Chapter 4

Modules and presentations

4.1 Definition, examples, and basic concepts

4.1.1 Modules

Let R be a commutative ring. An R- *module* `[der Modul, die Moduln]` M is a commutative group $(M, +, -, 0)$ (we denote the axioms by (V1-4)) together with an action of R on M, i.e. with each *scalar* $r \in R$ and $v \in M$ one has associated a unique $rv \in M$ such that

$$(V5) \quad \text{for all } r \text{ in } K \text{ and } v, w \text{ in } V \text{ it holds } r(v+w) = rv + rw$$

$$(V6) \quad \text{for all } v \text{ in } V \text{ it holds } 1v = v$$

$$(V7) \quad \text{for all } r, s \text{ in } K \text{ and } v \text{ in } V \text{ it holds } (r+s)v = rv + sv$$

$$(V5) \quad \text{for all } r, s \text{ in } K \text{ and } v \text{ in } V \text{ it holds } r(sv) = (rs)v\,.$$

For commutative R it does not matter whether we write rv or vr. For non-commutative is does matter and vr would fit better to the usual notations of Linear Algebra (in all countries where by law one drives and writes on the wrong side). If you see an vr, occasionally, in these notes, read it as rv.

Examples.

a. if K is a field then the K-modules are exactly the K-vector spaces

b. Each ring R is an R-module with $rv = r \cdot v$

c. R^n is an R-module for each ring R.

d. Each commutative group is a $\mathbb{Z}$-module with nv as defined defined recursively for $n \in \mathbb{N}$ by $0v = 0_V$, $(n+1)v = nv + v$ and with $(-n)v = -(nv)$.

One has the general associative-commutative law for addition and the distributive laws (Proof as exercise)

$$r(\sum_{i=1}^{n} \vec{v}_i) = \sum_{i=1}^{n} r\vec{v}_i, \quad [\sum_{i=1}^{n} r_i]\vec{v} = \sum_{i=1}^{n} r_i\vec{v}$$

$$0v = r0 = 0, \ (-r)v = -(rv).$$

In particular, each term is equivalent to one of the from

$$\sum_{i=1}^{n} r_i x_i$$

4.1.2 Submodules and homomorphisms

These are defined in analogy to vector spaces. U is an R- *submodule* of the R-module M if it is a subgroup and $r\boldsymbol{u} \in U$ for all $r \in R$ and $\boldsymbol{u} \in U$. The submodule *generated* by a subset E is

$$\mathsf{Span}_R(E) = \{\sum_{i=1}^{n} r_i \vec{v}_i \mid n \in \mathbb{N},\ r_i \in R,\ \vec{v}_i \in E\}$$

In particular,the submodule generated by a single element v is given as $Rv = \{rv \mid r \in R\}$ an called *cyclic*. If we consider R as an R module, then we also write $(v) = RV$.

A map ϕ between R-modules M and N is R- *linear* or an *homomorphims* if

$$\phi(\vec{x} + \vec{y}) = \phi\vec{x} + \phi\vec{y},\ \ \phi(r\vec{x}) = r\phi\vec{x} \ \text{ for all } \vec{x}, \vec{y} \in M,\ r \in R$$

Congruence relations are associated with submodules and homomorphism as for vector spaces (cf Ch.11) and direct sums and products behave as well. But be aware that not every submodule U gives rise to a direct decomposition $M - U \oplus W$ - consider the $\mathbb{Z}$-submodule $U = 2\mathbb{Z}$ of $\mathbb{Z}$, Thus, sect.11.2.5 and 11.3.7 do not extend to modules.

4.1.3 K[x]-modules

Every K-vector space V is an $\mathsf{End}(V)$-module

$$\phi \cdot \boldsymbol{x} = \phi(\boldsymbol{x})$$

if we allow a non-commutative ring. Here, writing scalars form K and endomorphisms on different sides of vectors would be reasonable - if is mandatory if K is not commutative.

Recall the polynomial ring $K[x]$ with coefficents from the field K. It elements are polynomials

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

Given a K-vector space V and an endomorphism ϕ we can evaluate $p(x)$ at ϕ (in the commutative K-subalgebra of $\mathsf{End}(V)$ generated by ϕ

$$p(\phi) = a_n \phi^n + \ldots + a_1 \phi + a_0 \mathsf{id}$$

The map $p(x) \mapsto p(\phi)$ is a K-algebra homomorphims of $K[x]$ into $\mathsf{End}(V)$. Also, given $A \in K^{n\times n}$ we can evaluate at A

$$p(A) = a_n A^n + \ldots + a_1 A + E$$

and if A corresponds to ϕ w.r.t. a give basis of V then $p(A)$ corresponds to $p(\phi)$.

Proposition 4.1.1 *Given a field K, there is a 1-1-correspondence between $K[x]$-modules and pairs (V, ϕ) where V is a K-vector space and ϕ and endomorphism of V. It is given by*

$$\phi(v) = xv \quad \textit{for all } v \in V$$

Moreover, the $K[x]$-submodules are exactly the ϕ-invariant subspaces.

Proof. Given a $K[x]$-module V, V is also a K-vector space since K is a subring of $K[x]$. Due to the module laws and commutativity of $K[x]$, $\phi(v) = xv$ defines an endomorphism of V. Conversely, given $\phi \in \mathsf{End}(V)$, define

$$p(x)\vec{v} = p(\phi)(\vec{v})$$

Since evaluation of polynomials is a homomorphism, this defines a $K[x]$-module. Observe that U is a $K[x]$-submodule if it is a K-vector subspace and $xv \in U$ for all $v \in U$. Indeed, $(\sum_i r_i x^i)v = \sum_i r_i x^i v \in U$ for all $v \in U$. $\square$

We denote this $K[x]$-module by ${}_{K[\phi]}V$. It contains all information about ϕ in a convenient form. In particular, it gives an easy access to transformations into canonical form.

4.2 Free modules and presentations

4.2.1 Modular philosophy of freeness

We need an understanding of module computations form a logic background - in the structural disguise this means to understand free modules. Recall the view of $K[x]$ as a free K-algebra with generator x.

The free R-module with generators $e_1, \ldots, e_n$ (fixed R) can (and should) be understood as follows: consider the algebraic structure T all terms which can be be constructed from $e_1, \ldots, e_n$ by addition, subtraction, constant 0, and multiplications with 'scalars from R, i.e.

- $e_1, \ldots, e_n$ and 0 are terms
- If s, t are terms, then so are $s + t$, $-t$ and rt for $r \in R$.

and compute modulo ($\sim$) the laws of R-modules - using the general rules of equational logic. In other words, $\sim$ is the coarsest congruence relation on T such that that for all $a, b, c \in T$ and $r, s \in R$

$$a + (b + c) \sim (a + b) + c,\ a + b \sim b + a,\ a + 0 \sim a,\ -a + a \sim 0$$

$$1a \sim a,\ r(a + b) \sim ra + rb,\ (rs)a \sim r(sa),\ (r + s)a \sim ra + sa$$

and we call $T/\sim$ the R-module *freely generated* by $e_1, \ldots, e_n$. Actally, if $\pi : T \to T/\sim$ is the canonical projection, then $\pi(e_1), \ldots, \pi(e_n)$ are *free generators* of the R-module $T/\sim$.

Principle 4.2.1 *For any ring R, the R-module $F = T/\sim$* freely generated *by $e_1, \ldots, e_n$ has the following* universal property

- *F is generated by $\pi(e)_1 m \ldots, \pi(e_n)$ and for any R-module N and $w_i \in N$ there is a (unique) homomorphism $\phi : F \to N$ such that $\phi(\pi(e_i)) = w_i$ for $i = 1, \ldots, n$.*

and is characterized by this property up to isomorphism (matching free generators). ϕ is surjective if N is generated by the w_i.

If there is no danger of confusion, we use e_i to denote $\pi(e_i)$. This does not mean that $e_i = e_j \Leftrightarrow \pi(e_i) = \pi(e_j)$.

Proof. Given $w_i \in N$ we may evaluate terms $t = t(e_1, \ldots, e_n) \in T$

$$\psi(t) = t(w_1, \ldots, w_n) \in N$$

$\psi : T \to N$ is a homomorphism and $a \sim b$ implies $\psi(a) = \psi(b)$ since N is an R-module. By the Homomorphims Theorems, there is a homomorphism $\phi : F \to N$ such that $\psi = \phi \circ \pi$.

Assume that we have R-modules F_i and $\pi_i : \{e_1, \ldots, e_n\} \to F_i$ both with the universal property, $i = 1, 2$ Then we have $\phi_{ij} : F_i \to F_j$ such that $\phi_{ij}(\pi_i(e_k)) = \pi_j(e_k)$ for all k. It follows $\phi_{ji}\phi_{ij}(\pi(e_k)) = \pi(e_k)$ whence $\phi_{ji}\phi_{ji} = \mathsf{id}_{F_i}$. Thus, ϕ_{12} and ϕ_{21} are mutually inverse isomorphisms. □

4.2.2 Bases

$e_1, \ldots, e_n$ is a *basis* of the R-module M if $E = \{e_1, \ldots, e_n\}$ generates M and and if they are *independent*

$$r_1 e_1 + \ldots r_n e_n = 0 \ \Rightarrow r_1 = \ldots = r_n = 0 \quad \text{for all } r_i \in R$$

Corollary 4.2.2 *For an R-module M and $e_1, \ldots, e_n$ in M t.f.a.e.*

(1) $e_1, \ldots, e_n$ is a basis of M

(2) The elements of M have unique representation $a = r_1 e_1 + \ldots + r_n e_n$ with $r_i \in R$

(3) There is an isomorphism $\phi : M \to R^n$ such that $\phi e_i = \boldsymbol{e}_i$ for $i = 1, \ldots, n$

(4) M is freely generated by $e_1, \ldots, e_n$ as an R-module

Proof. (1) $\Leftrightarrow$ (2) Existence of representation means that the e_i generate, uniqueness means independence. (3) $\Rightarrow$ (2) is obvious. (2) $\Rightarrow$ (3): One has well and necessarily so defined $\phi(a) = \sum_i r_i \boldsymbol{e}_i$. This is R-linear -as is easily checked. (4) $\Rightarrow$ (2): Choose $N = R^n$. There is linear $\phi : M \to R^n$ such that $\phi e_i = \boldsymbol{e}_i$. Now, if $\sum_i r_i e_i \sum_i s_i e_i$ in M then $\sum_i r_i \phi(e_i) = \sum_i s_u \phi(e_i)$, whence $r_i = s_i$. (3) $\Rightarrow$ (4). We know that a free module F with generators $v_1, \ldots, v_n$ exists. Let $\psi : F \to R^n$ the homomorphism with $\psi(v_i) = \boldsymbol{e}_i$. The elements of F have a representation $\sum_i r_i e_i$ and this is unique by the preceeeding argument. Thus, $F \cong R^n$ and R^n is freely generated by the $\boldsymbol{e}_i$. Then M is freely generated by the e_i due to the isomorphism $\phi^{-1}\psi : F \to M$. □.

4.2.3 Presentation of modules

If we say that the R-module M is given by *generators* $e_1, \ldots, e_n$ and *relations* $a_i \stackrel{!}{=} b_i\,(i \in I)$ (which together make a *presentation*) then we mean that we calculate with R-module terms in the $e_1, \ldots, e_n$

- using the general rules of equational logic
- modulo the R-module laws (this includes the tables describing the ring R), i.e. we may substitute in these lwas any terms for the variables
- modulo the equalities $a_i = b_i$ (no substitution for an e_i!!!)

If no relations are given, then we calculate only modulo R-module laws and obtain the free R-module with generators $e_1, \ldots, e_n$.

Actually, we should consider the e_i as *generator symbols* which are interpreted in modules N by elements e_i^N. The relation $a \stackrel{!}{=} b$ is given as a formal expression by a pair (a, b) of terms $a = a(e_1, \ldots, e_n)$ and $b = b(e_1, \ldots, e_n)$ and it is valid in the module N under the interpretation e_i^N if and only if

$$a(e_1^N, \ldots, e_n^N) = b(e_1^N, \ldots, e_n^N) \quad \text{holds in } N$$

Principle 4.2.3 *Let M be an R-module and $e_1, \ldots, e_n \in M$. Then the R-module M is given by the* generators $e_1, \ldots, e_n$ *and the* relations $a_i \stackrel{!}{=} b_i\,(i \in I)$ *if and only if M is generated by the e_i^M and if for any R-module N and interpretation e_i^N there is a homomorphism $\phi : M \to N$ such that $\phi(e_i^M) = e_i^N$ for $i = 1, \ldots, n$ (which is surjective if N is generated by the e_i^N). Moreover, M is determined by the presentation up to isomorphism and is obtained form the free R-module F with generators $e_1, \ldots, e_n$ as $F/\sim$ where $\sim$ is the finest congruence relation such that $a_i \sim b_i$ for all $i \in I$ - corresponding to the submodule U of F generated by the $a_i - b_i$ (more precisely, the elements $a_i(e_1^F, \ldots, e_n^F) - b(e_1^F, \ldots, e_n^F)$ of F $(i \in I)$. Thus, $M \cong F/U$.*

Observe that any relation $a \stackrel{!}{=} b$ may be equivalently replaced by $a - b \stackrel{!}{=} 0$.

Corollary 4.2.4 *The free R-mode with generators $e_1, \ldots, e_n$ and relations $w_i \stackrel{!}{=} 0$, $(i \in I)$ is obtained, up to isomorphism, as R^n/U with generators $\pi(\boldsymbol{e}_1), \ldots, \pi(\boldsymbol{e}_n)$ where $\pi : R^n \to R/U$ is the canonical homomorphism and*

$$U = \mathsf{Span}_R\{w_i(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \mid i \in I\}.$$

Recall, that $\pi(v) = U + v = v + U$ is a popular notation.

4.2.4 Cyclic one-relation $K[x]$-modules

We consider R-modules presented with single generator g and a single relation $w \stackrel{!}{=} 0$. Then w is equivalent to a term dg with $d \in R$. If $R = \mathbb{Z}$ we obtain $\mathbb{Z}/\mathbb{Z}d$, the integers modulo d.

Lemma 4.2.5 *$K[x]/(m(x))$ is a commutative K-algebra and the canonical homomorphism $\pi : K[x] \to K[x]/(m(x))$ is a K-algebra homomorphism.*

Theorem 4.2.6 *Let $m(x) = x^n + r_{n-1}x^{n-1} + \dots r_1 x + r_0 \in K[x]$. The $K[x]$-modules given by a presentation with one generator v_0 and the relation $m(x)v_0 \stackrel{!}{=} 0$ are exactly the K-vector spaces V with endomorphism ϕ where V has basis α w.r.t. to which the matrix of ϕ is*

$$\phi^\alpha = A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -r_0 \\ 1 & 0 & 0 & \dots & 0 & -r_1 \\ 0 & 1 & 0 & & 0 & -r_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & & 0 & -r_{n-2} \\ 0 & 0 & & \dots & 1 & -r_{n-1} \end{pmatrix}$$

Then this basis is

$$\alpha : v_0, \phi(v_0), \phi^2(v_0), \dots, \phi^{n-1}(v_0)$$

A is the *Frobenius*-matrix or *companion matrix* of the polynomial $m(x)$.

Proof. By Cor.4.2.4, up to isomorphism, the module given by the presentation is $V = K[x]/(m(x))$ with generator 1. The canonical homomorphism $\pi : K[x] \to V$ is surjective and with some precaution we may use the elements $p(x)$ of $K[x]$ to denote their images in V where $\pi(p(x)) = p(x) + (m(x))$ is meant.

By Prop.4.1.1, $\phi(v) = xv$ is an endomorphism of the K-vector space V. We claim that

$$1, x, \dots, x^{n-1}$$

is a basis of the K-vector space V. From

$$(*) \quad x^n = -(r_{n-1}x^{n-1} + \dots r_1 x + r_0)$$

is follows that $\mathsf{Span}_K\{1, x, \dots, x^{n-1}\}$ is a ϕ-invariant subspace, hence a $K[x]$ submodule and equal V since it contains the generator 1. Now, consider $s_i \in K$ with

$$s_0 + s_1 x + \dots + s_{n-1}x^{n-1} = 0$$

more precisely

$$s_0\pi(1) + s_1\pi(x) + \dots + s_{n-1}\pi(x^{n-1}) = \pi(0)$$

Since by Lemma 4.2.5 π is a K-algebra homomorphism, this implies

$$\pi(s_0 + s_1 x + \dots + s_{n-1}x^{n-1}) = \pi(0)$$

thus

$$s_0 + s_1 x + \dots + s_{n-1}x^{n-1} \sim 0 \text{ in } K[x]$$

i.e.

$$q(x) = s_0 + s_1 x + \dots + s_{n-1}x^{n-1} = p(x) \cdot m(x)$$

for some $p(x) \in K[x]$. Since $\deg q(x) < \deg m(x)$ this is possible only if $p(x) = q(x) = 0$ and so $s_i =$ for all i. This proves independence. The claim about the matrix is then obvious from $(*)$.

Conversely, given a K-vector space and an endomorphism ϕ having matrix $A = \phi^\alpha$ for some basis α, the basis looks as indicated (as one reads from the matrix). Thus, the $K[x]$-module V is generated by v_0. Moreover

$$\phi^n(v_0) = \phi(\phi^{n-1}(v_0)) = -(r_{n-1}\phi^{n-1}(v_0) + \dots r_1\phi(v_0) + r_0 v_0) = -(r_{n-1}\phi^{n-1} + \dots r_1\phi + r_0)(v_0)$$

i.e.

$$m(\phi)(v_0) = 0$$

Thus, the $K[x]$ module V with generator v_0 satisfies the relation $m(x)v_0 \stackrel{!}{=} 0$. By Principle 4.2.3 there is a homomorphism $\psi : K[x]/(m(x))$ into the $K[x]$-module V mapping 1 onto v_0. ψ is surjective, since v_0 is a generator. ϕ is also a K-linear map whence an isomorphism since $\dim K[x]/(m(x)) = n = \dim V$. $\square$

Corollary 4.2.7 *$m(x)$ is a polynomial $p(x)$ of minimal degree such $p(\phi) = 0$ and the unique normed such. $(-1)^n m(x)$ is the characteristic polynomial of ϕ.*

$m(x)$ is also called the *minimal polynomial* of ϕ. *Proof.* The first claim is obvious from the proof of the theorem, the second an exercise. $\square$

Corollary 4.2.8 *Here, for any $\lambda \in K$,*

$$\beta : (\phi - \lambda id)^{n-1}(v_0), \dots, (\phi - \lambda id)(v_0), v_0$$

is also a basis of V and $m(x) = (x-\lambda)^n$ if and only if

$$\phi^\beta = J_{\lambda,n} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & 0 & \lambda & 1 \\ 0 & & \dots & 0 & \lambda \end{pmatrix}.$$

The matrix $J_{\lambda,n}$ is a *Jordan-block* and the basis β a *Jordan-chain* (for ϕ and λ) with *startvector* v_0 and *eigenvector* $\phi - \lambda \mathsf{id}^{n-1}$.

