Symbolic Linear Algebra

—Special Lecture—

J. Middeke

Research Institute for Symbolic Computation

Summer Term 2017

Contents

Part I
Monoids, Groups, Rings, Fields, and Modules

1 Monoids and Groups

Definition 1 (Monoid). A *monoid* $(M, \diamond, \varepsilon)$ is a set M together with an operation $\diamond\colon M \times M \to M$ such that the following properties hold:

Associativity For every $a, b, c \in M$ we have $(a \diamond b) \diamond c = a \diamond (b \diamond c)$.

Neutral Element There is an $\varepsilon \in M$ such that $\varepsilon \diamond a = a$ and $a \diamond \varepsilon = a$ for all $a \in M$.

The element ε is called the *neutral element* of the monoid.

A monoid $(M, \diamond, \varepsilon)$ is called *Abelian* (or commutative) if additionally

Commutativity for all $a, b \in M$ we have $a \diamond b = b \diamond a$.

Example 2 (Monoids and Non-Monoids). Examples of monoids include:

(a) The natural numbers $\mathbb{N}$ with the usual addition. The neutral element is 0. Moreover, this monoid is Abelian.

(b) The natural numbers with the usual multiplication. The neutral element is now 1; the monoid is again Abelian.

(c) The natural numbers with the maximum max as operation. The neutral element is 0; the monoid is Abelian.

(d) The integers $\mathbb{Z}$ with the greatest common divisor gcd as operation. The neutral element is zero because $\gcd(0, a) = a$ for all $a \in \mathbb{Z}$ by definition (see Definition 125). Again, this is an Abelian monoid.

(e) The square n-by-n matrices ${}^{n}\mathbb{Q}^{n}$ with rational entries and multiplication. The neutral element is the identity matrix $\mathbf{1}_n$. This monoid is *not* Abelian for $n \geqslant 2$.

The following are *not* examples of monoids:

(a) The natural numbers with the minimum min as operation. Here, we do not have a neutral element.

(b) The integers with subtraction. In this case, the operation is not associative since, for example, $1 - (2 - 3) = 2 \neq -4 = (1 - 2) - 3$.

(c) The natural numbers with exponentiation, that is, $a \diamond b = a^b$. This operation is not associative since, for example, $2^{(3^2)} = 512 \neq 64 = (2^3)^2$.

Notation 3. Often the operation of an (abstract) monoid is simply denoted by the multiplication symbol $\cdot$ or by juxtaposition[1] and the neutral element is denoted by 1. In case of Abelian monoids, the operation is traditionally denoted by $+$ and the neutral element by 0. Consequently, we often just speak about the monoid M instead of the more correct $(M, \cdot, 1)$ (or $(M, +, 0)$ in the Abelian case). Moreover, since because of associativity products (or sums) can be computed in any order, we usually leave out parentheses and just write abc instead of $(ab)c$ (or $a(bc)$).

Exercise 4. Show that a monoid has only one neutral element.

Notation 5. Let M be a monoid. For $a \in M$ and $n \in \mathbb{N}$ we define

$$a^n = \underbrace{a \cdot a \cdots a}_{n \text{ times}}$$

if $n \geqslant 1$ and $a^0 = 1$. (Note that because of $a^n = aa^{n-1}$ for $n \geqslant 1$ we could also define the powers recursively.) If the we use additive notation for M, then we write

$$na = \underbrace{a + a + \ldots + a}_{n \text{ times}}$$

for $n \geqslant 1$ and $0a = 0$.[2]

Exercise 6. Let M be a monoid such that $a^2 = 1$ for all $a \in M$. Show that M is Abelian.

Definition 7 (Group). A monoid $(G, \cdot, 1)$ is a *group* if it fulfills the additional property that

Inverses for all $a \in G$ there exists an element $a^{-1} \in G$ such that $aa^{-1} = 1$ and $a^{-1}a = 1$.

We call a^{-1} the *inverse* of a.

Notation 8. For an Abelian group G the inverse of $a \in G$ is usually denoted by $-a$. Moreover, we normally write $a - b$ for $a + (-b)$.

Exercise 9. Show that the inverse of a group element is uniquely determined.

Example 10 (Groups and Non-Groups). Examples of groups are

(a) The integers $\mathbb{Z}$ with the usual addition. This is an Abelian group.

(b) The invertible n-by-n matrices $\mathrm{GL}_n(\mathbb{Q})$ with rational coefficients and the usual multiplication. This group is *not* Abelian for $n \geqslant 2$.

(c) The permutations $S_M = \{f \colon M \to M \mid f \text{ is bijective}\}$ of a non-empty set M together with the composition of functions $\circ$. This group is also *not* Abelian.

On the other hand, *none* of the monoids in Example 2 is a group.

Notation 11. For a group G we can define negative powers as $a^{-n} = (a^{-1})^n$ for $a \in G$ and $n \in \mathbb{N}$ (or in additive notation $(-n)a = n(-a)$).

[1] That is, writing just ab for the product $a \cdot b$.

[2] Note that in $0a = 0$ the first 0 is in $\mathbb{N}$ while the second 0 is the neutral element of M.

2 Rings and Fields

Definition 12 (Ring). A (unitary, commutative) *ring* $(R, +, 0, \cdot, 1)$ is a set R together with two operations $+\colon R \times R \to R$ and $\cdot\colon R \times R \to R$ such that

(a) $(R, +, 0)$ is an Abelian group,

(b) $(R, \cdot, 1)$ is an Abelian monoid,

and such that

Distributivity for all $a, b, c \in R$ we have $a \cdot (b + c) = a \cdot b + a \cdot c$.

Notation 13. In Definition 12 we made use of the convention that $\cdot$ binds more strongly than $+$. That is, we always read $a \cdot b + c$ as $(a \cdot b) + c$. Also, we will often just speak about a ring R and not name the operations explicitly. If we talk about several structures at the same time, we will sometimes write 0_R and 1_R in order to emphasise from which ring the (additive and multiplicative) neutral elements originate.

Example 14 (Rings and Non-Rings). The following sets are examples of rings

(a) The integers $\mathbb{Z}$ with the usual addition and multiplication.

(b) The square n-by-n matrices ${}^nR^n$ with entries from any ring R and with the usual addition and multiplication. This ring is *not* commutative (that is, $({}^nR^n, \cdot, \mathbf{1}_n)$ is not Abelian).

(c) The set of polynomials $R[X]$ over any ring R with the usual addition and multiplication.

(d) For any set S, we can make its power set $\mathfrak{P}(S)$ into a ring by setting

$$A + B := (A \cup B) \setminus (A \cap B) \qquad \text{and} \qquad A \cdot B = A \cap B$$

for $A, B \subseteq S$.

The following sets are *not* examples of rings

(a) The natural numbers $\mathbb{N}$ with the normal addition and multiplication.

Exercise 15. Prove that item (d) of Example 14 is indeed a ring.

Exercise 16. Prove that for any ring R we have

(a) $0a = 0$ for all $a \in R$;[3]

(b) $(-a)b = a(-b) = -(ab)$ for all $a, b \in R$; and

(c) $(-a)(-b) = ab$ for all $a, b \in R$.

Remark 17 (Zero Ring). If in a ring R we have $0 = 1$, then with Exercise 16 we obtain that $R = \{0\}$. Since this ring is not terribly interesting, we will from now on always assume that $0 \neq 1$.

Definition 18 (Units). If R is a ring, then the invertible elements in the monoid $(R, \cdot, 1)$ are called *units*. We denote the set of all units in R by R^*.

[3] Here we mean $0_R a = 0_R$ in contrast to the notation for exponents in additive notation introduced in Notation 5.

Remark 19. In general, we have $R^* \neq R \setminus \{0\}$. For instance, in the ring of integers the only units are 1 and -1; that is, $\mathbb{Z}^* = \{-1, 1\}$.

Definition 20 (Zero Divisors, Regular, Integral Domain). Let R be a ring. A *zero divisor* is an element $a \in R$ such that there exists $b \in R \setminus \{0\}$ with $ab = 0$.

An element $a \in R$ which is not a zero divisor is called *regular*.

A ring which does not have any zero divisors except for 0 is called an *integral domain*.

Exercise 21 (Cancellation Rule). Let R be an integral domain and let $a, b, c \in R$ such that $c \neq 0$. Show that $ac = bc$ or $ca = cb$ implies $a = b$.

Definition 22 (Field). A ring $(R, +, 0, \cdot, 1)$ is a *field* if

(a) $0 \neq 1$, and

(b) $(R \setminus \{0\}, \cdot, 1)$ is an Abelian group.

(That is, every non-zero element has a multiplicative inverse.)

Example 23 (Field and Non-Fields). The following are fields:

(a) The rational, real and complex numbers $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ with their usual addition and multiplication.

(b) The set $\mathbb{F}_2 = \{0, 1\}$ where addition and multiplication are given by the tables

$+$	0	1
0	0	1
1	1	0

and

$\cdot$	0	1
0	0	0 .
1	0	1

The following are *not* fields:

(a) The integers $\mathbb{Z}$.

Exercise 24. Prove that every field is an integral domain.

Exercise 25. Prove that every finite integral domain is a field.

Example 26 (Field of Fractions). If R is an integral domain, then we can form the *field of fractions* of R. The construction is exactly the same as for the rational numbers: Consider the set $S = R \times (R \setminus \{0\})$. We introduce an equivalence relation $\sim$ on S by setting $(a, b) \sim (x, y)$ if $ay = bx$. Let $Q(R) = S/\sim$ be the equivalence classes of $\sim$. We write the equivalence class of a pair $(a, b) \in S$ as fraction a/b. Addition and multiplication in $Q(R)$ are now defined as

$$\frac{a}{b} + \frac{x}{y} = \frac{ay + xb}{by} \qquad \text{and} \qquad \frac{a}{b} \cdot \frac{x}{y} = \frac{ax}{by}.$$

We can show (see Exercise 27) that this yields a field. Note that we can identify the original ring R with the subset $\{a/1 \mid a \in R\} \subseteq Q(R)$; we say that R is embedded in $Q(R)$. The field of fractions $Q(R)$ of R is the smallest field which contains R.

Exercise 27. Show that the operations in Example 26 are well-defined and that they do indeed make $Q(R)$ a field. What are the neutral elements? What are the inverses?

3 Modules and Ideals

Definition 28 (Module). Let $(R, +, 0, \cdot, 1)$ be a ring; and let $(M, +, 0)$ be an Abelian group. We call M a *(left) R-module* if there exists an action $\bullet\colon R \times M \to M$ such that

(a) $a \bullet (x + y) = (a \bullet x) + (a \bullet y)$ for all $a \in R$ and $x, y \in M$;

(b) $(a + b) \bullet x = (a \bullet x) + (b \bullet x)$ for all $a, b \in R$ and $x \in M$;

(c) $a \bullet (b \bullet x) = (ab) \bullet x$ for all $a, b \in R$ and $x \in M$;

(d) $1 \bullet x = x$ for all $x \in M$.

The action $\bullet$ is sometimes called the scalar multiplication of M by R.

Notation 29. Note that the $+$ in Definition 28 is used both for the operation of the ring R and the module M. We assume that $\bullet$ binds more strongly than $+$ but weaker than $\cdot$; that is, we would interprete $ab \bullet x$ as $(ab) \bullet x$ and $a \bullet x + b \bullet x$ as $(a \bullet x) + (b \bullet y)$. Sometimes we will omit the $\bullet$ and denote the scalar multiplication by juxtaposition.

Notation 30. We use ${}_R\mathrm{Mod}$ to denote the collection of all left R-modules. Thus, instead of saying that M is a left R-module, we will sometimes just write $M \in {}_R\mathrm{Mod}$. Some authors write ${}_RM$ to indicate that $M \in {}_R\mathrm{Mod}$; but we will not use that notation here.

Notation 31. In German, a module is called “der Modul” (with the stress on the first syllable), the plural is “die Moduln”.

Remark 32. Analogously, we can introduce *right modules* where the scalar multiplication is done from the right (that is, $M \times R \to M$) and the module laws are changed accordingly.

Example 33 (Vector Space). Every vector space over a field F is an F-module.

Example 34 (The Free Module). Let R be a ring. We consider the set

$$R^n = \underbrace{R \times \ldots \times R}_{n \text{ times}}$$

of all n-tuples over R. This becomes a module via

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

and

$$a \cdot (x_1, \ldots, x_n) = (ax_1, \ldots, ax_n)$$

for all $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in R^n$ and $a \in R$.

Exercise 35. Prove that Example 34 indeed yields a module.

Example 36 (Abelian Groups). Let G be an Abelian group. Then forming (additive) multiples

$$\bullet\colon \mathbb{Z} \times G \to G, \quad (n, a) \mapsto na$$

as in Notation 5 and Notation 11 makes G into a $\mathbb{Z}$-module.

Exercise 37. Prove that Example 36 is correct.

Example 38 (Linear Differential Operators). Consider the polynomials $R = \mathbb{R}[x]$ over the real numbers $\mathbb{R}$. We let

$$C^\infty(\mathbb{R}) = \{f \colon \mathbb{R} \to \mathbb{R} \mid f \text{ is differentiable infinitely often}\}$$

and define the action $\bullet\colon \mathbb{R}[x] \times C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ by

$$\left(\sum_{j=0}^{n} a_j x^j\right) \bullet f(t) = \sum_{j=0}^{n} a_j f^{(j)}(t).$$

We can easily check that this turns $C^\infty(\mathbb{R})$ into an $\mathbb{R}[x]$-module. This action makes $\mathbb{R}[x]$ an algebraic model for linear differential operators with constant coefficients.

Exercise 39. Verify that the definition in Example 38 indeed yields a module.

Notation 40. We will revisit the module $C^\infty(\mathbb{R})$ in other examples where we will then usually use the symbol ∂ as the indeterminate instead of x. Note that $\mathbb{R}[\partial]$ is still just the regular polynomial ring over $\mathbb{R}$ despite the funny symbol.

Exercise 41. Let M be an R-module. Prove that for all $x \in M$ and $a \in R$

(a) $0 \bullet x = 0$,[4]

(b) $a \bullet 0 = 0$, and

(c) $-1 \bullet x = -x$.

Definition 42 (Submodule). Let M be an R-module. A non-empty subset $N \subseteq M$ is called a *submodule* of M if for all $x, y \in N$ and all $a \in R$ we have

$$x + y \in N \qquad \text{and} \qquad a \bullet x \in N;$$

that is, N is an R-module in its own right. We will usually denote the fact that N is a submodule of M by writing $N \leqslant M$.

Remark 43. We want to show why in Definition 42 N is indeed an R-module: For this, we have to show that N is an Abelian group and that the scalar multiplication fulfills the properties of Definition 28. Looking at Definition 7, we see that the addition is associative and commutative on N since it is on the superset M. It remains to prove that $0 \in N$ and $-x \in N$ for every $x \in N$. Both follow from Exercise 41. The scalar multiplication must again have the desired properties since they hold for the larger set M.

Example 44 (Trivial Submodules). For every R-module M, both $\{0\}$ and M are submodules. Nontrivial submodules are called proper.

Example 45. Consider the $\mathbb{R}[\partial]$-module $C^\infty(\mathbb{R})$ (see Example 38). Define

$$N = \{f \in C^\infty(\mathbb{R}) \mid (\partial^2 + 1) \bullet f = 0\}.$$

Then N is a submodule of $C^\infty(\mathbb{R})$. This follows from the module laws: Let $f, g \in N$ and $a \in \mathbb{R}[\partial]$. Then

$$(\partial^2 + 1) \bullet (f + g) = (\partial^2 + 1) \bullet f + (\partial^2 + 1) \bullet g = 0$$

[4] That is, $0_R \bullet x = 0_M$.

and

$$(\partial^2+1)\bullet(a\bullet f) = \big((\partial^2+1)a\big)\bullet f = \big(a(\partial^2+1)\big)\bullet f = a\bullet\big((\partial^2+1)\bullet f\big) = a\bullet 0 = 0.$$

Thus, N is a submodule by Definition 42. We will see that this example is just a kernel of an R-linear map in Definition 73 and Example 68.

Definition 46 (Ideal). Let R be a ring. We can consider R as a module over itself.[5] The submodules of R are called *ideals*.[6]

Exercise 47. Prove that if $N, P \leqslant M$ are submodules of M, then so are $N+P$ and $N \cap P$.

Definition 48 (Generated Submodule). Let M be an R-module; and let $S \subseteq M$ be any set. The submodule of M *generated by* S is the set of all linear combinations of elements in S; that is, the set

$$\Big\{\sum_{s\in S} a_s s \;\Big|\; a_s \in R \text{ and } a_s = 0 \text{ for almost all } s \in S\Big\}.$$

We will denote it by RS (or SR in the case of right modules). If $S = \{x\}$ is a singleton set, we also just write Rx. We extend the same notation to ideals.

Remark 49. Alternatively, we could define

$$RS = \Big\{\sum_{s\in T} a_s s \;\Big|\; T \subseteq S \text{ finite and } a_s \in R \text{ for all } s \in T\Big\}.$$

The submodule generated by S is the smallest submodule of M which contains S. Since the empty sum is usually taken to be just 0, we have $R\varnothing = \{0\}$.

Remark 50. Another way to characterise RS is by saying that

$$RS = \bigcap_{N \leqslant M, S \subseteq N} N$$

that is, that RS is the intersection of all submodules of M which contain S.

Remark 51. A subset $N \subseteq M$ is a submodule (see Definition 42) if and only if $RN = N$.

Definition 52 (Generating Set). Let M be an R-module; and let $S \subseteq M$ be any set. We say that S *generates* M if $RS = M$.

Definition 53 (Cyclic Module/Principal ideal). A (sub-) module M which is generated by a single element x (that is, $M = Rx$) is called *cyclic*. A cyclic ideal is usually called *principal*.

Example 54. Consider the set $M = \mathbb{R}\{\sin, \cos\} = \{a\sin + b\cos \mid a, b \in \mathbb{R}\}$ which consists of all $\mathbb{R}$-linear combinations of the sine and the cosine. This is an $\mathbb{R}[\partial]$-module under the action introduced in Example 38. We will show that M is cyclic; in fact, we will prove the stronger claim that $M = \mathbb{R}[\partial]x$ for every non-zero element $x \in M$. Let $x = a\sin + b\cos \in M$ with $a, b \in \mathbb{R}$ not both zero; that is, $a^2 + b^2 \neq 0$. Then $\partial \bullet x = a\cos - b\sin$. Let now $y = u\sin + v\cos \in M$ be any element. We obtain

$$\frac{(av-bu)\partial + (bv+au)}{a^2+b^2}\bullet(a\sin + b\cos) = \frac{(av-bu)(a\cos - b\sin) + (bv+au)(a\sin + b\cos)}{a^2+b^2} = \frac{(b^2u - avb + bvaa^2u)\sin + (a^2v - bua + b^2v + aub)\cos}{a^2+b^2} = u\sin + v\cos;$$

that is, $y \in \mathbb{R}[\partial]x$.

[5] This is a special case of Example 34 with $n = 1$.

[6] Please note that we consider only commutative rings in this lecture. Thus, left and right ideals are the same.

Definition 55 (Linearly Independence). Let M be an R-module; and let $S \subseteq M$ be any set. Then the set S is called *(R-) linearly dependent* if there exists a family $(a_s)_{s \in S} \subseteq R$ such that

$$\sum_{s \in S} a_s s = 0$$

where $a_s = 0$ for almost all $s \in S$ but $a_s \neq 0$ for at least one $s \in S$. A subset $T \subseteq M$ which is not linearly dependent is called *linearly independent*.

Remark 56. Again, there an alternative definition: The set $S \subseteq M$ is linearly dependent if there exists a non-empty, finite subset $T \subseteq S$ and $a_s \in R \setminus \{0\}$ for all $s \in T$ such that

$$\sum_{s \in T} a_s s = 0.$$

Note that every set S which includes 0 is linearly dependent.

Remark 57. In a *vector space* V, if $S \subseteq V$ is linearly dependent, then there is $s \in S$ such that s can be written as a linear combination of vectors in $S \setminus \{s\}$. This is *not* true for a general module. As an example, consider the $\mathbb{Z}$-module $\mathbb{Z}^2$. Here $S = \{(2,0),(3,0)\}$ is linearly dependent; but no element can be written as a ($\mathbb{Z}$-linear) combination of the other.

Definition 58 (Basis). Let M be an R-module. A subset $B \subseteq M$ is called a *basis* of M if it is linearly independent and generates M.

Remark 59. When dealing with finite bases, we will usually assume that the elements are ordered in a specific way. In fact, we will often just write them as a family $(b_1, b_2, \ldots, b_n)$ instead of as a set. The implicit convention here is that two (finite) bases with the same elements are considered to be different if the order of the elements differs.

Example 60. For a ring R, consider the free module R^n (see Example 34). One possible basis is given by the family $(e_1, \ldots, e_n)$ consisting of the *unit vectors*

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \quad \ldots, \quad e_{n-1} = (0, \ldots, 0, 1, 0), \quad e_n = (0, \ldots, 0, 1).$$

Example 61. The $\mathbb{R}[\partial]$-module $M = \mathbb{R}\{\sin, \cos\}$ from Example 54 does not have a basis since $(\partial^2 + 1) \bullet (a \sin + b \cos) = 0$ for all $a, b \in \mathbb{R}$. That is, any non-empty subset of M is linearly dependent and can hence not be a basis. (Actually, M is just the submodule of Example 45.) A module M where for each $x \in M$ there is a regular $a \in R$ such that $a \bullet x = 0$ is called a *torsion module*.

Exercise 62. Show that $\mathbb{R}[x] \subseteq C^\infty(\mathbb{R})$ does not have a basis when considered as an $\mathbb{R}[\partial]$-module; but that it does have a basis if we consider it just as an $\mathbb{R}$-module.

Remark 63. In a *vector space* V, for $B \subseteq V$ the following statements are equivalent:

(a) B is a basis of V.

(b) B is a minimal generating set for V.

(c) B is a maximally linearly independent set in V.

However, this is generally *not true* for modules. As an example, consider $\mathbb{Z}^2$ as a $\mathbb{Z}$-module. The set

$$\{(2,0),(3,0),(0,2),(0,3)\}$$

is a minimal generating set (that is, every element of $\mathbb{Z}^2$ can be represented as a linear combinations of the elements of this set and removing any one of these will destroy that property); however, it is not a basis because the elements are linearly dependent. Similarly, the set

$$\{(2,0),(0,2)\}$$

is maximally linearly independent (that is, adding any other vector would make the set linearly dependent); but it is not a basis since it does not generate $\mathbb{Z}^2$.

Exercise 64. If an R-module M has a basis B, then every element $x \in M$ has a unique representation as a linear combination of basis elements.

4 Linear Maps

Definition 65 (Linear Map/Homomorphism/Endomorphism). Let M and N be two R-modules. A map $\varphi\colon M \to N$ is called *linear* over R or a *homomorphism* over R if

$$\varphi(x+y)=\varphi(x)+\varphi(y) \qquad \text{and} \qquad \varphi(ax)=a\varphi(x)$$

for all $x, y \in M$ and $a \in R$. We denote the sets of all R-linear maps between M and N by $\mathrm{Hom}_R(M,N)$.

If $M = N$, then the linear map φ is called an *endomorphism*. We write the set of all endomorphisms from M to itself as $\mathrm{End}_R(M)$.

Definition 66 (Isomorphism/Automorphism). An R-linear map $\varphi\colon M \to N$ is called an *isomorphism* if it is bijective. A bijective endomorphism is also called an automorphism. We say that two modules M and N are *isomorphic* if there exists an isomorphism between M and N. This is usually denoted by $M \cong N$.

Notation 67 (Identity). For every R-module M the identity map $\mathrm{id}_M\colon M \to M$ is an isomorphism. We will often leave out the index if it is clear to which module we are referring.

Example 68. Let M be an R-module, and let $a \in R$. Then the map φ given by $x \mapsto ax$ is linear. Indeed, let $x, y \in M$ and $b \in R$; then by the module laws we have

$$\varphi(x+y)=a(x+y)=ax+ay=\varphi(x)+\varphi(y)$$

and

$$\varphi(bx)=a(bx)=(ab)x=(ba)x=b(ax)=b\varphi(x).$$

(Note that we used the commutativity of R in the second computation; for a non-commutatve ring this kind of map is in general not linear.)

Exercise 69. Let $\varphi\colon M \to N$ be R-linear. Show that $\varphi(0)=0$.

Exercise 70. Let $\varphi, \hat{\varphi}\colon M \to N$ and $\psi\colon N \to P$ be R-linear maps; and let $a \in R$. Show that also $\varphi + \hat{\varphi}$, $a\varphi$, and $\psi \circ \varphi$ are R-linear, too. In that case that φ and ψ are isomorphisms, show that also $\psi \circ \varphi$ and φ^{-1} are isomorphisms.

Remark 71. Let M, N be R-modules. From Exercise 70 we can conclude that $\mathrm{Hom}_R(M,N)$ is also an R-module. Moreover, $\mathrm{End}_R(M)$ is a non-commutative ring with composition $\circ$ as multiplication and with units $\mathrm{End}_R(M)^* = \mathrm{Aut}_R(M)$.

Exercise 72. Let $\varphi\colon M \to N$ be an R-linear map; and let $U \leqslant M$ and $V \leqslant N$ be submodules. Then $\varphi(U)$ is a submodule of N and $\varphi^{-1}(V)$ is a submodule of M.

Definition 73 (Kernel/Image). For an R-linear map $\varphi\colon M \to N$ the *kernel* is $\ker\varphi = \varphi^{-1}(\{0\})$. Furthermore, we call $\operatorname{im}\varphi = \varphi(M)$ the *image* of φ. As Exercise 72 shows, $\ker\varphi \leqslant M$ and $\operatorname{im}\varphi \leqslant N$.

Theorem 74. *Let M be an R-module with a finite basis $B = (b_1, \ldots, b_n)$, let N be any R-module, and let $\varphi\colon M \to N$ be an R-linear map. Then φ is completely determined by the images of the basis elements. Conversely, any choice of images for the basis elements of M defines a homomorphism $\psi\colon M \to N$.*

Proof. Let $c_j = \varphi(b_j)$ for $j = 1, \ldots, n$. For any $x \in M$ we have the representation $x = x_1b_1 + \ldots + x_nb_n$ with $x_1, \ldots, x_n \in R$ since B is a basis. Then,

$$\varphi(x) = \varphi(x_1b_1 + \ldots + x_nb_n) = x_1\varphi(b_1) + \ldots + x_n\varphi(b_n) = x_1c_1 + \ldots + x_nc_n.$$

Thus, we can easily reconstruct φ from how it maps the basis elements.

Conversely, let $d_1, \ldots, d_n \in N$ and define

$$\psi(x) = \psi(x_1b_1 + \ldots + x_nb_n) = x_1d_1 + \ldots + x_nd_n.$$

It is easy to check that this is indeed a homomorphism. □

Corollary 75 (Free Modules)**.** *Let M be an R-module with a finite basis $B = (b_1, \ldots, b_n)$. Then we have $M \cong R^n$.*

Proof. It is easy to check that φ defined by $\varphi(e_j) = b_j$ is an isomorphism. □

Remark 76. Modules with a basis are usually called *free*. Moreover, Theorem 74 also holds for infinitely generated modules with basis B where we have to use the direct sum $R^{(B)} = \bigoplus_{b\in B} R$.

5 Matrices

Remark 77. Let M and N be two free R-modules with finite bases $B = (b_1, \ldots, b_m)$ for M and $C = (c_1, \ldots, c_n)$ for N. Let $\varphi\colon M \to N$. Then φ is completely determined by the images of the basis elements in B, and those images have a unique representation

$$\varphi(b_i) = \sum_{j=1}^{n} a_{ij}c_j$$

for $a_{ij} \in R$ with $i = 1, \ldots, m$ and $j = 1, \ldots, n$. In particular, the image of $x = \sum_{i=1}^m x_ib_i$ under φ is

$$\varphi(x) = \sum_{i=1}^{m} x_i\varphi(b_i) = \sum_{j=1}^{n}\sum_{i=1}^{m} x_ia_{ij}c_j.$$

Under the canonical identification of M with R^m with respect to B in Corollary 75 we write x as (row) vector $(x_1, \ldots, x_n)$. Then the vector representing $\varphi(x)$ with respect to C is

$$\begin{pmatrix}\sum_{i=1}^m x_ia_{i1} & \cdots & \sum_{i=1}^m x_ia_{in}\end{pmatrix} = \begin{pmatrix}x_1 & \cdots & x_m\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}.$$

That is, the matrix $A = (a_{ij})_{i,j=1,1}^{m,n}$ describes the effect of the homomorphism φ with respect to to the bases B and C. It is easy to see that conversely each matrix leads to a unique linear map. Thus, the set of all m-by-n matrices and the homomorphisms from M to N are in one-to-one correspondence (with respect to the two bases B and C).

Notation 78 (Matrices). We denote the set of m-by-n *matrices* by ${}^mR^n$. If $m = 1$, then we simply write R^n; and if $n = 1$, then we write ${}^mR$. We denote the n-by-n unit matrix by $\mathbf{1}_n$, and the m-by-n zero matrix by $\mathbf{0}_{m\times n}$. In both cases we will omit the indices if they are obvious from the context. We will denote the transpose of a matrix $A \in {}^mR^n$ by A^t. For elements $a_1, \ldots, a_n \in R$ we use $\operatorname{diag}(a_1, \ldots, a_n) \in {}^nR^n$ to denote a diagonal matrix with diagonal entries $a_1, \ldots, a_n$.[7]

Exercise 79. Let M with basis $B = (b_1, \ldots, b_m)$, N with basis $C = (c_1, \ldots, c_n)$, and P with basis $D = (d_1, \ldots, d_p)$ be free R-modules. We denote the map from Remark 77 which associates a homomorphism with its matrix with respect to the bases B and C by $\mathcal{M}_{B,C}\colon \operatorname{Hom}_R(M, N) \to {}^mR^n$ (and similarly $\mathcal{M}_{C,D}$ and $\mathcal{M}_{B,D}$). Show that

$$\mathcal{M}_{B,C}(\varphi + \psi) = \mathcal{M}_{B,C}(\varphi) + \mathcal{M}_{B,C}(\psi) \qquad \text{and} \qquad \mathcal{M}_{B,D}(\varrho \circ \varphi) = \mathcal{M}_{B,C}(\varphi)\mathcal{M}_{C,D}(\varrho)$$

for all $\varphi, \psi \in \operatorname{Hom}_R(M, N)$ and $\varrho \in \operatorname{Hom}_R(N, P)$.

