Chapter I

Modules and matrices

Apart from the general reference given in the Introduction, for this Chapter we refer in particular to [8] and [20].

Let R be a ring with $1 \neq 0$. We assume most definitions and basic notions concerning left and right modules over R and recall just a few facts.

If M is a left R-module, then for every $m \in M$ the set $\mathrm{Ann}\,(m) := \{r \in R \mid rm = 0_M\}$ is a left ideal of R. Moreover $\mathrm{Ann}\,(M) = \bigcap_{m \in M} \mathrm{Ann}\,(m)$ is an ideal of R. The module M is *torsion free* if $\mathrm{Ann}\,(m) = \{0\}$ for all non-zero $m \in M$.

The regular module ${}_RR$ is the additive group $(R, +)$ considered as a left R-module with respect to the ring product. The submodules of ${}_RR$ are precisely the left ideals of R.

A finitely generated R-module is *free* if it is isomorphic to the direct sum of n copies of ${}_RR$, for some natural number n. Namely if it is isomorphic to the module

$$(0.1) \qquad ({}_RR)^n := \underbrace{{}_RR \oplus \cdots \oplus_R R}_{n \text{ times}}$$

in which the operations are performed component-wise. If R is commutative, then $({}_RR)^n \cong ({}_RR)^m$ only if $n = m$. So, in the commutative case, the invariant n is called the *rank* of $({}_RR)^n$. Note that $({}_RR)^n$ is torsion free if and only if R has no zero-divisors.

The aim of this Chapter is to determine the structure of finitely generated modules over a principal ideal domain (which are a generalization of finite dimensional vector spaces) and to describe some applications. But we start with an important result, valid for modules over any ring.

1 The Theorem of Krull-Schmidt

(1.1) Definition *An R-module M is said to be* indecomposable *if it cannot be written as the direct sum of two proper submodules.*

For example the regular module ${}_{\mathbb{Z}}\mathbb{Z}$ is indecomposable since any two proper ideals $n\mathbb{Z}$ and $m\mathbb{Z}$ intersect non-trivially. E.g. $0 \neq nm \in n\mathbb{Z} \cap m\mathbb{Z}$.

(1.2) Definition *Let M be an R-module.*

(1) *M is* noetherian *if, for every ascending chain of submodules*

$$M_1 < M_2 < M_3 < \dots$$

there exists $n \in \mathbb{N}$ such that $M_n = M_{n+r}$ for all $r \geq 0$;

(2) *M is* artinian *if, for every descending chain of submodules*

$$M_1 > M_2 > M_3 < \dots$$

there exists $n \in \mathbb{N}$ such that $M_n = M_{n+r}$for all $r \geq 0$.

(1.3) Lemma *An R-module M is noetherian if and only if every submodule of M is finitely generated.*

(1.4) Examples

- *every finite dimensional vector space is artinian and noetherian;*
- *the regular $\mathbb{Z}$-modulo ${}_{\mathbb{Z}}\mathbb{Z}$ is noetherian, but it is not artinian;*
- *for every field $\mathbb{F}$, the polynomial ring $\mathbb{F}[x_1, \dots, x_n]$ is noetherian.*

(1.5) Theorem *(Krull-Schmidt) Let M be an artinian and noetherian R-module. Given two decompositions*

$$M = M_1 \oplus M_2 \oplus M_n = N_1 \oplus N_2 \oplus N_m$$

suppose that the M_i-s and the N_j-s are indecomposable submodules. Then $m = n$ and there exists a permutation of the N_i-s such that M_i is isomorphic to N_i for all $i \leq n$.

