Universidad Nacional Autónoma de México

Facultad de Ciencias

Formas canónicas

T E S I S

PARA OBTENER EL TÍTULO DE:

Matemático

PRESENTA:

Mario Alberto Rodríguez Cedeño

TUTOR

Dr. Valente Santiago Vargas

Ciudad Universitaria, CD. MX., 2024

0

Introducción

El álgebra lineal tiene como uno de sus principales objetivos el estudio de las transformaciones lineales entre espacios vectoriales. Debido al gran alcance que tienen las aplicaciones cuando nos restringimos a espacios vectoriales de dimensión finita, este tipo particular de transformaciones ha sido estudiado con detalle. Uno de los resultados de mayor importancia obtenidos en este estudio (véase [4], p. 267) nos permite entender de manera sencilla una transformación lineal haciendo uso de ciertos escalares que representan a la función, en el sentido de que toda la información que necesitamos de la transformación se encuentra en esos valores; dichos escalares son los llamados valores propios.

Por supuesto, una propiedad tan importante como la anterior no es válida más que bajo ciertas condiciones estrictas. Para poder analizar esos casos que quedan fuera tenemos que recurrir a otros resultados que, si bien es cierto que son más generales, no nos proporcionan una perspectiva tan simple como la de los valores propios. Las herramientas que obtenemos de dichos resultados son las formas canónicas, a saber, la forma canónica de Jordan y dos formas canónicas racionales, la de divisores elementales y la de factores invariantes.

La existencia de estas representaciones se sigue de un caso particular del teorema fundamental de la teoría de módulos sobre dominios de ideales principales, que nos permite descomponer cualquier módulo sobre un dominio de ideales principales en componentes cuyo entendimiento es más accesible (véase el teorema 1.3.11). La importancia de este resultado radica en la posibilidad de asociarle a cada transformación lineal un módulo sobre el anillo de polinomios (véase la definición 1.4.1). Este puente entre transformaciones y módulos actúa como diccionario entre ambos mundos. De esta manera, la existencia de las formas canónicas se traduce en la existencia de la descomposición mientras que la unicidad de las formas canónicas, en la unicidad de la descomposición.

El objetivo de este trabajo será desarrollar la idea anterior a profundidad, para ello empezaremos el capítulo 1 con definiciones y resultados básicos de la teoría de anillos y módulos como la noción de módulo, la suma directa de módulos, dominios de ideales principales, dominios de factorización única, anillos noetherianos, etc. Estos resultados nos permitirán desarrollar la teoría de módulos sobre dominios de ideales principales para llegar al teorema de descomposición de estos módulos. Después de desarrollar con detalle la idea del puente entre el mundo de las transformaciones lineales y el de los módulos, usaremos el teorema anterior para probar la existencia de la forma canónica de Jordan y otro resultado más (véase el teorema 1.5.1), aunado a la unicidad de la descomposición de los módulos, nos asegurará la unicidad.

En el capítulo 2 usaremos el puente entre el mundo de los módulos y el de las matrices, mencionado anteriormente, para demostrar la existencia y unicidad de las formas canónicas racionales (de divisores elementales y de factores invariantes) y finalizaremos con unos ejemplos en los que se expondrá una manera (poco eficiente) de calcular las

formas canónicas.

En el capítulo 3 estudiaremos los diagramas de puntos y con ellos obtendremos un mejor método de cálculo de las formas canónicas. Después demostraremos el teorema de descomposición de Jordan cuya motivación yace en la forma canónica de Jordan. Para terminar, presentaremos una manera de aplicar la existencia de la forma canónica de Jordan a sistemas de ecuaciones diferenciales lineales.

Índice general

Capítulo 1

Forma canónica de Jordan

1.1. Definiciones y resultados básicos de la teoría de anillos

Empezaremos recordando algunos conceptos de la teoría de anillos. En todo este trabajo la palabra anillo se referirá a un anillo conmutativo con 1.

Definición 1.1.1. *Un ideal de un anillo R es un subgrupo I de R tal que $\forall r \in R$, si $x \in I$, entonces $rx \in I$. Un ideal I es **primo** si $I \neq R$ y $ab \in I$ implica que $a \in I$ o $b \in I$; por otro lado, diremos que $p \in R \setminus \{0\}$ es **primo** si $\langle p \rangle$ es primo. Finalmente, un ideal $I \neq R$ es **maximal** si es maximal en el conjunto de ideales propios de R ordenados por contención; y un elemento $0 \neq q \in R \setminus R^*$, con $R^* = \{r \in R \mid r \text{ es unidad}\}$, es **irreducible** si $q = ab$ implica que a o b es una unidad.*

Definición 1.1.2. *Un anillo R es un **dominio entero** si $ab = 0$ implica que $a = 0$ o $b = 0$. Un **dominio de ideales principales (DIP)** R es un dominio entero con la propiedad de que cualquier ideal está generado por un elemento. Por otra parte, R es un **dominio de factorización única (DFU)** si, dado $x \in R$, existen $q_1, ..., q_n \in R$ elementos irreducibles y $u \in R$ una unidad tales que*

$$x = uq_1q_2 \cdots q_n$$

y esta factorización es única salvo por asociados (recordemos que a es asociado a b si $a = ub$ para alguna unidad $u \in R$).

Definición 1.1.3. *Sean $a, b \in R$. Un **máximo común divisor** de a y b es un elemento $d \in R$ tal que*

a) $d \mid a$ y $d \mid b$,

b) si $d' \mid a$ y $d' \mid b$, entonces $d' \mid d$,

donde $x \mid y$ significa que existe $z \in R$ tal que $y = zx$.

Podemos reescribir la definición anterior en términos de ideales como sigue. Dados un anillo R y $X = \{a_1, ..., a_n\} \subseteq R$, denotaremos por $\langle a_1, ..., a_n \rangle$ al ideal generado por el conjunto X.

Definición 1.1.4. *Sean $a, b \in R$. Un máximo común divisor de a y b es un elemento $d \in R$ tal que*

a) $\langle a, b\rangle \subseteq \langle d\rangle$,

b) si $\langle a, b\rangle \subseteq \langle d'\rangle$, *entonces* $\langle d\rangle \subseteq \langle d'\rangle$.

Decimos que a y b son ***coprimos (o primos relativos)*** *si 1 es un máximo común divisor de a y b.*

Si R es un DIP, entonces $\langle a, b\rangle$ es principal, i.e., existe $r \in R$ tal que $\langle r\rangle = \langle a, b\rangle$. El elemento r satisface la definición de máximo común divisor, pues $\langle r\rangle$ es el ideal principal más pequeño que contiene a $\langle a, b\rangle$. De la definición de coprimos y lo anterior se tiene la siguiente proposición.

Proposición 1.1.5. *Sean R un DIP y $a, b \in R$. Entonces las siguientes condiciones son equivalentes.*

a) Los elementos a y b son coprimos.

b) $\langle a, b\rangle = R$.

c) Existen $s, t \in R$ tales que $sa + tb = 1$.

Proposición 1.1.6. *Sea R un DIP. Entonces:*

a) $p \in R$ es primo si, y sólo si, es irreducible.

b) R es un DFU.

Demostración. Véase [3], p. 284, 287. □

Definición 1.1.7. *Decimos que un campo F es* ***algebraicamente cerrado*** *si cualquier polinomio con coeficientes en F tiene una raíz en F.*

Si F es un campo algebraicamente cerrado, de la definición anterior, se sigue que, para $p \in F[x]$, $p(x) = q(x)(x - \alpha_1)$, para algún $\alpha_1 \in F$ y algún $q \in F[x]$. Repitiendo el proceso llegamos a que $p(x) = (x - \alpha_1)\cdots(x - \alpha_n)$, donde $\text{grad}(p) = n$. Lo anterior se resume en la siguiente proposición.

Proposición 1.1.8. *Sean F un campo algebraicamente cerrado y $p \in F[x]$ no constante. Entonces p se factoriza como producto de polinomios de grado uno en $F[x]$.*

Proposición 1.1.9. *Sea F un campo algebraicamente cerrado. Entonces $p \in F[x]$ es irreducible si, y sólo si,* $\text{grad}(p) = 1$.

Demostración. Supongamos que p es irreducible y no constante. Como F es algebraicamente cerrado, $p(x) = (x - \alpha)q(x)$, con $\alpha \in F$ y $q \in F[x]$. Dado que p es irreducible, se sigue que $q(x)$ es una unidad. Por lo tanto, $\text{grad}(p) = 1$. El regreso es claro. □

Proposición 1.1.10. *Sea F un campo. Entonces $F[x]$ es un DIP.*

Demostración. Claramente, $F[x]$ es un dominio entero, pues F es un campo. Sean I un ideal de $F[x]$ y $q \in I$ tal que $\text{grad}(q)$ es mínimo en I. Si $p \in I$, entonces, por el algoritmo de la división, existen $s, r \in F[x]$ tales que $p = sq + r$, con $\text{grad}(r) < \text{grad}(q)$ ó $r = 0$. Si $r \neq 0$, como $p, q \in I$, se concluye que $r = p - sq \in I$, lo cual es una contradicción porque $\text{grad}(q)$ es mínimo en I. Por lo tanto, $r = 0$ y $p = sq$, i.e., $I = \langle q\rangle$. □

1.2. Definiciones y resultados básicos de la teoría de módulos

Comenzaremos con la definición de módulo.

Definición 1.2.1. *Sea R un anillo conmutativo con uno. Un **R-módulo** es un grupo abeliano M junto con una función (producto por escalares) $\cdot : R \times M \longrightarrow M$ con las siguientes propiedades, $\forall a, b \in R$ y $\forall m, m' \in M$*

a) $a \cdot (b \cdot m) = (ab) \cdot m$

b) $1 \cdot m = m$

c) $a \cdot (m + m') = a \cdot m + a \cdot m'$

d) $(a + b) \cdot m = a \cdot m + b \cdot m$.

*Si por el contexto es claro, llamaremos módulo a un R-módulo y utilizaremos am para denotar $a \cdot m$. Un **submódulo** N de un módulo M, denotado $N \leq M$, es un subgrupo de M con estructura de R-módulo dada por la restricción del producto por escalares de M a N. Dado $S \subseteq M$, denotaremos por*

$$\langle S \rangle = \{r_1 x_1 + \cdots + r_n x_n \in M \mid r_i \in R, x_i \in S, i = 1, ..., n, \text{ con } n \in \mathbb{N}\}$$

al submódulo más pequeño de M que contiene a S o el submódulo generado por S.

Ejemplos típicos de módulos son R^n, con $n \in \mathbb{N}$, y $R[x]$, con el producto por escalares usual. Como caso particular, si $n = 1$, los submódulos de $R^n = R$ son los ideales de R.

La siguiente proposición caracteriza a los submódulos de un módulo M.

Proposición 1.2.2. *Sean M un R-módulo y $\emptyset \neq N \subseteq M$. Entonces, N es un submódulo de M si, y sólo si, para todo $x, y \in N$ y $\lambda \in R$ se tiene que $\lambda x + y \in N$.*

Definición 1.2.3. *Sea M un R-módulo. Decimos que M es **finitamente generado** (f.g.) si existen $m_1, ..., m_n \in M$ tales que $M = \langle m_1, ..., m_n \rangle$. Si $n = 1$, diremos que M es **cíclico**.*

Definición 1.2.4. *Dados $M_1, ..., M_n$ R-módulos, definimos su suma directa (externa) como*

$$\coprod_{i=1}^{n} M_i := \{(x_1, ..., x_n) \mid x_i \in M_i, i = 1, ..., n\}.$$

Este conjunto junto con la suma y el producto escalar entrada a entrada es un R-módulo. De manera más general, si $\{M_i\}_{i \in I}$ es una familia de R-módulos, se define el R-módulo $\amalg_{i \in I} M_i$ como

$$\coprod_{i \in I} M_i := \{(x_i)_{i \in I} \mid \forall i \in I \ x_i \in M_i \text{ y } x_i = 0 \text{ para } i \in I \setminus \{i_1, ..., i_k\}, \text{ con } k \geq 1\}.$$

En el caso en el que $M_i = M$ para todo $i \in I$, escribiremos $M^{(I)}$ en lugar de $\amalg_{i \in I} M_i$.

Definición 1.2.5. *Sean M un R-módulo y $N_1, ..., N_k$ submódulos de M. Decimos que M es la suma directa (interna) de $N_1, ..., N_k$ y lo denotaremos por*

$$M = \bigoplus_{i=1}^{k} N_i$$

si

a) $M = \Sigma_{i=1}^{k} N_i = \{n_1 + \cdots + n_k \in M \mid n_i \in N_i, 1 \leq i \leq k\}$

b) $N_j \cap \Sigma_{i \neq j} N_i = \{0\},\ j = 1, ..., k.$

Al igual que en Álgebra lineal, en módulos también se tiene el concepto de independencia lineal y de generadores.

Definición 1.2.6. *Sea M un R-módulo. Decimos que un conjunto $S \subseteq M$ es **linealmente dependiente** si existen $x_1, ..., x_n \in S$ y escalares $r_1, ..., r_n \in R$ no todos cero tales que*

$$r_1 x_1 + \cdots + r_n x_n = 0.$$

*En caso contrario, S será **linealmente independiente**.*
*Si $S \subseteq M$ es tal que $\langle S \rangle = M$, diremos que S genera a M o que S es un **conjunto generador** de M.*

Al igual que en álgebra lineal, tenemos el concepto de base para un R-módulo.

Definición 1.2.7. *Sean M un R-módulo y $S \subseteq M$. Decimos que S es una **base** para M si S genera a M y S es un conjunto linealmente independiente.*

Definición 1.2.8. *Sean M, N dos R-módulos. Una función $f : M \longrightarrow N$ es un **morfismo de R-módulos** si es un morfismo de grupos y $\forall r \in R\ \forall m \in M,\ f(rm) = rf(m)$.*
Definimos el conjunto $\mathrm{Ker}(f)$ *como*

$$\mathrm{Ker}(f) := \{m \in M \mid f(m) = 0\}$$

y $\mathrm{Hom}_R(M, N)$ *como el conjunto de morfismos de R-módulos de M en N.*
*Si f es biyectiva, diremos que f es un **isomorfismo**, en este caso, se escribirá $M \cong N$.*

La siguiente noción es importante en la teoría de módulos y se utilizará en el teorema 1.3.8.

Definición 1.2.9. *Decimos que un R-módulo M es **libre** si $M \cong R^{(J)}$, con J un conjunto.*

Las bases de R-módulos tienen propiedades similares a las de espacios vectoriales, como veremos en la siguiente proposición. Antes de enunciarla daremos la siguiente definición que será de utilidad.

Definición 1.2.10. *Sean M un R-módulo y N un submódulo de M. Definimos el **módulo cociente**, denotado por M/N, como el grupo cociente de M entre N con producto por escalares dado por*

$$\begin{array}{rcl} R \times M/N & \longrightarrow & M/N \\ (r, m + N) & \mapsto & (rm) + N. \end{array}$$

Con este producto por escalares, M/N es un R-módulo.

Teorema 1.2.11 (Primer Teorema de Isomorfismo)**.** *Sean M, N dos R-módulos y $\phi : M \longrightarrow N$ un morfismo de R-módulos. Entonces* $\mathrm{Ker}(\phi)$ *es un submódulo de M y* $M/\mathrm{Ker}(\phi) \cong \phi(M)$.

Demostración. Véase [3], p. 349. □

Proposición 1.2.12. *Sean M, N dos R-módulos y $S \subseteq M$ una base para M. Entonces se cumple lo siguiente:*

a) Dada una función $f : S \longrightarrow N$, existe un único morfismo de R-módulos $\bar{f} : M \longrightarrow N$ tal que $\bar{f}|_S = f$.

b) Si S' es otra base para M, entonces $|S| = |S'|$.

Demostración. La demostración del primer inciso es análoga a la prueba en el caso de espacios vectoriales.
Para probar el segundo inciso, sea $\mathfrak{m} \subseteq R$ un ideal maximal de R y $\mathfrak{m}M \subseteq M$ el submódulo de M que consiste de todas las sumas de la forma $r_1m_1 + \cdots + r_nm_n$, con $r_i \in \mathfrak{m}$, $m_i \in M$ y $n \in \mathbb{N}$. Sabemos que $M/\mathfrak{m}M$ es un R-módulo, además, es un $R/\mathfrak{m}$-espacio vectorial con el siguiente producto por escalares,

$$\begin{array}{rcl} R/\mathfrak{m} \times M/\mathfrak{m}M & \longrightarrow & M/\mathfrak{m}M \\ (r + \mathfrak{m}, x + \mathfrak{m}M) & \mapsto & (rx) + \mathfrak{m}M. \end{array}$$

El producto está bien definido, pues si $r + \mathfrak{m} = r' + \mathfrak{m}$ y $x + \mathfrak{m}M = y + \mathfrak{m}M$, entonces $r - r' \in \mathfrak{m}$ y $x - y \in \mathfrak{m}M$, de modo que

$$rx - r'y = (r - r')x + r'(x - y) \in \mathfrak{m}M.$$

Las demás propiedades de espacio vectorial se heredan de la estructura de R-módulo.
El conjunto S genera a M lo que implica que $S + \mathfrak{m}M$ genera genera a $M/\mathfrak{m}M$. Si

$$(r_1 + \mathfrak{m})(s_1 + \mathfrak{m}M) + \cdots + (r_n + \mathfrak{m})(s_n + \mathfrak{m}M) = (r_1s_1 + \cdots + r_ns_n) + \mathfrak{m}M = 0,$$

entonces $r_1s_1 + \cdots + r_ns_n = r'_1m_1 + \cdots + r'_{n'}m_{n'}$, con $r'_i \in \mathfrak{m}$, $m_i \in M$ y $s_i \in S$. Pero, al expresar cada m_i como combinación lineal de elementos de S, llegamos a dos expresiones de un elemento en términos de la base, de donde se sigue que los coeficientes y los elementos de la base tienen que ser los mismos. Sin embargo, los coeficientes de la expresión del lado derecho de la igualdad están en $\mathfrak{m}$ y, por ende, $r_i \in \mathfrak{m}$, $i = 1, ..., n$; esto es, $r_i + \mathfrak{m} = 0$, $i = 1, ..., n$. Por lo tanto, $S + \mathfrak{m}M$ es una base del espacio vectorial $M/\mathfrak{m}M$ y $|S| = |S + \mathfrak{m}M|$. Aplicando lo mismo con S' y utilizando que en un espacio vectorial todas las bases tienen el mismo número de elementos, se sigue que $|S| = |S'|$. □

Corolario 1.2.13. *Sea M un R-módulo finitamente generado. Entonces existe $\phi \in \mathrm{Hom}_R(R^n, M)$ suprayectivo, para alguna $n \in \mathbb{N}$.*

Demostración. Sean $x_1, ..., x_n \in M$ generadores de M. Por el primer inciso de la proposición 1.2.12, existe un único morfismo de R-módulos $\phi : R^n \longrightarrow M$ tal que $\phi(e_i) = x_i$, donde $\{e_1, ..., e_n\}$ es la base canónica de R^n. Como $x_1, ..., x_n \in \phi(R^n)$, se concluye que ϕ es suprayectivo. □

A continuación daremos la definición de un módulo noetheriano y sus propiedades básicas.

Definición 1.2.14. *Sea M un R-módulo. Decimos que M es* ***noetheriano*** *si cualquier cadena ascendente de submódulos*

$$M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n \subseteq \cdots$$

se estabiliza, i.e., si existe $n \in \mathbb{N}$ tal que para toda $k \geq n$ se tiene que $M_n = M_k$.

Hay otras formas equivalentes de definir un módulo noetheriano.

Proposición 1.2.15. *Sea M un R-módulo. Las siguientes condiciones son equivalentes:*

a) M es noetheriano.

b) Cualquier conjunto $\Sigma \neq \emptyset$ de submódulos tiene un elemento maximal.

c) Cualquier submódulo de M es finitamente generado.

Demostración. Véase [1], p. 74 - 75. □

Como consecuencia inmediata tenemos el siguiente resultado.

Corolario 1.2.16. *Todo dominio de ideales principales es noetheriano.*

El siguiente resultado nos proporciona una manera de identificar R-módulos noetherianos, siempre que R sea noetheriano.

Proposición 1.2.17. *Sea M un R-módulo finitamente generado, con R noetheriano. Entonces M es noetheriano.*

Demostración. Véase [1], p. 76. □

1.3. Teoremas de estructura de módulos sobre dominios de ideales principales

Los resultados que se presentarán en esta sección serán de gran utilidad para la demostración de la existencia de la forma canónica de Jordan.

Definición 1.3.1. *Sean M un R-módulo y $m \in M$. El* ***anulador de m*** *está dado por*

$$\mathrm{Ann}(m) := \{r \in R : rm = 0\}.$$

El ***anulador de M*** *se define como*

$$\mathrm{Ann}(M) := \{r \in R : \forall m \in M \ rm = 0\}.$$

De la definición anterior se infiere que $\mathrm{Ann}(m) = \mathrm{Ker}(\phi)$, donde $\phi \in \mathrm{Hom}_R(R, M)$ está dado por $\phi(r) = rm$, por ello, $\mathrm{Ann}(m)$ es un ideal de R. El hecho de que $\mathrm{Ann}(M)$ sea un ideal se sigue de la igualdad $\mathrm{Ann}(M) = \bigcap_{m \in M} \mathrm{Ann}(m)$.

Definición 1.3.2. *Sean M un R-módulo y $m \in M$. Decimos que m es un elemento de* ***torsión*** *si* $\text{Ann}(m) \neq \{0\}$. *Si todo elemento de M es de torsión, entonces diremos que M es de* ***torsión***. *Si ningún elemento distinto de cero es de torsión, diremos que M es* ***libre de torsión***.

Definición 1.3.3. *Sean M un R-módulo y $r \in R$. Definimos $M(r)$ y M_r como sigue:*

$$\begin{aligned} M(r) &= \{m \in M |\ rm = 0\}, \\ M_r &= \{m \in M |\ \exists k \in \mathbb{N} : r^k m = 0\}. \end{aligned}$$

Proposición 1.3.4. *Sean M un R-módulo y $r \in R$. Entonces $M(r)$ y M_r son submódulos de M y $M(r) \subseteq M_r$.*

Demostración. Claramente, $0 \in M(r) \cap M_r$. Sean $x \in R$ y $m, m' \in M(r)$. Entonces $r(xm + m') = xrm + rm' = 0$, i.e., $xm + m' \in M(r)$.
Si $m, m' \in M_r$, entonces existen $k, k' \in \mathbb{N}$ tales que $r^k m = r^{k'} m' = 0$, de modo que $r^{k+k'}(xm + m') = xr^{k+k'}m + r^{k+k'}m' = 0$ y, por ende, $xm + m' \in M_r$. Por la proposición 1.2.2, $M(r)$ y M_r son submódulos de M. □

Definición 1.3.5. *Sea R un dominio de ideales principales. Sea M un R-módulo y $p \in R$ primo. Decimos que M es un* ***p-módulo*** *si $M = M_p$.*

Teorema 1.3.6. *Sean R un dominio de ideales principales y $0 \neq M$ un R-módulo de torsión finitamente generado. Entonces existen primos $p_1, ..., p_n \in R$, únicos salvo asociados y salvo el orden de los índices, con las siguientes dos propiedades:*

a) $M_{p_i} \neq 0 \ \ \forall i = 1, ..., n$ *y*

b) $M = \bigoplus_{i=1}^{n} M_{p_i}$.