Definition 80 (Singular/Regular/Unimodular). Let $B \in {}^mR^n$ be a matrix. We call B

singular if there exists $v \in R^m$ such that $v \neq 0$ and $vB = 0$;

regular if it is not singular; and

unimodular if $m = n$ and there exists $A \in {}^mR^m$ such that $AB = \mathbf{1}_m$.

(Strictly speaking we have defined left singular, left regular and left unimodular; however, as we will show below in Theorem 82, this does not matter.)

Notation 81. We write $\mathrm{GL}_m(R)$ for the set of all unimodular m-by-m matrices over R.

Theorem 82. *Let R be an integral domain.*

(a) A matrix in ${}^mR^n$ is left singular if and only if it is right singular.

(b) A matrix in ${}^mR^n$ is left regular if and only if it is right regular.

(c) A matrix in ${}^mR^m$ is left unimodular if and only if it is right unimodular.

Proof. Since R is an integral domain, we can form the field of fractions $Q(R)$. Assume that $B \in {}^mR^n$ is singular, then there is $v \in R^m$ such that $vB = 0$. This equation remains true if we consider B and v to have entries in $Q(R)$. Since $Q(R)$ is a field, we know from linear algebra that there exists $w \in {}^nQ(R)$ such that $Bw = 0$. We can bring the entries of w to a common denominator $d \in R$ and write $w = d^{-1}\hat{w}$ where $\hat{w} \in {}^nR$. Then $0 = Bw = B(d^{-1}w) = d^{-1}(Bw)$ which implies $Bw = 0$. Conversely, whenever B is right singular, a similar argument shows that B is also left singular. This proves part (a). Part (b) is equivalent to that.

Let now $U \in {}^mR^m$ be a left unimodular matrix. That is, there exists $V \in {}^mR^m$ such that $VU = \mathbf{1}$. Again, this relation remains true over $Q(R)$. It follows once more from linear algebra, that U has an inverse which must then be equal to V; that is, also $UV = \mathbf{1}$. The converse can be proved analogously. Thus, part (c) holds, too. □

[7] Later on, we will also abuse this notation for matrices which are not square.

Exercise 83. Let R be an integral domain, and let $a, b \in Q(R)$ be fractions over R. Show that we can write $a = d^{-1}\hat{a}$ and $b = d^{-1}\hat{b}$ where $\hat{a}, \hat{b}, d \in R$. (That is, show that we can bring fractions to a common denominator.)

6 Determinants

Definition 84 (Determinant). A function $\det\colon {}^mR^m \to R$ from the square matrices into the base ring is called a *determinant* if

(a) it is linear in each row, that is,

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ v+w \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ v \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ w \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} \qquad \text{and} \qquad \det\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ bv \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} = b\det\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ v \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix}$$

for all j and for all (rows) $a_1, \ldots, a_{j-1}, v, w, a_{j+1}, \ldots, a_m \in R^m$ and $b \in R$;

(b) $\det A = 0$ if the matrix $A \in {}^mR^m$ has two adjacent rows which are equal; and

(c) $\det \mathbf{1}_m = 1$.

Remark 85. From Definition 84 we get all the usual properties of determinants. Below, let $A \in {}^mR^m$ with rows $a_1, \ldots, a_m \in R^m$, and let $\det\colon {}^mR^m \to R$ be a determinant.

(a) If a row of A is zero, then $\det A = 0$. This follows from rule (a) of Definition 84 by using $b = 0$.

(b) Adding a linear multiple of a row to an adjacent row does not change the determinant since for $b \in R$ and $1 \leqslant j < m$

$$\det\begin{pmatrix} \vdots \\ a_j + ba_{j+1} \\ a_{j+1} \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ a_j \\ a_{j+1} \\ \vdots \end{pmatrix} + b\det\begin{pmatrix} \vdots \\ a_{j+1} \\ a_{j+1} \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ a_j \\ a_{j+1} \\ \vdots \end{pmatrix}$$

by first rule (a) and then rule (b).

(c) We can exchange two adjacent rows which will switch the sign of the determinant since by rule (b) and rule (a)

$$0 = \det \begin{pmatrix} \vdots \\ a_j + a_{j+1} \\ a_j + a_{j+1} \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_j \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ a_{j+1} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_{j+1} \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_{j+1} \\ a_{j+1} \\ \vdots \end{pmatrix}$$

$$= \det \begin{pmatrix} \vdots \\ a_j \\ a_{j+1} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_{j+1} \\ a_j \\ \vdots \end{pmatrix}$$

for any $1 \leqslant j < m$. By extension, since every permutation is a product of transpositions, we can permute the rows of A in any way where the determinant changes by the sign of of the permutation[8]. That is, if S_m denotes the set of all permutations of $\{1, \ldots, m\}$ and if $\pi \in S_m$, then

$$\det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \operatorname{sign}(\pi) \det \begin{pmatrix} a_{\pi(1)} \\ a_{\pi(2)} \\ \vdots \\ a_{\pi(m)} \end{pmatrix}.$$

(d) By the previous item, a determinant is zero if any two rows of A are equal. Moreover, adding a scalar multiple of any row to any other row does not change the determinant.

(e) Let $B = (b_{ij})_{ij} \in {}^mR^m$ be another matrix. We look now at the product BA. We have by linearity that

$$\det(BA) = \det \begin{pmatrix} \sum_{j_1=1}^m b_{1j_1} a_{j_1} \\ \vdots \\ \sum_{j_m=1}^m b_{mj_m} a_{j_m} \end{pmatrix} = \sum_{j_1=1}^m b_{1j_1} \det \begin{pmatrix} a_{j_1} \\ \vdots \\ \sum_{j_m=1}^m b_{mj_m} a_{j_m} \end{pmatrix}$$

$$= \sum_{j_1=1}^m b_{1j_1} \ldots \sum_{j_m=1}^m b_{mj_m} \det \begin{pmatrix} a_{j_1} \\ \vdots \\ a_{j_m} \end{pmatrix}.$$

The matrices in the last expression contain all possible combinations of rows of A. However, whenever in any of them a specific row appears twice, the determinant is zero by rule (b) of Definition 84. Thus, only terms survive where all the $j_1, \ldots, j_m$ are pairwise different. In other words, $j_1, \ldots, j_m$ is a permutation of $\{1, \ldots, m\}$. Moreover, any such permutation occurs in the sum. Thus, we obtain

$$\det(BA) = \sum_{\pi \in S_m} b_{1\pi(1)} \cdots b_{m\pi(m)} \begin{pmatrix} a_{\pi(1)} \\ \vdots \\ a_{\pi(m)} \end{pmatrix} = \sum_{\pi \in S_m} b_{1\pi(1)} \cdots b_{m\pi(m)} \operatorname{sign}(\pi) \det A$$

[8] The sign is the number of transpositions needed to express the permutation.

where we used item (c) in the last identity.

(f) We have not used rule (c) of Definition 84 so far. Letting $A = \mathbf{1}_m$ in the last identity of item (e) and using rule (c), we obtain the *Leibniz formula* for the determinant

$$\det B = \sum_{\pi \in S_m} \operatorname{sign}(\pi) b_{1\pi(1)} \cdots b_{m\pi(m)}.$$

In particular, this formular proves that there is only one determinant: As soon as we have the properties of Definition 84, we always arrive at the above formula.

(g) The Leibniz formula item (f) together with item (e) does also yield the formula for the product of determinants

$$\det(BA) = \Big(\sum_{\pi \in S_m} \operatorname{sign}(\pi) b_{1\pi(1)} \cdots b_{m\pi(m)} \Big) \det A = (\det B)(\det A).$$

(h) Consider a permutation $\pi \in S_n$. Then for each pair $(j, \pi(j))$, we have a corresponding pair $(\pi^{-1}(j), j)$. Thus, we can rewrite the product $b_{1\pi(1)} \cdots b_{m\pi(m)}$ as $b_{\pi^{-1}(1)1} \cdots b_{\pi^{-1}(m)m}$ by reordering the factors appropriately. If π now runs through all permutations, then so does π^{-1}. Moreover, $\operatorname{sign}(\pi) = \operatorname{sign}(\pi^{-1})$ since π^{-1} is a product of the same transpositions in the opposite order (because transpositions are their own inverses). Thus, we obtain

$$\det B = \sum_{\pi \in S_n} \operatorname{sign}(\pi) b_{1\pi(1)} \cdots b_{m\pi(m)} = \sum_{\pi^{-1} \in S_n} \operatorname{sign}(\pi^{-1}) b_{\pi^{-1}(1)1} \cdots b_{\pi^{-1}(m)m} = \det B^t$$

where $B^t = (b_{ji})_{ij} \in {}^m R^m$ is the transpose of B.

(i) By item (h) we see that the determinant is also linear in every column, vanishes if two columns are the same, changes sign if we permute the columns, and remains unchanged if we add scalar multiples of one column to another.

Remark 86. From linear algebra we remember that there is another way to define the determinant: Let $A = (a_{ij})_{ij} \in {}^m R^m$. Let $A_{\overline{jk}}$ denote the matrix A with the j^{th} row and the k^{th} column removed. Then for $k \leqslant m$ the *Laplace expansion* with respect to the k^{th} column is[9]

$$\det A = \sum_{j=1}^{m} (-1)^{j+k} a_{jk} \det A_{\overline{jk}}$$

or $\det A = (a_{11})$ if $m = 1$ where the determinants $\det A_{\overline{jk}}$ can be computed with any formula. In order to show that this definition does indeed yield the determinant, we just have to prove that the rules of Definition 84 hold. For this, we use induction over m. During the proof, we write $\det'$ for the determinant defined by the formula to distinguish it from the determinant of Definition 84. For $m = 1$ the rules are obviously true. Let now $m \geqslant 2$ and $k \leqslant m$. Pick any $j \leqslant m$ and assume that the j^{th} row of A is of the form $v + w$. Then for any $A_{\overline{ik}}$ with $i < j$, the $(j-1)^{\text{th}}$ row will be

[9] It would be more precise to use a different symbol until we have proved that this is indeed a determinant.

the sum of v and w with the k^{th} entry removed. Therefore, the i^{th} term in the sum for $\det' A$ is

$$\begin{aligned}(-1)^{i+k} a_{ik} \det A_{\overline{ik}} &= (-1)^{i+k} a_{ik} \det \begin{pmatrix} \vdots \\ v+w \\ \vdots \end{pmatrix}_{\overline{ik}} \\ &= (-1)^{i+k} a_{ik} \det \begin{pmatrix} \vdots \\ v \\ \vdots \end{pmatrix}_{\overline{ik}} + (-1)^{i+k} a_{ik} \det \begin{pmatrix} \vdots \\ w \\ \vdots \end{pmatrix}_{\overline{ik}}\end{aligned}$$

by induction. The same holds for $i > j$ and the j^{th} row. If $i = j$, then the term is

$$(-1)^{j+k}(v_k + w_k) \det A_{\overline{jk}} = (-1)^{j+k} v_k \det A_{\overline{jk}} + (-1)^{j+k} w_k \det A_{\overline{jk}}.$$

Thus, the sum splits into the determinant of A with k^{th} row equal to v and the determinant of A with k^{th} row equal to w. Similarly, we can see that multiplication by a scalar works as required. Let us assume that for $j < m$ the j^{th} row and the $(j+1)^{\text{th}}$ row of A are the same. Then all $A_{\overline{ik}}$ where $i < j$ or $i > j+1$ have two equal rows which means that their determinants vanish. On the other hand, $A_{\overline{jk}}$ and $A_{\overline{j+1,k}}$ are the same matrix. Thus,

$$\begin{aligned}\det' A &= (-1)^{j+k} a_{jk} \det A_{\overline{jk}} + (-1)^{j+1+k} a_{j+1,k} \det A_{\overline{j+1,k}} \\ &= (-1)^{j+k} a_{jk} \det A_{\overline{jk}} - (-1)^{j+k} a_{jk} \det A_{\overline{jk}} = 0.\end{aligned}$$

Finally, if $A = \mathbf{1}_m$ is the identity, then $A_{\overline{jk}}$ has a zero row unless $j = k$ in which case $A_{\overline{kk}} = \mathbf{1}_{m-1}$. Thus

$$\det' \mathbf{1}_m = (-1)^{k+k} \det \mathbf{1}_{m-1} = 1$$

as required. We see that $\det'$ is indeed a determinant in the sense of Definition 84.

Using the transpose, we obtain the Laplace expansion with respect to the k^{th} column.

Theorem 87. *Let $A = (a_{ij})_{ij} \in {}^mR^m$. Then the determinant $\det A$ is given by the*

Leibniz formula

$$\det A = \sum_{\pi \in S_m} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{m\pi(m)}$$

(where S_m are the permutations of $\{1, \dots, m\}$), or the

Laplace expansion with respect to the k^{th} column

$$\det A = \sum_{j=1}^{m} (-1)^{j+k} a_{jk} \det A_{\overline{jk}}$$

(where $A_{\overline{jk}}$ is A with the j^{th} row and k^{th} column removed), or the

Laplace expansion with respect to the k^{th} row

$$\det A = \sum_{j=1}^{m} (-1)^{j+k} a_{kj} \det A_{\overline{kj}}.$$

Moreover, the determinant is linear in every row and column, vanishes if any two rows or any two columns are the same, changes sign when any two rows or any two columns are swapped, and remains the same if a scalar multliple of any row is added to any other row or a scalar multliple of any column is added to any other column. Furthermore we have $\det A = \det A^t$ *and* $\det(AB) = (\det A)(\det B)$ *for any matrix* $B \in {}^mR^m$.

Example 88. Consider the polynomials $R[x]$ over R. Let $A = (a_{ij})_{ij} \in {}^nR^n$ be a constant matrix. Then $\det(\mathbf{1}_n x - A) \in R[x]$ is a polynomial of degree n of the form $x^n + r_1 x^{n-1} + \ldots + r_{n-1}x + r_n$ where $r_1 = -a_{11} - \ldots - a_{nn}$ and $r_n = (-1)^n \det A$. Formulae for the other coefficients also exist.

Definition 89 (Adjugate Matrix). Let $A \in {}^mR^m$. The *adjugate (matrix)* $\operatorname{adj} A$ is the m-by-m matrix with entries

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A_{\overline{ji}}$$

where as before $A_{\overline{ji}}$ is the matrix A with j^{th} row and i^{th} column removed.

Theorem 90. *For any matrix* $A \in {}^mR^m$ *we have*

$$A \operatorname{adj} A = (\operatorname{adj} A)A = \det A \cdot \mathbf{1}_m.$$

Proof. Indeed, for the first product we have

$$(A \operatorname{adj} A)_{ij} = \sum_{k=1}^{m} a_{ik}(\operatorname{adj} A)_{kj} = \sum_{k=1}^{m} a_{ik}(-1)^{k+j} \det A_{\overline{jk}}.$$

Now, if $i = j$ this is exactly the Laplace expansion for $\det A$ with respect to the i^{th} row. However, for $i \neq j$ this is equal to the Laplace expansion of a copy of A where the j^{th} row is replaced by the i^{th} row. So, $(A \operatorname{adj} A)_{ij} = 0$ in that case. Similarly, we can prove $(\operatorname{adj} A)A = \det A \cdot \mathbf{1}$. □

Theorem 91. *Let* R *be an integral domain and* $A \in {}^mR^m$ *be a square matrix. Then*

(a) A *is singular if and only if* $\det A = 0$*;*

(b) A *is regular if and only if* $\det A \neq 0$*; and*

(c) A *is unimodular if and only if* $\det A \in R^*$*. In that case* $(\det A)^{-1} = \det A^{-1}$.

Proof. Let $A \in {}^mR^m$ be singular. Then A is also singular over the field of fractions $Q(R)$ and thus $\det A = 0$. Conversely, if $\det A = 0$, then A is singular over $Q(R)$. Thus there is $v \in Q(R)^m$ with $vA = 0$. Bring the entries of v to a common denominator $v = d^{-1}w$ with $d \in R$ and $w \in R^m$. Then $0 = vA = d^{-1}(wA)$; that is, $wA = 0$ and we see that A is also singular over R. This proves part (a). Part (b) is equivalent to part (a).

Let now $A \in \mathrm{GL}_m(R)$ be unimodular. Then $A^{-1}A = \mathbf{1}$ and thus $(\det A^{-1})(\det A) = 1$ showing that $\det A \in R^*$ and also that $\det A^{-1} = (\det A)^{-1}$. On the other hand, let $\det A$ be a unit. Then we have $(\operatorname{adj} A)A = (\det A)\mathbf{1}$ and thus we see that A has the inverse $A^{-1} = (\det A)^{-1} \operatorname{adj} A$; that is, A is unimodular. □

Remark 92. If R is not an integral domain, then A can be singular even if $\det A \neq 0$. Consider $R = \mathbb{Z}_8$ (that is, the integers modulo 8) and

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in {}^2\mathbb{Z}_8{}^2.$$

Then $\det A \equiv 6 \not\equiv 0 \pmod 8$; but $v = (4,4) \in \mathbb{Z}_8^2 \setminus \{0,0\}$ fulfills $vA = (16,24) \equiv 0 \pmod 8$.

7 The Theorem of Caley-Hamilton

Exercise 93. Let $I \leqslant R$ be an ideal, and let M be an R-module. Show that

$$IM = \{a_1x_1 + \ldots + a_kx_k \mid k \geqslant 0, a_1, \ldots, a_k \in I, \text{ and } x_1, \ldots, x_k \in M\}$$

is a submodule of M. If M is generated by S, then show that IM is generated by IS.

Remark 94. Since any ideal $I \leqslant R$ is an R-submodule of R, we can apply Exercise 93 to $M = I$ and define the ideals $I^2 = II$, $I^3 = II^2$, and so on.

Exercise 95. Let M be an R-module, and let $\varphi \in \operatorname{End}_R(M)$ be an endomorphism. Show that the action

$$\bullet\colon R[x] \times M \to M, \quad (a_nx^n + \ldots + a_1x + a_0,\ m) \mapsto a_n\varphi^n(m) + \ldots + a_1\varphi(m) + a_0m$$

makes M into an $R[x]$-module.

Remark 96. Let M be an R-module. Then we can let the matrices ${}^mR^n$ act on (column) vectors ${}^nM$ over M in the following way: For $A = (a_{ij})_{ij} \in {}^mR^n$ and $x = (x_1, \ldots, x_n)^t \in {}^nM$ we let

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \bullet \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}.$$

This yields an R-linear map $A\bullet\colon {}^nM \to {}^mM$. (In order to prove $A(rx) = r(Ax)$ for all $r \in R$, we need to use commutativity.) It is obvious that $A \mapsto A\bullet$ is a R-linear map ${}^mR^n \to \operatorname{Hom}_R({}^nM, {}^mM)$.

Exercise 97. Show that with the action defined in Remark 96 we have $A(Bx) = (AB)x$ for all $A \in {}^mR^n$, $B \in {}^nR^p$, and $x \in {}^pR$.

Theorem 98 (Caley-Hamilton). *Let $I \leqslant R$ be an ideal, and let M be an R-module which is generated by n elements. Let $\varphi\colon M \to M$ be an endomorphism such that $\varphi(M) \subseteq IM$. Then there exists a monic polynomial*

$$f = x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in R[x]$$

such that $a_j \in I^j$ for $j = 1, \ldots, n$ and

$$f(\varphi) = \varphi^n + a_1\varphi^{n-1} + \ldots + a_{n-1}\varphi + a_n \operatorname{id}_M = 0.$$

Proof. Let $y_1, \ldots, y_n \in M$ be the generators of M. Then for each $i = 1, \ldots, n$ we have $\varphi(y_i) = a_{i1}y_1 + \ldots + a_{in}y_n$ for some $a_{i1}, \ldots, a_{in} \in I$. By Exercise 95, we can consider M as a $R[x]$-module. We can extend the $R[x]$-action to ${}^nR[x]^n$ as shown in Remark 96. Consider the matrix $A = (\delta_{ij}x - a_{ij})_{ij} \in {}^nR[x]^n$ and $y = (y_1, \ldots, y_n)^t \in {}^nM$. Then the equations above give

$$Ay = \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & x - a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \varphi(y_1) - a_{11}y_1 - \ldots - a_{1n}y_n \\ \vdots \\ \varphi(y_n) - a_{n1}y_1 - \ldots - a_{nn}y_n \end{pmatrix} = 0.$$

By Exercise 97, we have

$$0 = (\operatorname{adj} A)0 = (\operatorname{adj} A)(Ay) = ((\operatorname{adj} A)A)y = (\det A)\mathbf{1}_n y = (\det A)y.$$

Thus, $\det A$ annihilates every entry of y; that is, $(\det A)y_j = 0$ for $j = 1, \dots, n$. Consequently, since $y_1, \dots, y_n$ generate M, we obtain $(\det A)z = 0$ for every $z = z_1 y_1 + \ldots + z_n y_n \in M$. Using the Leibniz formula for the determinant (see Theorem 87), we obtain

$$\det A = x^n + a_1 x^{n-1} + \ldots + a_n$$

where $a_j \in I^j$ as required. □

Remark 99. In particular, we can use Theorem 98 for endormorphisms given by a matrix: Let $A \in {}^m R^m$ and consider the endormorphism $\bullet A \colon R^m \to R^m$ which is given by $v \mapsto vA$. We can use $I = R$ and obtain a polynomial $f = x^n + b_1 x^{n-1} + \ldots + b_n \in R[X]$ such that

$$0 = f(\bullet A) = (\bullet A)^n + b_1(\bullet A)^{n-1} + \ldots + b_n \operatorname{id} = \bullet(A^n + b_1 A^{n-1} + \ldots + b_n \operatorname{id}) = \bullet f(A)$$

by the way matrix multiplication and addition corresponds to the composition of endormorphisms (see Exercise 79). Applying the map $\bullet f(A)$ to the unit vectors extracts the rows of $f(A)$ which must thus all be zero. Hence we see that the Theorem of Caley-Hamilton as tought in basic linear algebra is just a special case of Theorem 98.

Corollary 100. *Let M be a finitely generated R-module, and $I \leqslant R$ be an ideal such that $IM = M$. Then there exists $r \in I$ such that $(1 - r)M = 0$.*

Proof. Consider the R-linear map $\operatorname{id} \colon M \to M$. By Theorem 98, there is $n \geqslant 1$ and $a_1, \dots, a_n \in I$ such that

$$0 = \operatorname{id}^n + a_1 \operatorname{id}^{n-1} + \ldots + a_n \operatorname{id} = \operatorname{id} + (a_1 + \ldots + a_n) \operatorname{id}.$$

Thus, with $r = -(a_1 + \ldots + a_n)$ we obtain $(1 - r)x = x - rx = 0$ for all $x \in M$. □

Theorem 101. *Let M be a finitely generated R-module. If $\varphi \colon M \to M$ is a surjective R-linear map, then φ is an isomorphism.*

Proof. We regard M as an $R[x]$ module with action $xy = \varphi(y)$ for all $y \in M$ as in Exercise 95. It is easy to see that M is also finitely generated as an $R[x]$-module. Consider the ideal $I = R[x]x$. Since φ is surjective, we have $xM = M$ and thus also $IM = M$. By Corollary 100 there exists $g \in I$ with $(1 - g)y = 0$ or $y = gy$ for all $y \in M$. Since we can write $g = \tilde{g}x = x\tilde{g}$ for some $\tilde{g} \in R[x]$, we have $\operatorname{id} = \tilde{g}(\varphi)\varphi = \varphi\tilde{g}(\varphi)$; that is, $\tilde{g}(\varphi) = \varphi^{-1}$. □

Theorem 102. *If $M \cong R^n$ for an R-module M, then every generating set of M with n elements is a basis of M. In particular, the number of elements of a basis of a free module is always the same.*

Proof. Let $y_1, \dots, y_n$ be n generators of M. Define the R-linear map $\varphi \colon R^n \to M$ by $\varphi(e_j) = y_j$ for $j = 1, \dots, n$ where $e_1, \dots, e_n$ are the unit vectors. Obviously, φ is surjective. By the assumption of the theorem, there is an isomorphism $\psi \colon M \to R^n$. Then $\varphi \circ \psi \colon M \to M$ is also surjective and R-linear. Thus, from Theorem 101 we obtain that $\varphi \circ \psi$ is an isomorphism. Consequently, also $\varphi = (\varphi \circ \psi) \circ \psi^{-1}$ is an isomorphism. This implies that $y_1, \dots, y_n$ is a basis of M.

Assume now that M had a basis $B = (b_1, \dots, b_m)$ with m and a basis $C = (c_1, \dots, c_n)$ with n elements where $m < n$. Then we can add elements of C to B obtaining a generating set

$\{b_1, \ldots, b_m, c_{m+1}, \ldots, c_n\}$ with n elements. However, this cannot be a basis of M since, for example, $c_n = a_1 b_1 + \ldots + a_m b_m$ for some $a_1, \ldots, a_m \in R$. This contradicts the first part of the theorem. □

Corollary 103. *By Theorem 102, $R^m \cong R^n$ if and only if $m = n$.* □

Remark 104. We remark that Corollary 103 does not in general hold for non-commutative rings. See [Lam99, (1.4) Example] for an example.

Remark 105. With Theorem 101 we also obtain a different proof that left unimodular matrices are also right unimodular. However, this version works even for rings with zero divisors. Let $A, B \in {}^mR^m$ such that $AB = \mathbf{1}_m$. Then $\cdot B \colon R^m \to R^m$ defined as $v \mapsto vB$ is surjective. Indeed, for every $x \in R^m$ we have $(xA)B = x$. By Theorem 101, $\cdot B$ is an isomorphism. Thus, there exists a matrix $C \in {}^mR^m$ corresponding to $(\cdot B)^{-1}$ such that $CB = BC = \mathbf{1}_m$. (Of course, $C = A$.)

Definition 106 (Rank). Let M be a finitely generated free R-module. The size of a (and thus any) basis of M is called the *rank* of M. We denote it by $\operatorname{rank}_R M$ (or just $\operatorname{rank} M$ if it is clear to which ring we are referring).

Example 107. Even if a module has a finite basis, its submodules do not need to be finitely generated: Let F be a field and $X = \{x_1, x_2, x_3, \ldots\}$ be an infinite set of indeterminates. Then $F[X]$ is finitely generated as a module over itself (for instance, $F[X] \cdot 1 = F[X]$). However, $F[X]X \leqslant F[X]$ is not finitely generated.

Exercise 108. Prove that the claim in Example 107 is correct.

Part II
Matrix Normal Forms

8 Basic Notations for Matrices

Remark 109 (Block Matrices). Let $m_1, \ldots, m_s$ and $n_1, \ldots, n_t$ be positive integers and $A_{ij} \in {}^{m_i}R^{n_j}$ be matrices for $i = 1, \ldots, s$ and $j = 1, \ldots, t$. Then the *block matrix*

$$\begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{st} \end{pmatrix} \in {}^{m_1 + \ldots + m_s}R^{n_1 + \ldots + n_t}$$

is defined as the matrix where the $(i,j)^{\text{th}}$ entry is the $(\tilde{\imath}, \tilde{\jmath})^{\text{th}}$ entry of $A_{k\ell}$ where

$$\tilde{\imath} = i - m_1 - \ldots - m_{k-1} \qquad \text{and} \qquad \tilde{\jmath} = j - n_1 - \ldots - n_{\ell-1}$$

and k and ℓ are such that

$$m_1 + \ldots + m_{k-1} < i \leqslant m_1 + \ldots + m_k \qquad \text{and} \qquad n_1 + \ldots + n_{\ell-1} < j \leqslant n_1 + \ldots + n_\ell.$$

For instance, if

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad A_{12} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}, \qquad A_{21} = \begin{pmatrix} 7 & 8 \end{pmatrix}, \qquad \text{and} \qquad A_{22} = \begin{pmatrix} 9 \end{pmatrix},$$

then the block matrix would be

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Furthermore, assume that we have positive integers $p_1, \dots, p_u$ and matrices $B_{ij} \in {}^{n_i}R^{p_j}$ for $i = 1, \dots, t$ and $j = 1, \dots, u$. Then

$$\begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{st} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1u} \\ \vdots & & \vdots \\ B_{t1} & \cdots & B_{tu} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{t} A_{1k}B_{k1} & \cdots & \sum_{k=1}^{t} A_{1k}B_{ku} \\ \vdots & & \vdots \\ \sum_{k=1}^{t} A_{sk}B_{k1} & \cdots & \sum_{k=1}^{t} A_{sk}B_{ku} \end{pmatrix}.$$

In other words, when the sizes of the blocks match, then we can multiply block matrices like normal matrices with the blocks as entries.

Finally, we are also going to use *block diagonal matrices* which are defined as follows: For matrices $C_1 \in {}^{m_1}R^{n_1}, \dots, C_s \in {}^{m_s}R^{n_s}$ we let $\operatorname{diag}(C_1, \dots, C_s)$ be the block matrix $(A_{ij})_{ij}$ with blocks $A_{ii} = C_i$ and $A_{ij} = \mathbf{0}_{m_i \times n_j}$ for $i \neq j$.

Exercise 110. Prove that the multiplication formula of Remark 109 is correct.

Exercise 111. Let $U_1 \in \mathrm{GL}_{n_1}(R), \dots, U_k \in \mathrm{GL}_{n_k}(R)$ be unimodular matrices. Prove that the block diagonal matrix $\operatorname{diag}(U_1, \dots, U_k)$ is also unimodular.

