Shandong University | Advanced Modern Algebra

Lecture 7: Modules over Principal Ideal Domains

Apr. 7, 2023

Lecturer: Bin Guan

This lecture refers to Chapter 12 in [1]. All the equation numbers without reference labels are from this book.

1 Preliminaries

1.1 Noetherian rings and Noetherian modules

Recall that we have the following inclusions among classes of commutative rings with identity:

$$\{\text{Fields}\} \subsetneqq \{\text{Euclidean Domains}\} \subsetneqq \{\text{P.I.D.s}\} \subsetneqq \{\text{U.F.D.s}\} \subsetneqq \{\text{Integral Domains}\}$$

with all containments being proper; a polynomial ring $F[x]$ in a variable x over a field F is a Euclidean Domain, and the polynomial ring $F[x_1, \ldots, x_n]$ is a U.F.D.(Unique Factorization Domain). However the latter ring is not a P.I.D.(Principal Ideal Domain) unless $n = 1$.

Actually, ideals in such polynomial rings, although not necessarily principal, are always finitely generated. General rings with this property are given a special name:

Theorem ([1] §12.1 Theorem 1)**.** *Let R be a ring and let M be a left R-module. Then TFAE:*

*(1) M satisfies the **ascending chain condition** on submodules (or **A.C.C.** on submodules), i.e., whenever*

$$M_1 \subseteq M_2 \subseteq M_3 \subseteq \cdots$$

is an increasing chain of submodules of M, then there is a positive integer m such that for all $k \geq m$, $M_k = M_m$ (so the chain becomes stationary at stage m:

$$M_1 \subseteq \cdots \subseteq M_{m-1} \subseteq M_m = M_{m+1} = M_{m+2} = \cdots).$$

(2) Every nonempty set of submodules of M contains a maximal element under inclusion.

(3) Every submodule of M is finitely generated.

The left R-module M is said to be a ***Noetherian*** *R-module if it satisfies any of the above equivalent conditions. The ring R is said to be* ***Noetherian*** *if it is Noetherian as a left module over itself, i.e., if there are no infinite increasing chains of left ideals in R.*

One can formulate analogous notions of A.C.C. on right and on two-sided ideals in a (possibly noncommutative) ring R. For noncommutative rings these properties need not be related.

Example. *Any P.I.D. R is a Noetherian ring due to condition (3) in the theorem with $M = R$. Then every nonempty set of ideals of R has a maximal element, and R satisfies the A.C.C. on two-sides ideals, which is equivalent to the descending chain condition (D.C.C.) on elements in this case.*

Example. *Even if M itself is a finitely generated R-module, submodules of M need not be finitely generated, so the condition that M be a Noetherian R-module is in general stronger than the condition that M be a finitely generated R-module.*

Take M to be the cyclic R-module R itself where R is the polynomial ring in infinitely many variables $x_1, x_2, \ldots$ with coefficients in some field F. The submodule (i.e. 2-sided ideal) generated by $\{x_1, x_2, \ldots\}$ cannot be generated by any finite set (note that one must show that no finite subset of this ideal will generate it).

Proof of Theorem 1. [(1) $\Rightarrow$ (2)] Assume M is Noetherian and let Σ be any nonempty collection of submodules of M. Choose any $M_1 \in \Sigma$. If M_1 is a maximal element of Σ then (2) holds, so assume M_1 is not maximal. Then there is some $M_2 \in \Sigma$ such that $M_1 \subsetneqq M_2$. If M_2 is maximal in Σ, (2) holds, so we may assume there is an $M_3 \in \Sigma$ properly containing M_2. Proceeding in this way one sees that if (2) fails we can produce an infinite strictly increasing chain of elements of Σ, contrary to (1).

[(2) $\Rightarrow$ (3)] Assume (2) holds and let N be any submodule of M. Let Σ be the collection of all finitely generated submodules of N. Since $0 \in \Sigma$, this collection is nonempty. By (2) Σ contains a maximal element N'. If $N' \subsetneqq N$, let $x \in N - N'$. Since $N' \in \Sigma$, the submodule N' is finitely generated by assumption, hence also the submodule generated by N' and x is finitely generated. This contradicts the maximality of N', so $N = N'$ is finitely generated.

[(3) $\Rightarrow$ (1)] Assume (3) holds and let $M_1 \subseteq M_2 \subseteq M_3 \subseteq \cdots$ be a chain of submodules of M. Let $N = \bigcup_{i=1}^{\infty} M_i$ and note that N is a submodule. By (3) N is finitely generated by, say, $a_1, a_2, \ldots, a_n$. Since $a_i \in N$ for all i, each a_i lies in one of the submodules in the chain, say M_{j_i}. Let $m = \max\{j_1, j_2, \ldots, j_n\}$. Then $a_i \in M_m$ for all i so the module they generate is contained in M_m, i.e., $N \subseteq M_m$. This implies $M_m = N = M_k$ for all $k \geq m$, which proves (1). $\square$

A polynomial ring in n variables can be considered as a polynomial ring in one variable with coefficients in a polynomial ring in $n-1$ variables. By following this inductive approach we can deduce that $F[x_1, x_2, \ldots, x_n]$ is Noetherian from the following more general result.

Theorem ([1] §9.6 Theorem 21, Hilbert's Basis Theorem)**.** *If R is a Noetherian ring then so is the polynomial ring $R[x]$.*

Sketch of proof. Let I be an ideal in $R[x]$ and let L be the set of all leading coefficients of the elements in I. One can verify that

$$L \text{ is an ideal of } R.$$

Since R is assumed Noetherian, the ideal L in R is finitely generated, say by $a_1, a_2, \ldots, a_n \in R$. For each $i = 1, \ldots, n$ let f_i be an element of I whose leading coefficient is a_i.

Let $N = \max\{\deg f_i\}_{i=1}^n$. For each $d = 0, 1, \ldots, N-1$, let L_d be the set of all leading coefficients of polynomials in I of degree d together with 0. One can verify that

$$\text{each } L_d \text{ is also an ideal of } R,$$

again finitely generated (since R is Noetherian). Let $b_{d,1}, b_{d,2}, \ldots, b_{d,n_d} \in R$ be a set of generators for each nonzero L_d, and let $f_{d,i}$ be a polynomial in I of degree d with leading coefficient $b_{d,i}$.