Proof. The $(x-\lambda)^k$ generate the K-vector space $K[x]$: inductively on obtains all x^k since $(x-\lambda)^k = x^k + p_k(x)$ with a polynomial $p_k(x)$ of degree $< k$. Thus, the $(x-\lambda)^k$, $k < n$, generate the K-vector space $K[x]/(m(x))$ and form a basis $\beta : (x-\lambda)^{n-1}, \dots, 1$ (since $\dim = n$). For $m(x) = (x-\lambda)^n$ the matrix of ϕ w.r.t. β is $J_{\lambda,n}$ as is seen from

$$x(x-\alpha)^k = (x-\lambda)^k(\lambda + x - \lambda) = \lambda(x-\lambda)^k + (x-\lambda)^{k+1}. \quad \square$$

4.2.5 Presentation matrix

Consider a presentation of an R-module with generators $e_1, \dots, e_n$ and relations $w_i \stackrel{!}{=} 0$ $(i \in I)$. Since module laws allow reduction of any term to a linear combination $\sum_i r_i e_i$,we may assume that the w_i are of this form. Thus

1 Any presentation of an R-module with generating set $E = \{e_1, \ldots, e_n\}$ may be equivalently given by

$$\sum_{i=1}^{n} r_{ij} e_i,\ j = 1, 2, 3, \ldots, \quad \mathcal{A} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \ldots \\ \vdots & & & \\ r_{n1} & r_{n2} & r_{n3} \ldots & \end{pmatrix} \in R^{n \times m}$$

2 The matrix $\mathcal{A}$ is called the *presentation matrix*

3 The module M is obtained as M R^n/U, U generated by the columns of $\mathcal{A}$

4 For $n = 1$ and $\mathcal{A} = (d_1)$, one has $M \cong R/Rd$ where $Rd = \{rd \mid r \in R\}$

5 If the presentation matrix is diagonal with entries $d_1, \ldots, d_n$, then M is isomorphic to

$$R/Rd_1 \times \ldots \times R/Rd_n$$

Ad 5: Let $U = \mathsf{Span}\{e_1 d_1, \ldots, e_n d_n\} \subseteq R^n$ and $\pi : R^n \to R^n/U$ and $\pi_i : R \to R/Rd_i$ canconical projections and

$$\psi : R^n \to R/Rd_1 \times \ldots \times R/Rd_n \text{ where } \psi \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} r_1[\ \bmod Rd_1] \\ \vdots \\ r_n[\ \bmod Rd_n] \end{pmatrix}$$

Then ψ is a surjective homomorphism. Moreover

$$\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in \mathsf{Kern}(\pi) \Leftrightarrow \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \sum d_i s_i e_i = \begin{pmatrix} d_1 s_1 \\ \vdots \\ d_n s_n \end{pmatrix} \Leftrightarrow\Leftrightarrow r_1 \in Rd_1, \ldots, r_n \in Rd_n \Leftrightarrow \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in \mathsf{Ker}\psi$$

Hence, by the Homomorphism Theorem, there is an isomorphism $\chi : R^n/U \to R/Rd_1 \times \ldots \times R/Rd_n$. $\square$

4.2.6 Characteristic matrix of an endomorphism

Theorem 4.2.9 *Given a K-vector space V with basis $\alpha : \vec{e}_1, \ldots, \vec{e}_n$ and endomorphism ϕ with matrix A w.r.t. α. Then w.r.t. the generators $\vec{e}_1, \ldots, \vec{e}_n$ of the $K[x]$-module ${}_{K[\phi]}V$*

$$\boxed{\textit{the } \text{characteristic matrix } A - xE \textit{ of } \phi \textit{ is a presentation matrix of } {}_{K[\phi]}V}.$$

Proof. The $\vec{e}_i$ satisfy the relations given by $A - xE$:

$$x\vec{e}_j = \phi(\vec{e}_j) = \sum_i a_{ij}\vec{e}_i \quad \Leftrightarrow \quad 0 = (a_{jj} - x)\vec{e}_j + \sum_{i \neq j} a_{ij}\vec{e}_i$$

Hence there is a surjective $K[x]$-linear map χ from the $K[x]$-module M with generators e_i and presentation matrix $A - xE$ onto ${}_{K[\phi]}V$ with $e_i \mapsto \vec{e}_i$. As a $K[x]$-module, M is

generated by the e_i. M is also a K-vector space. The K-vector subspace U generated by the $\pi(e_i)$ is a $K[x]$-submodule, since

$$xe_j = \sum_i a_{ij} e_i$$

due to the presentation. Thus, $U = M$ whence $\dim_K M \leq n$. χ is also K-linear and surjective, whence due to $\dim_K V = n$ an isomorphism. $\square$

4.3 Transformations of presentations

We will show that for rings $K[x]$ any presentation can be equivalently replaced by one given by a diagonal matrix in $K[x]^{n\times n}$. For that purpose we need two kinds of transformations of the presentation matrix

- Replacing relations by equivalent ones
- Change of basis

A matrix $\mathcal{S} \in R^{n\times n}$ is invertible if and only if there is $\mathcal{T} \in R^{n\times n}$ such that $\mathcal{ST} = \mathcal{TS} = \mathcal{E}$ the unit matrix. The invertible matrices form a subgroup of the multiplicative monoid $(R^{n\times n}, \cdot, \mathcal{E})$. In particular, the inverse is uniquely determined: $\mathcal{T} = \mathcal{S}^{-1}$.

4.3.1 Change of relations

Given a commutative ring R and a basis $\alpha : e_1, \ldots, e_n$ of a free R-module F, each $v \in F$ has unique representation

$$v = \sum_{i=1}^{n} x_i e_i$$

and we have the *coordinate column* of v

$$v^\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Now, given a matrix $\mathcal{A} = (a_{ij}) \in R^{n\times m}$, we write

$$\mathsf{Span}_R^\alpha(\mathcal{A}) = \mathsf{Span}_R\{\sum_{i=1}^{n} a_{ij} e_i \mid j = 1, \ldots, m\}$$

which is the span of those elements of F which have colums of $\mathcal{A}$ as coordinates.

Lemma 4.3.1 *Given a commutative ring R, a matrix $\mathcal{A} \in R^{n\times m}$, and $\mathcal{Q} \in R^{m\times m}$. Then for each basis α of a free R-module*

$$\mathsf{Span}_R^\alpha(\mathcal{A}) = \mathsf{Span}_R^\alpha(\mathcal{AQ})$$

Proof. Let $\mathcal{Q} = (q_{jk})$ and $\mathcal{AQ} = \mathcal{B} = (b_{ik})$. Then for all $k = 1, \ldots, m$

$$\sum_{i=1}^{n} b_{ik} e_i = \sum_{i=1}^{n} (\sum_{j=1}^{m} a_{ij} q_{jk}) e_i = \sum_{j=1}^{m} q_{jk} \sum_{i=1}^{n} a_{ij} e_i \in \mathsf{Span}_R^{\alpha}(\mathcal{A})$$

whence

$$\mathsf{Span}_R^{\alpha}(\mathcal{B}) \subseteq \mathsf{Span}_R^{\alpha}(\mathcal{A})$$

Since $\mathcal{A} = \mathcal{B}\mathcal{Q}^{-1}$ the converse inclusion is also valid. □.

4.3.2 Change of basis

Lemma 4.3.2 *Given a commutative ring R and a basis $\alpha : e_1, \ldots, e_n$ of a free R-module F, and invertible matrix $\mathcal{P} \in R^{n \times n}$, there is a unique basis $\beta : f_1, \ldots, f_n$ of F such that*

$$v^{\beta} = \mathcal{P} v^{\alpha} \quad \textit{for all } v \in F$$

Proof. Choose f_j such that f_j^{α} is the j-th column of $\mathcal{S} = \mathcal{P}^{-1}$, i.e.

$$f_j = \sum_{i=1}^{n} s_{ij} e_i$$

Then $f_1, \ldots, f_n$ is generating since

$$\sum_{j=1}^{n} p_{jk} f_j = \sum_{j=1}^{n} p_{jk} \sum_{i=1}^{n} s_{ij} e_i = \sum_{i=1}^{n} (\sum_{j=1}^{n} s_{ij} p_{jk}) e_i = e_k$$

They are independent, too: $\sum_{j=1}^{n} r_j f_j = 0$ implies

$$\sum_{i=1}^{n} (\sum_{j=1}^{n} r_j s_{ij}) e_i = \sum_{j=1}^{n} r_j (\sum_{i=1}^{n} s_{ij} e_i) = \sum_{j=1}^{n} r_j f_j = 0$$

whence by independence of the $e_1 \ldots, e_n$

$$\sum_{j=1}^{n} r_j s_{ij} = 0 \quad \text{for all } j = 1, \ldots, n$$

$$\mathcal{S} \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \mathcal{P}\mathcal{S} \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \mathbf{0}$$

Finally, we have $v^{\alpha} = \mathcal{S} v^{\beta}$ since for $v = \sum_{j=1}^{n} y_j f_j$ it follows

$$v = \sum_{j=1}^{n} y_j f_j = \sum_{j=1}^{n} y_j \sum_{i=1}^{n} s_{ij} e_i = \sum_{i=1}^{n} (\sum_{j=1}^{n} s_{ij} y_j) e_i \quad \square$$

Over non-commutative rings R this remains valid if we consider right R-modules, i.e. write vr and have the law $v(rs) = (vr)s$. The point is, that R^n is a left-$R^{n \times n}$ right-R bi-module: we have $\mathcal{A}(\boldsymbol{v}r) = (\mathcal{A}\boldsymbol{v})r$.

4.3.3 Transformation of presentations

Corollary 4.3.3 *If $\mathcal{A} \in R^{n\times M}$ is a presentation matrix of the R-module M w.r.t. the basis $\alpha : e_1, \ldots, e_n$ of the free R-module F, and if $\mathcal{P} \in R^{n\times n}$ and $\mathcal{Q} \in R^{m\times m}$ are invertible then $\mathcal{PAQ}$ is a presentation matrix for M w.r.t. the basis $\beta : f_1, \ldots, f_n$ of F the coordinates of which w.r.t. α are the columns of $\mathcal{P}^{-1}$.*

Proof. $M = F/U$ where

$$U = \mathsf{Span}_R^\alpha(\mathcal{A}) = \mathsf{Span}_R\{v \in F \mid v^\alpha \text{ colum of } \mathcal{A}\} =$$

$$= \mathsf{Span}_R\{v \in F \mid v^\beta \text{ colum of } \mathcal{PA}\} = \mathsf{Span}_R^\beta(\mathcal{PA}) = \mathsf{Span}_R^\beta(\mathcal{PAQ}) \quad \square$$

Lemma 4.3.4 *If $\mathcal{A} \in R^{n\times M}$ is a presentation matrix of the R-module M w.r.t. the basis $\alpha : e_1, \ldots, e_n$ of the free module F and if $\mathcal{B}$ arises from $\mathcal{A}$ by deleting zero colums, then $\mathcal{B}$ is a presentation matrix of M w.r.t. α.*

Proof. Obvious. $\square$

4.3.4 Elementary matrices

Given a commutative ring R, let $\mathcal{E}_{ij}$ the matrix with all entires 0 but 1 in position (i, j). The following matrices in $R^{n\times n}$ *elementary*

$$[Zi := Zi + rZj] = [Sj := Sj + rSi] = \mathcal{E} + r\mathcal{E}_{ij} \;\; r \in R$$

$$[Zi \leftrightarrow Zj] = Sj \leftrightarrow Si] = \mathcal{E} - \mathcal{E}_{ii} - \mathcal{E}_{jj} + \mathcal{E}_{ij} + \mathcal{E}_j i \;\; i \neq j$$

$$[Zi := uZ_i] = [S_i := uS_i] = \mathcal{E} + (u-1)\mathcal{E}_{ii} \;\; u \in R \text{ invertible}$$

The notation corresponds to the row transformations $[Z]$ of a matrix $\mathcal{A}$ induced by multiplying the elementary matrix on the left of $\mathcal{A}$ resp. column transformations $[S]$ on the right.

Lemma 4.3.5 *Elementary matrices are invertible with inverses*

$$[Zi := Zi + rZj]^{-1} = [Sj := Sj - rSi]$$

$$[Zi \leftrightarrow Zj]^{-1} = [Sj \leftrightarrow Si], \quad [Zi := uZ_i]^{-1} = [S_i := u^{-1}S_i]$$

Proof. Obvious. $\square$

Chapter 5

Euclidean rings

Subsections * are not needed for the for the main result: the theory of invariant divisors and rational canonical form.

5.1 Ideals

5.1.1 Ideals and congruences of rings

Given a commutative rings resp. K-algebra R, an *ideal* is submodule of the R-module R, i.e.

$$a,\, b \in I \Rightarrow a+b \in I \text{ and } a \in I \Rightarrow ra \in I \quad \text{for all } a,b,r \in R$$

There is a 1-1-correspondence between congruence relations and ideals given by

$$I = \{a \in R \mid a \sim 0\} \quad a \sim b \Leftrightarrow a-b \in I$$

Indeed, an equivalence relation $\sim$ is a congruence relation of the ring R if and only if it is a congruence relation of the R-module R: in both cases one has a congruence relation of the additive group (K-vector space) R satisfying

$$a \sim b \;\Rightarrow\; ra \sim rb$$

Consequently, a factor algebra $R/\sim$ may be written as R/I and the canonical projection as $\pi(a) = I + a = a + I$. The homomorphism theorems apply as well.

Form the description of spans in modules we obtain

$$(a) = Ra = \{ra \mid r \in\}$$

is an ideal, the *principal ideal* generated by a. The smallest ideal containing a, b is

$$\{ra + sb \mid r, s \in R\} = (a,b) = (a) + (b)$$

Corollary 5.1.1 ** For ideals I, J of a commutative ring or K-algebra, $R/I \cong R/J$ as R-modules if and only if $I = J$.*

Proof. Let $M = R/I$. $I = \{r \in R \mid rv = 0 \text{ for all } v \in M\}$ is the *annihilator* of M and invariant under linear isomorphisms. □.

5.1.2 Second Isomorphism Theorem *

Theorem 5.1.2 *Let $\pi : M \to N$ a surjective homomorphism between R-modules. Then for each submodule U of N, $\pi^{-1}(U)$ is a submodule of M and*

$$M/\pi^{-1}(U) \cong N/U$$

This estabilishes a 1-1- correspondence between submodules of N and submodules $\supseteq \ker \pi$ of M with inverse given by $V \mapsto \pi(V)$. Morever

$$U \subseteq W \Leftrightarrow \pi^{-1}(U) \subseteq \pi^{-1}(W)$$

$$\pi^{-1}(U+W) = \pi^{-1}(U) + \pi^{-1}(W), \;\; \pi^{-1}(U \cap W) = \pi^{-1}(U) \cap \pi^{-1}(W)$$

The analogous results hold for commutative rings and K-algebras and, formulated in terms of congruence relations for any algebraic structures.

Proof. $\pi^{-1}(U)$ is a submodule, obviously, and the kernel of $\pi_U \circ \pi$ where $\pi_U : N \to N/U$ is the canonical projection. Thus $M/\pi^{-1}(U) \cong N/U$.

By surjectivity, we have $\pi(\pi^{-1}(U)) = U$. If $V \supseteq \ker \pi$ and $w \in \pi^{-1}(\pi(V))$ then $\pi(w) = \pi(v)$ for some $v \in V$ whence $w - v \in V$ and $w \in V$. Since both maps $U \mapsto \pi^{-1}(U)$ and $V \mapsto \pi(V)$ preserve inclusion between submodules and since $+$ and $\cap$ may be characterized in these terms, the remaining claims follow. $\square$.

Example: Consider $\pi : \mathbb{Z} \to \mathbb{Z}/\mathbb{Z}n$ the canonical homomorphism. The submodules resp. ideal of $\mathbb{Z}/\mathbb{Z}n$ are given as $\mathbb{Z}\pi(m)$ where m divides n. For the canonical homomorphisms $\chi : \mathbb{Z}/\mathbb{Z}n \to (\mathbb{Z}/Zn)/\mathbb{Z}\pi(m)$ and $\psi : \mathbb{Z} \to \mathbb{Z}/\mathbb{Z}m$ we have $\psi = \chi \circ \pi$.

5.2 Integral domains

5.2.1 Definition and examples

An *integral domain* is commutative ring without *divisors of zero*, i.e. $ab = 0$ implies that $a = 0$ or $b = 0$. Equivalently, one has the *cancellation law*

- From $ax = ay$ and $a \neq 0$ it follows $x = y$

Examples. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

If K is an integral domain, the degree formula of 14.2.2. remains valid and it follows that $K[x]$ is an integral domain, too. Long division of $p(x)$ by $q(x)$ is possible if the leading coefficient b_m in $q(x)$ has an inverse in K. In the results about zeros in 14.2.4. it suffices to have K a subfield of the integral domain A.

5.2.2 Horner scheme *

Lemma 5.2.1 *Given an itegral domain K and $p(x) \in K[x]$ of degree n und $\alpha \in K$ there is $h(x) \in K[x]$ such that*

$$p(x) = h(x)(x - \alpha) + p[\alpha]$$

Proof. The idea is that

$$p(x) = (\dots((a_n x + a_{n-1})x + a_{n-2})\dots + a_1)x + a_0$$

and one may obtain $p[\alpha]$ with less multiplications as follows

a_n	a_{n-1}	$\dots$	a_1	a_0
α	$c_{n-1}\alpha$	$\dots$	$c_1\alpha$	$c_0\alpha$
$c_{n-1} = a_n$	$c_{n-2} = c_{n-1}\alpha + a_{n-1}$	$\dots$	$c_0 = c_1\alpha + a_1$	$p[\alpha] = c_0\alpha + a_0$

Now, put

$$h(x) = c_{n-1}x^{n-1} + \dots + c_1 x + c_0.$$

For verification consider

$$q(x) = (\dots(a_n x + a_{n-1})x + a_{n-2})\dots + a_1 \text{ i.e. } p(x) = q(x)x + a_0.$$

Computing $q[\alpha]$ one obtains the Horner-coefficients $c_{n-1}, \dots, c_1$ and it holds $q[\alpha] = c_0$ and $c_0\alpha + a_0 = p(\alpha)$. Applying inductive hypothesis to $q(x)$ one gets

$$\begin{aligned} p(x) = q(x)x + a_0 &= ((c_{n-1}x^{n-2} + \dots + a_1)(x - \alpha) + q[\alpha])x + a_0 \\ &= (c_{n-1}x^{n-1} + \dots c_1 x + c_0)(x - \alpha) - c_0 x + c_0\alpha + q[\alpha]x + a_0 \end{aligned}$$

and the last 4 summands add up to $p[\alpha]$.

5.2.3 Quotient fields *

A field Q is a *quotient field* of the ring R, if R is a subring of Q and

$$Q = \{ab^{-1} \mid a, b \in R,\ b \neq 0\}$$

Necessarily, R is an integral domain. Example: $\mathbb{Q}$ is quotient field of $\mathbb{Z}$.

