Linear algebra

Universität Wuppertal, WS 2017
Friedrich-Schiller-Universität Jena, WS 2015
Leibniz-Universität Hannover, SS 2013
Universität zu Köln, SS 2010
Universität Braunschweig, SS 2008
Friedrich-Schiller-Universität Jena, SS 2005

Jürgen Müller

Contents

I Linear algebra I

0 Introduction

(0.1) Example: Linear optimisation. A peasant has got an area of 40ha = 400.000m^2 of farm land, and a cowshed providing space to keep at most 10 cows. He is able to work 2400h per year, where to nurse a cow he needs 1ha of farm land to grow feeding grass and 200h working time, while to grow wheat on 1ha of farm land he needs 50h working time. He earns 260€ per cow, and 130€ per hectare of wheat. Now he wants to maximise his yearly profit, by choosing the appropriate number $x \in \mathbb{R}$ of cows to keep and area $y \in \mathbb{R}$ of land to grow wheat on; note that we might just weigh the cows instead of counting them.

Thus we have $[x, y] \in \mathcal{D}$, where $\mathcal{D} \subseteq \mathbb{R}^2$ is given by the following **constraints**:

$$\begin{aligned} y &\geq 0 && (1)\\ x &\geq 0 && (2)\\ x &\leq 10 && (3)\\ x+y &\leq 40 && (4)\\ 200x+50y &\leq 2400 && (5) \end{aligned}$$

Hence $\mathcal{D}$ is a **convex polygon**, see Table 1, where P is the intersection of the lines $\{[x, y] \in \mathbb{R}^2; x = 10\}$ and $\{[x, y] \in \mathbb{R}^2; 200x + 50y = 2400\}$ defined by (3) and (5), respectively, thus $P = [10, 8]$. Moreover, Q is the intersection of the lines $\{[x, y] \in \mathbb{R}^2; x + y = 40\}$ and $\{[x, y] \in \mathbb{R}^2; 200x + 50y = 2400\}$ defined by (4) and (5), respectively. Thus we have to solve a **system of linear equations**:

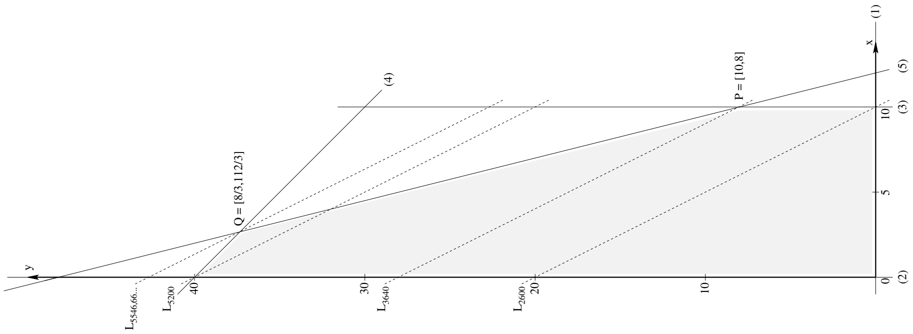
$$\left\{ \begin{array}{rcrcr} x & + & y & = & 40 \\ 200x & + & 50y & = & 2400 \end{array} \right.$$

Dividing the second equation by 50, and subtracting the first equation yields the equation $3x = 8$, hence $x = \frac{8}{3}$, and subtracting the latter from the first yields $y = 40 - \frac{8}{3} = \frac{112}{3}$, thus $Q = \frac{1}{3} \cdot [8, 112]$. Note that the system has **coefficients** in $\mathbb{Z}$, and we allow for solutions in $\mathbb{R}^2$, but the solution is in $\mathbb{Q}^2 \setminus \mathbb{Z}^2$.

We have to maximise the **cost function** $\varphi\colon \mathbb{R}^2 \to \mathbb{R}\colon [x, y] \mapsto 260x + 130y$ on $\mathcal{D}$. Since $\mathcal{D}$ is compact and φ is continuous, the maximum $c_0 := \max\{\varphi(x, y) \in \mathbb{R}; [x, y] \in \mathcal{D}\}$ is attained. To find it, we for $c \in \mathbb{R}$ consider the line $L_c := \varphi^{-1}(c) := \{[x, y] \in \mathbb{R}^2; \varphi(x, y) = c\}$. It is geometrically seen that $c_0 = \max\{c \in \mathbb{R}; L_c \cap \mathcal{D} \neq \emptyset\}$, hence c_0 is determined by the condition $Q \in L_{c_0}$ and we have $\varphi^{-1}(c_0) \cap \mathcal{D} = \{Q\}$. Thus we have $c_0 = 260 \cdot \frac{8}{3} + 130 \cdot \frac{112}{3} = \frac{16640}{3} = 5546,6\overline{6}$. In conclusion, the peasant at best earns $5546,6\overline{6}$€ per year, which happens if he keeps $\frac{8}{3} = 2,6\overline{6}$ cows and grows wheat on $\frac{112}{3}$ha $= 37,3\overline{3}$ha of farm land. Note that hence the problem of finding the maximum c_0 has been reduced to essentially solving a system of linear equations.

(0.2) Lines in $\mathbb{R}^2$. Having chosen a **coordinate system**, **geometrical** objects like lines L in the Euclidean plane $\mathbb{R}^2$ can be described **algebraically**. There

Table 1: Geometric picture of constraints and cost function.



are various ways to do so, where these descriptions are by no means unique, but are all equivalent, inasmuch it is possible to switch from either of these representations to any other one:

i) L can be given as $\{[x,y] \in \mathbb{R}^2; ax+by=c\}$, where $a,b,c \in \mathbb{R}$ such that $[a,b] \neq [0,0]$ are fixed, that is the points satisfying a certain **linear equation**.

ii) L can be given in **parametrised form** as $\{[x_0,y_0]+t\cdot[u,v] \in \mathbb{R}^2; t \in \mathbb{R}\}$, where $[x_0,y_0],[u,v] \in \mathbb{R}^2$ such that $[u,v] \neq [0,0]$ are fixed; that is $[x_0,y_0]$ is a fixed point belonging to L, and $[u,v]$ desribes the direction into which L runs.

iii) L can be given as by specifying **two points** $[x_0,y_0] \neq [x_1,y_1] \in \mathbb{R}^2$ belonging to it; note that here it becomes clear that we make use of the axiom of Euclidean geometry saying that two distinct points determine a unique line in the plane.

Here, the expression '$[x,y]+t\cdot[u,v]$' is comprised of a **scalar multiplication** and an **addition**, both performed entrywise on tuples. This is the algebraic translation of the geometric processes of dilating and negating 'point vectors', and of additing two 'point vectors', respectively.

For example, the lines in (0.1) are described as follows:

	(i)	(ii)	(iii)
(1)	$y=0$	$t\cdot[1,0]$	$[0,0],[1,0]$
(2)	$x=0$	$t\cdot[0,1]$	$[0,0],[0,1]$
(3)	$x=10$	$[10,0]+t\cdot[0,1]$	$[10,0],[10,1]$
(4)	$x+y=40$	$[20,20]+t\cdot[1,-1]$	$[40,0],[0,40]$
(5)	$4x+y=48$	$[12,0]+t\cdot[1,-4]$	$[12,0],[0,48]$
φ	$260x+130y=c$	$[0,\frac{c}{130}]+t\cdot[1,-2]$	$[\frac{c}{260},0],[0,\frac{c}{130}]$

For example, for line (5) we have the linear equation $4x+y=48$, hence for $x_0=0$ we get $y_0=48$, and for $y_1=0$ we get $x_1=12$, thus from $[x_1,y_1]-$

$[x_0, y_0] = [12, -48]$ we get the parametrisations $\{[0, 48] + t \cdot [1, -4] \in \mathbb{R}^2; t \in \mathbb{R}\}$, or equivalently $\{[12, 0] + t \cdot [1, -4] \in \mathbb{R}^2; t \in \mathbb{R}\}$. Conversely, given the latter, for any point $[x, y]$ on the line we have $x = 12 + t$ and $y = -4t$, for some $t \in \mathbb{R}$, which yields $y = -4(x - 12) = -4x + 48$, or equivalently $4x + y = 48$.

(0.3) Linear equations. Given **coefficients** $a_{ij} \in \mathbb{R}$, for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, known 'output' quantities $y_1, \ldots, y_m \in \mathbb{R}$, and unknown **indeterminate** 'input' quantities $x_1, \ldots, x_n \in \mathbb{R}$, where $m, n \in \mathbb{N}$, the associated **system of linear equations** is given as follows:

$$\left\{ \begin{array}{ccccccccccc} \sum_{j=1}^n a_{1j}x_j & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \sum_{j=1}^n a_{2j}x_j & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & y_2 \\ \vdots & & & & & & & & & & \vdots \\ \sum_{j=1}^n a_{mj}x_j & = & a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array} \right.$$

We combine the above quantities to **columns** $v := [x_1, \ldots, x_n]^{\text{tr}} \in \mathbb{R}^{n \times 1}$ and $w := [y_1, \ldots, y_m]^{\text{tr}} \in \mathbb{R}^{m \times 1}$, respectively, where by 'tr' we just indicate that the rows in question are considered as columns. Moreover, we write the coefficients as a $(m \times n)$**-matrix** with **entries** a_{ij}, that is as a **rectangular** scheme

$$A = [a_{ij}]_{ij} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

with m **rows** and n **columns**; if $m = n$ then A is called **quadratic**. Then the system can be written as $A \cdot v = w$, where the **matrix product** '$A \cdot v$' on the left hand side is just defined as the column $[\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j]^{\text{tr}} \in \mathbb{R}^{m \times 1}$. The set of **solutions** is defined as $\mathcal{L}(A, w) := \{v \in \mathbb{R}^{n \times 1}; A \cdot v = w\}$; if $\mathcal{L}(A, w) \neq \emptyset$ the system is called **solvable**. The aim is to understand when a system is solvable, and, in this case, how the set of solutions can be described.