Definition 112 (Elementary Matrix). Let R be a ring and $m \geqslant 1$. We define three types of *elementary matrices*:

Row/Column Addition Matrices For $q \in R$ and $j, k \leqslant m$ with $j \neq k$ let $\mathrm{Add}_{jk}(q) \in {}^{m}R^{m}$ be the identity matrix except with q at the $(j,k)^{\text{th}}$ position; that is, let

$$\begin{array}{c} \mathrm{Add}_{jk}(q) = \left(\begin{array}{ccccccc} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \ddots &&& \\ && q && 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \leftarrow j^{\text{th}}\text{ row} \\ \\ \\ \end{array} \\ \uparrow \\ k^{\text{th}}\text{ column} \end{array}$$

(where all the other entries are 0).

Row/Column Scaling Matrices For a unit in $u \in R^*$ and $j \leqslant m$ let $\mathrm{Mult}_j(u) \in {}^{m}R^{m}$ be the identity matrix except with u at the $(j,j)^{\text{th}}$ position; that is, $\mathrm{Mult}_j(u) = \operatorname{diag}(\mathbf{1}_{j-1}, u, \mathbf{1}_{m-j})$ using block matrix notation.

Row/Column Permuting Matrices For $j, k \leqslant m$ and $j \neq k$ let $\mathrm{Swap}_{jk} \in {}^{m}R^{m}$ be like the identity matrix except that the j^{th} row and the j^{th} column are exchanged; that is,

$$\operatorname{Swap}_{jk} = \begin{pmatrix} e_1 \\ \vdots \\ e_{j-1} \\ e_k \\ e_{j+1} \\ \vdots \\ e_{k-1} \\ e_j \\ e_{k+1} \\ \vdots \\ e_m \end{pmatrix} = \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 &&&& 1 &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& 1 &&&& 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow j^{\text{th}} \text{ row} \\ \\ \\ \\ \leftarrow k^{\text{th}} \text{ row} \\ \\ \\ \\ \end{array}$$

(using block matrix notation and showing the case that $j < k$).

Remark 113. Let $A \in {}^mR^n$ be a matrix.

(a) For $q \in R$, $\operatorname{Add}_{jk}(q)A$ equals A with q times the k^{th} row added to the j^{th} row, and $A \operatorname{Add}_{jk}(q)$ equals A with q times the j^{th} column added to the k^{th} column.

(b) For $u \in R^*$, $\operatorname{Mult}_j(u)A$ equals A with the j^{th} row multiplied by u, and $A \operatorname{Mult}_j(u)$ equals A with the j^{th} column multiplied by u.

(c) $\operatorname{Swap}_{jk} A$ equals A with the j^{th} and k^{th} rows interchanged, and $A \operatorname{Swap}_{jk}$ equals A with the k^{th} and j^{th} columns interchanged.

Remark 114. All elementary matrices are unimodular. More precisely, we have

$$\operatorname{Add}_{jk}(q)^{-1} = \operatorname{Add}_{jk}(-q), \qquad \operatorname{Mult}_j(u)^{-1} = \operatorname{Mult}_j(u^{-1}), \qquad \text{and} \qquad \operatorname{Swap}_{jk}^{-1} = \operatorname{Swap}_{jk}.$$

Moreover, it is

$$\det \operatorname{Add}_{jk}(q) = 1, \qquad \det \operatorname{Mult}_j(u) = u, \qquad \text{and} \qquad \det \operatorname{Swap}_{jk} = -1.$$

Exercise 115. Let M be an R-module and let $N, P \leqslant M$ be two submodules. The sum $N + P$ of N and P is defined as $N + P = \{a + b \mid a \in N \text{ and } b \in P\}$. Show that $N + P$ is a submodule and that $N \leqslant N + P$ and $P \leqslant N + P$. If S generates N and T generates P, then $S \cup T$ generates $N + P$.

Definition 116 (Row/Column Space). Let $A \in {}^mR^n$ be a matrix. The *row space* of A is the set of all R-linear combinations of the rows of A. We denote it by R^mA. Similarly, the *column space* is the R-linear combinations of all columns of A and we write $A\,{}^nR$ for that.

Remark 117. The idea of the notation in Definition 116 is that the row (or column) space is equal the sum of the cyclic modules generated by the rows (or columns) of M. Thus, in pseudo-notation, we have

$$R^mM = \begin{pmatrix} R & \cdots & R \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = Ra_1 + \ldots + Ra_n = R\{a_1, \ldots, a_n\}$$

where $a_1, \ldots, a_m \in R^n$ are the rows of M.

Remark 118. Another way to think about the row space of $M \in {}^mR^n$ is as the image of the linear map $\cdot M$ given by $v \mapsto vM$. Indeed, $R^m M = \operatorname{im}(\cdot M)$. Analogously, $M\,{}^nR = \operatorname{im}(M\cdot)$.

Remark 119. We will also apply Definition 116 to row vectors $v = (v_1, \ldots, v_m) \in R^m$ and column vectors $w = (w_1, \ldots, w_n)^t \in {}^nR$ regarding them as matrices with only one row or one column, respectively. For example, $R^n w = Rw_1 + \ldots + Rw_n$ is the *ideal* generated by $w_1, \ldots, w_n$.

9 Divisibility

Definition 120 (Divisor/Associates). Let $a, b \in R$. We say that a *divides* b and write $a \mid b$ if there exists $c \in R$ such that $ac = b$. We say that a and b are *associated* if a divides b and b divides a.

Exercise 121. Prove that in an integral domain a is an associate of b if and only if there exists a unit $u \in R^*$ such that $au = b$.

Exercise 122. Show that being associated is an equivalence relation on R.

Exercise 123. Let $a, b \in R$. Show that $a \mid b$ if and only if $Rb \leqslant Ra$.

Exercise 124. Show that the associates of 1 are precisely the units of R.

Definition 125 (Greatest Common Divisor). A *common divisor* of $a_1, \ldots, a_n \in R$ is an element $h \in R$ such that h divides a_j for each $j = 1, \ldots, n$. An element $g \in R$ is a *greatest common divisor* (or *GCD* for short) of $a_1, \ldots, a_n$ if g is a common divisor and every other common divisor $h \in R$ of $a_1, \ldots, a_n$ divides g. We write $g = \gcd(a_1, \ldots, a_n)$.

Exercise 126. Despite the use of the equality sign in the notation $g = \gcd(a_1, \ldots, a_n)$, greatest common divisors do not need to be unique. In fact, prove that every associate of a greatest common divisor of $a_1, \ldots, a_n$ is also a greatest common divisor of $a_1, \ldots, a_n$, and that conversely all greatest common divisors of $a_1, \ldots, a_n$ are associated.

Exercise 127. Let R be an integral domain, let $a_1, \ldots, a_n \in R$ be not all zero, and let $d = \gcd(a_1, \ldots, a_n)$. Write $a_j = d\tilde{a}_j$ for $j = 1, \ldots, n$. Show that $\gcd(\tilde{a}_1, \ldots, \tilde{a}_n) = 1$.

Definition 128 (Principal Ideal Domain). A ring R is called a *principal ideal domain* (or *PID* for short) if all its ideals are principal; that is, if all ideals $I \leqslant R$ are of the form $I = Ra$ for some $a \in R$ (see Definition 53).

Example 129. Every field is a principal ideal domain: A field F has only two ideals $\{0\} = F \cdot 0$ and $F = F \cdot 1$ which are both principal.

Example 130. The integers $\mathbb{Z}$ form a principal ideal domain. In order to prove this, we use integer long division (with remainder): If a and $b \in \mathbb{Z}$ with $b \neq 0$, then there are q and $r \in \mathbb{Z}$ with $a = qb + r$ and either $r = 0$ or $|r| < |b|$. Let now $I \leqslant \mathbb{Z}$ be any non-zero ideal. Then $I \setminus \{0\}$ is not empty. Choose an element $b \in I \setminus \{0\}$ with the smallest absolute value. Let $a \in I$ be any element. Then $a = qb + r$ where $r = 0$ or $|r| < |b|$. Let us assume that $r \neq 0$. Then $a - qb = r \in I$ was a member of I with a strictly smaller absolute value than b contradicting the choice of b. Thus, $r = 0$ and $a = qb \in Rb$. It follows that $I \leqslant Rb$; but of course we also have $Rb \leqslant I$, that is, $Rb = I$.

Example 131. If F is a field, then the univariate polynomials $F[X]$ are a principal ideal domain. The proof is exactly the same as in Example 130 using polynomial long division instead of integer long division.

Example 132. Multivariate polynomials are *not* principal ideal domains. For instance, for $R = \mathbb{Q}[x, y]$, the ideal $R\{x, y\}$ cannot be generated by one element. Also the polynomials $R = \mathbb{Z}[x]$ over the integers are also *not* a principal ideal domain. The ideal $R\{2, x\}$ cannot be generated by one element.

Definition 133. An integral domain R is a *Euclidean domain*, if there exists a map $\deg\colon R \setminus \{0\} \to \mathbb{N}$ (called a *degree function*) such that

(a) for every a and $b \in R$ with $b \neq 0$ there exist $q, r \in R$ such that $a = qb + r$ and either $r = 0$ or $\deg r < \deg b$; and

(b) for all non-zero $a, b \in R$ we have $\deg a \leqslant \deg(ab)$.

We call q in item (a) a *quotient* of a divided by b and r a *remainder*.

Example 134. The following rings are examples of Euclidean domains:

(a) The integers with $\deg a = |a|$.

(b) Univariate polynomial rings over fields with the usual degree function.

(c) Fields F with the degree function $\deg f = 0$ for all $f \neq 0$.

(d) The Gaussian integers $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$ where $i = \sqrt{-1}$. The degree function here is the square of the complex norm $\deg(a + bi) = a^2 + b^2$.

Exercise 135. Look up the division for the Gaussian integers (part (d) of Example 134) and implement it in a programming language of your choice.

Notation 136. Let R be a Euclidean domain and $a, b \in R$ with $b \neq 0$. Then there are $q, r \in R$ such that $a = qb + r$ with $r = 0$ or $\deg r < \deg b$. We will write $a \text{ quo } b = q$ for the quotient and $a \text{ rem } b = r$ for the remainder. Note that quotient and remainder do not need to be unique: For instance, considering the integers with $a = 7$ and $b = 4$ we have $7 = 1 \cdot 4 + 3 = 2 \cdot 4 - 1$. We will therefore assume that whenever a quotient and remainder of the same input a and b is computed, they will match each other. That is, we always assume $a = (a \text{ quo } b)b + (a \text{ rem } b)$ but do not bother what the exact choices might be.

Exercise 137. The univariate polynomials $F[x]$ over a field are a Euclidean domain with respect to the degree and the usual polynomial long division. Prove that for $F[x]$ quotient and remainder are always uniquely determined.

Theorem 138. *Every Euclidean domain is a principal ideal domain.*

Proof. The proof is the same as in Example 130. □

Remark 139. It is not easy to find a principal ideal domain which is not Euclidean. For an example see [And88] where it was shown that the ring $\mathbb{Q}[\![x, y]\!][(x^2 + y^3)^{-1}]$ is a principal ideal domain but not Euclidean.

Theorem 140. *Let R be a principal ideal domain, and let $a_1, \ldots, a_n \in R$. Then*

$$Ra_1 + Ra_2 + \ldots + Ra_n = R \gcd(a_1, \ldots, a_n).$$

Proof. Consider the ideal $Ra_1 + \ldots + Ra_n$. Since R is a principal ideal domain there must exist $g \in R$ such that $Ra_1 + \ldots + Ra_n = Rg$. Since then $Ra_j \leqslant Ra_1 + \ldots + Ra_n = Rg$ for every $j = 1, \ldots, n$ we get from Exercise 123 that g is a common divisor of $a_1, \ldots, a_n$. Assume that $h \mid a_1, \ldots, a_n$ for some h. This implies $Ra_j \leqslant Rh$ for all $j = 1, \ldots, n$, and thus also $Rg = Ra_1 + \ldots + Ra_n \leqslant Rh$. Consequently, $h \mid g$. That means that g is a greatest common divisor of $a_1, \ldots, a_n$.

Let conversely $\tilde{g}$ be any greatest common divisor of $a_1, \ldots, a_n$. Then g and $\tilde{g}$ are associated by Exercise 126. Thus, there exists $u \in R^*$ such that $\tilde{g} = ug$ and $R\tilde{g} = Rug = Rg = Ra_1 + \ldots + Ra_n$ because $Ru = R$. □

Corollary 141. *If R is a principal ideal domain, $a_1, \ldots, a_n \in R$, and $g \in R$ is a greatest common divisor of $a_1, \ldots, a_n$; then there are $s_1, \ldots, s_n \in R$ such that $g = s_1 a_1 + \ldots + s_n a_n$.* □

Definition 142 (Least Common Multiple). An element $m \in R$ is called a *common multiple* of $a_1, \ldots, a_n \in R$ if a_j divides m for $j = 1, \ldots, n$. We say that m is a *least common multiple* of $a_1, \ldots, a_n$ if it is a common multiple and m divides every other common multiple of $a_1, \ldots, a_n$. We write $m = \operatorname{lcm}(a_1, \ldots, a_n)$.

Exercise 143. Prove that any two least common multiples of $a_1, \ldots, a_n \in R$ are associated; and that any associate of a least common multiple of $a_1, \ldots, a_n$ is itself a least common multiple.

Exercise 144. Let R be a principal ideal domain. Prove that for $a_1, \ldots, a_n \in R$ we have $Ra_1 \cap \ldots \cap Ra_n = R \operatorname{lcm}(a_1, \ldots, a_n)$.

Notation 145. If $v = (v_1, \ldots, v_m) \in R^m$ or $w = (w_1, \ldots, w_n)^t \in {}^nR$, then write $\gcd(v) = \gcd(v_1, \ldots, v_m)$ and $\gcd(w) = \gcd(w_1, \ldots, w_n)$.

Remark 146. Using Notation 145 for vector greatest common divisors and Definition 116 for row spaces, we can write Theorem 140 in the more succinct form

$$R^n a = R \gcd(a)$$

where $a = (a_1, \ldots, a_n)^t \in {}^nR$ since $R^n a = \{s_1 a_1 + \ldots + s_n a_n \mid s_1, \ldots, s_n \in R\} = Ra_1 + \ldots + Ra_n$.

10 The Euclidean Algorithm

Theorem 147. *Let R be a principal ideal domain. Let $v \in {}^nR$ and $M \in {}^mR^n$. Then we have* $\gcd(v) \mid \gcd(Mv)$. *(A similar state holds for row vectors.)*

Proof. We can write the result of Theorem 140 in the form $R \gcd(v) = R^n v$. Since $R^m M \leqslant R^n$ is a subspace, we must have $R \gcd(Mv) = R^m Mv \leqslant R^n v = R \gcd(v)$ as ideals; and consequently $\gcd(v) \mid \gcd(Mv)$. □

Corollary 148. *If in particular $M \in \mathrm{GL}_n(R)$ in Theorem 147 is unimodular, then* $\gcd(v) = \gcd(Mv)$.

Proof. By Theorem 147, we have

$$\gcd(v) \mid \gcd(Mv) \mid \gcd(M^{-1}Mv) = \gcd(v).$$

Thus, $\gcd(v)$ and $\gcd(Mv)$ are associated. By Exercise 126, we obtain $\gcd(v) = \gcd(Mv)$. □

Algorithm 149 (Extended Euclidean Algorithm).

Input A column vector $v \in {}^nR$ where R is a Euclidean domain.

Output $\gcd(v)$ and a row vector $w \in R^n$ such that $wv = \gcd(v)$. Alternatively, return a unimodular matrix $Q \in \mathrm{GL}_n(R)$ such that $Qv = (\gcd(v), 0, \dots, 0)^t$.

Procedure

(a) Initialise $Q = \mathbf{1}_n$.

(b) If $v = (g, 0, \dots, 0)^t$, then return g and the first row of Q.

(c) Otherwise, choose a non-zero entry v_j of v of minimal degree.

(d) Interchange v_1 and v_j as well as the first and j^{th} row of Q.

(e) For $k = 2, \dots, n$, subtract v_k quo v_1 times v_1 from v_k and v_k quo v_1 times the first row of Q from the k^{th}.

(f) Go to step (b).

Theorem 150. *Algorithm 149 is correct and terminates.*

Proof. If the input vector is 0, the algorithm returns 0 (the correct greatest common divisor) and e_1. Else, there is at least one non-zero entry in the vector which will get swapped to position 1 in step (d). Therefore, the quotients in step (e) are well-defined. Similarly, when the termination condition in step (b) does not hold, there is a non-zero entry and the quotients are well-defined.

Let $v^{(0)}$ denote the input vector v, and let $v^{(k)}$ denote the input vector v after k iterations (that is, repetitions of steps (b) to (f)). Similarly, let Q^0 denote the matrix Q at the start of the algorithm and $Q^{(k)}$ be the same matrix after k iterations. We claim first that the invariants $Q^{(k)}v = v^{(k)}$ and $Q^{(k)} \in \mathrm{GL}_n(R)$ hold for every $k \geqslant 0$. Both claims are easy to see since they hold initially and we mimick all the transformations on v which we do in steps (d) and (e) on Q and all these transformations are elementary.

The algorithm terminates when $v^{(\ell)} = (g, 0, \dots, 0)$ for some $\ell \geqslant 0$. In that case we have

$$g = \gcd(v^{(\ell)}) = \gcd(Q^{(\ell)}v) = \gcd(v)$$

using Corollary 148. Thus, the algorithm returns the correct greatest common divisor. Moreover, by the first invariant we have $Q^{(\ell)}_{1,*}v = v^{(\ell)}_1 = g$. The algorithm thus returns the correct result.

It remains to show that the algorithm terminates. For this we remark that the degrees of the topmost entry $v^{(k)}_1$ form a strictly decreasing sequence for $k \geqslant 1$, that is,

$$\deg v^{(1)}_1 > \deg v^{(2)}_1 > \deg v^{(3)}_1 > \dots.$$

Let $k \geqslant 1$ be arbitrary such that the termination condition in step (b) does not (yet) hold and assume that we already did the swap in step (d). Doing the reductions in step (e) will replace $v^{(k)}_j$ by $v^{(k)}_j - (v^{(k)}_j \text{ quo } v^{(k)}_1)v^{(k)}_1 = v^{(k)}_j \text{ rem } v^{(k)}_1$ for all $j = 2, \dots, n$. Because $\deg(v^{(k)}_j \text{ rem } v^{(k)}_1) = 0$ or $\deg(v^{(k)}_j \text{ rem } v^{(k)}_1) < \deg v^{(k)}_1$ for each $j = 2, \dots, n$, in the next iteration we will either have reached the termination condition (if all the remainders are 0) or there will be at least one entry of strictly smaller degree which gets swapped to the topmost position. Thus, the chain is indeed strictly decreasing. Since the topmost entries (except in the case $v = 0$) are always non-zero, their degrees are natural numbers and the chain can only contain finitely many members. Thus, the algorithm must terminate after finitely many steps. □

Exercise 151. Implement the Extended Euclidean algorithm in a programming language of your choice. (It is sufficient if the implementation works for integers.)

Example 152. As an example for the application of Algorithm 149, we consider the ring $R = \mathbb{Z}$ of integers and the input vector

$$\begin{pmatrix} 15 \\ 6 \\ 10 \end{pmatrix} \in {}^3\mathbb{Z}.$$

In step (a) of Algorithm 149 we initialise Q (and v) to be

$$v = \begin{pmatrix} 15 \\ 6 \\ 10 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In step (b) we see that the termination condition is not yet reached. Thus, step (c) we search for the lowest degree entry of v. This is 6 in the second row. Thus, we swap the first two rows in step (d). Then, in step (e) we subtract 15 quo $6 = 2$ times the first row of v and Q from the second and 10 quo $6 = 1$ times the first row from the third. This yields

$$v = \begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

We go back to step (b). Again, the termination condition is false and thus, we choose again the lowest degree entry (3 in the second row) and bring it to the top row. Then we subtract 6 quo $3 = 2$ times the first row from the second and 4 quo $3 = 1$ times the first row from the third mimicking the transformations on Q. This gives us

$$v = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

Since the termination condition does still not hold, we do one more iteration. The lowest (no-zero) degree entry of v is 1 in the last row. Exchanging the first and last row and subtracting 3 quo $1 = 3$ times the first row from the last finally yields

$$v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} -1 & 1 & 1 \\ -2 & 5 & 0 \\ 4 & -5 & -3 \end{pmatrix}.$$

Here, the termination condition is reached and the algorithm returns 1 and $(-1, 1, 1)$. We can easily check that indeed

$$\begin{pmatrix} -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 15 \\ 6 \\ 10 \end{pmatrix} = -15 + 6 + 10 = 1.$$

We can also compute $\det Q = -1$ which is a unit. Thus, Q is indeed unimodular (by Theorem 91).

Exercise 153. Apply the Euclidean algorithm (Algorithm 149) to the following inputs

$$\begin{pmatrix} x^4+x^2+x+1 \\ x^3+1 \\ x^4+x^3+x^2+1 \end{pmatrix} \in {}^3\mathbb{F}_2[x]; \qquad \begin{pmatrix} 42 \\ 210 \\ 105 \end{pmatrix} \in {}^3\mathbb{Z}; \qquad \text{and} \qquad \begin{pmatrix} x^4+2x^3+x^2+x+1 \\ x^3+2x^2+2x+1 \\ x^4+x^3+x^2+2x+1 \end{pmatrix} \in {}^3\mathbb{Q}[x].$$

Exercise 154. Let R be a Euclidean domain, and let $a, b \in R$. Apply the extended Euclidean algorithm (Algorithm 149) but return Q as a whole instead of only the first row. We obtain an equation

$$\underbrace{\begin{pmatrix} s & t \\ u & v \end{pmatrix}}_{=Q} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \gcd(a,b) \\ 0 \end{pmatrix}.$$

Show that $ua = \operatorname{lcm}(a,b)$ (or, equivalently, $vb = \operatorname{lcm}(a,b)$).

Exercise 155. Let R be an Euclidean domainand $v = (v_1, \ldots, v_n)^t \in {}^nR$. Let $g = \gcd(v)$, and let $Q \in \mathrm{GL}_n(R)$ be the transformation matrix with

$$Qv = \begin{pmatrix} g \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

as computed by the Euclidean algorithm (Algorithm 149). Show that the first column of Q^{-1} is $v/g = (v_1/g, \ldots, v_n/g)^t$. Use that to prove that $\gcd(v/g) = 1$.

Application 156 (Linear Diophantine Equations). Let R be a Euclidean domain. Consider the *linear diophantine equation*

$$a_1x_1 + \ldots + a_nx_n = b$$

where $a_1, \ldots, a_n, b \in R$ and where we are looking for solutions $x_1, \ldots, x_n \in R$. We will show how to find all possible solutions.

Form the column vector $a = (a_1, \ldots, a_n)^t \in {}^nR$ and apply the Euclidean algorithm (Algorithm 149) obtaining $g = \gcd(a)$ and $Q \in \mathrm{GL}_n(R)$ such that $Qa = (g, 0, \ldots, 0)^t$. We claim that the equation has a solution if and only if $g \mid b$: For any choice of $x_1, \ldots, x_n \in R$, the left hand side of the equation is an element of $Ra_1 + \ldots + Ra_n = Rg$. Thus, a solution can only exist if $b \in Rg$; that is, if $g \mid b$. Conversely, if $b = cg$ for some $c \in R$, then $cQa = (cg, 0, \ldots, 0)^t = (b, 0, \ldots, 0)^t$. In other words, the entries of the first row of cQ are a possible solution.

It is obvious that adding linear combinations of the other rows of Q to w will also yield a solution to the equation: Write the rows of Q as $Q_{1,*}, \ldots, Q_{n,*}$. Then $Q_{j,*}a = 0$ for $j \neq 1$ and therefore for all $s_2, \ldots, s_n \in R$

$$(cQ_{1,*} + s_2Q_{2,*} + \ldots + s_nQ_{n,*})a = cQ_{1,*}a + s_2Q_{2,*}a + \ldots + s_nQ_{n,*}a = b.$$

Thus, $cQ_{1,*} + RQ_{2,*} + \ldots + RQ_{n,*}$ is contained in the set of all solutions.

Let conversely $x = (x_1, \ldots, x_n) \in R^n$ be any solution. Then

$$b = xa = (xQ^{-1})(Qa) = (xQ^{-1}) \begin{pmatrix} g \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Write $xQ^{-1} = (y_1, y_2, \ldots, y_n) \in R^n$. Then the equation implies that the first entry is $b = y_1 g$; that is, $y_1 = c$. Moreover,

$$x = \begin{pmatrix} c & y_2 & \cdots & y_2 \end{pmatrix} Q = cQ_{1,*} + y_2 Q_{2,*} + \ldots + y_n Q_{n,*} \in cQ_{1,*} + RQ_{2,*} + \ldots + RQ_{n,*}.$$

Thus, we see that $cQ_{1,*} + RQ_{2,*} + \ldots + RQ_{n,*}$ is indeed equal to the solution set.

Example 157. Consider the equation

$$2x + 3y + 7z = 5$$

over the integers $\mathbb{Z}$. Computing the greatest common divisor of $2, 3, 7$ with Algorithm 149 yields

$$\begin{pmatrix} -1 & 1 & 0 \\ -3 & 2 & 0 \\ -5 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{where} \quad Q = \begin{pmatrix} -1 & 1 & 0 \\ -3 & 2 & 0 \\ -5 & 1 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{Z})$$

is unimodular. Since $1 \mid 5$, we find that the solutions are

$$\begin{pmatrix} x & y & z \end{pmatrix} = 5 \begin{pmatrix} -1 & 1 & 0 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} -3 & 2 & 0 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} -5 & 1 & 1 \end{pmatrix};$$

or, using s, t as arbitrary integers, we can write this as

$$x = -5 - 3s - 5t, \qquad y = 5 + 2s + t, \qquad \text{and} \qquad z = t.$$

It is easy to check that this indeed solves the equation.

Remark 158 (Syzygy Module). Let R be a Euclidean domain, and let $a = (a_1, \ldots, a_n)^t \in {}^nR$. The *module of syzygies* of a is

$$\mathrm{Syz}(a) = \{w = (w_1, \ldots, w_n) \in R^n \mid wa = w_1 a_1 + \ldots + w_n a_n = 0\}.$$

We can see that this is actually just a special case of Application 156 with the right hand side $b = 0$. Thus, we obtain that

$$\mathrm{Syz}(a) = RQ_{2,*} + \ldots + RQ_{n,*}$$

where $Q \in \mathrm{GL}_n(R)$ is the transformation matrix computed by Algorithm 149 and $Q_{1,*}, \ldots, Q_{n,*}$ are its rows. Since Q is unimodular, its rows must be linearly independent (because the determinant is non-zero). Hence, $Q_{2,*}, \ldots, Q_{n,*}$ is actually a basis of $\mathrm{Syz}(a)$.

Remark 159. The approach of Application 156 can be employed in more general situations: Let R be a Euclidean domain and let M be an R-module. Consider the linear equation

$$a_1 \bullet x_1 + \ldots + a_n \bullet x_n = b$$

where $a_1, \ldots, a_n \in R$, $b \in M$ and we look for solutions $x_1, \ldots, x_n \in M$. Using the notation from Remark 96, we can rewrite the problem as $a \bullet x = b$ where $a = (a_1, \ldots, a_n) \in R^n$ and $x = (x_1, \ldots, x_n)^t \in {}^nM$. As before we apply Algorithm 149 to a^t obtaining $g = \gcd(a)$ and $Q^t \in \mathrm{GL}_n(R)$ such that $Q^t a^t = (g, 0, \ldots, 0)^t$. Thus, by Exercise 97 we can rewrite the equation as

$$b = ax = aQQ^{-1}x = \begin{pmatrix} g & 0 & \cdots & 0 \end{pmatrix} Q^{-1}x.$$

Setting $y = (y_1, \ldots, y_n)^t = Q^{-1}x$, we obtain that $b = g \bullet y_1$; and there are no conditions on $y_2, \ldots, y_n$. Assume that we can solve the single variable equation $g \bullet y_1 = b$; that is, that we find (one or all) $c \in M$ such that $g \bullet c = b$, then we can extend that to solutions of the original equation by just leaving $y_2, \ldots, y_n$ as variables and setting $x = Q(c, y_2, \ldots, y_n)^t$.

Example 160. We solve the differential system

$$y''' - y - z'' + z' = x$$

for $y, z \in C^\infty(\mathbb{R})$ (where x means the function $f(x) = x$). Modelling this as operator equation as in Example 38 and using the matrix notation from Remark 96, we obtain

$$\begin{pmatrix} \partial^3 - 1 & -\partial^2 + \partial \end{pmatrix} \bullet \begin{pmatrix} y \\ z \end{pmatrix} = x.$$

We apply the Euclidean algorithm to $a = (\partial^3 - 1, -\partial^2 + \partial)^t \in {}^2\mathbb{R}[\partial]$ obtaining

$$Q = \begin{pmatrix} 1 & \partial + 1 \\ \partial & \partial^2 + \partial + 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}[\partial]) \qquad \text{with} \qquad Q \begin{pmatrix} \partial^3 - 1 \\ -\partial^2 + \partial \end{pmatrix} = \begin{pmatrix} \partial - 1 \\ 0 \end{pmatrix}.$$

Thus, the original equation becomes

$$x = \begin{pmatrix} \partial^3 - 1 \\ -\partial^2 + \partial \end{pmatrix} \bullet \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \partial^3 - 1 \\ -\partial^2 + \partial \end{pmatrix} Q^t (Q^{-1})^t \bullet \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \partial - 1 & 0 \end{pmatrix} (Q^{-1})^t \bullet \begin{pmatrix} y \\ z \end{pmatrix}.$$

Let us denote $(Q^{-1})^t (y, z)^t = (\tilde{y}, \tilde{z})$. Then the equation reads now

$$\tilde{y}' - \tilde{y} = x$$

while there is no restriction on $\tilde{z}$. We solve for $\tilde{y}$ by first computing an integrating factor μ such that $\mu' = -\mu$ which implies $\mu = e^{-x}$. Thus the equation becomes $e^{-x} x = e^{-x} \tilde{y}' - e^{-x} \tilde{y} = (e^{-x} \tilde{y})'$ which leads to

$$e^{-x} \tilde{y} = \int x e^{-x} \, dx = -x e^{-x} + \int e^{-x} \, dx = -x e^{-x} - e^{-x} + C$$

(using partial integration with $u = x$ and $dv = e^{-x} \, dx$) where $C \in \mathbb{R}$ is an arbitrary constant. Thus, $\tilde{y} = -x - 1 + C e^x$ (while $\tilde{z} = f$ could be just any function in $C^\infty(\mathbb{R})$). We can now compute the solution in terms of the original variables obtaining

$$\begin{pmatrix} y \\ z \end{pmatrix} = Q^t \bullet \begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} 1 & \partial \\ \partial + 1 & \partial^2 + \partial + 1 \end{pmatrix} \bullet \begin{pmatrix} -x - 1 + C e^x \\ f \end{pmatrix} = \begin{pmatrix} -x - 1 + C e^x + f' \\ -2 + 2 C e^x - x + f'' + f' + f \end{pmatrix}$$

as the set of all solutions.