2 Finitely generated modules over a PID

We indicate by D a *principal ideal domain* (PID), namely a commutative ring, without zero-divisors, in which every ideal is of the form $Dd = \langle d \rangle$, for some $d \in D$.
Every euclidean domain is a PID. In particular we have the following

(2.1) Examples *of PID-s:*

- *the ring* $\mathbb{Z}$ *of integers;*
- *every field* $\mathbb{F}$*;*
- *the polynomial ring* $\mathbb{F}[x]$ *over a field.*

Let A be an $m \times n$ matrix with entries in D. Then there exist $P \in \mathrm{GL}_m(D)$ and $Q \in \mathrm{GL}_n(D)$ such that PAQ is a pseudodiagonal matrix in which the entry in position (i,i) divides the entry in position $(i+1, i+1)$ for all i-s. The matrix PAQ is called a normal form of A. A consequence of this fact is the following:

(2.2) Theorem *Let* V *be a free* D*-module of rank* n *and* W *be a submodule.*

(1) W *is free of rank* $t \leq n$*;*

(2) *there exist a basis* $\mathcal{B} = \{v_1, \cdots, v_n\}$ *of* V *and a sequence* $d_1, \cdots, d_t$ *of elements of* D *with the following properties:*

   i) d_i *divides* d_{i+1} *for* $1 \leq i \leq t-1$*,*

   ii) $\mathcal{C} = \{d_1 v_1, \cdots, d_t v_t\}$ *is a basis of* W*.*

We may now state the structure theorem of a finitely generated D-module M. To this purpose let us denote by $\mathrm{d}(M)$ the minimal number of generators of M as a D-module.

(2.3) Theorem *Let* M *be a finitely generated* D*-module, with* $\mathrm{d}(M) = n$*.*
There exists a descending sequence of ideals:

$$Dd_1 \geq \cdots \geq Dd_n \quad \text{(invariant factors of } M\text{)} \tag{2.4}$$

with $Dd_1 \neq D$*, such that:*

$$M \simeq \frac{D}{Dd_1} \oplus \cdots \oplus \frac{D}{Dd_n} \quad \text{(normal form of } M\text{)}. \tag{2.5}$$

Let $t \geq 0$ be such that $d_t \neq 0_D$ and $d_{t+1} = 0_D$. Then, setting:

$$T := \{0_M\} \quad \text{if } t = 0, \qquad T := \frac{D}{Dd_1} \oplus \cdots \oplus \frac{D}{Dd_t} \quad \text{if } t > 0, \tag{2.6}$$

we have that $\mathrm{Ann}\,(T) = Dd_t$ *and T is isomorphic to the torsion submodule of M.*

M is torsion free if and only if $t = n$, $M = T$. Indeed, by this Theorem:

$$M \simeq T \oplus D^{n-t}$$

where D^{n-t} is free, of rank $n - t$.

Proof (sketch) Let $m_1, \ldots, m_n$ be a set of generators of M as a D-module. Consider the epimorphism $\psi : D^n \to M$ such that

$$\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i m_i.$$

By Theorem 2.2, there exist a basis $\{v_1, \cdots, v_n\}$ of D^n and a sequence $d_1, \cdots, d_t$ of elements of D with the property that d_i divides d_{i+1} for $1 \leq i \leq t-1$, such that $\{d_1 v_1, \cdots, d_t v_t\}$ is a basis of Ker ψ. It follows $\frac{D^n}{\mathrm{Ker}\psi} \cong M$, whence:

$$\begin{aligned} \tfrac{Dv_1 \oplus \cdots \oplus Dv_t}{Dd_1 v_1 \oplus \cdots \oplus Dd_t v_t} \oplus \tfrac{Dv_{t+1} \oplus \cdots \oplus Dv_n}{\{0\} \oplus \cdots \oplus \{0\}} &\cong M \\ \tfrac{D}{Dd_1} \oplus \cdots \oplus \tfrac{D}{Dd_t} \oplus D \oplus \cdots \oplus D &\cong M. \end{aligned}$$

■

(2.7) Corollary *Let V be a vector space over $\mathbb{F}$, with $\mathrm{d}(V) = n$. Then $V \simeq \mathbb{F}^n$.*

(2.8) Corollary *Let M be a f.g. abelian group, with $\mathrm{d}(M) = n$. Then either:*

(1) *$M \simeq \mathbb{Z}^n$, or*

(2) *$M \simeq \mathbb{Z}_{d_1} \oplus \cdots \mathbb{Z}_{d_t} \oplus \mathbb{Z}^{n-t}$, $t \leq n$,*

where $d_1, \cdots, d_t$ is a sequence of integers ≥ 2, each of which divides the next one.