To solve the system we consider the **extended matrix** $[A|w] \in \mathbb{R}^{m \times (n+1)}$ obtained by concatenating the columns of A with the column w. Then multiplying the i-th equation with $0 \neq a \in \mathbb{R}$, and adding the a-fold of the i-th equation to the j-th equation for some $a \in \mathbb{R}$, translates into **row operations** on $[A|w]$, namely multiplying the i-th row entrywise with $0 \neq a \in \mathbb{R}$, and adding entrywise the a-fold of the i-th row to the j-th row, respectively. Since we are replacing equations by consequences of given equations, the set of solutions might become larger, but since all steps are reversible, the set of solutions actually remains the same, as desired. The aim then is to arrive at an equivalent system, whose solutions can be read off readily. For example:

i) We reconsider the system turning up in (0.1); note that this is an algebraic translation of the geometrical problem of finding the intersection of two lines in the plane: We have $m = n = 2$ and $A \cdot [x, y]^{\text{tr}} = w$, where

$$[A|w] = \left[\begin{array}{cc|c} 1 & 1 & 40 \\ 200 & 50 & 2400 \end{array} \right] \in \mathbb{R}^{2 \times (2+1)},$$

and suitable row operations, yielding $\mathcal{L}(A,w)=\{[\frac{8}{3},\frac{112}{3}]^{\text{tr}}\}$, are:

$$[A|w]\mapsto\left[\begin{array}{cc|c}1&1&40\\4&1&48\end{array}\right]\mapsto\left[\begin{array}{cc|c}1&1&40\\3&.&8\end{array}\right]\mapsto\left[\begin{array}{cc|c}1&1&40\\1&.&\frac{8}{3}\end{array}\right]\mapsto\left[\begin{array}{cc|c}.&1&\frac{112}{3}\\1&.&\frac{8}{3}\end{array}\right]$$

ii) For the system

$$\left\{\begin{array}{rcrcrcrcr}3x_1&+&6x_2&+&2x_3&+&10x_4&=&2\\10x_1&+&16x_2&+&6x_3&+&30x_4&=&6\\5x_1&+&14x_2&+&4x_3&+&14x_4&=&10\end{array}\right.$$

we have $m=3$ and $n=4$, and $A\cdot[x_1,x_2,x_3,x_4]^{\text{tr}}=w$ where

$$[A|w]=\left[\begin{array}{cccc|c}3&6&2&10&2\\10&16&6&30&6\\5&14&4&14&10\end{array}\right]\in\mathbb{R}^{3\times(4+1)}.$$

Adding the (-3)-fold of row 1 to row 2, and the (-2)-fold of row 1 to row 3 yields

$$[A|w]\mapsto\left[\begin{array}{cccc|c}3&6&2&10&2\\1&-2&.&.&.\\5&14&4&14&10\end{array}\right]\mapsto\left[\begin{array}{cccc|c}3&6&2&10&2\\1&-2&.&.&.\\-1&2&.&-6&6\end{array}\right].$$

Next, adding the (-3)-fold of row 2 to row 1, adding row 2 to row 3, and dividing row 1 by 2, and dividing row 3 by -6 yields

$$\left[\begin{array}{cccc|c}3&6&2&10&2\\1&-2&.&.&.\\-1&2&.&-6&6\end{array}\right]\mapsto\left[\begin{array}{cccc|c}.&12&2&10&2\\1&-2&.&.&.\\.&.&.&-6&6\end{array}\right]\mapsto\left[\begin{array}{cccc|c}.&6&1&5&1\\1&-2&.&.&.\\.&.&.&1&-1\end{array}\right].$$

Finally, adding the (-5)-fold of row 3 to row 1, and interchanging rows 1 and 2 yields

$$\left[\begin{array}{cccc|c}.&6&1&5&1\\1&-2&.&.&.\\.&.&.&1&-1\end{array}\right]\mapsto\left[\begin{array}{cccc|c}1&-2&.&.&.\\.&6&1&.&6\\.&.&.&1&-1\end{array}\right].$$

Hence we infer $x_4=-1$, and we may choose $x_2=t\in\mathbb{R}$ freely, then we get $x_3=6-6t$ and $x_1=2t$, implying that $\mathcal{L}(A,w)=\{[0,0,6,-1]^{\text{tr}}+t\cdot[2,1,-6,0]^{\text{tr}}\in\mathbb{R}^{4\times1};t\in\mathbb{R}\}$, a line in $\mathbb{R}^{4\times1}$.

(0.4) Inversion. We again come back to the system of linear equations in (0.1), given by $A=\begin{bmatrix}1&1\\200&50\end{bmatrix}\in\mathbb{R}^{2\times2}$. We have seen how to solve the specific system $A\cdot[x,y]^{\text{tr}}=w$, where $w=[40,2400]^{\text{tr}}$ was fixed. It is immediate that we could similarly solve any system with the same matrix A and varying right hand side w, but at the cost of doing the computation for any new w separately. We show that it suffices to do this only once, for a generic right hand side, that is for the system $A \cdot [x, y]^{\mathrm{tr}} = [c, d]^{\mathrm{tr}}$, where $c, d \in \mathbb{R}$ are indeterminates.

To see how $\mathcal{L}(A, [c, d]^{\mathrm{tr}})$ depends on $[c, d]$, we redo the above row operations:

$$\left[\begin{array}{cc|c} 1 & 1 & c \\ 200 & 50 & d \end{array}\right] \mapsto \left[\begin{array}{cc|c} 1 & 1 & c \\ 4 & 1 & \frac{d}{50} \end{array}\right] \mapsto \left[\begin{array}{cc|c} 1 & 1 & c \\ 3 & . & -c + \frac{d}{50} \end{array}\right]$$

$$\mapsto \left[\begin{array}{cc|c} 1 & 1 & c \\ 1 & . & -\frac{c}{3} + \frac{d}{150} \end{array}\right] \mapsto \left[\begin{array}{cc|c} . & 1 & \frac{4c}{3} - \frac{d}{150} \\ 1 & . & -\frac{c}{3} + \frac{d}{150} \end{array}\right] \mapsto \left[\begin{array}{cc|c} 1 & . & -\frac{c}{3} + \frac{d}{150} \\ . & 1 & \frac{4c}{3} - \frac{d}{150} \end{array}\right]$$

This shows that $\mathcal{L}(A, [c, d]^{\mathrm{tr}})$ always is a singleton set, consisting of $[x, y]^{\mathrm{tr}} = B \cdot [c, d]^{\mathrm{tr}}$, where $B \in \mathbb{R}^{2\times 2}$ is called the **inverse matrix** of A, and equals

$$B := \begin{bmatrix} -\frac{1}{3} & \frac{1}{150} \\ \frac{4}{3} & -\frac{1}{150} \end{bmatrix} = \frac{1}{150} \cdot \begin{bmatrix} -50 & 1 \\ 200 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Hence to solve the system $A \cdot [x, y]^{\mathrm{tr}} = [c, d]^{\mathrm{tr}}$, it now suffices to plug the right hand side $[c, d]^{\mathrm{tr}}$ into the matrix product $B \cdot [c, d]^{\mathrm{tr}}$; for example, for $[c, d] = [40, 2400]$ we indeed get $B \cdot [40, 2400]^{\mathrm{tr}} = \frac{1}{150} \cdot [400, 5600]^{\mathrm{tr}} = \frac{1}{3} \cdot [8, 112]^{\mathrm{tr}}$.

Moreover, observing that the answer is given in terms of a matrix product again, and that $[c, d]^{\mathrm{tr}} = E_2 \cdot [c, d]^{\mathrm{tr}}$ where $E_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ is the associated **identity matrix**, instead of performing row operations on the matrix $[A|[c, d]^{\mathrm{tr}}] \in \mathbb{R}^{2\times 3}$, we do so on the **extended matrix** $[A|E_2] \in \mathbb{R}^{2\times 4}$, which is obtained by concatenating the columns of A with the columns of E_2:

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & . \\ 200 & 50 & . & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & 1 & 1 & . \\ 4 & 1 & . & \frac{1}{50} \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & 1 & 1 & . \\ 3 & . & -1 & \frac{1}{50} \end{array}\right]$$

$$\mapsto \left[\begin{array}{cc|cc} 1 & 1 & 1 & . \\ 1 & . & -\frac{1}{3} & \frac{1}{150} \end{array}\right] \mapsto \left[\begin{array}{cc|cc} . & 1 & \frac{4}{3} & -\frac{1}{150} \\ 1 & . & -\frac{1}{3} & \frac{1}{150} \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & . & -\frac{1}{3} & \frac{1}{150} \\ . & 1 & \frac{4}{3} & -\frac{1}{150} \end{array}\right]$$

(0.5) Example: A simple cipher. As an application of this idea, we consider the following simple cipher: The letters $\{\mathsf{A}, \ldots, \mathsf{Z}\}$ of the latin alphabet are encoded into $\{0, \ldots, 25\}$, by $\mathsf{A} \mapsto 0$, $\mathsf{B} \mapsto 1$, ..., $\mathsf{Z} \mapsto 25$. Then pairs of letters are encrypted via

$$\mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1} \colon \begin{bmatrix} a \\ b \end{bmatrix} \mapsto \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \cdot \left(\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right);$$

for example this yields $\mathsf{TEXT} \mapsto [19, 4; 23, 19] \mapsto [55, 30; 108, 64]$.

Since encryption essentially is the matrix product with $A := \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, decryption amounts to solving the system of linear equations with coefficient matrix A, for various right hand sides. Hence decryption is given as

$$\mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1} \colon \begin{bmatrix} c \\ d \end{bmatrix} \mapsto \left(B \cdot \begin{bmatrix} c \\ d \end{bmatrix} \right) - \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $B := \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ is the inverse matrix of A, determined from:

$$\left[\begin{array}{cc|c} 2 & 3 & c \\ 1 & 2 & d \end{array}\right] \mapsto \left[\begin{array}{cc|c} . & -1 & c-2d \\ 1 & 2 & d \end{array}\right] \mapsto \left[\begin{array}{cc|c} . & -1 & c-2d \\ 1 & . & 2c-3d \end{array}\right]$$

Equivalently, using the identity matrix on the right hand side, we get:

$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & . \\ 1 & 2 & . & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} . & -1 & 1 & -2 \\ 1 & 2 & . & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} . & -1 & 1 & -2 \\ 1 & . & 2 & -3 \end{array}\right]$$

For example, the cipher text

$$\begin{gathered} 89, 52, 93, 56, 27, 15, 76, 48, 89, 52, 48, 26, 52, \\ 33, 30, 17, 52, 33, 23, 14, 77, 45, 17, 11, 114, 70 \end{gathered}$$

yields the plain text

$$[21, 14; 17, 18; 8, 2; 7, 19; 21, 14; 17, 3; 4, 13; 8, 3; 4, 13; 3, 4; 18, 12; 0, 4; 17, 25],$$

which decodes into ‘VORSICHT VOR DEN IDEN DES MAERZ’.

(0.6) Example: Football pool ‘Elferwette’. To describe the outcome of a soccer match, we identify ‘home team wins’ with 1, ‘guest team wins’ with 2, and ‘draw’ with 0. Hence letting $\mathbb{Z}_3 := \{0, 1, 2\} \subseteq \mathbb{Z}$, the outcome of $n \in \mathbb{N}$ matches is an **n-tuple** in $\mathbb{Z}_3^n$. We can add, subtract, and multiply in $\mathbb{Z}_3$ by computing in $\mathbb{Z}$, and subsequently taking the **residue modulo** 3. From $2 \cdot 2 = 1 \in \mathbb{Z}_3$ we conclude that the non-zero elements of $\mathbb{Z}_3$ have multiplicative inverses. Hence $\mathbb{Z}_3$ is a **finite field**, and $\mathbb{Z}_3^n$ is a **$\mathbb{Z}_3$-vector space**.

Now the task is to bet on the outcome of $n := 11$ matches, and the more guesses are correct the higher the reward is. To launch a systematic attack we observe that the **sphere** of **radius** 2 around any element of $\mathbb{Z}_3^{11}$ has cardinality $\sum_{i=0}^{2} \binom{11}{i} \cdot (3-1)^i = 1 + 11 \cdot 2 + 55 \cdot 4 = 243 = 3^5$. If we want to cover $\mathbb{Z}_3^{11}$ completely by such sphere, then at best these are pairwise disjoint. Since $|\mathbb{Z}_3^{11}| = 3^{11}$, this leads to the following question: Is there a subset $\mathcal{C} \subseteq \mathbb{Z}_3^{11}$ of cardinality $|\mathcal{C}| = \frac{3^{11}}{3^5} = 3^6 = 729$, such that any element of $\mathbb{Z}_3^{11}$ coincides in at least 9 positions with a (necessarily unique) element of $\mathcal{C}$? Note that if $\mathcal{C}$ exists at all then any two distinct elements of $\mathcal{C}$ differ in at least 5 positions; see (5.12).

1 Basics

(1.1) Sets. a) A **set** is a collection of well-defined distinct objects forming a new entity; in Cantor’s words [1895]: *Eine* **Menge** *ist eine gedankliche Zusammenfassung von bestimmten, wohlunterschiedenen Objekten der Anschauung oder des Denkens zu einem Ganzen.* Hence we stick to **naive set theory**, but Russell’s antinomy below shows that this generality leads to a contradiction.