Además, se satisface que $\text{Ann}(M) = \left\langle \prod_{i=1}^{n} p_i^{\alpha_i} \right\rangle$ *y* $\text{Ann}(M_{p_i}) = \langle p_i^{\alpha_i} \rangle$ *para todo* $i = 1, ..., n$, *donde α_i es un entero positivo.*

Demostración. Primero que nada, notemos que $\text{Ann}(M)$ no es cero ni todo R. En efecto, si $\text{Ann}(M) = R$, entonces para $x \in M$, $x = 1 \cdot x = 0$, pero $M \neq \{0\}$. Ahora, dado que M es finitamente generado y de torsión, existen $x_1, ..., x_k \in M$, $r_1, ..., r_k \in R \setminus 0$ tales que $M = \langle x_1, ..., x_k \rangle$ y $r_i x_i = 0$. Esto implica que $ax_i = 0$ para toda $i \in \{1, ..., k\}$, donde $a = \Pi_{i=1}^{k} r_i \neq 0$; es decir, a anula a todos los generadores, por lo tanto, a todo M, de donde concluimos que $\text{Ann}(M) \neq \{0\}$.
Sabemos que R es un DIP, entonces $\text{Ann}(M) = Rg = \{rg \in R \mid r \in R\} = \langle g \rangle$, para alguna $g \in R \setminus \{0\}$. Además, por la proposición 1.1.6, R es un dominio de factorización única y así

$$g = u\Pi_{i=1}^{n} p_i^{\alpha_i},$$

donde $p_i \in R$ es primo (todos no asociados entre sí) y $u \in R$ es una unidad. Como $g \in \text{Ann}(M)$, se sigue que

$$M = M(g) = M(u\Pi_{i=1}^{n} p_i^{\alpha_i}) = M(\Pi_{i=1}^{n} p_i^{\alpha_i}),$$

donde la última igualdad se da debido a que para toda $r \in R$, $urm = 0$ si, y sólo si, $rm = 0$, pues u es una unidad. Lo siguiente será probar que $M(g)$ "abre productos" bajo ciertas condiciones. Para ello recordemos que, por la proposición 1.1.5, si p y q son coprimos en R, entonces $sp + tq = 1$, para algunos $s, t \in R$. De donde

$$x = 1 \cdot x = spx + tqx, \text{ para todo } x \in M(pq). \tag{1.1}$$

Notemos que en la suma anterior se tiene que $spx \in M(q)$ ya que $q(spx) = spqx = 0$ y de manera similar $tqx \in M(p)$, i.e.,

$$M(pq) = M(p) + M(q).$$

Utilizando las igualdades (1.1) tenemos que si $x \in M(p) \cap M(q)$, entonces $x = spx + tqx = 0$; por lo tanto, $M(p) \cap M(q) = \{0\}$, de donde $M(pq) = M(p) \oplus M(q)$. Usando un argumento inductivo llegamos a que

$$M = M(g) = \bigoplus_{i=1}^{n} M(p_i^{\alpha_i}).$$

Ahora probaremos que $M_{p_i} = M(p_i^{\alpha_i})$. Veamos la contención $M_{p_i} \subseteq M(p_i^{\alpha_i})$. Supongamos que $x \in M_{p_i}$. Por la igualdad de arriba, tenemos que $x = x_1 + \cdots + x_n$, con $x_j \in M(p_j^{\alpha_j})$ y $j = 1, ..., n$. Entonces

$$p_i^k x = p_i^k x_1 + \cdots + p_i^k x_n = 0, \text{ para alguna } k \in \mathbb{N}.$$

Como $M = \bigoplus_{i=1}^{n} M(p_i^{\alpha_i})$ y $p_i^k x_j \in M(p_j^{\alpha_j})$, tenemos que:

$$p_i^k x_j = 0, \forall j \in \{1, ..., n\}.$$

Demostraremos que $x_j = 0$, si $j \neq i$. Notemos que p_i^k y $p_j^{\alpha_j}$ son primos relativos si $j \neq i$. Entonces, existen $r, s_j \in R$ tales que

$$1 = rp_i^k + s_j p_j^{\alpha_j}.$$

Por lo tanto, $x_j = 1x_j = (rp_i^k + s_j p_j^{\alpha_j})x_j = rp_i^k x_j + s_j p_j^{\alpha_j} x_j = 0 + 0 = 0$ pues $p_i^k x_j = 0$ y $x_j \in M(p_j^{\alpha_j})$. Por ello, $x = x_i \in M(p_i^{\alpha_i})$, probándose la contención $M_{p_i} \subseteq M(p_i^{\alpha_i})$. Claramente, $M(p_i^{\alpha_i}) \subseteq M_{p_i}$, por lo tanto, $M(p_i^{\alpha_i}) = M_{p_i}$ y

$$M = \bigoplus_{i=1}^{n} M_{p_i}.$$

Ningún $M(p_i^{\alpha_i}) = M_{p_i}$ es cero por el siguiente argumento. Si $M(p_j^{\alpha_j}) = 0$ para alguna j, entonces, para toda $x \in M$, tenemos que $0 = gx = (u\Pi_{i=1}^{n} p_i^{\alpha_i})x = p_j^{\alpha_j}(u\Pi_{i \neq j} p_i^{\alpha_i})x$ implica que $(u\Pi_{i \neq j} p_i^{\alpha_i})x = 0$, i.e., $u\Pi_{i \neq j} p_i^{\alpha_i} \in \text{Ann}(M) = Rg$. Como $p_j \mid g$, se tiene que $p_j \mid u\Pi_{i \neq j} p_i^{\alpha_i}$, lo cual es una contradicción, pues R es un DFU por la proposición 1.1.6. Con esto hemos probado la existencia de los primos $p_1, ..., p_n$ con las propiedades deseadas. Ahora veamos la unicidad.
Sean $q_1, ..., q_m$ primos distintos de R con $M_{q_i} \neq 0$ para $i = 1, ..., m$ y tales que

$$M = \bigoplus_{i=1}^{m} M_{q_i}.$$

De la suma anterior, tenemos que $M_{q_i} \neq M_{q_j}$ si $i \neq j$. Sea $0 \neq x \in M_{q_1}$. Entonces $q_1^k \in \mathrm{Ann}(x)$ para alguna $k \in \mathbb{N}$. Como R es un dominio de ideales principales, tenemos que $\mathrm{Ann}(x) = \langle r \rangle$ para alguna $r \in R$. Por lo tanto,

$$r | q_1^k \text{ y así } r = q_1^s \text{ para alguna } s \in \mathbb{N}.$$

Como $\langle g \rangle = \mathrm{Ann}(M)$, tenemos que $g \in \mathrm{Ann}(x) = \langle q_1^s \rangle$ y por lo tanto $q_1^s | g$, de donde concluimos que $q_1 | g$. Como R es un Dominio de Factorización Única, tenemos que q_1 es asociado a algún primo p_i de la factorización de $g = u \prod_{i=i}^n p_i^{\alpha_i}$. Después de reordenar, los factores primos p_i de g, podemos suponer que q_1 es asociado a p_1. Notemos que q_1 es asociado a un único p_i de la factorización de g. De esta forma, tenemos que $M_{q_1} = M_{p_1}$. Ahora si existieran dos primos q_i y q_j con $M_{q_i} \neq M_{q_j}$ tal que ambos son asociados al mismo divisor primo p de g, tendríamos que q_i es asociado a q_j y así $M_{q_i} = M_{q_j}$, lo cual es una contradicción.
Ahora, repitiendo el procedimiento anterior para q_2, tenemos que existe un primo p_i tal que q_2 es asociado a p_i y por lo dicho arriba, tenemos que $p_i \neq p_1$. De esta manera, después de reordenar los factores de g, podemos suponer que q_2 es asociado a p_2 y así $M_{q_2} = M_{p_2}$. Repitiendo el proceso para cada $i = 1, ..., m$, tenemos que $M_{q_i} = M_{p_i}$ para todo $i = 1, ..., m$. Por lo tanto $m \leq n$.
Supongamos que $m < n$. Como $M = \bigoplus_{j=1}^m M_{p_j} = \bigoplus_{i=1}^n M_{p_i}$, tenemos que para $m < i \leq n$:

$$M_{p_i} = M \cap M_{p_i} = \Big(\sum_{j=1}^m M_{p_j} \Big) \cap M_{p_i} \subseteq \Big(\sum_{j \in \{1,2,...,n\} - \{i\}} M_{p_j} \Big) \cap M_{p_i} = \{0\}.$$

Por lo tanto, $M_{p_i} = \{0\}$, lo cual es una contradición, por lo tanto $n = m$ y $M_{q_i} = M_{p_i}$ para $i = 1, ..., m$, probándose la unicidad.
Veamos que $\mathrm{Ann}(M_{p_i}) = \langle p_i^{\alpha_i} \rangle$ para todo $i = 1, ..., n$. En efecto, como R es un DIP, existe $h_i \in R$ tal que $\mathrm{Ann}(M_{p_i}) = \langle h_i \rangle$. Notemos que hemos visto que

$$M_{p_i} = M(p_i^{\alpha_i}) = \Big\{ m \in M \mid p_i^{\alpha_i} m = 0 \Big\}.$$

Por lo tanto, $p_i^{\alpha_i} \in \langle h_i \rangle$, es decir, tenemos que $h_i \mid p_i^{\alpha_i}$. Como R es un DFU, tenemos que $h_i = p_i^{\beta_i}$ con $\beta_i \leq \alpha_i$ para todo $i = 1, ..., n$. Ahora, tenemos que

$$M(p_i^{\alpha_i}) = M(p_i^{\beta_i}) \ \ \forall i = 1, ..., n.$$

En efecto, sea $m \in M(p_i^{\alpha_i})$, como $\mathrm{Ann}(M(p_i^{\alpha_i})) = \mathrm{Ann}(M_{p_i}) = \langle h_i \rangle = \langle p^{\beta_i} \rangle$, tenemos que $p_i^{\beta_i} m = 0$ y así $m \in M(p_i^{\beta_i})$. De donde concluimos que $M(p_i^{\alpha_i}) \subseteq M(p_i^{\beta_i})$.
Por otro lado, sea $m \in M(p_i^{\beta_i})$. Entonces $p_i^{\beta_i} m = 0$ y como $\beta_i \leq \alpha_i$, tenemos que $p_i^{\alpha_i} m = 0$ y así tenemos que $M(p_i^{\beta_i}) \subseteq M(p_i^{\alpha_i})$. Probándose que $M(p_i^{\alpha_i}) = M(p_i^{\beta_i})$.
Luego, tenemos que

$$M = M\Big(\prod_i^n p_i^{\alpha_i} \Big) = \bigoplus_{i=1}^n M(p_i^{\alpha_i}) = \bigoplus_{i=1}^n M(p_i^{\beta_i}) = M\Big(\prod_i^n p_i^{\beta_i} \Big).$$

Por lo tanto, $\prod_i^n p_i^{\beta_i} \in \mathrm{Ann}(M) = \langle \prod_{i=1}^n p_i^{\alpha_i} \rangle$. Y así $\prod_i^n p_i^{\alpha_i} \mid \prod_i^n p_i^{\beta_i}$. De donde concluimos que $\alpha_i = \beta_i$ para todo $i = 1, ..., n$. Probándose que $\mathrm{Ann}(M_{p_i}) = \langle p_i^{\alpha_i} \rangle$ para todo $i = 1, ..., n$. $\square$

A continuación daremos la definición del rango de un módulo.

Definición 1.3.7. *Sean R un dominio entero y M un R-módulo. Definimos el* ***rango*** *de M como el máximo número de elementos linealmente independientes en M (ver definición 1.2.6) y se denotará por* $\mathrm{ran}(M)$.

Lo siguiente será probar que todo p-módulo, con p primo, se descompone como suma directa de módulos cíclicos. Para ello necesitamos los siguientes resultados.

Teorema 1.3.8. *Sea M un R-módulo libre de rango n, con R un DIP, y $N \leq M$. Entonces*

I) *N es libre de rango $m \leq n$.*

II) *Existe $\{y_1, ..., y_n\}$ una base de M y $a_1, ..., a_m \in R \setminus \{0\}$, con $m \leq n$, tales que $a_1y_1, ..., a_my_m$ forman una base para N y*

$$a_1 \mid a_2 \mid ... \mid a_m.$$

Demostración. El resultado es trivial si $N = \{0\}$. Supongamos que $N \neq \{0\}$. Consideremos $\{x_1, ..., x_n\}$ una base de M. Sea $\phi : M \longrightarrow R$ un morfismo de R-módulos. Sabemos que $\phi(N)$ es un submódulo de R y los submódulos de R son ideales, por ello, $\phi(N)$ es un ideal de R. Como R es un DIP, $\phi(N) = \langle a_\phi \rangle$, con $a_\phi \in R$. Definimos

$$\Sigma := \{\langle a_\phi \rangle \mid \phi \in \mathrm{Hom}_R(M, R)\}.$$

Notemos que $\Sigma \neq \emptyset$, pues si $\phi = 0$, entonces $\phi(N) = \{0\} = \langle 0 \rangle$. Dado que R es noetheriano por el corolario 1.2.16, usando la proposición 1.2.15, se obtiene que Σ posee un elemento maximal $\gamma(N) = \langle a_\gamma \rangle$, para algún $\gamma \in \mathrm{Hom}_R(M, R)$. Sean $a_1 = a_\gamma$ y $y \in N$ tales que $\gamma(y) = a_1$.
Notemos que $a_1 \neq 0$, ya que como $x_1, ..., x_n$ es una base de M y $\pi_1, ..., \pi_n \in \mathrm{Hom}_R(M, R)$ son las respectivas proyecciones en R dadas por $\pi_i(r_1x_1 + \cdots + r_nx_n) = r_i$, entonces existe $i \in \{1, ..., n\}$ tal que $\pi_i(N) \neq \{0\}$, pues $N \neq \{0\}$, i.e. $\langle 0 \rangle \subsetneq \pi_i(N) \in \Sigma$, lo cual no sería posible si $\langle 0 \rangle$ fuera maximal en Σ.
Ahora veamos que $a_1 \mid \phi(y)$, $\forall \phi \in \mathrm{Hom}_R(M, R)$. Como R es un DIP, existe $d \in R$ tal que $\langle d \rangle = \langle a_1, \phi(y) \rangle$. Esto implica que $d = r_1a_1 + r_2\phi(y)$. Definimos

$$\psi := r_1\gamma + r_2\phi : M \longrightarrow R,$$

de esta forma $\psi(y) = r_1\gamma(y) + r_2\phi(y) = r_1a_1 + r_2\phi(y) = d$, de donde $d \in \psi(N)$. Como $\langle a_1 \rangle \subseteq \langle d \rangle \subseteq \psi(N)$ y $\langle a_1 \rangle$ es maximal, se tiene que $\langle a_1 \rangle = \langle d \rangle = \psi(N)$. Luego, como $\phi(y) \in \langle a_1, \phi(y) \rangle = \langle d \rangle = \langle a_1 \rangle$, tenemos que $ra_1 = \phi(y)$, para alguna $r \in R$, i.e., $a_1 \mid \phi(y)$. En particular, para cada proyección $\pi_i : M \longrightarrow R$ tenemos que $\pi_i(y) = a_1b_i$, para alguna $b_i \in R$, $i = 1, ..., n$. Es decir, $y = a_1b_1x_1 + a_1b_2x_2 + \cdots + a_1b_nx_n = a_1(b_1x_1 + \cdots + b_nx_n)$. Si definimos

$$y_1 := \sum_{i=1}^{n} b_ix_i \in M,$$

se sigue que $y = a_1y_1$. Por lo tanto, $a_1 = \gamma(y) = a_1\gamma(y_1)$ y como R es un dominio entero, concluimos que $\gamma(y_1) = 1$ pues $a_1 \neq 0$.
Ahora probaremos las siguientes dos igualdades:

$$M = Ry_1 \oplus \text{Ker}\gamma$$
$$N = Ra_1y_1 \oplus (N \cap \text{Ker}\gamma)$$

Para la primera, si $x \in M$, entonces $x = \gamma(x)y_1 + (x - \gamma(x)y_1)$ y $\gamma(x - \gamma(x)y_1) = \gamma(x) - \gamma(x)\gamma(y_1) = 0$. Esto es, $x \in Ry_1 + \text{Ker}\gamma$, de modo que $M = Ry_1 + \text{Ker}\gamma$. Si $x \in Ry_1 \cap \text{Ker}\gamma$, se tendría que $x = ry_1$ y $0 = \gamma(x) = \gamma(ry_1) = r\gamma(y_1) = r$. De donde $x = 0$. Por lo tanto, $M = Ry_1 \oplus \text{Ker}\gamma$.
Para la segunda igualdad, recordemos que $\gamma(N) = \langle a_1 \rangle$, esto implica que si $x \in N$, entonces $x = \gamma(x)y_1 + (x - \gamma(x)y_1) = ra_1y_1 + (x - \gamma(x)y_1)$. Notemos que $x - \gamma(x)y_1 = x - ra_1y_1 = x - ry$ y como $x, y \in N$, concluimos que $x - \gamma(x)y_1 \in N$. Por ello, $N = Ra_1y_1 + (N \cap \text{Ker}\gamma)$. Que la intersección sea cero se sigue de la misma manera que con la primera igualdad, probándose las dos igualdades deseadas.
Dado que M es f.g. y R es noetheriano (por ser un DIP), se sigue de la proposición 1.2.17 que M es un módulo noetheriano; por la proposición 1.2.15, tenemos que N es f.g y, por lo tanto, de rango finito. Por esta razón podemos aplicar inducción sobre m, el rango de N, para probar el primer inciso.
Primero, veamos que si $\text{ran}(N) = m$, entonces $\text{ran}(N \cap \text{Ker}(\gamma)) < m$. En efecto, sean $w_1, ..., w_k \in N \cap \text{Ker}(\gamma)$ elementos R-linealmente independientes. Afirmamos que el conjunto $\{a_1y_1, w_1, ..., w_k\}$ es R-linealmente independiente en N. Para ver esto, consideremos escalares $\lambda_i \in R$ tales que

$$\lambda_1a_1y_1 + \lambda_2w_1 + \cdots + \lambda_{k+1}w_k = 0.$$

Aplicando γ a la igualdad anterior y utilizando que cada $w_i \in \text{Ker}(\gamma)$, tenemos que $0 = \lambda_1a_1\gamma(y_1) = \lambda_1a_1$ y como R es un dominio entero y $a_1 \neq 0$, concluimos que $\lambda_1 = 0$. Luego, la combinación R-lineal anterior se convierte en $\sum_{i=1}^{k} \lambda_{i+1}w_i = 0$. Por lo tanto, $\lambda_{i+1} = 0$ para $i = 1, ..., k$, ya que $\{w_1, ..., w_k\}$ es linealmente independiente. Como $\text{ran}(N)$ es el máximo número de elementos linealmente independientes en N, tenemos que $k + 1 \leq \text{ran}(N)$. Por lo tanto, $k \leq \text{ran}(N) - 1$ y como esto se hace para cualquier conjunto de elementos linealmente independientes de $N \cap \text{Ker}(\gamma)$, concluimos que $\text{ran}(N \cap \text{Ker}(\gamma)) \leq \text{ran}(N) - 1 = m - 1$.
De la misma manera se prueba que $\text{ran}(\text{Ker}(\gamma)) < \text{ran}(M)$. Ahora, procedamos por inducción en $\text{ran}(N) = m$ para probar el primer inciso.
Si $m = 0$, quiere decir que para todo $z \in N$, el conjunto $\{z\}$ es linealmente dependiente. Por lo tanto, para $z \in N$, existe $0 \neq \lambda \in R$ tal que $\lambda z = 0$. Por otro lado, como $z \in M$ y $\{x_1, ..., x_n\}$ es una base de M, entonces existen únicos $r_i \in R$ tales que $z = \sum_{i=1}^{n} r_ix_i$. Luego, $0 = \lambda z = \sum_{i=1}^{n} (\lambda r_i)x_i = 0$, de donde concluimos que $\lambda r_i = 0$ para todo i ya que $\{x_1, ..., x_n\}$ es linealmente independiente. Como R es un dominio entero $\lambda \neq 0$, concluimos que $r_i = 0$ para todo i y así $z = \sum_{i=1}^{n} r_ix_i = 0$. Esto demuestra que $N = 0$ y, en este caso, ya acabamos.
Supongamos que el teorema es cierto para todo submódulo N de M con $0 < \text{ran}(N) < m$. Sea N un submódulo de M con $\text{ran}(N) = m$. Por lo hecho, en párrafos anteriores, tenemos que existen $\gamma : M \longrightarrow R$, $y_1 \in N$, y $0 \neq a_1 \in R$ tal que

$$N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma))$$

con $\text{ran}(N \cap \text{Ker}(\gamma)) < \text{ran}(N) = m$. Luego, por hipótesis de inducción concluimos que $N \cap \text{Ker}(\gamma)$ es un R-módulo libre con base $\{c_1, ..., c_k\}$. Notemos que $\{a_1y_1, c_1, ..., c_k\}$ es una base de N. En efecto, tenemos que $\{c_1, ..., c_k\}$ es linealmente independiente y de manera similar a como se hizo en párrafos anteriores, tenemos que $\{a_1y_1, c_1, ..., c_k\}$ es linealmente independiente. Además del hecho que $N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma))$, se tiene que cualquier $c \in N$ se escribe como $c = \lambda_1 a_1 y_1 + (\sum_{i=1}^{k} \lambda_{i+1} c_i)$ con $\lambda_i \in R$ para todo i. Así, tenemos que N es libre con base $\{a_1y_1, c_1, ..., c_k\}$. Probándose por inducción el inciso (I).
Ahora probaremos (II) por inducción sobre $n = \text{ran}(M)$. Supongamos $n = \text{ran}(M) = 1$ y que N es un submódulo de M. Sabemos que existen $\gamma : M \longrightarrow R$, $y_1 \in N$, y $0 \neq a_1 \in R$ tales que $M = Ry_1 \oplus \text{Ker}(\gamma)$ y $N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma))$. Como $\text{Ker}(\varphi)$ es un submódulo de M, por el inciso (I), tenemos que $\text{Ker}(\varphi)$ es libre. Ahora, como $M = Ry_1 \oplus \text{Ker}(\gamma)$, vimos en párrafos anteriores que $\text{ran}(\text{Ker}(\gamma)) < \text{ran}(M) = 1$ y de esta manera $\text{ran}(\text{Ker}(\gamma)) = 0$ y así $\text{Ker}(\gamma) = 0$ y, por lo tanto, $\{y_1\}$ es una base de M. Ahora, $N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma)) = Ra_1y_1$ y, por lo tanto, $\{a_1y_1\}$ es una base de N. Probándose el resultado para $n = 1$.
Supongamos que el teorema es cierto para todo módulo libre M con $1 \leq \text{ran}(M) < m$. Sea M un R-módulo tal que $\text{ran}(M) = m$. Sea N un submódulo de M. Sabemos que existen $\gamma : M \longrightarrow R$, $y_1 \in N$, y $0 \neq a_1 \in R$ tales que $M = Ry_1 \oplus \text{Ker}(\gamma)$ y $N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma))$. Por el inciso (I), tenemos que $\text{Ker}(\gamma)$ es libre; y además $\text{ran}(\text{Ker}(\gamma)) = \text{ran}(M) - 1 = m - 1$. Como $N \cap \text{Ker}(\gamma)$ es un submódulo de $\text{Ker}(\gamma)$, por hipótesis de inducción existe una base $\{y_2, y_3, ..., y_m\}$ de $\text{Ker}(\gamma)$ y $a_2, ..., a_r$ con $r \leq m$ tal que $\{a_2y_2, ..., a_ry_r\}$ es una base de $N \cap \text{Ker}(\gamma)$ y $a_2 \mid a_3 \mid \cdots \mid a_m$.
Como $M = Ry_1 \oplus \text{Ker}(\gamma)$ y $N = Ra_1y_1 \oplus (N \cap \text{Ker}(\gamma))$, tenemos que $\{y_1, y_2, ..., y_m\}$ es una base de M y $\{a_1y_1, a_2y_2, ..., a_ry_r\}$ es una base de N. Solo falta ver que $a_1 \mid a_2$.
Como $\{y_1, y_2, ..., y_m\}$ es una base de M, por la Proposición 1.2.12, podemos definir un morfismo de R-módulos $\phi : M \longrightarrow R$ tal que $\phi(y_1) = \phi(y_2) = 1$ y $\phi(y_i) = 0$ para $i > 2$. Tenemos que $\phi(a_1y_1) = a_1 \in \phi(N)$. Por lo tanto, $\langle a_1 \rangle \subseteq \phi(N)$. Como a_1 se elige tal que $\langle a_1 \rangle = \gamma(N)$ es maximal en el conjunto Σ, concluimos que $\langle a_1 \rangle = \gamma(N) = \phi(N)$. Como $a_2 = a_2\phi(y_2) = \phi(a_2y_2) \in \phi(N)$ pues $a_2y_2 \in N$, concluimos que $a_2 \in \langle a_1 \rangle$ y por lo tanto $a_1 | a_2$. Probándose el resultado. □

Sea R un DIP, por la proposición 1.1.6, tenemos que p es primo si, y sólo si, p es irreducible. Además, p es irreducible si, y sólo si, $\langle p \rangle$ es un ideal maximal y en este caso $R/\langle p \rangle$ es un campo.

Lema 1.3.9. *Sean R un DIP y $p \in R$ primo. Sea F el campo $R/\langle p \rangle$.*

a) Si $M = R^r$, entonces $M/pM \cong F^r$ como F-módulos.

b) Si $M = R/\langle a \rangle$, donde $a \in R \setminus \{0\}$, entonces

$$M/pM \cong \begin{cases} F & si\ p \mid a\ en\ R \\ 0 & si\ p \nmid a\ en\ R \end{cases} \text{ como } F\text{-módulos.}$$

c) Si $M = R/\langle a_1 \rangle \oplus R/\langle a_2 \rangle \oplus \cdots \oplus R/\langle a_k \rangle$, donde $p \mid a_i$, $i = 1, ..., k$, entonces $M/pM \cong F^k$ como F-módulos.