11 The Hermite Normal Form

Definition 161 (Hermite Normal Form). Let R be a Euclidean domain. A matrix $A = (a_{ij})_{ij} \in {}^m R^n$ is in (row) *Hermite normal form* (HNF) if there exist column indices $1 \leqslant j_1 < j_2 < \ldots < j_m \leqslant n$ such that for all $i = 1, \ldots, m$

(a) $a_{i j_i} \neq 0$,

(b) $a_{ik} = 0$ for $k < j_i$ (that is, A is in row echelon form), and

(c) $a_{kj_i} = 0$ or $\deg a_{ij_i} > \deg a_{kj_i}$ for $k < i$ (entries above the pivots have smaller degree).

The entries $a_{1j_1}, \ldots, a_{mj_m}$ are called the *pivots* of A and $j_1, \ldots, j_m$ are the pivot indices.

Remark 162. By Exercise 122 being associated is an equivalence relation. We can therefore pick a set of representatives for each equivalence class. For instance, for integers $\mathbb{Z}$ one usually chooses the absolute value, while for polynomials $F[x]$ over a field one usually chooses the monic polynomials[10]. Let $|a|$ denote the representative of $a \in R$. Then Definition 161 is usually extended by

(d) $a_{ij_i} = |a_{ij_i}|$ (that is, the pivots are the representatives of their class).

Moreover, assume that we can make the remainders with respect to to Euclidean division in R unique. For example, with the integers $\mathbb{Z}$ one can always choose the positive remainder[11]; while for polynomials $F[x]$ over a field the remainder is unique anyways by Exercise 137. In that case, we replace property (c) by

(c′) $a_{kj_i} = a_{kj_i} \operatorname{rem} a_{ij_i}$ for $k < i$ (entries above a pivot are reduced with respect to the pivot).

Remark 163. Some authors define the Hermite normal form to be a lower row echelon form; that is, with the pivot indices fulfilling $j_1 > \ldots > j_m$. Moreover, some authors use row indices instead of column indices obtaining a column echelon form. Since R is commutative, we can always switch between row and column Hermite normal form by simply transposing all the matrices.

Exercise 164. Prove that for a field F the Hermite normal form is the same as the reduced row echelon form.

Definition 165 (Row Equivalence). Two matrices A and $B \in {}^mR^n$ are said to be *row equivalent* if there exists a unimodular matrix $U \in \mathrm{GL}_m(R)$ such that $A = UB$.

Exercise 166. Show that row equivalence is indeed an equivalence relation.

Theorem 167. *Let $A \in {}^mR^n$ be in Hermite normal form.*

(a) The rows of A are linearly independent.

(b) Assume that we have fixed representatives of associate classes and unique remainders as in Remark 162. If there is a matrix $B \in {}^mR^n$ in Hermite normal form which is row equivalent to A, then $A = B$ (that is, a Hermite normal form is a unique representative for its class of row equivalent matrices).

Proof. Let $A = (a_{ij})_{ij}$ with pivot indices $j_1, \ldots, j_m$. Denote the rows of A by $A_{1,*}, \ldots, A_{m,*}$. Part (a) follows essentially because A is in row echelon form: If there are $s_1, \ldots, s_n$ such that $s_1 A_{1,*} + \ldots + s_m A_{m,*} = 0$, then at the j_1th position we have $s_{j_1} a_{1j_1} = 0$ since $a_{kj_1} = 0$ for $k > 1$ which follows from $j_1 < j_k$ and property (b) of Definition 161. Since $a_{1j_1} \neq 0$ because of property (a), we obtain $s_1 = 0$. Inductively, we can now show $s_2 = \ldots = s_m = 0$.

In order to prove part (b), write B as $B = (b_{ik})_{ik}$ with pivot indices $k_1, \ldots, k_m$ and with rows $B_{1,*}, \ldots, B_{m,*}$; and let $U = (U_{ij})_{ij} \in \mathrm{GL}_m(R)$ be such that $A = UB$. We are first going to prove that $j_1 = k_1, \ldots, j_m = k_m$. Assume that this was not the case. Then there exists a minimal row

[10] That is, polynomials with leading coefficient being 1.

[11] Some authors prefer to use the remainder with minimal absolute value instead.

index ℓ such that $j_1 = k_1, \ldots, j_{\ell-1} = k_{\ell-1}$ but $j_\ell \neq k_\ell$. Assume without loss of generality that $j_\ell < k_\ell$ (otherwise, we switch the roles of A and B). We have

$$A_{\ell,*} = U_{\ell,*}B = u_{\ell 1}B_{1,*} + \ldots + u_{\ell m}B_{m,*}.$$

Since $a_{\ell i} = 0$ for $i < j_\ell$, none of the rows $B_{1,*}, \ldots, B_{\ell-1,*}$ can contribute to that sum: If, for example, $\nu < \ell$ was minimal such that $u_{\ell\nu} \neq 0$, then the k_νth entry of $U_{\ell,*}B$ is $u_{\ell\nu}b_{\nu k_\nu} \neq 0$ (by property (a) of Definition 161 since B is in Hermite normal form). However, because $k_\nu = j_\nu < j_\ell$ and thus $a_{\ell k_\nu} = 0$ this cannot happen. On the other hand, $b_{\lambda j_\ell} = 0$ for $\lambda \geqslant \ell$ since $k_\lambda \geqslant k_\ell > j_\ell$. This implies $a_{\ell j_\ell} = 0$ which contradicts the assumption. Thus, the pivot indices of A and B must be the same.

Next, we show that the pivots are the same. For $i = 1, \ldots, m$ it is easly seen that $a_{ij_i} = u_{ii}b_{ij_i}$ by an argument similar to the one above (rows with smaller pivot index cannot contribute and rows with larger pivot index will not affect the j_ith entry). Thus b_{ij_1} divides a_{ij_i}. Switching the roles of A and B (using $B = U^{-1}A$) we obtain that also a_{ij_i} divides b_{ij_i}. Thus, the pivots of A are associated to their respective counterparts in B. Since we chose a unique representative, it follows that $a_{ij_i} = b_{ij_i}$. This shows also that $u_{ii} = 1$ and $u_{\nu i} = 0$ for $\nu < i$.

Now we prove that $A_{i,*} = B_{i,*}$ for $i = 1, \ldots, m$. Since $u_{ii} = 1$ and $u_{\nu i} = 0$ for $\nu < i$, we have

$$A_{i,*} = B_{i,*} + u_{i,i+1}B_{i+1,*} + \ldots + u_{im}B_{i,*}.$$

Consider the j_{i+1}th entry of this row. It must be $a_{i,j_{i+1}} = b_{i,j_{i+1}} + u_{i,i+1}b_{i+1,j_{i+1}}$ since the other rows $B_{i+2,*}, \ldots, B_{m,*}$ of B have 0 in that position. Since $\deg b_{i,j_{i+1}} < \deg b_{i+1,j_{i+1}}$ by property (c) of Definition 161, we see that $b_{i,j_{i+1}}$ is the remainder of $a_{i,j_{i+1}}$ of division by $b_{i+1,j_{i+1}}$ (or we simply use property (c′)); that is, $b_{i,j_{i+1}} = a_{i,j_{i+1}} \operatorname{rem} b_{i+1,j_{i+1}} = a_{i,j_{i+1}} \operatorname{rem} a_{i+1,j_{i+1}} = a_{i,j_{i+1}}$ since the pivots are equal and by property (c′) again. This implies further $u_{i,i+1}b_{i+1,j_{i+1}} = 0$ and thus $u_{i,i+1} = 0$ by property (a). Inductively, we can now prove that $u_{i,i+2} = \ldots = u_{im} = 0$. Thus, $A_{i,*} = B_{i,*}$ as required. □

Exercise 168. Let R be a Euclidean domain; and let $a, b \in R$ with $b \neq 0$. Show that if $b \mid a$, then $a \operatorname{rem} b = 0$.

Algorithm 169 (Hermite Division).

Input A matrix $H = (h_{ij})_{ij} \in {}^mR^n$ in Hermite normal form with pivot indices $j_1 < \ldots < j_m$; a row vector $w = (w_1, \ldots, w_n) \in R^n$.

Output A row vector matrix $q \in R^m$ and a row vector $r \in R^n$ such that $w = qH + r$ and such that $r_{j_i} = 0$ or $\deg r_{j_i} < \deg h_{ij_i}$ for all $i = 1, \ldots, m$.

Procedure

```
(a) Initialise $q \leftarrow 0$ and $r \leftarrow w$.
(b) For $i = 1, \ldots, m$ do
    (1) Let $q_i \leftarrow r_{j_i}$ quo $h_{ij_i}$.
    (2) Update $r \leftarrow r - q_i H_{i,*}$ (where $H_{i,*}$ is the $i$th row of $H$).
(c) Return $q$ and $r$.
```

Theorem 170. *(a) Algorithm 169 is correct and terminates.*

(b) A vector $w \in R^n$ is in the row space of a Hermite normal form $H \in {}^mR^n$ if and only if Algorithm 169 returns $r = 0$.

Proof. For part (a), please note first that the loop in step (b) of Algorithm 169 is over a finite range of numbers. Thus, the procedure always terminates. Following the loop, it is easy to see that

$$r = w - q_1 H_{1,*} - q_2 H_{2,*} - \ldots - q_m H_{m,*} = w - qH;$$

that is, indeed $w = qH + r$. Similarly, after the first iteration of the loop we have

$$r_{j_1} = w_{j_1} - (w_{j_1} \text{ quo } h_{1j_1})h_{1j_1} = w_{j_1} \text{ rem } h_{1j_1}.$$

Thus, $r_{j_1} = 0$ or $\deg r_{j_1} < \deg h_{1j_1}$. This does not change during the following iteration since $h_{kj_1} = 0$ for $k \geqslant 2$ by Definition 161. Similarly, in the next iteration we establish the required property for r_{j_2}. Going through all of the loop, it is easy to see that this works for all r_{j_i} where $i = 1, \ldots, m$.

For part (b), if $w = qH$, then obviously $w \in R^m H$. Assume now that $r \neq 0$. We want to show that there is no $v \in R^n$ such that $w = vH$. Since in this case we had $(v - q)H = r$, it is sufficient to show that $r \notin R^n H$. Let $k = 1, \ldots, n$ be minimal such that $r_k \neq 0$. If $k < j_1$, then all entries of the k^{th} column of H are zero by Definition 161; and we see that $r \notin R^n$. Otherwise, let $\ell = 1, \ldots, m$ be maximal such that $j_\ell \leqslant k$. If $\ell = k$, then first note that we cannot not use any rows $H_{i,*}$ with $i < \ell$ to generate r since all $r_{j_i} = 0$ for such i. Similarly, rows $H_{\nu,*}$ with $\nu > \ell$ cannot contribute to $r_k = r_{j_\ell}$. Thus, only $H_{\ell,*}$ could be used to generate $r_k = r_{j_\ell}$. However, since $0 \neq r_k = r_{j_\ell} = w_{j_\ell} \text{ rem } h_{\ell j_\ell}$, Exercise 168 implies that $h_{\ell j_\ell}$ does not divide r_k. Thus, also $H_{\ell,*}$ cannot be used and we see that $r \notin R^n H$. Finally, if $k \notin \{j_1, \ldots, j_m\}$, then the rows $R_{i,*}$ with $j_i < k$ cannot be used to generate r since the entries $r_{j_i} = 0$; and the rows $H_{\nu,*}$ with $j_\nu > k$ cannot be used to generate r_k. Thus, also here $r \notin R^m H$. □

Example 171. Consider the matrix

$$H = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \in {}^3\mathbb{Z}^4$$

which is in Hermite normal form. Let the two row vectors

$$\begin{pmatrix} 1 & 6 & 2 & -2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 2 & 0 & -1 & 13 \end{pmatrix} \in \mathbb{Z}^4$$

be given. Carrying out the division in Algorithm 169 for $(1, 6, 2, -2)$ we obtain

$$\begin{pmatrix} 1 & 6 & 2 & -2 \end{pmatrix} - 1 \cdot H_{1,*} = \begin{pmatrix} 0 & 5 & 2 & -5 \end{pmatrix}$$

since $1 \text{ quo } 1 = 1$. Now, $5 \text{ quo } 2 = 2$ and have

$$\begin{pmatrix} 0 & 5 & 2 & -5 \end{pmatrix} - 2 \cdot H_{2,*} = \begin{pmatrix} 0 & 1 & 0 & -5 \end{pmatrix}.$$

Finally, $-5 \text{ quo } 7 = -1$ (using our convention to choose the positive remainders) and we obtain

$$\begin{pmatrix} 0 & 1 & 0 & -5 \end{pmatrix} + 1 H_{3,*} = \begin{pmatrix} 0 & 1 & 0 & 2 \end{pmatrix}.$$

In total, we have computed that

$$\begin{pmatrix}1 & 6 & 2 & -2\end{pmatrix} = \begin{pmatrix}1 & 2 & -1\end{pmatrix} H + \begin{pmatrix}0 & 1 & 0 & 2\end{pmatrix}.$$

Thus, by Theorem 170 $(1, 6, 2, -2) \notin \mathbb{Z}^3 H$.

Applying Algorithm 169 to $(2, 0, -1, 13)$ yields

$$\begin{pmatrix}2 & 0 & -1 & 13\end{pmatrix} = \begin{pmatrix}2 & -1 & 1\end{pmatrix} H.$$

Thus, $(2, 0, -1, 13) \in \mathbb{Z}^3 H$.

Algorithm 172 (Hermite Normal Form)*.*

Input A matrix $A \in {}^mR^n$ where R is a Euclidean domain.

Output A unimodular matrix $Q \in \mathrm{GL}_m(R)$ and a matrix $H \in {}^rR^n$ in Hermite normal form such that $r \leqslant m$ and

$$QA = \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix}.$$

Procedure

(a) If $m = 0$ or $n = 0$ then stop and return the identity matrix $Q = \mathbf{1}_m$ and an empty matrix $H \in {}^0R^n$.

(b) If the first column of A is zero, that is, if $A = (0, \tilde{A})$ where $\tilde{A} \in {}^mR^{n-1}$; then apply the Hermite normal form algorithm recursively to $\tilde{A}$ obtainig $Q \in \mathrm{GL}_m(R)$ and a Hermite normal form $\tilde{H} \in {}^rR^{n-1}$, Return Q and $H = (0, \tilde{H}) \in {}^rR^n$.

(c) Otherwise:

(1) Apply the Euclidean algorithm (Algorithm 149) to the first column $A_{*,1}$ of A computing $g = \gcd(A_{*,1})$ and a unimodular matrix $U \in \mathrm{GL}_m(R)$ such that $UA_{*,1} = (g, 0, \ldots, 0)^t$. If we have unique representatives (see Remark 162), then choose g such that it is the unique representative of its associate class (if necessary multiply g and the first row of U by a unit).

(2) Partition UA as

$$UA = \begin{pmatrix} g & w \\ 0 & \tilde{A} \end{pmatrix}$$

where $\tilde{A} \in {}^{m-1}R^{n-1}$ and $w \in R^{n-1}$.

(3) Apply the Hermite normal form procedure recursively to $\tilde{A}$ obtainig a unimodular $\tilde{Q} \in \mathrm{GL}_{m-1}(R)$ and a Hermite normal form $\tilde{H} = (\tilde{h}_{ij})_{ij} \in {}^rR^{n-1}$.

(4) Apply Hermite division (Algorithm 169) in order to compute $w = \tilde{q}\tilde{H} + v$ with $\tilde{q} \in R^r$ and $v \in R^{n-1}$. Let $q = (\tilde{q}, 0) \in R^{m-1}$.

(5) Return

$$H = \begin{pmatrix} g & v \\ 0 & \tilde{H} \end{pmatrix} \in {}^{r+1}R^n \qquad \text{and} \qquad Q = \begin{pmatrix} 1 & -q\tilde{Q} \\ 0 & \tilde{Q} \end{pmatrix} U.$$

Remark 173. Some authors refer to

$$\begin{pmatrix} H \\ \mathbf{0} \end{pmatrix}$$

with H computed by Algorithm 172 as the Hermite normal form of A. We will sometimes adopt this terminology as a convenient short hand for the more precise formulation chosen in the algorithm.

Remark 174. In Algorithm 172, instead of constructing the transformation matrix $Q \in \mathrm{GL}_m(R)$ explicitly as a matrix product, we might also just mimick the elementary row transformations (and partitions) we apply to A on an identity matrix. In fact, if we apply the Hermite normal form algorithm (without explicitly computing Q) to the matrix $(A, \mathbf{1}_m)$, then we will obtain a matrix

$$\begin{pmatrix} H & \tilde{Q} \\ \mathbf{0} & \end{pmatrix} \qquad \text{such that} \qquad \tilde{Q}A = \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix}$$

and where H is in Hermite normal form and $\tilde{Q} \in \mathrm{GL}_m(R)$ is unimodular. (Note that $\tilde{Q}$ computed in this way is not necessarily exactly the same as Q computed by Algorithm 172; but it has the same properties.)

Remark 175. In step (a) of Algorithm 172 we use empty matrices for convenience. However, if one wishes to avoid that, one might replace this step by the following steps

(a′) If $A = \mathbf{0}_{m \times n}$, then return $Q = \mathbf{1}_m$ and and no H.

(a″) If $n = 1$ (that is, if A is a single column), then apply the Euclidean Algorithm 149 to A obtaining a matrix $U \in \mathrm{GL}_m(R)$ and $g = \gcd(A)$. Let $u \in R^*$ be such that ug is the unique representative of g (see Remark 162) and return the Hermite form $H = (g) \in {}^1R^1$ and $Q = \mathrm{diag}(u, \mathbf{1}_{n-1})U \in \mathrm{GL}_n(R)$.

(a‴) If $m = 1$ (that is, A is a single row), then let k be the minimal column index such that $A_{1r} \neq 0$. Let $u \in R^*$ be such that uA_{1r} is the unique representative of its class (see Remark 162) and return $Q = (q) \in \mathrm{GL}_1(R)$ and $H = qA$.

It is easy to see that in each of the three steps the matrices which are returned fulfil the output conditions of Algorithm 172.

Theorem 176. *Algorithm 172 is correct and terminates.*

Proof. It is easy to see that the algorithm terminates because in every recursive call the number of rows or the number of columns of the argument decreases. This can only happen finitely often.

There are three cases in which the algorithm returns a value. We will go through all of them and prove that they are correct. In step (a) of Algorithm 172 the base case of the reduction is handled. An empty matrix is obviously in Hermite form because all the conditions are trivially fulfilled. Moreover, $Q = \mathbf{1}_m$ is obviously unimodular and QA (which is again empty) fulfils the output conditions. (See also Remark 175 for an approach without empty matrices.)

In step (b) We have (by the rules for multiplying block matrices)

$$QA = \begin{pmatrix} 0 & Q\tilde{A} \end{pmatrix} = \begin{pmatrix} 0 & \tilde{H} \\ 0 & \mathbf{0} \end{pmatrix}$$

where $\tilde{H}$ is in Hermite normal form by the recursion. Thus, we only have to show that also $H = (0, \tilde{H})$ is in Hermite normal form. If $j_1, \ldots, j_r$ are the pivot indices of $\tilde{H}$, then we claim that $j'_k = j_k + 1$ for $k = 1, \ldots, r$ are the pivot indices of H: We can easily see that the properties (a), (b), and (c) hold for H since they hold for $\tilde{H}$ and the added first column contains only zeroes. Moreover, the properties of Remark 162 do also hold for H because they hold for $\tilde{H}$.

Finally, in step (c), we first note that the matrix Q given in substep (c.5) is the product of unimodular matrices: U in step (c.1) is unimodular, $\tilde{Q}$ in step (c.3) is unimodular, too, and we can write the left factor of Q as

$$\begin{pmatrix} 1 & -q\tilde{Q} \\ 0 & \tilde{Q} \end{pmatrix} = \begin{pmatrix} 1 & -q \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{pmatrix}$$

where both matrices are obviously unimodular by Exercise 111. Moreover, with

$$\tilde{Q}\tilde{A} = \begin{pmatrix} \tilde{H} \\ \mathbf{0} \end{pmatrix}$$

which we get from the recursion in step (c.3), we indeed obtain that

$$QA = \begin{pmatrix} 1 & -q\tilde{Q} \\ 0 & \tilde{Q} \end{pmatrix} UA = \begin{pmatrix} 1 & -q\tilde{Q} \\ 0 & \tilde{Q} \end{pmatrix} \begin{pmatrix} g & w \\ 0 & \tilde{A} \end{pmatrix} = \begin{pmatrix} g & v \\ 0 & \tilde{H} \\ 0 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix}.$$

Here, we used that $v = w - \tilde{q}\tilde{H} = w - (\tilde{q}, 0)\tilde{Q}\tilde{A} = w - q\tilde{Q}\tilde{A}$. Thus, the matrix Q fulfils the output condition.

It remains to check that

$$H = (h_{ij})_{ij} = \begin{pmatrix} u & v \\ 0 & \tilde{H} \end{pmatrix}$$

is indeed in Hermite normal form. By the recursive nature of the algorithm, we can assume that $\tilde{H} = (\tilde{h}_{ij})_{ij}$ is in Hermite normal form (recalling that we have already shown that this is true in the base case in step (a)). The pivot of the first row of H is obviously g at position $(1, 1)$. Since $\tilde{H}$ starts at the second column of H, the pivots which it contributes must be to the right of g. More precisely, if $\tilde{j}_1 < \ldots < \tilde{j}_r$ are the pivot indices of $\tilde{H}$; then $1 < \tilde{j}_1 + 1 < \ldots < \tilde{j}_r + 1$ are the pivot indices of H; we denote them by $j_1 = 1$ and $j_{k+1} = \tilde{j}_k + 1$ for $k = 1, \ldots, r$. Moreover, by the properties of Algorithm 169, we see that $v_{\tilde{j}_k}$ is a remainder of division by $\tilde{h}_{k\tilde{j}_k}$. Thus, for $\ell = 1, \ldots, r$ we obtain $h_{1j_{\ell+1}} = v_{\tilde{j}_\ell}$ is indeed a remainder of division by $h_{\ell+1, j_{\ell+1}} = \tilde{h}_{\ell, \tilde{j}_\ell}$. In total, all properties of Definition 161 are fulfilled. Moreover, by taking the unique representative of the greatest common divisor in step (c.1), we also fulfil property (d) (in Remark 162). □

Example 177. We use Algorithm 172 in order to compute the Hermite normal form of

$$A = \begin{pmatrix} -3 & 9 & 1 & 4 \\ 3 & -3 & 7 & -7 \\ -10 & 6 & 3 & -1 \\ -7 & 9 & 18 & -11 \end{pmatrix} \in {}^4\mathbb{Z}^4.$$

We will use superscripts in order to distinguish the matrices in the different recursive calls. Let $A^{(1)} = A$ for the original input. The first column of $A^{(1)}$ is

$$\begin{pmatrix} -3 \\ 3 \\ -10 \\ -4 \end{pmatrix}$$

which is non-zero. Thus, we are in case item (c) of Algorithm 172. Applying the Euclidean algorithm (Algorithm 149) yields the greatest common divisor $g^{(1)} = 1$ with the transformation matrix

$$U^{(1)} = \begin{pmatrix} -7 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 10 & 0 & -3 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix} \in \mathrm{GL}_4(\mathbb{Z}).$$

We obtain

$$U^{(1)}A^{(1)} = \left(\begin{array}{c|ccc} 1 & -51 & -1 & -30 \\ \hline 0 & 6 & 8 & -3 \\ 0 & 72 & 1 & 43 \\ 0 & 12 & 16 & -6 \end{array}\right) = \begin{pmatrix} g^{(1)} & w^{(1)} \\ 0 & \tilde{A}^{(1)} \end{pmatrix}.$$

We now call the algorithm recursively with $A^{(2)} = \tilde{A}^{(1)}$. Since the first column is not zero, we have again to apply the Euclidean algorithm (Algorithm 149) which yields the greatest common divisor $g^{(2)} = 6$, the transformation

$$U^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ -12 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad U^{(2)}A^{(2)} = \left(\begin{array}{c|cc} 6 & 8 & -3 \\ \hline 0 & -95 & 79 \\ 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} g^{(2)} & w^{(2)} \\ 0 & \tilde{A}^{(2)} \end{pmatrix}.$$

We continue recursively with $A^{(3)} = \tilde{A}^{(2)}$. Although we can see that matrix is already in Hermite normal form (except for the sign), we follow the algorithm. The first column of $A^{(3)}$ is non-zero and computing the greatest common divisor with the Euclidean algorithm (Algorithm 149) gives 95 with transformation

$$U^{(3)} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \left(\begin{array}{c|c} 95 & -79 \\ \hline 0 & 0 \end{array}\right) = \begin{pmatrix} g^{(3)} & w^{(3)} \\ 0 & \tilde{A}^{(3)}. \end{pmatrix}$$

We continue recursively with $A^{(4)} = \tilde{A}^{(3)}$. Since the first (and only) column of $A^{(4)}$ is zero, we are in case (b) of Algorithm 172. Thus, we skip the first column and continue recursively with the empty 1-by-0 matrix $A^{(5)}$. Now we are in case (a) of Algorithm 172; and we return the empty 0-by-0 matrix $H^{(5)}$ and the identity $Q^{(5)} = (1)$. Going up the recursive calls, we have now $\tilde{H}^{(4)} = H^{(5)}$ and $H^{(4)} = (0, \tilde{H}^{(4)})$ (which is the empty 0-by-1 matrix) and $Q^{(4)} = Q^{(5)}$. Up one level, we obtain $\tilde{H}^{(3)} = H^{(4)}$ and $\tilde{Q}^{(3)} = Q^{(4)}$. We now have to divide $w^{(3)}$ by $\tilde{H}^{(3)}$ using Algorithm 169. As $\tilde{H}^{(3)}$ is still the empty matrix, this step yields an empty row vector $\tilde{q}^{(3)}$ and $v^{(3)} = w^{(3)} = (-79)$. We otbain the row vector $q^{(3)} = (\tilde{q}^{(3)}, 0) = (0)$,

$$H^{(3)} = \begin{pmatrix} g^{(3)} & v^{(3)} \\ 0 & \tilde{H}^{(3)} \end{pmatrix} = \begin{pmatrix} 95 & -79 \end{pmatrix}, \qquad \text{and} \qquad Q^{(3)} = \begin{pmatrix} 1 & -q^{(3)}\tilde{Q}^{(3)} \\ 0 & \tilde{Q}^{(3)} \end{pmatrix} U^{(3)} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

One further level up, we have now $\tilde{H}^{(2)} = H^{(3)}$ and $\tilde{Q}^{(2)} = Q^{(3)}$. Dividing $w^{(2)}$ by $\tilde{H}^{(2)}$ using Algorithm 169 gives us $\tilde{q}^{(2)} = 0$ and $v^{(3)} = (8, -3)$. Thus, $q^{(2)} = (0, 0)$,

$$H^{(2)} = \begin{pmatrix} g^{(2)} & v^{(2)} \\ 0 & \tilde{H}^{(2)} \end{pmatrix} = \left(\begin{array}{c|cc} 6 & 8 & -3 \\ \hline 0 & 95 & -79 \end{array}\right),$$

as well as

$$Q^{(2)} = \begin{pmatrix} 1 & -q^{(2)}\tilde{Q}^{(2)} \\ 0 & \tilde{Q}^{(2)} \end{pmatrix} U^{(2)} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right) U^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 12 & -1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

Finally, we are back in the uppermost level where $\tilde{H}^{(1)} = H^{(2)}$ and $\tilde{Q}^{(1)} = Q^{(2)}$. We apply Hermite division (Algorithm 169) to $w^{(1)}$ and $\tilde{H}^{(1)}$ which gives $\tilde{q}^{(1)} = (-9, 0)$ and $v^{(1)} = (3, 71, -57)$. Then $q^{(1)} = (\tilde{q}^{(1)}, 0) = (-9, 0, 0)$,

$$H^{(1)} = \begin{pmatrix} g^{(1)} & v^{(1)} \\ 0 & \tilde{H}^{(1)} \end{pmatrix} = \left(\begin{array}{c|ccc} 1 & 3 & 71 & -57 \\ 0 & 6 & 8 & -3 \\ 0 & 0 & 95 & -79 \end{array}\right),$$

and

$$Q^{(1)} = \begin{pmatrix} 1 & -q^{(1)}\tilde{Q}^{(1)} \\ 0 & \tilde{Q}^{(1)} \end{pmatrix} U^{(1)} = \left(\begin{array}{c|ccc} 1 & 9 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 12 & -1 & 0 \\ 0 & -2 & 0 & 1 \end{array}\right) U^{(1)} = \begin{pmatrix} 2 & 9 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 12 & 3 & 0 \\ -1 & -2 & -1 & 1 \end{pmatrix}.$$

Having reached the top-most level of the recursive calls, the algorithm now returns $H^{(1)}$ and $Q^{(1)}$.

Remark 178 (Iterative Hermite Normal Form Computation). While the recursive Algorithm 172 for computing the Hermite normal form is easy to understand and easy to be proved correct, in practical implementations one would usually prefer the following iterative version.

Input A matrix $A = (a_{ij})_{ij} \in {}^{m}R^{n}$ where R is a Euclidean domain.