The objects collected are called the **elements** of the set. For any set M and any object x either $x \in M$ or $x \notin M$ holds. Moreover, any set is uniquely determined by its elements, hence contains a particular element only once, and we disregard the order of the elements. The **empty set** $\emptyset = \{\}$ is the set without elements.

For example, there are the **positive integers** $\mathbb{N} := \{1, 2, 3, \ldots\}$, the **non-negative integers** $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$, the **integers** $\mathbb{Z} := \{0, 1, -1, 2, -2, \ldots\}$, the **rational numbers** $\mathbb{Q}$, the **real numbers** $\mathbb{R}$, the **complex numbers** $\mathbb{C}$.

b) Let M and N be sets. If for all $x \in M$ we have $x \in N$ then M is called a **subset** of N, and N is called a **superset** of M; we write $M \subseteq N$. If $M \subseteq N$ and $M \neq N$ then M is called a **proper** subset of N; we write $M \subset N$. We have $M = N$ if and only if $M \subseteq N$ and $N \subseteq M$, that is we have $x \in M$ if and only if $x \in N$. In particular we have $\emptyset \subseteq M$ and $M \subseteq M$. For example, we have $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Here are a few statements concerning sets and their elements and subsets: Letting $M := \{1, \ldots, 4\}$, the statement **i)** '$1 \in M$' is true, **ii)** '$2 \subseteq M$' is false, **iii)** '$\{3\} \subseteq M$' is true, **iv)** '$\{4\} \in M$' is false, **v)** '$\{3, 4\} \subseteq M$' is true, **vi)** '$\{2, \{3, 4\}\} \subseteq M$' is false, **vii)** '$\{\} = \{\{\}\}$' is false.

c) Russell's antinomy. By the generality of naive set theory, there is the set of all sets. Hence let $\mathcal{M} := \{M \text{ set}; M \notin M\}$ be the set of all sets, which do not contain themselves as one of their elements, thus either $\mathcal{M} \in \mathcal{M}$ or $\mathcal{M} \notin \mathcal{M}$. Assume we have $\mathcal{M} \in \mathcal{M}$, then $\mathcal{M}$ does not contain $\mathcal{M}$ as one of its elements, hence $\mathcal{M} \notin \mathcal{M}$, a contradiction. Assume we have $\mathcal{M} \notin \mathcal{M}$, then $\mathcal{M}$ does contain $\mathcal{M}$ as one of its elements, hence $\mathcal{M} \in \mathcal{M}$, again a contradiction.

Hence the set of all sets cannot possibly exist. Thus we indeed have to impose restrictions on which objects we may collect to form a set. This leads to the general recipe that sets can be given by explicit enumeration or by logical description. The latter means that from a given set M a new set N, actually a subset, is formed by giving a logical formula saying which of the elements of M are elements of N and which are not; for example, we have $\{x \in \mathbb{Z}; x^2 = 1\} = \{1, -1\}$.

(1.2) Elementary constructions. a) Let M and N be sets. The sets $M \cap N := \{x; x \in M \text{ and } x \in N\} = \{x \in M; x \in N\} = \{x \in N; x \in M\}$ and $M \setminus N := \{x \in M; x \notin N\}$ are called **intersection** and **difference** of M and N, respectively. If $M \cap N = \emptyset$ then M and N are called **disjoint**. If $\mathcal{I} \neq \emptyset$ is a set, and M_i is a set for $i \in \mathcal{I}$, then their intersection is defined as $\bigcap_{i \in \mathcal{I}} M_i := \{x; x \in M_i \text{ for all } i \in \mathcal{I}\}$.

The set $M \cup N := \{x; x \in M \text{ or } x \in N\}$ is called the **union** of M and N; if additionally $M \cap N = \emptyset$ then the union of M and N is called **disjoint**, written as $M \dot\cup N$. If $\mathcal{I}$ is a set, and M_i is a set for $i \in \mathcal{I}$, then their union is defined as $\bigcup_{i \in \mathcal{I}} M_i := \{x; x \in M_i \text{ for some } i \in \mathcal{I}\}$.

Hence these constructions are translations of the logical operations and, or and non, where xor translates as $(M \cup N) \setminus (M \cap N) = \{x \in M \cup N; x \notin M \cap N\}$.

b) These operations fulfill various rules: For example, letting L be a set, we have the laws of **distributivity** $L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$ and $L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$:

To show the former, we have $x \in L \cap (M \cup N)$ if and only if $x \in L$ and ($x \in M$ or $x \in N$), that is ($x \in L$ and $x \in M$) or ($x \in L$ and $x \in N$), in other words $x \in L \cap M$ or $x \in L \cap N$, or equivalently $x \in (L \cap M) \cup (L \cap N)$. As for the latter, we have $L \subseteq (L \cup M)$ and $L \subseteq (L \cup N)$, hence $L \subseteq (L \cup M) \cap (L \cup N)$, as well as $(M \cap N) \subseteq M \subseteq (L \cup M)$ and $(M \cap N) \subseteq N \subseteq (L \cup N)$, thus $L \cup (M \cap N) \subseteq (L \cup M) \cap (L \cup N)$; conversely, for $x \in (L \cup M) \cap (L \cup N)$ we either have $x \in L \subseteq L \cup (M \cap N)$, or $x \notin L$, which implies that $x \in M$ and $x \in N$, that is $x \in M \cap N$, hence in any case we have $x \in L \cup (M \cap N)$, showing that $(L \cup M) \cap (L \cup N) \subseteq L \cup (M \cap N)$. ♯

Moreover, we have the **DeMorgan Rules** $L \setminus (M \cup N) = (L \setminus M) \cap (L \setminus N)$ and $L \setminus (M \cap N) = (L \setminus M) \cup (L \setminus N)$:

To show the former, we have $x \in L \setminus (M \cup N)$ if and only if $x \in L$ and $x \notin M \cup N$, the latter saying that ($x \notin M$ and $x \notin N$), hence equivalently we have ($x \in L$ and $x \notin M$) and ($x \in L$ and $x \notin N$), that is $x \in L \setminus M$ and $x \in L \setminus N$, in other words $x \in (L \setminus M) \cap (L \setminus N)$. As for the latter, we have $x \in L \setminus (M \cap N)$ if and only if $x \in L$ and $x \notin M \cap N$, the latter saying that ($x \notin M$ or $x \notin N$), hence equivalently we have ($x \in L$ and $x \notin M$) or ($x \in L$ and $x \notin N$), in other words $x \in (L \setminus M) \cup (L \setminus N)$. ♯

c) Let $\mathcal{P}(M) := \{L; L \subseteq M\}$ be the **power set** of M. For example, we have $\mathcal{P}(\emptyset) = \{\emptyset\}$ and $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$ and $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$.

For $x \in M$ and $y \in N$ let $[x, y] := \{x, \{x, y\}\}$ be the associated **ordered pair**; hence for $x \neq y$ we have $[x, y] \neq [y, x]$. Let $M \times N := \{[x, y]; x \in M, y \in N\}$ be the **Cartesian product** of M and N; hence $M \times N \neq \emptyset$ if and only if both $M \neq \emptyset$ and $N \neq \emptyset$. For $n \in \mathbb{N}$ and sets $M_1, \ldots, M_n$ the elements of $M_1 \times M_2 \times \cdots \times M_n := \{[x_1, \ldots, x_n]; x_i \in M_i \text{ for all } i \in \{1, \ldots, n\}\}$ are called **n-tuples**; if $M_1 = \cdots = M_n = M$ we also write $M^n := M \times M \times \cdots \times M$.

A **relation** between M and N is a subset $R \subseteq M \times N$, and $x \in M$ and $y \in N$ such that $[x, y] \in R$ are said to be **related** by R; we also write xRy. For example, if $M = N$, the **equality** relation is $\{[x, x] \in M^2; x \in M\}$, that is $x \in M$ and $y \in M$ are related if and only if $x = y$.

(1.3) Equivalence relations. A relation $R \subseteq M^2$ on a set M is called **reflexive** if $[x, x] \in R$ for all $x \in M$. Moreover, R is called **symmetric** if for all $[x, y] \in R$ we also have $[y, x] \in R$. Finally, R is called **transitive**, if for all $[x, y] \in R$ and $[y, z] \in R$ we also have $[x, z] \in R$. A reflexive, symmetric, and transitive relation is called an **equivalence relation**.

For $x \in M$ let $[x]_R := \{y \in M; [x, y] \in R\} \subseteq M$ be the associated **equivalence class**. Then since R is reflexive, we have $[x, x] \in R$ for all $x \in M$. Hence $x \in [x]_R$, so that $[x]_R \neq \emptyset$ and M is the union of the equivalence classes.

Moreover, M is the disjoint union of the distinct equivalence classes, that is these form a **partition** of M:

For $x, y \in M$ we have to show that $[x]_R \cap [y]_R \neq \emptyset$ already implies $[x]_R = [y]_R$. To do so, it suffices to show that for $z \in [x]_R$ we have $[z]_R = [x]_R$, since then for $z \in [x]_R \cap [y]_R$ we conclude that $[x]_R = [z]_R = [y]_R$: For $a \in [z]_R$ we have $[z, a] \in R$, hence from $[x, z] \in R$ and transitivity we conclude $[x, a] \in R$, thus $a \in [x]_R$, showing $[z]_R \subseteq [x]_R$. Conversely, for $b \in [x]_R$ we have $[x, b] \in R$, and since by symmetry we have $[z, x] \in R$, by transitivity we conclude $[z, b] \in R$, showing $[x]_R \subseteq [z]_R$ and thus $[z]_R = [x]_R$. ♯

For example, the equality relation $\{[x, x] \in M^2; x \in M\}$ is an equivalence relation; its equivalence classes are just the singleton sets $\{x\}$, for all $x \in M$.

(1.4) Example: Residue classes. a) Let $R := \{[x, y] \in \mathbb{Z}^2; x \mid y\}$ be the **divisibility relation**, that is we have $x \mid y$ if and only if there is $z \in \mathbb{Z}$ such that $xz = y$. Since $x \mid x$, and since $x \mid y \mid z$ implies $x \mid z$, for all $x, y, z \in \mathbb{Z}$, the relation R is reflexive and transitive, but since $x \mid y$ does not in general imply $y \mid x$, the relation R is not symmetric, hence not an equivalence relation.

b) Let $n \in \mathbb{N}$ and $R_n := \{[x, y] \in \mathbb{Z}^2; x \equiv y \pmod{n}\} = \{[x, y] \in \mathbb{Z}^2; n \mid (x-y)\}$. Since $n \mid 0 = (x-x)$, and $n \mid (-x)$ whenever $n \mid x$, for all $x \in \mathbb{Z}$, the relation R_n is reflexive and symmetric. From $n \mid (x-y)$ and $n \mid (y-z)$, for $x, y, z \in \mathbb{Z}$, we conclude $n \mid (x-y) + (y-z) = (x-z)$, hence R_n is transitive as well, and thus an equivalence relation.

The associated equivalence classes $\overline{x} = [x]_n = \{x + kn \in \mathbb{Z}; k \in \mathbb{Z}\} \subseteq \mathbb{Z}$, for $x \in \mathbb{Z}$, are called **residue classes modulo** n. Moreover, letting $\mathbb{Z}_n := \{0, \ldots, n-1\}$, taking residues modulo n shows that $[x]_n \cap \mathbb{Z}_n$ is a singleton set. Hence there are precisely n equivalence classes $\mathbb{Z}/n\mathbb{Z} := \{[0]_n, \ldots, [n-1]_n\} = \{\overline{0}, \ldots, \overline{n-1}\}$; for example, for $n = 2$ the equivalence classes are $[0]_2 = \{0, 2, -2, 4, -4, \ldots\}$ and $[1]_2 = \{1, -1, 3, -3, \ldots\}$, that is the even and odd integers, respectively.

