medio de operaciones elementales por bloques en la matriz $P^{k_1,\dots,k_q}_{\alpha_1,\dots,\alpha_q}(B,A)$ hasta obtener $\mathrm{diag}(P^{k_1}_{\alpha_1}(B,A), \dots, P^{k_q}_{\alpha_q}(B,A))$.

A modo de ejemplo, en el caso particular $q = 2$, $k_1 = k_2 = 2$ y $\alpha_1 = \alpha \neq \beta = \alpha_2$, tomando las matrices invertibles siguientes

$$P := \begin{bmatrix} (\alpha-\beta)^{-3}I_m & 0 & 0 & 0 \\ 0 & (\alpha-\beta)^{-3}I_m & 0 & 0 \\ (\alpha-\beta)^{-2}I_m & -(\alpha-\beta)^{-1}I_m & I_m & 0 \\ 2(\alpha-\beta)^{-3}I_m & -(\alpha-\beta)^{-2}I_m & 0 & I_m \end{bmatrix},$$

$$Q := \begin{bmatrix} (\alpha-\beta)^{3}I_n & 0 & 0 & 0 \\ 0 & (\alpha-\beta)^{3}I_n & 0 & 0 \\ -(\alpha-\beta)I_n & (\alpha-\beta)^{2}I_n & I_n & 0 \\ -2I_n & (\alpha-\beta)I_n & 0 & I_n \end{bmatrix},$$

se puede comprobar que

$$P\left[\begin{array}{cc|cc} \alpha B - A & 0 & 0 & 0 \\ B & \alpha B - A & 0 & 0 \\ \hline 0 & B & \beta B - A & 0 \\ 0 & 0 & B & \beta B - A \end{array}\right] Q = \left[\begin{array}{cc|cc} \alpha B - A & 0 & 0 & 0 \\ B & \alpha B - A & 0 & 0 \\ \hline 0 & 0 & \beta B - A & 0 \\ 0 & 0 & B & \beta B - A \end{array}\right],$$

de aquí que las matrices de partida, $P^{2,2}_{\alpha,\beta}(B,A)$ y $\mathrm{diag}(P^2_\alpha(B,A), P^2_\beta(B,A))$, tienen igual rango y nulidad.

□

**Lema 4.2.12.** [10, I. de Hoyos]. *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Sean cualesquiera $\alpha_1, \dots, \alpha_q \in \mathbb{C}$ distintos y $k_1, \dots, k_q \in \{1, 2, \dots\}$. Sea $k := k_1 + \dots + k_q$.*

*Si $\alpha_i \notin \sigma(\lambda B - A)$ para cada $i = 1, \dots, q$, entonces*

$$\mathrm{rg}(P^{k_1,\dots,k_q}_{\alpha_1,\dots,\alpha_q}(B,A)) = k \cdot \mathrm{rgn}(\lambda B - A).$$

Demostración. Por el Lema 4.2.11

$$\operatorname{rg}(P^{k_1,\ldots,k_q}_{\alpha_1,\ldots,\alpha_q}(B,A)) = \sum_{i=1}^{q} \operatorname{rg}(P^{k_i}_{\alpha_i}(B,A)).$$

Por la Observación 4.2.7 y por ser $\alpha_i \notin \sigma(\lambda B - A)$ se tiene

$$\operatorname{rg}(P^{k_i}_{\alpha_i}(B,A)) = k_i \operatorname{rgn}(\lambda B - A) \qquad (i = 1,\ldots,q).$$

Como $k = k_1 + \cdots + k_q$ la suma de rangos es $k \operatorname{rgn}(\lambda B - A)$.

□

Si en lugar de considerar $q$ elementos de $\mathbb{C}$ distintos, junto con el número de veces que aparece cada uno, consideramos $k$ números complejos cualesquiera $\beta_1,\ldots,\beta_k$ (no necesariamente distintos) y definimos

$$P_{\beta_1,\ldots,\beta_k}(B,A) := \begin{bmatrix} \beta_1 B - A & 0 & \cdots & \cdots & 0 \\ B & \beta_2 B - A & \ddots & & \vdots \\ 0 & B & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \beta_{k-1}B - A & 0 \\ 0 & \cdots & 0 & B & \beta_k B - A \end{bmatrix}$$

tenemos análogos resultados a los de los Lemas 4.2.11 y 4.2.12.

**Lema 4.2.13.** [10]. *Sea $\lambda B - A \in \mathcal{P}_{m\times n}$. Sea $\beta_1,\ldots,\beta_k \in \mathbb{C}$ (no necesariamente distintos) tal que $k_i$ de ellos son iguales a $\alpha_i$, $(i = 1,\ldots,q)$ y $\alpha_i \neq \alpha_j$ para $i \neq j$. Entonces se verifica*

*(i)* $\nu(P_{\beta_1,\ldots,\beta_k}(B,A)) = \nu(\operatorname{diag}(P^{k_1}_{\alpha_1}(B,A),\ldots,P^{k_q}_{\alpha_q}(B,A)))$,

*(ii) si $\beta_j \notin \sigma(\lambda B - A))$ para todo $j = 1,\ldots,k$ se tiene*

$$\operatorname{rg}(P_{\beta_1,\ldots,\beta_k}(B,A)) = k \operatorname{rgn}(\lambda B - A).$$

Demostración. Basta demostrar mediante transformaciones elementales que

$$\operatorname{rg}(P_{\beta_1,\ldots,\beta_k}(B,A)) = P^{k_1,\ldots,k_q}_{\alpha_1,\ldots,\alpha_q}(B,A)$$

y aplicar, entonces, los Lemas 4.2.11 y 4.2.12.