The last step is to show that, the polynomials $f_1, \ldots, f_n$ together with all the polynomials $f_{d,i}$ are a set of generators for I, i.e., that

$$I = (\{f_1, \ldots, f_n\} \cup \{f_{d,i} \mid 0 \leq d < N, 1 \leq i \leq n_d\}).$$

It follows that I is finitely generated, and since I was arbitrary, this completes the proof that $R[x]$ is Noetherian. □

Since a field is clearly Noetherian, Hilbert's Basis Theorem and induction immediately give:

Theorem ([1] §9.6 Corollary 22)**.** *The polynomial ring $F[x_1, x_2, \ldots, x_n]$ with coefficients from a field F is Noetherian, i.e., very ideal in this ring is finitely generated.*

If I is an ideal in $F[x_1, x_2, \ldots, x_n]$ generated by a (possibly infinite) set S of polynomials, the above corollary shows that I is finitely generated, and in fact I is generated by a finite number of the polynomials from the set S (cf. [1] §9.6 Exercise 1).

1.2 Linear dependency

Proposition ([1] §12.1 Proposition 3)**.** *Let R be an integral domain and let M be a free R-module of rank $n < \infty$. Then any $n+1$ elements of M are R-**linearly dependent**, i.e., for any $y_1, y_2, \ldots, y_{n+1} \in M$ there are elements $r_1, r_2, \ldots, r_{n+1} \in R$, not all zero, such that*

$$r_1 y_1 + r_2 y_2 + \cdots + r_{n+1} y_{n+1} = 0.$$

Proof. The quickest way of proving this is to embed R in its quotient field F (since R is an integral domain) and observe that since $M \cong R \oplus R \oplus \cdots \oplus R$ (n times) we obtain $M \subseteq F \oplus F \oplus \cdots \oplus F$. The latter is an n-dimensional vector space over F so any $n+1$ elements of M are F-linearly dependent. By clearing the denominators of the scalars (by multiplying through by the product of all the denominators, for example), we obtain an R-linear dependence relation among the $n+1$ elements of M. □

If R is any integral domain and M is any R-module recall that

$$\mathrm{Tor}(M) \coloneqq \{x \in M \mid rx = 0 \text{ for some nonzero } r \in R\}$$

is a submodule of M (called the **torsion submodule** of M) and if N is any submodule of $\mathrm{Tor}(M)$, N is called a torsion submodule of M (so the torsion submodule of M is the union of all torsion submodules of M, i.e., is the maximal torsion submodule of M). If $\mathrm{Tor}(M) = 0$, the module M is said to be **torsion free**.

For any integral domain R the **rank** of an R-module M is the maximum number of R-linearly independent elements of M. It is obvious that the rank of $\mathrm{Tor}(M)$ is 0, so that in particular any torsion R-module has rank 0.

Exercise ([1] §12.1 Exercise 1)**.** *Let M be a module over the integral domain R. Show that the rank of M is the same as the rank of the (torsion free) quotient $M/\operatorname{Tor} M$.*

The preceding proposition states that for a free R-module M over an integral domain the rank of a submodule is bounded by the rank of M. This notion of rank agrees with previous uses of the same term. If the ring $R = F$ is a field, then the rank of an R-module M is the dimension of M as a vector space over F and any maximal set of F-linearly independent elements is a basis for M.

For a general integral domain, however, an R-module M of rank n need not have a "basis", i.e., need not be a free R-module even if M is torsion free, so some care is necessary with the notion of rank, particularly with respect to the torsion elements of M.

Exercise ([1] §12.1 Exercise 20)**.** *Let R be an integral domain with quotient field F and let M be any R-module. Prove that* $\operatorname{rank}(M) = \dim_F(F \otimes_R M)$.

Exercise ([1] §12.1 Exercise 5)**.** *Torsion-free R-modules are not always free. Let $R = \mathbb{Z}[x]$ and let $M = (2, x)$ be the ideal generated by 2 and x, considered as a submodule of R. Show that $\{2, x\}$ is not a basis of M. Show that* $\operatorname{rank}(M) = 1$ *but that M is not free of rank* 1.

2 The Fundamental Theorem of finitely generated modules over a P.I.D.

2.1 The Invariant Factor Form

The next important result shows that if N is a submodule of a free module of finite rank over a P.I.D. then N is again a free module of finite rank and furthermore it is possible to choose generators for the two modules which are related in a simple way.

Theorem ([1] §12.1 Theorem 4)**.** *Let R be a P.I.D., let M be a free R-module of finite rank n and let N be a submodule of M. Then N is free of rank m, $m \leq n$, and there exists a basis $y_1, y_2, \ldots, y_n$ of M so that $a_1y_1, a_2y_2, \ldots, a_my_m$ is a basis of N, where $a_1, a_2, \ldots, a_m$ are nonzero elements of R with the divisibility relations $a_1 \mid a_2 \mid \cdots \mid a_m$.*

Sketch of proof. The theorem is trivial for $N = 0$, so assume $N \neq 0$.

For each R-module homomorphism φ of M into R, the image $\varphi(N)$ of N is a submodule of R, i.e., an ideal in R. Let

$$\Sigma \coloneqq \{\varphi(N) \mid \varphi \in \operatorname{Hom}_R(M, R)\}.$$

The collection Σ is certainly nonempty, since taking φ to be the trivial homomorphism shows that $(0) \in \Sigma$. Recall that any P.I.D. is a Noetherian ring, (by [1] §12.1 Corollary 2) Σ has at least one maximal element, i.e., there is at least one homomorphism $\nu \in \operatorname{Hom}_R(M, R)$ so that the principal ideal $\nu(N)$ is not properly contained in any other element of Σ.

Take a_1 in the P.I.D. R such that $\nu(N) = (a_1)$, and let $y \in N$ such that $\nu(y) = a_1$. One can verify that

$$\bullet\ a_1 \neq 0, \quad \text{and} \qquad \bullet\ a_1 \mid \varphi(y) \quad \text{for every } \varphi \in \operatorname{Hom}_R(M, R).$$

In particular, fix a basis $x_1, x_2, \ldots, x_n$ of the free module M, and we have $a_1 \mid \pi_i(y)$ for all i, where $\pi_i \in \operatorname{Hom}_R(M, R)$ is the natural projection homomorphism onto the i^{th} coordinate with respect to this basis. Write $\pi_i(y) = a_1b_i$ for some $b_i \in R$, $1 \leq i \leq n$ and define

$$y_1 \coloneqq \sum_{i=1}^{n} b_i x_i. \qquad \text{Note that} \quad y = \sum_{i=1}^{n} ab_i x_i$$

and $a_1y_1 = y$. Since $a_1 = \nu(y) = \nu(a_1y_1) = a_1\nu(y_1)$, $0 \neq a_1 \in R$ and R is an integral domain, this shows

$$\nu(y_1) = 1.$$

Next, one may verify that this element y_1 can be taken as one element in a basis for M and that a_1y_1 can be taken as one element in a basis for N, namely that we have

- $M = Ry_1 \oplus \ker \nu$ given by $x = \nu(x)y_1 + (x - \nu(x)y_1)$, and
- $N = Ra_1y_1 \oplus (N \cap \ker \nu)$ given by $x' = \nu(x')y_1 + (x' - \nu(x')y_1)$

(recall that $\nu(N) = (a_1)$ and hence $a_1 \mid \nu(x')$ for any $x' \in N$).