Theorem 5.2.2 *Each integral domain R admits an extension to a quotient field Q (unique up to isomorphism). Any embedding of R into a field K can be extended to an embbing of Q into K.*

Proof. As in the construction of $\mathbb{Q}$ from $\mathbb{Z}$ define on $Q' = \{(a, b) \mid a, b \in R,\ b \neq 0\}$

$$\begin{aligned} &(a, b) + (c, d) = (ad + bc, bd), \quad (a, b) \cdot (c, d) = (ac, bd) \\ &(a, b) \sim (c, d) \ \Leftrightarrow ad = bc \end{aligned}$$

which is a congruence relation. By factorizing, $\pi : Q' \to Q'/\sim$, one obtains an algebraic structure and even a commutative ring Q - $(Q, \cdot, 1)$ is a commutative monoid being a homomorphic image of $R \times (R \setminus \{0\})$, the other laws require some computation. The inverse of $\pi(a, b)$ is $\pi(b, a)$. The map $a \mapsto \pi(a, 1)$ is an embedding og R into Q and $\pi(a, b) = \pi(a, 1)\pi(1, b)$. Thus we may conceive R as a subring of Q and obtain the required representation. Given $\phi : R \to K$ define

$$\overline{\phi}(ab^{-1}) = \phi(a)\phi(b)^{-1}$$

This is well defined since $ad = bc$ implies $\phi(a)\phi(d) = \phi(b)\phi(c)$, also $\phi(a)\phi(b)^{-1} = \phi(c)\phi(d)^{-1}$. □

The quotient field of the polynomial ring $K[x]$ over a field is the field of *rational functions* over K and denoted by $K(x)$. Its elements are written as $\frac{f(x)}{g(x)}$

5.2.4 Units

A *Monoid* is a set M with associative multiplication and neutral element e. An element u of a monoid M is a *unit* or *invertible*, if there is $x \in M$ such that $xu = ux = 1$.

Lemma 5.2.3 *The units of a monoid M from a group M^*.*

Proof. The inverse x is unique since $yu = uy = 1$ implies $y = y1 = yux = 1x = x$. We may write $x = u^{-1}$; clearly $u^{-1} \in M^*$ and $(u^{-1})^{-1} = u$. If $v \in M^*$, then $uvv^{-1}u^{-1} = u1u^{-1} = uu^{-1} = 1$ and $v^{-1}u^{-1} = uv$, whenve $uv \in M^*$. □

The *group of units* R^* of a ring R consists of the units of the monoid $(R, \cdot, 1)$. Clearly

$$\mathbb{Z}^* = \{1, -1\}, \;\; (K[x])^* = K^* \text{ for fields } K$$

and for the direct product $R_1 \times R_2$ of rings (component wise addition and multiplication)

$$(R_1 \times R_2)^* = R_1^* \times R_2^*.$$

In the ring of $n \times n$-matrices over a commutative ring R one defines determinants by the explicte formula. Then $(\det A)\mathsf{ad}A = E$. If $\det A \in R^*$ then $A^{-1} = \det(A)^{-1}\mathsf{ad}A$ whence

$$(R^{n\times n})^* = \{A \mid \det A \in R^*\} \text{ for commutative } R.$$

Corollary 5.2.4 *For any commutative ring, $R = Ru$ if and only if u is a unit. Then, $M = Rug$ for any cyclic R-module M with generator g.*

5.2.5 Divisibility

In a commutative ring one defines

$$d|a \;\Leftrightarrow d \text{ divides } a \;\Leftrightarrow \exists r \in R.\, rd = a.$$

This ia a 'quasi order' on R

$$a|a \text{ (reflexive)}, \; a|b \text{ and } b|c \Rightarrow a|c \text{ (transitive)}$$

with compatibility

$$a|b \Rightarrow ac|bc, \; a|b \text{ und } a|c \Rightarrow a|(b \pm c)$$

One has

$$a \mid b \;\Leftrightarrow\; Ra \supseteq Rb$$

Lemma 5.2.5 *In a commutative ring, $a \mid b$ if and only if there is a surjective R-linear map $\chi : R/(a) \to R/(b)$.*

Proof. This is immediate by the Homomorphism Theorem. □

5.2.6 Associated elements

Let R be an integral domain. a and b are *associated*, $a \approx b$, iff one of the following conditions is satisfied

$$a|b \text{ and } b|a\ ,\ \ \exists r \in R^* : ra = b.$$

Indeed $ra = b$ and $sb = a$ imply $rsb = b$ whence $rs = 1$ by cancellation. $\approx$ is an equivalence relation since units form a subgroup. Moreover

- $a \approx a'$ und $b \approx b' \ \Rightarrow\ (a|b \Leftrightarrow a'|b')$
- $a \sim b$ w.r.t. the congruence associated with $(d) \Leftrightarrow a \equiv \mod d \ \Leftrightarrow\ d|(a-b)$
- $a|b \ \Leftrightarrow\ (b) \subseteq (a)$
- $a \approx b \Leftrightarrow\ (a) = (b)$
- $a \in R^* \ \Leftrightarrow\ (a) = (1) = R$

5.3 Principal ideals in euclidean rings

5.3.1 Definition and examples

An integral domain R is an *euclidean ring* if there is a map

$$\delta : R \setminus \{0\} \to \mathbb{N}$$

$$\begin{array}{ll} \forall a, b \in R \setminus \{0\} : & \delta(ab) \geq \delta(a) \\ \forall a, b \in R \setminus \{0\}\ \exists q, r \in R : & a = bq + r \qquad \text{and } \delta(r) < \delta(b) \text{ or } r = 0 \end{array}.$$

Define $\delta(0) = -\infty$. Examples

$$\mathbb{Z} \text{ with } \delta(a) = |a|, \quad K[x] \text{ with } \delta(f(x)) = \deg f(x),\ K \text{ a field.}$$

Lemma 5.3.1 *In any euclidean ring*

$$a|b \text{ und } \delta(a) = \delta(b) \ \Leftrightarrow a \approx b.$$

Proof. Let $b = ac$. Then $a = bq + r$ with $r = 0$ or $r = a - bq = a - acq = a(1 - cq)$ and $\delta(b) > \delta(r) = \delta(a(1 - cq)) \geq \delta(a)$. If $\delta(a) = \delta(b)$, then the latter may not occur whence also $b \mid a$, and $a \approx b$. □

5.3.2 Principal ideals

An integral domain in which every ideal is principal is a *principal ideal domain.*

Theorem 5.3.2 *Any euclidean ring is a principal ideal domain:* $0 \neq a \in I$ *with* $\delta(a)$ *minimal then* $I = (a)$.

Proof. Assume $I \neq (0)$. Choose $0 \neq a \in I$ with $\delta(a)$ minimal. Then $(a) \subseteq I$. Consider $0 \neq b \in I$. Then $b = aq + r$ mit $r = 0$ or $\delta(r) < \delta(a)$. In the second case $r = b - aq \in I$ contradicting minimality of $\delta(a)$. Thus $r = 0$ and $b \in (a)$. □

5.3.3 Cyclic modules

Recall that an R-module M is *cyclic* if it is generated by a single element g: $M = Rg$.

Proposition 5.3.3 *Any cyclic R-module $M = Rg$ has presentation given by generator g and relation $d \stackrel{!}{=} 0$ where $0 \neq d \in R$ with $\delta(d)$ minimal such that $dg = 0$ in M. Moreover, $dv = 0$ for all $v \in M$, there is a surjective R-linear map $\pi : R \to M$ such that $\pi(1) = g$, and M is a commutative ring (and K-algebra if so is R) and π a homomorphism w.r.t. the multiplication*

$$\pi(r) \cdot \pi(s) := \pi(rs)$$

The generators of M are exactly the units of this ring.

d is unique up to association and called a *minimal annihilator* of g resp. M.

Proof. $I = \{r \in R \mid rg = 0\}$ is an ideal of R: if $r, s \in I$ then $(r+s)g = rg + sg = 0$ and if $r \in R$ and $s \in R$ then $(sr)g = s(rg) = s0 = 0$. Since R is euclidean, we have $I = (d)$ with d as stated and unique up to association. Then $dv = drg = rdg = 0$ for all $v = rg \in M$. Since R is freely generated by 1, there is a unique homomorphism $\pi : R \to M$ such that $\pi(1) = g$. Then $I = (d) = \mathsf{Ker}\,\pi$ and it follows that M is presented by g, $d \stackrel{!}{=} 0$. Moreover, since I is an ideal, $M = R/I$ can be seen as the factor ring with the above multiplication. Now, $\pi(u)$ is a unit iff $\pi(s)\pi(u) = \pi(1) = g$ for some s iff $\pi(u)$ is a generator. $\square$.

5.4 Euclidean algorithm, GCD, and factorization

5.4.1 Bezout's Theorem

Consider an integral domain R. d is a *greatest common divisor* of a and b

$$d \approx GCD(a,b) \;\Leftrightarrow\; d|a, d|b \text{ and } \forall c : (c|a \text{ and } c|b) \Rightarrow c|d.$$

If such exists, it is unique up to association. Moreover

$$\mathrm{GCD}(a,b) \approx \mathrm{GCD}(b, a - qb)$$

since a, b and $a - qb$ have the same divisors.

Theorem 5.4.1 *In an euclidean ring GCDs exist and have* additive representation

$$d \approx GCD(a,b) \;\Leftrightarrow\; d|a, d|b \text{ and } \exists r, s : d = ra + sb \;\Leftrightarrow\; (d) = (a) + (b).$$

Proof. The extended euclidean algorithm produces d, r, s such that $d|a$, $d|b$, $d = ra + sb$. Then d is a GCD: $c|a$ and $c|b$ imply $c|ra$ and $c|sb$ whence $c|(ra + sb)$. Conversely, if d' is a GCD of a, b then $d' \approx d$ by uniqueness and $(d') = (a) + (b)$. $\square$

Algorithm 5.4.2 *(*Euklid+Bezout*). Given an euclidean ring R and a, b in R determine a $d \approx GCD(a,b)$ and x, y in R such that*

$$d = ax + by.$$

- *Put* $d_1 := a$, $x_1 := 1$, $y_1 := 0$; $d_2 := b$, $x_2 := 0$, $y_2 := 1$
- *Loop:* $n \rightsquigarrow n+1$
  - *Find* $d_n = d_{n-1}q + r$ *with* $0 \leq \delta r < \delta d_{n-1}$ *or* $r = 0$
  - *If* $r \neq 0$ *do* $d_{n+1} = r = d_n - qd_{n-1}$, $x_{n+1} = x_n - qx_{n-1}$ $y_{n+1} = y_n = qy_{n-1}$
  - *else* $d = r$, $x = x_n$, $y = y_n$ *stop*

Proof. Loop invariant: $d_n = ax_n + bx_n$, $GCD(d_n, d_{n-1}) = GCD(a, b)$ If $r = 0$, then $d_n | d_{n-1}$, whence $d = GCD(a, b)$. □

42	1	0
25	0	1
17	1	-1
8	-1	2
1	3	-5

$$1 = 3 \cdot 42 - 5 \cdot 25, \quad 25^{-1} \equiv -5 \equiv 20 \bmod 42$$

$x^{10}+1$	1	0	
x^6+1	0	1	
$-x^4+1$	1	$-x^4$	x^4
x^2+1	x^2	$-(x^6-1)$	$-x^2$
0	0	0	x^2+1

$$GCD(x^{10}+1,\, x^6+1) = x^2+1 = x^2(x^{10}+1) - (x^6-1)(x^6+1)$$

a and b are *relatively prime* or *coprime* if $GCD(a, b) \approx 1$ i.e. iff $ra + sb = 1$ for some r, s

Corollary 5.4.3 $a|(bc) \wedge GCD(a, b) = 1 \Rightarrow a|c$

Proof . $1 = ax + by$, whence $a|(axc + bcy) = c$.

Corollary 5.4.4 *If* $GCD(a, b) = 1$, *then* $\tilde{b} = b[\bmod a]$ *invertible in* $R/(a)$

$$by \equiv 1 \mod a \quad \textit{if } 1 = ax + by \textit{ for some } x$$

$$x^2 + 1 = (x+2)(x-2) + 5, \quad 1 = \frac{1}{5}(x^2 + 1 - (x+2)(x-2))$$

$$(x+2)^{-1} \equiv -\frac{1}{5}(x-2) \mod x^2+1$$

$$x^3 + x \equiv xx^2 + x \equiv x(x+1) + x \equiv x^2 \equiv x+1 \mod x^2+x+1$$

$$1 = x^2 + x + 1 - x(x+1), \quad (x+1)^{-1} \equiv x \mod x^2+x+1$$

5.4.2 Primes

Theorem 5.4.5 *For $0 \neq a \notin R^*$ in an euclidean ring t.f.a.e.*

- *a is* irreducible*, i.e. $a = bc \Rightarrow b \in R^*$ or $c \in R^*$*
- *(a) is a* maximal ideal*, i.e. for all ideals: $a \in I \Rightarrow (a) = I$ or $I = R$.*
- *$R/(a)$ is a field*
- *a is* prime*, i.e. $a|bc \Rightarrow a|b$ or $a|c$*

Proof. (1) $\Rightarrow$ (2): $I = (b)$ and $a = bc$, whence $I = R$ if $b \in R^*$ resp. $(a) = I$ if $c \in R^*$.
(2) $\Rightarrow$ (3): Let $\tilde{b} \neq 0$ in $R/(a)$, whence $b \notin (a)$. Thus $(a) \neq (a, b)$ and it follows $1 \in R = (a, b)$, i.e. there are r, c with $ar + bc = 1$. Thus $\tilde{b} \cdot \tilde{c} = \tilde{1}$, i.e. $\tilde{b}$ is invertible
(3) $\Rightarrow$ (4): $a|bc$ implies $\tilde{b} \cdot \tilde{c} = \widetilde{bc} = 0$ thus $\tilde{b} = 0$ or $\tilde{c} = 0$, i.e. $b \in (a)$ or $c \in (a)$.
(4) $\Rightarrow$ (1): If $a = bc$ is prime, then $a|b$ or $a|c$. On the other hand $c|a$ and $b|a$, whence $b \in R^*$ or $c \in R^*$. $\square$

Corollary 5.4.6 *Given a, b in an euclidean ring R, $b[\bmod a]$ is invertible in $R/(a)$ iff $GCD(a, b) \approx 1$.*

5.4.3 Factorization

An integral domain is *factorial* or an *UFD* if any non-unit $a \neq 0$ is a product

$$a = p_1 \cdot \ldots \cdot p_n$$

of primes, unique up to order and association.

Theorem 5.4.7 *Any euclidean ring is factorial.*

Proof of existence by order induction on $\delta(a)$: If a is not irreducible then $a = bc$ with $\delta(a) > \delta(b), \delta(c)$ and by induction $b = \prod_i p_i$ and $c = \prod_j q_j$ whence $a = \prod_i p_i \cdot \prod_j q_j$ with irreducible p_i and q_j.
Proof of uniqueness by induction on the number of factors. Let

$$p_1 \cdot \ldots \cdot p_n \approx q_1 \cdot \ldots \cdot q_m.$$

W.l.o.g. $p_1|q_1$, i.e. $q_1 \approx p_1$ since both are prime. It follows

$$p_2 \cdot \ldots \cdot p_n \approx q_2 \cdot \ldots \cdot q_m$$

and, by induction, $n = m$ and $p_i \approx q_i$ after renumbering. $\square$

5.4.4 Factorization algorithms *

At present, there are is a fast algorithms for testing primeness, but none for factorization of integers. For factorization of polynomials over finite fields there is a simple and efficient (GCD with test polynomials) cf. Berlekamp, Algebraic coding theory. For polynomials over $\mathbb{Q}$ there is a efficient but non-trivial procedure: Lenstra[2],Lovasz, Math.Ann 261, cf. Lenstra[2]: Algorithms in number theory, Handbook of Theoretical Computer Science A.

According to a theorem of Gauss, a polynomial in $\mathbb{Z}[x]$ is irreducible if and only if it is irreducible in $\mathbb{Q}[x]$ and the GCD of it coefficients is 1. This can be used for a brute force factorization method due to Kornecker: Given $p(x) \in \mathbb{Z}[x]$ of degree n, choose $z_0, \ldots, z_m \in \mathbb{Z}$ where $m = \frac{n}{2}$ and determine for each k the set D_k of divisors of $p(z_k)$. Thus if $f(x) \in \mathbb{Z}[x]$ divides $p(x)$ then $f(z_k) \in D_k$. By interpolation, construct all polynomials $f(x)$ of degree $\leq m$ with $f(z_k) \in D_k$ and carry out long division of $p(x)$ by $f(x)$. If $p(x) = q(x)f(x)$ is a proper decomposition continue with both $q(x)$ and $f(x)$ in place of $p(x)$.

5.4.5 LCM

m is an *least common multiple*, $m \approx LCM(a,b)$ if

$$a \mid m,\ b \mid m \ \text{ and if } a|c \text{ and } b|c \text{ implies } m|c \text{ for all } c$$

In an integral domain. LCMs are unique up to association, if they exist, In an euclidean ring they exist and

$$m \approx LCM(a,b) \ \Leftrightarrow\ (m) = (a) \cap (b)$$

Given factorizations

$$a = p_1^{k_1} \cdot \ldots \cdot p_n^{k_n},\ b = p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$$

it follows

$$GCD(a,b) \approx p_1^{\min\{k_1,l_1\}} \cdot \ldots \cdot p_n^{\min\{k_n,l_n\}}$$

$$LCM(a,b) \approx p_1^{\max\{k_1,l_1\}} \cdot \ldots \cdot p_n^{\max\{k_n,l_n\}}, \quad m \approx LCM(a,b) \ \Leftrightarrow\ (a) \cap (b) = (m).$$

$$GCD(a,b) \cdot LCM(a,b) \approx ab$$

5.5 Invariant and elementary divisors

5.5.1 Invariant divisors

Theorem 5.5.1 *For any euclidean ring R and $\mathcal{A} \in R^{n\times m}$ there are products $\mathcal{P} \in \mathbb{R}^{n\times n}$ and $\mathcal{Q} \in \mathbb{R}^{m\times m}$ of elementary matrices such that*

$$\mathcal{P}\mathcal{A}\mathcal{Q} = \mathcal{D} = \begin{pmatrix} d_1 & 0 & 0 & \ldots \\ 0 & d_2 & 0 & \\ \vdots & & \ddots & \\ & & \ldots & \end{pmatrix} \quad \textit{with } d_i|d_{i+1} \textit{ for } i < \min\{m,n\}.$$

The $d_1, d_2, \ldots$ form a system of *invariant divisors* of $\mathcal{A}$. Later we shall show uniqueness up to association. For many purposes it is good enough to have $\mathcal{D}$ a diagonal matrix. Anyway, one should derive such, first.

Proof by the following algorithm. $\mathcal{P}$ is the product $\mathcal{P}_k \cdots \mathcal{P}_1$ of elementary matrices associated with the row transformations used, $\mathcal{Q}$ is the product $\mathcal{Q}_1 \cdots \mathcal{Q}_l$ of of the elementary matrices associated with the column transformations used. □

Algorithm 5.5.2 *dynamic:* $\mathcal{A} \rightsquigarrow \mathcal{A}_{new} =: \mathcal{A}$.