(1.5) Maps. a) Let M and N be sets. A relation $f \subseteq M \times N$ such that for all $x \in M$ there is a unique $y \in N$ satisfying $[x, y] \in f$ is called a **map** or **function** from M to N; we write $f \colon M \to N \colon x \mapsto y$, and $y = f(x)$. The element y is called the **image** of x, while x is called a **preimage** of y, with respect to f. Hence we have $f = \{[x, f(x)] \in M \times N; x \in M\}$, showing that, by interpreting pairs $[x, y]$ as coordinates, this is reminiscent of drawing the **graph** of f.

The sets M and N are called the **source** and the **domain** of f, respectively. The set $\mathrm{im}(f) := \{y \in N; y = f(x) \text{ for some } x \in M\}$ is called the **image** of f. For a subset $N' \subseteq N$, the set $f^{-1}(N') := \{x \in M; f(x) \in N'\}$ is called the **preimage** of N', with respect to f; note that despite the notation introduced further below here 'f^{-1}' does not denote a map. For a subset $M' \subseteq M$, the **restriction** of f to M' is defined as $f|_{M'} \colon M' \to N \colon x \mapsto f(x)$.

The map $f \colon M \to N$ is called **surjective** if $\mathrm{im}(f) = N$, that is for all $y \in N$ there is some $x \in M$ such that $y = f(x)$. Moreover, f is called **injective** if for all $y \in N$ the preimage $f^{-1}(\{y\})$ has at most one element, that is for all $y \in N$ there is at most one element $x \in M$ such that $y = f(x)$, or equivalently we have $f(x) \neq f(x') \in N$ whenever $x \neq x' \in M$. Finally, f is called **bijective** if it is both surjective and injective, that is $f^{-1}(\{y\})$ is a singleton set for all $y \in N$, or equivalently for all $y \in N$ there is a unique $x \in M$ such that $y = f(x)$.

b) Let $\mathrm{Maps}(M, N) := \{f\colon M \to N\}$, where maps $f, f'\colon M \to N$, considered as subsets of $M \times N$, are called **equal** if we have $f(x) = f'(x)$ for all $x \in M$. The map $\mathrm{id}_M\colon M \to M\colon x \mapsto x$ is called the **identity map**. The **composition** of maps $f\colon M \to N$ and $g\colon L \to M$, where L is a set, is defined as $fg = f \cdot g = f \circ g\colon L \to N\colon x \mapsto f(g(x))$. In particular, we have $f \cdot \mathrm{id}_M = f$ and $\mathrm{id}_N \cdot f = f$.

Moreover, fg is surjective whenever f and g are surjective, but conversely the surjectivity of fg only implies the surjectivity of f; and fg is injective whenever f and g are injective, but conversely the injectivity of fg only implies the injectivity of g. Hence we infer that fg is bijective whenever f and g are bijective, but still conversely the bijectivity of fg only implies the surjectivity of f and the injectivity of g.

c) If $f\colon M \to N$ is bijective, the relation $f^{-1} := \{[y, x] \in N \times M; [x, y] \in f\}$ is a map as well. Thus, f is said to be **invertible**, and $f^{-1}\colon N \to M$ is called the **inverse map** of f. Hence we have $f(f^{-1}(y)) = y$ for all $y \in N$, thus $ff^{-1} = \mathrm{id}_N$, and $f^{-1}(f(x)) = x$ for all $x \in M$, thus $f^{-1}f = \mathrm{id}_M$. Moreover, f^{-1} is bijective as well, and we have $(f^{-1})^{-1} = f$.

Conversely, if the map $f\colon M \to N$ has an inverse $f^{-1}\colon N \to M$, that is we have $f^{-1}f = \mathrm{id}_M$ and $ff^{-1} = \mathrm{id}_N$, then f is bijective: From $f^{-1}f = \mathrm{id}_M$ we infer that f is injective, and from $ff^{-1} = \mathrm{id}_N$ we infer that f is surjective.

For example, the map $q = \{[x, x^2] \in \mathbb{R}^2; x \in \mathbb{R}\}$, that is $q\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2$, is neither injective nor surjective: We have $\mathrm{im}(q) = \mathbb{R}_{\geq 0} := \{y \in \mathbb{R}; y \geq 0\}$, and for all $y \geq 0$ the preimage of $\{y\}$ equals $q^{-1}(\{y\}) = \{\pm\sqrt{y}\}$, which is a singleton set if and only if $y = 0$. Restricting the source and the domain of q to $\mathbb{R}_{\geq 0}$, we get the bijective map $q' = \{[x, x^2] \in \mathbb{R}^2_{\geq 0}; x \in \mathbb{R}_{\geq 0}\}$, that is $q'\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}\colon x \mapsto x^2$, whose inverse is given as $\rho = \{[x^2, x] \in \mathbb{R}^2_{\geq 0}; x \in \mathbb{R}_{\geq 0}\} = \{[y, \sqrt{y}] \in \mathbb{R}^2_{\geq 0}; y \in \mathbb{R}_{\geq 0}\}$, that is $\rho\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}\colon y \mapsto \sqrt{y}$.

(1.6) Dedekind-Peano axioms. a) A set N fulfilling the following conditions is called a **set of positive integers**: There is an element $1 \in N$ and an injective **successor map** $\sigma\colon N \to N \setminus \{1\}\colon n \mapsto n'$, in other words any element $n \in N \setminus \{1\}$ has at most one **predecessor**, such that the **principle of induction** holds: For any subset $M \subseteq N$ such that $1 \in M$, and such that for any $n \in M$ we also have $n' = \sigma(n) \in M$, we already have $M = N$.

From considering $M := \{1\} \,\dot{\cup}\, \mathrm{im}(\sigma)$, it follows by induction that $M = N$, hence we have $\mathrm{im}(\sigma) = N \setminus \{1\}$, that is σ is bijective; in other words any $n \in N \setminus \{1\}$ has a unique predecessor, while 1 has none.

b) The set $\mathbb{N} := \{1, 2, \ldots\}$ of **positive integers** together with the successor map $\mathbb{N} \to \mathbb{N} \setminus \{1\} \colon n \mapsto n+1$ fulfills the above conditions; we take the existence of $\mathbb{N}$ for granted. The set $\mathbb{N}$ is the unique model of a set of positive integers N: There is a unique map $f \colon \mathbb{N} \to N$ fulfilling $f(1) = 1$ and $f(n+1) = f(n)'$ for all $n \in \mathbb{N}$, and f is bijective:

Let $M \subseteq \mathbb{N}$ be the set of all $n \in \mathbb{N}$ such that there is a unique $f_n \colon \{1, \ldots, n\} \to N$ fulfilling $f_n(1) = 1$ and $f_n(m+1) = f_n(m)'$ for all $m < n$; hence for $1 \neq n \in M$ we have $f_n|_{\{1,\ldots,n-1\}} = f_{n-1}$. The injectivity of the successor map of $\mathbb{N}$ implies that f_n is uniquely extendible to an admissible map $f_{n+1} \colon \{1, \ldots, n+1\} \to N$, thus by induction for $\mathbb{N}$ we have $M = \mathbb{N}$.

Moreover, for $\mathrm{im}(f) \subseteq N$ we have $1 \in \mathrm{im}(f)$, and for $f(n) \in \mathrm{im}(f)$ we have $f(n)' = f(n+1) \in \mathrm{im}(f)$, thus by induction for N we have $\mathrm{im}(f) = N$, that is f is surjective. Finally, let $M \subseteq \mathbb{N}$ be the set of all $n \in \mathbb{N}$ such that $f^{-1}(\{f(n)\}) = \{n\}$. Then from $f(\mathbb{N} \setminus \{1\}) \subseteq N \setminus \{1\}$ we get $f^{-1}(\{f(1)\}) = f^{-1}(\{1\}) = \{1\}$, and for $n \in \mathbb{N}$ we have $f^{-1}(\{f(n+1)\}) = f^{-1}(\{f(n)'\}) = f^{-1}(\{f(n)\})+1 = \{n+1\}$, thus by induction for $\mathbb{N}$ we have $M = \mathbb{N}$, that is f is injective. $\sharp$

c) The principle of induction is used as a method to prove (arithmetical) statements which are 'governed' by $\mathbb{N}$. This is at best explained by an example:

For $n \in \mathbb{N}$ we have $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$: For $n = 1$ we have $\sum_{i=1}^{1} i = 1 = \frac{1 \cdot 2}{2}$, and assuming correctness for n the statement for $n+1$ reads $\sum_{i=1}^{n+1} i = (n+1) + \sum_{i=1}^{n} i = (n+1) + \frac{n(n+1)}{2} = \frac{2(n+1)+n(n+1)}{2} = \frac{(n+1)(n+2)}{2}$. $\sharp$

Moreover, in preparation of the discussion of finiteness issues below, we consider a couple of maps on $\mathbb{N}$: Modifiying the successor function by enlarging its domain yields $\tau \colon \mathbb{N} \to \mathbb{N} \colon n \mapsto n+1$, which is injective but not surjective. Letting $\rho \colon \mathbb{N} \to \mathbb{N} \colon \begin{cases} 1 & \mapsto & 1, \\ n & \mapsto & n-1, \quad \text{for } n \geq 2, \end{cases}$ which is surjective but not injective, we get $\rho\tau = \mathrm{id}_{\mathbb{N}}$, that is ρ is a 'left-sided inverse' of τ. But since τ is not surjective, it cannot possibly be an inverse, and indeed we have $\sigma\rho \colon \mathbb{N} \to \mathbb{N} \colon \begin{cases} 1 & \mapsto & 2, \\ n & \mapsto & n, \quad \text{for } n \geq 2, \end{cases}$ hence $\sigma\rho \neq \mathrm{id}_{\mathbb{N}}$.

(1.7) Cardinality. a) Sets M and N are called to be **equicardinal** if there is a bijection $f \colon M \to N$. In particular, if there is $n \in \mathbb{N}_0$ such that there is a bijection $\{1, \ldots, n\} \to M$, then M is called **finite** of **cardinality** $|M| := n$; otherwise M is called **infinite** and we write $|M| = \infty$. Indeed, for a finite set M the cardinality is uniquely determined:

Let $M = \{1, \ldots, n\}$, for some $n \in \mathbb{N}_0$, and let $g \colon \{1, \ldots, m\} \to \{1, \ldots, n\}$ be a bijection, for some $m \in \mathbb{N}_0$ such that $m \leq n$. To show that $m = n$, we proceed by induction on $n \in \mathbb{N}_0$: The statement is trivial for $n = 0$; indeed the only set being in bijection with $\{1, \ldots, 0\} = \emptyset$ is $\emptyset$ itself. Hence let $n \geq 1$. Letting $k := g(m)$ we get a bijection $g \colon \{1, \ldots, m-1\} \to \{1, \ldots, k-1, k+1, \ldots, n\}$. Since $\{1, \ldots, n-1\} \to \{1, \ldots, k-1, k+1, \ldots, n\} \colon \begin{cases} i & \mapsto & i, & \text{for } i < k, \\ i & \mapsto & i+1, & \text{for } i \geq k, \end{cases}$ is a bijection, we may assume that $g \colon \{1, \ldots, m-1\} \to \{1, \ldots, n-1\}$ is a bijection as well, and by induction we have $n-1=m-1$. $\sharp$

b) Let M be finite, and $N \subseteq M$. Then restricting a bijection $\{1,\ldots,|M|\} \to M$ to N shows that N is finite as well such that $|N| \leq |M|$, and we have $|N|=|M|$ if and only if $N=M$. We prove a few further helpful observations:

i) Any map $f\colon M \to M$ is injective if and only if it is surjective:

There is a subset $M' \subseteq M$ such that $f|_{M'}\colon M' \to \mathrm{im}(f)$ is a bijection. Then we have $M'=M$ if and only f is injective, and from $|M'|=|\mathrm{im}(f)| \leq |M|$ we conclude that $|M'|=|M|$ if and only if f is surjective. $\sharp$

ii) We have $|\mathcal{P}(M)|=2^{|M|}$:

We may assume that $M=\{1,\ldots,n\}$, where $n:=|M| \in \mathbb{N}_0$, and proceed by induction on n: For $n=0$ we have $|\mathcal{P}(\emptyset)|=|\{\emptyset\}|=1=2^0$. Hence let $n \geq 1$ and $M':=M\setminus\{n\}$. Then $|M'|=n-1$ by induction yields $|\mathcal{P}(M')|=2^{n-1}$. We have $\mathcal{P}(M)=\{N \subseteq M; n \notin N\} \,\dot\cup\, \{N \subseteq M; n \in N\}$, where the former coincides with $\mathcal{P}(M')$, and for the latter we have the bijection $\{N \subseteq M; n \in N\} \to \mathcal{P}(M')\colon N \mapsto N\setminus\{n\}$ with inverse $\mathcal{P}(M') \to \{N \subseteq M; n \in N\}\colon N' \mapsto N' \,\dot\cup\, \{n\}$. Hence we have $|\mathcal{P}(M)|=|\mathcal{P}(M')|+|\mathcal{P}(M')|=2\cdot 2^{n-1}=2^n$. $\sharp$

iii) There are $(|M|)!$ bijective maps $M \to M$, where for $n \in \mathbb{N}$ we let $n!:=n(n-1)\cdots 1$, and $0!:=1$, being called the associated **factorial**:

We proceed by induction on $n:=|M| \in \mathbb{N}_0$, where for $n=0$ there is a unique map $\emptyset \to \emptyset$, which is bijective. Hence letting $n \in \mathbb{N}$ we may assume that $M=\{1,\ldots,n\}$, and let $M':=M \setminus \{n\}$. Given a bijective map $f\colon M \to M$, we have $f(n)=m$ for some $m \in M$, and hence $f\colon M' \to M \setminus \{m\}$ is bijective as well. Since there are n possibilities to choose m, and we have $|M'|=n-1$, by induction there are $n \cdot (n-1)! = n!$ possibilities for f. $\sharp$

(1.8) Monoids. a) A set G together with a **multiplication** $\cdot\colon G \times G \to G$ fulfilling the following conditions is called a **monoid**: There is a **neutral element** $1 \in G$ such that $1 \cdot a = a = a \cdot 1$ for all $a \in G$, and we have **associativity** $(ab)c=a(bc)$ for all $a,b,c \in G$. If additionally $ab=ba$ holds for all $a,b \in G$, then G is called **commutative** or **abelian**.

In particular, we have $G \neq \emptyset$. The neutral element is uniquely defined: If $1' \in G$ also is a neutral element, then we have $1=1\cdot 1'=1'$. For all $a_1,\ldots,a_n \in G$ the product $a_1a_2\cdots a_n \in G$ is well-defined, independently from the bracketing, and if G is commutative, then $a_1a_2\cdots a_n \in G$ is independent from the order of the factors. For $a \in G$ let $a^0:=1$ and $a^{n+1}:=a^n \cdot a$ for all $n \in \mathbb{N}_0$. Then we have $a^m a^n = a^{m+n}$ and $(a^m)^n=a^{mn}$ for all $m,n \in \mathbb{N}_0$. If $a,b \in G$ **commute**, that is $ab=ba$, then we have $(ab)^n=a^nb^n=b^na^n$ for all $n \in \mathbb{N}_0$.

A subset $H \subseteq G$ is called a **submonoid**, if $1 \in H$ and multiplication restricts to a map $\cdot\colon H \times H \to H$. Then H again is a monoid; for example $\{1\}$ and G are submonoids of G.

b) For example, the set $\mathbb{N}_0 := \mathbb{N} \,\dot\cup\, \{0\}$ of **non-negative integers** becomes a commutative monoid with respect to **addition** $+\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, with neutral element 0. For all $a, b, c \in \mathbb{N}_0$ we have $a = b$ if and only if $a + c = b + c$, and the **order relation** $\leq$ on $\mathbb{N}_0$ is defined by letting $a \leq b$ if there is $c \in \mathbb{N}_0$ such that $a + c = b$. For example, $n\mathbb{N}_0 := \{nk \in \mathbb{N}_0; k \in \mathbb{N}_0\} \subseteq \mathbb{N}_0$ is an additive submonoid, for all $n \in \mathbb{N}_0$.

Moreover, $\mathbb{N}_0$ becomes a commutative monoid with respect to **multiplication** $\cdot\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, with neutral element 1. For all $a, b \in \mathbb{N}_0$ and $c \in \mathbb{N}$ we have $a = b$ if and only if $ac = bc$. Finally, we have distributivity $a(b + c) = ab + ac$ for all $a, b, c \in \mathbb{N}_0$. For example, $\{n^k \in \mathbb{N}; k \in \mathbb{N}_0\} \subseteq \mathbb{N} \subseteq \mathbb{N}_0$ are multiplicative submonoids, for all $n \in \mathbb{N}$.

(1.9) Groups. a) Let G be a monoid. An element $a \in G$ is called **right invertible** if there is a **right inverse** $b' \in G$ such that $ab' = 1$, it is called **left invertible** if there is a **left inverse** $b'' \in G$ such that $b''a = 1$. If $a \in G$ is both right and left invertible then it is called **invertible** or a **unit**; if $b' \in G$ is a right inverse and $b'' \in G$ is a left inverse, then we have $b'' = b'' \cdot 1 = b''ab' = 1 \cdot b' = b'$, thus there is a unique **inverse** $a^{-1} := b' = b'' \in G$ such that $aa^{-1} = 1 = a^{-1}a$.

Let $G^* \subseteq G$ be the set of units; for example, as additive and multiplicative monoids we have $\mathbb{N}_0^* = \{0\}$ and $\mathbb{N}_0^* = \{1\}$, respectively. For $a \in G^*$ and $n \in \mathbb{N}$ we let $a^{-n} := (a^{-1})^n$. Then we have $a^m a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$ for all $m, n \in \mathbb{Z}$, and if $a, b \in G^*$ commute then $(ab)^n = a^n b^n = b^n a^n$ for all $n \in \mathbb{Z}$.

We have $1 \in G^*$, where $1^{-1} = 1$. For all $a, b \in G^*$ we from $ab(b^{-1}a^{-1}) = 1 = (b^{-1}a^{-1})ab$ conclude $(ab)^{-1} = b^{-1}a^{-1}$ and thus $ab \in G^*$, hence $G^* \subseteq G$ is a submonoid. Moreover, we have $(a^{-1})^{-1} = a$, thus $a^{-1} \in G^*$, hence $(G^*)^* = G^*$.

b) A monoid G such that $G = G^*$ is called a **group**, and G is called **commutative** if the underlying monoid is. For any monoid G the set G^* is a group, called the **group of units** of G.

If G is a group, then for all $a, b, c \in G$ we have $a = b$ if and only if $ac = bc$, which holds if and only if $ca = cb$: From $ac = bc$ we deduce $a = a \cdot 1 = acc^{-1} = bcc^{-1} = b \cdot 1 = b$, and from $ca = cb$ we deduce $a = 1 \cdot a = c^{-1}ca = c^{-1}cb = 1 \cdot b = b$.

A submonoid H of a group G such that for all $a \in H$ we also have $a^{-1} \in H$, is called a **subgroup**; we write $H \leq G$. Then H again is a group; for example, we have $\{1\} \leq G$ and $G \leq G$.

c) Alternatively, we may define groups more directly as follows, where imposing an appropriate invertibility condition allows for weakening other conditions:

A set G together with a **multiplication** $\cdot\colon G \times G \to G$ fulfilling the following conditions is called a **group**: We have **associativity** $(ab)c = a(bc)$ for all $a, b, c \in G$; there is a **right neutral element** $1 \in G$ such that $a \cdot 1 = a$ for all $a \in G$; and for any $a \in G$ there is a **right inverse** $a^{-1} \in G$ such that $a \cdot a^{-1} = 1$.

Indeed, we only have to show that the right neutral element is left neutral as well, and that right inverses are left inverses as well:

Firstly, $aa = a$ implies $a = 1$: We have $a = a \cdot 1 = a(aa^{-1}) = (aa)a^{-1} = aa^{-1} = 1$. Now, from $(a^{-1}a)(a^{-1}a) = a^{-1}(aa^{-1})a = (a^{-1} \cdot 1)a = a^{-1}a$ we infer $a^{-1}a = 1$, and we get $a = a \cdot 1 = a(a^{-1}a) = (aa^{-1})a = 1 \cdot a$. $\sharp$

(1.10) Symmetric groups. a) Let M be a set. Then $\mathrm{Maps}(M, M)$ becomes a monoid with respect to composition of maps, having neutral element id_M: We have $(f(gh))(x) = f(g(h(x))) = ((fg)h)(x)$ for all $x \in M$, thus $(fg)h = f(gh)$ for all $f, g, h \in \mathrm{Maps}(M, M)$.

The map $f \in \mathrm{Maps}(M, M)$ is left invertible if and only if f is injective: If $g \in \mathrm{Maps}(M, M)$ is a left inverse, then for $x, y \in M$ such that $f(x) = f(y)$ we have $x = \mathrm{id}_M(x) = g(f(x)) = g(f(y)) = \mathrm{id}_M(y) = y$, hence f is injective. If conversely f is injective, then defining $g \in \mathrm{Maps}(M, M)$ by $g(y) := x \in M$ whenever $y = f(x) \in \mathrm{im}(f)$, and $g(y) := y$ whenever $y \in M \setminus \mathrm{im}(f)$, we have $g(f(x)) = x$ for all $x \in M$, hence $gf = \mathrm{id}_M$.

The map $f \in \mathrm{Maps}(M, M)$ is right invertible if and only if f is surjective: If $g \in \mathrm{Maps}(M, M)$ is a right inverse, then for all $x \in M$ we have $x = \mathrm{id}_M(x) = f(g(x))$, hence $x \in \mathrm{im}(f)$, thus f is surjective. If conversely f is surjective, then for all $y \in M$ we by the **Axiom of Choice** may choose $x_y \in M$ such that $f(x_y) = y$. This defines a map $g\colon M \to M\colon y \mapsto x_y$, and for all $y \in M$ we have $f(g(y)) = f(x_y) = y$, hence $fg = \mathrm{id}_M$.