At last, one can prove the freeness of N by induction on the rank, m, of N (the main task is to show that $N \cap \ker \nu$ has rank $m - 1$); and prove the rest of the theorem by induction on n, the rank of M (by the induction assumption, there is a basis $y_2, y_3, \ldots, y_n$ of $\ker \nu$ such that $a_2y_2, a_3y_3, \ldots, a_my_m$ is a basis of $N \cap \ker \nu$ for some $a_2, a_3, \ldots, a_m \in R$ with $a_2 \mid a_3 \mid \cdots \mid a_m$; the main step is to show $a_1 \mid a_2$). □

For any submodule N of M, the **annihilator** of N is defined by

$$\mathrm{Ann}(N) \coloneqq \{r \in R \mid rn = 0 \text{ for all } n \in N\}.$$

Note that:

- $\mathrm{Ann}(N)$ is the ideal of R;
- if N is not a torsion submodule of M then $\mathrm{Ann}(N) = (0)$;
- if N, L are submodules of M with $N \subseteq L$, then $\mathrm{Ann}(L) \subseteq \mathrm{Ann}(N)$;
- if R is a P.I.D. and $N \subseteq L \subseteq M$ with $\mathrm{Ann}(N) = (a)$ and $\mathrm{Ann}(L) = (b)$, then $a \mid b$, in particular the annihilator of any element x of M divides the annihilator of M (this is implied by Lagrange's Theorem when $R = \mathbb{Z}$).

Recall that the left R-module C is a **cyclic** R-module (for any ring R, not necessarily commutative nor with 1) if there is an element $x \in C$ such that $C = Rx$. We can then define an R-module homomorphism

$$\pi : R \to C \qquad \text{by} \qquad \pi(r) \coloneqq rx,$$

which will be surjective by the assumption $C = Rx$. The First Isomorphism Theorem gives an isomorphism of (left) R-modules $R/\ker \pi \cong C$ where $\ker \pi = \mathrm{Ann}(x)$. If R is a P.I.D., $\ker \pi$ is a principal ideal (a), so we see that the cyclic R-modules C are of the form $R/(a)$ where $(a) = \mathrm{Ann}(C)$.

The cyclic modules are the simplest modules (since they require only one generator). The existence portion of the Fundamental Theorem states that any finitely generated module over a P.I.D. is isomorphic to the direct sum of finitely many cyclic modules.

Theorem ([1] §12.1 Theorems 5 & 9, Fundamental Theorem: Invariant Factor Form)**.** *Let R be a P.I.D. and let M be a finitely generated R-module. Then*

(1) (Existence) *M is isomorphic to the direct sum of finitely many cyclic modules. More precisely,*

$$M \cong R^r \oplus \mathrm{Tor}(M), \qquad \mathrm{Tor}(M) \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_m)$$

for some integer $r \geq 0$ and nonzero elements $a_1, a_2, \ldots, a_m \in R$ which are not units in R and which satisfy the divisibility relations $a_1 \mid a_2 \mid \cdots \mid a_m$.

(2) M is torsion free if and only if M is free; M is a torsion module if and only if $r = 0$, and in this case the annihilator of M is the ideal (a_m).

(3) (Uniqueness) *If we have*

$$M \cong R^{r'} \oplus R/(b_1) \oplus R/(b_2) \oplus \cdots \oplus R/(b_{m'})$$

for some integer $r' \geq 0$ and nonzero elements $b_1, b_2, \ldots, b_{m'} \in R$ which are not units with $b_1 \mid b_2 \mid \cdots \mid b_{m'}$, then $r = r'$, $m = m'$ and $(a_i) = (b_i)$ (so $a_i = b_i$ up to units) for all i. (It is precisely the divisibility condition $a_1 \mid a_2 \mid \cdots \mid a_m$ which gives this uniqueness.)

*The integer r is called the **free rank** or the **Betti number** of M, and the elements $a_1, a_2, \ldots, a_m \in R$ (defined up to multiplication by units in R) are called the **invariant factors** of M.*

Proof of Existence. The module M can be generated by a finite set of elements by assumption, so let $x_1, \ldots, x_n$ be a set of generators of M of minimal cardinality.

Let R^n be the free R-module of rank n with basis $b_1, \ldots, b_n$ and define the homomorphism $\pi : R^n \to M$ by defining $\pi(b_i) \coloneqq x_i$ for all i, which is automatically surjective since $x_1, \ldots, x_n$ generate M. By the First Isomorphism Theorem for modules we have $R^n/\ker\pi \cong M$.

Now, by [1] §12.1 Theorem 4 applied to R^n and the submodule $\ker\pi$, we can choose another basis $y_1, \ldots, y_n$ of R^n, so that $a_1y_1, \ldots, a_my_m$ is a basis of $\ker\pi$ for some elements $a_1, \ldots, a_m$ of R with $a_1 \mid a_2 \mid \cdots \mid a_m$. This implies

$$M \cong R^n/\ker\pi = (Ry_1 \oplus Ry_2 \oplus \cdots \oplus Ry_n)/(Ra_1y_1 \oplus Ra_2y_2 \oplus \cdots \oplus Ra_my_m).$$

To identify the quotient on the right hand side we use the natural surjective R-module homomorphism

$$\begin{array}{rcl} Ry_1 \oplus Ry_2 \oplus \cdots \oplus Ry_n & \longrightarrow & R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_m) \oplus R^{n-m} \\ \big(\alpha_1y_1, \alpha_2y_2, \ldots, \alpha_ny_n\big) & \longmapsto & \big(\alpha_1 + (a_1), \ldots, \alpha_m + (a_m), \alpha_{m+1}, \ldots, \alpha_n\big). \end{array}$$

The kernel of this map is clearly the set of elements where a_i divides α_i, $i = 1, 2, \ldots, m$, i.e. $Ra_1y_1 \oplus \cdots \oplus Ra_my_m$. Hence we obtain

$$M \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_m) \oplus R^{n-m}.$$

If a is a unit in R then $R/(a) = 0$, so in this direct sum we may remove any of the initial a_i which are units. This gives the decomposition with $r = n - m$.