It can be shown that the normal form (2.5) of a f.g. D-module M is unique. Thus:

(2.9) Theorem *Two finitely generated D-modules are isomorphic if and only if they have the same normal form* (2.5) *or, equivalently, the same invariant factors* (2.4).

In the notation of Theorem 2.3, certain authors prefer to call invariant factors the elements $d_1, \ldots, d_n$ instead of the ideals generated by them. In this case the invariant factors are determined up to unitary factors.

(2.10) Example *Every abelian group of order p^3, with p prime, is isomorphic to one and only one of the following:*

- $\mathbb{Z}_{p^3}$, $t = 1$, $d_1 = p^3$;
- $\mathbb{Z}_p \oplus \mathbb{Z}_{p^2}$, $t = 2$, $d_1 = p$, $d_2 = p^2$;
- $\mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \mathbb{Z}_p$, $t = 3$, $d_1 = d_2 = d_3 = p$.

(2.11) Example *Every abelian group of order 20 is isomorphic to one and only one of the following:*

- $\mathbb{Z}_{20}$, $t = 1$, $d_1 = 20$;
- $\mathbb{Z}_2 \oplus \mathbb{Z}_{10}$, $t = 2$, $d_1 = 2$, $d_2 = 10$.

3 The primary decomposition

We recall that D is a PID. For any $a, b \in D$ we have $Da + Db = Dd$, whence $d =$ G.C.D.(a, b). It follows easily that D is a unique factorization domain.
The results of this Section are based on the previous facts and the well known Chinese remainder Theorem, namely:

(3.1) Theorem *Let $a, b \in D$ such that* M.C.D.$(a, b) = 1$. *For all $b_1, b_2 \in D$, there exists $c \in D$ such that*

$$\begin{cases} c \equiv b_1 \pmod{a} \\ c \equiv b_2 \pmod{b}. \end{cases} \tag{3.2}$$

Proof There exist $y, z \in D$ such that $ay + bz = 1$. Multiplying by b_1 and b_2:

$$\begin{matrix} ayb_1 + bzb_1 = b_1 \\ ayb_2 + bzb_2 = b_2 \end{matrix}.$$

It follows

$$\begin{matrix} bzb_1 \equiv b_1 \pmod{a} \\ ayb_2 \equiv b_2 \pmod{b} \end{matrix}.$$

We conclude that $c = bzb_1 + ayb_2$ satisfies (3.2). ■

(3.3) Theorem *Let $d = p_1^{m_1} \dots p_k^{m_k}$, where each p_i is an irreducible element of D and $p_i \neq p_j$ for $1 \leq i \neq j \leq k$. Then:*

$$\frac{D}{Dd} \simeq \frac{D}{Dp_1^{m_1}} \oplus \cdots \oplus \frac{D}{Dp_k^{m_k}} \quad \text{(primary decomposition)}. \tag{3.4}$$

$Dp_1^{m_1}, \cdots, Dp_k^{m_k}$ (or simply $p_1^{m_1}, \cdots, p_k^{m_k}$) are the elementary divisors *of $\frac{D}{Dd}$.*

Proof Setting $a = p_1^{m_1}$, $b = p_2^{m_2} \dots p_k^{m_k}$, we have $d = ab$ with G.C.D.$(a, b) = 1$. The map

$$f : D \to \frac{D}{Da} \oplus \frac{D}{Db} \quad \text{such that} \quad x \mapsto \begin{pmatrix} Da + x \\ Db + x \end{pmatrix}$$

is a D-homomorphism. Moreover it is surjective by theorem 3.1. Finally Ker $f = Da \cap Db = Dd$. We conclude that

$$\frac{D}{Dd} \simeq \frac{D}{Da} \oplus \frac{D}{Db} = \frac{D}{Dp_1^{m_1}} \oplus \frac{D}{D\left(p_2^{m_2} \dots p_k^{m_k}\right)}$$