> *a pair* (i, j) *of indices is* active *in* $\mathcal{A}$*, if* $a_{ij} \neq 0$ *and if in the* i*-th row or* j*-th column there is an entry* $\neq 0$

Now, we proceed induction/recursion on

$$\delta(\mathcal{A}) = \begin{cases} \min\{\delta(a_{ij}) \mid (i,j) \text{ active in } \mathcal{A}\} & \text{if non-empty} \\ -\infty & \text{else} \end{cases}$$

to obtain a transformation of $\mathcal{A}$ *to a diagonal matrix:*

- *If* $\delta(\mathcal{A}) \geq 0$ *do*
  - $[Sk := Sk - qSi]$ *with* $\delta(a_{ik} - qa_{ij}) < \delta(a_{ij})$
  - $[Zk := Zk - qZj]$ *with* $\delta(a_{kj} - qa_{ij}) < \delta(a_{ij})$
  - *such that* $\delta(\mathcal{A}_{new}) < \delta(\mathcal{A})$
- *If* $\delta(\mathcal{A}) = -\infty$ *apply permutation to transform* $\mathcal{A}$ *into diagonal form*

Given $d = GCD(a, b) = ra + sb$ *the tranformations* $[S2 := S2 + rS1]$, $[Z1 := Z1 + sZ2]$, $[S1 := S1 - \frac{a}{d}S2]$, $[Z2 := Z2 - \frac{b}{d}Z1]$, $[S1 \leftrightarrow S2]$, $[S2 := (-1)S2]$ *are used to obtain*

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \rightsquigarrow \begin{pmatrix} a & ra \\ 0 & b \end{pmatrix} \rightsquigarrow \begin{pmatrix} a & d \\ 0 & b \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & d \\ -\frac{a}{d}b & b \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 0 & d \\ -\frac{a}{d}b & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} d & 0 \\ 0 & -\frac{a}{d}b \end{pmatrix} \rightsquigarrow \begin{pmatrix} d & 0 \\ 0 & \frac{a}{d}b \end{pmatrix}$$

Given diagonal $\mathcal{D}$ *proceed as follows*

- *If there is* $i < j$ *such that* d_i *does not divide* d_j *choose first* i *mininal and then* j *minimal and apply the above transformations to the minor* $\begin{pmatrix} d_i & 0 \\ 0 & d_j \end{pmatrix}$.

5.5.2 Scheme of computation

For applications, the matrix $\mathcal{P}$ is of no interest, but of interest is

$$\mathcal{P}^1 = \mathcal{P}_1^{-1} \cdots \mathcal{P}_k^{-1}$$

The following scheme of computation can be used

- Start with $\mathcal{E} \mid \mathcal{A} \mid \mathcal{E}$
- Given $\mathcal{L} \mid \mathcal{B} \mid \mathcal{R}$ apply a column transformation $\mathcal{T}$ to $\mathcal{B}$ and $\mathcal{R}$ simultaneously, leave $\mathcal{L}$ unchanged, i.e. produce
$$\mathcal{L} \mid \mathcal{B}\mathcal{T} \mid \mathcal{R}\mathcal{T}$$
- Given $\mathcal{L} \mid \mathcal{B} \mid \mathcal{R}$ apply a row transformation to $\mathcal{B}$ and the inverse column transformation to $\mathcal{L}$, leave $\mathcal{R}$ unchanged, i.e. produce
$$\mathcal{L}\mathcal{T}^{-1} \mid \mathcal{T}\mathcal{B} \mid \mathcal{R}$$
- If $\mathcal{B}$ in $\mathcal{L} \mid \mathcal{B} \mid \mathcal{R}$ is in the required form then $\mathcal{P}^{-1} = \mathcal{L}$ and $\mathcal{Q} = \mathcal{R}$

5.5.3 Example: Presentation of an abelian group

$$\mathcal{A} = \begin{pmatrix} 4 & 0 & 4 \\ 6 & 12 & 16 \\ 0 & 6 & 6 \end{pmatrix}, \quad \mathcal{P} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} 0 & 1 & -6 \\ 1 & 0 & -5 \\ -1 & 0 & 6 \end{pmatrix}$$

It follows that in the free commutative group with generators e_1, e_2, e_3 the subgroup

$$U = \mathsf{Span}_{\mathbb{Z}}\{4e_1 + 6e_2,\ 12e_2 + 6e_3,\ 4e_3 + 16e_2 + 6e_3\}$$

is given w.r.t. the basis

$$f_1 = e_1 + e_2,\ 2e_1 + 3e_2,\ e_3$$

as

$$U = \mathbb{Z}2f_2 \oplus \mathbb{Z}4f_2 \oplus \mathbb{Z}6f_3$$

and that commutative group G with generators e_1, e_2, e_3 and relations

$$4e_1 + 6e_2 \stackrel{!}{=} 0, \quad 12e_2 + 6e_3 \stackrel{!}{=} 0, \quad 4e_3 + 16e_2 + 6e_3 \stackrel{!}{=} 0$$

is isomorphic to

$$G \cong \mathbb{Z}/\mathbb{Z}2 \times \mathbb{Z}/\mathbb{Z}4 \times \mathbb{Z}/Z6$$

Further transformation yields the invariant divisors

$$2,\ 2,\ 12$$

whence

$$G \cong \mathbb{Z}/\mathbb{Z}2 \times \mathbb{Z}/\mathbb{Z}2 \times \mathbb{Z}/Z12$$

	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$\begin{matrix}4&0&4\\6&12&16\\0&6&6\end{matrix}$	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	
	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$\begin{matrix}4&0&0\\6&12&10\\0&6&6\end{matrix}$	$\begin{matrix}1&0&-1\\0&1&0\\0&0&1\end{matrix}$	$[S3 := S3 - S1]$
$[Z2 := Z2 - Z1]$ $[S1 := S1 + S2]$	$\begin{matrix}1&0&0\\1&1&0\\0&0&1\end{matrix}$	$\begin{matrix}4&0&0\\2&12&10\\0&6&6\end{matrix}$	$\begin{matrix}1&0&-1\\0&1&0\\0&0&1\end{matrix}$	
$[Z1 := Z1 - 2Z2]$ $[S2 := S2 + 2S1]$	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}0&-24&-20\\2&12&10\\0&6&6\end{matrix}$	$\begin{matrix}1&0&-1\\0&1&0\\0&0&1\end{matrix}$	
	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}0&-24&-20\\2&0&10\\0&6&6\end{matrix}$	$\begin{matrix}1&-6&-1\\0&1&0\\0&0&1\end{matrix}$	$[S2 := S2 - 6S1]$
	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}0&-24&-20\\2&0&0\\0&6&6\end{matrix}$	$\begin{matrix}1&-6&-6\\0&1&0\\0&0&1\end{matrix}$	$[S3 := S3 - 5S1]$
	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}0&-4&-20\\2&0&0\\0&0&6\end{matrix}$	$\begin{matrix}1&0&-6\\0&1&0\\0&-1&1\end{matrix}$	$[S2 := S2 - S3]$
	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}0&-4&0\\2&0&0\\0&0&6\end{matrix}$	$\begin{matrix}1&0&-6\\0&1&-5\\0&-1&6\end{matrix}$	$[S3 := S3 - 5S2]$
	$\begin{matrix}1&2&0\\1&3&0\\0&0&1\end{matrix}$	$\begin{matrix}2&0&0\\0&-4&0\\0&0&6\end{matrix}$	$\begin{matrix}0&1&-6\\1&0&-5\\-1&0&6\end{matrix}$	$[S1 \leftrightarrow S2]$

5.5.4 Example: Presentation of an endomorphism

For our principal application, the matrix $\mathcal{Q}$ is not needed. Also, transformations in which one row resp. column is used to change others may be carried out simultaneously. Applying row transformations, there is no need to list also the inverse column transformations. Thus, the row and column transformations listed are applied to the right hand matrix and with each row transformation the inverse column transformation has to be applied to the left hand matrix. Example

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -x & -1 & 1 \\ 0 & 1 & 0 & 1 & -2-x & 1 \\ 0 & 0 & 1 & 0 & 0 & -1-x \end{array}\right) = \mathcal{A}$$

$$S1 : +S1 + xS3,\ S2 := S2 + S_3$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & x+1 & -(x+1) & 1 \\ 0 & 0 & 1 & -x(x+1) & -(x+1) & -1-x \end{array}\right)$$

$$S2 := -S2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & x+1 & x+1 & 1 \\ 0 & 0 & 1 & -x(x+1) & x+1 & -1-x \end{array}\right)$$

$$Z2 := Z2 - Z1,\ \ Z3 := Z3 + (1+x)Z1$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & x+1 & x+1 & 0 \\ -x-1 & 0 & 1 & -x(x+1) & x+1 & 0 \end{array}\right)$$

$$S1 := S1 - S2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & x+1 & 0 \\ -x-1 & 0 & 1 & -(x+1)^2 & x+1 & 0 \end{array}\right)$$

$$Z3 := Z3 - Z2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & x+1 & 0 \\ -x-1 & 1 & 1 & -(x+1)^2 & 0 & 0 \end{array}\right)$$

$$S1 \leftrightarrow S3,\ \ S3 := -S3$$

$$\mathcal{P}^{-1} = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & x+1 & 0 \\ -x-1 & 1 & 1 & 0 & 0 & (x+1)^2 \end{array}\right) = \mathcal{D}$$

Here, given a $\mathbb{Q}$-vector space V with basis $\alpha : \vec{e}_1, \vec{e}_2, \vec{e}_3$ we may consider $\mathcal{A}$ the presentation matrix of the endomorphism ϕ given by

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

In the free $\mathbb{Q}[x]$-module with basis e_1, e_2, e_3 (and canonical homomorphism $\pi : F \to V$ with $\pi(e_i) = \vec{e_i}$) the submodule

$$U = \mathsf{Span}_{\mathbb{Q}[x]}\{(x-1)e_1,\ -e_1 + (-2-x)e_2,\ e_1 + e_2 + (-1-x)e_3\}$$

corresponding to this presentation is given w.r.t. the basis

$$f_1 = e_1 + e_2 + (-1-x)e_3,\ f_2 = e_2 + e_3,\ f_3 = e_3$$

as

$$U = \mathbb{Q}[x]1f_1 \oplus \mathbb{Q}[x](x+1)f_2 \oplus \mathbb{Q}[x](x+1)^2 f_3$$

Consequently, as a $\mathbb{Q}[x]$-module

$$V \cong \mathbb{Q}[x]/\mathbb{Q}[x] \times \mathbb{Q}[x]/\mathbb{Q}[x](x+1) \times \mathbb{Q}[x]/\mathbb{Q}[x](x+1)^2 \cong \mathbb{Q}[x]/\mathbb{Q}[x](x+1) \times \mathbb{Q}[x]/\mathbb{Q}[x](x+1)^2$$

namely

$$V = \mathbb{Q}[x]\vec{f_2} \oplus \mathbb{Q}[x]\vec{f_3}$$

where

$$\pi(f_1) = \vec{f_1} = \vec{e_1} + \vec{e_2} + (-x-1)\vec{e_3} = \vec{0},\ \pi(f_2) = \vec{f_2} = \vec{e_2} + \vec{e_3},\ \ \pi(f_3) = \vec{f_3} = \vec{e_3}$$

Thus, w.r.t. the basis

$$\beta : \vec{f_2},\ \vec{f_3},\ \phi(\vec{f_3}) = \vec{e_1} + \vec{e_2} - \vec{e_3}$$

of the $\mathbb{Q}$-vector space V we have

$$\phi^\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -2 \end{pmatrix}$$

5.5.5 Solving systems of linear equations *

To solve a sytem $\mathcal{A}\boldsymbol{x} = \boldsymbol{b}$ over an euclidean ring compute $\mathcal{P}$, $\mathcal{P}^{-1}$, $\mathcal{D}$ and $\mathcal{Q}$ such that $\mathcal{P}\mathcal{A}\mathcal{Q} = \mathcal{Q}$. Substitute $\boldsymbol{y} = \mathcal{Q}\boldsymbol{x}$. Then the system is equivalent to

$$\mathcal{D}\boldsymbol{y} = \mathcal{P}\boldsymbol{b} =: \boldsymbol{c}$$

and solvable if and only if $c_i \in Rd_i$ for all i. The solution set is then given as

$$\{\mathcal{Q}\begin{pmatrix} r_1 q_1 \\ \vdots \\ r_m q_m \end{pmatrix} \mid r_i \in R\} \quad \text{where } d_i q_i = c_i$$

5.5.6 Elementary divisors

Consider two types of list of elements of an euclidean ring

- $d_1, \ldots, d_k$ such that $d_i \mid d_{i+1}$
- all members are 0, units, or prime powers and no two powers of the same or asscoiated prime are not associated

Lemma 5.5.3 *Up to association and order there is a 1-1-correspondence between the two types of list given by*

- *Both have the same number of units and the same number of zeros*
- *For any $d_i = p_1^{n_1} \cdot \ldots \cdot p_l^{n_l}$ of the first list, $p_1^{n_1}, \ldots, p_l^{n_l}$ belongs to the second*

Proof. To produce a list of the first type from the second, assume that associated primes are equal. For each prime p_j choose the highest exponent n_j with $p_j^{n_j}$ in the list and let

$$d_0 = \prod_j p_j^{n_j}$$

Cancel these $p_j^{n_j}$ from the list and continue with the remaining prime powers in the same way. This yields

$$d_l, \ldots, d_0 \quad \text{with } d_i | d_{i+1}$$

Add the units on the left, the zeros on the right end of the list and renumber if you like. □

5.6 Direct products and Chinese Remainder Theorem

5.6.1 Direct products

Given ideals I_i of a commutative ring resp. K-algebra R we say that I_1 and I_2 are *coprime* if $1 = r_1 + r_2$ for some $r_i \in I_i$. We also write $R = I_1 + I_2$

Lemma 5.6.1 *If I_1, I_2 are coprime ideals of R then there is a canonical surjective homomorphism $\phi : R \to R/I_1 \times R/I_2$ which is also R-linear with kernel $\ker \phi = I_1 \cap I_2$*

Proof. Define

$$\phi(a) = (\pi_1(a),\, \pi_2(a)) = (a + I_1,\; a + I_2)$$

Then ϕ is a homomorphism and $\phi(a) = 0 \Leftrightarrow \pi_1(a) = \pi_2(a) = 0 \Leftrightarrow a \in I_1 \cap I_2$. Consider $(\pi_1(a_1),\, \pi_2(a_2)) \in R/I_1 \times R/I_2$. We need $a \in R$ such that

$$\pi_i(a) = \pi_i(a_i) \quad \text{i.e. } a - a_i \in I_i \;\text{ for } i = 1, 2$$

By hypothesis, there are $r_i \in I_i$ such that $1 = r_1 + r_2$. In particular, $\pi_i(r_i b) = 0$ for all $b \in R$. Thus

$$\pi_i(r_j a_i) = \pi_i(r_j a_i + r_i a_i) = \pi_i((r_j + r_i)a_i) = \pi_i(a_i) \;\text{ and } \pi_j(r_j a_i) = 0 \quad \text{for } i \neq j$$

Thus, choose

$$a = r_2 r a_1 + r_1 a_2 \quad \square$$

Corollary 5.6.2 *If I_1, I_2 are coprime ideals of R then, canonically,*

$$R/(I_1 \cap I_2) \cong R/I_1 \times R/I_2$$

Corollary 5.6.3 *$R \cong R_1 \times R_2$ if and only if there are ideals I_1 and I_2 of R such that*

$$R_i \cong R/I_i, \;\; I_1 \cap I_2 = 0, \;\; I_1 + I_2 = R$$

Proof. If $R = R_1 \times R_2$ choose $I_1 = \{0\} \times R_2$ and $I_2 = R_1 \times \{0\}$. □

5.6.2 Chinese Remainder

In an euclidean ring, principal ideals (d_1) and (d_2) are coprime iff the elements d_1, d_2 are coprime.

Theorem 5.6.4 *Let R be an euclidean ring (and a K-algebra) and $d = LCM(d_1, d_2) \in R$. Then there is a canonical injective R-linear map which is also a ring (and a K-algebra) homomorphism*

$$\chi : R/Rd \to R/Rd_1 \times R/Rd_2, \;\; \phi(a + Rd) = (a + Rd_1, \; a + Rd_2)$$

and ϕ is an isomorphism if d_1, d_2 are coprime. If $R/Rd_0 \cong R/Rd_1 \times R/Rd_2$ as R-modules then $d_0 = d_1 d_2$ and d_1, d_2 are coprime. then

In particular, all *simultaneous congruences*

$$x \equiv b_1 \mod d_1, \;\; x \equiv b_2 \mod d_2$$

have unique solution modulo $d = d_1 d_2$ if and only if d_1, d_2 are coprime. Namely,

$$(*) \;\; x = b_1 a_2 d_2 + b_2 a_1 d_1 \quad \text{if } 1 = a_1 d_1 + a_2 d_2$$

Proof. Let $I_i = Rd_i$ and observe that $r(a + I_i) = ra + I_i$ so in the above lemma ϕ is R-linear and then so is χ. Also $I_1 \cap I_2 = LCM(d_1, d_2)$, obviuosly. Now, if $GCD(d_1, d_2) = 1$ then $LCM(d_1, d_2) = d_1 d_2 = d$ and we may apply Cor.5.6.2.

Now, assume $R/Rd_0 \cong R/Rd_1 \times R/Rd_2$ as R-modules. Then there is a surjective R-linear map of R/Rd_0 onto R/Rd_i whence $d_i \mid d_0$ by Lemma 5.2.5. Also, this implies that all simultaneous congruences $(*)$ have unique solution modulo d_0. Consider $b_1 = b_2 = 0$. Any multiple of $LCM(d_1, d_2)$ is a solution, in particular $d_1 d_2$ and d_0. By uniqueness it follows $d_0 = d_1 d_2 = LCM(d_1, d_2)$. But then $GCD(d_1, d_2) = 1$. □ In an more abstract approach, this isomorphism means in view of Cor.5.6.2 that there are ideals I_i of R/Rd_0 such that

$$I_1 + I_2 = R/Rd_0, \;\; I_1 \cap I_2 = \{0\}, \;\; (R/Rd_0)/I_i \cong R/Rd_i$$

Let $\pi : R \to R/Rd_0$ the canonical homomorphism. Then by the Isomorphism Theorem

$$J_i = \pi^{-1}(I_i) = \{r \in R \mid \pi(r) \in I_i$$

are ideals of R and $J_i = Rd_i$ by Cor.5.1.1. Moreover

$$J_1 + J_2 = \pi^{-1}(I_1 + I_2) = \pi^{-1}(R/Rd_0), \;\; J_1 \cap J_2 = \pi^{-1}(I_1 \cap I_2) = \pi^{-1}(\{0\}) = \mathsf{Ker}\, \pi = Rd_0$$

Thus, d_1, d_2 are coprime and $d_0 = LCM(d_1, d_2) = d_1 d_2$. □

Corollary 5.6.5 *Let M be an R-module, $g, g_1, g_2 \in M$ and $d, d_1, d_2 \in R$.*

(i) If $M = Rg$ with minimal annihilator d and $d = d_1 d_2$ with coprime d_1, d_2 then $M = Rd_2 g \oplus Rd_1 g$ with minimal annihilators d_i of $d_j g$ $(i \neq j)$.

(ii) If d_i is a minimal annihilator of g_i for $i = 1, 2$ and $M = Rg_1 \oplus Rg_2$ then M is cyclic if and only if d_1, d_2 are coprime and then $M = R(g_1 + g_2)$ with minimal annihilator $d_1 d_2$.