Hence $f \in \mathrm{Maps}(M, M)$ is invertible if and only if f is bijective, the inverse being the inverse map $f^{-1} \in \mathrm{Maps}(M, M)$. The group of units $\mathcal{S}_M := \mathrm{Maps}(M, M)^* = \{\pi\colon M \to M; \pi \text{ bijective}\}$ is called the **symmetric group** on M; it is in general non-commutative, and its elements are called **permutations**.

b) For $n \in \mathbb{N}$ we write $\mathcal{S}_n := \mathcal{S}_{\{1,\ldots,n\}}$, and we let $\mathcal{S}_0 := \mathcal{S}_\emptyset$; recall that $|\mathcal{S}_n| = n!$. For example, for $n = 1$ there the unique permutation $\mathrm{id}_{\{1\}}\colon 1 \mapsto 1$. For $n = 2$ there are two permutations $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, where we write out the pairs $[i, \pi(i)]$ in 'downward notation'. For $n = 3$ there are six permutations:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

It of course suffices to record the second rows $[\pi(1), \pi(2), \pi(3)]$ only:

$$[2,3,1],\ [3,2,1],\ [1,3,2],\ [3,1,2],\ [1,2,3],\ [2,1,3]$$

Then composition of maps is given by concatenation; for example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \underline{1} & \underline{3} & \underline{2} \\ 2 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ \underline{1} & \underline{3} & \underline{2} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} \underline{2} & \underline{1} & \underline{3} \\ 3 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ \underline{2} & \underline{1} & \underline{3} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$$

In terms of 'row notation' this reduces to list access; for the above example, by abuse of notation writing $\pi := [1,3,2]$ and $\rho := [2,1,3]$, respectively, we get:

$$[2,1,3] \cdot [1,3,2] = \rho \cdot \pi = [\rho(\pi(1)), \rho(\pi(2)), \rho(\pi(3))] = [2,3,1]$$

$$[1,3,2] \cdot [2,1,3] = \pi \cdot \rho = [\pi(\rho(1)), \pi(\rho(2)), \pi(\rho(3))] = [3,1,2]$$

Inversion is given by swapping the rows, and subsequently standard notation is achieved by sorting the rows in parallel; similarly, in 'row notation' this amounts to sorting, and keeping track of the operations performed; for example:

$$\left(\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}\right)^{-1} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$$

c) Any permutation $\pi \in \mathcal{S}_n$ can be written as a product of **disjoint cycles**: We consider the **directed graph** with **vertex** set $\{1, \ldots, n\}$ having an **edge** $i \to j$ if $\pi(i) = j$. Since π is a map, from any vertex precisely one edge emanates, and since π is bijective, at any vertex precisely one edge ends. Hence the **connected components** of this graph are **directed circles**. This also shows that the cycle decomposition of π is unique up to reordering the factors, where the order of the factors does not matter. Moreover, **fixed points**, that is cycles of **length** 1, are typically left out; and inverses are given by reading cycles backwardly.

For example, we have $\mathcal{S}_1 = \{()\}$ and $\mathcal{S}_2 = \{(), (1,2)\}$; these groups are commutative. Next, we have $\mathcal{S}_3 = \{(1,2,3), (1,3), (2,3), (1,3,2), (), (1,2)\}$, where $(1,2,3)^{-1} = (1,3,2)$ and $(1,3,2)^{-1} = (1,2,3)$, while the other elements of $\mathcal{S}_3$ are their own inverses; from $(1,2,3) \cdot (1,2) = (1,3)$ and $(2,3) \cdot (1,2,3) = (1,3)$ we deduce that $\mathcal{S}_n$ is not commutative whenever $n \geq 3$.

d) We consider the following **perfect shuffles** of a deck of an even number $n \in \mathbb{N}$ of cards: Divide the deck into its top and bottom halves of the same size, and then interleave the halves perfectly. Then the top card of either the top or the bottom half ends up at the top of the final deck, where these cases are called the **out-shuffle** and the **in-shuffle**, respectively. Recording the position of the various cards before and after the shuffling yields permutations $\omega_n, \iota_n \in \mathcal{S}_n$. For example, for $n = 8$ we get:

$$\begin{aligned} \omega_8 &= \begin{bmatrix} 1 & 5 & 2 & 6 & 3 & 7 & 4 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (2,3,5)(4,7,6) \\ \iota_8 &= \begin{bmatrix} 5 & 1 & 6 & 2 & 7 & 3 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (1,2,4,8,7,5)(3,6) \end{aligned}$$

The **Mongean shuffles** are given as follows: Start with the topmost card, and then put every other card on the top and on the bottom. Then the last card ends up at the top or the bottom, yielding permutations $\mu_n, \mu'_n \in \mathcal{S}_n$. For example, for $n = 8$ we get:

$$\begin{aligned}
\mu_8 &= \begin{bmatrix} 8 & 6 & 4 & 2 & 1 & 3 & 5 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} &= (1,5,7,8)(2,4,3,6) \\
\mu'_8 &= \begin{bmatrix} 7 & 5 & 3 & 1 & 2 & 4 & 6 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} &= (1,4,6,7)(2,5)
\end{aligned}$$

Iterating the shuffling corresponds to multiplying the associated permutations: For example, for $n = 8$ performing an in-shuffle followed by an out-shuffle yields $\omega_8\iota_8 = (2,3,5)(4,7,6) \cdot (1,2,4,8,7,5)(3,6) = (1,3,4,8,6,5)(2,7)$, while the other way around we get $\iota_8\omega_8 = (1,2,4,8,7,5)(3,6) \cdot (2,3,5)(4,7,6) = (1,2,6,8,7,3)(4,5)$. This translates back into decks of cards as follows:

$$\begin{aligned}
\omega_8\iota_8 &= (1,3,4,8,6,5)(2,7) &= \begin{bmatrix} 5 & 7 & 1 & 3 & 6 & 8 & 2 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \\
\iota_8\omega_8 &= (1,2,6,8,7,3)(4,5) &= \begin{bmatrix} 3 & 1 & 7 & 5 & 4 & 2 & 8 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix}
\end{aligned}$$

2 Rings

(2.1) Rings. a) A set R together with an addition $+\colon R \times R \to R$ and a multiplication $\cdot\colon R \times R \to R$ fulfilling the following conditions is called a **ring**: The set R is a commutative additive group with neutral element 0, and a multiplicative monoid with neutral element 1, such that **distributivity** $a(b+c) = ab+ac$ and $(b+c)a = ba+ca$ holds, for all $a, b, c \in R$.

If additionally $ab = ba$ holds, for all $a, b \in R$, then R is called **commutative**. A subset $S \subseteq R$ being an additive subgroup and a multiplicative submonoid is called a **subring**; then S is again a ring. For example, $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are commutative rings, but $\mathbb{N}_0$ is not a ring.

We have $0 \cdot a = 0 = a \cdot 0$ and $(-1) \cdot a = -a = a \cdot (-1)$, as well as $(-a)b = -(ab) = a(-b)$, for all $a, b \in R$: From $0 + 0 = 0$ we get $0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$ and hence $0 = 0 \cdot a - (0 \cdot a) = (0 \cdot a + 0 \cdot a) - (0 \cdot a) = 0 \cdot a$; for $a \cdot 0 = 0$ we argue similarly. Moreover, we have $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1+1) \cdot a = 0 \cdot a = 0$; for $a \cdot (-1) = -a$ we argue similarly. Finally, we have $-(ab) = (-1) \cdot ab = (-a)b$ and $-(ab) = ab \cdot (-1) = a(-b)$.

For example, let $R := \{0\}$ with addition $0 + 0 = 0$, and multiplication $0 \cdot 0 = 0$, and $1 := 0$. Then R is a commutative ring, called the **zero ring**. Conversely, if R fulfils $1 = 0$, then we have $a = a \cdot 1 = a \cdot 0 = 0$, for all $a \in R$, hence $R = \{0\}$.

b) The multiplicative submonoid R^* is again called the **group of units** of R; for example we have $\mathbb{Z}^* = \{\pm 1\}$. If $R \neq \{0\}$ then we have $0 \notin R^*$: Assume that $0 \in R^*$, then there is $0^{-1} \in R$ such that $1 = 0 \cdot 0^{-1} = 0$, a contradiction.

A ring $R \neq \{0\}$ such that $R^* = R \setminus \{0\}$ is called a **skew field** or **division ring**, a commutative skew field is called a **field**. A subring $S \subseteq R$ of a (skew) field

R, such that for all $0 \neq a \in S$ we have $a^{-1} \in S$, is called **sub(skew)field**; then S is again a (skew) field. For example, $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are fields, but $\mathbb{Z}$ is not a (skew) field.

c) Let $R \neq \{0\}$ be commutative. An element $0 \neq a \in R$ such that $ab = 0$ for some $0 \neq b \in R$ is called a **zero-divisor**. If there are no zero-divisors, that is for all $0 \neq a, b \in R$ we have $ab \neq 0$, then R is called an **integral domain**.

For $0 \neq a \in R$ the map $\lambda_a \colon R \to R \colon x \mapsto ax$ is injective if and only if a is not a zero-divisor: If λ_a is injective, then $ax = 0 = a \cdot 0$ implies $x = 0$, for all $x \in R$, thus a is not a zero-divisor; and if a is not a zero-divisor, then $ax = ax'$, where $x, x' \in R$, implies $a(x - x') = 0$ and thus $x = x'$, hence λ_a is injective.

Any $a \in R^*$ is not a zero-divisor: From $ax = ax'$, where $x, x' \in R$, we get $x = a^{-1}ax = a^{-1}ax' = x'$, hence λ_a is injective. In particular, any field is an integral domain; but for example $\mathbb{Z}$ is an integral domain but not a field. Moreover, if R is finite then any $0 \neq a \in R$ either is a zero-divisor or a unit: If λ_a is injective, then it is surjective as well, hence there is $b \in R$ such that $ab = 1$, hence $a \in R^*$. In particular, any finite integral domain is a field.

(2.2) Proposition. a) The relation Δ on $\mathbb{N}_0^2$ defined by $[a, b]\Delta[a', b']$ if $a + b' = b + a'$, for all $a, b, a', b' \in \mathbb{N}_0$, is an equivalence relation.

Let $\mathbb{Z} := \{[a, b]_\Delta \subseteq \mathbb{N}_0^2; a, b \in \mathbb{N}_0\}$ be the associated set of equivalence classes.

b) Then $\mathbb{Z}$ is an integral domain, called the **integers**, with addition $[a, b]_\Delta + [c, d]_\Delta := [a + c, b + d]_\Delta$ and multiplication $[a, b]_\Delta \cdot [c, d]_\Delta := [ac + bd, ad + bc]_\Delta$, for all $a, b, c, d \in \mathbb{N}_0$, additive neutral element $[0, 0]_\Delta$, the additive inverse of $[a, b]_\Delta$ being $[b, a]_\Delta$, and multiplicative neutral element $[1, 0]_\Delta$.

c) The map $\mathbb{N}_0 \to \mathbb{Z} \colon n \mapsto [n, 0]_\Delta$ is injective, via which $\mathbb{N}_0 \subseteq \mathbb{Z}$ is both an additive and multiplicative submonoid. Still writing $0 := [0, 0]_\Delta \in \mathbb{Z}$ and $1 := [1, 0]_\Delta$, we have $\mathbb{Z} = \mathbb{N} \,\dot\cup\, \{0\} \,\dot\cup\, (-\mathbb{N})$ and $\mathbb{Z}^* = \{\pm 1\}$; there is a unique extension of $\leq$ on $\mathbb{N}_0$ so that $\mathbb{Z}$ becomes an **ordered ring**.

Proof. a) We have $a + b = b + a$, thus $[a, b]\Delta[a, b]$, hence Δ is reflexive. From $[a, b]\Delta[a', b']$ we have $a + b' = b + a'$, hence $a' + b = b' + a$, thus $[a', b']\Delta[a, b]$, hence Δ is symmetric. From $[a, b]\Delta[a', b']$ and $[a', b']\Delta[a'', b'']$ we have $a + b' = b + a'$ and $a' + b'' = b' + a''$, thus $(a + b'') + (a' + b') = (a + b') + (a' + b'') = (b + a') + (b' + a'') = (b + a'') + (a' + b')$, hence $a + b'' = b + a''$, thus $[a, b]\Delta[a'', b'']$, hence Δ is transitive.

b) We have to show that addition and multiplication are independent from the choice of representatives of the equivalence classes: Let $[a, b]\Delta[a', b']$ and $[c, d]\Delta[c', d']$. Then we have $(a + c) + (b' + d') = (a + b') + (c + d') = (a' + b) + (c' + d) = (b + d) + (a' + c')$. Moreover, we have $(ac + bd + a'd' + b'c') + (a'c + b'd) = (ac + bd) + a'(d' + c) + b'(c' + d) = (ac + bd) + a'(c' + d) + b'(c + d') = (ac + bd) + (a'd + a'c' + b'c + b'd') = (a + b')c + (a' + b)d + (a'c' + b'd') = (a' + b)c + (a + b')d + (a'c' + b'd') = (bc + ad + a'c' + b'd') + (a'c + b'd)$, hence $(ac + bd) + (a'd' + b'c') = (bc + ad) + (a'c' + b'd')$.