See [1] §12.1 Theorem 9 for the proof of Uniqueness. □

2.2 The Elementary Divisor Form

Using the Chinese Remainder Theorem it is possible to decompose the cyclic modules in the above theorem further, so that M is the direct sum of cyclic modules whose annihilators are as simple as possible (namely (0) or generated by powers of primes in R).

Theorem ([1] §7.6 Theorem 17, Chinese Remainder Theorem)**.** *Let $A_1, A_2, \ldots, A_k$ be ideals in R. The map*

$$R \to R/A_1 \times R/A_2 \times \cdots \times R/A_k \quad \textit{defined by} \quad r \mapsto (r + A_1, r + A_2, \ldots, r + A_k)$$

is a ring homomorphism with kernel $A_1 \cap A_2 \cap \cdots \cap A_k$.

*If for each $i, j \in \{1, 2, \ldots, k\}$ with $i \neq j$ the ideals A_i and A_j are **comaximal** (i.e. $A_i + A_j = R$), then this map is surjective and $A_1 \cap A_2 \cap \cdots \cap A_k = A_1A_2 \cdots A_k$, so*

$$R/(A_1A_2 \cdots A_k) = R/(A_1 \cap A_2 \cap \cdots \cap A_k) \cong R/A_1 \times R/A_2 \times \cdots \times R/A_k.$$

Corollary ([1] §7.6 Corollary 18)**.** *Let n be a positive integer and let $p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$ be its factorization into powers of distinct primes. Then*

$$\mathbb{Z}/(n) \cong \mathbb{Z}/(p_1^{\alpha_1}) \times \mathbb{Z}/(p_2^{\alpha_2}) \times \cdots \times \mathbb{Z}/(p_k^{\alpha_k})$$

as rings, so in particular we have the following isomorphism of multiplicative groups:

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})^\times \times (\mathbb{Z}/p_2^{\alpha_2}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_k^{\alpha_k}\mathbb{Z})^\times.$$

If we compare orders on the two sides of this last isomorphism, we obtain the formula

$$\varphi(n) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2})\cdots\varphi(p_k^{\alpha_k})$$

for the Euler φ-function.

Suppose a is a nonzero element of the Principal Ideal Domain R. Then since R is also a Unique Factorization Domain we can write

$$a = up_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s}$$

where the p_i are distinct primes in R and u is a unit. This factorization is unique up to units, so the ideals (p_i), $i = 1, \ldots, s$ are uniquely defined.

For $i \neq j$ we have $(p_i) + (p_j) = R$ since the sum of these two ideals is generated by a greatest common divisor, which is 1 for distinct primes p_i, p_j. Put another way, the ideals (p_i), $i = 1, \ldots, s$, are comaximal in pairs. The intersection of all these ideals is the ideal (a) since a is a least common multiple of $p_1^{\alpha_1}, p_2^{\alpha_2}, \ldots, p_s^{\alpha_s}$.

The Chinese Remainder Theorem shows that

$$R/(a) \cong R/(p_1^{\alpha_1}) \times R/(p_2^{\alpha_2}) \times \cdots \times R/(p_s^{\alpha_s})$$

as rings and also as R-modules (cf. [1] §10.3 Exercise 17). Applying this to the modules in the Invariant Factor Form allows us to write each of the direct summands $R/(a_i)$ for the invariant factor a_i of M as a direct sum of cyclic modules whose annihilators are the prime power divisors of a_i. This proves the Existence part of the following:

Theorem ([1] §12.1 Theorems 6 & 9, Fundamental Theorem: Elementary Divisor Form)**.** *Let R be a P.I.D. and let M be a finitely generated R-module. Then M is the direct sum of a finite number of cyclic modules whose annihilators are either (0) or generated by powers of primes in R, i.e.,*

$$M \cong R^r \oplus R/(p_1^{\alpha_1}) \oplus R/(p_2^{\alpha_2}) \oplus \cdots \oplus R/(p_t^{\alpha_t})$$

*where $r \in \mathbb{Z}_{\geq 0}$ and $p_1^{\alpha_1}, \ldots, p_t^{\alpha_t}$ are positive powers of (not necessarily distinct) primes in R. The prime powers (defined up to multiplication by units in R) are called the **elementary divisors** of M.*

Two finitely generated R-modules M_1 and M_2 are isomorphic if and only if they have the same free rank and the same list of elementary divisors.

2.3 The Primary Decomposition

Suppose M is a finitely generated torsion module over a P.I.D. If for the distinct primes $p_1, p_2, \ldots, p_n$ occurring in the Elementary Divisor Form we group together all the cyclic factors corresponding to the same prime p_i, we see in particular that M can be written as a direct sum $M = N_1 \oplus N_2 \oplus \cdots \oplus N_n$ where N_i consists of all the elements of M which are annihilated by some power of the prime p_i. This result holds also for modules over R which may not be finitely generated:

Theorem ([1] §12.1 Theorem 7, the Primary Decomposition Theorem)**.** *Let R be a P.I.D. and let M be a nonzero torsion R-module (not necessarily finitely generated) with nonzero annihilator a. Suppose the factorization of a into distinct prime powers in R is*

$$a = up_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$$

and let $N_i := \{x \in M \mid p_i^{\alpha_i} x = 0\}$, $1 \leq i \leq n$. Then

- *N_i is a submodule of M with annihilator $p_i^{\alpha_i}$, and is the submodule of M of all elements annihilated by some power of p_i.*
- *We have $M = N_1 \oplus N_2 \oplus \cdots \oplus N_n$, and we call N_i the p_i-**primary component** of M.*
- *If M is finitely generated then each N_i is the direct sum of finitely many cyclic modules whose annihilators are divisors of $p_i^{\alpha_i}$.*
- *In particular, if M is a finite abelian group of order $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$ then, considered as a $\mathbb{Z}$-module, M is annihilated by (a), the p_i-primary component of M is the unique Sylow p_i-subgroup of M and M is isomorphic to the direct product of its Sylow subgroups.*

Proof. Exercise ([1] §10.3 Exercise 18). For reference see the proof of [1] §6.1 Theorem 3. □

Notice that with this terminology the elementary divisors of a finitely generated module M are just the invariant factors of the primary components of $\mathrm{Tor}(M)$.