Proof. Ad (i): We have by Chinese remainder

$$M \cong R/d \cong R/Rd_1 \times R/Rd_2, \quad g \mapsto 1 \mapsto (1, 1)$$

Now, d_j is u unit modulo d_i for $i \neq j$ thus a generator of R/d_i - corresponding to $g_i = d_j g \in M$ under this isomorphism. And $r d_j g = 0$ iff $d_i d_j | r d_j$ iff $d_i | r$ so d_i is the minimal annihilator of g_i.

Ad (ii). We have $Rg_i \cong R/Rd_i$ whence by Chinese Remainder

$$M = Rg_1 \oplus Rg_2 \cong R/Rd_1 \times R/Rd_2$$

cyclic if and only if d_1, d_2 are coprime and $M \cong R/Rd$ where $d = d_1 d_2$. Now, under these isomorphisms, g_i corresponds to a unit u_i of R/Rd_i whence $g_1 + g_2$ to the unit (u_1, u_2) of $R/Rd_1 \times R/Rd_2$ and this to a unit u of R/Rd. □

5.6.3 Example

Consider an 8-dimensional $\mathbb{R}$-vector space V with basis $\alpha : \vec{e}_1, \ldots, \vec{e}_8$. The endomorphism ϕ with matrix $\phi^\alpha = A$ turns V into an $\mathbb{R}[x]$-module

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

The minimal polynomials of the blocks and their factorization in $\mathbb{R}[x]$ are given as

$$x^4 - 2x^3 + 2x^2 - 2x + 1 = (x-1)^2(x^2+1), \;\; x^2 - 2x + 1 = (x-1)^2, \; x^2 + 1$$

Generators of the invariant subspaces associated with the blocks are

$$\vec{e}_1, \vec{e}_6, \vec{e}_7$$

The prime powers factors give the list of elementary divisors of A

$$(x-1)^2, \; x^2+1, \; (x-1)^2, \; x^2+1$$

Generators of the associated invariant subspaces are

$$(x^2+1)\vec{e}_1,\ (x-1)^2\vec{e}_1,\ \vec{e}_6, \vec{e}_7$$

observe that the third block is Jordan, so its generator is the last vector $\vec{e}_6$. W.r.t. the basis

$$(x^2+1)\vec{e}_1 = \vec{e}_1 + \vec{e}_3,\ x(x^2+1)\vec{e}_1 = \vec{e}_2 + \vec{e}_4,\ (x-1)^2\vec{e}_1 = \vec{e}_1 - 2\vec{e}_2 + \vec{e}_3,$$

$$x(x-1)^2\vec{e}_1 = \vec{e}_2 - 2\vec{e}_3 + \vec{e}_4,\ \vec{e}_6,\ x\vec{e}_6 = \vec{e}_5 + \vec{e}_6,\ \vec{e}_7, \vec{e}_8$$

ϕ has matrix

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

The invariant divisors of A are obtained first multiplying as large as possible coprime elementary divisors. Here, this amounts to the minimal polynomial of the first and the produtc of the minimal polynomials of the second and third block.

$$(x-1)^2(x^2+1),\ (x-1)^2(x^2+1)$$

Thus, we have generators for invariant subspaces

$$\vec{e}_1,\ \vec{e}_6 + \vec{e}_7$$

and w.rt.t. the basis

$$\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4,\ \vec{e}_6 + \vec{e}_7,\ x(\vec{e}_6 + \vec{e}_7) = \vec{e}_5 + \vec{e}_6 + \vec{e}_8,$$

$$x^2(\vec{e}_6 + \vec{e}_7) = 2\vec{e}_5 + \vec{e}_6 - \vec{e}_7,\ x^3(\vec{e}_6 + \vec{e}_7) = 3\vec{e}_5 + \vec{e}_6 - \vec{e}_8$$

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

The minimal polynomial of A is the invariant divisor of highest degree, the characteristic polynomial the product of all elementary divisors

$$(x-1)^2(x^2+1),\ \ (x-1)^4(x^2+1)^2$$

Passing to the complexification, we can further factorize $x^2+1 = (x+i)(x-i)$. We obtain Jordan-basis and matrix

$$(x-1)(\vec{e}_1+\vec{e}_3) = -\vec{e}_1+\vec{e}_2+\vec{e}_4,\ \vec{e}_1+\vec{e}_3,$$

$$(x+i)(\vec{e}_1-2\vec{e}_2+\vec{e}_3) = i\vec{e}_1+(1-2i)\vec{e}_2+(-2+i)\vec{e}_3+\vec{e}_4$$

$$(x-i)(\vec{e}_1-2\vec{e}_2+\vec{e}_3) = i\vec{e}_1+(1+2i)\vec{e}_2+(-2-i)\vec{e}_3+\vec{e}_4,$$

$$\vec{e}_5,\ \vec{e}_6,\ (x+i)\vec{e}_7 = i\vec{e}_7+\vec{e}_8,\ (x-i)\vec{e}_7 = -i\vec{e}_7+\vec{e}_8$$

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

5.6.4 Multiple GCDs

Let R be an euclidean ring (and a K-algebra). $d \approx GCD(k_1,\ldots,k_n)$ if and only if $d|k_i$ for all i and $c \mid d$ for all c such that $c \mid k_i$ for all i. Clearly

$$GCD(k_1,\ldots,k_n) = GCD(GCD(k_1,\ldots,k_{n-1}),k_n)$$

If follows that there are

$$a_i \in R \text{ such that } \ d = a_1k_1+\ldots+a_nk_n$$

- Determine c_i with d_{n-1} =GCD$(k_1,\ldots,k_{n-1}) = c_1k_1+\ldots+c_{n-1}k_{n-1}$
- d =GCD(d_1,k_n). Determine b, a_n mit $d = bd_{n-1}+a_nk_n$
- $a_1 = bc_1,\ldots,a_{k-1} = bc_{k-1}$

Elements $m_1,\ldots,m_n$ of an euclidean ring are *pairwise coprime* if $GCD(m_i,m_j) \approx 1$ for all $i \neq j$ Equivalently

$$GCD(m_i,\ \frac{m}{m_i}) \approx 1 \ \text{ where } m = \prod_{j\neq i} m_j$$

5.6.5 Partial fractions *

Theorem 5.6.6 *If Q is the quotient field of an euclidean ring R then any $\frac{f}{g} \in Q$ with $\delta f < \delta g$ can be written as a sum of* partial fractions *of the form $\frac{a}{p^k}$ with irreducible $p|g$, $\delta a < \delta p$ and $\delta q < \delta f$.*

Proof. Let $g = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}$ with prime p_i and

$$q_i = p_1^{k_1} \cdot \ldots \cdot p_{i-1}^{k_{i-1}} \cdot p_{i+1}^{k_{i+1}} \cdot \ldots \cdot p_m^{k_m}$$

By Chinese Remainder we have

$$f = a_1 q_1 + \ldots + a_m q_m, \quad \frac{f}{g} = \frac{a_1}{p_1^{k_1}} + \ldots + \frac{a_m}{p_m^{k_m}}$$

Thus, we have to deal only with with quotients $\frac{a}{p^k}$. This is done by recursion on k. Long division yields

$$a = bp + r, \quad \frac{a}{p^k} = \frac{r}{p^k} + \frac{b}{p^{k-1}} \quad \text{with } \delta r < \delta p \quad \square$$

5.6.6 Chinese Remainder Theorem in multiple factors *

Given pairwise coprime m_i it follows

a. $GCD(\prod_{i\in I} m_i,\ \prod_{j\in j} m_j) \approx 1$ for $I \cap J = \emptyset$

b. There are $a_i \in R$ such that $1 = a_1 \frac{m}{m_1} + \ldots + a_n \frac{m}{m_n}$

c. There is a canonical isomorphism

$$\chi : R/(m) \to R/(m_1) \times \ldots \times R/(m_n) \ \text{ mit } a[\bmod m] \mapsto (a[\bmod m_1], \ldots, a[\bmod m_n])$$

d. The following *simultaneuos congruences* are solvable

$$x \equiv b_1 (\mod m_1), \ldots, x \equiv b_n (\mod m_n)$$

with solution given by

$$a = b_1 a_1 \frac{m}{m_1} + \ldots + b_n a_n \frac{m}{m_n}$$

Concerning d) compute

$$b_i a_i \frac{m}{m_i} \equiv 0 \equiv b_1 a_i \frac{m}{m_i} (\mod m_1) \quad \text{for } i > 1$$

$$a \equiv b_1 a_1 \frac{m}{m_1} + b_1 a_2 \frac{m}{m_2} + \ldots + b_1 a_n \frac{m}{m_n} \equiv b_1 (a_1 \frac{m}{m_1} + \ldots + a_n \frac{m}{m_n}) \equiv b_1 (\mod m_1)$$

and similarly for $m_2, \ldots, m_n$. This gives surjectivity of ϕ as in Lemma 5.6.1. Clearly, $(m) = \mathsf{Ker}\, \phi$. Thus, χ is an isomorphism. $\square$

5.6.7 Decomposition Lemma

The following can be obtained via the isomorphism given by the Chinese Remainder Theorem. Though, we formulate and prove it independently.

Lemma 5.6.7 *Let $d_0 \neq 0$ and $d_0 = d_1 d_2$ in the euclidean ring R and $1 = r_1 d_1 + r_2 d_2$. Let M be an R-module and $M_i = \{v \in M \mid d_i v = 0$. Let $\{i, j\} = \{1, 2\}$. Then*

(i) $M_0 = M_1 \oplus M_2$ *with submodules of* M

(ii) $v \mapsto d_j v$ *is an* R*-linear automorphism of* M_i *with inverse* $v \mapsto r_j v$

(iii) If $M_0 = Rg$ *then* $M_i = Rd_j g$

(iv) If $M_1 = Rg_1$ *and* $M_2 = Rg_2$ *then* $M_0 = R(g_1 + g_2)$

Proof.

a. By commutativity of R, the M_i are submodules.

b. $d_j v \in M_i$ for all $v \in M_0$, since $d_0 = d_1 d_2$

c. For all $v \in M$: $v = 1v = r_2 d_2 v + r_1 d_1 v \in M_1 + M_2$ whence $M = M_1 + M_2$

d. For $v \in M_i$: $r_j d_j v = 0 + r_j d_j v = r_i d_i v + r_j d_j v = 1v = v$ whence (ii)

e. Let $v_i \in M_i$ and $v_1 + v_2 = 0$. Then $0 = r_j d_j (v_1 + v_2) = r_j d_j v_1 + r_j d_j v_2 = r_j d_j v_i = v_i$. Thus $M_0 = M_1 \oplus M_2$.

f. Let $M_0 = Rg$ and $v_1 \in M_1$. Then $v_1 = rg = r1g = r(r_1 d_1 + r_2 d_2)g = rr_1 d_1 g + rr_2 d_2 g$ with $rr_1 d_1 g \in M_2$ whence $v_1 = rr_2 d_2 g$.

g. Let $M_1 = Rg_1$ and $M_2 = Rg_2$. For any v there are $v_i \in M_i$ and $s_i \in R$ such that $v = v_1 + v_2 = s_1 g_1 + s_2 g_2 = s_1 r_2 d_2 g_1 + s_2 r_1 d_1 g_2 = (s_1 r_2 d_2 + s_2 r_1 d_1)(g_1 + g_2)$ □

5.5.7 Minimal annihilators and Cayley-Hamilton

Given an R-module M over an euclidean ring

$$\mathsf{Ann}(M) = \{r \in R \mid rv = 0 \text{ for all } v \in M\}$$

is an ideal of R, obviuosly, whence $\mathsf{Ann}(M) = (d)$ where $d \in \mathsf{Ann}(M)$ with $\delta(d)$ minimal. d is unique up to association and called *minimal annihilator* for M. In the case of an $K[x]$-module given by an endomorphism ϕ of a finite dimensional vector space V, a minimal annhilator is given as the normed $d(x) \in K[x]$ of minimal degree suh that $d(\phi) = 0$ and called the *minimal polynomial* of ϕ. By finite dimension, $d(x)$ *neq*0 exists (otherwise, $K[x]$ would be isomorphic to a subspace of V).

Proposition 5.5.8 *If the module* M *is presented by a diagonal matrix with entries* $d_1, \ldots, d_n$ *then any minimal annihilator is* $d \approx LCM(d_1, \ldots, d_n)$

Proof. We have $M = \bigoplus_i Rv_i$ with $Rv_i \cong R/(d_i)$. Let d' the LCM of the d_i. Then $d' \in \mathsf{Ann}(M)$ wgence $d \mid d'$. On the other hand, $dv_i = 0$ so $d_i \mid d$ for all i and $d' \mid d$. □

Corollary 5.5.9 Cayley-Hamilton. *For any endomorphism* ϕ *of a finite dimensional* K*-vector space, the minimal polynomial* $d(x)$ *divides the characteristic polynomial* $\chi(x)$ *in* $K[x]$. *In particular,* $\chi(]phi) = 0$

Proof. Assume ϕ given by the matrix $A \in K^{n\times n}$.then $A - xE$ is a presentation matrix for the $K[x]$-module V given by ϕ. There are invertible $\mathcal{P}$ and $\mathcal{Q} \in K[x]^{n\times n}$, in particular $\det \mathcal{P} \in K^*$ and $\det \mathcal{Q} \in K^*$) such that $\mathcal{P}(A - xE)\mathcal{Q}$ is diagonal with $d_i(x) \mid d_{i+1}(x)$. It follows

$$\chi(x) = \det(A - xE) = \det \mathcal{P} \det(A - xE) \det \mathcal{Q} \approx d_1(x) \cdot \ldots \cdot d_n(x)$$

where $d_n(x) \approx d(x)$. □

5.5.8 Extension to principal ideal domains*

Recall that a *principal ideal domain* is an integral domain in which every ideal is principal. In particular, for all a, b there is d such that

$$(a) + (b) = (d)$$

equivalently, there are a_1, b_1, d, x, y such that

$$a = a_1 d,\ b = b_1 d,\ a_1 x + b_1 y = 1$$

the latter obtained from $ax + by = d$ by cancellation. It follows

$$\begin{pmatrix} a_1 & b_1 \\ -y & x \end{pmatrix} \begin{pmatrix} x & -b_1 \\ y & a_1 \end{pmatrix} = E$$

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x & -b_1 \\ y & a_1 \end{pmatrix} = d \begin{pmatrix} a_1 & b_1 \end{pmatrix} \begin{pmatrix} x & -b_1 \\ y & a_1 \end{pmatrix} = d \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} d & 0 \end{pmatrix}$$

and by transposing

$$\begin{pmatrix} x & y \\ -b_1 & a_1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} d \\ 1 \end{pmatrix}$$

Thus, the Theorem on Invariant Divisors extends to principal ideal domains and so does its corollary: Any finitely generated module over an principal ideal domain is a direct sum of finitely many cyclic ones. Also, principal ideal domains have unique factorization.

5.6.7 Addenda et corrigenda

a. Lemma 22.2.5 applies to ring homomorphisms as well as to R-linear maps. Proof immediate by Homomorphism Theorem

b. proof of 22.3.3 read $r, s \in I$

c. proof of 22.4.1 read $c|(ra + sb)$

d. Algorithm 22.5.2: Induction on $(\delta(\mathcal{A}), n$ where n is th number of row. Read $[Sk := Sk - qS_j]$ and $[Zk := Zk - qZ_i]$ with $\delta(a_{kj} - q_{ij}) < \delta(a_{ij})$.

e. In Lemma 22.6.1 ϕ is also R-linear. In the proof: Choose $a = r_2 a_1 + r_1 a_2$.

f. In Thm. 22.6.4. Let $d = LCM(d_1, d_2)$ not $d = d_1 d_2$.

Chapter 6

Canonical forms of matrices

A general assumption for this chapter is that V is an n-dimensional K-vector space with an endomorphism ϕ such that the minimal polynomial is a product of linear factors. We consider V as the $K[x]$-module where $x\vec{v} = \phi(\vec{v})$.

6.1 Jordan matrices and bases

6.1.1 Jordan-chains and Jordan-blocks

For any eigenvalue λ of ϕ define

$$\phi_\lambda = \phi - \lambda\mathsf{id}, \quad \text{i.e. } \phi_\lambda(\vec{x}) = \phi(\vec{x}) - \lambda\vec{x}$$

The λ- *Jordan-chain* $J_\lambda(\vec{v})$ of the vector $\vec{v}$ with *head* or *start vector* $\vec{v}$ and *tail* or *eigenvector* $\sigma(\vec{v})$ consists of the vectors

$$\vec{0} \neq \sigma(\vec{v}) = \phi_\lambda^{k-1}(\vec{v}),\ \phi_\lambda^{k-2}(\vec{v}), \ldots, \phi_\lambda(\vec{v}),\ \vec{v} \quad \text{with } \phi_\lambda\sigma(\vec{v}) = \phi_\lambda^k(\vec{v}) = \vec{0}$$

and has *length* k.

$$\vec{0} \underset{\phi_\lambda}{\longleftarrow} \phi_\lambda^{k-1}(\vec{v}) \underset{\phi_\lambda}{\longleftarrow} \phi_\lambda^{k-2}(\vec{v}) \underset{\phi_\lambda}{\longleftarrow} \phi_\lambda^{k-3}(\vec{v}) \longleftarrow \cdots \longleftarrow \phi_\lambda(\vec{v}) \underset{\phi_\lambda}{\longleftarrow} \vec{v}$$

The vectors in the chain form an independent list: If $\sum_{i=0}^{k-1} r_i\phi_\lambda^i(\vec{v}) = \vec{0}$ then then $\sum_{i=0}^{k-2} r_i\phi_\lambda^{i+1}(\vec{v}) = \phi(\vec{0}) = \vec{0}$ so by induction $r_i = 0$ for $i > k-1$ and then $r_{k-1}\phi^{k-1}\vec{v} = \vec{0}$ and $r_{k-1} = 0$.

A λ- *Jordan-block* is a matrix

$$= J_{\lambda,n} = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & 0 & \lambda & 1 \\ 0 & & \ldots & 0 & \lambda \end{pmatrix}.$$

Lemma 6.1.1 *Let $\phi \in \mathsf{End}(V)$ and $\beta : \vec{v}_1, \ldots \vec{v}_k$ a basis of V. Then the matrix ϕ_β of ϕ w.r.t. β is a λ-Jordan-block if and only if $\vec{v}_1, \ldots, \vec{v}_K$ is a λ-Jordan-chain.*

Proof. To have a λ-Jordan block as matrix means that

$$\phi(\vec{v}_1) = \lambda\vec{v}_1, \;\; \phi(\vec{v}_i) = \lambda\vec{v}_i + \vec{v}_{i-1}, \;\; \text{i.e. } \vec{v}_{i-1} = \phi(\vec{v}_i) - \lambda\vec{v}_i \quad \square$$

Proposition 6.1.2 *V is a cyclic $K[x]$-module with minimal polynomial $(x-\lambda)^n$ degree n if and only if the K-vector space V admits a basis which is a λ-Jordan-chain (of length n). If so, then*

(i) The $K[x]$-module V is isomorphic to $K[x]/((x-\lambda)^n)$.