and our claim follows by induction on k. ■

(3.5) Examples

- $\mathbb{Z}_6 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3$, *elementary divisors* $2, 3$;
- $\mathbb{Z}_6 \oplus \mathbb{Z}_6 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3$, *elementary divisors* $2, 2, 3, 3$;
- $\mathbb{Z}_{40} \cong \mathbb{Z}_8 \oplus \mathbb{Z}_5$, *elementary divisors* $8, 5$;
- $\frac{\mathbb{C}[x]}{\langle x^3-1\rangle} \cong \frac{\mathbb{C}[x]}{\langle x-1\rangle} \oplus \frac{\mathbb{C}[x]}{\langle x-\omega\rangle} \oplus \frac{\mathbb{C}[x]}{\langle x-\overline{\omega}\rangle}$, *el. div.* $x-1,\ x-\omega,\ x-\overline{\omega}$ *where* $\omega = e^{\frac{i2\pi}{3}}$.

4 Modules over $\mathbb{F}[x]$ defined by matrices

Let $\mathbb{F}$ be a field. We recall that two matrices $A, B \in \mathrm{Mat}_n(\mathbb{F})$ are *conjugate* if there exist $P \in \mathrm{GL}_n(\mathbb{F})$ such that $P^{-1}AP = B$. The conjugacy among matrices is an equivalence relation in $\mathrm{Mat}_n(\mathbb{F})$, whose classes are called *conjugacy classes*. Our goal here is to find representatives for these classes.

The additive group $(\mathbb{F}^n, +)$ of column vectors is a left module over the ring $\mathrm{Mat}_n(\mathbb{F})$, with respect to the usual product of matrices. For a fixed matrix $A \in \mathrm{Mat}_n(\mathbb{F})$, the map: $\varphi_A : \mathbb{F}[x] \to \mathrm{Mat}_n(\mathbb{F})$ such that

$$f(x) \mapsto f(A)$$

is a ring homomorphism. It follows that $\mathbb{F}^n$ is an $\mathbb{F}[x]$-module with respect to the product:

$$f(x)\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} := f(A)\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}. \tag{4.1}$$

The $\mathbb{F}[x]$-module defined by (4.1) will be denoted by ${}_A\mathbb{F}^n$. Identifying $\mathbb{F}$ with the subring $\mathbb{F}x^0$ of $\mathbb{F}[x]$, the module ${}_A\mathbb{F}^n$ is a vector space over $\mathbb{F}$ in the usual way. Indeed, for all $\alpha \in \mathbb{F}$ and all $v \in {}_A\mathbb{F}^n$, we have: $(\alpha x^0)v = (\alpha A^0)v = \alpha v$.
Clearly, if V is any $\mathbb{F}[x]$-module, the map $\mu_x : V \to V$ such that

$$v \mapsto xv, \quad \forall\, v \in V \tag{4.2}$$

is an $\mathbb{F}[x]$-homomorphism. In particular μ_x is $\mathbb{F}$-linear.

(4.3) Theorem *Let V be an $\mathbb{F}[x]$-module, $dim_{\mathbb{F}}(V) = n$, and let $A, B \in \mathrm{Mat}_n(\mathbb{F})$.*

(1) *$V \simeq {}_A\mathbb{F}^n$ if and only if μ_x has matrix A with respect to a basis $\mathcal{B}$ of V;*

(2) *${}_A\mathbb{F}^n \simeq {}_B\mathbb{F}^n$ if and only if B is conjugate to A.*

Proof
(1) Suppose that μ_x has matrix A with respect to a basis $\mathcal{B}$ and call η the map which assigns to each $v \in V$ its coordinate vector $v_{\mathcal{B}}$ with respect to $\mathcal{B}$. We have:

$$Av_{\mathcal{B}} = (\mu_x(v))_{\mathcal{B}} = (xv)_{\mathcal{B}}, \quad \forall\, v \in V.$$

Clearly $\eta : V \to {}_A\mathbb{F}^n$ is an isomorphism of $\mathbb{F}$-modules. Moreover:

$$\eta(xv) = (xv)_{\mathcal{B}} = Av_{\mathcal{B}} = x\, v_{\mathcal{B}} = x\, \eta(v).$$