Demostración. a) Definimos un morfismo de R-módulos como sigue,

$$\begin{array}{rcl} \phi : M & \longrightarrow & F^r \\ (x_1, ..., x_r) & \mapsto & (x_1 + \langle p\rangle, ..., x_r + \langle p\rangle). \end{array}$$

Claramente, ϕ es suprayectiva, de modo que el Primer Teorema de Isomorfismo nos asegura que

$$F^r \cong M/\mathrm{Ker}(\phi) \text{ como } R\text{-módulos.}$$

Ahora, si $(x_1, ..., x_r) \in \mathrm{Ker}(\phi)$, entonces $(x_1 + \langle p\rangle, ..., x_r + \langle p\rangle) = 0$, lo que implica que $x_i = pa_i$ para alguna $a_i \in R$, $i = 1, ..., r$. Por ello,

$$(x_1, ..., x_r) = p(a_1, ..., a_r) \in pM.$$

Inversamente, si $p(a_1, ..., a_r) \in pM$, entonces

$$\phi(p(a_1, ..., a_r)) = (pa_1 + \langle p\rangle, ..., pa_r + \langle p\rangle) = 0.$$

Por lo tanto, $\mathrm{Ker}(\phi) = pM$ y $M/pM \cong F^r$ como R-módulos. Ahora, notemos que M/pM tiene estructura de F-módulo con el siguiente producto escalar

$$\begin{array}{rcl} F \times M/pM & \longrightarrow & M/pM \\ (x + \langle p\rangle, m + pM) & \mapsto & (xm) + pM, \end{array}$$

pues $x - y = pt$ implica que $xm - ym = ptm \in pM$, i.e., el producto está bien definido. Recordemos que el isomorfismo dado por el Primer Teorema de Isomorfismo está dado por

$$\begin{array}{rcl} \bar{\phi} : M/pM & \longrightarrow & F^r \\ m + pM & \mapsto & \phi(m). \end{array}$$

De modo que

$$\bar{\phi}((x + \langle p\rangle)(m + pM)) = \bar{\phi}((xm) + pM) = \phi(xm).$$

Si $m = (a_1, ..., a_r)$, entonces $\phi(xm) = (xa_1 + \langle p\rangle, ..., xa_r + \langle p\rangle) = (x + \langle p\rangle)\phi(m)$ y, por lo tanto, $M/pM \cong F^r$ como F-módulos.

b) Supongamos que $p \mid a$ en R. Definimos ϕ como sigue,

$$\begin{array}{rcl} M & \longrightarrow & F \\ x + \langle a\rangle & \mapsto & x + \langle p\rangle. \end{array}$$

Notemos que si $x + \langle a\rangle = y + \langle a\rangle$, entonces $x - y = ab$, para alguna $b \in R$. Como $p \mid a$, se tiene que $a = pc$, para alguna $c \in R$. En consecuencia, $x - y = pcb \in \langle p\rangle$, i.e., $\phi(x + \langle a\rangle) = \phi(y + \langle a\rangle)$ y ϕ está bien definida. Igual que en el inciso a), ϕ es suprayectiva y, por ende,

$$F \cong M/\mathrm{Ker}(\phi).$$

Además,

$$\begin{aligned}\text{Ker}(\phi) &= \{x+\langle a\rangle | \phi(x+\langle a\rangle)=0\}\\ &= \{x+\langle a\rangle | x\in\langle p\rangle\}\\ &= \{p(y+\langle a\rangle)|y\in R\}\\ &= pM.\end{aligned}$$

Por lo tanto, $F\cong M/pM$. De la misma manera que en el inciso a), M/pM tiene una estructura natural de F-módulo y con ésta el isomorfismo de R-módulos $\bar{\phi}: M/pM\longrightarrow F$ es un isomorfismo de F-módulos.
Veamos que pasa si $p\nmid a$. En este caso, tenemos que p y a son coprimos, entonces existen $s,t\in R$ tales que $1=sp+ta$. Sea $x+\langle a\rangle\in M$ con $x\in R$, entonces $x=xsp+xta$, consecuentemente

$$x+\langle a\rangle=(xsp+xta)+\langle a\rangle=xsp+\langle a\rangle=p(xs+\langle a\rangle)\in pM.$$

Lo que nos permite concluir que $M=pM$, i.e., $M/pM\cong 0$.

c) Como $p\mid a_i$, para $i=1,...,k$, el inciso b) asegura que existe un F-isomorfismo $\phi_i: R/\langle a_i\rangle\Big/p\left(R/\langle a_i\rangle\right)\longrightarrow F$, $i=1,...,k$. Esto nos permite definir el siguiente morfismo,

$$\phi:\left(R/\langle a_1\rangle\Big/p\left(R/\langle a_1\rangle\right)\right)\oplus\left(R/\langle a_2\rangle\Big/p\left(R/\langle a_2\rangle\right)\right)\oplus\cdots\oplus\left(R/\langle a_k\rangle\Big/p\left(R/\langle a_k\rangle\right)\right)\longrightarrow F^k$$

$$\phi=(\phi_1,\phi_2,...,\phi_k).$$

Dado que ϕ se compone, coordenada a coordenada, de F-isomorfismos, se tiene que ϕ también es un F-isomorfismo. Para simplificar la notación definimos lo siguiente,

$$M':=\left(R/\langle a_1\rangle\Big/p\left(R/\langle a_1\rangle\right)\right)\oplus\left(R/\langle a_2\rangle\Big/p\left(R/\langle a_2\rangle\right)\right)\oplus\cdots\oplus\left(R/\langle a_k\rangle\Big/p\left(R/\langle a_k\rangle\right)\right).$$

Ahora vamos a definir el siguiente morfismo,

$$\begin{aligned}\psi: M &\longrightarrow M'\\ (x_1+\langle a_1\rangle,...,x_k+\langle a_k\rangle) &\mapsto \left((x_1+\langle a_1\rangle)+p\left(R/\langle a_1\rangle\right),...,(x_k+\langle a_k\rangle)+p\left(R/\langle a_k\rangle\right)\right).\end{aligned}$$

El morfismo así definido es claramente suprayectivo. El Primer Teorema de Isomorfismo asegura que $M'\cong M/\text{Ker}(\psi)$.
Si $(x_1+\langle a_1\rangle,...,x_k+\langle a_k\rangle)\in\text{Ker}(\psi)$, entonces $x_i+\langle a_i\rangle\in p\left(R/\langle a_i\rangle\right)$, para cada $i=1,...,k$. Lo que implica que $x_i+\langle a_i\rangle=p(y_i+\langle a_i\rangle)$, $i=1,...,k$. De modo que

$$(x_1+\langle a_1\rangle,...,x_k+\langle a_k\rangle)=p\left(y_1+\langle a_1\rangle,...,y_k+\langle a_k\rangle\right)\in pM.$$

Por otro lado, si $p\left(y_1+\langle a_1\rangle,...,y_k+\langle a_k\rangle\right)\in pM$, entonces para cada entrada se tiene que

$$p(y_i+\langle a_i\rangle)\in p\left(R/\langle a_i\rangle\right),\ i=1,...,k,$$

i.e., $\psi(p\left(y_1+\langle a_1\rangle,...,y_k+\langle a_k\rangle\right))=0$. Por lo tanto, $\text{Ker}(\psi)=pM$. Con un procedimiento análogo al del inciso a) se sigue que M/pM es un F-módulo y $F^k\cong M'\cong M/pM$ como F-módulos.

□

Teorema 1.3.10. *Sean R un DIP, $p \in R$ primo y $0 \neq M$ un R-módulo finitamente generado. Supongamos que M es un p-módulo (ver definición 1.3.5). Entonces existen únicos $\alpha_1, ..., \alpha_k \in \mathbb{N} \setminus \{0\}$ tales que $p^{\alpha_1} \mid p^{\alpha_2} \mid \cdots \mid p^{\alpha_k}$ (llamados divisores de M) y*

$$M \cong \bigoplus_{i=1}^{k} R/\langle p^{\alpha_i} \rangle.$$

Además, $\mathrm{Ann}(R/\langle p^{\alpha_i} \rangle) = \langle p^{\alpha_i} \rangle$ *para todo i.*

Demostración. Dado que M es finitamente generado, por el corolario 1.2.13, existe $\phi : R^n \longrightarrow M$ morfismo de R-módulos suprayectivo; lo que a su vez implica que $M \cong R^n/\mathrm{Ker}(\phi)$. Por el teorema 1.3.8, sabemos que existe $\{y_1, ..., y_n\}$ base de R^n y $a_1, ..., a_m \in R \setminus \{0\}$ tales que $a_1 y_1, ..., a_m y_m$ forman una base para $\mathrm{Ker}(\phi)$ y $a_1 \mid a_2 \mid \cdots \mid a_m$. De donde

$$M \cong \bigoplus_{i=1}^{n} Ry_i \Big/ \bigoplus_{i=1}^{m} Ra_i y_i$$

Se puede ver que la siguiente función

$$\varphi : \bigoplus_{i=1}^{n} Ry_i \longrightarrow \left(\bigoplus_{i=1}^{m} R/\langle a_i \rangle \right) \bigoplus R^{n-m}$$

dada por

$$\varphi\Big(r_1 y_1, ..., r_m y_m, r_{m+1} y_{m+1}, ..., r_n y_n\Big) = \Big(r_1 + \langle a_1 \rangle, ..., r_m + \langle a_m \rangle, r_{m+1}, ..., r_n\Big)$$

es un morfismo suprayectivo de R-módulos con $\mathrm{Ker}(\varphi) = \bigoplus_{i=1}^{m} Ra_i y_i$, de donde por el Primer teorema de Isomorfismo, tenemos que

$$\bigoplus_{i=1}^{n} Ry_i \Big/ \bigoplus_{i=1}^{m} Ra_i y_i \cong \left(\bigoplus_{i=1}^{m} R/\langle a_i \rangle \right) \bigoplus R^{n-m},$$

de tal manera que

$$M \cong \bigoplus_{i=1}^{n} Ry_i \Big/ \bigoplus_{i=1}^{m} Ra_i y_i \cong \left(\bigoplus_{i=1}^{m} R/\langle a_i \rangle \right) \bigoplus R^{n-m}.$$

Pero M es de torsión (por ser un p-módulo) y $rr' = 0$ implica que $r = 0$ o $r' = 0$, pues R es un DIP, entonces $n = m$ y

$$M \cong \bigoplus_{i=1}^{n} R/\langle a_i \rangle.$$

En lo anterior existe la posibilidad de que algún a_i sea una unidad; en este caso, $R/\langle a_i \rangle = 0$. Notemos que si a_i es unidad, entonces a_j es unidad para todo $j = 1, ..., i$ pues $a_j \mid a_i$ si $j = 1, ..., i$. Entonces, en la expresión anterior, podemos quitar tales términos y, dado que $M \neq 0$, podemos suponer que $M \cong \bigoplus_{i=1}^{k} R/\langle a_i \rangle$ con $1 \leq k \leq n$ y a_i no unidad para todo $i = 1, ..., k$ tal que $a_1 \mid a_2 \mid \cdots \mid a_{k-1} \mid a_k$. Como M es p-módulo finitamente generado,

tenemos que $\text{Ann}(M) = \langle p^\alpha \rangle$ para algún $\alpha \in \mathbb{N} \setminus \{0\}$. Para $1 \leq i \leq k$, consideremos $u_i := \Big(0 + \langle a_1 \rangle, ..., 1 + \langle a_i \rangle, ..., 0 + \langle a_k \rangle\Big) \in \bigoplus_{i=1}^k R/\langle a_i \rangle$. Como M es un p-módulo, existe $p^\beta \in R$ con $\beta \in \mathbb{N} \setminus \{0\}$ tal que $0 = p^\beta u_i$. Es decir, $p^\beta(1 + \langle a_i \rangle) = p^\beta + \langle a_i \rangle = 0 + \langle a_i \rangle$ y así $p^\beta \in \langle a_i \rangle$. Por lo tanto, $a_i \mid p^\beta$ y como R es un dominio de factorización única y a_i no es unidad, tenemos que $a_i = p^{\alpha_i}$ para algún $\alpha_i \in \mathbb{N} \setminus \{0\}$. Probándose que

$$M \cong \bigoplus_{i=1}^{k} R/\langle p^{\alpha_i} \rangle$$

con $p^{\alpha_1} \mid p^{\alpha_2} \mid \cdots \mid p^{\alpha_k}$.
Veamos que $\text{Ann}\Big(R/\langle p^{\alpha_i} \rangle\Big) = \langle p^{\alpha_i} \rangle$ para todo i. En efecto, como R es un DIP tenemos que $\text{Ann}\Big(R/\langle p^{\alpha_i} \rangle\Big) = \langle h \rangle$ para algún $h \in R$. Como $\langle p^{\alpha_i} \rangle \Big(R/\langle p^{\alpha_i} \rangle\Big) = 0$, tenemos que $p^{\alpha_i} \in \text{Ann}\Big(R/\langle p^{\alpha_i} \rangle\Big) = \langle h \rangle$ y así $h \mid p^{\alpha_i}$. Como R es un DFU, tenemos que $h = p^\alpha$ con $\alpha \leq \alpha_i$. Dado que $\langle p^{\alpha_i} \rangle = 0_{R/\langle p^{\alpha_i} \rangle} = p^\alpha \Big(1 + \langle p^{\alpha_i} \rangle\Big) = p^\alpha + \langle p^{\alpha_i} \rangle$, se tiene que $p^\alpha \in \langle p^{\alpha_i} \rangle$. Por lo tanto, $p^{\alpha_i} \mid p^\alpha$ y entonces $\alpha_i \leq \alpha$. Por lo tanto, $h = p^{\alpha_i}$.
Ahora veamos la unicidad. Supongamos que

$$M \cong \bigoplus_{i=1}^{k} R/\langle p^{\alpha_i} \rangle \cong \bigoplus_{i=1}^{r} R/\langle p^{\beta_i} \rangle,$$

donde $p^{\beta_1} \mid p^{\beta_2} \mid \cdots \mid p^{\beta_r}$. Probemos que $k = r$ y que $(p^{\alpha_1}, ..., p^{\alpha_k}) = (p^{\beta_1}, ..., p^{\beta_r})$. Calculando el anulador, tenemos que $\text{Ann}(M) = \langle p^{\alpha_k} \rangle = \langle p^{\beta_r} \rangle$, de donde concluimos que $\alpha_k = \beta_r$.
La demostración la haremos por inducción sobre α_k. Si $\alpha_k = 1$, entonces $\alpha_i = \beta_j = 1$, $i = 1, ..., k$ y $j = 1, ..., r$. Lo que implica que

$$F^k = \bigoplus_{i=1}^{k} R/\langle p \rangle \cong \bigoplus_{i=1}^{r} R/\langle p \rangle = F^r,$$

donde $F = R/\langle p \rangle$. Sea $\phi : F^k \longrightarrow F^r$ un R-isomorfismo. Entonces

$$\begin{aligned} \phi((x + \langle p \rangle)(y_1 + \langle p \rangle, ..., y_k + \langle p \rangle)) &= \phi(xy_1 + \langle p \rangle, ..., xy_k + \langle p \rangle) \\ &= \phi(x(y_1 + \langle p \rangle, ..., y_k + \langle p \rangle)) \\ &= x\phi(y_1 + \langle p \rangle, ..., y_k + \langle p \rangle). \end{aligned}$$

Si $\phi(y_1 + \langle p \rangle, ..., y_k + \langle p \rangle) = (a_1 + \langle p \rangle, ..., a_r + \langle p \rangle)$, se obtiene que

$$\begin{aligned} x(a_1 + \langle p \rangle, ..., a_r + \langle p \rangle) &= (xa_1 + \langle p \rangle, ..., xa_r + \langle p \rangle) \\ &= ((x + \langle p \rangle)(a_1 + \langle p \rangle), ..., (x + \langle p \rangle)(a_r + \langle p \rangle)) \\ &= (x + \langle p \rangle)(a_1 + \langle p \rangle, ..., a_r + \langle p \rangle) \\ &= (x + \langle p \rangle)\phi(y_1 + \langle p \rangle, ..., y_k + \langle p \rangle). \end{aligned}$$

Es decir, ϕ es un F-isomorfismo y, puesto que F es un campo, se concluye que $k = r$ (pues $F^k \cong F^r$ como espacios vectoriales).
Supongamos que el resultado es cierto para cualquier p-módulo N con $\text{Ann}(N) = \langle p^\gamma \rangle$ con $1 \leq \gamma < m$.
Sea M un p-módulo tal que

$$M \cong M_1 \cong M_2$$

$$M_1 := \bigoplus_{i=1}^{k} R/\langle p^{\alpha_i} \rangle, \qquad M_2 = \bigoplus_{i=1}^{r} R/\langle p^{\beta_i} \rangle.$$

con $\text{Ann}(M_1) = \langle p^{\alpha_k} \rangle = \langle p^m \rangle = \langle p^{\beta_r} \rangle = \text{Ann}(M_2)$. Supongamos que los divisores de M_1 y M_2 son respectivamente:

$$(p^{\alpha_1}, p^{\alpha_2}, ..., p^{\alpha_k}) = (\underbrace{p, p, ..., p}_{a-veces}, p^{\alpha_{a+1}}, ..., p^{\alpha_k})$$
$$(p^{\beta_1}, p^{\beta_2}, ..., p^{\beta_r}) = (\underbrace{p, p, ..., p}_{b-veces}, p^{\beta_{b+1}}, ..., p^{\beta_r})$$

con $2 \leq \alpha_{a+1} \leq \cdots \leq \alpha_k$ y $2 \leq \beta_{b+1} \leq \cdots \leq \beta_r$. Notemos que

$$\langle p \rangle \cdot M_1 = \bigoplus_{i=1}^{k} \frac{\langle p \rangle \cdot R}{\langle p^{\alpha_i} \rangle}.$$

Como el morfismo $\varphi : R \longrightarrow \frac{\langle p \rangle \cdot R}{\langle p^{\alpha_i} \rangle}$ dado por $\varphi(r) = pr + \langle p^{\alpha_i} \rangle$ es tal que $\text{Ker}(\varphi) = \langle p^{\alpha_i - 1} \rangle$, tenemos que

$$\langle p \rangle \cdot M_1 = \bigoplus_{i=1}^{k} \frac{\langle p \rangle \cdot R}{\langle p^{\alpha_i} \rangle} \cong \bigoplus_{i=1}^{k} \frac{R}{\langle p^{\alpha_i - 1} \rangle} \cong \bigoplus_{i=a+1}^{k} \frac{R}{\langle p^{\alpha_i - 1} \rangle}.$$

Luego, tenemos el siguiente arreglo:

$$(\underbrace{1, 1, ..., 1}_{a-veces}, p^{\alpha_{a+1}-1}, ..., p^{\alpha_k - 1}).$$

Es decir, $\langle p \rangle \cdot M_1$ tiene divisores asociados $(p^{\alpha_{a+1}-1}, ..., p^{\alpha_k - 1})$ (hemos quitado los unos). Además, $\text{Ann}(\langle p \rangle \cdot M_1) = \langle p^{\alpha_k - 1} \rangle$.
De manera similar,

$$\langle p \rangle \cdot M_2 = \bigoplus_{i=1}^{r} \frac{\langle p \rangle \cdot R}{\langle p^{\beta_i} \rangle} \cong \bigoplus_{i=1}^{r} \frac{R}{\langle p^{\beta_i - 1} \rangle} \cong \bigoplus_{i=b+1}^{r} \frac{R}{\langle p^{\beta_i - 1} \rangle}.$$

Luego, tenemos el siguiente arreglo:

$$(\underbrace{1, 1, ..., 1}_{b-veces}, p^{\beta_{b+1}-1}, ..., p^{\beta_r - 1}).$$

Es decir, $\langle p \rangle \cdot M_2$ tiene divisores asociados $(p^{\beta_{b+1}-1}, ..., p^{\beta_r - 1})$ (hemos quitado los unos). Además, $\text{Ann}(\langle p \rangle \cdot M_2) = \langle p^{\beta_r - 1} \rangle$. Notemos que tenemos el isomorfismo

$$\langle p \rangle \cdot M_1 \cong \langle p \rangle \cdot M_2$$

con $\text{Ann}(\langle p\rangle \cdot M_2) = \langle p^{\beta_r - 1}\rangle = \langle p^{\alpha_k - 1}\rangle = \text{Ann}(\langle p\rangle \cdot M_1)$, pues $\alpha_k = \beta_r$. Luego, por hipótesis de inducción, concluimos que $k - a = r - b$ y $\alpha_{a+1} = \beta_{b+1}, ..., \alpha_k = \beta_r$.
Finalmente, consideremos el campo $R/\langle p\rangle = F$ (esto pasa pues p es primo en un D.F.U). Ahora, notemos que por el lema 1.3.9, se tiene que

$$F^k \cong \frac{M_1}{\langle p\rangle \cdot M_1} \cong \frac{M_2}{\langle p\rangle \cdot M_2} \cong F^r.$$

Por lo tanto, $k = r$ y como $k - a = r - b$, tenemos que $a = b$ y de esta manera concluimos que

$$\begin{aligned}(p^{\alpha_1}, p^{\alpha_2}, ..., p^{\alpha_k}) &= (\underbrace{p, p, ..., p}_{a-veces}, p^{\alpha_{a+1}}, ..., p^{\alpha_k}) \\ &= (\underbrace{p, p, ..., p}_{b-veces}, p^{\beta_{b+1}}, ..., p^{\beta_r}) \\ &= (p^{\beta_1}, p^{\beta_2}, ..., p^{\beta_r}).\end{aligned}$$

Probándose que el resultado es cierto para M. □

Teorema 1.3.11. *Sean R un dominio de ideales principales y $0 \neq M$ un R-módulo finitamente generado y de torsión. Entonces existen primos únicos $p_1, ..., p_n$ en R, no asociados entre sí, tales que*

$$M \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}}\rangle,$$

y para cada i se tiene que $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$. Además, tal decomposición es única, en el sentido de que si

$$M \cong \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}}\rangle$$

es otra descomposición de M, donde los $q_1, ..., q_m$ son primos de R no asociados entre sí y tal que para cada i se tiene que $\beta_{i,1} \leq \beta_{i,2} \leq \cdots \leq \beta_{i,l_i}$, entonces $n = m$ y existe un reordenamiento de los q_i tal que $p_i = q_i$ para todo $i = 1, ..., n$ que cumple que $l_i = k_i$ y para cada i se tiene que $\alpha_{i,j} = \beta_{i,j}$ $\forall 1 \leq j \leq k_i$.

Demostración. La existencia de tal descomposición se sigue de los teoremas 1.3.6 y 1.3.10, donde, por construcción, tenemos que $M = \bigoplus_{i=1}^{n} M_{p_i}$ con $M_{p_i} \cong \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}}\rangle$. Ahora supongamos que existe un isomorfismo $\varphi : M \longrightarrow N$ con

$$N = \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}}\rangle,$$

donde los $q_1, ..., q_m$ son primos de R no asociados entre sí y tales que para cada i se tiene que $\beta_{i,1} \leq \beta_{i,2} \leq \cdots \leq \beta_{i,l_i}$. Luego, tenemos que $N_{q_i} = \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}}\rangle$ y $N = \bigoplus_{i=1}^{m} N_{q_i}$. Vía el isomorfismo φ, la descomposición de N induce una descomposición de M de la forma $M = \bigoplus_{i=1}^{m} M_{q_i}$. Luego, por la unicidad del teorema 1.3.6, concluimos que $n = m$ y

que $p_i = q_i$ para todo $i = 1, ..., n$.
Entonces para cada $i = 1, ..., n$, tenemos que

$$\bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}} \rangle \cong M_{p_i} \cong N_{p_i} \cong \bigoplus_{j=1}^{l_i} R/\langle p_i^{\beta_{i,j}} \rangle.$$

Luego, por teorema 1.3.10, tenemos que $l_i = k_i$ y $\alpha_{i,j} = \beta_{i,j}$ $\forall 1 \leq j \leq k_i$. Probándose la unicidad. □

La demostración del siguiente resultado básicamente se sigue de la unicidad de la expresión dada en el teorema 1.3.11, pero escribiremos una prueba detallada para beneficio del lector.