(ii) W.r.t. some Jordan-chain β, the matrix of of ϕ is a Jordan-block $J_{\lambda,n}$.

(iii) $\vec{v}$ is a generator of the $K[x]$-module V if and only if $\vec{v}$ is a start vector of a Jordan-chain.

Proof. Assume that V is cyclic with minimal polynomial $(x-\lambda)^n$, then $V \cong K[x]/((x-\lambda)^n)$. The $(x-\lambda)^k$ generate the K-vector space $K[x]$: inductively on obtains all x^k since $(x-\lambda)^k = x^k + p_k(x)$ with a polynomial $p_k(x)$ of degree $< k$. Thus, the $(x-\lambda)^k$, $k < n$, generate the K-vector space and form a basis $\beta : (x-\lambda)^{n-1}, \ldots, 1$ (since dim $= n$). This tranfers ot V via the isomorphism. That ϕ has w.r.t. β matrix $J_{\lambda,n}$ is seen from

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

Conversely, given a basis consisting of a Jordan-chain, the start-vector $\vec{v}$ is a generator, obviously. Also $(x-\lambda)^m(\vec{v}) = \vec{0} \Leftrightarrow m \geq n$ so $(x-\lambda)^n$ is the minimal polynomial. $\square$

6.1.2 Jordan-matrices and bases

A *Jordan basis* for ϕ is a basis which is a list-concatenation of λ-Jordan chains, λ ranging over the eigenvalues of ϕ. A *Jordan matrix* is a block-diagonal matrix with λ-Jordan-blocks on the diagonal.

Corollary 6.1.3 *β is a Jordan basis for the endomorphism ϕ if and only if ϕ_β is a Jordan matrix.*

Theorem 6.1.4 *Let ϕ be an endomorphism of a finite dimensional K-vector space V such that the minimal polynomial is a product of linear factors from $K[x]$ (which is guaranteed by the Fundamental Theorem of Algebra if $K = \mathbb{C}$). Then V admits a Jordan-basis of ϕ. The associated Jordan-matrix J is uniquely determined up to permutation of blocks, For each EW λ, the number of λ-Jordan blocks is the geometric multplicity, the sum of the block sizes the algebraic multiplicity, i.e. the number of occurences of the EW λ on the diagonal.*

Proof. Let $\alpha : \vec{e}_1, \ldots, \vec{e}_n$ be a basis of V and ϕ be given by A w.r.t. α. Then $A - xE$ is a presentation matrix for the $K[x]$-module V w.r.t. the basis $e_1, \ldots, e_n$of the free $K[x]$-module F. In particular, there is a canonical $K[x]$-linear $\pi : F \to V$ with $\pi(e_i) = \vec{e}_i$. By the Theorem on Invariant Divisors 5.5.1 there are invertible matrices $\mathcal{P}$ and $\mathcal{Q}$ in $K[x]^{n\times n}$ such that $\mathcal{PAQ} = \mathcal{D}$ is diagonal and presentation matrix of V w.r.t. the basis $f_1, \ldots, f_n$

of the free $K[x]$-module given by the columns of $\mathcal{P}^{-1}$ cf Cor.4.3.3. Then we have a direct decomposition into cyclic submodules

$$V = K[x]\vec{f_1} \oplus \ldots \oplus K[x]\vec{f_n}$$

where $\vec{f_i} = \pi(f_i)$.

Now, by Prop.??lcmann the LCM of the $d_i(x)$ is the minimal polynomial, so by assumption a product of linear factors. Thus each $d_i(x)$ is a product with pairwise distinct λ_j

$$(x - \lambda_1)^{k_1} \cdot \ldots \cdot (x - \lambda_l)^{k_l}$$

By iterated application of the Decomposition Lemma 5.6.7 we get

$$K[x]\vec{f_i} = \bigoplus_{h=1}^{l} K[x]\vec{f}_{ih} \quad \text{with } \vec{f}_{ih} = \prod_{j \neq h}(x - \lambda_j)^{k_j}\vec{f_i}$$

- $K[x]\vec{f}_{ih}$ cyclic with minimal polynomial $(x - \lambda_h)^{k_h}$

Thus, each of these invariant subspaces admits a basis which is a λ_h-Jordan-chain of length k_h. Taken together, these bases form a Jordan basis of V. □

6.1.3 Canonical forms

Corollary 6.1.5 *Any endomorphism of a finite dimensional K-vector space admits a basis such that the matrix is block diagonal with companion matrices of polynomials $m_i(x) \in K[x]$. One may require*

- $m_i(x)|m_{i+1}(x)$ *for all i,* Frobenius- *or* rational canonical form *with* invariant divisors $m_i(x)$
- $m_i(x) = p_i(x)^{k_i}$ *with prime $p_i(x)$,* Weierstrass canonical form *with* elementary divisors $p_i(x)^{k_i}$
- *If $p_i(x) = (x - \lambda_i)^{k_i}$ for all i (e.g. if K is algebraically closed, say $K = \mathbb{C}$), then in the Weiertrass canonical form on may replace the companion matrices by Jordan blocks (changing the basis) to obtain* Jordan canonical form

Proof. Follow the proof of Jordan canonical from to the diagonal presentation with $d_i(x) \mid d_{i+1}(x)$. This yields Frobenius. Use Decomposition Lemma to pass to Weierstrass.

6.1.4 Example

Let V a $\mathbb{Q}$-vector space with basis $\vec{e}_1, \ldots, \vec{e}_5$ and ϕ the endomorphism with matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 1 & 1 & 0 & 1 & -4 \\ -1 & 0 & 0 & 1 & -2 \end{pmatrix}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 2-x & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1-x & -4 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -2-x
\end{array}$$

$$\boxed{S1 := S1 - (2-x)S2}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -(2-x)^2 & 2-x & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & -1+x & 1 & 0 & 1-x & -4 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -2-x
\end{array}$$

$$\boxed{Z2 := Z2 - (2-x)Z1,\ Z4 := Z4 - Z1,\ S1 \leftrightarrow S2}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
2-x & 1 & 0 & 0 & 0 & 0 & -(2-x)^2 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & -1+x & 0 & 1-x & -4 \\
0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & -2-x
\end{array}$$

$$\boxed{S2 := S2 + S4,\ S5 := -(S5 + (2+x)S4)}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
2-x & 1 & 0 & 0 & 0 & 0 & -(2-x)^2 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1-x & x^2+x+2 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0
\end{array}$$

$$\boxed{Z4 := Z4 - (1-x)Z5}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
2-x & 1 & 0 & 0 & 0 & 0 & -(2-x)^2 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
1 & 0 & 0 & 1 & 1-x & 0 & 0 & 0 & 0 & x^2+x+2 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0
\end{array}$$

$$\boxed{S4 \leftrightarrow S2,\ Z5 \leftrightarrow Z2,\ Z5 \leftrightarrow Z4,\ S4 := -S4}$$

$$\begin{array}{ccccc||ccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
2-x & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 2-x & 0 & 0 \\
1 & 1-x & 0 & 0 & 1 & 0 & 0 & 0 & (2-x)^2 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x^2+x+2
\end{array}$$

Diagonalize the characteristic matrix $E \mid A - xE \quad \rightsquigarrow \quad (f_1, \ldots, f_5) \mid \mathcal{A}'$

$$d_1 = d_2 = 1,\ d_3 = 2 - x,\ d_4 = (2-x)^2,\ d_5 = x^2 + x + 2$$

Characteristic polynomial	$(x-2)^3(x^2+x+2)$
Minimal polynomial	$(x-2)^2(x^2+x+2)$
Eigenvalues	2 (geom. multipl. 2, alg. multipl. 3)
	$-\frac{1}{2} \pm \frac{\sqrt{7}}{2}i$

The matrix transformation yields the decomposition of V into cyclic submodules: generators are the images $\vec{f_1}, \ldots, \vec{f_5}$ in V of the new basis vectors $f_1, \ldots, f_5$ of the free modul $\mathbb{Q}[x]^5$. Those which are $\vec{0}$ may be discarded, here $\vec{f_1}$ and $\vec{f_2}$ since at the associated position in the diagonal presentation matrix $\mathcal{A}'$ one has a 1 (i.e. $1f_1 \stackrel{!}{=} 0$). We verify $\vec{f_1} = 0$

$$\vec{f_1} = \vec{e_1} + (2-x)\vec{e_2} + \vec{e_4} = \vec{e_1} + 2\vec{e_2} - \phi(\vec{e_2}) + \vec{e_4} = \vec{e_1} + 2\vec{e_2} - (\vec{e_1} + 2\vec{e_2} + \vec{e_4}) + \vec{e_4} = 0$$

This leaves us with (in general, applications of A would be required)

$$\vec{f_3} = \vec{e_3},\ \vec{f_4} = \vec{e_2},\ \vec{f_5} = \vec{e_4}$$

and relations

$$(2-x)f_3 \stackrel{!}{=} 0,\ (2-x)^2 f_4 \stackrel{!}{=} 0,\ (x^2+x+2)f_5 \stackrel{!}{=} 0$$

From there we can read the structure of V as $\mathbb{Q}[x]$-module and choose suitable bases

$$\begin{array}{rcccccc}
V & \cong & \mathbb{Q}[x]/(2-x) & \times & \mathbb{Q}[x]/(2-x)^2 & \times & \mathbb{Q}[x]/(x^2+x+2) \\
 & & 1 & & 1, x & & 1, x \\
V & = & \mathbb{Q}[x]\vec{f_3} & \oplus & \mathbb{Q}[x]\vec{f_4} & \oplus & \mathbb{Q}[x]\vec{f_5} \\
 & & \vec{f_3} & & \vec{f_4}, x\vec{f_4} & & \vec{f_5}, x\vec{f_5} \\
 & & \vec{e_3} & & \vec{e_2}, \phi(\vec{e_2}) & & \vec{e_4}, \phi(\vec{e_4}) \\
 & & \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} & & \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\2\\0\\1\\0\end{pmatrix} & & \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\1\end{pmatrix}
\end{array}$$

The associated matrix of ϕ is

$$A' = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

Since the polynomials are powers of primes, A' is in Weierstrass canonical form. Other generators for the cyclic submodules may be obtained multiplying with a polynomial invertible modulo the mimimal polynomial of the submodule.

Frobenius canonical form is obtained combining submodules into sums according to Chinese Remainder. Here, this applies to $(2-x)^2$ and x^2+x+1. Adding the generators $\vec{f_4}$ and $\vec{f_5}$ one gets a generator $\vec{f_4}+\vec{f_5}$ of the direct sum

$$\begin{array}{rcccc}
V & \cong & \mathbb{Q}[x]/(2-x) & \times & \mathbb{Q}[x]/((2-x)^2(x^2+x+2) \\
 & & 1 & & 1, x, x^2, x^3 \\
V & = & \mathbb{Q}[x]\vec{f_3} & \oplus & \mathbb{Q}[x](\vec{f_4}+\vec{f_5}) \\
 & & \vec{f_3} & & \vec{f_4}+\vec{f_5}, x(\vec{f_4}+\vec{f_5}), x^2(\vec{f_4}+\vec{f_5}), x^3(\vec{f_4}+\vec{f_5}) \\
 & & \vec{e_3} & & \vec{e_2}+\vec{e_4}, \phi(\vec{e_2}+\vec{e_4}), \phi^2(\vec{e_2}+\vec{e_4}), \phi^3(\vec{e_2}+\vec{e_4}) \\
 & & \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} & & \begin{pmatrix}0\\1\\\\1\\0\end{pmatrix}, \begin{pmatrix}1\\2\\0\\2\\1\end{pmatrix}, \begin{pmatrix}4\\4\\0\\1\\-1\end{pmatrix}, \begin{pmatrix}12\\8\\0\\13\\-1\end{pmatrix}
\end{array}$$

with matrix in Frobenius canonical form

$$A'' = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -8 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\text{da } (2-x)^2(x^2+x+2) = x^4 - 3x^3 - 2x^2 - 4x + 8.$$

Again there are other generators of $\mathbb{Q}[x](\vec{f_4}+\vec{f_5})$: all

$$b(x)\vec{f_4} + a(x)\vec{f_5}$$

$$\text{where } b(x) \text{ invertible} \mod (2-x)^2, \quad a(x) \text{ invertible} \mod x^2+x+2$$

To obtain Jordan canonical form, do everything over $\mathbb{C}$. The decomposition into cyclics remains valid but can be refined due to

$$x^2+x+2 = (x-\lambda)(x-\overline{\lambda}) \text{ with } \lambda = \frac{-1}{2} + \frac{\sqrt{7}}{2}i,\ \overline{\lambda} = \frac{-1}{2} - \frac{\sqrt{7}}{2}i$$

$$\begin{array}{ccccccccc}
\mathbb{C}^5 & \cong & \mathbb{C}[x]/(2-x) & \times & \mathbb{C}[x]/(2-x)^2 & \times & \mathbb{C}[x]/(x-\lambda) & \times & \mathbb{C}[x]/(x-\overline{\lambda}) \\
 & & 1 & & x-2,1 & & 2i\Im\lambda & & -2i\Im\lambda \\
\mathbb{C}^5 & = & \mathbb{C}[x]\vec{f_3} & \oplus & \mathbb{C}[x]\vec{f_4} & \oplus & \mathbb{C}[x]\vec{f_{51}} & \oplus & \mathbb{C}[x]\vec{f_{52}} \\
 & & \vec{f_3} & & (x-2)\vec{f_4}, \vec{f_4} & & (x-\overline{\lambda})\vec{f_5} & & (x-\lambda)\vec{f_5} \\
 & & \vec{e_3} & & (\phi-2id)(\vec{e_2}), \vec{e_2} & & (\phi-\overline{\lambda}id)\vec{e_4} & & (\phi-\lambda id)\vec{e_4} \\
 & & \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} & & \begin{pmatrix}1\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} & & \begin{pmatrix}0\\0\\0\\\frac{-1}{2}+\frac{\sqrt{7}}{2}i\\1\end{pmatrix} & & \begin{pmatrix}0\\0\\0\\\frac{-1}{2}-\frac{\sqrt{7}}{2}i\\1\end{pmatrix}
\end{array}$$

Jordan matrix von ϕ

$$\begin{pmatrix}
2 & 0 & 0 & 0 & 0 \\
0 & 2 & 1 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & \frac{-1}{2}+\frac{\sqrt{7}}{2}i & 0 \\
0 & 0 & 0 & 0 & \frac{-1}{2}-\frac{\sqrt{7}}{2}i
\end{pmatrix}$$

Namely $x-\overline{\lambda} \equiv 2i\Im\lambda \mod x-\lambda$ and $x-\lambda \equiv -2i\Im\lambda \mod x-\overline{\lambda}$ since

$$1 = \frac{-1}{2i\Im\lambda}(x-\lambda) + \frac{1}{2i\Im\lambda}(x-\overline{\lambda})$$

with eigenvectors

$$\vec{f_3}, (x-2)\vec{f_4} \text{ w.r.t. } EW2$$

$$\vec{f_{51}} \text{ w.r.t. } EW\ \frac{-1}{2}+\frac{\sqrt{7}}{2}i$$

$$\vec{f_{52}} \text{ w.r.t. } EW\ \frac{-1}{2}-\frac{\sqrt{7}}{2}i$$

6.1.5 Review: Structure of an endomorphism

a. The basis α of a free $K[x]$-module and matrix $A-xE$ may be transformed into a basis $f_1,\ldots,f_n$ and diagonal matrix with normed diagonal entries $1,\ldots,1,d_s,\ldots,d_n \in K[x]$, $d_i \not\approx 1$ forr $i \geq s$. This yields a direct decomposition into cyclic submoduls with minimal polynomials d_i

$$V = K[x]\vec{f_s} \oplus \ldots \oplus K[x]\vec{f_n}$$

$$\text{where } \vec{f_j} = \sum_{i,k} b_{jik}\phi^k(\vec{e_i}) \text{ in } V \text{ if } f_j = \sum_i (\sum_k b_{jik}x^k)e_i \text{ in the free module.}$$

b. The K-vector space V has basis

$$\vec{f_i}, \phi(\vec{f_i}), \ldots, \phi^{n_i-1}(\vec{f_i}),\ \ i = s,\ldots,n,\ n_i = \deg d_i.$$

W.r.t. this basis ϕ has matrix A', block diagonally composed from *Frobenius matrices* or *companion matrices* of the polynomials d_i.

$$A'_i = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -r_{i0} \\ 1 & 0 & 0 & \dots & 0 & -r_{i1} \\ 0 & 1 & 0 & & 0 & -r_{i2} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & & 0 & -r_{in_i-2} \\ 0 & 0 & & \dots & 1 & -r_{in_i-1} \end{pmatrix} \quad \text{where } d_i = \sum_{k=0}^{n_i} r_{ik}x^k.$$

c. $\det(A - xE) \approx d_1 \cdot \ldots \cdot d_n$, $n = \dim_K V = \deg \det(A - xE) = \sum_i \deg d_i$.

d. One may achieve $d_i \mid d_{i+1}$ for $i < n$. These d_i are unique up to $\approx$ and called *invariant divisors* of ϕ, also d_n is the minimal polynomial of ϕ. A' is a *Frobenius* or *rational canonical form* of A.

e. (Cayley-Hamilton) The $LCM(d_1, \dots, d_n)$ is associated to the minimal polynomial $d(x)$ of ϕ and divides the characeristic polynomial of ϕ. In particular $d(\phi) = 0$.

f. Factorizing the d_i into powers d_{ij} of coprime irreducible polynomials $d_i = d_{i1} \cdot \ldots \cdot d_{im_i}$, one obtains a direct decomposition into primary cyclic submodules with minimal polynomial d_{ih}

$$V = K[x]f_{s1} \oplus \dots \oplus K[x]f_{sm_s} \oplus \dots \oplus K[x]f_{n1} \oplus \dots \oplus K[x]f_{nm_n}$$

where $f_{ih} = (d_i/d_{ih})f_i = (d_{i1}(\phi) \circ \dots d_{i,h-1}(\phi) \circ d_{i,h+1}(\phi) \circ \dots \circ d_{im_i}(\phi))(f_i)$ in V

g. The K-vector space V has basis

$$f_{ih}, \phi(f_{ih}), \dots, \phi^{n_{ih}-1}(f_{ih}), \quad i = s, \dots, n, \ h = 1, \dots, m_i, \ n_{ih} = \deg d_{ih}.$$

W.r.t. this basis ϕ has matrix A', block diagonally composed from the companions of the d_ih.

$$A'_{ih} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -r_{ih0} \\ 1 & 0 & 0 & \dots & 0 & -r_{ih1} \\ 0 & 1 & 0 & & 0 & -r_{i2} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & & 0 & -r_{ihn_{ih}-2} \\ 0 & 0 & & \dots & 1 & -r_{ihn_{ih}-1} \end{pmatrix} \quad \text{where } d_{ih} = \sum_{k=0}^{n_{ih}} r_{ihk}x^k.$$

h. The *elementary divisors* d_{ih} of ϕ are unqiue up to $\approx$. A' is a *Weierstrass canonical form* of A.

i. If the d_{ih} are powers of linear polynomials $d_{ih} = (x - \lambda_{ih})^{n_{ih}}$ then the K-vector space V has *Jordan basis*

$$(\phi - \lambda_{ih}id)^{n_{ih}-1}(f_{ih}), \dots, (\phi - \lambda_{ih}id)(f_{ih}), f_{ih}, \quad i = s, \dots, n, \ h = 1, \dots, m_i.$$

W.r.t. this basis ϕ has matrix A' block diagonally composed from $n_{ih} \times n_{ih}$ *Jordan blocks*

$$J_{\lambda_{ih},n_{ih}} = \begin{pmatrix} \lambda_{ih} & 1 & 0 & \dots & 0 \\ 0 & \lambda_{ih} & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \lambda_{ih} & 1 \\ 0 & & \dots & 0 & \lambda_{ih} \end{pmatrix}.$$

j. This matrix is unique, if it exists, up to order of blocks and called *Jordan canonical form* of A. It can be achieved passing from K to its algebraic closure.