It follows easily that η is an isomorphism of $\mathbb{F}[x]$-modules. Thus $V \simeq {}_A\mathbb{F}^n$.
Vice versa, suppose that there exists an $\mathbb{F}[x]$-isomorphism $\gamma : V \to {}_A\mathbb{F}^n$. Set $\mathcal{B} = \{\gamma^{-1}(e_1),\ \dots,\ \gamma^{-1}(e_n)\}$, where $\{e_1, \dots, e_n\}$ is the canonical basis of $\mathbb{F}^n$. Then

$$\gamma(v) = \gamma\left(\sum_{i=1}^{n} k_i\, \gamma^{-1}(e_i)\right) = \sum_{i=1}^{n} k_i e_i = v_{\mathcal{B}}, \quad \forall\, v \in V.$$

Now $\gamma(xv) = x\gamma(v)$ gives $(\mu_x(v))_{\mathcal{B}} = Av_{\mathcal{B}}$. So μ_x has matrix A with respect to $\mathcal{B}$.

(2) Take $V = {}_A\mathbb{F}^n$, the $\mathbb{F}[x]$-module for which $\mu_x = \mu_A$. By the previous point ${}_A\mathbb{F}^n \simeq {}_B\mathbb{F}^n$ if and only if the linear map μ_A, induced by A with respect to the canonical basis, has matrix B with respect to an appropriate basis $\mathcal{B}$ of V. By elementary linear algebra this happens if and only if B is conjugate to A. ■

5 The rational canonical form of matrices

(5.1) Theorem *Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The $\mathbb{F}[x]$-module ${}_A\mathbb{F}^n$ defined in* (4.1) *is finitely generated and torsion free.*

Proof $\mathbb{F}^n$ is finitely generated as a $\mathbb{F}$-module. Hence, a fortiori, as a $\mathbb{F}[x]$-module. In order to show that it is torsion free we must show that, for all $v \in \mathbb{F}^n$, there exists a non-zero polynomial $f(x) \in \mathbb{F}[x]$ such that $f(x)v = f(A)v = 0_{\mathbb{F}^n}$. This is clear if $A^i v = A^j v$ for some non-negative $i \neq j$. Because, in this case, we may take $f(x) = x^i - x^j$. Otherwise the subset $\{v,\ Av,\ \cdots,\ A^n v\}$ of $\mathbb{F}^n$ has cardinality $n+1$. It follows that there exist $k_0, \cdots, k_n$ in $\mathbb{F}$, not all zero, such that $k_0 v + k_1 Av + \cdots k_n A^n v = 0_{\mathbb{F}^n}$. So we may take $f(x) = k_0 + k_1 x + \cdots + k_n x^n$. ■

By Theorem 2.3 there exists a chain of ideals $\langle d_1(x)\rangle \geq \cdots \geq \langle d_t(x)\rangle \neq \{0\}$ such that

$$ {}_A\mathbb{F}^n \quad \simeq \quad \frac{\mathbb{F}[x]}{\langle d_1(x)\rangle} \oplus \ \cdots \ \oplus \frac{\mathbb{F}[x]}{\langle d_t(x)\rangle}. \tag{5.2}$$

Clearly $\langle d_t(x)\rangle = \mathrm{Ann}({}_A\mathbb{F}^n) = \mathrm{Ker}\ \varphi_A$. Moreover each $d_i(x)$ can be taken monic.