□

## 4.3. Ejercicios

**4.1.** Sea $\lambda B - A \in \mathbb{C}[\lambda]^{m\times n}$.

(i) Si $\lambda\tilde{B} - \tilde{A} \in \mathbb{C}[\lambda]^{(m+1)\times n}$ es un haz resultante de añadir una fila al haz $\lambda B - A$, siendo respectivamente $r_0 := \operatorname{ci}(\lambda B - A)$, $(r_1, r_2, \ldots) := \operatorname{r}(\lambda B - A)$ y $\tilde{r}_0 := \operatorname{ci}(\lambda\tilde{B} - \tilde{A})$, $(\tilde{r}_1, \tilde{r}_2, \ldots) := \operatorname{r}(\lambda\tilde{B} - \tilde{A})$, entonces se verifica

$$\tilde{r}_0 \leq r_0; \quad (\tilde{r}_0 - \tilde{r}_1, \tilde{r}_0 - \tilde{r}_2, \ldots) \prec\!\!\prec (r_0 - r_1, r_0 - r_2, \ldots).$$

(ii) Si $\lambda\tilde{B}-\tilde{A} \in \mathbb{C}[\lambda]^{m\times(n+1)}$ es un haz resultante de añadir una columna al haz $\lambda B-A$, siendo respectivamente $s_0 := \mathrm{ri}(\lambda B-A)$, $(s_1, s_2, \ldots) := \mathrm{s}(\lambda B-A)$ y $\tilde{s}_0 := \mathrm{ri}(\lambda\tilde{B}-\tilde{A})$, $(\tilde{s}_1, \tilde{s}_2, \ldots) := \mathrm{s}(\lambda\tilde{B}-\tilde{A})$, entonces se verifica

$$\tilde{s}_0 \leq s_0; \quad (\tilde{s}_0-\tilde{s}_1, \tilde{s}_0-\tilde{s}_2, \ldots) \prec\!\!\prec (s_0-s_1, s_0-s_2, \ldots).$$

## Notas

Los métodos de cálculo de los invariantes enteros de un haz mediante rangos fueron publicados por Karcanias y Kalogeropoulos, y por Pokrzywa en 1986.

# Bibliografía

[1] I. Baragaña: *Prescripción parcial de haces con una aplicación a la teoría de control*, Tesis doctoral, Universidad del País Vasco, Bilbao, 1990.

[2] D. L. Boley: The algebraic structure of pencils and block Toeplitz matrices, *Linear Algebra Appl.*, **279**, (1998) 255–279.

[3] A. Edelman, E. Elmroth and B. Kågström. A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm, *SIAM J. Matrix Analysis*, **20** , no. 3 (1999)667–699.

[4] G. D. Forney: Minimal bases of rational vector spaces, with applications to multivariable linear systems, *SIAM J. Control,* **13**, No. 3 (1975) 493–520.

[5] S. Friedland: *Spectral theory of matrices I. General matrices*, MRC, TSR # 2082, University of Wisconsin, Madison, 1980.

[6] F. R. Gantmacher: *Théorie des matrices*, Tome 2, Dunod, Paris, 1966.

[7] J.-M.Gracia, L. Ortiz de Elguea y M.-J. Sodupe: Matrix pencil generated by a tensor product from two matrix pencils, *Linear Algebra Appl.*, **249**, (1996) 289–310.

[8] I. Gohberg, P. Lancaster y L. Rodman: *Invariant subspaces of matrices with applications*, Wiley, New York, 1986.

[9] I. N. Herstein: *Algebra moderna*, Trillas, México, 1970.

[10] I. de Hoyos: *Perturbación de matrices rectangulares y haces de matrices*, Tesis doctoral, Universidad del País Vasco, Bilbao, 1990.

[11] N. Jacobson: *Basic algebra I*, W. H. Freeman & Company, San Francisco, 1974.

[12] T. Kailath: *Linear systems*, Prentice-Hall, 1980.

[13] N. Karcanias, G. Kalogeropoulos: On the Segré, Weyr characteristics of right (left) regular matrix pencils, *Int. J. Control,* **44**, no. 4 (1986) 991–1015.

[14] N. Karcanias, G. Kalogeropoulos: Right, left characteristic sequences and column, row minimal indices of a singular pencil, *Int. J. Control,* **47**, no. 4 (1988) 937–946.

[15] G. Kalogeropoulos, N. Karcanias: Relationships between invariant subspaces and deflating subspaces of a regular pair $(F, G)$, *Mathematica Balkanica*, New Series, Vol, **2**, Fasc. 1, (1988) 54–63.

[16] I.G. Macdonald: *Symmetric functions and Hall polynomials*, Clarendon, Oxford, 1979.

[17] M. Marcus: *Introduction to modern algebra*, Marcel Dekker, Inc, New York, 1978.

[18] A. Pokrzywa: On the perturbations and the equivalence orbit of a matrix pencil, *Linear Algebra Appl.*, **82**, (1986) 99–121.

[19] X. Puerta: *Contribucions a l'estudi geomètric de sistemes lineals multivariables*, Tesis Doctoral, Universitat Politècnica de Catalunya, Barcelona, 1998.

[20] A. Roca: *Asignación de invariantes en sistemas de control*, Tesis Doctoral, Universidad Politécnica de Valencia, Valencia, 2003.

[21] W.G. Stewart: On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* **9** (1972) 669–686.

[22] H. W. Turnbull, A. C. Aitken: *An introduction to the theory of canonical matrices*, Dover, New York.

[23] J. H. M. Wedderburn: *Lectures on Matrices*. American Mathematical Society Colloquium Publications, volume 17. New York, American Mathematical Society, 1934.

# Índice alfabético