6.2 Primary decomposition

6.2.1 Primary decomposition for modules

An R-module M is a *torsion module* if $T(M) = M$. In a particular, any $K[x]$-module which is a finite dimensional K-vector space is a torsion module.

Given a prime element p of R, a torsion module is *p- primary* if for any $v \in M$ there is $k \geq 0$ such that $p^k v = 0$. Of course, then M is q-primary for any q associated with p.

Theorem 6.2.1 *Let R be an euclidean ring and M a a finitely generated torion R-module. Then*

(i) There is $d \neq 0$ with $\delta(d)$ minimal such that $dv = 0$ for all $v \in M$. d is unique up to association and called the minimal annihilator *of M.*

(ii) Given a factorization $d \approx \prod_{i=1}^{l} p_i^{k_i}$ into primes $p_i \not\approx p_j$ for $i \neq j$.

$$M = M_1 \oplus \ldots \oplus M_l, \quad M_i = \{v \in M \mid p_i^{k_i} v = 0\}$$

(iii) M has unique direct decomposition into p_i-primary submodules $N_i \neq 0$ with non-associated primes.

(iv) In (iii) one has minimal annihilators $p_i^{k_i}$ of N_i $(i \leq l)$ if and only if $d = \prod_{i=1}^{l} p_i^{k_i}$ is a minimal annhilator of M. In particular, l is unique and the $p_i^{k_i}$ are unique up to order and association.

(v) d in (iv) is an invariant divisor of M of highest degree.

If $R = K[x]$ and if $p_i = x - \lambda_i$ then M_i is the *generalized eigenspace* w.r.t. eigenvalue λ_i and $d = d(x)$ is called the *minimal polynomial* of M.

Proof of the Thm. Ad (i). For each generator v_i, choose $r_i \neq 0$ with $r_i v_i = 0$. Then $rv = 0$ for all v where $r = \prod_i r_i$. Now, d is a generator of the ideal $\{r \in R \mid rv = 0 \text{ for all } v \in M\}$. (ii) follows with Lemma5.6.7. If $M = \bigoplus N_i$ with p_i primary N_i then $N_i \subseteq M_i$ and so $N_i = M_i$ since both sums are direct. Ad (iv). Clearly $dv = 0$ for $v \in N_i$ so for all $v \in M$. d is minimal, since $p_i^{k_i - 1} v \neq 0$ for some $v \in N_i$. The uniqueness of the $p_i^{k_i}$ follows from unique factorization. (v) is immediate by Thm.6.5.1(iii). $\square$

6.2.2 Example: Generalized eigenspaces

$\mathbb{Q}[x]$-module $\mathbb{Q}^6$ given by A w.r.t. canonical basis.

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

$$A - 3E = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \quad (A - 3E)^2 = \begin{pmatrix} 0 & 0 & 0 & -1 & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathsf{rk}(A - 3E) = 4,\ \mathsf{rk}(A - 3E)^2 = \mathsf{rk}(A - 3E)^3 = 3$$

Minimal polynomial for generalized eigenspace w.r.t. $\lambda = 3$ is $(x - 3)^2$

$$\text{basis } \beta_3 \text{ of } \mathsf{ker}(A - 3E)^2 : \quad \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$$

$$A - 2E = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 9 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (A - 2E)^2 = \begin{pmatrix} 1 & 2 & 2 & 1 & 2 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathsf{rk}(A - 2E) = \mathsf{rk}(A - 3E)^2 = 3$$

Minimal polynomial for generalized eigenspace w.r.t. $\lambda = 2$ is $(x - 2)^2$

$$\text{basis } \beta_2 \text{ of } \mathsf{ker}(A - 2E)^2 :$$

$$-\boldsymbol{e}_3 + \boldsymbol{e}_6,\ -\boldsymbol{e}_2 + \boldsymbol{e}_5,\ -\boldsymbol{e}_1 + \boldsymbol{e}_4$$

Minimal polynomial of A is $(x - 3)^2(x - 2)^2$

$${}_\alpha T_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad A\text{"} = {}_\alpha T_\beta^{-1} A \, {}_\alpha T_\beta = \begin{pmatrix} 3 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} A'|V_3 & O \\ O & A'|V_2 \end{pmatrix}$$

$$A'|V_3 = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad A'|V_2 = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\text{J-basis } V_3 : \gamma_3 : \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3 - \boldsymbol{e}_2$$

$$\text{J-basis: } V_2 : \gamma_2$$

$$-\boldsymbol{e}_3 + \boldsymbol{e}_6,\ -\boldsymbol{e}_2 + \boldsymbol{e}_5,\ -\boldsymbol{e}_1 + \boldsymbol{e}_4 - (-\boldsymbol{e}_2 + \boldsymbol{e}_5)$$

$${}_\beta T_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad {}_\alpha T_\gamma = {}_\alpha T_\beta \, {}_\beta T_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$J = {}_\alpha T_\gamma^{-1} A \, {}_\alpha T_\gamma = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

6.3 Nilpotent matrices

6.3.1 Shift

Lemma 6.3.1 *Let $\mu \in K$ and $\psi = \phi_\mu = \phi - \mu\mathsf{id}$ Then λ is an EW of ϕ if and only if $\lambda - \mu$ is an EW of ψ. Moreover β is a Jordan-basis for ϕ if and only if it is so for ψ. U is ψ-invariant if and only if U is ϕ-invariant.*

Proof. $\psi_{\lambda-\mu} = \phi_\lambda$. □

This reduces the case where the minimal polynomial is $(x-\lambda)^m$, i.e. $(\phi - \lambda\mathsf{id})^m = 0$ to the *nilpotent* case: we may assume that

- $\phi^m = 0$ for some $m \leq n$

6.3.2 Module versus vector space

Let $R = K[x]$. A list $\vec{v}_1, \ldots, \vec{v}_k$, Since primary decomposition is most simply dealt with as in the general case of modules, of vectors in V, all $\neq \vec{0}$, shall be called J- *indepenent* if the sum $\sum_{i=1}^k R\vec{v}_i$ is direct, J- *generating* if $V = \sum_{i=1}^k R\vec{v}_i$, and a J- *basis* if it is both. Given $\vec{0} \neq \vec{v} \in V$ its J- *chain* is the list

$$J(\vec{v}) : \ \vec{v}, \phi^1\vec{v}, \ldots, \phi^l\vec{v} \neq \vec{0}, \quad \text{where } \phi^{l+1}\vec{v} = \vec{0}$$

- $\vec{v}_1, \ldots, \vec{v}_k$ is J-independent (J-generating for V) if and only if the concatenated list $J(\vec{v}_1), \ldots J(\vec{v}_k)$ is K-independent (K-generating for V).

Here, we refer to independence and generators in the K-vector space V. Indeed

$$\sum_{i=1}^{v} p_i(\phi)\vec{v}_i = \sum_{i=1}^{v}\sum_{j=1}^{k_i} r_{ij}\phi^j\vec{v}_i \quad \text{where } p_i(x) = \sum_{j=0}^{k_i} r_{ij}x^j$$

$\vec{0}$

V_1 V_2 V_3 V_4 V_5 V_6 V_7

Lemma 6.3.2 *Let $\phi^h\vec{v}_1, \ldots, \phi^h\vec{v}_k$ be K-independent. Then*

(i) $\vec{v}_1, \ldots, \vec{v}_k$ are K-independent, the sum $W = \mathsf{span}_K\{\vec{v}_1, \ldots, \vec{v}_k\} + \mathsf{ker}\,\phi^h$ is direct, and $\dim W = k + \dim \mathsf{ker}\,\phi^h$

(ii) T.f.e.a.

- *$\phi^{h+1}\vec{v}_i = \vec{0}$ for all i*
- *$\phi\vec{v}_1, \ldots, \phi\vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_l$ is a J-basis of $\mathsf{ker}\,\phi^h$*
- *$\vec{v}_1, \ldots, \vec{v}_l$ is a J-basis of W*

Proof. If $\vec{w} \in \mathsf{ker}\,\phi^h$ and $\sum_i r_i\vec{v}_i + \vec{w} = \vec{0}$ then $\sum_i r_i\phi^h\vec{v}_i + \vec{0} = \vec{0}$ whence r_i for all i and $\vec{w} = \vec{0}$. □

6.3.3 Uniqueness of Jordan canonical form for nilpotent maps

Theorem 6.3.3 *For a nilpotent endomorphism, the Jordan canonical form is unique up to permutation of blocks.*

Proof. The Jordan matrix is determined by the number of J-chains of given lengths. Ordering by decraesing length $k_1 \geq k_2 \ldots$ we claim that these numbers are obtained as follows (where ϕ is given by A)

$$|\{i \mid k_i > k\}| = \dim \ker \phi^{k+1} - \dim \ker \phi^k = \operatorname{rank} A^k - \operatorname{rank} A^{k+1}$$

This is shown by induction (compare the general primary case) applied to $U = \ker \phi^{k_1-1}$. Let $k_1 = k_t > k_{t+1}$. By the Lemma,

$$\phi\vec{v}_1, \ldots, \phi\vec{v}_t, \vec{v}_{t+1}, \ldots, \vec{v}_s$$

is a J-basis of U and so the $k_1 - 1, \ldots, k_t - 1, k_{t+1}, \ldots, k_s$ are obtained form the data for U - which are part of that for V. Finally, $t = \dim V - \dim U$. $\square$.

6.3.4 Existence and computation of J-bases for nilpotent maps

Theorem 6.3.4 *Let m be minimal with $\phi^{m+1} = 0$ and $\phi^m\vec{v}_1, \ldots, \phi^m\vec{v}_j$ be K independent. Then there is a J-basis $\vec{v}_1, \ldots, \vec{v}_j, \ldots \vec{v}_l$ of V. It can be computed iterating the following two steps*

- Preparation*: Determine $\vec{v}_1, \ldots, \vec{v}_j, \ldots \vec{v}_k$ with $\phi^m\vec{v}_1, \ldots, \phi^m\vec{v}_k$ a K-basis of $\operatorname{im} \phi^m$*
- Recursion*: Determine a J-basis $\phi\vec{v}_1, \ldots, \phi\vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_l$ of $\ker \phi^m$*

Proof. The preparation step can be carried out, obviously. For the recursion step apply inductive hypothesis (w.r.t. m) to the ϕ-invariant subspace

$$U = \ker \phi^m \text{ and } \phi\vec{v}_1, \ldots, \phi\vec{v}_k$$

According to (ii) of the Lemma, $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots \vec{v}_l$ is a J-basis of W. But by the dimension formula for the endomorphism ϕ^m we have $\dim V = \dim U + \dim \operatorname{im} \phi^m = \dim U + k$ and with (i) of the lemma it follows $W = V$. $\square$ The following observations are of use in the computation. Given

$$V = \operatorname{span}_R(\{\vec{v}_1, \ldots, \vec{v}_j\} \cup X)$$

a. $\vec{v}_{j+1}, \ldots, \vec{v}_k$ may be choosen from X

b.
$$\ker \phi^m = \operatorname{span}_R(\{\phi\vec{v}_1, \ldots, \phi\vec{v}_k\} \cup \{\vec{x}' \mid \vec{x} \in X\})$$
where for $\vec{x} \in X$
$$\vec{x}' = \vec{x} - \sum_{i=1}^{k} r_i\vec{v}_i \quad \text{with} \quad \phi^m\vec{x} = \sum_{i=1}^{k} r_i\phi^m\vec{v}_i \text{ if } \phi^m\vec{x} \neq \vec{0}, , \quad \vec{x}' = \vec{x} \text{ else}$$

c. Start with $j = 0$ and X any K-basis of V.

d. If ϕ is given by A w.r.t. the basis $\vec{e}_1, \ldots, \vec{e}_n$ choose the $\vec{v}_1, \ldots, \vec{k}$ as a maximal subset such that the corresponding colums of A^m are independent

e. From the ranks of the powers of A one can determine the Jordan canonical form, i.e. the structure of a Jordan basis. This then can be used to verify in each step that the proper number of vectors needed to build the basis has been found.

Proof. For (a) observe that

$$\mathsf{im}\,\phi^m = \mathsf{span}_K\{\phi^m\vec{w}_i \mid i \in I\} \quad \text{if } V = \mathsf{span}_R\{\vec{w}_i \mid i \in I\}$$

Indeed, if $\vec{v} = \sum_i p_i(\phi)\vec{w}_i = \sum_i \sum_j r_{ij}\phi^j\vec{w}_i$ then $\phi^m\vec{v} = \sum_i r_{i0}\phi^m\vec{w}_i$.

Concerning (b) observe that $\phi^m\vec{x}' = \phi^m\vec{x} - \sum_i r_i\phi^m\vec{v}_i = \vec{0}$ and, by defintion of $\vec{x}'$

$$V = \mathsf{span}_R(\{\vec{v}_1, \ldots, \vec{v}_k\} \cup X') = \mathsf{span}_K\{\vec{v}_1, \ldots, \vec{v}_k\} + U'$$

where

$$U' = \mathsf{span}_R(\{\phi\vec{v}_1, \ldots, \phi\vec{v}_k\} \cup X') \subseteq U = \mathsf{ker}\,\phi^m$$

By (i) of the lemma, this sum is direct, whence $\dim U' = \dim V - k = \dim U$ and $U = U'$. □

The unique eigenvector in the J-chain of $\vec{v}$ is $\sigma\vec{v} = \phi^l\vec{v}$ where l is maximal with $\phi^l\vec{v} \neq \vec{0}$.

Corollary 6.3.5 *If the $\sigma\vec{v}_1, \ldots, \sigma\vec{v}_k$ are K-independent, then the $\vec{v}_1, \ldots, \vec{v}_k$ are J-independent.*

Proof. Choose h minimal with $\phi^{h+1}\vec{v}_i = \vec{0}$ for all i. Define

$$\vec{w}_i = \begin{cases} \phi\vec{v}_i & \text{if } \phi^h\vec{v}_i \neq \vec{0} \\ \vec{v}_i & \text{else} \end{cases}$$

Applying inductive hypothesis to $\mathsf{ker}\,\phi^h$, the $\vec{w}_1, \ldots \vec{w}_k$ are J-independent. By (i) of the lemma, $J(\vec{v}_1), \ldots, J(\vec{v}_k)$ is a basis of W as defined, there, whence a J-basis of W. □.

6.3.5 Example

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$A^3 = O$, $\mathsf{rank}(A^2) = 2$. $A^2\boldsymbol{e}_5 = 2\boldsymbol{e}_1$ and $A^2\boldsymbol{e}_6 = 2\boldsymbol{e}_2$ are independent, thus $\boldsymbol{e}_5$ and $\boldsymbol{e}_6$ are suitable heads. The associated Jordan-chains contain $A\boldsymbol{e}_5 = \boldsymbol{e}_3 + \boldsymbol{e}_4$ and $A\boldsymbol{e}_6 = \boldsymbol{e}_3 - \boldsymbol{e}_4$, hence their span U has basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_6$. Basis completion with vectors from $\mathsf{ker}A^2$ e.g. $\boldsymbol{e}_7, \boldsymbol{v} = 2\boldsymbol{e}_8 - 3\boldsymbol{e}_5$. Now $A\boldsymbol{e}_7 = \boldsymbol{e}_1 + 2\boldsymbol{e}_2 \in U$ but $A\boldsymbol{v} = -3\boldsymbol{e}_2 + \boldsymbol{e}_4 + 2\boldsymbol{e}_7 \notin U$, whence $\boldsymbol{v}$ is the wanted head.

6.3.6 Uniqueness of Jordan canonical form

Theorem 6.3.6 *Given $A \in K^{n\times n}$, there exists invertible $S \in K^{n\times n}$ such that $J = S^{-1}AS$ is block-diagonal with Jordan blocks if and only if the minimal or the characteristic polynomial of A is a product of linear factors $x - \lambda_i$ in $K[x]$. J is unique up to the order of blocks.*

Proof. Existence: Decompose the $K[x]$-module V given by A (resp. ϕ) into primary components V_{λ_i} , then generalized eigenspaces of the λ_i. For each, construct a Jordan basis for the nilpotent $(\phi - \lambda_i \mathsf{id})|V_{\lambda_i}$. The concatenation of these yields a Jordan basis of V and the colums of S. Conversely, $\det(A - xE)$ is a product of linear factors, claerly.

Uniqueness follows from uniqueness of the generalized eigenspaces and uniqueness in the nilpotent case. □

Corrigenda et addenda.

a. In the proof of Thm.23.1.4 read: $K[x]\vec{f_i}$ in place of $K[x[d_i(x)\vec{f_i}$ and recall that $d_i(x)$ is the minimal polynomial of the module $K[x]\vec{f_i} \cong K[x]/(d_i(x))$

b. The generalized eigenspace V_λ w.r.t. λ is given as

$$\mathsf{ker}((\phi - \lambda\mathsf{id})^k)$$

k mimimal such that

$$\dim \mathsf{im}(\phi - \lambda\mathsf{id})^k) = \dim \mathsf{im}(\phi - \lambda\mathsf{id})^{k+1})$$

$$\text{i.e. } \mathsf{rank}((A - \lambda E)^k) = \mathsf{rank}((A - \lambda E)^{k+1})$$

and then

$$\dim \mathsf{im}(\phi - \lambda\mathsf{id})^k) = \dim \mathsf{im}(\phi - \lambda\mathsf{id})^k) \text{ for all } l \geq k$$

Indeed, $V = V_\lambda \oplus W$ with invariant subspace W such that $(\phi - \lambda\mathsf{id})|W$ is bijective and $\mathsf{ker}((\phi - \lambda\mathsf{id})^k \subseteq V_\lambda$ whence

$$\dim \mathsf{im}(\phi - \lambda\mathsf{id})^k = \dim \mathsf{im}((\phi - \lambda\mathsf{id})|V_\lambda)^k + \dim W$$

$$= \dim V_\lambda - \dim \mathsf{ker}((\phi - \lambda\mathsf{id})^k + \dim W$$

and this is $\dim W$ if and only $V_\lambda - \mathsf{ker}((\phi - \lambda\mathsf{id})^k$

6.4 Jordan-Chevalley decomposition

6.4.1 Existence

Theorem 6.4.1 *For $A \in \mathbb{C}^{n\times n}$ there are $H, N \in \mathbb{C}^{n\times n}$ such that*

$$A = H + N,\ HN = NH,\ \ H \textit{ diagonalizable }, N \textit{ nilpotent}$$

Proof. There is an invertible matrix S such that $S^{-1}AS = J$ is in Jordan form. Obviously, $J = H_0 + N_0$ where H_0 is diagonal and N_0 nilpotent. Moreover, $J_i = H_i + N_i$ in the block decomposition into Jordan-blocks with $H_i = \lambda_i E_{k_i}$ whence $H_i N_i = N_i H_i$, It follows $H_0 N_0 = N_0 H_0$. Now, put $H = SH_0S^{-1}$ and $N = SN_0S^{n-1}$. □.