(5.3) Definition

(1) $d_1(x), \cdots,\ d_t(x)$ *are called the* similarity invariants *of A;*

(2) $d_t(x)$ *is called the* minimal polynomial *of A.*

(5.4) Definition *For a given monic polynomial of degree s*

$$d(x) = k_0 + k_1 x + k_2 x^2 \cdots + k_{s-1} x^{s-1} + x^s \in \mathbb{F}[x]$$

its companion matrix $C_{d(x)}$ *is defined as the matrix of* $\mathrm{Mat}_s(\mathbb{F})$ *whose columns are respectively $e_2, \ldots, e_s, [-k_0, \ldots, -k_{s-1}]^T$, namely the matrix:*

$$C_{d(x)} \ := \ \begin{pmatrix} 0 & 0 & \cdots & -k_0 \\ 1 & 0 & \cdots & -k_1 \\ 0 & 1 & \cdots & -k_2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & -k_{s-1} \end{pmatrix}. \tag{5.5}$$

(5.6) Lemma *The companion matrix $C_{d(x)}$ has $d(x)$ as characteristic polynomial and as minimal polynomial.*

The first claim can be shown by induction on s, the second noting that

$$C_{d(x)}e_i = e_{i+1}, \quad i \leq s-1.$$

(5.7) Theorem *Consider the $\mathbb{F}[x]$-module $V = \frac{\mathbb{F}[x]}{\langle d(x)\rangle}$ and the map $\mu_x : V \to V$.*

(1) $\mathcal{B} := \left\{\langle d(x)\rangle + x^0,\ \langle d(x)\rangle + x,\ \cdots,\ \langle d(x)\rangle + x^{s-1}\right\}$ *is a basis of V over $\mathbb{F}$;*

(2) μ_x *has matrix $C_{d(x)}$ with respect to $\mathcal{B}$.*

Proof Routine calculation, noting that $\mu_x\left(\langle d(x)\rangle + f(x)\right) = \langle d(x)\rangle + xf(x)$. ■

We may now consider the general case. Let

$$V = \frac{\mathbb{F}[x]}{\langle d_1(x)\rangle} \oplus \cdots \oplus \frac{\mathbb{F}[x]}{\langle d_t(x)\rangle} = V_1 \oplus \cdots \oplus V_t$$

where each $d_i(x)$ is a monic, non-constant polynomial, and

$$d_i(x) \text{ divides } d_{i+1}(x),\ 1 \leq i \leq t-1. \tag{5.8}$$

With respect to the basis $\mathcal{B}_1 \times \{0_{V_2\oplus\cdots\oplus V_t}\}\ \dot\cup \ldots\ \dot\cup\ \mathcal{B}_t \times \left\{0_{V_1\oplus\cdots\oplus V_{t-1}}\right\}$, where each $\mathcal{B}_i$ is the basis of $\frac{\mathbb{F}[x]}{\langle d_i(x)\rangle}$ defined in Theorem 5.7, the map μ_x has matrix:

$$C = \begin{pmatrix} C_{d_1(x)} & & \\ & \cdots & \\ & & C_{d_t(x)} \end{pmatrix}. \tag{5.9}$$

(5.10) Definition *Every matrix C as in* (5.9)*, with $d_1(x),\ \ldots,\ d_t(x)$ satisfying* (5.8)*, is called a* rational canonical form.

(5.11) Lemma *The rational canonical form C in* (5.9) *has characteristic polynomial $\prod_1^t d_i(x)$ and minimal polynomial $d_t(x)$.*

From the above results we may conclude the following

(5.12) Theorem *For any field $\mathbb{F}$, every matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ is conjugate to a unique rational canonical form.*

Clearly conjugate matrices have the same characteristic polynomial and the same minimal polynomial. So Lemma 5.11 has the following:

(5.13) Corollary (Theorem of Hamilton-Cayley). *Let $f(x)$ be the characteristic polynomial of a matrix A. Then $f(A) = 0$.*

(5.14) Example *The rational canonical forms in* $\mathrm{Mat}_2(\mathbb{F})$ *are of the following types:*
a) $t = 2$, $d_1(x) = d_2(x) = x - k$,

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

b) $t = 1$, $d_1(x) = x^2 + k_1x + k_0$,

$$\begin{pmatrix} 0 & -k_0 \\ 1 & -k_1 \end{pmatrix}.$$

(5.15) Example *The rational canonical forms in* $\mathrm{Mat}_3(\mathbb{F})$ *are of the following types:*
a) $t = 3$, $d_1(x) = d_2(x) = d_3(x) = x - k$,