2.4 The Fundamental Theorem of finitely generated Abelian groups

By taking $R = \mathbb{Z}$ in the above theorems we derive the following Fundamental Theorem for Abelian groups:

Theorem ([1] §5.2 Theorem 3, the Fundamental Theorem of Finitely Generated Abelian Groups)**.** *Let G be a finitely generated abelian group. Then*

$$G = \mathbb{Z}^r \times \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_s\mathbb{Z}$$

for some integers $r, n_1, n_2, \ldots, n_s$ satisfying that $r \geq 0$, $n_j \geq 2$ for all j, and $n_i \mid n_{i+1}$ for $1 \leq i \leq s-1$. The decomposition is unique.

Theorem ([1] §5.2 Theorem 5, the Primary Decomposition Theorem for finite abelian groups)**.** *Let G be an abelian group of order $n > 1$ and let the unique factorization of n into distinct prime powers be*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}.$$

Then

(1) $G \cong A_1 \times A_2 \times \cdots \times A_k$, where $|A_i| = p_i^{\alpha_i}$, i.e., A_i is the (unique) Sylow p_i-subgroups of G, and G is isomorphic to the direct product of its Sylow subgroups.

(2) For each $A \in \{A_1, A_2, \ldots, A_k\}$ with $|A| = p^\alpha$,

$$A \cong \mathbb{Z}/p^{\beta_1}\mathbb{Z} \times \mathbb{Z}/p^{\beta_2}\mathbb{Z} \times \cdots \mathbb{Z} \times \mathbb{Z}/p^{\beta_t}$$

with $1 \leq \beta_1 \leq \cdots \leq \beta_t$ and $\beta_1 + \cdots + \beta_t = \alpha$ (where t and $\beta_1, \ldots, \beta_t$ depend on i).

(3) The decomposition in (1) and (2) is unique, i.e., if $G \cong B_1 \times B_2 \times \cdots \times B_{k'}$ with $|B_i| = p_i^{\alpha_i}$ for all i, then $k = k'$, $B_i \cong A_i$, and B_i and A_i have the same invariant factors.

By the above theorem, in order to find all abelian groups of order $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, one must find for each i, $1 \leq i \leq k$, all possible lists of invariant factors for groups of order $p_i^{\alpha_i}$. Each list of invariant factors in this case is simply a partition of α_i (ordered in ascending order). In particular, the number of nonisomorphic abelian groups of order p^α equals the number of partitions of α. This number is independent of the prime p.

The set of elementary divisors of each abelian group is then obtained by taking one set of invariant factors from each of the k lists. The abelian groups are the direct products of the cyclic groups whose orders are the elementary divisors (and distinct lists of elementary divisors give non-isomorphic groups).

Example. *If $n = 72 = 2^3 3^2$ we list the abelian groups of this order as follows:*

Order p^α	**Partitions of α**	**Abelian Groups of Order p^α**
2^3	1, 1, 1	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$
	1, 2	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$
	3	$\mathbb{Z}/8\mathbb{Z}$
3^2	1, 1	$\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$
	2	$\mathbb{Z}/9\mathbb{Z}$

We obtain the abelian groups of order 72 by taking one abelian group from each of the two lists (right hand column above) and taking their direct product. Doing this in all possible ways gives all isomorphism types:

Abelian Groups of Order 72	**Elementary Divisors**	**Invariant Factors**
$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z})$	$2, 2, 2, 3, 3$	$2, 6, 6$
$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/9\mathbb{Z})$	$2, 2, 2, 9$	$2, 2, 18$
$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z})$	$2, 4, 3, 3$	$6, 12$
$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/9\mathbb{Z})$	$2, 4, 9$	$2, 36$
$(\mathbb{Z}/8\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z})$	$8, 3, 3$	$3, 24$
$(\mathbb{Z}/8\mathbb{Z}) \times (\mathbb{Z}/9\mathbb{Z})$	$8, 9$	72

By the Fundamental Theorems above, this is a complete list of all abelian groups of order 72 — every abelian group of this order is isomorphic to precisely one of the groups above and no two of the groups in this list are isomorphic.

We emphasize that the elementary divisors of G are not invariant factors of G (but invariant factors of subgroups of G).

Note that if a finitely generated module M is written as a direct sum of cyclic modules of the form $R/(a)$, then the ideals (a) which occur are not in general unique ($\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \cong \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ for example), unless some additional conditions are imposed (such as the divisibility condition for the invariant factors, or the condition that a be the power of a prime in the case of the elementary divisors). To decide whether two modules are isomorphic it is necessary to first write them in such a standard form.

Exercise ([1] §5.2 Exercise 9)**.** *Let $A = \mathbb{Z}/60\mathbb{Z} \times \mathbb{Z}/45\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/36\mathbb{Z}$. Find the number of elements of order 2 and the number of subgroups of index 2 in A.*

3 Canonical forms of matrices

We now apply our results on finitely generated modules in the special case where the P.I.D. is the ring $F[x]$ of polynomials in x with coefficients in a field F. Recall that (cf. [1] §10.1) there is a bijection between the collection of $F[x]$-modules and the collection of pairs (V, T)

$$\left\{ \begin{array}{c} V \text{ an } F[x]\text{-module} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} V \text{ a vector space over } F \\ \text{and} \\ T : V \to V \text{ a linear transformation} \end{array} \right\}$$

given by

$$\text{the element } x \text{ acts on } V \text{ as the linear transformation } T.$$

In terms of this bijection,

$$\left\{ \begin{array}{c} W \text{ an } F[x]\text{-submodule of } V \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} W \text{ a subspace of } V \\ \text{and} \\ W \text{ is } T\text{-stable} \end{array} \right\}.$$

Let V be a finite dimensional vector space over F of dimension n and let T be a fixed linear transformation of V. Since V has finite dimension over F by assumption, it is by definition finitely generated as an F-module, hence certainly finitely generated as an $F[x]$-module, so the classification theorems of the preceding section apply.

Any nonzero free $F[x]$-module (being isomorphic to a direct sum of copies of $F[x]$) is an infinite dimensional vector space over F, so if V has finite dimension over F then it must in fact be a torsion $F[x]$-module (i.e., its free rank is 0). It follows from the Fundamental Theorem that then V is isomorphic as an $F[x]$-module to the direct sum of cyclic, torsion $F[x]$-modules.