Output The Hermite normal form $H \in {}^{r}R^{n}$ of A where $r \leqslant m$.

Procedure

```
(a) Initialise $r \leftarrow 1$ and $c \leftarrow 1$.
(b) While $r \leqslant m$ and $c \leqslant n$ do
    (1) Let $P = \{i \mid r \leqslant i \leqslant m \text{ and } a_{ic} \neq 0\}$.
    (2) If $P = \varnothing$, then set $c \leftarrow c + 1$ and continue the loop in step (b).
    (3) Else, if $P = \{i\}$, then:
        (i) Swap the $r^{\text{th}}$ and the $i^{\text{th}}$ row of $A$.
        (ii) Multiply the $r^{\text{th}}$ row of $A$ by a unit such that $a_{rc} = |a_{rc}|$.
        (iii) For $k = 1, \ldots, r - 1$ subtract ($a_{kc}$ quo $a_{rc}$) times the $r^{\text{th}}$ row of $A$ from the $k^{\text{th}}$ row.
        (iv) Set $r \leftarrow r + 1$ and $c \leftarrow c + 1$, and continue the loop in step (b).
```

(4) Else:

(i) Let $i \in P$ be such that $\deg a_{ic}$ is minimal.
(ii) Swap the r^{th} and the i^{th} row of A.
(iii) For $k = r+1, \ldots, m$ subtract (a_{kc} quo a_{rc}) times the r^{th} row of A from the k^{th} row.
(iv) Continue the loop in step (b).

(c) Remove all rows which are zero from A and return A.

(We have left out computation of the transformation matrix; it can be obtained as explained in Remark 174.)

Exercise 179. Prove that the algorithm explained in Remark 178 does indeed compute the Hermite normal form of A.

Exercise 180. Apply the algorithm in Remark 178 to the matrix

$$\begin{pmatrix} 0 & x^2 & x & 0 \\ x^3+x^2+1 & 1 & 1 & 0 \\ 0 & x^3+x & 0 & 0 \\ 1 & 1 & x^2+x+1 & x \end{pmatrix} \in {}^4\mathbb{F}_2[x]^4.$$

Exercise 181. Implement the algorithm in Remark 178 in a programming language of your choice. (It is sufficient if it works for integer matrices.)

Remark 182. The Hermite normal form is implemented in several computer algebra systems. Examples include:

Maple The command is called HermiteForm and is located in the LinearAlgebra package. It works for both integer and polynomial matrices. We do an example for integer matrices:

```
with(LinearAlgebra):
A := RandomMatrix(4,3, generator=-9..9);
```

$$A := \begin{bmatrix} 6 & 2 & 4 \\ 2 & 0 & -7 \\ -4 & 0 & -7 \\ 9 & -1 & -8 \end{bmatrix}$$

```
U, H := HermiteForm(A, output=['U', 'H']);
```

$$U, H := \begin{bmatrix} 3 & -13 & 9 & 5 \\ 3 & -11 & 8 & 4 \\ 5 & -26 & 17 & 10 \\ 7 & -36 & 24 & 14 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

```
Equal(U.A, H);
```

true

Mathematica Here, we have the command HermiteDecomposition. This seems to work only for integer matrices. An example would be:

```
A = RandomInteger[{-9,9}, {4,3}];
A // MatrixForm
```

$$\begin{pmatrix} 4 & -2 & 2 \\ 5 & -8 & 1 \\ -8 & 5 & -9 \\ 0 & -6 & 2 \end{pmatrix}$$

```
{U,H} = HermiteDecomposition[A];
MatrixForm /@ {U,H}
```

$$\left\{\begin{pmatrix} 8 & -3 & 2 & 3 \\ 20 & -8 & 5 & 8 \\ 9 & -4 & 2 & 4 \\ 75 & -28 & 20 & 29 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}\right\}$$

```
U.A == H
```

True

Sage Here the method is called hermite_form. It works for integer and polynomial matrices. We do an example for integers:

```
A = random_matrix(ZZ, 4, 3)
A
```

$$\begin{bmatrix} -1 & 1 & -3 \\ 8 & 0 & -1 \\ -2 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$$

```
H, U = A.hermite_form(transformation=True)
H, U
```

$$\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 6 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 4 & 0 \\ 2 & 3 & 9 & 2 \end{bmatrix}\right)$$

For polynomial matrices, we have to declare the ring first. Then we can define the matrix and compute the Hermite normal form.

```
R.<x>=QQ[]
A = matrix(3,3, [1, x-1, x, x, x+1, -1, -1, 0, x+1])
A
```

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

```
A.hermite_form()
```

$$\begin{pmatrix} 1 & x-1 & 1 \\ 0 & 1 & \frac{1}{2}x^2 - \frac{3}{2} \\ 0 & 0 & -x^3 + x^2 + 5x + 1 \end{pmatrix}$$

Note, however, that Sage does not reduce the entries above the pivots here.

Remark 183. We can use the Hermite normal form implementations in the various computer algebra systems to simulate the Euclidean algorithm (Algorithm 149): Simply apply the Hermite normal form to a matrix consisting of a single column. For instance, in Example 160 we could have used Maple for the computations

```
with(LinearAlgebra):
v := <d^3 - 1, -d^2 + d>;
```

$$v := \begin{bmatrix} d^3 - 1 \\ -d^2 + d \end{bmatrix}$$

```
HermiteForm(v, output=['U', 'H']);
```

$$\begin{bmatrix} 1 & d+1 \\ d & d^2+d+1 \end{bmatrix}, \begin{bmatrix} d-1 \\ 0 \end{bmatrix}$$

Theorem 184. *The (non-zero) rows of the Hermite normal form $H \in {}^rR^n$ of $A \in {}^mR^n$ are a basis for the row space $R^m A$. In particular, the row space of A is free with rank r.*

Proof. By Theorem 167 (part (a)), the rows of H are linearly independent. Further, if $u = vA \in R^m A$ for some $v \in R^m$, then $u = vQ^{-1}(QA)$. Since zero-rows in

$$QA = \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix}$$

do not contribute to u, we see that $u \in R^r H$ is in the row space of H. Conversely, if $x = yH$ is in the row space of H for some $y \in R^r$, then

$$x = yH = \begin{pmatrix} y & 0 \end{pmatrix} \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} y & 0 \end{pmatrix} QA$$

is also in the row space of A. Thus, the rows of H generate $R^m A$. In total, we see that they form a basis. □

Corollary 185. *For a Euclidean domain R, finitely generated submodules of R^n have a rank which is less than or equal to n.*

Proof. Let $x_1, \ldots, x_m$ be the generators of a submodule N of R^n. We form the matrix

$$A = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in {}^mR^n.$$

Then $R^m A = N$, and by Theorem 184 the Hermite normal form of A will be a basis of N. The Hermite normal form has at most as many rows as A, which implies $\operatorname{rank} N \leqslant m$. Moreover, the Hermite normal form cannot have more rows than columns since it is in upper echelon form. (By Definition 161 we can have at most as many pivots as columns.) Thus, $n \geqslant m \geqslant \operatorname{rank} N$ as desired. □

Definition 186 (Rank of a Matrix). Let R be a Euclidean domain, and let $A \in {}^mR^n$. We define the (row) *rank* of A as the rank of the row space $R^m A$ of A (or, equivalently, as the number of non-zero rows in the Hermite normal form of A). We denote it by $\operatorname{rank} A$.

Definition 187 (Rank-Revealing Transformation). Let $A \in {}^mR^n$ be a matrix over the Euclidean domain R. A transformation $Q \in \mathrm{GL}_m(R)$ is called *rank-revealing* if

$$QA = \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix}$$

for some regular $B \in {}^rR^n$ (where $r \leqslant m$).

Remark 188. The Hermite normal form computation Algorithm 172 yields a rank-revealing transformation.

Theorem 189. *In the situation of Definition 187 we have* $r = \operatorname{rank} A$.

Proof. Since B is regular, its rows are linearly independent. At the same time, they generate the row space of B. Thus, they must be a basis of $R^r B$. Since

$$R^r B = R^m \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix} = R^m QA = R^m A,$$

we see that they are also a basis for $R^m A$. Thus, $r = \operatorname{rank} A$. □

12 Applications of the Hermite Normal Form

Definition 190 (Left/Right Kernel). Let $A \in {}^mR^n$. The *left kernel* of A is

$$\ker \cdot A = \{v \in R^m \mid vA = 0\} = \ker (v \mapsto vA)$$

while the *right kernel* is

$$\ker A \cdot = \{w \in {}^nR \mid Aw = 0\} = \ker (w \mapsto Aw).$$

Theorem 191. *Let R be a Euclidean domain, let $A \in {}^mR^n$, and let $Q \in \mathrm{GL}_m(R)$ be a rank-revealing transformation for A. If* $\operatorname{rank} A = r$, *then the last $m - r$ rows of Q are a basis for the left kernel of A.*

Proof. Partition Q into

$$Q = \begin{pmatrix} V \\ W \end{pmatrix} \qquad \text{where} \quad V \in {}^rR^m \quad \text{and} \quad W \in {}^{m-r}R^m.$$

Since Q is rank-revealing, there exist a regular $B \in {}^rR^n$ such that

$$QA = \begin{pmatrix} V \\ W \end{pmatrix} A = \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix};$$

that is, $VA = B$ and $WA = \mathbf{0}$. Thus, $R^{m-r}W \subseteq \ker \cdot A$.

Let conversely $u \in \ker \cdot A$. Then

$$0 = uA = uQ^{-1}QA = (uQ^{-1}) \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix}.$$

Let $uQ^{-1} = (x, y)$ with $x \in R^r$ and $y \in R^{m-r}$. Then the equation above implies $0 = xB + y\mathbf{0} = xB$. Since the rows of B are linearly independent, we must have $x = 0$. Thus, we conclude that $u = (x, y)Q = (0, y)Q = yW \in R^{m-r}W$; that is, $\ker \cdot A \subseteq R^{m-r}W$. □

Application 192. Let R be a Euclidean domain, and let $A \in {}^mR^n$ and $b \in {}^mR$. We want to solve the diophantine linear system

$$Ax = b$$

for $x \in {}^nR$. Compute the Hermite normal form of the transpose A^t of A (or, equivalently, compute the column Hermite normal form of A). With Algorithm 172 we find $Q^t \in \mathrm{GL}_n(R)$ such that Q^tA^t is in Hermite normal form. Let $\Phi^t \in \mathrm{GL}_m(R)$ be a column permutation such that the pivot indices of $Q^tA^t\Phi^t$ are $1, 2, \dots, \min\{m, n\}$. Thus, after transposing everything we can write

$$\Phi AQ = \begin{pmatrix} L & \mathbf{0} \\ M & \mathbf{0} \end{pmatrix}$$

where $L \in {}^rR^r$ is a lower triangular matrix with a non-zero diagonal, $M \in {}^{m-r}R^r$ is an arbitrary matrix and $r = \operatorname{rank} A$. We can rewrite the original equation as

$$\Phi b = \Phi AQQ^{-1}x = \begin{pmatrix} L & \mathbf{0} \\ M & \mathbf{0} \end{pmatrix} Q^{-1}x$$

multiplying by Φ from the left and inserting $\mathbf{1} = QQ^{-1}$. We partition $\Phi b = (c, d)^t$ and $Q^{-1}x = (y, z)^t$ to match the partition of ΦAQ. The equation becomes

$$Ly = c \qquad \text{and} \qquad My = d$$

with no conditions on z. Since L is lower triangular, we can inductively compute a solution for $Ly = c$: Let $L = (\ell_{ij})_{ij}$, $c = (c_1, \dots, c_r)$, and $y = (y_1, \dots, y_r)^t$. Then the first entry of the equation is $\ell_{11}y_1 = c_1$. This has a solution if and only if $\ell_{11} \mid c_1$. Assuming that this is the case, the unique solution is $y_1 = c_1/\ell_{11}$. The next entry of the equation is $\ell_{21}y_1 + \ell_{22}y_2 = c_2$ or $\ell_{22}y_2 = c_2 - \ell_{21}(c_1/\ell_{11})$. If ℓ_{22} divides the right hand side, then we can also find a unique solution for y_2. We proceed in this way until we find solutions for all the $y_1, \dots, y_r$ or until we fail to find a solution for one of the rows. If we fail to find a solution, then the original equation $Ax = b$ can likewise not have a solution: Assume that this was not the case and that x was a solution, then the first r entries of $Q^{-1}x$ would be a solution of $Ly = c$; contradicting the fact that we did not find one. Let us assume that we found a solution $Ly = c$. As stated above, it must be unique since in each row there is only one new variable. The equation $My = d$ provides us wih further conditions on y;

we sometimes call these the *compatibility conditions*. As above we find that $My = d$ if and only if $Ax = b$ has a solution. Assuming that also the compatibility conditions hold, we find all solutions to the original system by setting $x = Q(y,z)^t$ where y is the partial solution and $z = (z_{r+1}, \ldots, z_n)$ are some variables.

Example 193. Consider $R = \mathbb{Z}$ and the linear diophantine system

$$\begin{array}{rrrrl} -5x_1 + & 3x_2 - & 2x_3 + & 9x_4 &= 7 \\ -47x_1 + & 31x_2 - & 18x_3 + & 87x_4 &= 65 \\ -73x_1 + & 51x_2 - & 28x_3 + & 138x_4 &= 101 \\ -47x_1 + & 32x_2 - & 18x_3 + & 88x_4 &= 65 \end{array}.$$

In matrix form this becomes

$$\underbrace{\begin{pmatrix} -5 & 3 & -2 & 9 \\ -47 & 31 & -18 & 87 \\ -73 & 51 & -28 & 138 \\ -47 & 32 & -18 & 88 \end{pmatrix}}_{=A\in{}^4\mathbb{Z}^4} x = \underbrace{\begin{pmatrix} 7 \\ 65 \\ 101 \\ 65 \end{pmatrix}}_{=b\in{}^4\mathbb{Z}}.$$

Computing the Hermite normal form of A^t yields

$$AQ = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ \hline 1 & 1 & 1 & 0 \end{array}\right) \qquad \text{where} \quad Q = \begin{pmatrix} 5 & -3 & 3 & 2 \\ -1 & -1 & 0 & -2 \\ -10 & 6 & -3 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix} \in \mathrm{GL}_4(\mathbb{Z}).$$

We do not need any permutations in this example; and we have already indicated the block structure of AQ. We try to solve

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 65 \\ 101 \end{pmatrix}.$$

From the first row, we see that $y_1 = 7$. In the second row we thus obtain $65 = y_1 + 2y_2 = 7 + 2y_2$ or, equivalently, $58 = 2y_2$. This equation has the solution $y_2 = 29$. Finally, in the last row we get $101 = 2y_1 + 3y_3 = 14 + 3y_3$ or $87 = 3y_3$. This has the solution $y_3 = 29$. We now have to check the compatibility condition $y_1 + y_2 + y_3 = 65$ which holds true for our solution $(y_1, y_2, y_3) = (7, 29, 29)$. Thus, the original system has the solution set

$$x = Q\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 5 & -3 & 3 & 2 \\ -1 & -1 & 0 & -2 \\ -10 & 6 & -3 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 29 \\ 29 \\ z \end{pmatrix} = \begin{pmatrix} 35 + 2z \\ -36 - 2z \\ 17 + z \\ 36 + 2z \end{pmatrix}$$

where $z \in \mathbb{Z}$ is arbitrary.

Exercise 194. Solve the linear diophantine system

$$\begin{array}{rrrrl} 39x_1 - & 27x_2 - & 3x_3 - & 71x_4 &= 98 \\ 69x_1 - & 47x_2 - & 5x_3 - & 125x_4 &= 172 \\ 130x_1 - & 89x_2 - & 10x_3 - & 236x_4 &= 325 \\ 68x_1 - & 46x_2 - & 5x_3 - & 123x_4 &= 169 \end{array}$$

over the integers.

Exercise 195. Implement the method described in Application 192 in a programming language of your choice.

Example 196. We can use a similar strategy as in Application 192 in order to solve operator equations. Let $R = \mathbb{R}[\partial]$ and $M = C^\infty(\mathbb{R})$ as in Example 38. We consider the system

$$\begin{array}{lclcll} 2f'' - f' - f & - & g'' + g' & - & 2h'' + h' + h & = 0 \\ -2f'' + f' + 4f & + & g'' - g' - 2g & + & 2h'' - h' - 4h & = 0 \\ -2f' + 5f & + & g' - 3g & + & 2h' - 5h & = 0 \end{array}$$

in the unknown functions $f, g, h \in C^\infty(\mathbb{R})$. In matrix notation (see Remark 96) this system becomes $A \bullet y = 0$ where

$$A = \begin{pmatrix} 2\partial^2 - \partial - 1 & -\partial^2 + \partial & -2\partial^2 + \partial + 1 \\ -2\partial^2 + \partial + 4 & \partial^2 - \partial - 2 & 2\partial^2 - \partial - 4 \\ -2\partial + 5 & \partial - 3 & 2\partial - 5 \end{pmatrix} \qquad \text{and} \qquad y = \begin{pmatrix} f \\ g \\ h \end{pmatrix}.$$

The column Hermite normal form of A is

$$H = \left(\begin{array}{cc|c} \partial - 1 & 0 & 0 \\ 2 & \partial + 2 & 0 \\ \hline 2 & 3 & 0 \end{array}\right) = AQ \qquad \text{where} \quad Q = \begin{pmatrix} -\partial + 1 & -\partial & 1 \\ -2\partial + 1 & -2\partial - 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{R}[\partial])$$

and where we partitioned H into the lower triangular part and the compatibility conditions. Letting $(u, v, w)^t = Q^{-1}y$, we thus have to solve the system

$$\begin{array}{lcll} u' - u & & & = 0 \\ 2u & + & v' + 2v & = 0 \\ 2u & + & 3v & = 0 \end{array}$$

with no condition on h. The first equation yields $u = C_1 e^x$ with $C_1 \in \mathbb{R}$ arbitrary. Substituting this into the second equation[12] we have to solve

$$v' + 2v = -2C_1 e^x.$$

This is a first order linear equation with integrating factor $\mu = e^{\int 2dx} = e^{2x}$. We obtain $-2C_1 e^{3x} = e^{2x}(v' + 2v) = (e^{2x}v)'$, and thus

$$v = -2C_1 e^{-2x} \int e^{3x} dx = -2C_1 e^{-2x}\left(\frac{1}{3}e^{3x} + C_2\right) = -\frac{2}{3}C_1 e^x - 2\underbrace{C_1 C_2}_{=\tilde{C}_2} e^{-2x}.$$

with $\tilde{C}_2 \in \mathbb{R}$ arbitrary. We still have to check the compatibility condition $2u + 3v = 0$. Substituting our solutions for u and v this equation becomes

$$0 = 2C_1 e^x \; - \; 2C_1 e^x - 6\tilde{C}_2 e^{-2x} = -6\tilde{C}_2 e^{-2x}.$$

[12] For this specific system, using the third equation next would have been easier. However, we want to follow the general method as closely as possible.

Since e^{-2x} is not zero, we see that $\tilde{C}_2$ must be zero in order to make the compatibility condition work. Thus, we obtain conditions on the constants. This is different to Application 192 where we simply have to check whether the compatibility conditions hold or not. In total, we have

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} C_1 e^x \\ -\frac{2}{3} C_1 e^x \\ w \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} f \\ g \\ h \end{pmatrix} = Q \bullet \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{2}{3} C_1 e^x + w \\ C_1 e^x \\ w \end{pmatrix}$$

where $w \in C^\infty(\mathbb{R})$ and $C_1 \in \mathbb{R}$ are arbitrary.

Exercise 197. Solve the system of linear differential equations

$$\begin{array}{lclcll} -2f' + 2f & - & 2g' + 2g & - & h' + h & = 0 \\ -2f'' - f' - 5f & - & 2g'' - g' - 5g & - & h'' - h' - 3h & = 0 \\ -2f'' + 3f' - 2f & - & 2g'' + 3g' - 2g & - & h'' + h' & = 0 \end{array}$$

for $f, g, h \in C^\infty(\mathbb{R})$.

Application 198. Using the Hermite normal form, we can design a method for computing the inverse of a matrix $A \in {}^nR^n$ over an Euclidean domain R if it exists, or prove that A is not unimodular: Compute first the Hermite normal form H with transformation matrix $Q \in \mathrm{GL}_n(R)$ and the rank of A. If the rank is not n, then A cannot be unimodular as the determinant is zero (see Theorem 91). If the rank is n, then

$$QA = H = \begin{pmatrix} h_{11} & & * \\ & \ddots & \\ 0 & & h_{nn} \end{pmatrix}$$

is an upper triangular matrix with pivots $h_{11}, \ldots, h_{nn}$. Since $A = Q^{-1}H$ and therefore $\det A = (\det Q)^{-1} \det H = (\det Q)^{-1} h_{11} \cdots h_{nn}$ we see that A is unimodular if and only if $h_{11}, \ldots, h_{nn}$ are units; again by Theorem 91. Assuming that this is the case, by property (c) of Definition 161 that implies that all entries above the $h_1, \ldots, h_n$ are zero. Thus, $QA = \operatorname{diag}(h_{11}, \ldots, h_{nn})$ which implies $A^{-1} = \operatorname{diag}(h_{11}^{-1}, \ldots, h_{nn}^{-1})Q$.

Example 199. Consider $R = \mathbb{Q}[x]$ and let

$$A = \begin{pmatrix} x & x^2 + x - 1 & x^2 - 1 \\ -1 & -x - 2 & -1 - 2x \\ -1 & -1 - x & -1 - x \end{pmatrix}.$$

Then the Hermite normal form of A is $H = \mathbf{1}$ and the transformation matrix is

$$A^{-1} = Q = \begin{pmatrix} -x^2 + 1 & x^2 + x & -x^3 - x^2 - 1 \\ x & -1 - x & x^2 + x + 1 \\ -1 & 1 & -1 - x \end{pmatrix}.$$

(In this case the diagonal entries of H are all already 1.)

Exercise 200. Compute the inverses of

$$\begin{pmatrix} -2x^2 + 3x - 1 & -2x^3 + 4x^2 - 3x + 1 & x^3 - x^2 + 2x - 1 \\ -4x + 4 & -4x^2 + 6x - 3 & 2x^2 - x + 3 \\ -5 + 6x & 6x^2 - 8x + 4 & -3x^2 + x - 5 \end{pmatrix} \in {}^3\mathbb{Q}[x]^3$$

and

$$\begin{pmatrix} -2 & -7 & -6 & -5 \\ -3 & -11 & -10 & -8 \\ -4 & -17 & -15 & -12 \\ -5 & -20 & -18 & -14 \end{pmatrix} \in {}^4\mathbb{Z}^4$$

or show that they do not exist.

Corollary 201. *Let R be a Euclidean domain. Then every unimodular matrix is the product of elementary matrices.*

Proof. Following Algorithm 172, it is easy to check that all transformations are elementary. Thus, the inverse of $A \in \mathrm{GL}_n(R)$ as computed by Application 198 is a product of elementary matrices. But then also A itself must be a product of elementary matrices since the inverses of elementary matrices are again elementary matrices by Remark 114. □

Application 202. Let $U \in {}^mR^n$ where R is a Euclidean domain. We want to determine whether U can be completed to a unimodular matrix. That is, in the case $m \geqslant n$ we want know whether there is $A \in {}^mR^{m-n}$ such that $(U, A) \in \mathrm{GL}_m(R)$; and in the case $m < n$ we want to know whether there exists $B \in {}^{n-m}R^n$ such that

$$\begin{pmatrix} U \\ B \end{pmatrix} \in \mathrm{GL}_n(R).$$

Without loss of generality, we concentrate on the case that $m \geqslant n$ (for the other case simply use the transposed matrices). Compute the Hermite normal form H of U with transformation matrix $Q \in \mathrm{GL}_m(R)$. Then

$$QU = \begin{pmatrix} H \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} x_1 & & * \\ & \ddots & \\ & & x_n \\ & \mathbf{0} & \end{pmatrix}$$

for some $x_1, \ldots, x_n \in R$. (We do not require H to have rank n; some of the x's could be part of the zero block below H.) First assume that $x_1, \ldots, x_n$ are all units. Then as in Application 198, it follows that the entries above the diagonal must be zero. That is, after dividing by $x_1, \ldots, x_n$ we have

$$\mathrm{diag}(x_1^{-1}, \ldots, x_n^{-1}, \mathbf{1})QU = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix}, \qquad \text{or, equivalently,} \qquad U = \left(Q^{-1}\,\mathrm{diag}(x_1, \ldots, x_n, \mathbf{1})\right)\begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix}.$$

Thus, U equals the first n columns of the unimodular matrix $Q^{-1}\,\mathrm{diag}(x_1, \ldots, x_n, \mathbf{1})$.

Assume now that at least one of $x_1, \ldots, x_n$ is not a unit. (This includes the case that H does have a rank strictly less than n.) Assume there was a unimodular completion $(U, A) \in \mathrm{GL}_n(R)$ for some $A \in {}^mR^{m-n}$. We have

$$Q\begin{pmatrix} U & A \end{pmatrix} = \left(\begin{array}{ccc|c} x_1 & & * & \\ & \ddots & & \mathbf{*} \\ 0 & & x_n & \\ \hline & \mathbf{0} & & W \end{array}\right)$$

for some matrix $W \in {}^{m-n}R^{m-n}$. Thus,

$$x_1 \cdots x_n(\det W) = (\det Q)(\det \begin{pmatrix} U & A \end{pmatrix}) \in R^*$$

since Q and (U, A) are unimodular. However, this contradicts our assumption about the $x_1, \ldots, x_n$. Thus, there cannot be a unimodular completion if any of the $x_1, \ldots, x_n$ is not a unit.

Example 203. We want to complete the vectors

$$\begin{pmatrix} 5 \\ 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} \in {}^3\mathbb{Z}$$

to a basis of ${}^3\mathbb{Z}$. This is the same task as finding a unimodular completion of the matrix

$$U = \begin{pmatrix} 5 & 3 \\ 4 & 2 \\ 7 & 4 \end{pmatrix} \in {}^3\mathbb{Z}^2.$$

Following Application 202, we compute the Hermite normal form obtaining

$$QU = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{where} \quad Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -3 & 1 \\ -2 & -1 & 2 \end{pmatrix}.$$

Since the diagonal entries of the Hermite normal form are units, the completion is

$$Q^{-1} = \left(\begin{array}{cc|c} 5 & 3 & -4 \\ 4 & 2 & -3 \\ 7 & 4 & -5 \end{array} \right).$$

(Note that since the diagonal entries are just 1 in this example, we do not have the diagonal matrix of Application 202 here.)

Application 204. Let $A, B \in {}^nR^n$ where R is a Euclidean domain. We are looking for a right greatest common divisor of A and B; that is, a matrix $G \in {}^nR^n$ such that $A = \tilde{A}G$ and $B = \tilde{B}G$ for some $\tilde{A}, \tilde{B} \in {}^nR^n$ and such that whenever $A = \hat{A}H$ and $B = \hat{B}H$ for some $H \in {}^nR^n$ and $\hat{A}, \hat{B} \in {}^nR^n$ then $G = \hat{G}H$ for some $\hat{G} \in {}^nR^n$. Form the block matrix

$$\begin{pmatrix} A \\ B \end{pmatrix} \in {}^{2n}R^n$$

and compute the Hermite normal form

$$Q \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} G \\ \mathbf{0} \end{pmatrix}$$

where $G \in {}^nR^n$ and $Q \in \mathrm{GL}_n(R)$. Decompose

$$Q = \begin{pmatrix} M & N \\ S & T \end{pmatrix} \qquad \text{and} \qquad Q^{-1} \begin{pmatrix} U & V \\ X & Y \end{pmatrix}$$

into n-by-n blocks. Then

$$\begin{pmatrix} A \\ B \end{pmatrix} = Q^{-1} \begin{pmatrix} G \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} U & V \\ X & Y \end{pmatrix} \begin{pmatrix} G \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} UG \\ XG \end{pmatrix}.$$

That is, G is a right divisor of A and B. Let now $H \in {}^nR^n$ be another right divisor, say, $A = \hat{A}H$ and $B = \hat{B}H$ for some $\hat{A}, \hat{B} \in {}^nR^n$. We have

$$\begin{pmatrix} G \\ \mathbf{0} \end{pmatrix} = Q \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} M & N \\ S & T \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} MA + NB \\ SA + TB \end{pmatrix}.$$

Thus, $G = MA + NB = (M\hat{A} + N\hat{B})H$; that is, H is a right factor of G.

Remark 205. By transposing everything, we can use the approach of Application 204 to compute also left greatest common divisors of two matrices.

Remark 206. In Application 204 (and Remark 205), instead of using the Hermite normal form we could employ any rank-revealing transformation $U \in \mathrm{GL}_n(R)$: The important part is that the lower n rows of

$$U \begin{pmatrix} A \\ B \end{pmatrix}$$

are zero. However, since the rank of

$$\begin{pmatrix} A \\ B \end{pmatrix}$$

is at most n, this must be the case for the rank-revealing transformation U.

Example 207. Consider the matrices

$$A = \begin{pmatrix} -12 & 80 & 19 \\ 25 & 53 & 31 \\ 68 & 8 & 12 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} -82 & 39 & 15 \\ -24 & 23 & 25 \\ 0 & 17 & -23 \end{pmatrix} \in {}^3\mathbb{Z}^3.$$

Computing the Hermite normal form yields

$$\begin{pmatrix} -44 & 263 & -342 & -197 & 0 & 0 \\ -12 & 38 & -39 & -19 & 0 & 0 \\ -37 & 204 & -260 & -148 & 0 & 0 \\ -44 & 332 & -453 & -268 & 0 & 0 \\ -30 & 230 & -315 & -187 & 1 & 0 \\ -20 & 142 & -192 & -113 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc} -12 & 80 & 19 \\ 25 & 53 & 31 \\ 68 & 8 & 12 \\ \hline -82 & 39 & 15 \\ -24 & 23 & 25 \\ 0 & 17 & -23 \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 258 \\ 0 & 1 & 197 \\ 0 & 0 & 281 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

That means, a right greatest common divisor of A and B is

$$G = \begin{pmatrix} 1 & 0 & 258 \\ 0 & 1 & 197 \\ 0 & 0 & 281 \end{pmatrix}.$$

13 The Smith–Jacobson Normal Form

Definition 208 (Equivalence). Two matrices A and $B \in {}^mR^n$ are said to be *equivalent* if there exist unimodular matrices $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$ such that $PAQ = B$.