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

b) $t = 2$, $d_1(x) = x - k$, $d_2(x) = (x - h)(x - k)$,

$$\begin{pmatrix} k & 0 & 0 \\ 0 & 0 & -kh \\ 0 & 1 & k + h \end{pmatrix},$$

c) $t = 1$, $d_1(x) = x^3 + k_2x^2 + k_1x + k_0$,

$$\begin{pmatrix} 0 & 0 & -k_0 \\ 1 & 0 & -k_1 \\ 0 & 1 & -k_2 \end{pmatrix}.$$

6 Jordan canonical forms

The rational canonical forms of matrices have the advantage of parametrizing the conjugacy classes of $\mathrm{Mat}_n(\mathbb{F})$ for any field $\mathbb{F}$. The disadvantage is that they say very little about eigenvalues and eigenspaces. For this reason, over an algebraically closed field, the Jordan canonical forms are more used and better known. They can be deduced from the primary decomposition of the $\mathbb{F}[x]$-modules associated to the rational canonical forms.

(6.1) Definition *For every* $\lambda \in \mathbb{F}$ *and every integer* $s \geq 0$ *we define inductively the* Jordan block $J(s, \lambda)$ *setting:*

$$J(0,\lambda) := \emptyset, \qquad J(1,\lambda) := (\lambda), \qquad J(s,\lambda) := \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 1 & & & \\ 0 & & & \\ \cdots & & J(s-1,\lambda) & \\ 0 & & & \end{pmatrix}, \ s > 1.$$

So, for example:

$$J(2,\lambda) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}, \quad J(3,\lambda) = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}, \quad J(4,\lambda) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix}.$$

(6.2) Lemma *$J(s,\lambda)$ has λ as unique eigenvalue and corresponding eigenspace of dimension 1 in $\mathbb{F}^s$.*

Proof $J(s,\lambda)$ has characteristic polynomial $(x-\lambda)^s$, hence λ as unique eigenvalue. Elementary calculation shows that $\langle e_s \rangle$ is the corresponding eigenspace. ■

(6.3) Lemma *Let us consider the $\mathbb{F}[x]$-module*

$$V := \frac{\mathbb{F}[x]}{\langle (x-\lambda)^s \rangle}.$$

The Jordan block $J(s,\lambda)$ is the matrix of $\mu_x : V \to V$ with respect to the basis:

$$\mathcal{B}' := \left\{ I + (x-\lambda)^0, \quad I + (x-\lambda)^1, \quad \cdots, \quad I + (x-\lambda)^{s-1} \right\}.$$

In particular $J(s,\lambda)$ is conjugate to the companion matrix $C_{(x-\lambda)^s}$.

Proof For all $i \geq 0$ the following identity holds:

$$x(x-\lambda)^i = \lambda(x-\lambda)^i - \lambda(x-\lambda)^i + x(x-\lambda)^i = \lambda(x-\lambda)^i + (x-\lambda)^{i+1}.$$

It follows that, for $i \leq s-2$:

$\mu_x \left(I + (x-\lambda)^i \right) = I + x(x-\lambda)^i = \lambda \left(I + (x-\lambda)^i \right) \; + \; I + (x-\lambda)^{i+1}$,

$\mu_x \left(I + (x-\lambda)^{s-1} \right) = I + x(x-\lambda)^{s-1} = I + \lambda(x-\lambda)^{s-1} = \lambda \left(I + (x-\lambda)^{s-1} \right)$.

The last claim follows from Theorem 4.3. ■

(6.4) Corollary

(1) *Let* $d(x) = (x-\lambda_1)^{s_1} \dots (x-\lambda_m)^{s_m}$ *where* $\lambda_i \neq \lambda_j$ *for* $i \neq j$.