We shall see that this decomposition of V will allow us to choose a basis for V with respect to which the matrix representation for the linear transformation T is in a specific simple form. When we use the **invariant factor decomposition** of V we obtain the **rational canonical form** for the matrix for T; when we use the **elementary divisor decomposition** (and when F contains all the eigenvalues of T) we obtain the **Jordan canonical form**, as the matrix representing T which is as close to being a diagonal matrix as possible. The uniqueness portion of the Fundamental Theorem ensures that the rational and Jordan canonical forms are unique (which is why they are referred to as canonical).

One important use of these canonical forms is to classify the distinct linear transformations of V. In particular they allow us to determine when two matrices represent the same linear transformation, i.e., when two given $n \times n$ matrices are similar.

3.1 The rational canonical form

If we fix a basis $\mathcal{B}$ of V, then any linear transformation T of V has an associated $n \times n$ matrix A. Conversely, if A is any $n \times n$ matrix then the map T defined by $T(v) := Av$ for $v \in V$, where the v on the right is the $n \times 1$ vector consisting of the coordinates of v with respect to the fixed basis $\mathcal{B}$ of V, is a linear transformation of V.

3.1.1 The annihilator and minimal polynomial

Recall that λ is an **eigenvalue** of the linear transformation T (i.e. there is a nonzero vector $v \in V$ such that $Tv = \lambda v$), if and only if $\lambda I - T$ is a singular linear transformation of V, if and only if $\det(\lambda I - T) = 0$. Let x be an indeterminate over F. The polynomial $\det(xI - T)$ is called the

characteristic polynomial of T and will be denoted $c_T(x)$. If A is an $n \times n$ matrix with coefficients in F, $\det(xI - A)$ is called the **characteristic polynomial** of A and will be denoted $c_A(x)$. It is easy to see by expanding the determinant that the characteristic polynomial of either T or A is a monic polynomial of degree $n = \dim V$, and the set of eigenvalues of T (or A) is precisely the set of roots of the characteristic polynomial of T (of A, respectively). In particular, T has at most n distinct eigenvalues.

We have seen that V considered as a module over $F[x]$ via the linear transformation T is a torsion $F[x]$-module. Let $m(x) \in F[x]$ be the unique monic polynomial generating $\mathrm{Ann}(V) \subseteq F[x]$. Equivalently, $m(x)$ is the unique monic polynomial of minimal degree annihilating V, i.e., such that $m(T)$ is the 0 linear transformation, and if $f(x) \in F[x]$ is any polynomial annihilating V, $m(x)$ divides $f(x)$.

Since the ring of all $n \times n$ matrices over F is isomorphic to the collection $\mathrm{End}(V)$ of all linear transformations of V to itself (an isomorphism is obtained by choosing a basis for V), it follows that for any $n \times n$ matrix A over F there is similarly a unique monic polynomial of minimal degree with $m(A)$ the zero matrix.

The unique monic polynomial which generates the ideal $\mathrm{Ann}(V)$ in $F[x]$ is called the **minimal polynomial** of T and will be denoted $m_T(x)$. The unique monic polynomial of smallest degree which when evaluated at the matrix A is the zero matrix is called the **minimal polynomial** of A and will be denoted $m_A(x)$.

We shall shortly prove that $c_T(x) \in \mathrm{Ann}(V)$, i.e. the minimal polynomial $m_T(x)$ for T is a divisor of the characteristic polynomial $c_T(x)$ for T (this is the Cayley–Hamilton Theorem), and similarly for A, so in fact the degrees of minimal polynomials are at most n.

The invariant factor decomposition of V gives an isomorphism

$$V \cong F[x]/\big(a_1(x)\big) \oplus F[x]/\big(a_2(x)\big) \oplus \cdots \oplus F[x]/\big(a_m(x)\big) \tag{12.1}$$

of $F[x]$-modules where $a_1(x), a_2(x), \ldots, a_m(x)$ are polynomials in $F[x]$ of degree at least one with the divisibility conditions $a_1(x) \mid a_2(x) \mid \ldots \mid a_m(x)$. These invariant factors $a_i(x)$ are only determined up to a unit in $F[x]$, but since the units of $F[x]$ are precisely the nonzero elements of F (i.e., the nonzero constant polynomials), we may make these polynomials unique by stipulating that they be **monic**.

Since the annihilator of V is the ideal $(a_m(x))$ (cf. the Fundamental Theorem), we immediately obtain:

Proposition ([1] §12.2 Proposition 13)**.** *The minimal polynomial $m_T(x)$ is the largest invariant factor of V. All the invariant factors of V divide $m_T(x)$.*

3.1.2 Cyclic submodules and companion matrices

We now choose a basis for each of the direct summands $F[x]/\big(a_i(x)\big)$ for V in the decomposition (12.1) above for which the matrix for T is quite simple. Recall that the linear transformation T acting on the left side of (12.1) is the element x acting by multiplication on each of the factors on the right side of the isomorphism in (12.1).

Recall that the elements $1, \bar{x}, \bar{x}^2, \ldots, \bar{x}^{k-1}$ give a basis for the vector space $F[x]/(a(x))$, where $a(x) = x^k + b_{k-1}x^{k-1} + \cdots + b_1 x + b_0$ is any monic polynomial in $F[x]$, and $\bar{x} \coloneqq x \bmod (a(x))$.

With respect to this basis the linear transformation of multiplication by x acts in a simple manner:

$$\times x: \quad \begin{array}{lcl} 1 & \mapsto & \bar{x} \\ \bar{x} & \mapsto & \bar{x}^2 \\ \bar{x}^2 & \mapsto & \bar{x}^3 \\ & \vdots & \\ \bar{x}^{k-2} & \mapsto & \bar{x}^{k-1} \\ \bar{x}^{k-1} & \mapsto & \bar{x}^k = -b_0 - b_1\bar{x} - \cdots - b_{k-1}\bar{x}^{k-1} \end{array}$$

where the last equality is because $\bar{x}^k + b_{k-1}\bar{x}^{k-1} + \cdots + b_1\bar{x} + b_0 = 0$ since $a(\bar{x}) = 0$ in $F[x]/(a(x))$. With respect to this basis, the matrix for multiplication by x is determined by

$$(\times x)\begin{bmatrix} 1 & \bar{x} & \bar{x}^2 & \cdots & \bar{x}^{k-1} \end{bmatrix} = \begin{bmatrix} 1 & \bar{x} & \bar{x}^2 & \cdots & \bar{x}^{k-1} \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & \cdots & \cdots & -b_0 \\ 1 & 0 & \cdots & \cdots & \cdots & -b_1 \\ 0 & 1 & \cdots & \cdots & \cdots & -b_2 \\ 0 & 0 & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & -b_{k-1} \end{bmatrix}.$$

This is the **companion matrix** of the monic polynomial $a(x)$, which is the $k \times k$ matrix with 1's down the first subdiagonal, $-b_0, -b_1, \ldots, -b_{k-1}$ down the last column and zeros elsewhere. We denote the companion matrix of $a(x)$ by $\mathcal{C}_{a(x)}$.