Exercise 209. Proof that equivalence is indeed an equivalence relation.

Definition 210 (Smith–Jacobson Normal Form). A matrix $A \in {}^mR^n$ is in *Smith–Jacobson normal form* if $A = \operatorname{diag}(a_1, a_2, \ldots, a_k)$ where $k = \min\{m, n\}$, $a_1, \ldots, a_k \in R$, and $a_1 \mid a_2 \mid \ldots \mid a_k$. The non-zero elements among $a_1, \ldots, a_k$ are called the *invariant factors* of A.

Remark 211. The Smith–Jacobson normal form was first described by Henry J. S. Smith in 1861 for the integers. Consequently, it is also often called the *Smith normal form* in that context. Later it was generalised to other domains with the most general version being given by Nathan Jacobson. (His normal form works in non-commutative principal ideal domains.)

Algorithm 212. *Input* A matrix $A = (a_{ij})_{ij} \in {}^mR^n$ where R is a Euclidean domain.

Output A matrix $N \in {}^mR^n$ in Smith–Jacobson normal form which is equivalent to A.

Procedure

(a) If A is empty (that is, $m = 0$ or $n = 0$) or $A = \mathbf{0}$, then return $N = A$.

(b) Choose a non-zero entry in A and swap it to position $(1, 1)$.

(c) Apply the Euclidean algorithm (Algorithm 149) to the first column of A obtaining a matrix of the form

$$\begin{pmatrix} f & * \\ 0 & * \end{pmatrix}$$

where f is non-zero.

(d) Apply the Euclidean algorithm (Algorithm 149) to the first row of A obtaining a matrix of the form

$$\begin{pmatrix} g & 0 \\ w & * \end{pmatrix}$$

where g is non-zero.

(e) If $g \nmid w$,[13] then go to step (c).

(f) Else, use g to eliminate each entry of w. Now, the matrix A is of the form

$$\begin{pmatrix} g & 0 \\ 0 & \tilde{A} \end{pmatrix}$$

with $g \neq 0$.

(1) If there is any entry $\tilde{a}$ in $\tilde{A}$ (in the i^{th} column of A) such that g does not divide $\tilde{a}$, then add the i^{th} column of A to the first column and go to step (c).

(2) Else, apply the algorithm recursively to $\tilde{A}$ obtaining $\tilde{N}$ and return $N = \operatorname{diag}(g, \tilde{N})$.

[13] That is, g does not divide (every entry of) w.

Example 213. We are going to compute the[14] Smith–Jacobson normal form of the matrix

$$\begin{pmatrix} 0 & 0 & 12 & 10 \\ 6 & -3 & 12 & 9 \\ 2 & -5 & 10 & 7 \end{pmatrix} \in {}^{3}\mathbb{Z}^{4}.$$

Since the top-left entry is zero, we swap the first and the third column of A which gives us

$$\begin{pmatrix} 12 & 0 & 0 & 10 \\ 12 & -3 & 6 & 9 \\ 10 & -5 & 2 & 7 \end{pmatrix}.$$

Now, we apply the Euclidean algorithm (Algorithm 149) to the first column of A. This yields

$$\begin{pmatrix} 2 & -40 & 40 & -10 \\ 0 & -48 & 48 & -14 \\ 0 & -45 & 42 & -13 \end{pmatrix}.$$

Next, we apply the Euclidean algorithm to the first row of A obtaining

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -48 & 48 & -14 \\ 0 & -45 & 42 & -13 \end{pmatrix}.$$

At this point, we have achieved a block diagonalisation. We see that the lower-right block contains entries which are not divisible by the top-left entry 2. We choose the -13 in the last column and add this column to the first which yields

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ -14 & -48 & 48 & -14 \\ -13 & -45 & 42 & -13 \end{pmatrix}.$$

Now, we once more apply the Euclidean algorithm to the first column and get

$$\begin{pmatrix} 1 & -93 & 90 & -27 \\ 0 & -186 & 180 & -54 \\ 0 & -138 & 132 & -40 \end{pmatrix}.$$

With another application of the Euclidean algorithm on the first row we reach again a block diagonalisation

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -186 & 180 & -54 \\ 0 & -138 & 132 & -40 \end{pmatrix}.$$

This time, it is clear that the top-left entry divides every other entry. Thus, we continue now recursively with to lower-right block

$$\begin{pmatrix} -186 & 180 & -54 \\ -138 & 132 & -40 \end{pmatrix}.$$

[14] We will show in Corollary 231 later that the Smith–Jacobson normal form of a matrix is in fact unique.

We do not need to do any row or column swaps, but can immediately apply the Euclidean algorithm to the first column. We get

$$\begin{pmatrix} 6 & -12 & 2 \\ 0 & -48 & 2 \end{pmatrix}.$$

Now, we apply the Euclidean algorithm to the first row, which gives us

$$\begin{pmatrix} 2 & 0 & 0 \\ -46 & -138 & -90 \end{pmatrix}.$$

In this situation we have to once more apply the Euclidean algorithm to the first column. This yields another block diagonalisation

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -138 & -90 \end{pmatrix}.$$

Here, all entries are divisible by the top-left entry 2. Thus, we continue recursively on the lower-right block

$$\begin{pmatrix} -138 & -90 \end{pmatrix}.$$

Also here, we do not need to swap any columns or rows. Moreover, since we only have one row, we can skip the application of the Euclidean algorithm on the first column. Using it instead on the first row yields the matrix

$$\begin{pmatrix} 6 & 0 \end{pmatrix}.$$

This is a block diagonalisation with an empty lower-right block. Thus, the matrix is now in Smith–Jacobson normal form. Putting everything back together, the Smith–Jacobson normal form of A is

$$\operatorname{diag}(1, \operatorname{diag}(2, \begin{pmatrix} 6 & 0 \end{pmatrix})) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{pmatrix}.$$

Remark 214. We can easily modify Algorithm 212 to compute also the transformation matrices $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$ with $PAQ = N$. Simply initialise $P = \mathbf{1}_m$ and $Q = \mathbf{1}_n$, and then mirror every row transformation done during the algorithm on P and every column transformation on Q.

Lemma 215. *Let R be a principal ideal domain. Then every ascending chain of ideals $Ra_1 \subseteq Ra_2 \subseteq Ra_3 \subseteq \dots$ must become stationary. That is, there exists an $n \geqslant 1$ such that $Ra_n = Ra_{n+1} = \dots$.*

Exercise 216. Prove that the union $I = \bigcup_{i \geqslant 1} Ra_i$ of all the ideals $Ra_1, Ra_2, \dots$ is an ideal.

Proof. By Exercise 216, the union $I = \bigcup_{i \geqslant 1} Ra_i$ of all the ideals $Ra_1, Ra_2, \dots$ is an ideal. Since R is a principal ideal domain there must thus exist a $b \in R$ such that $I = Rb$. However, we must have $b \in Ra_n$ for some $n \geqslant 1$. It follows that $Rb \subseteq Ra_n \subseteq Ra_{n+k} \subseteq I = Rb$ for all $k \geqslant 1$; hence $Ra_n = Ra_{n+1} = \dots$ as desired. □

Theorem 217. *Algorithm 212 is correct and terminates.*

Proof. We prove the correctness first. We note first, that the algorithm applies only elementary row and column operations—some of them hidden in the application of the Euclidean algorithm (Algorithm 149)—to A. Thus, whatever matrix N is returned in the end will be equivalent to A. (This also holds through the recursive calls which we can simply imagine to operate on all of A instead of just a block.) Thus, we just need to proof that the returned matrix N is indeed in Smith–Jacobson normal form. This is obvious, if we return already in step (a) of Algorithm 212. Assume that we have reach step (f.2). We only arrive there if we have turned A into a block diagonal matrix $A = \operatorname{diag}(g, \tilde{A})$ where g divides every entry of $\tilde{A}$. In other words $\tilde{A} = g\hat{A}$ for some $\hat{A} \in {}^{m-1}R^{n-1}$. Now, the recursive call will yields a matrix $\tilde{N}$ in Smith–Jacobson normal form which is equivalent to $\tilde{A}$. That is, $\tilde{N} = \operatorname{diag}(a_2, a_3, \ldots, a_k)$ for some $a_2, \ldots, a_k \in R$ with $a_2 \mid \ldots \mid a_k$ and $\tilde{P}\tilde{A}\tilde{Q} = \tilde{N}$ for some unimodular matrices $\tilde{P} \in \mathrm{GL}_{m-1}(R)$ and $\tilde{Q} \in \mathrm{GL}_{n-1}(R)$. Since $\tilde{N} = \tilde{P}\tilde{A}\tilde{Q} = \tilde{P}g\hat{A}\tilde{Q} = g\tilde{P}\hat{A}\tilde{Q}$, we see that g also divides every entry in $\tilde{N}$. Consequently, $g \mid a_2 \mid \ldots \mid a_k$ and the matrix $N = \operatorname{diag}(g, \tilde{N}) = \operatorname{diag}(g, a_2, \ldots, a_k)$ which we return is indeed in Smith–Jacobson normal form.

It remains to prove that the algorithm does indeed terminate. For this it is sufficient to prove that we can reach step (c) of Algorithm 212 only a finite number of times. First look at the loop between the steps (c) and (e). When we apply the Euclidean algorithm (Algorithm 149) alternatingly to the first column and the first row, the top-left entry of A is always involved. This implies that the greatest common divisor which is computed is always a divisor of the top-left entry. Thus, in the top-left entry we obtain a chain of elements $g_1, g_2, g_3, \ldots$ such that $g_{j+1} \mid g_j$ for $j \geqslant 1$. This corresponds to a chain of ideals $Rg_1 \subseteq Rg_2 \subseteq \ldots$ in R. By Lemma 215 the chain must become stationary, that is, there is an $n \geqslant 1$ such that $Rg_n = Rg_{n+1} = \ldots$. Assume now that in step (e) g_n was not a divisor of w. Then, we go back to step (c) and do another Euclidean algorithm on the first column ending up with a greatest common divisor g_{n+1} of g and the entries of w. However, since g_n and g_{n+1} are associated, also g_n is a greatest common divisor of g_n and the entries of w; in particular $g_n \mid w$ contradicting our assumption. Hence, we must eventually get out of the inner loop.

Similarly, the loop between the steps (c) and (f.1) of Algorithm 212 can only be run finitely often. Whatever entry $\tilde{a}$ we bring to the first column of A by doing the addition, after returning to step (c) we will still have a greatest common divisor of the (previous) top-left entry in the top-left position. This implies that also here we obtain an ascending chain of ideals which again must become stationary. If that happens, the top-left entry g must divide everything in $\tilde{A}$; for otherwise, if there was a $\tilde{a}$ which was not divisible by g then step (c) would replace g by a common divisor of $\tilde{a}$ and g. Since the chain of ideals is stationary, this common divisor would be associated to g, contradicting the assumption that g does not divide $\tilde{a}$. □

Exercise 218. Compute the Smith–Jacobson normal form of the following matrices

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 3 & 1 & 2 \\ 8 & 2 & 0 & 8 \\ 9 & 3 & 1 & 8 \end{pmatrix} \in {}^4\mathbb{Z}^4 \qquad \text{and} \qquad \begin{pmatrix} x-5 & 1 & -6 \\ -6 & x-2 & 1 \\ -3 & -9 & x+7 \end{pmatrix} \in {}^3\mathbb{Q}[x]^3.$$

Exercise 219. Implement the Smith–Jacobson normal form in a programming language of your choice. (It is sufficient if the implementation works for the integers.)

Remark 220. The Smith–Jacobson normal form is implemented in most major computer algebra systems. For instance, Maple The command is called SmithForm and its contained in the LinearAlgebra package. It works with both integer and polynomial matrices. Optionally, also the transformation matrices can be computed.

```
with(LinearAlgebra):
A := <1,9,1,1; 4,3,4,1; 4,-9,4,1; -3,6,-3,0>;
```

$$A := \begin{bmatrix} 1 & 9 & 1 & 1 \\ 4 & 3 & 4 & 1 \\ 4 & -9 & 4 & 1 \\ -3 & 6 & -3 & 0 \end{bmatrix}$$

```
U, S, V := SmithForm(A, output=['U', 'S', 'V']);
```

$$U, S, V := \begin{bmatrix} 13 & -12 & 9 & 0 \\ 12 & -11 & 8 & 0 \\ -32 & 29 & -21 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -10 & -10 & -1 \\ 0 & -2 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

```
Equal(U . A . V, S);
```

$$\mathit{true}$$

Mathematica Here the command is called SmithDecomposition. It only works with integer matrices and will always return the transformation matrices.

```
A = {{1,9,1,1}, {4,3,4,1}, {4,-9,4,1}, {-3,6,-3,0}}
A // MatrixForm
```

$$\begin{pmatrix} 1 & 9 & 1 & 1 \\ 4 & 3 & 4 & 1 \\ 4 & -9 & 4 & 1 \\ -3 & 6 & -3 & 0 \end{pmatrix}$$

```
({U,R,V} = SmithDecomposition[A]) // MatrixForm
```

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 9 & 8 & 11 & 0 \\ 13 & 12 & 16 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -10 & -3 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

```
U . A . V == R
```

$$\mathit{True}$$

There is also the Smith Normal Forms package which has commands to deal with polynomial matrices, too.

Sage Here the command is called smith_form. It works for integer and for polynomial matrices.

```
A = matrix([[1,9,1,1],[4,3,4,1],[4,-9,4,1],[-3,6,-3,0]])
A
```

$$\begin{pmatrix} 1 & 9 & 1 & 1 \\ 4 & 3 & 4 & 1 \\ 4 & -9 & 4 & 1 \\ -3 & 6 & -3 & 0 \end{pmatrix}$$

```
S,U,V = A.smith_form()
S,U,V
```

$$\left(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & -4 \\ 1 & -1 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -3 & -2 & -1 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 12 & 11 & 0 \end{pmatrix} \right)$$

```
U * A * V == S
```

$$\mathit{True}$$

For polynomial matrices we have to define the proper ring first. Then we can define the matrix and compute the Smith–Jacobson normal form.

```
R.<x>=QQ[]
A = matrix(3,3, [x-1,2,3, 4,x-5,6, 7,8,x-9])
A.smith_form()
```

$$\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & x^3 - 15x^2 - 18x + 360 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{4} & 0 \\ -\frac{49}{382} & -\frac{28}{191} & \frac{7}{382}x + \frac{25}{382} \\ -7x + 67 & -8x + 22 & x^2 - 6x - 3 \end{pmatrix}, \begin{pmatrix} 1 & -\frac{1}{4}x + \frac{5}{4} & \frac{7}{1528}x^3 - \frac{73}{1528}x^2 - \frac{259}{764}x + \frac{156}{191} \\ 0 & 1 & -\frac{7}{382}x^2 + \frac{19}{191}x + \frac{354}{191} \\ 0 & 0 & 1 \end{pmatrix} \right)$$

Notation 221 (Submatrix). Let $A = (a_{ij})_{ij} \in {}^mR^n$, and let $1 \leqslant i_1 < \ldots < i_r \leqslant m$ and $1 \leqslant j_1 < \ldots < j_s \leqslant n$ where $1 \leqslant r \leqslant m$ and $1 \leqslant s \leqslant n$. Let $I = \{i_1, \ldots, i_r\}$ and $J = \{j_1, \ldots, j_s\}$. Then with A_{IJ} we denote the *submatrix*

$$A_{IJ} = \begin{pmatrix} a_{i_1,j_1} & \ldots & a_{i_1,j_s} \\ \vdots & & \vdots \\ a_{i_r,j_1} & \ldots & a_{i_r,j_s} \end{pmatrix} \in {}^rR^s.$$

If $I = \{i\}$ we also write A_{iJ}, and if $I = \{1, \ldots, m\}$ we write A_{*J}. Similarly for $J = \{j\}$ we write A_{Ij} and for $J = \{1, \ldots, n\}$ we write A_{I*}. If $I = \{1, \ldots, m\} \setminus K$ and $J = \{1, \ldots, n\} \setminus L$, we also write $A_{IJ} = A_{\overline{KL}}$.

Definition 222 (Minor). Let $A \in {}^mR^n$. Then for $k = 1, \ldots, \min\{m, n\}$ a *k-by-k minor* of A is $\det A_{IJ}$ where $I \subseteq \{1, \ldots, m\}$ and $J \subseteq \{1, \ldots, n\}$ fulfil $|I| = |J| = k$.

Definition 223 (Determinantal Divisor). Let $A \in {}^mR^n$. For $k = 1, \ldots, \min\{m, n\}$ the *k^{th} determinantal divisor* of A is the greatest common divisor of all k-by-k minors of A. We will denote it by $d_k(A)$.

Example 224. Consider

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in {}^3\mathbb{Z}^3.$$

Then the 1-by-1 submatrices are

$$(1), \quad (2), \quad (3), \quad (4), \quad (5), \quad (6), \quad (7), \quad (8), \quad \text{and} \quad (9).$$

The 1-by-1 minors are thus $1, 2, 3, 4, 5, 6, 7, 8, 9$ and their greatest common divisor is 1. Consequently, the first determinantal divisor is $d_1(A) = 1$. The 2-by-2 submatrices are

$$\begin{aligned}
A_{\{1,2\},\{1,2\}} &= \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, & A_{\{1,3\},\{1,2\}} &= \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}, & A_{\{2,3\},\{1,2\}} &= \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}, \\
A_{\{1,2\},\{1,3\}} &= \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}, & A_{\{1,3\},\{1,3\}} &= \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}, & A_{\{2,3\},\{1,3\}} &= \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}, \\
A_{\{1,2\},\{2,3\}} &= \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}, & A_{\{1,3\},\{2,3\}} &= \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}, & A_{\{2,3\},\{2,3\}} &= \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix};
\end{aligned}$$

and their determinants and thus the 2-by-2 minors are

$$\begin{aligned}
\det A_{\{1,2\},\{1,2\}} &= -3 & \det A_{\{1,3\},\{1,2\}} &= -6 & \det A_{\{2,3\},\{1,2\}} &= -3 \\
\det A_{\{1,2\},\{1,3\}} &= -6 & \det A_{\{1,3\},\{1,3\}} &= -12 & \det A_{\{2,3\},\{1,3\}} &= -6 \\
\det A_{\{1,2\},\{2,3\}} &= -3 & \det A_{\{1,3\},\{2,3\}} &= -6 & \det A_{\{2,3\},\{2,3\}} &= -3.
\end{aligned}$$

The greatest common divisor of the 2-by-2 minors and therefore the second determinantal divisor is $d_2(A) = 3$. There is only one 3-by-3 submatrix which is A itself. The third determinantal divisor is hence $d_3(A) = \det A = 0$.

Lemma 225. *Let $A \in {}^mR^n$ and $P \in {}^mR^m$, then $d_k(A) \mid d_k(PA)$ for all $k = 1, \ldots, \min\{m, n\}$.*

Proof. Let $I \subseteq \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$ with $|I| = |J| = k$. Write $I = \{i_1, \ldots, i_k\}$ with $i_1 < \ldots < i_k$ and let $P = (p_{rs})_{rs}$. Then

$$\begin{aligned}
\det(PA)_{IJ} = \det(P_{I*}A_{*J}) = \det \begin{pmatrix} P_{i_1*}A_{*J} \\ \vdots \\ P_{i_k*}A_{*J} \end{pmatrix} &= \det \begin{pmatrix} \sum_{\ell_1=1}^m p_{i_1,\ell_1} A_{\ell_1,J} \\ \vdots \\ \sum_{\ell_k=1}^m p_{i_k,\ell_k} A_{\ell_k,J} \end{pmatrix} \\
= \sum_{\ell_1=1}^m \cdots \sum_{\ell_k=1}^m p_{i_1,\ell_1} \cdots p_{i_k,\ell_k} \det \begin{pmatrix} A_{\ell_1,J} \\ \vdots \\ A_{\ell_k,J} \end{pmatrix} &= \sum_{1 \leqslant \ell_1 < \ldots < \ell_k \leqslant m} C_{\ell_1,\ldots,\ell_k} p_{i_1,\ell_1} \cdots p_{i_k,\ell_k} \det \begin{pmatrix} A_{\ell_1,J} \\ \vdots \\ A_{\ell_k,J} \end{pmatrix} \\
&= \sum_{L \subseteq \{1,\ldots,m\}} p_L \det A_{LJ}
\end{aligned}$$

where $C_{\ell_1,\dots,\ell_k}$ is a constant which arises from adding all the determinants of the same matrix with the correct signs obtained from sorting the rows, and where $p_L = C_{\ell_1,\dots,\ell_k} p_{i_1,\ell_1} \cdots p_{i_k,\ell_k}$ for $L = \{\ell_1, \dots, \ell_k\}$ with $1 \leqslant \ell_1 < \dots < \ell_k \leqslant m$. Here, we used the linearity in each row of the determinant in the fourth equality; while for the fifth we reordered the rows of the matrices (possibly changing the signs of the determinants) and removed matrices with identical rows (as their determinants would be zero). The equation shows that the k-by-k minors of PA are linear combinations of the k-by-k minors of A. Thus, every common divisor of the k-by-k minors of A must also be a common divisor of the k-by-k minors of PA. In particular, this is true for the greatest common divisor. We thus obtain $d_k(A) \mid d_k(PA)$ as desired. □

Exercise 226. Let $A \in {}^mR^n$. Show that $d_k(A) = d_k(A^t)$ for all $k = 1, \dots, \min\{m, n\}$.

Exercise 227. Let $A \in {}^mR^n$ and $Q \in {}^nR^n$. Show that $d_k(A) \mid d_k(AQ)$ for all $k = 1, \dots, \min\{m, n\}$.

Exercise 228. Let $A \in {}^mR^n$, $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$. Prove that $d_k(A) = d_k(PAQ)$ for $k = 1, \dots, \min\{m, n\}$.

Theorem 229. *Let $A \in {}^mR^n$ have the Smith–Jacobson normal form $N = \mathrm{diag}(x_1, \dots, x_r, 0, \dots, 0)$ where $x_1, \dots, x_r \neq 0$. Then $r = \operatorname{rank} A$ and $x_1 = d_1(A)$ and $x_j = d_j(A)/d_{j-1}(A)$ for $j = 2, \dots, r$.*

Proof. Let $N = PAQ$ for some $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$. Since $R^m P = R^m$ and therefore $R^m(PAQ) = R^m(AQ)$, we know that $\operatorname{rank}(AQ) = \operatorname{rank}(PAQ)$. The map $v \mapsto vQ$ is a (left) R-module automorphism of R^n. Thus, for any submodule M of R^n we have $\operatorname{rank} M = \operatorname{rank} MQ$. In particular; $\operatorname{rank} A = \operatorname{rank} R^m A = \operatorname{rank} R^m AQ = \operatorname{rank}(AQ)$. In total, we have $\operatorname{rank} A = \operatorname{rank} N$, and it is easy to see that $\operatorname{rank} N = r$.

By Exercise 228, we have $d_k(A) = d_k(N)$ for $k = 1, \dots, \min\{m, n\}$. We will now show that $d_k(N) = x_1 x_2 \cdots x_k$ for $k = 1, \dots, r$. This will then imply $x_k = d_k(N)/d_{k-1}(N) = d_k(A)/d_{k-1}(A)$ for $k \geqslant 2$. The only k-by-k minors of N which are non-zero are those which do not include any zero rows or zero columns. Thus, they are the determinants of submatrices of the form $\mathrm{diag}(x_{i_1}, \dots, x_{i_k})$ where $1 \leqslant i_1 < \dots < i_k \leqslant r$. Thus, the k-by-k minors are products $x_{i_1} \cdots x_{i_k}$. Recall that $x_1 \mid x_2 \mid \dots \mid x_r$. This implies that x_1 divides every k-by-k minor, x_2 divides every k-by-k minor and so. Since $x_1 x_2 \cdots x_k$ is one of the k-by-k minors as well, we thus obtain $d_k(N) = x_1 x_2 \cdots x_k$ as claimed. □

Exercise 230. For $n \geqslant 1$, compute the Smith–Jacobson normal form of

$$A = \begin{pmatrix} x-1 & 0 & \cdots & 0 \\ 0 & x-2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x-n \end{pmatrix} \in {}^n\mathbb{Q}[x]^n.$$

Corollary 231. *The Smith–Jacobson normal form of a matrix over a principal ideal domain is unique (except for the multiplication of its rows by units).*

Proof. Let $M = \mathrm{diag}(x_1, \dots, x_r, 0, \dots, 0)$ and $N = \mathrm{diag}(y_1, \dots, y_s, 0, \dots, 0)$ be two matrices in Smith–Jacobson normal form with $x_1, \dots, x_r, y_1, \dots, y_s \neq 0$; and assume that M and N are equivalent. Then $r = s$ since both matrices must have the same rank by Theorem 229. Moreover, Theorem 229 implies that $x_1 \cdots x_k = d_k(M) = d_k(N) = y_1 \cdots y_k$ for $k = 1, \dots, r$. Inductively, one can now prove that x_1 and y_1 are associated, x_2 and y_2 are associated, and so on. Thus, the diagonal entries of M and N differ only by multiplication with units. □

Remark 232. In light of Corollary 231, we can make the Smith–Jacobson normal form of a matrix unique by picking unique representatives for each class of associates.

Corollary 233. *Let $A \in {}^mR^n$ where R is a Euclidean domain. Then* $\operatorname{rank} R^m A = \operatorname{rank} A\, {}^nR$. *In particular, we could have defined* $\operatorname{rank} A$ *equivalently as the column rank of A.*

Proof. Similar to the proof of Theorem 229 one shows that equivalent matrices have the same column rank. Now consider the Smith–Jacobson normal form $N = \operatorname{diag}(x_1, \ldots, x_r, 0, \ldots, 0)$ of A where $x_1, \ldots, x_r \neq 0$. It has the same row rank and the same column rank as A. However, we can easily see that $\operatorname{rank} R^m N = r = \operatorname{rank} N\, {}^nR$. □

Application 234. We can use the Smith–Jacobson normal form to solve diophantine systems. Consider the system

$$Ax = b$$

where $A \in {}^mR^n$ and $b \in {}^mR$. We want to solve for $x \in {}^nR$. Let $N = \operatorname{diag}(a_1, \ldots, a_r, 0, \ldots, 0)$ be the Smith–Jacobson normal form of A, and let $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$ be the transformation matrices; that is, $PAQ = N$. We abbreviate the non-zero part of N by $\Delta = \operatorname{diag}(a_1, \ldots, a_r)$. Then we can transform the orginal equation into the equivalent equation

$$Pb = PAQQ^{-1}x = N(Q^{-1}x).$$

We write $y = Q^{-1}x$ and $c = Pb$. Then $Ny = c$ has a solution if and only if $Ax = b$ has a solution. Write now $y = (u, v)^t$ and $c = (f, g)^t$ where the upper blocks have r entries. Then $Ny = c$ has the shape

$$\begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

which is equivalent to saying that $\Delta u = f$ (with no conditions on v) and $g = 0$. Thus, if the compatibility conditions $g = 0$ are fulfilled, we have to solve $a_i u_i = f_i$ for $i = 1, \ldots, r$. This is possible if and only if $a_i \mid f_i$ for each i in which case the (unique) solution is $u_i = f_i/a_i$ for $i = 1, \ldots, r$. Since there are no conditions of v, its entries provide free variables. If the process succeds, the original system has the solution $x = Qy$.

Example 235. Consider the system

$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= -4 \\ 2x_1 + 6x_2 + 6x_3 + 12x_4 &= 18 \\ x_1 + 4x_2 + 7x_3 + 11x_4 &= 22 \\ 2x_1 + 8x_2 + 4x_3 + 12x_4 &= 14 \end{aligned}$$

over the integers. The coefficient matrix and the right hand side are

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 6 & 6 & 12 \\ 1 & 4 & 7 & 11 \\ 2 & 8 & 4 & 12 \end{pmatrix} \in {}^4\mathbb{Z}^4 \qquad \text{and} \qquad b = \begin{pmatrix} -4 \\ 18 \\ 22 \\ 14 \end{pmatrix} \in {}^4\mathbb{Z}.$$

The Smith–Jacobson normal form of A is

$$PAQ = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} \operatorname{diag}(a_1, a_2, a_3) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

with the transformation matrices

$$P = \begin{pmatrix} 1 & 1 & 0 & -1 \\ -2 & 1 & 0 & 0 \\ 8 & -3 & 0 & -1 \\ 1 & -1 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & -3 & -4 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_4(\mathbb{Z}).$$

We have

$$Pb = \left(\begin{array}{c} 0 \\ 26 \\ -100 \\ \hline 0 \end{array}\right) = \begin{pmatrix} f \\ g \end{pmatrix}.$$

Thus, the compatibility conditions $g = 0$ are fulfilled. Moreover, $a_i \mid f_i$ for $i = 1, 2, 3$. Thus, we find the solution

$$y = \begin{pmatrix} 0/1 \\ 26/2 \\ -100/10 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \\ -10 \\ v \end{pmatrix}$$

for the transformed system where $v \in \mathbb{Z}$ is arbitrary. This yields the solution

$$x = Qy = \begin{pmatrix} -3 \\ 1 - v \\ 3 - v \\ v \end{pmatrix}$$

of the original system.

Remark 236. An approach similar to Application 234 works for more general equations over modules. Let R be a Euclidean domain and let M be a left R module. Moreover, let $A \in {}^{m}R^{n}$ and $b \in {}^{m}M$. Then we can find solutions $x \in {}^{n}M$ of

$$A \bullet x = b$$

by computing the Smith–Jacobson normal form $N = PAQ$ of A with transformation matrices $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$. As in Application 234 let us decompose $N = \operatorname{diag}(\Delta, \mathbf{0})$ where $\Delta = \operatorname{diag}(a_1, \ldots, a_r)$ as well as $y = Q^{-1}x = (u, v)^t$ and $c = Pb = (f, g)^t$. Then the original system has a solution if and only if $\Delta \bullet u = f$ has a solution and the compatibility condition $g = 0$ holds. The difference to Application 234 is that determining whether there exists $u_1, \ldots, u_r \in M$ with $a_i \bullet u_i = f_i$ for all $i = 1, \ldots, r$ can be much more difficult.