The companion matrix $C_{d(x)}$ *is conjugate to the matrix:*

$$J_{d(x)} := \begin{pmatrix} J(s_1,\lambda_1) & & \\ & \cdots & \\ & & J(s_m,\lambda_m) \end{pmatrix}. \tag{6.5}$$

(2) *Every rational canonical form* $C = \begin{pmatrix} C_{d_1(x)} & & \\ & \cdots & \\ & & C_{d_t(x)} \end{pmatrix}$ *is conjugate to*

$$J = \begin{pmatrix} J_{d_1(x)} & & \\ & \cdots & \\ & & J_{d_t(x)} \end{pmatrix} \quad \text{(Jordan form of } C\text{)}.$$

(6.6) Definition *In the above notation let* $\lambda_1, \dots, \lambda_m$ *be the distinct roots of* $d_t(x)$. *Set:*

$$d_i(x) = (x-\lambda_1)^{s_{i_1}} \ \dots \ (x-\lambda_m)^{s_{i_m}}, \quad 1 \leq i \leq t.$$

The factors of positive degree among

$$(x-\lambda_1)^{s_{1_1}}, \ \cdots, \ (x-\lambda_m)^{s_{1_m}}, \ \cdots, \ (x-\lambda_1)^{s_{t_1}}, \ \cdots, \ (x-\lambda_m)^{s_{t_m}}$$

(counted with their multiplicities) are called the elementary divisors *of* J.

(6.7) Example *If* $d_1(x) = (x-4), \quad d_2(x) = (x-3)(x-4)^2, \quad d_3(x) = (x-3)(x-4)^3$, *then the elementary divisors are:* $(x-4),\ (x-3),\ (x-4)^2,\ (x-3),\ (x-4)^3$.

So we have proved the following:

(6.8) Theorem *Let* $\mathbb{F}$ *be an algebraically closed field. Two matrices* A, B *in* $\mathrm{Mat}_n(\mathbb{F})$ *are conjugate if and only if they have the same Jordan form (up to a permutation of the blocks) or, equivalently, the same elementary divisors (counted with their multiplicities).*

We conclude this Section stating a useful result, not difficult to prove.

(6.9) Theorem *Let* $\mathbb{F}$ *be algebraically closed and let* $A \in \mathrm{Mat}_n(\mathbb{F})$. *The following conditions are equivalent:*

(1) A *is diagonalizable;*

(2) *the minimal polynomial of* A *has no multiple roots;*

(3) *every Jordan form of* A *is diagonal;*

(4) $\mathbb{F}^n$ *has a basis of eigenvectors of* A.

7 Exercises

(7.1) Exercise *Let $f : S \to R$ be a ring homomorphism. Show that every R-module M becomes an S-module by setting $sm := f(s)m, \quad \forall\ s \in S,\ m \in M$.*

(7.2) Exercise *Let p be a prime. Determine, up to isomorphisms, the abelian groups of order p^4.*

(7.3) Exercise *Determine, up to isomorphisms, the abelian groups of order 24 and order 100.*

(7.4) Exercise *Show that an euclidean domain is a principal ideal domain.*

(7.5) Exercise *Determine the primary decomposition and the normal form of the abelian group M having elementary divisors $2, 2, 4, 5, 5, 3, 9$. What is $Ann(M)$? What is the minimal number $d(M)$ of generators?*

(7.6) Exercise *Let D be a principal ideal domain and let d_1, d_2 be non-zero elements in D. Show that $Dd_1 = Dd_2$ if and only if $d_2 = \lambda d_1$ with λ invertible in D.*

(7.7) Exercise *Let M_1 and M_2 be R modules and $N_1 \leq M_1$, $N_2 \leq M_2$ be submodules. Show that:*

$$\frac{M_1 \oplus M_2}{N_1 \oplus N_2} \ \cong\ \frac{M_1}{N_1} \oplus \frac{M_2}{N_2}.$$

(7.8) Exercise *Suppose that R is a commutative ring. Let M be an R-module, m an element of M such that* Ann $(m) = \{0_R\}$. *Show that, for every ideal J of R:*

- *$Jm := \{jm \mid j \in J\}$ is a submodule of M;*
- *$\frac{Rm}{Jm} \cong \frac{R}{J}$ as R-modules.*

(7.9) Exercise *Calculate eigenvalues, eigenspaces, Jordan form and rational canonical form of each of the following matrices:*

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ 2 & 1 & 1 & -1 \end{pmatrix}$$