3.1.3 The rational canonical form

We apply the above result to each of the cyclic modules on the right side of (12.1) above and let $\mathcal{B}_i$ be the elements of V corresponding to the basis chosen above for the cyclic factor $F[x]/(a_i(x))$ under the isomorphism in (12.1). Then by definition the linear transformation T acts on $\mathcal{B}_i$ by the companion matrix for $a_i(x)$ (since we have seen that this is how multiplication by x acts).

The union $\mathcal{B}$ of the $\mathcal{B}_i$'s gives a basis for V (since the sum on the right of (12.1) is direct), and with respect to this basis, the linear transformation T has matrix

$$\begin{bmatrix} \mathcal{C}_{a_1(x)} & & & \\ & \mathcal{C}_{a_2(x)} & & \\ & & \ddots & \\ & & & \mathcal{C}_{a_m(x)} \end{bmatrix} \tag{12.2}$$

i.e., the direct sum of the companion matrices for the invariant factors.

Notice that this matrix is uniquely determined from the invariant factors of the $F[x]$-module V, and the list of invariant factors uniquely determines the module V up to isomorphism as an $F[x]$-module.

Theorem ([1] §12.2 Theorem 14, Rational Canonical Form)**.** *Let V be a finite dimensional vector space over the field F and let T be a linear transformation of V. Then there is a basis for V, with respect to which the matrix for T is in **rational canonical form**, i.e., is a **block diagonal** matrix whose diagonal blocks are the companion matrices for monic polynomials $a_1(x), a_2(x), \ldots, a_m(x)$ of degree at least one with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$. The rational canonical form for T is unique.*

The use of the word **rational** is to indicate that this canonical form is calculated entirely within the field F and exists for any linear transformation T. This is not the case for the Jordan canonical form (considered later), which only exists if the field F contains the eigenvalues for T. Moreover, [1] §12.2 Corollary 18 shows that the rational canonical form for an $n \times n$ matrix A is an $n \times n$ matrix with entries in the smallest field containing the entries of A; further, this canonical form is the same matrix even if we allow conjugation of A by nonsingular matrices whose entries come from larger fields.

The following result translates the notion of similar linear transformations (i.e., the same linear transformation up to a change of basis) into the language of modules and relates this notion to rational canonical forms.

Theorem ([1] §12.2 Theorem 15)**.** *Let S and T be linear transformations of V. Then TFAE:*

(1) S and T are similar linear transformations

(2) the $F[x]$-modules obtained from V via S and via T are isomorphic $F[x]$-modules

(3) S and T have the same rational canonical form.

Exercise (cf. [1] §12.2 Example (5) p.487)**.** *Find all similarity classes of 3×3 matrices A over $\mathbb{Q}$ and matrices over $\mathbb{F}_2$ satisfying $A^3 = I$.*

Exercise ([1] §12.2 Exercise 15)**.** *Determine up to similarity all 2×2 rational matrices (i.e., $\in M_2(\mathbb{Q})$) of precise order 4 (multiplicatively). Do the same if the matrix has entries from $\mathbb{C}$.*

3.1.4 Invariant factors

We shall see below how to calculate not only the minimal polynomial for T but also the other invariant factors.

Let $a(x) \in F[x]$ be any monic polynomial. By direct calculation one can verify that the characteristic polynomial $c_{\mathcal{C}_{a(x)}}(x)$ of the companion matrix $\mathcal{C}_{a(x)}$ is $a(x)$; and if M is the block diagonal matrix

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

given by the direct sum of matrices $A_1, A_2, \ldots, A_k$, then the characteristic polynomial $c_M(x)$ of M is $c_{A_1}(x)c_{A_2}(x)\cdots c_{A_k}(x)$.

Proposition ([1] §12.2 Proposition 20)**.** *Let A be an $n \times n$ matrix over the field F.*

(1) The characteristic polynomial of A is the product of all the invariant factors of A.

(2) (The Cayley–Hamilton Theorem) *The minimal polynomial $m_A(x)$ of A divides the characteristic polynomial $c_A(x)$ of A. In particular $c_A(A) = 0$ as matrices.*

(3) The characteristic polynomial of A divides some power of the minimal polynomial of A. In particular these polynomials have the same roots, not counting multiplicities.

The same statements are true if the matrix A is replaced by a linear transformation T of an n-dimensional vector space over F.

The above proposition are frequently quite useful in the determination of the invariant factors for a matrix A, particularly for matrices of small degree. The following result computes the invariant factors in general.

Theorem ([1] §12.2 Theorem 21)**.** *Let A be an $n \times n$ matrix over the field F. The three operations*
(a) *interchanging two rows or columns*
(b) *adding a multiple (in $F[x]$) of one row or column to another*
(c) *multiplying any row or column by a unit in $F[x]$ (i.e., by a nonzero element in F),*
*are called **elementary row and column operations** in $M_n(F[x])$. Using these elementary operations, the $n \times n$ matrix $xI - A$ with entries from $F[x]$ can be put into the diagonal form (called the **Smith Normal Form** for A)*

$$\begin{bmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & a_1(x) & & & & \\ & & & & a_2(x) & & & \\ & & & & & \ddots & & \\ & & & & & & a_m(x) \end{bmatrix}$$

with monic elements $a_1(x), a_2(x), \ldots, a_m(x) \in F[x]$ with degrees at least one and satisfying $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$. The polynomials $a_1(x), \ldots, a_m(x)$ are the invariant factors of A.

Proof. Cf. [1] §12.1 Exercises 16 to 19. □

3.2 The Jordan canonical form

In this section we use the elementary divisor form of the Fundamental Theorem to obtain the Jordan canonical form. Matrices in this canonical form are as close to being diagonal matrices as possible, so the matrices are simpler than in the rational canonical form (but we lose some of the "rationality" results).