Then $\mathbb{Z}$ is a commutative additive group with neutral element $[0,0]_\Delta$. Moreover, we have $([a,b]_\Delta[c,d]_\Delta)[e,f]_\Delta = [ac+bd, ad+bc]_\Delta[e,f]_\Delta = [ace+bde+adf+bcf, acf+bdf+ade+bce]_\Delta = [a,b]_\Delta[ce+df, cf+de]_\Delta = [a,b]_\Delta([c,d]_\Delta[e,f]_\Delta)$, hence $\mathbb{Z}$ is a commutative multiplicative monoid with neutral element $[1,0]_\Delta \neq [0,0]_\Delta$. We have distributivity $([a,b]_\Delta + [c,d]_\Delta)[e,f]_\Delta = [a+c, b+d]_\Delta[e,f]_\Delta = [ae+ce+bf+df, af+cf+be+de]_\Delta = [ae+bf, af+be]_\Delta + [ce+df, cf+de]_\Delta = ([a,b]_\Delta[e,f]_\Delta) + ([c,d]_\Delta[e,f]_\Delta)$, thus $\mathbb{Z}$ is a commutative ring.

Moreover, $\mathbb{Z}$ does not have zero-divisors: Let $[0,0]_\Delta \neq [a,b]_\Delta, [c,d]_\Delta \in \mathbb{Z}$ such that $[0,0]_\Delta = [a,b]_\Delta[c,d]_\Delta = [ac+bd, ad+bc]_\Delta$, that is we have $a \neq b$, and $c \neq d$, and $ac+bd = ad+bc$. We may assume $b > a$ and $d > c$, otherwise we consider $[b,a]_\Delta = -[a,b]_\Delta$ and $[d,c]_\Delta = -[c,d]_\Delta$, respectively. Hence there are $x, y \in \mathbb{N}$ such that $b = a+x$ and $d = c+y$. This yields $ac+(a+x)(c+y) = a(c+y)+(a+x)c$, hence $2ac+ay+cx+xy = 2ac+ay+cx$, thus $xy = 0$, a contradiction.

c) Let $a, b \in \mathbb{N}_0$. If $a > b$ then there is $c \in \mathbb{N}$ such that $a = b+c$, implying $[a,b]_\Delta = [b+c, b]_\Delta = [c,0]_\Delta \in \mathbb{N}$, if $a = b$ then $[a,b]_\Delta = [a,a]_\Delta = [0,0]_\Delta$, and if $a < b$ then there is $c \in \mathbb{N}$ such that $b = a+c$, implying $[a,b]_\Delta = [a, a+c]_\Delta = [0,c]_\Delta = -[c,0]_\Delta \in (-\mathbb{N})$.

Let $[0,0]_\Delta \neq [a,b]_\Delta, [c,d]_\Delta \in \mathbb{Z}$ such that $1 = [a,b]_\Delta[c,d]_\Delta = [ac+bd, ad+bc]_\Delta$, that is we have $a \neq b$, and $c \neq d$, and $ac+bd = ad+bc+1$. We may assume that $b > a$ and $d > c$, hence there are $x, y \in \mathbb{N}$ such that $b = a+x$ and $d = c+y$. This yields $ac+(a+x)(c+y) = a(c+y)+(a+x)c+1$, hence $2ac+ay+cx+xy = 2ac+ay+cx+1$, thus $xy = 1$. This implies $x = 1 = y$, hence $[a,b]_\Delta = [a, a+1]_\Delta = [0,1]_\Delta = -[1,0]_\Delta$ and $[c,d]_\Delta = [c, c+1]_\Delta = -[1,0]_\Delta$. ♯

(2.3) Theorem. Let R be an integral domain.
a) The relation Γ on $R \times (R \setminus \{0\})$ defined by $[a,b]\Gamma[a',b']$ if $ab' = ba'$, for all $a, a', b, b' \in R$ such that $b, b' \neq 0$, is an equivalence relation.

Let $\mathrm{Q}(R) := \{\frac{a}{b} \subseteq R \times (R \setminus \{0\}); a, b \in R, b \neq 0\}$ be the associated set of equivalence classes, where we write $\frac{a}{b} := [a,b]_\Gamma$.

b) Then $\mathrm{Q}(R)$ is a field, called the **field of fractions** of R, with addition $\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd}$ and multiplication $\frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}$, for all $a, b, c, d \in R$ such that $b, d \neq 0$, additive neutral element $\frac{0}{1}$, the additive inverse of $\frac{a}{b}$ being $\frac{-a}{b}$, and multiplicative neutral element $\frac{1}{1}$, the multiplicative inverse of $\frac{a}{b} \neq \frac{0}{1}$ being $\frac{b}{a}$.

c) The map $R \to \mathrm{Q}(R)\colon a \mapsto \frac{a}{1}$ is injective, via which $R \subseteq \mathrm{Q}(R)$ is a subring; we still write $0 := \frac{0}{1} \in \mathrm{Q}(R)$ and $1 := \frac{1}{1} \in \mathrm{Q}(R)$. Moreover, we have $R = \mathrm{Q}(R)$ if and only if R is a field.

In particular, $\mathbb{Q} := \mathrm{Q}(\mathbb{Z})$ is called the **rational numbers**; there is a unique extension of $\leq$ on $\mathbb{Z}$ so that $\mathbb{Q}$ becomes an **ordered field**.

Proof. a) We have $ab = ba$, hence Γ is reflexive. From $[a,b]\Gamma[a',b']$ we have $ab' = ba'$, hence $a'b = b'a$, thus $[a',b']\Gamma[a,b]$, hence Γ is symmetric. From $[a,b]\Gamma[a',b']$ and $[a',b']\Gamma[a'',b'']$ we have $ab' = ba'$ and $a'b'' = b'a''$, thus $(ab'' -$

$ba'')b' = ab' \cdot b'' - b \cdot b'a'' = ba' \cdot b'' - b \cdot a'b'' = 0$. Since $b' \neq 0$ and R is an integral domain, we have $ab'' - ba'' = 0$, and thus $[a,b]\Gamma[a'',b'']$, hence Γ is transitive.

b) We have to show that addition and multiplication are independent from the choice of representatives of the equivalence classes: Let $\frac{a}{b} = \frac{a'}{b'}$ and $\frac{c}{c} = \frac{c'}{d'}$. Since R is an integral domain, we have $bd \neq 0$ and $b'd' \neq 0$. From $(ad+bc)b'd' - bd(a'd'+b'c') = ab'dd' + bb'cd' - ba'dd' - bb'dc' = (ab'-ba')dd' + bb'(cd'-dc') = 0$ and $acb'd' - bda'c' = ab' \cdot cd' - ba' \cdot dc' = 0$ we get $\frac{ad+bc}{bd} = \frac{a'd'+b'c'}{b'd'}$ and $\frac{ac}{bd} = \frac{a'c'}{b'd'}$.

The set $\mathrm{Q}(R)$ is a commutative multiplicative monoid with neutral element $\frac{1}{1}$, and from $(\frac{a}{b}+\frac{c}{d})+\frac{e}{f} = \frac{ad+bc}{bd}+\frac{e}{f} = \frac{adf+bcf+bde}{bdf} = \frac{a}{b}+\frac{cf+de}{df} = \frac{a}{b}+(\frac{c}{d}+\frac{e}{f})$ we get that $\mathrm{Q}(R)$ is a commutative additive group with neutral element $\frac{0}{1} \neq \frac{1}{1}$. We have distributivity $(\frac{a}{b}+\frac{c}{d})\frac{e}{f} = \frac{ad+bc}{bd} \cdot \frac{e}{f} = \frac{ade+bce}{bdf} = \frac{adef+bcef}{bdf^2} = \frac{ae}{bf}+\frac{ce}{df} = \frac{a}{b} \cdot \frac{e}{f} + \frac{c}{d} \cdot \frac{e}{f}$, thus $\mathrm{Q}(R)$ is a commutative ring. For $\frac{a}{b} \neq \frac{0}{1}$ we have $a = a \cdot 1 \neq b \cdot 0 = 0$, thus $\frac{b}{a} \in \mathrm{Q}(R)$ fulfils $\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ba} = \frac{1}{1}$, hence $\mathrm{Q}(R)$ is a field. Finally, if R is a field then for any $b \neq 0$ we have $\frac{a}{b} = \frac{ab^{-1}}{1} \in \mathrm{Q}(R)$. ♯

(2.4) Theorem. a) Let $n \in \mathbb{N}$. Then $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \ldots, \overline{n-1}\}$ is a commutative ring, called the associated **residue class ring**, with addition $\overline{a} + \overline{b} := \overline{a+b}$ and multiplication $\overline{a} \cdot \overline{b} := \overline{ab}$, for all $a, b \in \mathbb{Z}$, with additive neutral element $\overline{0}$, the additive inverse of $\overline{a}$ being $\overline{-a}$, and multiplicative neutral element $\overline{1}$.
b) Moreover, $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if n is a **prime**. In this case $\mathbb{Z}/n\mathbb{Z}$ is a field, called the **finite prime field** of **order** n.

Proof. a) We only have to show that addition and multiplication are independent from the choice of representatives of the equivalence classes; then the rules of arithmetic are inherited via the map $\mathbb{Z} \to \mathbb{Z}_n \colon a \mapsto \overline{a}$: Let $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$, that is we have $k, l \in \mathbb{Z}$ such that $a' = a + kn$ and $b' = b + ln$. Thus we have $a' + b' = (a + kn) + (b + ln) = (a + b) + (k + l)n$ and $a'b' = (a + kn)(b + ln) = ab + (al + bk + kln)n$, hence $\overline{a' + b'} = \overline{a + b}$ and $\overline{a'b'} = \overline{ab}$.

b) Note that $\mathbb{Z}/n\mathbb{Z}$ is the zero ring if and only if $n = 1$. If n is **composite** then there are $a, b \in \{2, \ldots, n-1\}$ such that $n = ab$, hence we have $\overline{a}, \overline{b} \neq \overline{0}$, but $\overline{ab} = \overline{n} = \overline{0}$, thus $\mathbb{Z}/n\mathbb{Z}$ is not an integral domain.