Example 237. Consider the following system of linear ordinary differential equations

$$\begin{array}{llll} 2f'' + 3f & + 2g'' + g & + 4h'' + 4h & = 7x^2 - 2x + 11 \\ f'' + f & + g'' + g & + 2h'' + 2h & = 3x^2 + 5 \\ f'' + 2f & + g'' & + 2h'' + 2h & = 4x^2 - 2x + 6 \end{array}.$$

We want to solve for $f, g, h \in C^\infty(\mathbb{R})$. We rewrite the equation in the language of rings and modules using Example 38. Define the matrix

$$A = \begin{pmatrix} 2\partial^2 + 3 & 2\partial^2 + 1 & 4\partial^2 + 4 \\ \partial^2 + 1 & \partial^2 + 1 & 2\partial^2 + 2 \\ \partial^2 + 2 & \partial^2 & 2\partial^2 + 2 \end{pmatrix} \in {}^{3}\mathbb{R}[\partial]^{3}$$

and the vectors $x = (f, g, h)^t$ and

$$b = \begin{pmatrix} 7x^2 - 2x + 11 \\ 3x^2 + 5 \\ 4x^2 - 2x + 6 \end{pmatrix} \in {}^3C^\infty(\mathbb{R}).$$

Then the system can be written as $A \bullet x = b$. The Smith–Jacobson normal form of A is

$$N = PAQ = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & \partial^2 + 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

where the transformation matrices are

$$P = \begin{pmatrix} 1 & -2 & 0 \\ -\frac{1}{2}\partial^2 - 1 & \partial^2 + 2 & \frac{1}{2} \\ -2 & 2 & 2 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{R}[\partial]).$$

Applying P to the right hand side b yields

$$P \bullet b = \left(\begin{array}{c} x^2 - 2x + 1 \\ x^2 + x + 1 \\ \hline 0 \end{array}\right).$$

Hence, the compatibility conditions are fulfilled. Let now $y = (u, v, w)^t = Q^{-1}x$. Then the original system is equivalent to

$$u = x^2 - 2x + 1 \qquad \text{and} \qquad v'' + v = x^2 + x + 1$$

with no conditions on w. The first equation is already solved. For the second, we note that the general solution of $z'' + z = 0$ is $z = C_1 \cos x + C_2 \sin x$. Thus, a fundamental system for the equation is $z_1 = \cos x$ and $z_2 = \sin x$. The Wronskian matrix is

$$Z = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \qquad \text{with inverse} \qquad Z^{-1} = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}.$$

Using variation of constants (see Remark 286), we obtain

$$v = -\cos x \int (x^2 + x + 1) \sin x \, dx + \sin x \int (x^2 + x + 1) \cos x \, dx.$$

With integration by parts, the integrals evaluate to

$$\int (x^2 + x + 1) \sin x \, dx = (2x + 1) \sin x - (x^2 + x - 1) \cos x$$

and

$$\int (x^2 + x + 1) \cos x \, dx = (x^2 + x - 1) \sin x + (2x + 1) \cos x.$$

This yields

$$v = x^2 \cos^2 x + x^2 \sin^2 x + x \cos^2 x + x \sin^2 x - \cos^2 x - \sin^2 x = x^2 + x - 1.$$

Thus, we have the solution

$$y = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} x^2 - 2x + 1 \\ x^2 + x - 1 + C_1 \cos x + C_2 \sin x \\ w \end{pmatrix}$$

where $w \in C^\infty(\mathbb{R})$ is arbitrary. This leads to the solution

$$x = \begin{pmatrix} f \\ g \\ h \end{pmatrix} = Q \bullet y = \begin{pmatrix} 2x^2 - x + C_1 \cos x + C_2 \sin x - w \\ x^2 + x - 1 + C_1 \cos x + C_2 \sin x - w \\ w \end{pmatrix}$$

of the original system.

Remark 238. For solving diophantine systems as in Application 234 (and also in Remark 236) diagonalising the system matrix is sufficient. That is, we only need to transform A using elementary row and column transformations into a matrix of the form $\operatorname{diag}(a_1, \ldots, a_r, 0, \ldots, 0)$; but it is not necessary that $a_1 \mid \ldots \mid a_r$. We can compute such a diagonalisation with Algorithm 212 where we simply omit step (f.1) and immediately go into the recursive call in step (f.2) instead. Of course, this diagonal form will no longer be unique.

Remark 239. If we do compute a full Smith–Jacobson normal form; then in Remark 236 we can make use of the fact that the invariant factors divide each other: If we are solving differential equations and we compute fundamental systems for the homogenous equations given by the invariant factors using Remark 284, finding the factorisations will become easier if we do it iteratively. In this way, instead of factoring a_j from scratch, we only have to factor a_j/a_{j-1}.

Exercise 240. Solve the system of linear ordinary differential equations

$$\begin{aligned} f'' - 3f' + 2f - g'' + 2g - h' + h &= 3x - 5 \\ -f' + f - g'' + 2g - h' + h &= 2x - 2 \\ f' - f + g'' - g + h' - h &= -2x + 2 \end{aligned}$$

for $f, g, h \in C^\infty(\mathbb{R})$.

Exercise 241. Let $a, b \in R$ where R is a Euclidean domain. Show that the matrices

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \gcd(a,b) & 0 \\ 0 & \operatorname{lcm}(a,b) \end{pmatrix}$$

are equivalent.

Exercise 242. Use Exercise 241 to derive a method to compute the Smith–Jacobson normal form of a diagonal matrix.

14 The Popov Normal Form

Notation 243. In this whole section K be a field and let $R = K[x]$ be the univariate polynomial ring over K in the indeterminate x.

Remark 244. In this section we will present the Popov normal form as a normal form with respect to row equivalence and we introduce related topics such as row reduction. Some authors prefer to do the Popov normal form as a normal form with respect to column equivalence. In this case, we can simply work on the transpose.

Notation 245. Let $A = (a_{ij})_{ij} \in {}^m R^n$ be a matrix. We define

$$\deg A = \max\{\deg a_{ij} \mid i = 1, \ldots, m \text{ and } j = 1, \ldots, n\}.$$

Further, for $k \geqslant 0$ we use $\operatorname{coeff}_k(p)$ to denote the coefficient of x^k in $p \in R$; and we extend this to matrices by

$$\operatorname{coeff}_k(A) = \big(\operatorname{coeff}_k(a_{ij})\big)_{ij}.$$

In order to avoid exceptions later, we also define the special case $\operatorname{coeff}_{-\infty}(p) = 0$. We will use the same notations for (row and column) vectors regarding them as (single row or single column) matrices.

Definition 246 (Row Degrees). Consider the matrix

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in {}^m R^n$$

with rows $a_1, \ldots, a_m \in R^n$. For $k = 1, \ldots, m$ we define the k^{th} *row degree* of A to be $\operatorname{rdeg}_k(A) = \deg a_k$.

Definition 247 (Order). The *order* of $A \in {}^m R^n$ is defined as

$$\operatorname{ord} A = \sum_{\substack{j=1 \\ A_{j,*} \neq 0}}^{m} \operatorname{rdeg}_j(A);$$

that is, as the sum of the degrees of all non-zero rows of A.

Definition 248 (Leading Coefficient Matrix). Let

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in {}^m R^n$$

be a matrix with rows $a_1, \ldots, a_m \in R^n$. Denote the row degrees by $\nu_1 = \operatorname{rdeg}_1(A), \ldots, \nu_m = \operatorname{rdeg}_m(A)$. We define the *leading coefficient matrix* of A as

$$\operatorname{LCM}(A) = \begin{pmatrix} \operatorname{coeff}_{\nu_1}(a_1) \\ \vdots \\ \operatorname{coeff}_{\nu_m}(a_m) \end{pmatrix} = \Big(\operatorname{coeff}_{\deg a_i}(a_{ij})\Big)_{ij} \in {}^m K^n.$$

Example 249. Consider $K = \mathbb{R}$, and let

$$A = \begin{pmatrix} x^2 & x-1 & x+2 \\ x-1 & 0 & 3 \\ x^2 & x^2 - x & x^2+1 \end{pmatrix} \in {}^3\mathbb{R}[x]^3.$$

Then the row degrees of A are

$$\operatorname{rdeg}_1(A) = 2, \qquad \operatorname{rdeg}_2(A) = 1, \qquad \text{and} \qquad \operatorname{rdeg}_3(A) = 2;$$

and the order of A is $\operatorname{ord} A = 5$. The leading coefficient matrix of A is

$$\mathrm{LCM}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \in {}^{3}\mathbb{R}^{3}.$$

Definition 250 (Row Reduced). A matrix $A \in {}^{m}R^{n}$ is called *row reduced* if $\mathrm{LCM}(A)$ has full row rank.

Example 251. The matrix A from Example 249 is not row reduced. On the other hand, the matrix

$$B = \begin{pmatrix} x & x-1 & -2x+2 \\ x-1 & 0 & 3 \\ x^2 & x^2-x & x^2+1 \end{pmatrix} \in {}^{3}\mathbb{R}[x]^{3} \qquad \text{with} \qquad \mathrm{LCM}(B) = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

is row reduced. Note that B is row equivalent to A.

Algorithm 252 (Row Reduction). *Input* A matrix $A \in {}^{m}R^{n}$.

Output A row reduced matrix $B \in {}^{r}R^{n}$ with $r \leqslant m$ and a unimodular matrix $Q \in \mathrm{GL}_m(R)$ such that

$$QA = \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix}.$$

Procedure

(a) Initialise $Q \leftarrow \mathbf{1}_m$.

(b) Swap all non-zero rows of A to the top and mimick the transformations on Q. Then delete all zero rows from A.

(c) Compute $L = \mathrm{LCM}(A)$.

(d) If L has full row rank, then stop and return $B = A$ and Q.

(e) Else,

  (1) Find a vector $v \in K^m \setminus \{0\}$ such that $vL = 0$.

  (2) Let i be such that $v_i \neq 0$ and $\mathrm{rdeg}_i(A)$ is maximal.

  (3) Let

$$U = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ v_1 x^{\mathrm{rdeg}_i(A) - \mathrm{rdeg}_1(A)} & & v_i & & v_m x^{\mathrm{rdeg}_i(A) - \mathrm{rdeg}_m(A)} \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \in \mathrm{GL}_m(R).$$

  (4) Update $A \leftarrow UA$ and $Q \leftarrow \mathrm{diag}(U, \mathbf{1})Q$.

  (5) Go to step (b).

Theorem 253. *Algorithm 252 is correct and terminates.*

Proof. When the algorithm terminates, the resulting matrix $B = A$ will have a regular leading coefficient matrix and hence be row reduced. Moreover, since we always transform A and Q in the same way (and since Q starts out as an identity matrix), we will have

$$QA = \begin{pmatrix} B \\ \mathbf{0} \end{pmatrix}.$$

It remains to show that the transformations are unimodular. We change A in two ways: Swapping and deleting rows in step (b) and multiplication by U in step (e.4). Deleting rows basically amounts to simply ignoring them; so we can view this as elementary row operation. The matrix U defined in step (e.3) is a polynomial matrix since $\operatorname{rdeg}_i(A) - \operatorname{rdeg}_k(A) \geqslant 0$ for $k = 1, \ldots, m$ by the choice of i in step (e.2). Also, U has determinant $\det U = v_i \in K \setminus \{0\} = R^*$, again by the choice of i. Thus U is indeed unimodular.

Now we prove that Algorithm 252 terminates. Consider the order $\operatorname{ord} A$ of A. In each iteration of the algorithm, we (potentially) delete zero rows from A which does not change the order and we multiply A by U defined in step (e.3). Multiplication by U replaces the i^{th} row of A by

$$v_1 x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_1(A)} A_{1,*} + \ldots + v_i A_{i,*} + \ldots + v_m x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_m(A)} A_{m,*}.$$

We first note that every row vector $x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_k(A)} A_{k,*}$ in that sum has degree $\operatorname{rdeg}_i(A)$: Since $A_{k,*}$ has degree $\operatorname{rdeg}_k(A)$ there must be at least one entry of degree $\operatorname{rdeg}_k(A)$. Multiplication by $x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_k(A)}$ raises that degree to $\operatorname{rdeg}_i(A)$. Obviously, also the term $v_i A_{i,*}$ has degree $\operatorname{rdeg}_i(A)$. We now consider the coefficients of $x^{\operatorname{rdeg}_i(A)}$ in the sum. Since multiplication by $x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_k(A)}$ does not change the leading coefficients, we have that

$$\operatorname{coeff}_{\operatorname{rdeg}_i(A)}(x^{\operatorname{rdeg}_i(A) - \operatorname{rdeg}_k(A)} A_{k,*}) = \operatorname{coeff}_{\operatorname{rdeg}_k(A)}(A_{k,*}).$$

Thus the coefficient of $x^{\operatorname{rdeg}_i(A)}$ of the sum is

$$\sum_{k=1}^{m} v_k \operatorname{coeff}_{\operatorname{rdeg}_k(A)}(A_{k,*}) = v \operatorname{LCM}(A) = 0$$

by the choice of v in step (e.1). It follows that multiplication with U replaces the i^{th} row of A with a row of strictly lower degree. Thus, the order $\operatorname{ord} A$ of A strictly decreases in each iteration. However, since the order is a non-negative integer, this can only happen finitely often. Consequently, the algorithm terminates. □

Example 254. Consider $K = \mathbb{Q}$ and the matrix

$$A = \begin{pmatrix} 12 + 12x & 6x + 18 & -12 - 6x \\ 6x - 24 & 3x - 6 & 9 - 3x \\ -20x^2 + 20 & -10x^2 - 20x - 20 & 10x^2 + 10x - 20 \end{pmatrix} \in {}^{3}\mathbb{Q}[x]^{3}.$$

We follow Algorithm 252 to compute the row reduced form of A. None of the rows of A are zero, and thus we do not delete any of them. The leading coefficient matrix is

$$L = \operatorname{LCM}(A) = \begin{pmatrix} 12 & 6 & -6 \\ 6 & 3 & -3 \\ -20 & -10 & 10 \end{pmatrix}$$

An entry in the left kernel of L is for example $v = (1, 8, 3) \in \mathbb{Q}^3$. The row degrees of A are 1, 1, and 2. Thus, for the third entry $v_3 = 3$ we have $v_3 \neq 0$ and $\mathrm{rdeg}_3(A)$ is maximal. This leads to the matrix

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x & 8x & 3 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{Q}[x]).$$

Next, we update A and obtain

$$A \leftarrow UA = \begin{pmatrix} 12 + 12x & 6x + 18 & -12 - 6x \\ 6x - 24 & 3x - 6 & 9 - 3x \\ -180x + 60 & -90x - 60 & 90x - 60 \end{pmatrix}.$$

Note that the degree of the last row decreased. There are no zero rows and the new leading coefficient matrix is

$$L = \mathrm{LCM}(A) = \begin{pmatrix} 12 & 6 & -6 \\ 6 & 3 & -3 \\ -180 & -90 & 90 \end{pmatrix}.$$

Computing the kernel of L, we find that for instance $v = (0, 30, 1)$ is an entry of the kernel. Here, $v_2, v_3 \neq 0$ and the second and third row degree of A are the same. We choose to reduce the third row. This leads to

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 30 & 1 \end{pmatrix} \qquad \text{and} \qquad A \leftarrow UA = \begin{pmatrix} 12 + 12x & 6x + 18 & -12 - 6x \\ 6x - 24 & 3x - 6 & 9 - 3x \\ -660 & -240 & 201 \end{pmatrix}.$$

The leading coefficient matrix of the new A is

$$L = \mathrm{LCM}(A) = \begin{pmatrix} 12 & 6 & -6 \\ 6 & 3 & -3 \\ -660 & -240 & 210 \end{pmatrix}$$

and the left kernel of L is spanned by $v = (1, -2, 0)$. The first two entries are non-zero and the first two row degrees of A are the same. We choose to reduce the first row of A leading to the transformations matrix

$$U = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A \leftarrow UA = \begin{pmatrix} 60 & 30 & -30 \\ 6x - 24 & 3x - 6 & 9 - 3x \\ -660 & -240 & 201 \end{pmatrix}.$$

We obtain

$$L = \mathrm{LCM}(A) = \begin{pmatrix} 60 & 30 & -30 \\ 6 & 3 & -3 \\ -660 & -240 & 210 \end{pmatrix}$$

and the left kernel of L contains $v = (1, -10, 0)$. Here, the first two entries of v are non-zero, but the second row degree of A is higher than the first row degree. Thus, we have to reduce the second row. This leads to

$$U = \begin{pmatrix} 1 & 0 & 0 \\ x & -10 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A \leftarrow UA = \begin{pmatrix} 60 & 30 & -30 \\ 240 & 60 & -90 \\ -660 & -240 & 201 \end{pmatrix}.$$

At this point, we have $L = \mathrm{LCM}(A) = A$ and $\det L = 270000 \neq 0$. Thus, the algorithm terminates with

$$B = A = \begin{pmatrix} 60 & 30 & -30 \\ 240 & 60 & -90 \\ -660 & -240 & 201 \end{pmatrix}$$

as its result. Multiplying all the matrices U together, we obtain the transformation matrix

$$Q = \begin{pmatrix} 1 & -2 & 0 \\ x & -2x - 10 & 0 \\ x & 8x + 30 & 3 \end{pmatrix}$$

which fulfils $QA = B$ (for the original input A).

Exercise 255. Apply Algorithm 252 to the matrix

$$\begin{pmatrix} 2 - 2x & 2x^2 - x + 1 & -1 & 1 - 2x \\ -2x^2 + 2x + 2 & -1 + x & -2 - x & 2x^3 + 2x - 1 \\ -2x^3 + 2x + 1 & x^2 - 2x - 1 & x^2 - 2x - 2 & 2x^2 - x - 2 \\ 2 + 2x & -2x^3 - 2 & 2 & -2x^2 + 2 \end{pmatrix} \in {}^4\mathbb{Q}[x]^4.$$

Exercise 256. Implement Algorithm 252 in a computer algebra system of your choice.

Definition 257 (Leading Coefficient). Let $p \in R \setminus \{0\}$. Then $\mathrm{lc}(p) = \mathrm{coeff}_{\deg p}(p)$ is the *leading coefficient* of p. Similarly, for a matrix (or vector) $A \in {}^mR^n$ we let $\mathrm{lc}(A) = \mathrm{coeff}_{\deg A}(A)$.

Remark 258. An equivalent way of defining the leading coefficient is to write $p \in R$ as $p = p_d x^d + p_{d-1}x^{d-1} + \ldots + p_1 x + p_0$ with coefficients $p_0, \ldots, p_d \in K$. If $p_d \neq 0$, then $\mathrm{lc}(p) = p_d$. Similarly, we can write a matrix $A \in {}^mR^n$ as $A = A_d x^d + A_{d-1}x^{d-1} + \ldots + A_1 x + A_0$ with coefficients $A_0, \ldots, A_d \in {}^mK^n$. Again, if $A_d \neq \mathbf{0}$, then $\mathrm{lc}(A) = A_d$.

Remark 259. With the leading coefficient we have

$$\mathrm{LCM}(A) = \begin{pmatrix} \mathrm{lc}(A_{1,*}) \\ \vdots \\ \mathrm{lc}(A_{m,*}) \end{pmatrix}$$

for all $A \in {}^mR^n$.

Exercise 260. Let $p \in R$ and $v \in R^n$. Show that $\mathrm{lc}(pv) = \mathrm{lc}(p)\,\mathrm{lc}(v)$.

Theorem 261. *Let $G \in {}^mR^n$ be a matrix where all rows are non-zero. Then the following statements are equivalent:*

(a) G is row reduced.

(b) G has full row rank and $\operatorname{ord} G$ is minimal among all row equivalent matrices.

(c) The K-dimension of $V_d = \{v \in R^n G \mid \deg v < d\}$ is

$$\dim_K V_d = \sum_{\substack{i=1 \\ \mathrm{rdeg}_i(G) \leqslant d}}^{m} \big(d - \mathrm{rdeg}_i(G)\big)$$

for all $d \geqslant 0$.

(d) If $v \in R^m$ *and* $w = vG$, *then* $\deg w = \max\{\deg v_i + \mathrm{rdeg}_i(G) \mid i = 1, \ldots, m\}$.

Proof. We show first that property (b) implies (a). Assume that $\operatorname{ord} G$ is minimal. Then $\mathrm{LCM}(G)$ must have full rank because else we could construct a transformation matrix as in Algorithm 252 (in step (e.3)) which when multiplied to G yielded a matrix of smaller order.

We prove now that (a) implies (d). We denote the maximum by $\mu = \max\{\deg v_i + \mathrm{rdeg}_i(G) \mid i = 1, \ldots, m\}$. We first note that $w = v_1 G_{1,*} + \ldots + v_m G_{m,*}$ and thus $\deg w \leqslant \deg(v_i G_{i,*}) = \deg v_i + \mathrm{rdeg}_i(G)$ for all $i = 1, \ldots, m$. That is, $\deg w \leqslant \mu$. Assume now that $\deg w$ was strictly less than the maximum μ. Let $1 \leqslant i_1 < \ldots < i_\ell \leqslant m$ be the row indices for which $v_{i_j} \neq 0$ and $\deg(v_{i_j} G_{i_j,*}) = \mu$. Since $\deg w < \mu$ we must have $\deg(v_{i_1} G_{i_1,*} + \ldots + v_{i_\ell} G_{i_\ell,*}) < \mu$; and thus

$$\mathrm{coeff}_\mu(v_{i_1} G_{i_1,*} + \ldots + v_{i_\ell} G_{i_\ell,*}) = 0.$$

Using Exercise 260, we obtain

$$\mathrm{lc}(v_{i_1})\,\mathrm{lc}(G_{i_1,*}) + \ldots + \mathrm{lc}(v_{i_\ell})\,\mathrm{lc}(G_{i_\ell,*}) = 0.$$

However, this yields a non-trivial relation of the rows of $\mathrm{LCM}(G)$ (using Remark 259). Thus, G cannot be row reduced. The claim follows by contraposition.

Next, we show that (d) implies (c). By property (d), if $w = vG \in V_d$, then $\max\{\deg v_i + \mathrm{rdeg}_i(G) \mid i = 1, \ldots, m\} < d$ or, equivalently, $\deg v_i < d - \mathrm{rdeg}_i(G)$ for all $i = 1, \ldots, m$. The vector space of polynomials of degree less than $d - \mathrm{rdeg}_i(G)$ is either empty and has thus dimension 0 if $d < \mathrm{rdeg}_i(G)$, or it has dimension equal to $d - \mathrm{rdeg}_i(G)$. This implies (c).

Finally, we demonstrate that (c) implies (b). We can without loss of generality reorder the rows of G such that $\mathrm{rdeg}_1(G) \leqslant \ldots \leqslant \mathrm{rdeg}_m(G)$. By the assumption (c), the subspace V_d consists of the vectors vG with $v \in R^m$ and $\deg v_i < d - \mathrm{rdeg}_i(G)$ for $i = 1, \ldots, m$. Consequently, there are fewer than i linearly independent (over R) vectors in $R^m G$ of degree less than i. Thus, by induction on i, any i linearly independent vectors in $R^m G$ must have a sum of degrees at least $\sum_{j=1}^{i} \mathrm{rdeg}_j(G)$. □

Remark 262. Because of property (b) of Theorem 261 (the rows of) row reduced matrices are referred to as *minimal basis* by some authors.

Remark 263. Property (d) of Theorem 261 is also known as the *predictable degree property*.

Remark 264. As part (b) of Theorem 261 shows, the rows of any row reduced matrix are linearly independent over R. Thus, computing a row reduced matrix via Algorithm 252 is a rank-revealing transformation. Consequently, we can use row reduction to compute kernels as in Theorem 191 or matrix greatest common divisors as in Remark 206.

Application 265. We can use the row reduction in Algorithm 252 in order to invert polynomial matrices. Let $A \in \mathrm{GL}_n(R)$. Apply row reduction to obtain

$$QA = B$$

where $B \in {}^nR^n$ is row reduced and $Q \in \mathrm{GL}_n(R)$ is unimodular. We cannot have any zero rows in B since A has full rank. Moreover, we have $R^n B = R^n A = R^n$ since A is unimodular. Therefore the identity matrix forms another possible basis for the row space of B. We have $\operatorname{ord} \mathbf{1}_n = 0$. By Theorem 261 this implies $\operatorname{ord} B = 0$ as well. Thus, $B \in {}^nK^n$ and $B = QA$ is invertible; that is, $B \in \mathrm{GL}_n(K)$. In total, we obtain $A^{-1} = B^{-1} Q$.

Example 266. Example 251 provides an example for Application 265: We had

$$A = \begin{pmatrix} 12x+12 & 6x+18 & -6x-12 \\ 6x-24 & 3x-6 & -3x+9 \\ -20x^2+20 & -10x^2-20x-20 & 10x^2+10x-20 \end{pmatrix} \in {}^3\mathbb{Q}[x]^3$$

and we computed

$$QA = B = \begin{pmatrix} 60 & 30 & -30 \\ 240 & 60 & -90 \\ -660 & -240 & 210 \end{pmatrix} \qquad \text{where} \quad Q = \begin{pmatrix} 1 & -2 & 0 \\ x & -2x-10 & 0 \\ x & 30+8x+3 & \end{pmatrix} \in \mathrm{GL}_3(\mathbb{Q}[x]).$$

The matrix $B \in {}^3\mathbb{Q}^3$ is invertible and we obtain

$$\begin{aligned} A^{-1} = B^{-1}Q = \frac{1}{300} \begin{pmatrix} -10 & 1 & -1 \\ 10 & -8 & -2 \\ -20 & -6 & -4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 0 \\ x & -2x-10 & 0 \\ x & 30+8x & 3 \end{pmatrix} \\ = \frac{1}{300} \begin{pmatrix} -10 & -20-10x & -3 \\ 10-10x & 0 & -6 \\ -20-10x & -20-20x & -12 \end{pmatrix}. \end{aligned}$$

Exercise 267. Compute the inverse of

$$\begin{pmatrix} -x-1 & -x-1 & 1 \\ 2 & 1 & 0 \\ 2-2x & 2x-1 & -6 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{Q}[x])$$

using the method from Application 265.

Exercise 268. Prove that a matrix $A \in {}^nR^n$ is invertible if and only if row reduction (Algorithm 252) yields a matrix $B \in \mathrm{GL}_n(K)$. That is, prove that Application 265 can be used to decide whether a matrix is invertible or not.

Definition 269 (Popov Normal Form). A matrix $A = (a_{ij})_{ij} \in {}^mR^n$ is in *Popov normal form*[15] if

(a) $\mathrm{rdeg}_i(A) \leqslant \mathrm{rdeg}_{i+1}(A)$ for all $i = 1, \ldots, m-1$;

(b) there exist column indices $j_1, \ldots, j_m$ (the *pivot indices*) such that

   (1) a_{i,j_i} is monic and $\mathrm{rdeg}_i(A) = \deg a_{i,j_i}$ for all $i = 1, \ldots, m$,

   (2) $\deg a_{ik} < \mathrm{rdeg}_i(A)$ if $k < j_i$,

   (3) $\deg a_{k,j_i} < \mathrm{rdeg}_i(A)$ if $k \neq i$, and

   (4) if $\mathrm{rdeg}_i(M) = \mathrm{rdeg}_k(M)$ and $i < k$, then $j_i < j_k$.

If A is in Popov normal form with pivot indices $j_1, \ldots, j_m$, then we call $a_{1,j_1}, \ldots, a_{m,j_m}$ the *pivots* of A.

Theorem 270. *If $A \in {}^mR^n$ is in Popov normal form, then up to permutation of rows the leading coefficient matrix* $\mathrm{LCM}(A)$ *of A is in row echelon form. In particular, A is row reduced.*

[15] Also called "polynomial-echelon form" by some authors.

Proof. We remark first, that the pivot indices must be pairwise different: If $i \neq k$ and $j_i = j_k$, then by properties (b.1) and (b.3) of Definition 269 we had

$$\deg a_{k,j_i} < \operatorname{rdeg}_i(A) = \deg a_{i,j_i} = \deg a_{i,j_k} < \operatorname{rdeg}_k(A)$$

and similarly $\operatorname{rdeg}_k(A) < \operatorname{rdeg}_i(A)$ which cannot both be true at the same time.

Since permutations are allowed, we permute the rows of A in such a way that that $j_1 < \ldots < j_m$. This will potentially violate property (a) of Definition 269; however, we do not need that property for the proof. By property (b.1), the entries at position (i, j_i) of $\operatorname{LCM}(A)$ will be simply 1 for all $i = 1, \ldots, m$. By property (b.2), everything to the left in the same row of such an entry will be 0. Since the pivots are in different columns, this concludes the proof. □

Remark 271. Theorem 270 explains the choice of the names "pivot" and "pivot indices" in Definition 269.

Exercise 272. The converse of Theorem 270 is not true. Find a counter example.