Corollary 6.4.2 *If l is the maximal block size in the Jordan canonical form then*

$$A^m = \sum_{k=0}^{m} \binom{m}{k} N^k H^{m-k} = \sum_{k=0}^{\min\{m,l-1\}} \binom{m}{k} N^k H^{m-k}$$

Proof. Since $HN = NH$ we may apply the polynomial formula. But, $N^k = O$ for $k \geq l$. □.

6.4.2 Matrix exponential function

For each $A \in \mathbb{C}^{n\times n}$ there is a uniquely determined matrix $\exp(A)$ such that

$$\exp(A) = \lim_{h\to\infty} \sum_{m=0}^{h} \frac{1}{m!} A^m$$

To prove this, let ϕ be the endomorphism determined by A. Since the limit is to be understood column wise, the claim amounts to the existence of $\lim_{h\to\infty} \sum_{m-0}^{h} \frac{1}{m!}\phi^h(\boldsymbol{x}) = \exp(\phi)(\boldsymbol{x})$ for the canonical basis vectors, i.e. for all vectors in $\mathbb{C}^n$. .Thus, we may assume $A = H + N$ in Jordan canonical form. It follows according to the corollary, computing with series, formally,

$$\exp(A) = \sum_{m=0}^{\infty} \frac{1}{m!} A^m = \sum_{m=0}^{\infty}\sum_{k=0}^{m} \binom{m}{k} \frac{1}{m!} N^k H^{m-k} = \sum_{m=0}^{\infty}\sum_{k=0}^{m} \frac{1}{k!(m-k)!} N^k H^{m-k}$$

$$= (\sum_{k=0}^{l-1} \frac{1}{k!} N^k) \cdot (\sum_{j=0}^{\infty} \frac{1}{j!} H^j) = \exp(N)\exp(H) = \exp(H)\exp(N)$$

since $HN = NH$. But, if H is diagonal with diagonal entries λ_i then H^k has diagoal entries λ_i^k whence $\exp(H)$ erists and

$$\exp(H) = \begin{pmatrix} e^{\lambda_1} & 0 & \dots \\ 0 & e^{\lambda_2} & \dots \\ 0 & 0 & \ddots \end{pmatrix}$$

This meas that for any $\varepsilon > 0$ there is an h_0 such that for all $k \geq h \geq h_0$ one has $|\sum_{m=h}^{k} \frac{1}{m!} H^m| < \varepsilon$ which readily transfers to prove existence of $\exp(A)$.

Now, consider the vector valued function

$$\boldsymbol{y}(t) = \exp(At)\boldsymbol{y}_0 \ (t \in \mathbb{R}) \text{ with fixed } \boldsymbol{y}_0 \in \mathbb{C}^n$$

We claim that one has derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{y}(t) = A\boldsymbol{y}(t)$$

i.e. that $\boldsymbol{y}(t)$ is a solution of the system of first order linear differential equations with constant coefficients given by A. Again, this claim is invariant under basis transformation,

hence we may assume A in Jordan canonical form and even consisting of a single Jordan-block of size l and EW λ, i.e $A = \lambda E + N$. In that case

$$\exp(At) = exp(\lambda Et + Nt) = \exp(\lambda Et)\exp(Nt) = \sum_{k=0}^{l-1} \frac{e^{\lambda t}}{k!} t^k N^k$$

Differentiating entry-wise we get

$$\sum_{k=}^{l-1} \frac{1}{k!}(\lambda e^{\lambda t}t^k + e^{\lambda t}kt^{k-1})N^k = (\lambda E)\sum_{k=}^{l-1} \frac{1}{k!}(\lambda e^{\lambda t}t^k)N^k + N(\sum_{k=}^{l-1} \frac{1}{(k-1)!}e^{\lambda t}t^{k-1}N^{k-1}) = A\exp(A)$$

6.4.3 Uniqueness of Jordan-Chevalley decomposition

Theorem 6.4.3 *Given $A \in K^{n\times n}$ the minimal polynomial of which is a product of linear factors in $K[x]$, there are unique diagonalizable H and nilpotent N such that $A = H + N$ and $HN = NH$. Moreover, $H, N \in K[A]$.*

Lemma 6.4.4 *Let V be a finite dimensional vector space with endomorpism ϕ and minimal polynomial $d(x) = d_1(x)d_2(x)$ with coprime $d_1(x), d_2(x)$. Then $V = V_1 \oplus V_2$ where $V_i = \{v \in V \mid d_i(\phi)v = 0\}$ and $\pi \in K[\phi]$ for the projections $\pi : V \to V_i$.*

Proof. $V = V_1 \oplus V_2$ by Lemma 5.6.7. Also $1 = r_1(x)d_1(x) + r_2(x)d_2(x)$ and $(r_j d_j)(\phi)$ is identity on V_i and 0 on V_j. Thus, $\pi = (r_j d_j)(\phi) \in K[\phi]$. □

Lemma 6.4.5 *Let R be a K-algebra and $\alpha_1, \ldots, \alpha_m \in R$ such that $\alpha_i\alpha_j = \alpha_j\alpha_i$ for all i, j. Then the smallest K-subalgebra containing all α_i is given as*

$$\{\sum_{i_1,\ldots,i_m} a_{i_1,\ldots,i_m}\alpha^{i_1} \cdot \ldots \cdot \alpha^{i_m} \mid i_i \in \mathbb{N}.\ a_{i_1,\ldots,i_m} \in K\}$$

and is, in particular, commutative.

This is then the K-subalgebra *generated by the commuting $A_1, \ldots, A_m$*. Proof. Straightforeward computation. □

Lemma 6.4.6 *If the matrices $N_1, \ldots, N_k \in K^{n\times n}$ are nilpotent and $N_iN_j = N_jN_i$ for all i, j then each A in the K-algebra generated by the A_i is nilpotent.*

Proof follows from the exercise: sum and product of two commuting nilpotent matrices is nilpotent. □

Proof of the Thm. We first show that there are H, N in $K[A]$ with $A = H + N$, H diagonalizable, N nilpotent. Then also $HN = NH$. The statement can also be formulated for endomorphism - and is basis invariant. thus. So we may assume that $A = J$ is in Jordan canonical form and we have the obvious decomposition $J = H + N$. We claim that $H \in K[A]$ - then also $N = A - H$ in $K[A]$ and so $HN = NH$. But the endomorphism defined by H is $\sum_i \lambda_i\pi_i$ where the π_i are the projections associated with decomposition

into generalized eigenspaces, so $\pi \in K[\phi]$ by iterated application of Lemma 6.4.4 and so $H \in K[A]$ whence $N = A - H \in K[A]$.

Now, given H, N in $K[A]$ with $A = H + N$, H diagonalizable, N nilpotent, consider $A = H' + N'$ with diagonalizable H', nilpotent N', and $H'N' = N'H'$. It follows $H'A = H'(H' + N') = H'H' + H'N' = H'H' + N'H' = AH'$. Similarly, $N'A = AN'$. Since $H, N \in K[A]$, it follows from Lemma 6.4.5 that $HH' = H'H$ and $NN' = N'N$.

By Thm.?? we have H and H' simultaneously diagonalizable and so $H - H'$ diagonalizable. On the other hand, by Lemma 6.4.6, $N' - N$ is nilpotent. From $A = H + N = H' + N'$ it follows $H - H' = N' - N$ which is a matrix which is both diagonalizable and nilpotent. So it has to be O, since a nilpotent diagonal matrix is O. □

6.5 Rational canonical form

In this section, R denotes an euclidean ring.

6.5.1 Structure theorem

Theorem 6.5.1 *Given an R-module M on n generators $\vec{e}_1, \ldots, \vec{e}_n$ over an euclidean ring R and canonical homomorphism $\pi : F \to M$, $\pi(e_i) = \vec{e}_i$, where $\alpha : e_1, \ldots, e_n$ is a basis of a free R-module F. Then*

(i) M is isomorphic to a direct product of cyclic modules

(ii) Given any presentation matrix $\mathcal{A}$ there are invertible $\mathcal{P}$ and $\mathcal{Q}$ such that $\mathcal{PAQ}$ is diagonal with diagonal entries d_i (with $d_i \mid d_{i+1}$)

(iii) Given matrices as in (ii) and $\mathcal{A}$ w.r.t. the basis α there is a basis $\beta : f_1, \ldots, f_n$ of F such that $\mathcal{P}^{-1}$ gives the α-coordinates of the f_i and

$$M = R\pi(f_1) \oplus \ldots \oplus R\pi(f_n) \quad \text{with } R\pi(f_i) \cong R/Rd_i$$

(iv) Requiring $d_i \mid d_{i+1}$, the $d_i \not\approx 1$ are unique up to association (and called the invariant divisors *of M) resp. the nonzero summands in (iii) are unique up to isomorphism.*

Proof. Given a system a_j $(j \in J)$ of generators of U, let $\mathcal{A}$ be a matrix with colums $(a_j)^\alpha$. By Thm.5.5.1 on invariant divisors there are invertible $\mathcal{P}$ and $\mathcal{Q}$ such that $\mathcal{D} = \mathcal{PAQ}$ is diagonal with diagonal entries $d_i \mid d_{i+1}$ - this also applies if $\mathcal{A}$ has infinitely many colums, since the column operations in a step reducing the format of the matrix may be carried out, simultaneously. Now, Cor.4.3.3 and sect. 21.2.5. point 5 apply. Uniqueness of the invariant divisors follows from that of the elementary divisors - see below). □

Corollary 6.5.2 *For any submodule of a free R-module F on n-free generators there is a basis $f_1, \ldots, f_n$ von F, an $r \leq n$ and $d_i \in R$ with $d_i | d_{i+1}$, $i < r$ such that $d_1 f_1, \ldots, d_r f_r$ is a basis of U.*

6.5.2 Torsion free part

Given an R-module M, its *torsion submodule* is

$$T(M) = \{v \in M \mid rv = 0 \text{ for some } r \neq 0\}$$

M is *torsion free* if $T(M) = 0$.

Theorem 6.5.3 *Let R be an euclidean ring and M an R-module with n generators. Then $M = T(M) \oplus U$ with some U freely generated by m elements for some uniquely determined $0 \leq m \leq n$*

Proof. In Thm.6.5.1 let $d_i \mid d_{i+1}$ and $d_i \neq 0 \Leftrightarrow i \leq s$. Then we have $M = W \oplus U$ where $W = \mathsf{Span}_R\{\pi(f_i) \mid i \leq s\}$ and $U = \mathsf{Span}_R\{\pi(f_i) \mid i > s\}$. In particular, U has basis $\pi(f_{s+1}), \ldots, \pi(f_n)$ and $d_s w = 0$ for all $w \in W$. Thus $U \cong R^m$ with $m = n - s - 1$. If $w \in W$ and $0 \neq u \in U$ then $ru \neq 0$ whence $r(w + u) \neq 0$ for all $r \neq 0$. Thus, $T(M) = W$ and $R^m \cong U \cong M/W$.

Now, it suffices to show that $R^m \cong R^k$ implies $m = k$. Assume $k \leq m$, let Q be the quotient field of R, and consider R^m as a subset (and R-submodule) of Q^m. Then the canonical basis of R^m is a basis of the Q-vector space Q^m. The canonical basis of R^k corresponds under the isomorphism to a k-element generating set of R^m. But this is then also a generating set of the Q-vector space Q^m, whence $k = m$. □

6.5.3 Structure of primary modules

Theorem 6.5.4 *Let R be an euclidean ring and M a finitely generated p-primary R-module. Then*

(1) $M = Rv_1 \oplus \ldots \oplus Rv_s$ *with* $Rv_i \cong R/(p_i^k)$ *and* $k_1 \geq \ldots \geq k_s > 0$

(2) In (1), s and the p_i^k are uniquely determined by M and called the elementary divisors *of M*

(3) $\phi_p(v) = pv$ *is an R-linear map* $\phi_p : M \to M$ *and k_1 the minimal k with* $\ker \phi_p^k = M$ *resp.* $\mathsf{im}\, \phi_p^k = 0$

(4) $(\ker \phi_p^{k+1})/(\ker \phi_p^k)$ *is canonically a $R/(p)$-vector space. M is determined up to isomorphism by the dimensions of these spaces for* $0 \leq k < k_1$

$$|\{i \mid k_i > k\}| = \dim\, (\ker \phi_p^{k+1})/(\ker \phi_p^k)$$

(5) $(\mathsf{im}\, \phi_p^k)/(\mathsf{im}\, \phi_p^{k+1})$ *is canonically a $R/(p)$-vector space. M is determined up to isomorphism by the dimensions of these spaces for* $0 \leq k < k_1$.

(6) $\mathsf{im}\, \phi_p^{k_1 - 1}$ *is uniquely determined,* $\neq 0$ *but* $\phi_p(\mathsf{im}\, \phi_p^{k_1-1}) = 0$.

In the case of the Jordan canonical form of an endomorphism ϕ with unique eigenvalue λ, we have $p = x - \lambda$, i.e. $\phi_p = \phi - \lambda\mathsf{id}$. $\mathsf{im}\,\phi_p^{k_1-1}$ is then a subspace of the eigenspace of ϕ and any Jordan basis has to contain a subset which is a basis of this subspace.

Proof. Given p-primary M, the minimal annhilator is a power of p. Thus, (i) and (iii) are obvious from the Theorem on primary decomposition.

Now, let $U = \mathsf{ker}\,\phi_p^{k+1}$ and $W = \mathsf{ker}\,\phi_p^k$. Claerly $W \subseteq U$. Given $u \in U$ and $r \in R$ we define

$$(r + (p))(u + W) = ru + W$$

This is well defined: if $r + (p) = r' + (p)$ then $r - r' = sp$ a whence $(r - r')u = spu \in W$ and so

$$ru + W = r'u + W$$

On the other hand, if $u - u' \in W$ then $r'u - r'u' = r'(u - u') \in W$ since W is a submodule. Thus

$$ru + W = r'u + W = r'u' + W$$

The module laws are inherited, obviuosly. Thus U/W is a vector space over the field $R/(p)$. Observe that

$$\phi_p^{k_1-1}(r_1v_1 + \ldots + r_sv_s) = 0 \;\Leftrightarrow\; p^{k_1-1}r_iv_i = 0 \;\text{ for all i}$$

and the latter holds a priori for all $i > m$ and for $i \leq m$ if and only if $p \mid r_i$. Thus

$$N := \mathsf{ker}\,\phi_p^{k_1-1} = Rpv_1 \oplus \ldots \oplus Rpv_m \oplus Rv_{m+1} \oplus \ldots \oplus Rv_s$$

and N is p-primary submodule of M with elementary divisors determined by those of M as the p^{k_i-1} with $k_i > 1$. Assuming uniqueness as inductive hypothesis (proceeding by induction on k_1) uniqueness for M follows provided we have the number m of the $k_i = k_1$.

$$M/N \cong Rv_1/Rpv_1 \oplus \ldots \oplus Rv_m/Rpv_m \cong (R/(p))^m$$

so m is the dimension of the $R/(p)$-vector space M/N. This proves (2) and (4) follows by induction, too. The proof of (5) is similar, (6) is obvious. □

6.5.4 Uniqueness of elementary and invariant divisors of a matrix

Corollary 6.5.5 *The invariant divisors as well as the elementary divisors of a matrix over a euclidean ring are unique up to association and order.*

Proof. Consider $\mathcal{A}$ a presentation matrix of an R-module M. The number of elementary or invariant divisors $d_i = 0$ is the size of a basis of $M/T(M)$. The elementary divisors p^k are determined up to association by the p-primary components of $T(M)$. From these we combine the invariant divisors $d_i \not\approx 0, 1$ beginning with the highest powers. Having these, the number of invariant (and elementary) divisors $d_i \approx 1$ just has to fill up to the number of rows of $\mathcal{A}$. □

6.5.5 Similar matrices

Theorem 6.5.6 *For $n \times n$-matrices A and A' over a field K t.f.a.e.*

(1) A and A' are similar*, i.e. there is an invertible matrix S over K such that $A' = S^{-1}AS$*

(2) $A - xE$ and $A' - xE$ are equivalent*, i.e. there are invertible matrices P and Q over $K[x]$ such that $A' - xE = P(A - xE)Q$*

(3) The $K[x]$-modules ${}_{K[A]}K^n$ and ${}_{K[A']}K^n$ defined by A resp. A' are isomorphic

(4) A and A' (i.e. $A - xE$ and $A' - xE$)) have the 'same' invariant divisors

(5) A and A' (i.e. $A - xE$ and $A' - xE$) have the 'same' elementary divisors

(6) A and A' (i.e. $A - xE$ and $A' - xE$) have the 'same' determinantal divisors

Proof. $1 \Rightarrow 2$: $S^{-1}(A - xE)S = S^{-1}AS - xE = A' - xE$. $2 \Rightarrow 3$: by Cor.4.3.3. $3 \Rightarrow 1$: The module isomorphism $\sigma : {}_{K[A']}K^n \to {}_{K[A]}K^n$ and the matrix S are related by

$$\sigma(\boldsymbol{v}) = S\boldsymbol{v}.$$

Given σ is bijective and K-linear and one can find S. By $K[x]$-linearity, for all $\boldsymbol{v}$

$$AS\boldsymbol{v} = A\sigma\boldsymbol{v} = x\sigma\boldsymbol{v} = \sigma(x\boldsymbol{v}) = SA'\boldsymbol{v}.$$

$3 \Rightarrow 2$ can be shown, directly: Given S, one obtains a module isomorphsim

$$\omega : {}_{K[A]}K^n \to {}_{K[A']}K^n, \ \ \omega\boldsymbol{v} = S^{-1}\boldsymbol{v}$$

Indeed, for all $\boldsymbol{v}$ and $f(x) = \sum_k r_k x^k$

$$f(x)\omega(\boldsymbol{v}) = \sum_k r_k A'^k \omega(\boldsymbol{v}) = \sum_k r_k (S^{-1}AS)^k S^{-1}\boldsymbol{v} = \sum_k r_k S^{-1}A^k\boldsymbol{v} =$$

$$= S^{-1}\sum_k r_k A^k\boldsymbol{v} = \omega(f(x)\boldsymbol{v}).$$

(3) is equivalent to (4) resp. (5) by existence and uniqueness of divisors. The k-th *determinantal divisor* is defined as the normed GCD of all determinants of $k \times k$-minors of $A - xE$. This is unchanged under transformation. In the diagonal matrix having the invariant divisors on the diagonal, the k-th determinantal divisor is the product of the fisst k invariant divisors. Hence, these determine each other. □

Corollary 6.5.7 *For any $A \in K^{n\times n}$ there is invertible $S \in K^{n\times n}$ such that $S^{-1}AS = A^t$.*

Contents