The elementary divisors of a module are the prime power divisors of its invariant factors. For the $F[x]$-module V the invariant factors were monic polynomials $a_1(x), a_2(x), \ldots, a_m(x)$ of degree at least one (with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$), so the associated elementary divisors are the powers of the irreducible polynomial factors of these polynomials. These polynomials are only defined up to multiplication by a unit and, as in the case of the invariant factors, we can specify them uniquely by requiring that they be monic.

To obtain the simplest possible elementary divisors we shall assume that the polynomials $a_1(x), a_2(x), \ldots, a_m(x)$ factor **completely into linear factors**, i.e., that the elementary divisors of V are powers $(x - \lambda)^k$ of linear polynomials. Since the product of the elementary divisors is the characteristic polynomial, this is equivalent to the assumption that **the field F contains all the eigenvalues** of the linear transformation T (equivalently, of the matrix A representing the linear transformation T).

Under this assumption on F, it follows immediately from Elementary Divisor Form of the Fundamental Theorem that, V is the direct sum of finitely many cyclic $F[x]$-modules of the form $F[x]/(x - \lambda)^k$, where $\lambda \in F$ is one of the eigenvalues of T, corresponding to the elementary divisors of V.

We now choose a vector space basis for each of the direct summands corresponding to the elementary divisors of V for which the corresponding matrix for T is particularly simple. Recall

that by definition of the $F[x]$-module structure the linear transformation T acting on V is the element x acting by multiplication on each of the direct summands $F[x]/(x-\lambda)^k$.

Consider the elements

$$(\bar{x}-\lambda)^{k-1},\ (\bar{x}-\lambda)^{k-2},\ \ldots\ ,\ (\bar{x}-\lambda),\ 1 \in F[x]/(x-\lambda)^k.$$

Expanding each of these polynomials in $\bar{x}$ we see that the matrix relating these elements to the F-basis $\bar{x}^{k-1}, \bar{x}^{k-2}, \ldots, \bar{x}, 1$ of $F[x]/(x-\lambda)^k$ is upper triangular with 1's along the diagonal. Since this is an invertible matrix, it follows that the elements above are an F-basis for $F[x]/(x-\lambda)^k$.

With respect to this basis the linear transformation of multiplication by x acts in a particularly simple manner (note that $x = \lambda + (x-\lambda)$ and that $(\bar{x}-\lambda)^k = 0$ in the quotient $F[x]/(x-\lambda)^k$):

$$\begin{array}{lllllll}
 & (\bar{x}-\lambda)^{k-1} & \mapsto & \lambda\cdot(\bar{x}-\lambda)^{k-1} & + & (\bar{x}-\lambda)^k & = \quad \lambda\cdot(\bar{x}-\lambda)^{k-1} \\
 & (\bar{x}-\lambda)^{k-2} & \mapsto & \lambda\cdot(\bar{x}-\lambda)^{k-2} & + & (\bar{x}-\lambda)^{k-1} & \\
\times x: & & \vdots & & & & \\
 & (\bar{x}-\lambda) & \mapsto & \lambda\cdot(\bar{x}-\lambda) & + & (\bar{x}-\lambda)^2 & \\
 & 1 & \mapsto & \lambda\cdot 1 & + & (\bar{x}-\lambda). &
\end{array}$$

With respect to this basis, the matrix for multiplication by x is therefore given by

$$(\times x)\begin{bmatrix}(\bar{x}-\lambda)^{k-1} & (\bar{x}-\lambda)^{k-2} & \cdots & 1\end{bmatrix} = \begin{bmatrix}(\bar{x}-\lambda)^{k-1} & (\bar{x}-\lambda)^{k-2} & \cdots & 1\end{bmatrix}\begin{bmatrix}\lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda\end{bmatrix},$$

where the blank entries are all zero. The $k\times k$ matrix with λ along the main diagonal and 1 along the first superdiagonal depicted above is called the $k\times k$ **elementary Jordan matrix with eigenvalue** λ or the **Jordan block of size** k **with eigenvalue** λ.

Applying this to each of the cyclic factors of V in its elementary divisor decomposition we obtain a vector space basis for V with respect to which the linear transformation T has as matrix the direct sum of the Jordan blocks corresponding to the elementary divisors of V, i.e., is block diagonal with Jordan blocks along the diagonal:

$$\begin{bmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t\end{bmatrix}.$$

Notice that this matrix is uniquely determined up to permutation of the blocks along the diagonal by the elementary divisors of the $F[x]$-module V, and conversely the list of elementary divisors uniquely determines the module V up to $F[x]$-module isomorphism.

Theorem ([1] §12.3 Theorem 22, Jordan Canonical Form)**.** *Let V be a finite dimensional vector space over the field F and let T be a linear transformation of V. Assume F contains all the eigenvalues of T. Then there is a basis for V with respect to which the matrix for T is in* ***Jordan canonical form****, i.e., is a block diagonal matrix whose diagonal blocks are the Jordan blocks for the elementary divisors of V. The Jordan canonical form for T is unique up to a permutation of the Jordan blocks along the diagonal.*

The Jordan canonical form differs from a diagonal matrix only by the possible presence of some 1's along the first superdiagonal (and then only if there are Jordan blocks of size greater than one), hence is close to being a diagonal matrix. The following result shows in particular that the Jordan canonical form for a matrix A is as close to being a diagonal matrix as possible.

Corollary ([1] §12.3 Corollarys 24 & 25)**.**

(1) If a matrix A is similar to a diagonal matrix D, then D is the Jordan canonical form of A.

(2) Two diagonal matrices are similar if and only if their diagonal entries are the same up to a permutation.

(3) If A is an $n \times n$ matrix with entries from F and F contains all the eigenvalues of A, then A is similar to a diagonal matrix over F if and only if the minimal polynomial of A has no repeated roots.

Exercise ([1] §12.3 Exercise 49)**.** *Let A be an $n \times n$ matrix with entries from the field K, where K is either the real or complex numbers. Define the **exponential** of A by the convergent series (entry by entry)*

$$\exp A = e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots .$$

Prove that $\det(e^A) = e^{\mathrm{tr}(A)}$, where $\mathrm{tr}(A)$ is the trace of A (the sum of the diagonal entries of A). Hint: prove this for upper triangular matrices first.

Other related exercises in [1]

§12.1 2 4 6 7 8 10 11 14
§12.2 4 8 10 12 15 17
§12.3 2 21 22 24 26

References

[1] D. S. Dummit and R. M. Foote. *Abstract algebra*. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.