Proposición 1.3.12. *Sean R un dominio de ideales principales y M, N, N' tres R-módulos finitamente generados y de torsión. Si $M \oplus N \cong M \oplus N'$, entonces $N \cong N'$.*

Demostración. Podemos asumir que $M \neq 0$; y en tal caso, sabemos que

$$M \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}} \rangle,$$

donde la descomposición está dada como en el teorema 1.3.11. Ahora, tal descomposición es única salvo el orden de los primos y entonces el número de sumandos en tal descomposición es único. Para probar el resultado, procedamos por inducción sobre el número de sumandos de tal descomposición de M, el cual denotaremos por $\omega(M)$.
Supongamos que $\omega(M) = 1$, es decir, $M \cong R/\langle p^\beta \rangle$ para algún primo $p \in R$ y β entero positivo. Sean N y N' módulos finitamente generados y de torsión tales que $M \oplus N \cong M \oplus N'$. Supongamos que $N \cong 0$ ó $N' \cong 0$. Podemos asumir sin pérdida de generalidad que $N \cong 0$. En este caso, se tiene que

$$R/\langle p^\beta \rangle \cong R/\langle p^\beta \rangle \oplus N'.$$

De la unicidad del teorema 1.3.11 concluimos que $N' \cong 0$.
Lo anterior nos permite suponer que $N \ncong 0$ y $N' \ncong 0$, entonces N y N' tienen las siguientes descomposiciones,

$$N \cong \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}} \rangle,$$

$$N' \cong \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{t_i} R/\langle w_i^{\gamma_{i,j}} \rangle,$$

como en el teorema 1.3.11. Tenemos 4 casos:

a) El primo p no es asociado a ningún q_i de la descomposición de N; y p no es asociado a ningún w_i de la descomposición de N'.
En este caso, tenemos que

$$\bigoplus_{i=1}^{m+1} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}} \rangle \cong N \oplus M \cong N' \oplus M \cong \bigoplus_{i=1}^{r+1} \bigoplus_{j=1}^{t_i} R/\langle w_i^{\gamma_{i,j}} \rangle,$$

donde $q_{m+1} = p = w_{r+1}$, $l_{m+1} = t_{r+1} = 1$, $\beta_{m+1,1} = \gamma_{r+1,1} = \beta$ con $\beta_{i,1} \leq \cdots \leq \beta_{i,l_i}$ y $\gamma_{i,1} \leq \cdots \leq \gamma_{i,t_i}$ para toda i. Es decir, tenemos dos descomposiciones como en el teorema 1.3.11. De esta manera, por la unicidad dada en el teorema 1.3.11, tenemos que $m+1 = r+1$ y por lo tanto $m = r$; y después de reordenar los primos podemos suponer que $q_i = w_i$, $l_i = t_i$ para todo i y que $\beta_{i,j} = \gamma_{i,j}$ para todo i, j. Así, concluimos que $N \cong N'$.

b) El primo p es asociado algún q_i de la descomposición de N; y p no es asociado a ningún w_i de la descomposición de N'.
Supongamos, sin perdida de generalidad, que $p = q_1$. Ahora, como $\beta_{1,1} \leq \beta_{1,2} \leq \cdots \leq \beta_{1,l_1}$ con $l_1 \geq 1$, podemos insertar β en tal cadena para formar una cadena de longitud más grande (de longitud al menos 2)

$$\beta'_{1,1} \leq \beta'_{1,2} \leq \cdots \leq \beta'_{1,l_1+1}$$

tal que existe h con $1 \leq h \leq l_1 + 1$ que cumple que $\beta'_{1,h} = \beta$. De esta manera, tenemos la siguiente descomposición de $M \oplus N$:

$$M \oplus N \cong \bigoplus_{j=1}^{l_1+1} R/\langle q_1^{\beta'_{1,j}} \rangle \bigoplus \Big(\bigoplus_{i=2}^{m} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}} \rangle \Big)$$

como en el teorema 1.3.11. Por otro lado, tenemos una descomposición de $M \oplus N'$ como sigue:

$$M \oplus N' \cong \bigoplus_{i=1}^{r+1} \bigoplus_{j=1}^{t_i} R/\langle w_i^{\gamma_{i,j}} \rangle,$$

donde $p = w_{r+1}$, $t_{r+1} = 1$ y $\gamma_{r+1,1} = \beta$ y por lo tanto, tal descomposición es como la del teorema 1.3.11. Como $p = q_1 = w_{r+1}$ y $M \oplus N \cong M \oplus N'$, después de reordenar los sumandos, concluimos del teorema 1.3.11, que $l_1 + 1 = 1$ y así $l_1 = 0$. Lo cual es una contradicción y, por lo tanto, este caso no existe.

c) El primo p no es asociado a ningún q_i de la descomposición de N; y p es asociado a algún w_i de la descomposición de N'. Este caso no es posible de manera similar a (b).

d) El primo p es asociado algún q_i de la descomposición de N y p es asociado a algún w_i de la descomposición de N'.
Sin perdida de generalidad, suponemos que $p = q_1 = w_1$. Ahora, como $\beta_{1,1} \leq \beta_{1,2} \leq \cdots \leq \beta_{1,l_1}$ con $l_1 \geq 1$, podemos insertar β en tal cadena para formar una cadena de longitud más grande (de longitud al menos 2):

$$\beta'_{1,1} \leq \beta'_{1,2} \leq \cdots \leq \beta'_{1,l_1+1}.$$

De la misma manera, como $\gamma_{1,1} \leq \gamma_{1,2} \leq \cdots \leq \gamma_{1,t_1}$ con $t_1 \geq 1$, podemos insertar β en tal cadena para formar una cadena de longitud más grande (de longitud al menos 2):

$$\gamma'_{1,1} \leq \gamma'_{1,2} \leq \cdots \leq \gamma'_{1,t_1+1}.$$

De esta manera, tenemos las siguientes descomposiciones

$$M \oplus N \cong \bigoplus_{j=1}^{l_1+1} R/\langle q_1^{\beta'_{1,j}} \rangle \bigoplus \Big(\bigoplus_{i=2}^{m} \bigoplus_{j=1}^{l_i} R/\langle q_i^{\beta_{i,j}} \rangle \Big)$$

$$M \oplus N' \cong \bigoplus_{j=1}^{t_1+1} R/\langle q_1^{\gamma'_{1,j}} \rangle \bigoplus \Big(\bigoplus_{i=2}^{r} \bigoplus_{j=1}^{t_i} R/\langle w_i^{\gamma_{i,j}} \rangle \Big)$$

como en el teorema 1.3.11. Por el teorema 1.3.11, tenemos que $m = r$, $l_1+1 = t_1+1$, $\beta'_{1,j} = \gamma'_{1,j}$ para todo $1 \leq j \leq t_1 + 1 = l_1 + 1$; y después de reordenar los sumandos tenemos que $q_i = w_i$, $l_i = t_i$ para todo $i = 2, ..., m$ y $\beta_{i,j} = \gamma_{i,j}$ para todo i, j. De esta manera tenemos que $N \cong N'$.

Esto concluye el caso en el que $\omega(M) = 1$.
Ahora, para el caso general, consideremos un módulo M que tiene descomposición como en el teorema 1.3.11 con $\omega(M) = a > 1$:

$$M \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}} \rangle.$$

Tenemos que $M \cong \Big(\bigoplus_{i=1}^{n-1} \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}} \rangle \Big) \oplus \Big(\bigoplus_{j=1}^{k_n-1} R/\langle p_n^{\alpha_{n,j}} \rangle \Big) \bigoplus R/\langle p_n^{n,k_n} \rangle$. Tomando $M' := \Big(\bigoplus_{i=1}^{n-1} \bigoplus_{j=1}^{k_i} R/\langle p_i^{\alpha_{i,j}} \rangle \Big) \oplus \Big(\bigoplus_{j=1}^{k_n-1} R/\langle p_n^{\alpha_{n,j}} \rangle \Big)$, tenemos que

$$M \cong M' \bigoplus R/\langle p_n^{n,k_n} \rangle,$$

donde $\omega(M') = a - 1$. Sean N y N' dos R-módulos finitamente generados y de torsión tales que $M \oplus N \cong M \oplus N'$. Luego, tenemos que

$$M' \bigoplus R/\langle p_n^{n,k_n} \rangle \bigoplus N \cong M' \bigoplus R/\langle p_n^{n,k_n} \rangle \bigoplus N'. \tag{1.2}$$

Por el caso $\omega(M) = 1$, podemos cancelar $R/\langle p_n^{n,k_n} \rangle$ de (1.2) para obtener $M' \bigoplus N \cong M' \bigoplus N'$, donde $\omega(M') = a - 1$ y así, por hipótesis de inducción, podemos cancelar M' para obtener que $N \cong N'$. □

1.4. Existencia de la forma canónica de Jordan

Para probar la existencia de la forma canónica de Jordan, empezaremos dándole estructura de $F[x]$-módulo a un espacio vectorial de dimensión finita.

Definición 1.4.1. *Dado un espacio vectorial V de dimensión finita sobre un campo F y una transformación lineal $f : V \longrightarrow V$, definimos la siguiente función $\cdot_f : F[x] \times V \longrightarrow V$ como*

$$p(x) \cdot_f v = p(f)(v), \text{ para } v \in V,$$

donde $p(f) : V \longrightarrow V$ es la transformación lineal que está definida como

$$p(f) := a_0 Id_V + a_1 f + a_2 f^2 + \cdots + a_n f^n,$$

con $p(x) = a_0 + a_1 x + \cdots + a_n x^n$.

Es claro que V ya es un grupo abeliano, además, la función definida arriba satisface las condiciones necesarias para ser un producto por escalares, pues $1 \cdot_f v = Id(v) = v$. Las propiedades distributivas y la asociatividad del producto se siguen de la linealidad de las transformaciones y del hecho de que $(p(x)q(x))(f) = p(f) \circ q(f)$.

Definición 1.4.2. *V con el producto mencionado anteriormente es un $F[x]$-módulo que denotaremos V_f y, en caso de que no haya ambigüedad, escribiremos $p(x) \cdot_f v = p(x)v$ o, simplemente, pv.*

Definición 1.4.3. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Definimos $F[f] := \{p(f) \mid p(x) \in F[x]\}$. El conjunto $F[f]$ es un subespacio vectorial de $\mathrm{End}_F(V)$, donde $\mathrm{End}_F(V)$ denota el conjunto de todas las transformaciones lineales de V en V.*

Lo siguiente será probar que V_f es un módulo de torsión, para ello enunciaremos el Teorema de Cayley-Hamilton.

Teorema 1.4.4 (Cayley-Hamilton)**.** *Sean $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita y $p(x)$ su polinomio característico, entonces $p(f) = 0$.*

Demostración. Véase [4], p. 315. □

Proposición 1.4.5. *Sea $V \neq \{0\}$ un F-espacio vectorial de dimensión finita, entonces V_f es un $F[x]$-módulo de torsión.*

Demostración. Dado $v \in V_f$, por el teorema de Cayley-Hamilton, $0 = p(f)(v) = p \cdot_f v$ y $p(x) \neq 0$, pues $\mathrm{grad}(p(x)) = dim(V) \geq 1$. □

Definición 1.4.6. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Definimos el **polinomio mínimo de f** como el único polinomio mónico $g \in F[x]$ tal que $\mathrm{Ann}(V_f) = \langle g \rangle$.*

La prueba de la existencia de la forma canónica de Jordan recae en tres resultados, dos de los cuales hablan de la estructura de los módulos finitamente generados sobre dominios de ideales principales. El primero es el teorema 1.3.6 que afirma que es posible descomponer un módulo de torsión con las características anteriores en una suma directa de p-módulos, con p primo; el segundo es el teorema 1.3.10 que asegura la descomposición de los p-módulos en suma directa de módulos cíclicos. Finalmente, el tercer resultado nos proporciona una representación matricial de transformaciones lineales sobre V_f para el caso particular en el que V_f es cíclico (ver teorema 1.4.9).
Este último resultado es el que se verá en esta sección.

Teorema 1.4.7. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal tal que V_f es cíclico con generador c. Consideremos el F-espacio vectorial de la definición 1.4.3, entonces $\nu_c : F[f] \longrightarrow V$ definida como $\nu_c(p) = p(f)(c)$ es un isomorfismo de F-espacios vectoriales. Además, si*

$$g = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n$$

es el polinomio mínimo de f, entonces $\text{grad}(g) = \dim(V)$ *y* $\{c, f(c), ..., f^{n-1}(c)\}$ *es una base para V. La representación matricial de f con respecto a la base anterior está dada por*

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Demostración. Claramente, la función ν_c es lineal, ya que el producto por escalares en V_f es distributivo y $\nu_c(\lambda p) = (\lambda p)(f)(c) = \lambda(p(f)(c))$.
Sabemos que V_f es cíclico y está generado por c, por ello $V_f = \{p \cdot_f c : p \in F[x]\} = \{p(f)(c) : p \in F[x]\}$ y, así, ν_c es suprayectiva. Sea $p(f) \in F[f]$ tal que $\nu_c(p(f)) = p(f)(c) = 0$. Entonces, dado $v = q(f)(c) \in V$, se tiene que $p(f)(v) = p \cdot_f v = p \cdot_f (q \cdot_f c) = (pq) \cdot_f c = (qp) \cdot_f c = q \cdot_f (p \cdot_f c) = 0$. Por ende, $p(f) = 0$ y ν_c es inyectiva, esto es, ν_c es un isomorfismo.
El conjunto $\{c, f(c), ..., f^{n-1}(c)\}$ genera a V como F-espacio vectorial. En efecto, cualquier elemento $v \in V$ se puede expresar como

$$v = p(f)(c) = \sum_{i=0}^{k} \alpha_i f^i(c),$$

ya que V_f es cíclico, y como $0 = g(f)(c) = a_0 c + a_1 f(c) + \cdots + a_{n-1} f^{n-1}(c) + f^n(c)$, se tiene que $\langle c, f(c), ..., f^{n-1}(c)\rangle = \langle f^i(c) : i \in \mathbb{N} \rangle$. Supongamos que $b_0 c + b_1 f(c) + \cdots + b_{n-1} f^{n-1}(c) = 0 = p(f)(c)$, con $b_i \in F$ no todos cero. Entonces el polinomio $p(x) = \sum_{i=0}^{n-1} b_i x_i \in F[x]$ es no nulo de, a lo más, grado $n-1$ que anula al generador c y, como se vio en la prueba de la inyectividad de ν_c, se sigue que $p(f) = 0$. De la definición de polinomio mínimo se obtiene que $p \in \text{Ann}(V_f) = \langle g \rangle$, i.e., $g \mid p$, una contradicción, pues $\text{grad}(g) = n$. Por ello, $\{c, f(c), ..., f^{n-1}(c)\}$ es una base para V y $\dim(V) = n = \text{grad}(g)$. Que la representación matricial de f en la base $\{c, f(c), ..., f^{n-1}(c)\}$ sea la que está dada en el enunciado del teorema 1.4.7, se sigue directamente de la naturaleza de la base y de la identidad

$$0 = g(f)(c) = a_0 c + a_1 f(c) + \cdots + a_{n-1} f^{n-1}(c) + f^n(c).$$

□

Definición 1.4.8. *Sea $g = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n \in F[x]$ un polinomio mónico. La siguiente matriz*

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

se llama la matriz ***compañera de g****.*

Teorema 1.4.9. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal tal que V_f es cíclico generado por c. Si el polinomio mínimo de f es g^r, con*

$$g(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n,$$

entonces los siguientes elementos

$$\begin{array}{cccc} c & f(c) & \cdots & f^{n-1}(c) \\ g(f)(c) & g(f)(f(c)) & \cdots & g(f)(f^{n-1}(c)) \\ \vdots & \vdots & \ddots & \vdots \\ g(f)^{r-1}(c) & g(f)^{r-1}(f(c)) & \cdots & g(f)^{r-1}(f^{n-1}(c)) \end{array}$$

forman una base de V y la representación matricial de f en ésta es la matriz de $rn \times rn$

$$\begin{pmatrix} [g] & & & & \\ A & [g] & & & \\ & A & [g] & & \\ & & & \ddots & \ddots \\ & & & & A & [g] \end{pmatrix}$$

donde

$$[g] = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \quad y \quad A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

son matrices de tamaño $n \times n$.

Demostración. Por el teorema 1.4.7, sabemos que $\dim(V) = \text{grad}(g^r) = rn$. Para $0 \leq m \leq r-1$, consideremos el polinomio

$$g^m = b_0 + b_1x + \cdots + b_{mn-1}x^{mn-1} + x^{mn} \in F[x],$$

(en realidad los coeficientes de g^m dependen de los coeficientes $a_1, ..., a_{n-1}$ y m, pero como no queremos sobrecargar la notación, lo dejaremos así, sobreentendiendo eso). Entonces, tenemos la siguiente transformación lineal:

$$g(f)^m = g^m(f) = b_0 \cdot \text{Id} + b_1f + \cdots + b_{mn-1}f^{mn-1} + f^{mn} : V \longrightarrow V.$$

Ahora, para $0 \leq j \leq n-1$, tenemos que

$$g^m(f)(f^j(c)) = b_0f^j(c) + b_1f^{j+1}(c) + \cdots + b_{mn-1}f^{mn-1+j}(c) + f^{mn+j}(c).$$

Notemos que al considerar $g^m(f^j(c))$ con $0 \leq m \leq r-1$ y $0 \leq j \leq n-1$, el sumando de grado mayor de la forma $bf^l(c)$ que obtenemos, es cuando $m = r-1$ y $j = n-1$, es decir,

$f^{(r-1)n+(n-1)}(c) = f^{rn-1}(c)$, que proviene del término $g^{r-1}(f)(f^{n-1}(c))$. Sea $T = g(f)$, al considerar una combinación lineal que nos de cero de los elementos de

$$A := \left\{ g^m(f^j(c)) : \ 0 \leq m \leq r-1, 0 \leq j \leq n-1 \right\}$$

obtenemos la ecuación:

$$\sum_{i=0}^{n-1} B_{0,i} \cdot f^i(c) + \sum_{i=0}^{n-1} B_{1,i} \cdot T(f^i(c)) + \cdots + \sum_{i=0}^{n-1} B_{r-1,i} \cdot T^{r-1}(f^i(c)) = 0, \qquad (1.3)$$

con $B_{j,i} \in F$ para todo j, i. Después de agrupar todos los términos de la misma forma $f^l(c)$ con $l \leq rn-1$, obtenemos un polinomio $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{rn-1}x^{rn-1} \in F[x]$ con $\alpha_{rn-1} = B_{r-1,n-1}$ tal que

$$0 = p(f)(c) = \alpha_0 + \alpha_1 f(c) + \cdots + \alpha_{rn-2}f^{rn-2}(c) + \alpha_{rn-1}f^{rn-1}(c).$$

Al igual que en la prueba del teorema 1.4.7, el hecho de que $p(f)(c) = 0$ implica que $p(f) = 0 \in F[f]$, de donde $p \in \mathrm{Ann}(V_f) = \langle g^r \rangle$, i.e, $g^r \mid p$. Como $\mathrm{grad}(p) = rn - 1 < rn = \mathrm{grad}(g^r)$, concluimos que $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{rn-1}x^{rn-1} = 0$ y así, $\alpha_{rn-1} = B_{r-1,n-1} = 0$.
Luego, la ecuación (1.3) se convierte en la siguiente suma (con un sumando menos):

$$\sum_{i=0}^{n-1} B_{0,i} \cdot f^i(c) + \sum_{i=0}^{n-1} B_{1,i} \cdot T(f^i(c)) + \cdots + \sum_{i=0}^{n-2} B_{r-1,i} \cdot T^{r-1}(f^i(c)) = 0. \qquad (1.4)$$

Luego, al considerar

$$\left\{ g^m(f^j(c)) : \ 0 \leq m \leq r-1, 0 \leq j \leq n-2 \right\}$$

el término de grado mayor de la forma $f^l(c)$, que obtenemos es cuando $m = r-1$ y $j = n-2$, es decir, $f^{(r-1)n+(n-2)}(c) = f^{rn-2}(c)$ que proviene de $g^{r-1}(f^{n-2}(c))$. Por lo tanto, en la ecuación (1.4), después de agrupar todos los términos de la misma forma $f^l(c)$ con $l \leq rn-2$, obtenemos un polinomio $q(x) = \beta_0 + \beta_1 x + \cdots + \beta_{rn-2}x^{rn-2} \in F[x]$ con $\beta_{rn-2} = B_{r-1,n-2}$ tal que

$$0 = q(f)(c) = \beta_0 + \beta_1 f(c) + \cdots + \beta_{rn-3}f^{rn-3}(c) + \beta_{rn-2}f^{rn-2}(c).$$

De la misma manera que arriba, concluimos que $q \in \mathrm{Ann}(V_f) = \langle g^r \rangle$, i.e, $g^r \mid q$. Por lo tanto, $q(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{rn-1}x^{rn-1} = 0$ y así $\alpha_{rn-1} = B_{r-1,n-1} = 0$.
Podemos proseguir de la misma manera y así concluir que $B_{j,i} = 0$ para todo j, i y, por lo tanto,

$$A = \left\{ g^m(f^j(c)) : \ 0 \leq m \leq r-1, 0 \leq j \leq n-1 \right\} \subseteq V$$

es linealmente independiente. Como $|A| = rn = \dim(V)$, se tiene que A es una base de V. Si nos fijamos en los elementos de la base, exceptuando los de la última columna, al aplicarles f, todos van a dar al siguiente elemento. Ahora, notemos que

$$f^n(c) = -a_0 c - a_1 f(c) - \cdots - a_{n-1}f^{n-1}(c) + g(f)(c).$$

Dado que el polinomio g satisface que $g^m x = x g^m$, entonces $f g^m(f) = g^m(f) f$, de modo que

$$\begin{aligned} f(g^m(f))(f^{n-1}(c)) &= g^m(f)(f^n(c)) \\ &= g^m(f)(-a_0 c - a_1 f(c) - \cdots - a_{n-1} f^{n-1}(c) + g(f)(c)) \\ &= -a_0 g^m(f)(c) - \cdots - a_{n-1} g^m(f)(f^{n-1}(c)) + g^{m+1}(f)(c), \end{aligned}$$

con $0 \leq m \leq r-1$. Por lo tanto, la representación matricial de f con la base del enunciado es la que se menciona en el enunciado. □

Observación 1.4.10. *a) Para que quede un poco más clara la forma de la matriz asociada a f con respecto a la base del teorema 1.4.9, tenemos que si $r = 3$, la matriz es de la siguiente forma:*

$$\left(\begin{array}{ccccc|ccccc|ccccc} 0 & 0 & \cdots & 0 & -a_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \mathbf{1} & 0 & \cdots & 0 & -a_1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{1} & \cdots & 0 & -a_2 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & -a_{n-1} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & \mathbf{1} & 0 & 0 & \cdots & 0 & -a_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \mathbf{1} & 0 & \cdots & 0 & -a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \mathbf{1} & \cdots & 0 & -a_2 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & \mathbf{1} & -a_{n-1} & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \mathbf{1} & 0 & 0 & \cdots & 0 & -a_0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \mathbf{1} & 0 & \cdots & 0 & -a_1 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \mathbf{1} & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{array}\right)$$

b) Notemos que en la matriz del teorema 1.4.9, cada entrada abajo de la diagonal es 1 (ver matriz de arriba).

Notación 1.4.11. *Sea $V \neq 0$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal tal que V_f es cíclico con generador c. Supongamos que el polinomio mínimo de f es g^r con*

$$g(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n$$

y así, g^r es de la forma

$$g^r = b_0 + b_1 x + \cdots + b_{rn-1} x^{rn-1} + x^{rn}.$$

a) Por el Teorema 1.4.7 existe una base β_1 de V tal que f tiene la siguiente matriz asociada respecto de β_1 y la cual denotaremos por $[g]^r$ (la matriz compañera de g^r):

$$[g]^r = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -b_0 \\ 1 & 0 & 0 & \cdots & 0 & -b_1 \\ 0 & 1 & 0 & \cdots & 0 & -b_2 \\ 0 & 0 & 1 & \cdots & 0 & -b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -b_{rn-1} \end{pmatrix}$$

b) Por el Teorema 1.4.9 existe una base β_2 de V tal que f tiene la siguiente matriz asociada respecto de β_2 y la cual denotaremos por $[g]_r$

$$[g]_r = \begin{pmatrix} [g] & & & & \\ A & [g] & & & \\ & A & [g] & & \\ & & & \ddots & \\ & & & A & [g] \end{pmatrix}$$

donde $[g]$ es la matriz compañera de g y

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Notemos que la primera matriz $[g]^r$ tiene a r como superíndice y la matriz $[g]_r$ tiene a r como subíndice.

Ahora, para definir la forma Canónica de Jordan, necesitamos las siguientes nociones preliminares.

Definición 1.4.12. *Sea $p = a_0 + a_1x + \cdots + a_nx^n \in F[x]$ un elemento primo mónico de grado n. Una **p-matriz clásica**, es una matriz A de la siguiente forma:*

$$A = \begin{pmatrix} [p]_{m_1} & & & \\ & [p]_{m_2} & & \\ & & \ddots & \\ & & & [p]_{m_n} \end{pmatrix}$$

donde $0 < m_1 \leq m_2 \leq \cdots \leq m_n$ y cada $[p]_{m_i}$ es una matriz cuadrada de tamaño nm_i como la del teorema 1.4.9. Es decir, cada $[p]_{m_i}$ es de la forma

$$[p]_{m_i} = \begin{pmatrix} [p] & & & & \\ B_i & [p] & & & \\ & B_i & [p] & & \\ & & & \ddots & \\ & & & B_i & [p] \end{pmatrix}$$

donde $[p]$ es la matriz compañera de p de tamaño $n \times n$ y

$$B_i = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

también es de tamaño $n \times n$.

Definición 1.4.13. *Una **matriz canónica clásica** es una matriz cuadrada de la forma*

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix},$$

donde cada A_i es una p_i-matriz clásica para distintos primos $p_1, \ldots, p_n \in F[x]$.

Notemos que si $p = x - \lambda \in F[x]$, la matriz compañera de $[p]$ es de 1×1 de la forma (λ) y en este caso $[p]_{m_i}$ es una matriz de tamaño m_i donde todas las entradas de la diagonal son λ, bajo la diagonal hay puros unos y cero en las demás entradas. Así, tenemos las siguiente definición.