Theorem 273. *Let $A, B \in {}^mR^n$ be both in Popov normal form and assume that there exists $Q \in \mathrm{GL}_m(R)$ such that $QA = B$. Then $A = B$.*

Proof. We denote the pivot indices of A by $j_1, \ldots, j_m$ and those of B by $k_1, \ldots, k_m$. By property (a) of Definition 269 we have

$$\operatorname{rdeg}_1(A) \leqslant \ldots \leqslant \operatorname{rdeg}_m(A) \qquad \text{and} \qquad \operatorname{rdeg}_1(B) \leqslant \ldots \leqslant \operatorname{rdeg}_m(B).$$

Decompose A and B into blocks where the row of each block have the same degree

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_s \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} B_1 \\ \vdots \\ B_t \end{pmatrix}.$$

Let A_i consist of μ_i rows for $i = 1, \ldots, s$ and let B_ℓ have ν_ℓ rows for $\ell = 1, \ldots, t$. Also, decompose $Q = (q_{i\ell})_{i\ell}$ into the same blocks as B and Q^{-1} into the same blocks as A.

$$Q = \begin{pmatrix} Q_1 \\ \vdots \\ Q_t \end{pmatrix} \qquad \text{and} \qquad Q^{-1} = \begin{pmatrix} W_1 \\ \vdots \\ W_s \end{pmatrix}.$$

We first claim that $\deg A_1 = \deg B_1$. Assume that was not the case and $\deg A_1 > \deg B_1$. Then, since $Q_{1,*}A = B_{1,*}$ we had a non-zero R-linear combination of rows of A with a smaller degree than any of the rows of A. Since A is row reduced by Theorem 270, this violates the predictable degree property (property (d) of Theorem 261). Similarly, by reversing the roles of A and B, we see that $\deg B_1 > \deg A_1$ is not possible. Invoking again the predictable degree property, we see that the rows of B_1 must be linear combinations of the rows of A_1 and vice versa (none of the other rows of A or B respectively can contribute since their degrees are too high). Thus, the rows of A_1 and B_1 span the same space. Since they are linearly independent, both sets of rows are bases for this space. That implies, that A_1 and B_1 have the same number of rows; that is, $\mu_1 = \nu_1$. Thus, we can write

$$A_1 = A_{1,d_1}x^{d_1} + \ldots + A_{1,1}x + A_{1,0} \qquad \text{and} \qquad B_1 = B_{1,d_1}x^{d_1} + \ldots + B_{1,1}x + B_{1,0}$$

where $d_1 = \deg A_1 = \deg B_1$ and $A_{1,0}, \ldots, A_{1,d_1}, B_{1,0}, \ldots, B_{1,d_1} \in {}^{\mu_1}K^n$. Moreover, the predictable degree property actually implies that the rows of B_1 are K-linear combinations of the rows of A_1 and vice versa (again since otherwise the degrees would not match). Thus,

$$Q_1 = \begin{pmatrix} Q_{11} & \mathbf{0} \end{pmatrix}$$

where $Q_{11} \in {}^{\mu_1}K^{\mu_1}$. Write now

$$Q = \begin{pmatrix} Q_{11} & \mathbf{0} \\ * & W \end{pmatrix}$$

for some $W \in {}^{m-\mu_1}R^{m-\mu_1}$. Then using the Leibniz formula we obtain $\det Q = \det Q_{11} \det W \in R^*$ which implies that $\det Q_{11}$ is a unit. Consequently, Q_{11} is invertible. From $Q_{11}A_1 = B_1$ we conclude that $Q_{11}A_{1,d_1} = B_{1,d_1}$ since Q_{11} is a constant matrix. In other words, A_{1,d_1} and B_{1,d_1} are row equivalent. Since all rows of A_1 have the same degrees, we find that $\operatorname{LCM}(A_1) = A_{1,d_1}$. Moreover, the rows of A_1 are still in Popov normal form since it is a submatrix of A. By Theorem 270 this means that A_{1,d_1} is in row echelon form. Looking closer at the proof, we see that property (b.3) of Definition 269 implies that A_{1,d_1} is actually in reduced row echelon form. Similarly B_{1,d_1} is in reduced row echelon form. By the uniqueness of the reduced row echelon form, we must have $A_{1,d_1} = B_{1,d_1}$ and $Q_{11} = \mathbf{1}_{\mu_1}$. But this also implies $A_1 = B_1$.

Consider now B_2. First assume that $\deg B_2 < \deg A_2$. By the predictable degree property (d) of Theorem 261 and by $QA = B$, this implies that the rows of B_2 are in the row space of A_1. However, since we already know that $A_1 = B_1$ and since the rows of B are linearly independent by part (b) of Theorem 261, this is impossible. Similarly, by switching the roles of A and B, we find that $\deg B_2 > \deg A_2$ is also not possible. Thus, we must have $\deg A_2 = \deg B_2$. Moreover, the rows of B_2 are generated by the rows of A_1 and A_2 since the other blocks of A cannot contribute. Thus, for some matrices Q_{21} and Q_{22} we have

$$B_2 = Q_{21}A_1 + Q_{22}A_2; \qquad \text{that is,} \qquad Q_2 = \begin{pmatrix} Q_{21} & Q_{22} & \mathbf{0} \end{pmatrix}.$$

We consider the leading coefficient matrix of B_2. By the predictable degree property, $\operatorname{LCM}(B_2)$ is generated by the rows of $\operatorname{LCM}(A_2)$ and $\operatorname{LCM}(A_1)$. Since $A_1 = B_1$ and thus $\operatorname{LCM}(A_1) = \operatorname{LCM}(B_1)$, the pivots of A_1 and B_2 cannot be in the same columns. Moreover, since $\deg A_1 < \deg B_2$ and by property (b.3), every column of $\operatorname{LCM}(B_2)$ where A_1 has a pivot must be zero. In other words, $\operatorname{LCM}(A_1)$ cannot contribute to $\operatorname{LCM}(B_2)$. That means that the rows of $\operatorname{LCM}(A_2)$ generate those of $\operatorname{LCM}(B_2)$. Conversely, switching the roles of A and B, we see that also the rows of $\operatorname{LCM}(B_2)$ generate those of $\operatorname{LCM}(A_2)$. Since both matrices are in reduced row echelon form they must thus be equal. That means, $Q_{22} = \mathbf{1}$. This in turn implies $B_2 = A_2 + Q_{21}A_1$.

Assume now that $Q_{21} \neq \mathbf{0}$. Then at least one row of B_2 which we will call b is partly generated from the rows of A_1; that is,

$$b = a + \sum_{i=1}^{\mu_1} c_i r_i$$

where a is the corresponding row of A_2, $c_1, \ldots, c_{\mu_1} \in R$ are polynomials, and $r_1, \ldots, r_{\mu_1}$ are the rows of A_1. Choose the smallest index $1 \leqslant \ell \leqslant \mu_1$ such that c_ℓ is of maximal degree. Then the left-most entry of highest degree of $c_1r_1 + \ldots + c_{\mu_1}r_{\mu_1}$ originates from the pivot of r_ℓ; that is, it will be at position j_ℓ and have a degree of $\deg c_\ell + \operatorname{rdeg}_\ell(A_1) \geqslant \operatorname{rdeg}_\ell(A_1)$. Since the entry of a at position j_ℓ has a degree strictly smaller than $\operatorname{rdeg}_\ell(A_1)$ by property (b.3) of Definition 269, this implies that the entry of b at position j_ℓ has degree larger or equal to $\operatorname{rdeg}_\ell(A_1)$. However, since we have $j_\ell = k_\ell$ and $\operatorname{rdeg}_\ell(A_1) = \operatorname{rdeg}_\ell(B_1)$ because of $A_1 = B_1$, this violates property (b.3) of Definition 269 for B. Thus, this is impossible and we must have $Q_{21} = \mathbf{0}$. Consequently, $A_2 = B_2$.

We can now apply the same argument to B_3 and then to B_4 and so on. This will show that $B_3 = A_3$, $B_4 = A_4$ and so forth. In total this leads to $A = B$. □

Algorithm 274 (Popov Normal Form). *Input* A matrix $A \in {}^mR^n$.

Output A matrix $P \in {}^mR^n$ in Popov normal form and a matrix $Q \in \mathrm{GL}_m(R)$ such that $QA = P$.

Procedure

```
(a) Use Algorithm 252 in order to row reduce the matrix $A$. Call the result $A_1$ and call the
    transformation matrix $Q_1$.
(b) Sort the rows of $A_1$ with respect to their degrees in order to obtain
        $$A_2 = \begin{pmatrix} B_1 \\ \vdots \\ B_\ell \\ \mathbf{0}_{m_0 \times n} \end{pmatrix}$$
    where the blocks $B_1 \in {}^{m_1}R^n, \ldots, B_\ell \in {}^{m_\ell}R^n$ consist of non-zero rows of equal degree
    and where $\deg B_1 < \ldots < \deg B_\ell$. Mimick the same transformation of $Q_1$ obtaining $Q_2$.
(c) For each $j = 1, \ldots, \ell$:
    (1) Compute $L_j = \mathrm{LCM}(B_j) \in {}^{m_j}K^n$.
    (2) Compute a matrix $W_j \in \mathrm{GL}_{m_j}(K)$ such that $W_j L_j$ is in reduced row echelon form.
    (3) Let
            $$D_j = \operatorname{diag}(\mathbf{1}_{m_1+\ldots+m_{j-1}}, W_j, \mathbf{1}_{m_{j+1}+\ldots+m_\ell+m_0})$$
        and set $A_2 \leftarrow D_j A_2$ (updating the blocks $B_1, \ldots, B_\ell$ accordingly) and $Q_2 \leftarrow D_j Q_2$.
    (4) Let $\nu_{j1}, \ldots, \nu_{jm_j}$ be the pivot indices of $L_j$.
    (5) For $i = m_1 + \ldots + m_{j+1} + 1, \ldots, m$ and for $k = 1, \ldots, m_j$:
        (i) Subtract
                $$\frac{\operatorname{lc}((A_2)_{i,\nu_{jk}})}{\operatorname{lc}((B_j)_{k,\nu_{jk}})} x^{\operatorname{rdeg}_i(A_2) - \deg B_j}$$
            times the $(m_1 + \ldots + m_j + k - 1)^{\text{th}}$ row of $A_2$ from the $i^{\text{th}}$ row. Do the same
            for $Q_2$.
            (This eliminates the highest degree term in the $i^{\text{th}}$ row and $\nu_{jk}{}^{\text{th}}$ column of $A_2$.)
(d) Return $P = A_2$ and $Q = Q_2$.
```

Theorem 275. *Algorithm 274 terminates and is correct.*

Proof. The termination of the algorithm is obvious. Moreover, since the matrix Q just records all the row transformations, we obviously have $Q \in \mathrm{GL}_m(R)$ and $QA = P$. Thus, we only have to show that P is in Popov normal form.

For this, we check the conditions of Definition 269. We start with property (a) of the definition. At the same time, we will prove that the matrix A_1 remains row reduced during the transformations in step (c) of Algorithm 274. Both statements are true when we enter the loop in step (c) since the matrix A_2 is just a row permutation of the row reduced matrix A_1 where the rows are sorted with respect to their degree.

We claim that the properties continue to be true during step (c) of the algorithm. Since A_1 is row reduced, also all the submatrices $B_1, \ldots, B_\ell$ must be row reduced. For a given j, step (c.3) replaces B_j with $W_j B_j$. Write $B_j = B_{j,d_j} x^{d_j} + \ldots + B_{j,0}$ where $d_j = \deg B_j$ and where $B_{j,0}, \ldots, B_{j,d_j} \in {}^{m_j}K^n$ are constant matrices. Obviously, we have $L_j = B_{j,d_j}$ since every row of B_j has dgeree d_j. Then $W_j B_j = (W_j L_j) x^{d_j} + (W_j B_{j,d_j-1}) x^{d_j-1} \ldots + (W_j B_{j,0})$. Consequently, the new leading coefficient matrix is $\operatorname{LCM}(W_j B_j) = W_j L_j$ because L_j has full rank (as B_j is row reduced) and W_j is invertible and thus no row of $W_j L_j$ is zero. In particular, the degree of all rows of $W_j B_j$ is still d_j such that the rows of the modified matrix A_2 are still sorted with respect to to their degrees. Moreover, the (non-zero) rows of the leading coefficient matrix of A_2 are still independent meaning that A_2 is still row reduced.

In step (c.5) of Algorithm 274. We subtract a multiple of the k^{th} row of B_j from a row of a block B_h where $h > j$. Denote the k^{th} row of B_j by v and the modified row of B_h by w; also denote the degree of B_h by $d_h = \deg B_h$. The multiplier $\mu = \operatorname{lc}(w_{\nu_{jk}}) / \operatorname{lc}(v_{\nu_{jk}}) x^{d_h - d_j}$ is chosen in such a way that the highest degree term in the ${\nu_{jk}}^{\text{th}}$ position of w is eliminated. Since $\deg(\mu v) = d_j + (d_h - d_j) = \deg w$, the degree of $w - \mu v$ is at most d_h. However, since A_2 is row reduced, the transformation cannot lower the order of A_2 which is minimal by Theorem 261. Thus, the degree of $w - \mu v$ must be equal to that of w. In particular, are the row of A_2 after the elimination still sorted with respect to their degrees. In addition, the transformation changes the leading coefficient matrix of A_2 by adding a multiple of one row to another. Since $\operatorname{LCM}(A_2)$ had full rank initially, the same is true after the transformation. Consequently, A_2 remains row reduced during the entire loop.

We turn now to property (b) of Definition 269. We will show that within step (c) of Algorithm 274 each block of the matrix A_2 is transformed in such a way that for each block B_j with $j = 1, \ldots, \ell$ (a) the leading coefficient matrix is in reduced row echelon form, and (b) the entries in the other blocks B_k with $j < k$ in columns corresponding to the pivots of $\operatorname{LCM}(B_j)$ have a degree less than $\deg B_k$. The latter point implies immediately property (b.3) of Definition 269; while the former point implies the properties (b.1), (b.2), and (b.4).

We proceed by induction on the index j of the block. For $j = 1$, the first claim is easy to verify since in step (c.3) of Algorithm 274 we convert the leading matrix of block B_1 to reduced row echelon form. In step (c.5), for $k > j$ we eliminate the terms of degree $\deg B_k$ in the columns corresponding to the pivots of B_j from the block B_k. Since for each higher degree block we always start with the leftmost pivot, the later eliminations cannot undo the first eliminations. Thus, we will indeed have that all entries in the other blocks B_k in those columns corresponding to pivots of B_j have a degree lower than $\deg B_k$.

Let now $j \geqslant 2$. Again, after step (c.3) of Algorithm 274 the leading coefficient matrix of B_j will be in reduced row echelon form. Moreover, its columns corresponding to pivots of lower degree block will be zero because they have been zero before step (c.3). During the reduction step (c.5) for each $k > j$ we remove the terms of degree $\deg B_k$ from the block B_k in the columns corresponding to the pivots of B_j. This cannot introduce entries of degree $\deg B_k$ in the columns corresponding to the pivots of lower degree blocks since the corresponding entries in B_j have a degree strictly less than $\deg B_j$ by the induction hypothesis. Thus, the second claim also holds in this case. □

Example 276. Consider the matrix

$$A = \left(\begin{array}{cccc} x^2 & x & x^2+x & 0 \\ x^2+x & x^2+x+1 & -x^2 & x^2+x+2 \\ \hline x^3+2x & 2x^3+x^2 & 2x^2 & x^3+x^2+1 \\ 0 & 2x^3+x^2+2x+1 & x^3+x^2 & x+2 \end{array}\right) \in {}^4\mathbb{Q}[x]^4.$$

The rows of A are already sorted with respect to their degree and we have two blocks, one of degree 2 and one of degree 3. Moreover since the leading coefficient matrix

$$\mathrm{LCM}(A) = \left(\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 1 \\ \hline 1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{array}\right)$$

has full rank (the determinant is -1), A is also already row reduced. Thus, we are in the situation after step (b) of Algorithm 274 with $A_2 = A$.

We enter the loop in step (c). First we consider the block of the degree 2 rows. Its leading coefficient matrix is

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

with reduced row echelon form

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

and transformation matrix

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

We apply the same transformation to the degree 2 block of the matrix A_2. This is the same as doing the multiplication

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A_2 = \left(\begin{array}{cccc} x^2 & x & x^2+x & 0 \\ x & x^2+1 & -2x^2-x & x^2+x+2 \\ \hline x^3+2x & 2x^3+x^2 & 2x^2 & x^3+x^2+1 \\ 0 & 2x^3+x^2+2x+1 & x^3+x^2 & x+2 \end{array}\right).$$

Now, we need to eliminate the degree 3 entries in the lower block for those columns corresponding to the pivots of the upper block, that is, for the first and second column. For this, we subtract x the first row of A_2 from the third and then we subtract $2x$ times the second row from the third. This yields the new matrix

$$\left(\begin{array}{cccc} x^2 & x & x^2+x & 0 \\ x & x^2+1 & -2x^2-x & x^2+x+2 \\ \hline -2x^2+2x & -2x & 3x^3+3x^2 & -x^3-x^2-4x+1 \\ 0 & 2x^3+x^2+2x+1 & x^3+x^2 & x+2 \end{array}\right).$$

Similarly, we eliminate the degree 3 entries in the first two columns of the the last row obtaining

$$\left(\begin{array}{cccc} x^2 & x & x^2+x & 0 \\ x & x^2+1 & -2x^2-x & x^2+x+2 \\ \hline -2x^2+2x & -2x & 3x^3+3x^2 & -x^3-x^2-4x+1 \\ -2x^2 & x^2+1 & 5x^3+3x^2 & -2x^3-2x^2-3x+2 \end{array}\right).$$

The leading coefficient matrix of the lower block is now

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

which transforms into the reduced row echelon form

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

Applying this to the lower block of A_2 yields the Popov normal form

$$\left(\begin{array}{cccc} x^2 & x & x^2+x & 0 \\ x & x^2+1 & -2x^2-x & x^2+x+2 \\ \hline -2x^2+4x & -x^2-4x-1 & x^3+3x^2 & -5x \\ -4x^2+10x & -3x^2-10x-3 & 6x^2 & x^3+x^2-11x-1 \end{array}\right).$$

We can check that indeed the properties of Definition 269 hold for this matrix.

Exercise 277. Use Algorithm 274 in order to compute the Popov normal form of

$$\begin{pmatrix} 2x^3+28x^2+16x & 10x^3+25x^2+38x+23 & 4x^3-6x^2+8x & x^3+19x^2+27x+45 \\ x^3-x^2-x & -x^3-x^2-3x-2 & -2x^3-x^2-x & x^3-4 \\ -2x^3+3x^2+3x & 2x^3+3x^2+7x+5 & 4x^3+x^2+2x & -2x^3+x^2+x+10 \\ -4x^3-18x^2-7x & -8x^3-16x^2-24x-12 & 2x^2-5x & -3x^3-12x^2-20x-23 \end{pmatrix}$$

with $R = \mathbb{Q}[x]$.

Exercise 278. Implement Algorithm 274 in a computer algebra system of your choice.

Remark 279. We can use computer algebra systems to compute the Popov normal form:

Maple The command is called PopovForm and it resides in the LinearAlgebra package. It computes the *column* Popov normal form instead of the row Popov normal form from Definition 269.

Sage In Sage we have the weak_popov_form method.

A Solving Linear Ordinary Differential Equations

In this chapter we do not worry about the analytical implications of differential equations but rather present some simplified solution methods. For more details, please see a textbook on differential equations.

Notation 280. In this section, we will denote the derivative of f with respect to x by $f' = df/dx$. Higher derivatives are denoted by $f'' = d^2f/dx^2$, $f''' = d^3f/dx^3$, and $f^{(n)} = d^nf/dx^n$ for $n \geqslant 0$.

Remark 281 (Integrating Factor). Consider the inhomogeneous first order equation

$$f' + pf = q$$

where p, q are $C^\infty(\mathbb{R})$ functions. Consider

$$\mu = e^{\int p\, dx}.$$

(We only need one particular solution here, it is not necessary to introduce the constant.) Note that

$$\mu' = p\mu$$

by the chain rule. Multiplying the original equation by μ yields

$$\mu q = \mu f' + p\mu f = \mu f' + \mu' f = (\mu f)'$$

using the product rule. Thus,

$$\mu f = \int \mu q \, dx$$

and consequently

$$f = \mu^{-1} \int \mu q \, dx.$$

Example 282. Consider the equation

$$x f' = f + x^3 \sin x.$$

We rewrite the function into the form which we have in Remark 281 obtaining

$$f' - \frac{1}{x} f = x^2 \sin x.$$

Here, $p = -1/x$ and $q = x^2 \sin x$. Thus, the integrating factor is

$$\mu = e^{-\int \frac{dx}{x}} = e^{-\ln x} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

This implies that the solution is

$$\begin{aligned} f = \mu^{-1} \int \mu q \, dx = x \int x \sin x \, dx &= x\Big(-x\cos x + \int \cos x \, dx\Big) \\ &= x\Big(-x\cos x + \sin x + C\Big) = x\sin x - x^2 \cos x + Cx \end{aligned}$$

where C is an arbitrary constant. Note that we used integration by parts (with $u = x$ and $dv = \sin x \, dx$) in order to do the integral.

Definition 283 (Fundamental System). Let $a_0, \dots, a_{n-1}$ be functions. A *fundamental system* for the equation

$$f^{(n)} + a_{n-1} f^{(n-1)} + \ldots + a_1 f' + a_0 f = 0$$

of order n is a family $y_1, \dots, y_n$ of n functions which are linearly independent over the constants.

Remark 284. Consider the homogenous n^{th} order linear ordinary differential equation

$$c_n f^{(n)} + c_{n-1} f^{(n-1)} + \ldots + c_1 f' + c_0 f = 0$$

where the coefficients $c_0, \dots, c_n \in \mathbb{R}$ are real constants. We write the left hand side as operator $\chi = c_n \partial^n + \ldots + c_1 \partial + c_0$; that is, the equation is $\chi \bullet f = 0$ using the action of $\mathbb{R}[\partial]$ on $C^\infty(\mathbb{R})$

as defined in Example 38. (This χ is also called the *characteristic polynomial* of the equation.) Assume that we have a factorisation

$$\chi = c(\partial - a_1)^{e_1} \dots (\partial - a_k)^{e_k}$$

where $c \in \mathbb{R}$, $a_1, \dots, a_k \in \mathbb{C}$ are the distinct roots and their multiplicities are $e_1, \dots, e_k \geqslant 1$. Consider a single factor $(x - a_j)^{e_j}$ where we assume for the moment that $a_j \in \mathbb{R}$ is real. Obviously, if f fulfills $(x - a_j) \bullet f = 0$, then f also fulfills $\chi \bullet f = 0$. We claim that $(\partial - a_j)^{e_j} \bullet x^k e^{a_j x} = 0$ for $k = 0, \dots, e_j - 1$. For this, note that

$$(\partial - a_j) \bullet x^k e^{a_j x} = a_j x^k e^{a_j} + k x^{k-1} e^{a_j x} - a_j x^k e^{a_j x} = k x^{k-1} e^{a_j x}$$

and that $(\partial - a_j) \bullet e^{a_j x} = 0$. From this, the claim follows by induction. Assume now that $a_j = u + iv \in \mathbb{C}$ was a complex root. The same computation as above applies; that is, $x^k e^{(u+iv)x}$ for $k = 0, \dots, e_j - 1$ are solutions. However, these are complex valued functions, while we are searching for real valued solutions. Recall that since χ is a real polynomial, also the conjugate $\overline{a_j} = u - iv$ must be a root of χ of the same multliplicity. That is, also $x^k e^{(u-iv)x}$ for $k = 0, \dots, e_j - 1$ are solutions. We know combine two solutions with the same power of x using Euler's formula $e^{ix} = \cos x + i \sin x$. Let $c_+, c_- \in \mathbb{C}$ be complex numbers. Then

$$\begin{aligned}
c_+ x^k e^{(u+iv)x} + c_- x^k e^{(u+iv)x} &= x^k \big(c_+ e^{ux} e^{i(vx)} + c_- e^{ux} e^{i(-vx)} \big) \\
&= x^k \big(c_+ e^{ux} (\cos vx + i \sin vx) + c_- e^{ux} (\cos vx + i \sin(-vx)) \big) \\
&= x^k \big(c_+ e^{ux} (\cos vx + i \sin vx) + c_- e^{ux} (\cos vx - i \sin vx) \big) \\
&= x^k \big(c_+ e^{ux} \cos vx + i c_+ e^{ux} \sin vx + c_- e^{ux} \cos vx - i c_- e^{ux} \sin vx) \big) \\
&= x^k \big((c_+ + c_-) e^{ux} \cos vx + i (c_+ - c_-) e^{ux} \sin vx \big)
\end{aligned}$$

where used the fact that $\sin(-x) = -\sin x$ for all x in the third equation. Now, we set $c_+ = c_- = 1/2$. Then the expression becomes

$$x^k e^{ux} \cos vx$$

which is a real valued function. Similarly, setting $c_+ = i/2$ and $c_- = -i/2$ leads to another real valued function

$$-x^k e^{ux} \sin vx.$$

In total, we have found n different real solutions to the equation. It is possible to show that they are all linearly independent over the real numbers. That means, we have found a fundamental system.

Example 285. Consider the equation

$$f^{(4)} + f'' + 36 f' + 52 f = 0.$$

The characteristic polynomial is

$$\chi = \partial^4 + \partial^2 + 36\partial + 52 = (\partial - 2 - 3i)(\partial - 2 + 3i)(\partial + 2)^2$$

Thus, we have the real root -2 with multliplicity 2 and the conjugate complex roots $2 - 3i$ and $2 + 3i$. According to Remark 284, this means that a fundamental system is

$$e^{-2x}, \qquad x e^{-2x}, \qquad e^{2x} \cos 3x, \qquad \text{and} \qquad e^{2x} \sin 3x.$$

Remark 286 (Variation of Constants)*.* Consider the equation

$$f^{(n)} + a_{n-1}f^{(n-1)} + \ldots + a_1 f' + a_0 f = b$$

with coefficients $a_0, \ldots, a_{n-1}$ and right hand side b. In order to find a solution, we first rewrite the equation as a first order linear system. Let

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

This is the companion matrix of the characteristic polynomial $\chi = \partial^n + a_{n-1}\partial^{n-1} + \ldots + a_1\partial + a_0$. Then $\chi \bullet f = b$ if and only if

$$(\partial \mathbf{1} - A) \bullet \begin{pmatrix} f \\ f' \\ \vdots \\ f^{(n-1)} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix}$$

using the matrix on vector action described in Remark 96. Assume now that we have a fundamental system $z_1, \ldots, z_n$ of the corresponding homogenous equation $\chi \bullet f = 0$. Then we can define the so-called *Wronskian matrix* $Z = (z_j^{(i)})_{ij}$ of $z_1, \ldots, z_n$. It is well-known that Z is invertible. Moreover, $(\partial - A) \bullet Z = \mathbf{0}$; that is, $Z' = AZ$. Consider now the operator

$$G(Y) = Z \int Z^{-1} Y \, dx$$

where the integration is applied to every component of $Z^{-1}Y$. Let Y be any vector, then

$$\begin{aligned} (\partial - A) \bullet G(Y) = (\partial - A) \bullet Z \int Z^{-1} Y \, dx &= (Z\partial + Z' - AZ) \bullet \int Z^{-1} Y \, dx \\ &= Z\partial \bullet \int Z^{-1} Y \, dx = Z\Big(\int Z^{-1} Y \, dx\Big)' = Z Z^{-1} Y = Y \end{aligned}$$

where the second identity comes from the product rule $\partial \bullet (MN) = M\partial \bullet N + (\partial \bullet M)N = (M\partial + M') \bullet N$ for matrices and the fifth identity it the fundamental theorem of analysis $(\int g \, dx)' = g$ applied to every entry of $Z^{-1}Y$. So, as operators $(\partial - A)G = \mathrm{id}$. In particular,

$$(\partial - A) \bullet G\Bigg(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix}\Bigg) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix}.$$

That means, $G((0, \ldots, 0, b)^t)$ is a solution of the inhomogeneous system. Hence, G maps right hand sides to solutions; that is, G is the *Green's operator* for the system. For our right hand side we obtain the solution

$$\begin{pmatrix} f \\ f' \\ \dots \\ f^{(n-1)} \end{pmatrix} = Z \int Z^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix} \, dx$$

or, in other words,

$$f = \begin{pmatrix} z_1 & \cdots & z_n \end{pmatrix} \int (Z^{-1})_{*,n} b \, dx = \sum_{j=1}^{n} z_j \int (Z^{-1})_{jn} b \, dx$$

is a particular solution of the original equation.

Example 287. Consider the equation

$$f''' - 3f'' + 3f' - f = e^x.$$

Using Remark 284, we find that a fundamental system of the corresponding homogenous equation $(\partial - 1)^3 \bullet f = 0$ is e^x, xe^x, x^2e^x. Thus, as in Remark 286 we form the fundamental matrix

$$Z = \begin{pmatrix} e^x & xe^x & x^2e^x \\ e^x & (x+1)e^x & (x^2+2x)e^x \\ e^x & (x+2)e^x & (x^2+4x+2)e^x \end{pmatrix}.$$

The last column of Z^{-1} is

$$\begin{pmatrix} \frac{1}{2}x^2e^{-x} \\ -xe^{-x} \\ \frac{1}{2}e^{-x} \end{pmatrix}.$$

Thus, we obtain

$$\int Z^{-1}_{*,n} e^x \, dx = \int \begin{pmatrix} \frac{1}{2}x^2 \\ -x \\ \frac{1}{2} \end{pmatrix} \, dx = \begin{pmatrix} \frac{1}{6}x^3 \\ -\frac{1}{2}x^2 \\ \frac{1}{2}x \end{pmatrix}$$

which yields the solution

$$\frac{1}{6}x^3e^x - \frac{1}{2}x^2xe^x + \frac{1}{2}xx^2e^x = \frac{1}{6}x^3e^x.$$

Exercise 288. Let $p = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \in \mathbb{R}[x]$; and let

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

be the companion matrix of p. Show that the Smith–Jacobson normal form of $x\mathbf{1} - A \in {}^n\mathbb{R}[x]^n$ is $\operatorname{diag}(1, \ldots, 1, p)$.

Remark 289. Exercise 288 gives us another way to see that the n^{th} order equation $p \bullet f = b$ and the first order system $(\partial\mathbf{1} - A) \bullet y = be_n$ are equivalent.

References

[And88] Dan D. Anderson, *An existence theorem for non-Euclidean PID's*, Communications in Algebra **16** (1988), no. 6, 1221–1229.

[Lam99] Tsit-Yuen Lam, *Lecture on modules and rings*, Graduate Texts in Mathematics, no. 189, Springer, 1999.

Symbols