Definición 1.4.14. *Sea $p = x - \lambda \in F[x]$. Una λ-**matriz elemental de Jordan** de tamaño $m_i \times m_i$ es una matriz de la forma*

$$[x - \lambda]_{m_i} = \begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & & \ddots & \\ & & & 1 & \lambda \end{pmatrix}.$$

Definición 1.4.15. *Una $(x - \lambda)$-**matriz de Jordan** es una $(x - \lambda)$-matriz clásica, es decir, una matriz A de la siguiente forma:*

$$A = \begin{pmatrix} [x - \lambda]_{m_1} & & & \\ & [x - \lambda]_{m_2} & & \\ & & \ddots & \\ & & & [x - \lambda]_{m_n} \end{pmatrix}$$

donde $0 < m_1 \leq m_2 \leq \cdots \leq m_n$ y cada $[x - \lambda]_{m_i}$ es una λ-matriz elemental de Jordan de tamaño $m_i \times m_i$ de la forma:

$$[x - \lambda]_{m_i} = \begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & & \ddots & \\ & & & 1 & \lambda \end{pmatrix}$$

*Una **forma canónica de Jordan** es una matriz canónica clásica*

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}$$

donde cada A_i es una $(x-\lambda_i)$-matriz de Jordan para distintos primos $x-\lambda_1, ..., x-\lambda_n \in F[x]$.
*Sea V un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación F-lineal. Decimos que f **tiene una forma canónica de Jordan** si su representación matricial en alguna base es una forma canónica de Jordan.*

Ejemplo 1.4.16. *Por ejemplo, tenemos la siguiente forma canónica de Jordan*

$$J = \left(\begin{array}{c|ccc|ccc|cc} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Lema 1.4.17. *Sean V un F-espacio vectorial de dimensión finita, $f : V \longrightarrow V$ una transformación lineal y $W \subseteq V$ un F-subespacio de V. Entonces*

a) W es un $F[x]$-submódulo de V_f si y sólo si W es f-invariante. En este caso, escribiremos W_f cuando consideremos a W como submódulo de V_f.

b) Supongamos que W es f-invariante y consideremos $g = f|_W : W \longrightarrow W$. Tenemos que W_f es un $F[x]$-submodulo de V_f y, además,

$$W_g = W_f,$$

donde W_g es la estructura de $F[x]$-modulo inducida por g en W.

Demostración. a) Primero supongamos que W es un $F[x]$-submódulo de V_f. En este caso, tenemos que, para todo $w \in W$, $f(w) = x \cdot_f w \in W$, i.e., $f(W) \subseteq W$. Por lo tanto, W es f-invariante.
Ahora supongamos que W es f-invariante. Como W es un F-subespacio de V, se tiene que, para todo $w, z \in W$ y $\lambda \in F$, $\lambda w + z \in W$. Falta probar que, para todo $p \in F[x]$, $p \cdot_f w \in W$ si $w \in W$. Para ello, tomemos $w \in W$ y $p(x) \in F[x]$, con

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n.$$

Entonces

$$p(x) \cdot_f w = p(f)(w) = a_0 Id(w) + a_1 f(w) + \cdots + a_n f^n(w) \in W,$$

pues W es un F-subespacio y, para toda $k \in \mathbb{N}$, $f^k(w) \in W$ (ya que $f(W) \subseteq W$). Con ello concluimos que W es un $F[x]$-submódulo de V_f.

b) Por el inciso a), basta probar que $W_g = W_f$. Notemos que el conjunto subyacente de ambos módulos es W y la operación de suma en ambos está dada por la suma de W como F-subespacio de V. Falta ver que el producto por escalares es el mismo.
Sean $w \in W$ y $p \in F[x]$. Entonces

$$p \cdot_g w = p(g)(w) = p(f|_W)(w) = p(f)(w) = p \cdot_f w.$$

Por lo tanto,

$$W_g = W_f.$$

□

Teorema 1.4.18. *Sea $V \neq 0$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Sea $g = \prod_{i=1}^{n} p_i^{\alpha_i}$ el polinomio mínimo de f donde $p_1, ..., p_n$ son elementos primos no asociados entre sí en $F[x]$ y $\alpha_1, ..., \alpha_n$ son enteros positivos y consideremos*

$$(V_f)_{p_i} := \Big\{ x \in V_f \mid p_i^r \cdot_f x = p_i^r(f)(x) = 0, \text{ para algún } r \in \mathbb{N} \Big\}.$$

Las siguientes condiciones se satisfacen.

a) Cada $(V_f)_{p_i}$ es un $F[x]$-submódulo de V_f y

$$V_f = \bigoplus_{i=1}^{n} (V_f)_{p_i}$$

como $F[x]$-módulos. Además, la transformación lineal $f|_{(V_f)_{p_i}} : (V_f)_{p_i} \longrightarrow (V_f)_{p_i}$ tiene como polinomio mínimo a $p_i^{\alpha_i}$.

b) Para cada $i = 1, ..., n$ existe una familia $\{V_{i,j}\}_{j=1}^{k_i}$ de $F[x]$-submódulos de $(V_f)_{p_i}$ tales que $(V_f)_{p_i} = \bigoplus_{j=1}^{k_i} V_{i,j}$ donde $V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}} \rangle$ y $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$. Además el polinomio mínimo de $f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$ es $p_i^{\alpha_{i,j}}$.

Demostración. a) Por la proposición 1.4.5, tenemos que V_f es un $F[x]$-módulo de torsión y además es finitamente generado ya que como V es de dimensión finita como F-espacio vectorial, entonces es finitamente generado como $F[x]$-módulo pues $F \subseteq F[x]$. La proposición 1.1.10 nos permite utilizar el teorema 1.3.6 para concluir que existen primos $p_1, ..., p_n \in F[x]$ únicos salvo asociados y salvo el orden de los índices tales que $(V_f)_{p_i} \neq 0$ para todo $i = 1, ..., n$ y tales que

$$V_f = \bigoplus_{i=1}^{n} (V_f)_{p_i}.$$

Por definición del polinomio mínimo sabemos que $\mathrm{Ann}(V_f) = \langle g \rangle$. Luego, del teorema 1.3.6, sabemos que $p_1, ..., p_n$ son los elementos que aparecen en la descomposición en primos (irreducibles) del polinomio mínimo g de f. El polinomio mínimo es mónico, por ello $g = p_1^{\alpha_1} \cdots p_n^{\alpha_n}$. En la descomposición en primos de g puede suceder que algún p_i no sea mónico, i.e., $p_i = uq_i$ con q_i primo y mónico y $u \in F$; en este caso,

dado que g es mónico, u se tiene que cancelar con algún otro elemento $p_j = u^{-1}q_j$ con q_j primo y mónico. Por ende, siempre podemos suponer que los primos p_i son mónicos.
Ahora, sabemos que $(V_f)_{p_i}$ es un $F[x]$-submódulo de V_f por la proposición 1.3.4 y así $(V_f)_{p_i}$ es f-invariante por el lema 1.4.17. Entonces tenemos una transformación lineal

$$f|_{(V_f)_{p_i}} : (V_f)_{p_i} \longrightarrow (V_f)_{p_i}.$$

Por el teorema 1.3.6, sabemos que $\mathrm{Ann}\Big((V_f)_{p_i}\Big) = \langle p_i^{\alpha_i}\rangle$ y así, $p_i^{\alpha_i}$ es el polinomio mínimo de $f|_{(V_f)_{p_i}}$.

b) Usando el mismo argumento que para V_f tenemos que $(V_f)_{p_i}$ es un $F[x]$-módulo finitamente generado y además por definición de $(V_f)_{p_i}$ tenemos que $(V_f)_{p_i}$ es un p_i-módulo. Por el teorema 1.3.10, tenemos que

$$(V_f)_{p_i} \cong \bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\alpha_{i,j}}\rangle$$

con $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$. Por lo tanto, existe una familia $\{V_{i,j}\}_{j=1}^{k_i}$ de $F[x]$-submódulos de $(V_f)_{p_i}$ tal que $(V_f)_{p_i} = \bigoplus_{j=1}^{k_i} V_{i,j}$ donde $V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}}\rangle$. Por el lema 1.4.17, concluimos que $V_{i,j}$ es f-invariante, es decir, $f_{i,j} := f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$. Por el lema 1.4.17 tenemos que $(V_{i,j})_{f_{i,j}} = (V_{i,j})_f$. Por el teorema 1.3.10, tenemos que $\mathrm{Ann}\Big(F[x]/\langle p_i^{\alpha_{i,j}}\rangle\Big) = \langle p_i^{\alpha_{i,j}}\rangle$ y, por lo tanto, $\mathrm{Ann}(V_{i,j}) = \langle p_i^{\alpha_{i,j}}\rangle$. De donde concluimos que $p_i^{\alpha_{i,j}}$ es el polinomio mínimo de $f_{i,j} = f|_{V_{i,j}}$ (ver definición 1.4.6).

□

Observación 1.4.19. *Como* $(V_f)_{p_i} = \bigoplus_{j=1}^{k_i} V_{i,j}$, *donde* $V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}}\rangle$ *y* $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$, *tenemos que*

$$\mathrm{Ann}((V_f)_{p_i}) = \langle p_i^{\alpha_{i,k_i}}\rangle.$$

Por otro lado, sea $g = \prod_{i=1}^{n} p_i^{\alpha_i}$ *el polinomio mínimo de* f*; y así, por definición del polinomio mínimo tenemos que* $\mathrm{Ann}(V_f) = \langle g\rangle$. *Luego, por el teorema 1.3.6, tenemos que*

$$\mathrm{Ann}((V_f)_{p_i}) = \langle p_i^{\alpha_i}\rangle.$$

Por lo tanto, $\alpha_{i,k_i} = \alpha_i$ *para todo* $i = 1, ..., n$,

Observación 1.4.20. *Notemos que si* $p(x)$ *es un polinomio en* $F[x]$, *entonces* $F[x]/\langle p(x)\rangle$ *es un* $F[x]$*-módulo cíclico, pues está generado por* $1 + \langle p(x)\rangle$.

Con las definiciones anteriores y los resultados previos ya podemos demostrar la existencia de la forma canónica de Jordan.

Teorema 1.4.21 (Existencia de la forma canónica de Jordan)**.** *Sea* F *un campo algebraicamente cerrado. Sean* $V \neq \{0\}$ *un espacio vectorial de dimensión finita sobre* F *y* $f : V \longrightarrow V$ *una transformación lineal. Entonces existe una forma canónica de Jordan para* f.

Demostración. Sea $g = \prod_{i=1}^{n} p_i^{\alpha_i}$ el polinomio mínimo de f donde los p_i son primos en $F[x]$ no asociados entre sí. Por el teorema 1.4.18, tenemos que existen F-subespacios vectoriales $V_{i,j}$ de V que son f-invariantes (es decir, son $F[x]$-submódulos de V_f) tales que tenemos la siguiente descomposición de V_f como $F[x]$-módulos

$$V_f = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big)$$

donde $V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}} \rangle$ para $i = 1, ..., n$ y $1 \leq j \leq k_i$. Además, se tiene que $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$ para todo $i = 1, ..., n$ y el polinomio mínimo de $f_{i,j} := f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$ es $p_i^{\alpha_{i,j}}$. Notemos que la descomposición de arriba nos da una descomposición de V como F-espacio vectorial a través de subespacios f-invariantes:

$$V = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big).$$

Observemos que como $F[x]$ es un DFU, tenemos que $p(x)$ es primo si, y sólo si, $p(x)$ es irreducible y como F es algebraicamente cerrado, tenemos que los únicos polinomios mónicos irreducibles son de la forma $x - \lambda$ con $\lambda \in F$ (ver proposición 1.1.9). Por lo tanto, tenemos que $p_i = x - \lambda_i$ para algún $\lambda_i \in F$. Por el lema 1.4.17(b), tenemos que $(V_{i,j})_{f_{i,j}} = (V_{i,j})_f \cong F[x]/\langle p_i^{\alpha_{i,j}} \rangle$ y así, por la observación 1.4.20, tenemos que $(V_{i,j})_{f_{i,j}}$ es un $F[x]$-módulo cíclico. Luego, utilizando el teorema 1.4.9 en la transformación lineal

$$f_{i,j} : V_{i,j} \longrightarrow V_{i,j}$$

con polinomio mínimo $p_i^{\alpha_{i,j}} = (x - \lambda_i)^{\alpha_{i,j}}$, tenemos que existe una base $\beta_{i,j}$ de $V_{i,j}$ tal que la matriz asociada a $f_{i,j}$ respecto de esta base es de tamaño $\alpha_{i,j} \times \alpha_{i,j}$ y tiene la forma

$$[x - \lambda_i]_{\alpha_{i,j}} = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \lambda_i & & \\ & & & \ddots & \\ & & & 1 & \lambda_i \end{pmatrix}.$$

Como

$$V = \Big(\bigoplus_{j=1}^{k_1} F[x]/\langle p_1^{\alpha_{1,j}} \rangle \Big) \bigoplus \Big(\bigoplus_{j=1}^{k_2} F[x]/\langle p_2^{\alpha_{2,j}} \rangle \Big) \cdots \bigoplus \Big(\bigoplus_{j=1}^{k_n} F[x]/\langle p_n^{\alpha_{n,j}} \rangle \Big)$$

Tenemos que

$$\beta := \Big(\bigcup_{j=1}^{k_1} \beta_{1,j} \Big) \bigcup \Big(\bigcup_{j=1}^{k_2} \beta_{2,j} \Big) \bigcup \cdots \bigcup \Big(\bigcup_{j=1}^{k_n} \beta_{n,j} \Big)$$

es una base tal que la matriz asociada a f respecto de esta base es una forma canónica de Jordan. □

1.5. Unicidad de la forma canónica de Jordan

Para demostrar la unicidad de la forma canónica de Jordan necesitaremos el siguiente teorema.

Teorema 1.5.1. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Supongamos que existe un polinomio mónico g de grado n, con*

$$g(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n,$$

y una base ordenada $\gamma = \{v_1, ..., v_{rn}\}$ tal que la representación matricial de f en esta base es la que está dada en el teorema 1.4.9. Bajo estas condiciones V_f es cíclico, generado por v_1 y el polinomio mínimo de f es g^r.

Demostración. Dado que la representación matricial de f en la base γ es de la forma

$$\begin{pmatrix} [g] & & & & \\ A & [g] & & & \\ & A & [g] & & \\ & & & \ddots & \ddots \\ & & & & A \quad [g] \end{pmatrix}$$

donde

$$[g] = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

se tiene que $f(v_i) = v_{i+1}$, siempre que i no sea múltiplo de n. Si $0 \leq k \leq r-2$, entonces

$$\begin{aligned} f^n(v_{kn+1}) &= f(f^{n-1}(v_{kn+1})) \\ &= f(v_{(k+1)n}) \\ &= -a_0v_{kn+1} - a_1v_{kn+2} - \cdots - a_{n-1}v_{(k+1)n} + v_{(k+1)n+1} \\ &= -a_0v_{kn+1} - a_1f(v_{kn+1}) - \cdots - a_{n-1}f^{n-1}(v_{kn+1}) + v_{(k+1)n+1}. \end{aligned}$$

De donde $v_{(k+1)n+1} = g(f)(v_{kn+1})$ y, por inducción, $v_{(k+1)n+1} = g(f)^{k+1}(v_1)$, para $0 \leq k \leq r-1$. Lo anterior prueba que cualquier elemento de γ es de la forma $f^\beta(g(f)^\alpha(v_1))$, con $\alpha, \beta \in \mathbb{N}$ adecuados, donde $g(f)^\alpha = Id$ y $f^\beta = Id$ si $\alpha = \beta = 0$. Consideremos el polinomio $h(x) = x^\beta g(x)^\alpha$. Así, tenemos que $h(x) \cdot_f v_1 = f^\beta(g(f)^\alpha(v_1))$ y, por lo tanto, $V_f = \langle v_1 \rangle$. Además,

$$\begin{aligned} g(f)^r(v_1) &= g(f)(g(f)^{r-1}(v_1)) \\ &= g(f)(v_{(r-1)n+1}) \\ &= a_0v_{(r-1)n+1} + \cdots + a_{n-1}f^{n-1}(v_{(r-1)n+1}) + f^n(v_{(r-1)n+1}) \\ &= a_0v_{(r-1)n+1} + a_1v_{(r-1)n+2} + \cdots + a_{n-1}v_{rn} + f(v_{rn}) \\ &= 0. \end{aligned}$$

La última igualdad se sigue de la última columna de la representación matricial de f con respecto a γ. Como $V_f = \langle v_1 \rangle$ y $g(f)^r(v_1) = 0$, se tiene que $g(f)^r \equiv 0$. Por el teorema 1.4.7, el grado del polinomio mínimo de f coincide con $\dim(V) = rn$. Sabemos que $\text{grad}(g^r) = rn$ y g^r es mónico (porque g lo es), entonces, por definición de polinomio mínimo, se concluye que g^r es el polinomio mínimo de f. □

Teorema 1.5.2 (Unicidad de la forma canónica de Jordan)**.** *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Si J y J' son dos formas canónicas de Jordan para f, entonces es posible reordenar la base asociada a la representación matricial J' de tal forma que la nueva representación matricial coincida con J.*

Demostración. Supongamos que J es una forma canónica de Jordan de f, es decir, existe una base B tal que la representación matricial de f respecto a B tiene la forma:

$$J = \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_n \end{pmatrix}$$

donde cada A_i es una $[x - \lambda_i]$-matriz de Jordan para ciertos $\lambda_1, ..., \lambda_n \in F$ distintos:

$$A_i = \begin{pmatrix} [x-\lambda_i]_{\alpha_{i,1}} & & & \\ & [x-\lambda_i]_{\alpha_{i,2}} & & \\ & & \ddots & \\ & & & [x-\lambda_i]_{\alpha_{i,k_i}} \end{pmatrix}.$$

Además, cada $[x - \lambda_i]_{\alpha_{i,j}}$ es una λ_i-matriz elemental de Jordan de tamaño $\alpha_{i,j} \times \alpha_{i,j}$ de la forma:

$$[x - \lambda_i]_{\alpha_{i,j}} = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \lambda_i & & \\ & & & \ddots & \\ & & & 1 & \lambda_i \end{pmatrix},$$

y las matrices $[x - \lambda_i]_{\alpha_{i,j}}$ de A_i están ordenadas de tal forma que $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$, esto para cada i con $1 \leq i \leq n$.
Luego, por la definición de cómo se construye la matriz asociada a f, existe $B_{i,j} \subseteq B$ tal que si $V_{i,j} := \langle B_{i,j} \rangle \subseteq V$, entonces f se restringe bien a $V_{i,j}$ y la matriz asociada a $f|_{V_{i,j}}$ respecto de $B_{i,j}$ es $[x - \lambda_i]_{\alpha_{i,j}}$. Aplicando el teorema 1.5.1, a la transformación lineal $f_{i,j} := f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$, al polinomio $x - \lambda_i$ y a la base ordenada $B_{i,j}$, tenemos que $V_{i,j}$ es cíclico como $F[x]$-módulo generado por el primer vector de la base ordenada $B_{i,j}$, digamos $\mathbf{v_{i,j}}$, y $f|_{V_{i,j}}$ tiene polinomio mínimo $p_{i,j} := (x - \lambda_i)^{\alpha_{i,j}}$. De lo anterior, tenemos un morfismo de $F[x]$-módulos que es suprayectivo

$$\varphi_{i,j} : F[x] \longrightarrow V_{i,j}$$

dado por $\varphi_{i,j}(g(x)) = g(x)\cdot_{f_{i,j}} \mathbf{v_{i,j}}$ (ver definición 1.4.1); y además $\mathrm{Ker}(\varphi_{i,j}) = \langle (x-\lambda_i)^{\alpha_{i,j}} \rangle$. Por lo tanto, tenemos un isomorfismo de $F[x]$-módulos:

$$V_{i,j} \cong F[x]/\langle (x-\lambda_i)^{\alpha_{i,j}} \rangle.$$

De la representación matricial de f y del isomorfismo de arriba, tenemos que

$$V \cong \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big) \cong \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} F[x]/\langle (x-\lambda_i)^{\alpha_{i,j}} \rangle \Big).$$

Notemos que para i fijo, se tiene que $\bigoplus_{j=1}^{k_i} F[x]/\langle (x-\lambda_i)^{\alpha_{i,j}} \rangle$ es un $(x-\lambda_i)$-módulo (ver definición 1.3.5). Supongamos que J' es otra forma canónica de Jordan de f

$$J' = \begin{pmatrix} \boxed{B_1} & & & \\ & \boxed{B_2} & & \\ & & \ddots & \\ & & & \boxed{B_m} \end{pmatrix},$$

donde cada B_i es una $(x-\gamma_i)$-matriz de Jordan para $\gamma_1, ..., \gamma_m \in F$ distintos y cada γ_i-matriz elemental de Jordan $[x-\gamma_i]_{\beta_{i,j}}$ que pertenece a B_i es de tamaño $\beta_{i,j} \times \beta_{i,j}$ con $\beta_{i,1} \leq \beta_{i,2} \leq \cdots \leq \beta_{i,t_i}$. De una manera similar a como le hicimos para J, tenemos una descomposición de V:

$$V \cong \bigoplus_{i=1}^{m} \Big(\bigoplus_{j=1}^{t_i} V'_{i,j} \Big) \cong \bigoplus_{i=1}^{m} \Big(\bigoplus_{j=1}^{t_i} F[x]/\langle (x-\gamma_i)^{\beta_{i,j}} \rangle \Big).$$

Luego por teorema 1.3.11, tenemos que $n = m$, $x - \lambda_i = x - \gamma_i$ para todo $i = 1, ..., m$, $k_i = t_i$ y $\alpha_{i,j} = \beta_{i,j}$ para $1 \leq j \leq k_i$. Por lo tanto, $J = J'$, probándose la unicidad de la forma canónica de Jordan. □

Observación 1.5.3. *Sean F un campo algebraicamente cerrado, $V \neq 0$ un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación lineal. Como F es algebraicamente cerrado, tenemos que los elementos primos no nulos de $F[x]$ son de la forma $x - \lambda_i$ con $\lambda_i \in F$. Como V_f es de torsión, como $F[x]$-módulo, tenemos que $\mathrm{Ann}(V_f) = \langle g \rangle$, donde g es el polinomio mínimo de f.*
Por la demostración del teorema 1.3.6, tenemos que existen únicos primos $p_i = x - \lambda_i \in F[x]$ tales que

$$V = \bigoplus_{i=1}^{n} V_{x-\lambda_i}$$

donde $g(x) = (x-\lambda_1)^{a_1} \cdots (x-\lambda_n)^{a_n}$ para ciertos enteros positivos $a_i > 1$.

a) Por la demostración de la existencia de la forma canónica de Jordan en el teorema 1.4.21, tenemos que las matrices λ_i-elementales que aparecen en la $[x-\lambda_i]$-matriz de Jordan son de la forma

$$\begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \lambda_i & & \\ & & & \ddots & \\ & & & 1 & \lambda_i \end{pmatrix}.$$

Al observar la última columna de la matriz anterior, podemos notar que λ_i es un valor propio para $i = 1, ..., n$ (esto por cómo se construye la matriz asociada a f). Inversamente, si $\lambda \in F$ es un valor propio de f, entonces $\lambda = \lambda_i$ para alguna raíz λ_i de $g(x)$. Para ver esta afirmación, sea $\lambda \in F$ un valor propio de f y $0 \neq v \in V$ un vector propio asociado a λ. Entonces, si $g(x) = a_0 + a_1 x + \cdots + x^r$, se tiene que

$$\begin{aligned} 0 &= g(f)(v) \\ &= a_0 v + a_1 f(v) + \cdots + a_{n-1} f^{r-1}(v) + f^r(v) \\ &= a_0 v + a_1 \lambda v + \cdots + a_{r-1} \lambda^{r-1} v + \lambda^r v \\ &= (a_0 + a_1 \lambda + \cdots + a_{r-1} \lambda^{r-1} + \lambda^r) v \\ &= g(\lambda) v. \end{aligned}$$

Como $v \neq 0$ se concluye que $g(\lambda) = 0$; y por lo tanto $\lambda = \lambda_i$ para alguna i.

b) *Consideremos el* ***polinomio característico*** *$p(x) = (x - \lambda_1)^{b_1} \cdots (x - \lambda_n)^{b_n}$ de f. Sea K_{λ_i} el espacio propio generalizado asociado al valor propio λ_i, por el teorema 3.1.3 del capítulo 3, tenemos que $\dim(K_{\lambda_i}) = b_i$. Notemos que en la notación que estamos utilizando en este capítulo, tenemos que*

$$K_{\lambda_i} := (V_f)_{x - \lambda_i}.$$

Del polinomio característico, podemos determinar el tamaño de las $[x - \lambda_i]$-matrices A_i de la forma canónica de Jordan de f:

$$J = \begin{pmatrix} \boxed{A_1} & & \\ & \boxed{A_2} & \\ & & \ddots \\ & & & \boxed{A_n} \end{pmatrix}.$$

En efecto, tenemos que A_i es de tamaño $b_i \times b_i$ para cada $i = 1, ..., n$. Ahora, como cada A_i es de la forma:

$$A_i = \begin{pmatrix} \boxed{[x - \lambda_i]_{\alpha_{i,1}}} & & \\ & \boxed{[x - \lambda_i]_{\alpha_{i,2}}} & \\ & & \ddots \\ & & & \boxed{[x - \lambda_i]_{\alpha_{i,k_i}}} \end{pmatrix},$$

tenemos que para cada $i = 1, ..., n$ se debe cumplir:

$$b_i := \sum_{j=1}^{k_i} \alpha_{i,j}.$$

En el capítulo 3, veremos cómo determinar los $\alpha_{i,j}$ para cada i, j, sólo conociendo f y el polinomio característico p de f.
Finalmente, notemos que si g es el polinomio mínimo de f, entonces g divide al polinomio característico de f. De hecho, por la observación 1.4.19, tenemos que $\mathrm{Ann}((V_f)_{x - \lambda_i}) = \langle (x - \lambda_i)^{\alpha_{i,k_i}} \rangle$ y así concluimos que el polinomio mínimo de f es $g = (x - \lambda_1)^{\alpha_{1,k_1}} \ldots (x - \lambda_n)^{\alpha_{n,k_n}}$.

Capítulo 2

Forma canónica racional y ejemplos

2.1. Forma canónica racional de divisores elementales

Para probar la existencia de la forma canónica de Jordan utilizamos fuertemente que el campo F fuera algebraicamente cerrado. Esta condición es esencial como podemos ver en el siguiente ejemplo.

Ejemplo 2.1.1. *Consideremos la transformación $L_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $L_A(x) = Ax$, para $x \in \mathbb{R}^2$, con A la siguiente matriz,*

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

Si esta transformación tuviera una forma canónica de Jordan, entonces existiría una base para $\mathbb{R}^2$ tal que la representación matricial de A en esa base está dada por

$$\begin{pmatrix} a & 0 \\ b & c \end{pmatrix},$$

con $a, c \in \mathbb{R}$ y $b \in \{0, 1\}$. Pero eso implicaría que c es un valor propio de A, una contradicción, pues el polinomio característico de A es $p(x) = x^2 + 1$ que, en $\mathbb{R}$, no tiene raíces.

Ahora probaremos la existencia y unicidad de otras formas canónicas para las que no se necesita que el campo sea algebraicamente cerrado.

Definición 2.1.2. *Sean $V \neq 0$ un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación lineal. Por el teorema 1.4.18, tenemos que*

$$V \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\alpha_{i,j}} \rangle.$$

donde $p_i \in F[x]$ son primos mónicos distintos para $i = 1, ..., n$ y $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$ para cada i. Los polinomios $p_i \in F[x]$ de tal descomposición son llamados los ***divisores elementales de f****.*

Definición 2.1.3. *Sean F un campo y $p \in F[x]$ un polinomio mónico primo. Una* ***matriz p-racional*** *sobre F es una matriz de la siguiente forma*

$$\begin{pmatrix} [p]^{m_1} & & & \\ & [p]^{m_2} & & \\ & & \ddots & \\ & & & [p]^{m_k} \end{pmatrix}$$

para alguna sucesión de enteros $m_1 \leq m_2 \leq \cdots \leq m_k$, donde acorde a la notación dada en 1.4.11, tenemos que $[p]^{m_j}$ denota la matriz compañera de p^{m_j} para cada $1 \leq j \leq k$.

Definición 2.1.4. *Una* ***forma canónica racional de divisores elementales*** *sobre un campo F es una matriz A de la forma*

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}$$

donde cada A_i es una matriz p_i-racional para distintos primos $p_1, ..., p_n$ de $F[x]$.

Definición 2.1.5. *Sea V un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Decimos que una matriz A es una* ***forma canónica racional de divisores elementales de f*** *si existe una base β de V tal que la matriz asociada respecto a esta base es una matriz de la forma dada en la definición 2.1.4.*

Con las definiciones anteriores podemos probar el siguiente teorema.

Teorema 2.1.6. *Sea $f : V \longrightarrow V$ una transformación lineal, con $V \neq 0$ y de dimensión finita sobre un campo F. Entonces existe una forma canónica racional de divisores elementales.*

Demostración. Sea $g = \prod_{i=1}^{n} p_i^{\alpha_i}$ el polinomio mínimo de f donde los p_i son primos en $F[x]$ no asociados entre sí. Note que, como g es mónico, podemos asumir que cada p_i es mónico. Por el teorema 1.4.18, tenemos que existen F-subespacios vectoriales $V_{i,j}$ de V que son f-invariantes (es decir, son $F[x]$-submódulos de V_f) tales que tenemos la siguiente descomposición de V_f como $F[x]$-módulos

$$V_f = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big)$$

donde $V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}} \rangle$ para $i = 1, ..., n$ y $1 \leq j \leq k_i$. Además, se tiene que $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$ para todo $i = 1, ..., n$ y el polinomio mínimo de $f_{i,j} := f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$ es $p_i^{\alpha_{i,j}}$.
Notemos que la descomposición de arriba nos da una descomposición de V como F-espacio vectorial a través de subespacios f-invariantes:

$$V = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big).$$

Dado que $F[x]/\langle p_i^{\alpha_{i,j}}\rangle$ es cíclico como $F[x]$-módulo, también $V_{i,j}$ lo es, es decir, podemos aplicar el teorema 1.4.7 a la transformación lineal $f_{i,j} : V_{i,j} \longrightarrow V_{i,j}$ y de esta manera existe una base $\beta_{i,j}$ de $V_{i,j}$ tal que la matriz asociada a $f_{i,j}$ es $[p_i]^{\alpha_{i,j}}$ (ver notación 1.4.11). Así, tenemos que $\bigcup_{j=1}^{k_i} \beta_{i,j}$ es una base de $(V_f)_{p_i} = \bigoplus_{j=1}^{k_i} V_{i,j}$ tal que la matriz de $f|_{(V_f)_{p_i}} : (V_f)_{p_i} \longrightarrow (V_f)_{p_i}$ es de la forma

$$A_i = \begin{pmatrix} [p_i]^{\alpha_{i,1}} & & & \\ & [p_i]^{\alpha_{i,2}} & & \\ & & \ddots & \\ & & & [p_i]^{\alpha_{i,k_i}} \end{pmatrix}$$

con $\alpha_{i,1} \leq \cdots \leq \alpha_{i,k_i}$.
Finalmente, $\beta = \bigcup_{i=1}^{n} \left(\bigcup_{j=1}^{k_i} \beta_{i,j} \right)$ es una base de V tal que la matriz asociada es de la forma

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}$$

donde cada A_i es una matriz p_i-racional para distintos primos $p_1, ..., p_n$ de $F[x]$.

□

Para probar la unicidad necesitaremos el siguiente teorema.

Teorema 2.1.7. *Sea $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F. Si existe una base $\{v_1, ..., v_n\}$ tal que la representación matricial de f con respecto a esta base es de la forma*

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix},$$

entonces $V_f = \langle v_1 \rangle$ y el polinomio mínimo de f es

$$g(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n.$$

Demostración. Observemos que por la naturaleza de la representación matricial de f, se tiene que $f^k(v_1) = v_{k+1}$, $1 \leq k \leq n-1$. Como $x^k \cdot_f v_1 = f^k(v_1)$, concluimos que $\{v_1, ..., v_n\} \subseteq \langle v_1 \rangle \subseteq V_f$ y dado que $\{v_1, ..., v_n\}$ es un conjunto generador para V_f, se infiere que $\langle v_1 \rangle = V_f$. La última columna de la representación matricial nos da la siguiente identidad,

$$0 = a_0 v_1 + a_1 v_2 + \cdots + a_{n-1} v_n + f(v_n) = a_0 v_1 + a_1 f(v_1) + \cdots + a_{n-1} f^{n-1}(v_1) + f^n(v_1) = g \cdot_f v_1,$$

lo que implica que $g \in \mathrm{Ann}(V_f)$. Ahora, como V_f es cíclico, podemos aplicar el teorema 1.4.7 y concluir que el grado del polinomio mínimo de f es $\dim(V) = n$. Como el grado de g es n, concluimos que g es el polinomio mínimo de f. □

Teorema 2.1.8. *Sea $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F. Si A y B son dos formas canónicas racionales de divisores elementales, entonces podemos reordenar la base asociada a B para que la representación matricial con ese orden coincida con A.*

Demostración. La demostración es análoga a como se hizo en la demostración del teorema 1.5.2, daremos un bosquejo para beneficio del lector.
Sea A una forma racional de divisores elementales de f. Luego, existe una base β de V tal que la matriz A de f respecto de esta base es una forma canónica racional de divisores elementales de f de la forma

$$A = \begin{pmatrix} A_1 & & \\ & \boxed{A_2} & \\ & & \ddots \\ & & & \boxed{A_n} \end{pmatrix}$$

donde cada A_i es una matriz p_i-racional para distintos primos $p_1, ..., p_n$ de $F[x]$:

$$A_i = \begin{pmatrix} [p_i]^{\alpha_{i,1}} & & \\ & \boxed{[p_i]^{\alpha_{i,2}}} & \\ & & \ddots \\ & & & \boxed{[p_i]^{\alpha_{i,k_i}}} \end{pmatrix}$$

para alguna sucesión de enteros $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$.

Luego, por la definición de cómo se construye la matriz asociada a f, existe $\beta_{i,j} \subseteq \beta$ tal que si $V_{i,j} := \langle \beta_{i,j} \rangle \subseteq V$, entonces f se restringe bien a $V_{i,j}$ y la matriz asociada a $f|_{V_{i,j}}$ respecto de $\beta_{i,j}$ es $[p_i]^{\alpha_{i,j}}$ (la matriz compañera de $p_i^{\alpha_{i,j}}$). Aplicando el teorema 2.1.7, a la transformación lineal $f_{i,j} := f|_{V_{i,j}} : V_{i,j} \longrightarrow V_{i,j}$, al polinomio $p_i^{\alpha_{i,j}}$ y a la base ordenada $\beta_{i,j}$, concluimos que $V_{i,j}$ es cíclico como $F[x]$-módulo, generado por el primer vector de la base ordenada $\beta_{i,j}$, digamos $\mathbf{v_{i,j}}$, y $f|_{V_{i,j}}$ tiene polinomio mínimo $p_{i,j}^{\alpha_{i,j}}$. De lo anterior, tenemos un morfismo de $F[x]$-módulos que es suprayectivo

$$\varphi_{i,j} : F[x] \longrightarrow V_{i,j}$$

dado por $\varphi_{i,j}(g(x)) = g(x) \cdot_{f_{i,j}} \mathbf{v_{i,j}}$ (ver definición 1.4.1); y además $\mathrm{Ker}(\varphi_{i,j}) = \langle p_i^{\alpha_{i,j}} \rangle$. Por lo tanto, tenemos un isomorfismo de $F[x]$-módulos:

$$V_{i,j} \cong F[x]/\langle p_i^{\alpha_{i,j}} \rangle.$$

De la representación matricial de f y del isomorfismo de arriba, tenemos que

$$V_f = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} V_{i,j} \Big) \cong \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\alpha_{i,j}} \rangle \Big).$$

Supongamos que B es otra forma canónica racional de divisores elementales de f:

$$B = \begin{pmatrix} B_1 & & \\ & \boxed{B_2} & \\ & & \ddots \\ & & & \boxed{B_m} \end{pmatrix},$$

donde cada B_i es una matriz q_i-racional para distintos primos $q_1, ..., q_m \in F[x]$. De una manera similar a como le hicimos para A, tenemos una descomposición de V_f:

$$V_f = \bigoplus_{i=1}^{m} \Big(\bigoplus_{j=1}^{t_i} V'_{i,j} \Big) \cong \bigoplus_{i=1}^{m} \Big(\bigoplus_{j=1}^{t_i} F[x]/\langle q_i^{\beta_{i,j}} \rangle \Big).$$

Entonces, por el teorema 1.3.11, tenemos que $n = m$ y $q_i = p_i$ para $i = 1, ..., m$ y también concluimos que $k_i = t_i$ y $\alpha_{i,j} = \beta_{i,j}$ para $1 \leq j \leq k_i$. Por lo tanto, $A = B$, probándose la unicidad de la forma canónica racional de divisores elementales. □

2.2. Forma canónica racional de factores invariantes

Definición 2.2.1. *Una **forma canónica racional de factores invariantes** sobre un campo F es una matriz de la forma*

$$A = \begin{pmatrix} [f_1] & & & \\ & [f_2] & & \\ & & \ddots & \\ & & & [f_n] \end{pmatrix}$$

donde cada $[f_i]$ es la matriz compañera de cierto $f_i \in F[x]$ mónico y además $f_{i+1} \mid f_i$ para $i = 1, ..., n-1$.

Definición 2.2.2. *Sea V un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Decimos que una matriz A es una **forma canónica racional de factores invariantes de f** si existe una base β de V tal que la matriz asociada respecto a esta base es una matriz de la forma dada en la definición 2.2.1.*

Teorema 2.2.3. *Sea $V \neq 0$ un espacio vectorial de dimensión finita sobre un campo F y $f : V \longrightarrow V$ una transformación lineal. Entonces existe una forma canónica racional de factores invariantes de f.*

Demostración. Por el teorema 1.4.18, tenemos que

$$V \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\alpha_{i,j}} \rangle.$$

donde $\alpha_{i,1} \leq \alpha_{i,2} \leq \cdots \leq \alpha_{i,k_i}$ para cada i. Definiendo $\beta_{i,j} := \alpha_{i,k_i-(j-1)}$. Tenemos que

$$V \cong \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\beta_{i,j}} \rangle.$$

con $\beta_{i,1} \geq \beta_{i,2} \geq \cdots \geq \beta_{i,k_i}$ para cada i. Luego, podemos ordenar los ideales de la siguiente forma:

$$\begin{aligned} &\{0\} \subseteq \langle p_1^{\beta_{1,1}} \rangle \subseteq \langle p_1^{\beta_{1,2}} \rangle \subseteq \cdots \subseteq \langle p_1^{\beta_{1,k_1}} \rangle \\ &\{0\} \subseteq \langle p_2^{\beta_{2,1}} \rangle \subseteq \langle p_2^{\beta_{2,2}} \rangle \subseteq \cdots \subseteq \langle p_2^{\beta_{2,k_2}} \rangle \\ &\quad \vdots \\ &\{0\} \subseteq \langle p_n^{\beta_{n,1}} \rangle \subseteq \langle p_n^{\beta_{n,2}} \rangle \subseteq \cdots \subseteq \langle p_n^{\beta_{n,k_n}} \rangle \end{aligned}$$

$$\begin{array}{lllllllll}
\{0\} & \subseteq \langle p_1^{\beta_{1,1}}\rangle & \subseteq \langle p_1^{\beta_{1,2}}\rangle \subseteq & \cdots & \subseteq \cdots\cdots & \subseteq \cdots\cdots & \subseteq \langle p_1^{\beta_{1,k_1-1}}\rangle & \subseteq \underline{\langle p_1^{\beta_{1,k_1}}\rangle} \\
\{0\} & \subseteq \langle p_2^{\beta_{2,1}}\rangle & \subseteq \langle p_2^{\beta_{2,2}}\rangle \subseteq & \cdots & \subseteq \underline{\langle p_2^{\beta_{2,k_2}}\rangle} & \subseteq \langle p_2^0\rangle & \subseteq \cdots\cdots & \subseteq \langle p_2^0\rangle \\
\vdots & & & & & & & \\
\{0\} & \subseteq \langle p_n^{\beta_{n,1}}\rangle & \subseteq \langle p_n^{\beta_{n,2}}\rangle \subseteq & \cdots \subseteq \underline{\langle p_n^{\beta_{n,k_n}}\rangle} & \subseteq \langle p_n^0\rangle & \subseteq \langle p_n^0\rangle & \subseteq \cdots\cdots & \subseteq \langle p_n^0\rangle
\end{array}$$

Figura 2.1: Arreglo de divisores elementales de f

Sea $t := \max\{k_i \mid 1 \leq i \leq n\}$. Para cada $i = 1, ..., n$, definimos $\beta_{i,j} = 0$ para $k_i + 1 \leq j \leq t$. Sin pérdida de generalidad, podemos suponer que $k_1 = t$ y $k_1 \geq k_2 \geq \cdots \geq k_n$.

Esto lo podemos visualizar en el arreglo de la figura 2.1, que tiene n renglones y t columnas. Ahora, para j con $1 \leq j \leq t$ definimos:

$$q_j := \prod_{i=1}^{n} p_i^{\beta_{i,j}}.$$

Es decir, q_j es el producto de los elementos de la columna j del arreglo de la figura 2.1. Ahora veremos que

$$F[x]/\langle q_j\rangle \cong \bigoplus_{i=1}^{n} F[x]/\langle p_i^{\beta_{i,j}}\rangle.$$

Para ello notemos que si $a, b \in F[x]$ son elementos coprimos, entonces

$$(*): \quad F[x]/\langle ab\rangle \cong F[x]/\langle a\rangle \bigoplus F[x]/\langle b\rangle.$$

En efecto, definimos

$$\phi : F[x] \longrightarrow F[x]/\langle a\rangle \bigoplus F[x]/\langle b\rangle,$$

como $\phi(d) = (d + \langle a\rangle, d + \langle b\rangle)$ para todo $d \in F[x]$. Es fácil ver que ϕ es un morfismo de $F[x]$-módulos y que $\mathrm{Ker}(\phi) = \langle a\rangle \cap \langle b\rangle$. Además, como a y b son coprimos, tenemos que existen $r, s \in F[x]$ tales que

$$1 = ra + sb.$$

Por lo tanto, dado $(d + \langle a\rangle, d' + \langle b\rangle) \in F[x]/\langle a\rangle \bigoplus F[x]/\langle b\rangle$, tenemos que $d = dra + dsb$ y $d' = d'ra + d'sb$. Por lo tanto, considerando $dsb + d'ra \in F[x]$ tenemos que

$$\begin{aligned}
\phi(dsb + d'ra) &= \Big((dsb + d'ra) + \langle a\rangle, (dsb + d'ra) + \langle b\rangle\Big) \\
&= \Big(dsb + \langle a\rangle, d'ra + \langle b\rangle\Big) \\
&= \Big((d - dra) + \langle a\rangle, (d' - d'sb) + \langle b\rangle\Big) \\
&= (d + \langle a\rangle, d' + \langle b\rangle).
\end{aligned}$$

Por lo tanto, ϕ es suprayectiva y por el Primer Teorema de Isomorfismo, tenemos que

$$F[x]/\langle a\rangle \cap \langle b\rangle \cong F[x]/\langle a\rangle \bigoplus F[x]/\langle b\rangle.$$

Como a y b son coprimos se sigue que $\langle a\rangle \cap \langle b\rangle = \langle ab\rangle$. Probándose el isomorfismo $(*)$.
Ahora, como $q_j := \prod_{i=1}^{n} p_i^{\beta_{i,j}}$, donde los p_i son primos no asociados entre sí, tenemos que $p_i^{\beta_{i,j}}$ es coprimo con $p_l^{\beta_{l,j}}$ si $i \neq l$. De donde, utilizando $(*)$ podemos argumentar inductivamente para ver que

$$F[x]/\langle q_j\rangle \cong \bigoplus_{i=1}^{n} F[x]/\langle p_i^{\beta_{i,j}}\rangle.$$

Notemos que si algún $\beta_{i,j} = 0$ para algún j, entonces $F[x]/\langle p_i^{\beta_{i,j}}\rangle = 0$.
Ahora, como $V_f \cong \bigoplus_{i=1}^{n}\bigoplus_{j=1}^{k_i} F[x]/\langle p_i^{\beta_{i,j}}\rangle$ y haciendo la suma directa por columnas en el arreglo de la figura 2.1, tenemos que

$$V_f \cong \bigoplus_{j=1}^{t} F[x]/\langle q_j\rangle.$$

Notemos que como $q_j = \prod_{i=1}^{n} p_i^{\beta_{i,j}}$ y $q_{j+1} = \prod_{i=1}^{n} p_i^{\beta_{i,j+1}}$ y cada $\beta_{i,j} \geq \beta_{i,j+1}$, entonces $q_{j+1} \mid q_j$.
Luego, existen subespacios f-invariantes V_j con $V_j \cong F[x]/\langle q_j\rangle$ tales que $V = \bigoplus_{j=1}^{t} V_j$. Veamos que el polinomio mínimo de $f_j := f|_{V_j} : V_j \longrightarrow V_j$ es q_j. Para esto es suficiente ver que

$$\mathrm{Ann}(F[x]/\langle q_j\rangle) = \langle q_j\rangle.$$

En efecto,

$$\begin{aligned}\mathrm{Ann}(F[x]/\langle q_j\rangle) = \mathrm{Ann}\Big(\bigoplus_{i=1}^{n} F[x]/\langle p_i^{\beta_{i,j}}\rangle\Big) = \bigcap_{i=1}^{n}\mathrm{Ann}\Big(F[x]/\langle p_i^{\beta_{i,j}}\rangle\Big) &= \bigcap_{i=1}^{n}\langle p_i^{\beta_{i,j}}\rangle \\ &= \langle\prod_{i=1}^{n} p_i^{\beta_{i,j}}\rangle \\ &= \langle q_j\rangle.\end{aligned}$$

Por la observación 1.4.20, tenemos que $F[x]/\langle q_j\rangle$ es cíclico como $F[x]$-módulo y así, por el teorema 1.4.7, existe una base B_j de V_j tal que la matriz de f_j es la matriz compañera de q_j. Luego $B = \bigcup_{j=1}^{t} B_j$ es una base de V tal que la matriz de f tiene la forma

$$A = \begin{pmatrix} [q_1] & & & \\ & [q_2] & & \\ & & \ddots & \\ & & & [q_t] \end{pmatrix}$$

donde cada $[q_i]$ es la matriz compañera de $q_i \in F[x]$ y además $q_{i+1} \mid q_i$ para $i = 1, ..., t-1$.
Probándose la existencia de la forma canónica de factores invariantes de f. □

Teorema 2.2.4. *Sea V un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación lineal. Si A y B son dos forma canónicas racionales de factores invariantes de f, entonces $A = B$.*

Demostración. Sean

$$A = \begin{pmatrix} \boxed{[q_1]} & & \\ & \boxed{[q_2]} & \\ & & \ddots \\ & & & \boxed{[q_t]} \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} \boxed{[p_1]} & & \\ & \boxed{[p_2]} & \\ & & \ddots \\ & & & \boxed{[p_s]} \end{pmatrix}$$

dos formas canónicas de factores invariantes de f. Entonces tenemos las siguientes dos cadenas de ideales

$$\langle q_1 \rangle \subseteq \langle q_2 \rangle \subseteq \cdots \subseteq \langle q_t \rangle \subseteq F[x]$$
$$\langle p_1 \rangle \subseteq \langle p_2 \rangle \subseteq \cdots \subseteq \langle p_s \rangle \subseteq F[x].$$

Como $V \cong \bigoplus_{i=1}^{t} F[x]/\langle q_i \rangle$ y $V \cong \bigoplus_{i=1}^{s} F[x]/\langle p_i \rangle$. Tenemos que

$$\mathrm{Ann}(V) \cong \mathrm{Ann}\Big(\bigoplus_{i=1}^{t} F[x]/\langle q_i \rangle\Big) = \bigcap_{i=1}^{t} \mathrm{Ann}\Big(F[x]/\langle q_i \rangle\Big) = \bigcap_{i=1}^{t} \langle q_i \rangle = \langle q_1 \rangle.$$

De manera similar, $\mathrm{Ann}(V) \cong \mathrm{Ann}\Big(\bigoplus_{i=1}^{s} F[x]/\langle p_i \rangle\Big) = \langle p_1 \rangle$ y así concluimos que $\langle p_1 \rangle = \langle q_1 \rangle$. Luego, tenemos que:

$$F[x]/\langle q_1 \rangle \bigoplus \Big(\bigoplus_{i=2}^{t} F[x]/\langle q_i \rangle\Big) \cong F[x]/\langle p_1 \rangle \bigoplus \Big(\bigoplus_{i=2}^{s} F[x]/\langle p_i \rangle\Big).$$

Por la proposición 1.3.12, podemos cancelar $F[x]/\langle q_1 \rangle$ para obtener que

$$\bigoplus_{i=2}^{t} F[x]/\langle q_i \rangle \cong \bigoplus_{i=2}^{s} F[x]/\langle p_i \rangle.$$

Luego, podemos repetir el proceso previo con el $F[x]$-módulo anterior para concluir inductivamente que $t = s$ y $\langle q_i \rangle = \langle p_i \rangle$ para todo i. De esta manera como los polinomios p_i y q_i son mónicos, concluimos que $p_i = q_i$ para todo i, probándose la unicidad. □

2.3. Ejemplos

Ahora daremos algunos ejemplos del cálculo de las diferentes formas canónicas, sin embargo, debido a la naturaleza puramente teórica de los teoremas de existencia, necesitaremos los siguientes resultados.

Lema 2.3.1. *Sean V un F-espacio vectorial de dimensión finita y $f, g : V \longrightarrow V$ transformaciones lineales. Entonces existen bases β y γ de V tales que $[f]_\beta = [g]_\gamma$ si, y sólo si, existe un isomorfismo de espacios vectoriales $\theta : V \longrightarrow V$ tal que el siguiente diagrama conmuta*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle\theta}\downarrow & & \downarrow{\scriptstyle\theta} \\ V & \xrightarrow[g]{} & V \end{array}$$

donde $[f]_\beta$ y $[g]_\gamma$ denotan las matrices de f y g respecto de β y γ, respectivamente.

Demostración. Sea β una base de n elementos de V y consideremos la transformación lineal tomar coordenadas $\phi_\beta : V \longrightarrow F^n$. Recordemos que $A := [f]_\beta$ es la única matriz tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle \phi_\beta}\downarrow & & \downarrow{\scriptstyle \phi_\beta} \\ F^n & \xrightarrow[L_A]{} & F^n \end{array}$$

donde L_A es la transformación lineal tal que $L_A(x) = Ax$ para todo $x \in F^n$.
Ahora, sea $g : V \longrightarrow V$ tal que existe una base γ de V tal que la matriz asociada respecto de esta base es A. Es decir, tenemos un isomorfismo $\phi_\gamma : V \longrightarrow F^n$ tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} V & \xrightarrow{g} & V \\ {\scriptstyle \phi_\gamma}\downarrow & & \downarrow{\scriptstyle \phi_\gamma} \\ F^n & \xrightarrow[L_A]{} & F^n \end{array}$$

Luego, podemos formar el isomorfismo $\phi_\gamma^{-1}\phi_\beta : V \longrightarrow V$ tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle \phi_\gamma^{-1}\phi_\beta}\downarrow & & \downarrow{\scriptstyle \phi_\gamma^{-1}\phi_\beta} \\ V & \xrightarrow[g]{} & V. \end{array}$$

Recíprocamente, supongamos que existe un isomorfismo de espacios vectoriales $\theta : V \longrightarrow V$ tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle \theta}\downarrow & & \downarrow{\scriptstyle \theta} \\ V & \xrightarrow[g]{} & V. \end{array}$$

Sea γ una base de V tal que la matriz de g respecto de esta base es B, es decir, $B := [g]_\gamma$. Luego podemos formar el siguiente diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle \theta}\downarrow & & \downarrow{\scriptstyle \theta} \\ V & \xrightarrow[g]{} & V \\ {\scriptstyle \phi_\gamma}\downarrow & & \downarrow{\scriptstyle \phi_\gamma} \\ F^n & \xrightarrow{L_B} & F^n. \end{array}$$

Consideremos la base canónica $\{e_i\}_{i=1}^n$ de F^n y consideremos

$$v_i := \theta^{-1}(\phi_\gamma^{-1}(e_i)) \in V.$$

Luego, tenemos que $\beta = \{v_1, ..., v_n\}$ es una base de V y además se satisface que $(\phi_\gamma \circ \theta)(v_i) = e_i$ para todo $i = 1, ..., n$. Por lo tanto, concluimos que el morfismo tomar coordenadas respecto de β coincide con $\phi_\gamma \circ \theta$, es decir, $\phi_\beta = \phi_\gamma \circ \theta$. Por lo tanto, del diagrama de arriba, concluimos que $B = [f]_\beta$. □

Proposición 2.3.2. *Sean $f, g : V \longrightarrow V$ transformaciones lineales entre espacios vectoriales de dimensión finita sobre un campo F. Entonces f y g son similares (i.e., existe un isomorfismo de espacios vectoriales $\phi : V \longrightarrow V$ tal que $f = \phi g \phi^{-1}$) si, y sólo si, $V_f \cong V_g$.*

Demostración. Supongamos que f y g son similares. Entonces existe $\phi : V \longrightarrow V$ un isomorfismo de espacios vectoriales tal que $f = \phi g \phi^{-1}$. Probaremos que ϕ es un isomorfismo de $F[x]$-módulos entre V_f y V_g. Basta probar que saca escalares. De la igualdad $f = \phi g \phi^{-1}$, tenemos que $f\phi = \phi g$, lo que nos dice que hay una especie de conmutatividad. Igual que con la conmutatividad, aplicando inducción, se tiene que $f^n \phi = \phi g^n$, $n \in \mathbb{N}$. Sean $v \in V_g$ y $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in F[x]$. De lo anterior se sigue que

$$\begin{aligned} \phi(p \cdot_g v) &= \phi(a_0 v + a_1 g(v) + \cdots + a_n g^n(v)) \\ &= a_0\phi(v) + a_1\phi g(v) + \cdots + a_n \phi g^n(v) \\ &= a_0\phi(v) + a_1 f\phi(v) + \cdots + a_n f^n \phi(v) \\ &= p \cdot_f \phi(v). \end{aligned}$$

Por lo tanto, $V_f \cong V_g$.
Supongamos que existe un $F[x]$-isomorfismo $\phi : V_g \longrightarrow V_f$ (en particular, ϕ es un F-isomorfismo). Esto implica que, $\forall v \in V_g$,

$$\phi g(v) = \phi(x \cdot_g v) = x \cdot_f \phi(v) = f\phi(v).$$

En consecuencia, $\phi g = f\phi$ y f y g son similares. □

Teorema 2.3.3. *Dado $g(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n \in F[x]$, se tiene la siguiente igualdad,*

$$(-1)^n g = \det([g] - xId),$$

donde Id es la matriz identidad y $[g]$ es la matriz compañera de g

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Demostración. El caso $n = 1$ es trivial. Supongamos que el resultado se cumple para $n-1$, con $n \geq 2$. Realizando el cálculo del determinante por menores a través de la primera

columna se consigue lo siguiente,

$$\begin{aligned}
\det([g]-xId) &= \det\begin{pmatrix} -x & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & -x & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & -x & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}-x \end{pmatrix} \\
&= -x\det\begin{pmatrix} -x & 0 & 0 & \cdots & 0 & -a_1 \\ 1 & -x & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & -x & \cdots & 0 & -a_3 \\ 0 & 0 & 1 & \cdots & 0 & -a_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}-x \end{pmatrix} \\
&\quad -\det\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & -x & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & -x & \cdots & 0 & -a_3 \\ 0 & 0 & 1 & \cdots & 0 & -a_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}-x \end{pmatrix} \\
&= -x(-1)^{n-1}(a_1+a_2x+...+a_{n-1}x^{n-2}+x^{n-1})+(-1)^{n+1}(-a_0) \\
&= (-1)^n g(x),
\end{aligned}$$

donde la penúltima igualdad se da por la hipótesis inductiva aplicada a la primera matriz de tal desarrollo y calculando el determinante de la segunda matriz por el método de menores desarrollando por el primer renglón y del hecho de que el determinante de una matriz triangular es el producto de los elementos de la diagonal. □

Con ayuda de los resultados previos podemos probar el siguiente resultado.

Proposición 2.3.4. *Sea $p(x) \in F[x]$ el polinomio característico de $f : V \longrightarrow V$ una transformación lineal, con $V \neq 0$ y de dimensión finita sobre un campo F. Entonces p es producto de todos los factores invariantes de f, i.e., $p = (-1)^n q_1 \cdots q_t$, donde q_j es el j-ésimo factor invariante y $n = \dim(V)$.*

Demostración. Por el teorema 2.2.3 y la proposición 2.2.4, existe una única forma canónica racional de factores invariantes de f

$$A = \begin{pmatrix} [q_1] & & & \\ & [q_2] & & \\ & & \ddots & \\ & & & [q_m] \end{pmatrix}$$

donde cada $[q_i]$ es la matriz compañera de cierto $q_i \in F[x]$ mónico y además $q_{i+1} \mid q_i$ para

$i = 1, ..., m-1$. Es decir, existe una base β de V tal que la matriz asociada es A. Además

$$A - x\mathrm{Id} = \left(\begin{array}{c|c|c|c} [q_1] - x\mathrm{Id} & & & \\ \hline & [q_2] - x\mathrm{Id} & & \\ \hline & & \ddots & \\ \hline & & & [q_m] - x\mathrm{Id} \end{array}\right)$$

Sabemos de álgebra lineal que el polinomio característico $p(x)$ de f es el determinante de $f - x\mathrm{Id}_V$ respecto de cualquier base de V. Como $A - x\mathrm{Id} = [f - x\mathrm{Id}_V]_\beta$ y esta matriz está formada por matrices cuadradas más pequeñas, tenemos que:

$$p(x) = \det(A - x\mathrm{Id}) = \det([q_1] - x\mathrm{Id}) \cdots \det([q_n] - x\mathrm{Id}).$$

Ahora, como cada $[q_i]$ es la matriz compañera de q_i, por el teorema 2.3.3, tenemos que $\det([q_i] - x\mathrm{Id}) = (-1)^{\mathrm{grad}(q_i)} q_i$ para todo $i = 1, ..., n$. Por lo tanto, $p(x) = (-1)^n q_1 \cdots q_n$, donde $n = \sum_{i=1}^{m} \mathrm{grad}(q_i) = \dim(V)$. □

La última proposición que necesitamos afirma que el polinomio mínimo es un factor invariante. Su demostración esencialmente está en la prueba de la proposición 2.2.4, pero lo damos aparte para conveniencia del lector.

Lema 2.3.5. *Sean $V \neq 0$ un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación lineal. Sean $q_1, ..., q_m$ los factores invariantes de f, entonces q_1 es el polinomio mínimo de f.*

Demostración. Como $q_1, ..., q_m$ son los factores invariantes de f tenemos que

$$V_f \cong \bigoplus_{i=1}^{m} F[x]/\langle q_i \rangle$$

con $\langle q_1 \rangle \subseteq \cdots \subseteq \langle q_m \rangle$. Sea g el polinomio mínimo de f, sabemos por definición que $\langle g \rangle = \mathrm{Ann}(V_f)$. Por otro lado, notemos que

$$\mathrm{Ann}(V_f) \cong \mathrm{Ann}\Big(\bigoplus_{i=1}^{m} F[x]/\langle q_i \rangle\Big) = \bigcap_{i=1}^{m} \mathrm{Ann}\Big(F[x]/\langle q_i \rangle\Big) = \bigcap_{i=1}^{m} \langle q_i \rangle = \langle q_1 \rangle.$$

Como q_1 y g son mónicos, tenemos que $g = q_1$. □

Observación 2.3.6. *Sean V un F-espacio vectorial de dimensión finita y $f : V \longrightarrow V$ una transformación lineal. Notemos que, del lema anterior tenemos que, todos los factores invariantes de f son divisores del polinomio mínimo de f.*

Con estos resultados ya tenemos todas las herramientas necesarias para hacer unos ejemplos.

Ejemplo 2.3.7. *Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la transformación lineal dada por la siguiente matriz,*

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}.$$

Su polinomio característico está dado por

$$\begin{aligned} p(x) &= \det(A - xId) \\ &= \det\begin{pmatrix} 1-x & 1 & 3 \\ 5 & 2-x & 6 \\ -2 & -1 & -3-x \end{pmatrix} \\ &= (1-x)(-(2-x)(3+x)+6) - (-5(3+x)+12) + 3(-5+2(2-x)) \\ &= -(1-x)(2-x)(3+x) + 6 - 7x \\ &= -x^3. \end{aligned}$$

Sabemos que el polinomio mínimo tiene las mismas raíces que el polinomio característico. Si el polinomio mínimo fuera $g(x) = x$, entonces $g(f) = g(A) = A = 0$, una contradicción. Si $g(x) = x^2$, entonces $g(f) = g(A) = A^2 = 0$, pero la entrada inferior derecha de A^2 es -3. Por lo tanto, $g(x) = x^3$, ya que $g \mid p$. La proposición 2.3.4 y el lema 2.3.5 aseguran que el único factor invariante de f es $g(x) = x^3$. Por cómo se construyeron los factores invariantes a través de los divisores elementales en la prueba del teorema 2.2.3 y del arreglo de la figura 2.1, se deduce que el único divisor elemental de f es $g(x) = x^3$. De las definiciones de las formas canónicas racionales, se tiene que la forma de divisores elementales y la de factores invariantes coinciden y son de la siguiente forma,

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

que, a su vez, es la forma canónica de Jordan.

Ejemplo 2.3.8. *Para este ejemplo, la transformación lineal $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ estará dada por la siguiente matriz,*

$$B = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}.$$

Su polinomio característico es

$$\begin{aligned} p(x) &= \det\begin{pmatrix} 2-x & 2 & 1 \\ 2 & 2-x & 1 \\ 2 & 2 & 1-x \end{pmatrix} \\ &= (2-x)((2-x)(1-x)-2) - 2(2(1-x)-2) + 4 - 2(2-x) \\ &= (2-x)(-3x+x^2) + 6x \\ &= -x^3 + 5x^2 \\ &= -x^2(x-5). \end{aligned}$$

Dado que el polinomio mínimo y el polinomio característico tienen las mismas raíces, se obtienen dos casos, el polinomio mínimo es $g(x) = x(x-5)$ o $g(x) = -p(x)$. Para poder determinar en cuál de los dos casos nos encontramos hagamos el siguiente cálculo sencillo.

$$\begin{aligned} B(B-5Id) &= \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 2 & 2 & -4 \end{pmatrix} \\ &= 0. \end{aligned}$$

Lo que implica que $x(x-5) \in \mathrm{Ann}(\mathbb{R}^3_f)$ *y, como* $x(x-5)$ *es el polinomio de menor grado de entre las dos opciones que tenemos, se concluye que* $g(x) = x(x-5)$. *La proposición 2.3.4 y el lema 2.3.5 afirman que los factores invariantes de* f *son* $x(x-5)$ *y* x. *En este caso, el arreglo de divisores elementales de* f *de la figura 2.1 es de la forma*

$$\begin{array}{cc} x & x \\ x-5 & 1 \end{array}$$

Y la forma canónica racional de factores invariantes es

$$\left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 5 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

donde $\begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}$ *es la matriz compañera del factor invariante* $q_1 = x(x-5)$ *y* (0) *es la matriz compañera del factor invariante* $q_2 = x$. *Del arreglo de arriba obtenido de la figura 2.1, concluimos que la forma canónica racional de divisores elementales* f *es de la forma*

$$\left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 5 \end{array}\right)$$

donde $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ *es la matriz* x*-racional de la definición 2.1.3 asociada al primo* $x \in F[x]$ *(tiene dos bloques de tamaño* 1×1*) y* (5) *es la matriz* $(x-5)$*-racional de la definición 2.1.3 asociada al primo* $x - 5 \in F[x]$. *Como todos los primos considerados son de grado uno y las potencias de los primos en el arreglo de arriba de los divisores elementales son todos* 1*, tenemos que la forma canónica racional de divisores elementales coincide con la forma canónica de Jordan (pues en este caso se tiene la siguiente igualdad de matrices* $[p_i]^r = [p_i]_r$ *si* $r = 1$*, de acuerdo a la notación 1.4.11).*

Ejemplo 2.3.9. *En este ejemplo trabajaremos con la transformación lineal* $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ *dada por la matriz*

$$C = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

El polinomio característico de f *es*

$$\begin{aligned} p(x) &= \det \begin{pmatrix} -x & 0 & 1 & 1 \\ 0 & -x & 1 & 1 \\ 1 & 1 & -x & 0 \\ 1 & 1 & 0 & -x \end{pmatrix} \\ &= -x \det \begin{pmatrix} -x & 1 & 1 \\ 1 & -x & 0 \\ 1 & 0 & -x \end{pmatrix} + \det \begin{pmatrix} 0 & -x & 1 \\ 1 & 1 & 0 \\ 1 & 1 & -x \end{pmatrix} - \det \begin{pmatrix} 0 & -x & 1 \\ 1 & 1 & -x \\ 1 & 1 & 0 \end{pmatrix} \\ &= -x(-x^3 + x + x) - x^2 - x^2 \\ &= x^4 - 4x^2 = x^2(x+2)(x-2). \end{aligned}$$

Dado que el polinomio mínimo y el polinomio característico tienen las mismas raíces, se sigue que el polinomio mínimo es $g = x(x+2)(x-2)$ *o* $g = p$. *Para descartar una de las opciones hagamos el siguiente cálculo.*

$$\begin{aligned} C(C+2Id)(C-2Id) &= C(C^2-4Id) \\ &= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -2 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 2 & -2 \end{pmatrix} \\ &= 0. \end{aligned}$$

Por lo tanto, $g = x(x+2)(x-2)$. *La proposición 2.3.4 y el lema 2.3.5 afirman que los factores invariantes de* f *son* $x(x+2)(x-2)$ *y* x. *En este caso, el arreglo de divisores elementales de* f *de la figura 2.1 es de la forma*

$$\begin{array}{cc} x & x \\ x+2 & 1 \\ x-2 & 1 \end{array}$$

Como $q_1 = x(x-2)(x+2) = x(x^2-4) = x^3 - 4x$ *la matriz compañera de* q_1 *es de la forma*

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 4 \\ 0 & 1 & 0 \end{pmatrix}.$$

Luego, la forma canónica racional de factores invariantes es

$$\left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

donde (0) *es la matriz compañera del factor invariante* $q_2 = x$.
Por otro lado, la forma canónica racional de divisores elementales es

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

donde $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ *es la matriz* x*-racional de la definición 2.1.3 asociada al primo* $x \in F[x]$, *la cual tiene dos bloques de tamaño* 1×1, (-2) *es la matriz* $(x+2)$*-racional de la definición 2.1.3 asociada al primo* $(x+2) \in F[x]$ *y* (2) *es la matriz* $(x-2)$*-racional de la definición 2.1.3 asociada al primo* $(x-2) \in F[x]$. *Como todos los primos considerados son de grado uno y las potencias de los primos en el arreglo de arriba de los divisores elementales son todos* 1, *tenemos que la forma canónica racional de divisores elementales coincide con la forma canónica de Jordan (pues en este caso* $[p_i]^r = [p_i]_r$ *si* $r = 1$, *de acuerdo a la notación 1.4.11).*

En los ejemplos anteriores la forma de Jordan coincide con la de divisores elementales, veamos ahora un ejemplo en el que las tres formas canónicas son distintas. Para ello utilizaremos el campo $\mathbb{C}$, en lugar de $\mathbb{R}$.

Ejemplo 2.3.10. *Consideremos la transformación lineal $f : \mathbb{C}^5 \longrightarrow \mathbb{C}^5$ dada por la siguiente matriz,*

$$D = \begin{pmatrix} -i & 0 & 0 & 0 & 0 \\ 0 & -i & i & 1-i & -1+i \\ 0 & 0 & 1 & 2i & -2i \\ 0 & 0 & 1 & -1+2i & 2-2i \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

El polinomio característico de D es

$$\begin{aligned} p(x) &= \det \begin{pmatrix} -i-x & 0 & 0 & 0 & 0 \\ 0 & -i-x & i & 1-i & -1+i \\ 0 & 0 & 1-x & 2i & -2i \\ 0 & 0 & 1 & -1+2i-x & 2-2i \\ 0 & 0 & 0 & 0 & 1-x \end{pmatrix} \\ &= -(x+i)\cdot \det \begin{pmatrix} -i-x & i & 1-i & -1+i \\ 0 & 1-x & 2i & -2i \\ 0 & 1 & -1+2i-x & 2-2i \\ 0 & 0 & 0 & 1-x \end{pmatrix} \\ &= (x+i)(x-1)\cdot \det \begin{pmatrix} -i-x & i & 1-i \\ 0 & 1-x & 2i \\ 0 & 1 & -1+2i-x \end{pmatrix} \\ &= -(x+i)^2(x-1)\cdot \det \begin{pmatrix} 1-x & 2i \\ 1 & -1+2i-x \end{pmatrix} \\ &= -(x+i)^2(x-1)((1-x)(-1+2i-x)-2i) \\ &= -(x+i)^2(x-1)(-1-2ix+x^2) \\ &= -(x-i)^2(x+i)^2(x-1) \end{aligned}$$

Dado que el polinomio característico y el polinomio mínimo comparten raíces, tenemos cuatro opciones para el polinomio mínimo, a saber,

$$\begin{aligned} g(x) &= (x-i)(x+i)(x-1) \ o \\ g(x) &= (x-i)^2(x+i)(x-1) \ o \\ g(x) &= (x-i)(x+i)^2(x-1) \ o \\ g(x) &= (x-i)^2(x+i)^2(x-1) \end{aligned}$$

Empecemos con el siguiente cálculo,

$$
\begin{aligned}
(D - iId) &= \begin{pmatrix} -2i & 0 & 0 & 0 & 0 \\ 0 & -2i & i & 1-i & -1+i \\ 0 & 0 & 1-i & 2i & -2i \\ 0 & 0 & 1 & -1+i & 2-2i \\ 0 & 0 & 0 & 0 & 1-i \end{pmatrix}, \\
(D + iId) &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 1-i & -1+i \\ 0 & 0 & 1+i & 2i & -2i \\ 0 & 0 & 1 & -1+3i & 2-2i \\ 0 & 0 & 0 & 0 & 1+i \end{pmatrix}, \\
(D - Id) &= \begin{pmatrix} -i-1 & 0 & 0 & 0 & 0 \\ 0 & -i-1 & i & 1-i & -1+i \\ 0 & 0 & 0 & 2i & -2i \\ 0 & 0 & 1 & -2+2i & 2-2i \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \\
(D - iId)(D + iId)(D - Id) &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2+2i & -4i & 4i \\ 0 & 0 & -4 & 4-4i & -4+4i \\ 0 & 0 & -2-2i & 4 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \neq 0.
\end{aligned}
$$

Por lo tanto, $g \neq (x-i)(x+i)(x-1)$. *Ahora,*

$$(D-i)^2(D+i)(D-1) = 0,$$

de modo que

$$(x-i)(x+i)(x-1) \mid g \mid (x-i)^2(x+i)(x-1) \quad y \quad g \neq (x-i)(x+i)(x-1).$$

Es decir, $g(x) = (x-i)^2(x+i)(x-1)$.
La proposición 2.3.4 y el lema 2.3.5 nos aseguran que los factores invariantes son $(x-i)^2(x+i)(x-1)$ *y* $x+i$*; en este caso, el arreglo de divisores elementales de* f *de la figura 2.1 es de la forma*

$$\begin{matrix} x+i & x+i \\ (x-i)^2 & \\ x-1 & \end{matrix}$$

Como $q_1 = (x-i)^2(x+i)(x-1) = x^4 - (1+i)x^3 + (1+i)x^2 - (1+i)x + i$, *tenemos que la matriz compañera del factor invariante* q_1 *es*

$$\begin{pmatrix} 0 & 0 & 0 & -i \\ 1 & 0 & 0 & 1+i \\ 0 & 1 & 0 & -1-i \\ 0 & 0 & 1 & 1+i \end{pmatrix}.$$

Por lo tanto, la forma racional de factores invariantes es

$$\left(\begin{array}{cccc|c} 0 & 0 & 0 & -i & 0 \\ 1 & 0 & 0 & 1+i & 0 \\ 0 & 1 & 0 & -1-i & 0 \\ 0 & 0 & 1 & 1+i & \\ \hline 0 & 0 & 0 & 0 & -i \end{array}\right)$$

donde $(-i)$ *es la matriz compañera del factor invariante* $q_2 = x + i$.
Por otro lado, la matriz de divisores elementales es

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

donde $\begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}$ *es la matriz* $(x+i)$*-racional de la definición 2.1.3 asociada al primo* $(x+i) \in \mathbb{C}[x]$ *(tiene dos bloques de tamaño* 1×1*), y como* $(x-i)^2 = x^2 - 2ix - 1$*, se tiene que* $\begin{pmatrix} 0 & 1 \\ 1 & 2i \end{pmatrix}$ *es la matriz* $(x-i)$*-racional de la definición 2.1.3 asociada al primo* $(x-i) \in \mathbb{C}[x]$ *y* (1) *es la matriz* $(x-1)$*-racional de la definición 2.1.3 asociada al primo* $(x-1) \in \mathbb{C}[x]$.
Ahora, para calcular la forma canónica de Jordan (por la construcción en el teorema 1.4.21) tenemos que en lugar de calcular la matriz compañera de $(x-i)^2$ *calcular la matriz* $[(x-i)]_2$ *de la notación 1.4.11. En este caso la matriz* $[(x-i)]_2$ *es* $\begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$ *y por lo tanto la forma canónica de Jordan es*

$$J = \left(\begin{array}{c|c|cc|c} -i & 0 & 0 & 0 & 0 \\ \hline 0 & -i & 0 & 0 & 0 \\ \hline 0 & 0 & i & 0 & 0 \\ 0 & 0 & 1 & i & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Capítulo 3

Cálculo de formas canónicas y aplicaciones

3.1. Diagrama de puntos

En los ejemplos del capítulo anterior pudimos calcular los factores invariantes y los divisores elementales debido a que el polinomio característico era de grado bajo. En esta sección expondremos otro método utilizando los diagramas de puntos y veremos qué relación tiene con la teoría de módulos desarrollada anteriormente (véase la discusión después del teorema 3.1.10).
Empecemos con las definiciones y resultados necesarios para entender el diagrama de puntos.

Definición 3.1.1. *Sean $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F y $\lambda \in F$. Decimos que $x \in V \setminus \{0\}$ es un* ***vector propio generalizado de f correspondiente a λ*** *si existe $0 < p \in \mathbb{N}$ tal que*

$$(f - \lambda \cdot \mathrm{Id})^p(x) = 0.$$

Definición 3.1.2. *Sean $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F y λ un valor propio de f. El* ***espacio propio generalizado de f correspondiente a λ*** *está definido como*

$$K_\lambda := \{x \in V : (f - \lambda \cdot \mathrm{Id})^p(x) = 0, \text{ para algún } 0 < p \in \mathbb{N}\}.$$

Con estas definiciones podemos enunciar el siguiente teorema, que será de ayuda para calcular el número de puntos en los diagramas.

Teorema 3.1.3. *Sea $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F, cuyo polinomio característico se factoriza como producto de factores lineales $(x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$. Entonces $\dim(K_{\lambda_i}) = m_i$.*

Demostración. Véase [4], p. 480. □

Para empezar con los diagramas de puntos requerimos la siguiente definición y tres teoremas.

Definición 3.1.4. *Sean $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F y x un vector propio generalizado asociado a λ. Si $0 < p \in \mathbb{N}$ es el mínimo tal que $(f - \lambda \cdot \mathrm{Id})^p(x) = 0$, entonces decimos que*

$$\{(f - \lambda \cdot \mathrm{Id})^{p-1}(x), ..., (f - \lambda \cdot \mathrm{Id})(x), x\}$$

*es un **ciclo de vectores propios generalizados de f correspondiente a λ**. El **vector inicial** del ciclo es $(f - \lambda \cdot \mathrm{Id})^{p-1}(x)$ y el **vector final**, x. En este caso el ciclo es de longitud p.*

Teorema 3.1.5. *Sea $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F, cuyo polinomio característico se factoriza como producto de factores lineales. Sea β una base para V compuesta por la unión disjunta de ciclos de vectores propios generalizados de f (ordenados como en la definición 3.1.4). Entonces*

a) La transformación f se restringe bien a cualquier ciclo contenido en β y la representación matricial de las restricciones es una matriz de la forma

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}.$$

b) La representación matricial de f con respecto a β es de la forma

$$[T]_\beta = \begin{pmatrix} J_{\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2} & \dots & 0 \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & J_{\lambda_k} \end{pmatrix}$$

donde J_{λ_i} es una matriz cuadrada (de 1×1) de la forma (λ_i) o bien de una matriz cuadrada de $n \times n$ ($n > 1$) de la forma J_{λ_i} del inciso (a), para algún valor propio λ_i de T.

Demostración. Véase [4], p. 482. □

La siguiente proposición nos da una propiedad importante de los ciclos de vectores propios generalizados.

Proposición 3.1.6. *Cualquier ciclo de vectores propios generalizados es linealmente independiente.*

Demostración. Véase [4], p. 483. □

Teorema 3.1.7. *Sean $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F y λ un valor propio de f. Entonces K_λ tiene una base compuesta por la unión disjunta de ciclos de vectores propios generalizados asociados a λ.*

Demostración. Véase [4], p. 484. □

Teorema 3.1.8. *Sea $f : V \longrightarrow V$ una transformación lineal entre espacios vectoriales de dimensión finita sobre un campo F, cuyo polinomio característico se factoriza como producto de factores lineales. Entonces*

$$V = \bigoplus_{i=1}^{k} K_{\lambda_i},$$

donde $\lambda_1, ..., \lambda_k$ son los distintos valores propios de f.

Demostración. Véase [4], p. 487. □

Observación 3.1.9. *Por el teorema 3.1.5, si β es una base para V compuesta por la unión disjunta de ciclos de vectores propios generalizados de f, ordenados de la siguiente manera*

$$\{(f - \lambda \cdot \mathrm{Id})^{p-1}(x), ..., (f - \lambda \cdot \mathrm{Id})(x), x\},$$

entonces la matriz asociada a T está formada por bloques de la forma

$$\begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & & \ddots & 1 \\ & & & & & \lambda \end{pmatrix}.$$

Sin embargo, para ir acorde con los bloques de Jordan como lo estamos manejando en este trabajo, los ciclos de β los ordenaremos de la siguiente manera

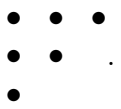
$$\{x, (f - \lambda \cdot \mathrm{Id})(x), ..., (f - \lambda \cdot \mathrm{Id})^{p-1}(x)\},$$

y así obtendremos el bloque de Jordan (λ-matriz elemental de Jordan, ver definición 1.4.14) de la forma

$$[x - \lambda]_p = \begin{pmatrix} \lambda & & & & & \\ 1 & \lambda & & & & \\ & 1 & \lambda & & & \\ & & & & \ddots & \\ & & & & 1 & \lambda \end{pmatrix}.$$

Luego, en todo lo que sigue, las construcciones presentadas en [4] para crear el diagrama de puntos no cambian. Lo único que cambiaremos es el orden de los vectores en los ciclos de la base ordenada β al momento de obtener la matriz asociada a T respecto a esta base.

A cada K_λ le vamos a asociar un diagrama de puntos. Por el teorema 3.1.7, sabemos que existe una base β para K_λ compuesta por la unión disjunta de ciclos de vectores propios generalizados asociados a λ, digamos $\gamma_1, ..., \gamma_{n_i}$, respectivamente de longitud $p_1, ..., p_{n_i}$. El diagrama tendrá $\dim(K_\lambda)$ puntos, agrupados por columnas. La i-ésima columna corresponde a los vectores del i-ésimo ciclo, ordenados de arriba para abajo empezando por el vector inicial del ciclo y terminando con el final. Los ciclos están ordenados por longitud, del más grande al más pequeño.
Si tuviéramos tres ciclos, a saber, $\{(f-\lambda\cdot \mathrm{Id})^2(v_1), (f-\lambda\cdot \mathrm{Id})(v_1), v_1\}$, $\{(f-\lambda\cdot \mathrm{Id})(v_2), v_2\}$ y $\{v_3\}$, entonces el diagrama de puntos de K_λ se vería de la siguiente manera,



Los teoremas 3.1.5, 3.1.7 y 3.1.8 nos permiten encontrar una base para cada K_λ compuesta por la unión disjunta de ciclos de vectores propios generalizados asociados a λ, de tal forma que la unión sea una base para V y la representación matricial en esa base sea la forma canónica de Jordan. Si supiéramos cómo es el diagrama de puntos para cada K_λ, podríamos reconstruir las $(x-\lambda)$-matrices clásicas (ver definición 1.4.15) de las que se compone la forma de Jordan. Cabe notar que la unicidad de la forma canónica de Jordan nos asegura que el diagrama para K_λ es único, puesto que su estructura depende totalmente de la $(x-\lambda)$-matriz clásica (que es parte de la forma de Jordan). El siguiente teorema nos permite reconstruir los diagramas calculando el rango de ciertas matrices.

Teorema 3.1.10. *Sea r_j el número de puntos en el j-ésimo renglón del diagrama para K_λ. Entonces*

1. $r_1 = \dim(V) - \mathrm{rank}(f - \lambda \cdot \mathrm{Id})$.
2. $r_j = \mathrm{rank}((f - \lambda \cdot \mathrm{Id})^{j-1}) - \mathrm{rank}((f - \lambda \cdot \mathrm{Id})^j)$, *para* $j > 1$.

Demostración. Véase [4], p.493. □

Como se discutió anteriormente, cada ciclo γ_i en el diagrama de puntos de K_λ corresponde a una matriz como la del teorema 1.4.9. En la demostración de la unicidad de la forma canónica de Jordan se puede ver que el subespacio generado por los vectores de γ_i es isomorfo a $F[x]/\langle (x-\lambda)^{p_i}\rangle$. La suma de todos estos subespacios, variando sobre los valores propios y los ciclos de cada K_λ, nos da una descomposición de V_f en p-módulos cíclicos. Por los teoremas 1.3.6 y 1.3.10 sabemos que la descomposición de V_f es única y, por cómo están definidos los divisores elementales, se concluye que los divisores elementales asociados al primo $x-\lambda$ son $(x-\lambda)^{p_i}$, $i = 1, ..., n_i$.
Por lo tanto, conocer los diagramas de puntos para cada K_λ no nada más nos permite encontrar la forma de Jordan, sino también ambas formas racionales.
Ahora utilizaremos los diagramas de puntos para encontrar la forma canónica de Jordan y los divisores elementales de la transformación que aparece en el ejemplo 2.3.10.

Ejemplo 3.1.11. *Recordemos que la transformación f del ejemplo 2.3.10 tiene polinomio característico $p(x) = -(x-i)^2(x+i)^2(x-1)$. El teorema 3.1.3 afirma que los diagramas de*

K_i y K_{-i} tienen dos puntos, el de K_1, en cambio, sólo tiene un punto y, consecuentemente, la $(x-1)$-matriz clásica de la forma de Jordan de f es

$$(1).$$

Para construir los diagramas de K_i y K_{-i} aplicaremos el teorema 3.1.10. Tomando $\lambda = i$ se tiene que

$$\begin{aligned} r_1 &= \dim(\mathbb{C}^5) - \operatorname{rank}(f - i\cdot \mathrm{Id}) \\ &= 5 - \operatorname{rank}\begin{pmatrix} -2i & 0 & 0 & 0 & 0 \\ 0 & -2i & i & 1-i & -1+i \\ 0 & 0 & 1-i & 2i & -2i \\ 0 & 0 & 1 & -1+i & 2-2i \\ 0 & 0 & 0 & 0 & 1-i \end{pmatrix} \\ &= 5-4=1. \end{aligned}$$

De modo que el diagrama de K_i está dado por

$$\begin{matrix} \bullet \\ \bullet \end{matrix}$$

y, por ende, la $(x-i)$-matriz clásica de la forma de Jordan de f es

$$\begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}.$$

Para $\lambda = -i$ tenemos lo siguiente,

$$\begin{aligned} r_1 &= \dim(\mathbb{C}^5) - \operatorname{rank}(f + i\cdot \mathrm{Id}) \\ &= 5 - \operatorname{rank}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 1-i & -1+i \\ 0 & 0 & 1+i & 2i & -2i \\ 0 & 0 & 1 & -1+3i & 2-2i \\ 0 & 0 & 0 & 0 & 1+i \end{pmatrix} \\ &= 5-3=2. \end{aligned}$$

Por lo tanto, el diagrama de puntos para K_{-i} es de la forma

$$\bullet \quad \bullet$$

y la $(x+i)$-matriz clásica de la forma de Jordan de f es

$$\begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}.$$

Juntando las componentes de la forma de Jordan se consigue la forma canónica de Jordan de f, a saber,

$$\begin{pmatrix} -i & 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 & 0 \\ 0 & 0 & i & 0 & 0 \\ 0 & 0 & 1 & i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

La discusión anterior a este ejemplo nos asegura que los divisores elementales obtenidos por este método coinciden con los obtenidos en el ejemplo 2.3.10.

3.2. Descomposición de Jordan

Recordemos que la forma canónica de Jordan es una matriz diagonal con la excepción de algunos 1 debajo de la diagonal. Esto nos permite pensar a esta matriz como suma de una matriz diagonal y una matriz con unos o ceros debajo de la diagonal y ceros en lo demás. El teorema de descomposición de Jordan nos habla de esta característica y, además, nos da información de los sumandos en esta descomposición.
La demostración del teorema utiliza la noción de diagonalización simultánea, cuya definición se enuncia a continuación.

Definición 3.2.1. *Decimos que dos transformaciones $f, g : V \longrightarrow V$ diagonalizables son* ***simultáneamente diagonalizables*** *si existe una base de vectores propios de f que, a su vez, es una base de vectores propios de g.*

El siguiente teorema nos ayudará a demostrar el teorema de descomposición de Jordan.

Teorema 3.2.2. *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F y $f, g : V \longrightarrow V$ transformaciones diagonalizables. Entonces f y g son simultáneamente diagonalizables si, y sólo si, $f \circ g = g \circ f$.*

Demostración. Si f y g son simultáneamente diagonalizables, entonces existe una base de vectores propios de f y g, $\{v_1, ..., v_n\}$. Esto implica que

$$f(g(v_i)) = f(\lambda_i^g v_i) = \lambda_i^g \lambda_i^f v_i = \lambda_i^f \lambda_i^g v_i = g(f(v_i)),$$

donde $\lambda_i^g, \lambda_i^f \in F$ son los valores propios de g y f, respectivamente. Por ello $f \circ g = g \circ f$. Ahora supongamos que $f \circ g = g \circ f$. Al ser f diagonalizable, se tiene que V es suma directa de sus espacios propios,

$$V = \bigoplus_{i=1}^{n} E_{\lambda_i}.$$

La conmutatividad de f y g implica lo siguiente,

$$\forall v \in E_{\lambda_i},\ f(g(v)) = g(f(v)) = \lambda_i g(v).$$

Es decir, g se restringe bien a E_{λ_i}. Claramente, la restricción de g también es diagonalizable y, por ende, existe una base de vectores propios de la restricción para E_{λ_i}. Uniendo las bases para cada E_{λ_i}, obtenemos una base para V compuesta por vectores propios de f y g. □

Teorema 3.2.3 (Descomposición de Jordan)**.** *Sean $V \neq \{0\}$ un espacio vectorial de dimensión finita sobre un campo F, algebraicamente cerrado, y $f : V \longrightarrow V$ una transformación. Entonces existen una transformación diagonalizable $\delta : V \longrightarrow V$ y una transformación nilpotente $\eta : V \longrightarrow V$ (i.e., $\eta^k = 0$ para alguna $k \in \mathbb{N}$) tales que $f = \delta + \eta$ y $\delta \circ \eta = \eta \circ \delta$, δ y η son únicas con estas características. Además, existen $p, q \in F[x]$ tales que $\delta = p(f)$ y $\eta = q(f)$.*

Demostración. Sean $\lambda_1, ..., \lambda_n \in F$ los distintos valores propios de f. Por el teorema 3.1.8 podemos definir $\delta(v_1 + \cdots + v_n) := \lambda_1 v_1 + \cdots + \lambda_n v_n$, donde $v_i \in K_{\lambda_i}$. Esto implica que

los espacios propios de δ son los K_{λ_i} y, por el teorema 3.1.8, concluimos que δ es diagonalizable. Ahora, definimos $\eta := f - \delta$. Antes de probar que η es nilpotente observemos lo siguiente. En un ciclo de vectores propios generalizados de longitud k, cualquier elemento es anulado por $(f - \lambda \cdot \mathrm{Id})^k$. De modo que el teorema 3.1.7 nos asegura que cualquier elemento de K_{λ_i} es anulado por $(f - \lambda_i \cdot \mathrm{Id})^{p_i}$, donde p_i es el máximo de las longitudes de los ciclos de K_{λ_i}. Con esto en mente, notemos que

$$\forall v_i \in K_{\lambda_i},\ \eta^{p_i}(v_i) = (f(v_i) - \delta(v_i))^{p_i} = (f(v_i) - \lambda_i v_i)^{p_i} = 0.$$

Por lo tanto, $\eta^{\max\{p_i|\ i=1,\dots,n\}} = 0$.
Sea $m_f(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_n)^{m_n}$ el polinomio mínimo de f. Definimos

$$h_i(x) := \frac{m_f}{(x - \lambda_i)^{m_i}}.$$

Claramente, $\mathrm{mcd}\{h_1, ..., h_n\} = 1$, por ello existen $a_i \in F[x]$ tales que

$$a_1 h_1 + \cdots + a_n h_n = 1.$$

Por el teorema 3.1.3 y por la forma en la que se construyeron los diagramas de puntos se sigue que $(f - \lambda_i \cdot \mathrm{Id})^{m_i}(v_i) = 0$ para $v_i \in K_{\lambda_i}$, pues $\dim(K_{\lambda_i}) = m_i \geq p_i$. En consecuencia, $a_i h_i(f)(v_j) = 0$, con $i \neq j$ y $v_j \in K_{\lambda_j}$. Lo que implica que por la igualdad de arriba:

$$a_i h_i(f)(v_i) = a_1 h_1(f)(v_i) + \cdots + a_n h_n(f)(v_i) = v_i.$$

Por lo tanto,

$$\delta = \lambda_1 a_1 h_1(f) + \cdots + \lambda_n a_n h_n(f) = p(f) \text{ donde } p(x) = \sum_{i=1}^{n} \lambda_i a_i h_i \in F[x],$$

de donde $\eta = f - \delta = f - p(f) = q(f)$, con $q(x) = x - p(x) \in F[x]$.
La conmutatividad de δ y η se sigue del hecho de que ambas son polinomios en f. Supongamos que existen $\delta', \eta' : V \longrightarrow V$ tales que $f = \delta' + \eta'$ y $\delta' \circ \eta' = \eta' \circ \delta'$, con δ' diagonalizable y η' nilpotente. Las condiciones anteriores implican que δ' y f conmutan y, por ende, δ' conmuta con $\delta = p(f)$. Lo mismo se puede decir de η', i.e. $\eta' \circ \eta = \eta \circ \eta'$. El teorema del binomio asegura que

$$(\eta' - \eta)^t = \sum_{i=0}^{t} c_i \eta^{t-i} (-\eta')^i.$$

Por ello, si $\eta^k = 0$ y $\eta^{k'} = 0$, entonces $(\eta' - \eta)^{2\max\{k,k'\}} = 0$. Es decir, $\eta' - \eta$ es nilpotente. Por otro lado, el teorema 3.2.2 afirma que existe una base de vectores propios de δ y δ' para V y, por consiguiente, es una base de vectores propios de $\delta - \delta'$. Ahora, sabemos que $\delta + \eta = f = \delta' + \eta'$, lo que implica que $\delta - \delta' = \eta' - \eta$. Entonces $\delta - \delta'$ es diagonalizable y nilpotente. Claramente, al aplicar un cambio de base, la propiedad de ser nilpotente se preserva. Además, elevar una matriz diagonal a una potencia no es más que elevar cada una de sus entradas diagonales a dicha potencia. De modo que si nos fijamos en la diagonalización de $\delta - \delta'$ y escogemos una potencia que la anule, podemos concluir que las entradas diagonales son cero. Por lo tanto, $0 = \delta - \delta' = \eta' - \eta$, lo que concluye la demostración. □

3.3. Sistemas de ecuaciones diferenciales lineales

La forma canónica de Jordan se puede usar para resolver sistemas de ecuaciones diferenciales lineales utilizando el método que expondremos en esta sección. Para ello, empecemos ocupándonos del caso en el que la matriz asociada al sistema es un bloque de Jordan, de modo que el sistema en forma matricial sería

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \\ \vdots \\ x_{n-1}' \\ x_n' \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}. \tag{3.1}$$

Teorema 3.3.1. *Las soluciones del sistema de ecuaciones diferenciales lineales (3.1) están dadas por*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ t & 1 & 0 & \cdots & 0 & 0 & 0 \\ \frac{t^2}{2!} & t & 1 & \cdots & 0 & 0 & 0 \\ \frac{t^3}{3!} & \frac{t^2}{2!} & t & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-3}}{(n-3)!} & \frac{t^{n-4}}{(n-4)!} & \cdots & t & 1 & 0 \\ \frac{t^{n-1}}{(n-1)!} & \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-3}}{(n-3)!} & \cdots & \frac{t^2}{2!} & t & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ \vdots \\ c_{n-1} \\ c_n \end{pmatrix}. \tag{3.2}$$

Demostración. La prueba es por inducción sobre n. Para $n = 1$, el sistema es simplemente la ecuación $x_1' = \lambda x_1$, cuya conocida solución es $x_1 = e^{\lambda t} c_1$. Supongamos el resultado cierto para n. Si tenemos el sistema

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \\ \vdots \\ x_n' \\ x_{n+1}' \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \\ x_{n+1} \end{pmatrix},$$

entonces los primeros n renglones nos dan un sistema de ecuaciones diferenciales lineales que no depende de x_{n+1} y, consecuentemente, las soluciones de ese sistema son las que se muestran en la ecuación (3.2). Para x_{n+1} se tiene la ecuación

$$x_{n+1}' = x_n + \lambda x_{n+1}.$$

Sustituyendo la solución para x_n, se tiene que

$$x_{n+1}' = e^{\lambda t} \sum_{k=0}^{n-1} c_{n-k} \frac{t^k}{k!} + \lambda x_{n+1},$$

de donde

$$e^{-\lambda t}x'_{n+1} - \lambda e^{-\lambda t}x_{n+1} = \sum_{k=0}^{n-1} c_{n-k}\frac{t^k}{k!}.$$

Integrando de ambos lados tenemos que

$$e^{-\lambda t}x_{n+1} = \int \sum_{k=0}^{n-1} c_{n-k}\frac{t^k}{k!}\,dt = \sum_{k=0}^{n-1} c_{n-k}\int \frac{t^k}{k!}\,dt = \sum_{k=0}^{n-1} c_{n-k}\frac{t^{k+1}}{(k+1)!} + c_{n+1}.$$

Lo que implica que

$$x_{n+1} = e^{\lambda t}\sum_{k=0}^{n} c_{n+1-k}\frac{t^k}{k!}.$$

Con esto se concluye la prueba por inducción. □

Ahora, recordemos que la forma canónica de Jordan se compone de varios bloques de Jordan acomodados diagonalmente. Si la matriz asociada al sistema de ecuaciones diferenciales fuera una forma canónica de Jordan, de la composición de la forma de Jordan podríamos inferir que cada bloque da lugar a un sistema de ecuaciones diferenciales lineales independiente de los demás. Utilizando el teorema 3.3.1 podemos entonces encontrar las soluciones para cada sistema por separado.
Si estamos en el caso más general, en el que la matriz asociada al sistema, A, no está en su forma canónica de Jordan, lo que podemos hacer es lo siguiente.
Encontramos una matriz invertible Q tal que $A = QJQ^{-1}$, donde J es la forma canónica de Jordan de A. Si nuestro sistema en forma matricial es $X' = AX$, entonces, con lo anterior, se vuelve $X' = QJQ^{-1}X$. Lo que implica que

$$Q^{-1}X' = (Q^{-1}X)' = J(Q^{-1}X),$$

la primera igualdad se debe a que Q^{-1} es una matriz constante. Tomando $Z = Q^{-1}X$, el sistema se convierte en

$$Z' = JZ,$$

cuyas soluciones podemos encontrar con el procedimiento descrito anteriormente. Ya una vez conociendo Z, podemos encontrar a X con la igualdad $Z = Q^{-1}X$. Multiplicando por Q de ambos lados, se tiene que

$$X = QZ.$$

A continuación ejemplificaremos el procedimiento anterior.

Ejemplo 3.3.2. *Encontraremos las soluciones del siguiente sistema de ecuaciones diferenciales,*

$$\begin{aligned} x'_1 &= x_1 + x_2 + 3x_3 \\ x'_2 &= 5x_1 + 2x_2 + 6x_3 \\ x'_3 &= -2x_1 - x_2 - 3x_3, \end{aligned}$$

cuya matriz asociada es

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}.$$

Del ejemplo 2.3.7 sabemos que su forma canónica de Jordan está dada por

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

De modo que existe Q, una matriz de 3×3 invertible, tal que $A = QJQ^{-1}$. Siguiendo el procedimiento explicado anteriormente resolvemos el sistema $Z' = JZ$, que, por el teorema 3.3.1, tiene soluciones de la forma

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ \frac{t^2}{2} & t & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Por lo tanto, las soluciones del sistema de ecuaciones diferenciales están dadas por

$$X = QZ.$$

Lo único que nos falta es encontrar la matriz Q de cambio de base. Para ello buscaremos un ciclo de vectores propios generalizados correspondientes al valor propio 0. Recordemos que el polinomio característico de A es $p(x) = -x^3$, consecuentemente, $A^3 = 0$ (Teo. de Cayley-Hamilton). Esto implica que si $y \in \mathbb{R}^3\backslash\{0\}$ y $A^2y \neq 0 \neq Ay$, entonces $\{A^2y, Ay, y\}$ es un ciclo de vectores propios generalizados correspondientes a 0 y, por la proposición 3.1.6, el ciclo es una base. Claramente la representación matricial en esta base (en orden inverso) es la forma canónica de Jordan. Es decir, el cambio de base que necesitamos está dado por la matriz (ver observación 3.1.9)

$$\begin{pmatrix} y & Ay & A^2y \end{pmatrix}.$$

Tomemos $y = e_1 = (1, 0, 0)$. En este caso,

$$\begin{aligned} Ae_1 &= (1, 5, -2) \\ A^2e_1 &= (0, 3, -1). \end{aligned}$$

Es decir, podemos utilizar a e_1 para el ciclo de vectores propios generalizados, de donde, la matriz de cambio de coordenadas está dada por

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & -2 & -1 \end{pmatrix}$$

y, por lo tanto, las soluciones del sistema son

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ \frac{t^2}{2} & t & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Para finalizar este ejemplo, observemos que es posible encontrar ciclos de vectores propios generalizados totalmente distintos (tómese, por ejemplo, $y = e_2 = (0, 1, 0)$) y, por consiguiente, la matriz Q de cambio de base será distinta. Sin embargo, las soluciones serán las mismas, pues el cambio de coordenadas altera también a $X(0)$, las condiciones iniciales, y, con ello, se compensa la diferencia entre las distintas matrices Q.

Bibliografía

[1] M. F. Atiyah and I. G. MacDonald. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969.

[2] T. S. Blyth. *Module Theory, An approach to linear algebra.* University of St Andrews, 2 edition, 1977.

[3] D. S. Dummit and R. M. Foote. *Abstract Algebra.* John Wiley and Sons, Inc., 3 edition, 2004.

[4] S. H. Friedberg, A. J. Insel, and L. E. Spence. *Linear Algebra.* Pearson, 5 edition, 2019.

[5] Steven H. Weintraub. *Jordan Canonical Form: Application to Differential Equations.* Morgan & Claypool, 1 edition, 2008.