Let n be a prime. Then for any $\overline{0} \neq \overline{a} \in \mathbb{Z}/n\mathbb{Z}$ the map $\lambda_{\overline{a}} \colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \colon \overline{x} \mapsto \overline{ax}$ is injective: From $\overline{ax} = \overline{ax'}$, where $\overline{x}, \overline{x'} \in \mathbb{Z}/n\mathbb{Z}$, we get $n \mid a(x - x')$, hence since n is a prime we conclude $n \mid a$ or $n \mid x - x'$, and since $\overline{a} \neq \overline{0}$ amounts to saying $n \not| \, a$ this implies $n \mid x - x'$, that is $\overline{x} = \overline{x'}$. Hence $\mathbb{Z}/n\mathbb{Z}$ is a finite integral domain, thus is a field. ♯

(2.5) Example: Fermat numbers. For $n \in \mathbb{N}_0$ let $F_n := 2^{2^n} + 1 \in \mathbb{N}$ be the n-th **Fermat number**, where $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, $F_4 = 65537$ are primes; it was conjectured [Fermat, 1640] that F_n always is a prime. But for $F_5 := 2^{2^5} + 1 = 4\,294\,967\,297 \sim 4 \cdot 10^9$ we have $F_5 = 641 \cdot 6\,700\,417$ [Euler, 1732]; all F_n for $n \in \{5, \ldots, 30\}$ are known to be composite, but it is still an open problem whether $\{F_0, \ldots, F_4\}$ are the only Fermat primes:

We have $641 = 640 + 1 = 5 \cdot 2^7 + 1 \in \mathbb{Z}$, thus $\overline{5 \cdot 2^7} = -\overline{1} \in \mathbb{Z}/641\mathbb{Z}$, and $641 = 625 + 16 = 5^4 + 2^4 \in \mathbb{Z}$, thus $\overline{2}^4 = -\overline{5}^4 \in \mathbb{Z}/641\mathbb{Z}$, hence $\overline{F_5} = \overline{2}^{32} + \overline{1} = \overline{2}^4 \overline{2}^{28} + \overline{1} = -\overline{5 \cdot 2^7}^4 + \overline{1} = -(-\overline{1})^4 + \overline{1} = -\overline{1} + \overline{1} = \overline{0} \in \mathbb{Z}/641\mathbb{Z}$. ♯

(2.6) Example: Finite prime fields. Using the bijection $\mathbb{Z}_n \to \mathbb{Z}/n\mathbb{Z} \colon a \mapsto \overline{a}$, for $n \in \mathbb{N}$, we may transport addition and multiplication from $\mathbb{Z}/n\mathbb{Z}$ to $\mathbb{Z}_n$, yielding operations given by adding respectively multiplying in $\mathbb{Z}$ first, and then taking residues modulo n. This reduces computing in $\mathbb{Z}/n\mathbb{Z}$ to computing in $\mathbb{Z}$.

We consider the cases $p = 2, 3, 5$, and by way of comparison $n = 4$. Then addition and multiplication in $\mathbb{Z}_n$ are described as given below. Note that the case $p = 2$ is reminiscent of boolean algebra, by identfiying 0 and 1 with the logical values false and true, respectively, and '+' and '·' with the logical operations xor and and, respectively:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

(2.7) Real and complex numbers. a) We take the existence of the **real numbers** $\mathbb{R}$ for granted, where $\mathbb{R}$ is the **ordered completion** of its subfield $\mathbb{Q} \subseteq \mathbb{R}$ with respect to the order relation $\leq$; in particular $\mathbb{R}$ is an ordered field.

Let the **absolute value** $|\cdot| \colon \mathbb{R} \to \mathbb{R}_{\geq 0} := \{x \in \mathbb{R}; x \geq 0\}$ be defined by $|x| := x$ for $x \in \mathbb{R}_{\geq 0}$ and $|x| := -x$ for $x \in \mathbb{R}_{<0} := \{x \in \mathbb{R}; x < 0\}$. Hence we have **definiteness**, that is $|x| = 0$ if and only if $x = 0$, **multiplicativity** $|xx'| = |x| \cdot |x'|$, and the **triangle inequality** $|x + x'| \leq |x| + |x'|$, for all $x, x' \in \mathbb{R}$.

b) Let $\mathbb{C} := \mathbb{R}^2$ be the set of **complex numbers**, which we may identify $\mathbb{C}$ with the **Gaussian plane**. We define an addition $+ \colon \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and a multiplication $\cdot \colon \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ by $[x, y] + [x', y'] := [x + x', y + y']$ and $[x, y] \cdot [x', y'] := [xx' - yy', xy' + yx']$, respectively. Then $\mathbb{C}$ becomes a commutative ring with additive neutral element $[0, 0]$, the additive inverse of $[x, y]$ being $[-x, -y]$, and multiplicative neutral element $[1, 0]$.

Moreover, the multiplicative inverse of $[x, y] \neq [0, 0]$ is given as $[\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}]$; note that $x^2 + y^2 > 0$: We have $[x, y] \cdot [\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}] = [\frac{x^2}{x^2+y^2} - \frac{-y^2}{x^2+y^2}, \frac{-xy}{x^2+y^2} + \frac{yx}{x^2+y^2}] = [1, 0]$. Hence $\mathbb{C}$ is a field. Moreover, the map $\mathbb{R} \to \mathbb{C} \colon x \mapsto [x, 0]$ is injective, and fulfills $x + x' \mapsto [x + x', 0] = [x, 0] + [x', 0]$ and $x \cdot x' \mapsto [xx', 0] = [x, 0] \cdot [x', 0]$, hence $\mathbb{R}$ can be viewed as a subfield of $\mathbb{C}$.

Let $i := [0, 1] \in \mathbb{C}$ be the **imaginary unit**, fulfilling $i^2 = [0, 1] \cdot [0, 1] = [-1, 0] = -1 \in \mathbb{C}$; thus in particular $\mathbb{C}$ cannot be ordered. Hence, using the above identification, for $z = [x, y] \in \mathbb{C}$ we have $z = [x, 0] + [0, 1] \cdot [y, 0] = x + iy \in \mathbb{C}$, where $\mathrm{Re}(z) := x \in \mathbb{R}$ and $\mathrm{Im}(z) := y \in \mathbb{R}$ are called the **real** and **imaginary part** of z, respectively. Let $\overline{z} := x - iy \in \mathbb{C}$ be the **complex conjugate** of z;

Thus for $z' = x' + iy' \in \mathbb{C}$ we have $z + z' = (x + iy) + (x' + iy') = (x + x') + i(y + y') \in \mathbb{C}$, and multiplication is given by distributivity as $zz' = (x + iy)(x' + iy') = xx' + ixy' + iyx' + i^2yy' = (xx' - yy') + i(xy' + yx') \in \mathbb{C}$. Moreover, we have $z\overline{z} = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2 \in \mathbb{C}$, and hence for $z \neq 0$ we have $z^{-1} = (x + iy)^{-1} = \frac{1}{x+iy} = \frac{x-iy}{(x+iy)(x-iy)} = \frac{x-iy}{x^2+y^2} = \frac{x}{x^2+y^2} + i \cdot \frac{-y}{x^2+y^2} \in \mathbb{C}$; for example, we have $\frac{1-i}{1+i} = \frac{(1-i)^2}{(1+i)(1-i)} = \frac{1-2i+i^2}{1-i^2} = \frac{-2i}{2} = -i$.

Let $|z| := \sqrt{x^2 + y^2} = \sqrt{z\overline{z}} \in \mathbb{R}_{\geq 0}$ be its **absolute value**, extending the absolute value on $\mathbb{R}$; thus for $z \neq 0$ we have $z^{-1} = \frac{1}{|z|^2} \cdot \overline{z} \in \mathbb{C}$. Hence we have definiteness, as well as multiplicativity: We have $zz' = (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx')$ and thus $zz' \cdot \overline{zz'} = (xx' - yy')^2 + (xy' + yx')^2 = (x^2x'^2 - 2xx'yy' + y^2y'^2) + (x^2y'^2 + 2xy'yx' + y^2x'^2) = (x^2 + y^2)(x'^2 + y'^2) = z\overline{z} \cdot z'\overline{z'}$, hence $|zz'|^2 = |z|^2|z'|^2$, thus $|zz'| = |z||z'|$; note that this also implies $\overline{zz'} = \overline{z} \cdot \overline{z'}$.

Moreover, the **triangle inequality** still holds: Note that for $a, b \in \mathbb{R}_{\geq 0}$ we have the inequality between **arithmetic** and **geometric mean**: From $(a - b)^2 \geq 0$ we get $a^2 + 2ab + b^2 \geq 4ab$, hence $\frac{a+b}{2} \geq \sqrt{ab}$. Let $z, z' \in \mathbb{C}$, where we may assume that $z, z', z + z' \neq 0$. Then we have $\frac{|x||x'|+|y||y'|}{|z||z'|} = \sqrt{\frac{|x|^2|x'|^2}{|z|^2|z'|^2}} + \sqrt{\frac{|y|^2|y'|^2}{|z|^2|z'|^2}} \leq \frac{1}{2} \cdot (\frac{|x|^2}{|z|^2} + \frac{|x'|^2}{|z'|^2} + \frac{|y|^2}{|z|^2} + \frac{|y'|^2}{|z'|^2}) = \frac{1}{2} \cdot \left(\frac{|x|^2+|y|^2}{|z|^2} + \frac{|x'|^2+|y'|^2}{|z'|^2}\right) = 1$, thus $|x||x'| + |y||y'| \leq |z||z'|$, a special case of the **Cauchy-Schwarz inequality**. This yields $|z + z'|^2 = (x + x')^2 + (y + y')^2 \leq |x||x + x'| + |x'||x + x'| + |y||y + y'| + |y'||y + y'| \leq |z||z + z'| + |z'||z + z'| = (|z| + |z'|) \cdot |z + z'|$, thus $|z + z'| \leq |z| + |z'|$, in terms of real and imaginary parts a special case of the **Minkowski inequality**.

In the Gaussian plane, complex multiplication becomes a **rotation-dilatation**, which is at best explained by an example: Letting $\zeta := \frac{1}{2} + i \cdot \frac{\sqrt{3}}{2} \in \mathbb{C}$ we get $|\zeta|^2 = (\frac{1}{2})^2 + (\frac{\sqrt{3}}{2})^2 = 1$, and hence $\zeta^{-1} = \overline{\zeta} = \frac{1}{2} - i \cdot \frac{\sqrt{3}}{2}$. Moreover, $\zeta^2 = (\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2})^2 = (\frac{1}{2} \cdot \frac{1}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{3}}{2}) + 2i \cdot \frac{1}{2} \cdot \frac{\sqrt{3}}{2} = -\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2}$, and thus $\zeta^3 = \zeta \cdot \zeta^2 = (\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2}) \cdot (-\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2}) = (-\frac{1}{2} \cdot \frac{1}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{3}}{2}) + i \cdot (\frac{1}{2} \cdot \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{1}{2}) = -1$, which yields $\zeta^4 = -\zeta = -\frac{1}{2} - i \cdot \frac{\sqrt{3}}{2}$, and from $\zeta^6 = (\zeta^3)^2 = 1$ we get $\zeta^5 = \zeta^6 \cdot \zeta^{-1} = \zeta^{-1} = \overline{\zeta} = \frac{1}{2} - i \cdot \frac{\sqrt{3}}{2}$.

3 Vector spaces

(3.1) Modules. a) Let R be a ring. An additive commutative group V, such that there is a **scalar multiplication** $\cdot\colon R \times V \to V$ fulfilling the following conditions is called a **(left)** R**-module**: We have **unitarity** $1 \cdot v = v$ and R-**linearity** $a(v+w) = av + aw$, as well as $(ab)v = a(bv)$ and $(a+b)v = av + bv$, for all $v, w \in V$ and $a, b \in R$. If R is a skew field, then V is called a **(left)** R**-vector space**, and its elements are called **vectors**.

We have $a \cdot 0 = 0 \in V$, and $0 \cdot v = 0 \in V$, and $a(-v) = (-a)v = -(av) \in V$, for all $v \in V$ and $a \in R$; note that we write both $0 \in R$ and $0 \in V$: We have $a \cdot 0 = a(0+0) = a \cdot 0 + a \cdot 0$, hence $a \cdot 0 = 0$, as well as $0 \cdot v + 0 \cdot v = (0+0)v = 0 \cdot v$, hence $0 \cdot v = 0$, and finally $a(-v) + av = a(-v+v) = a \cdot 0 = 0$ and $(-a)v + av = (-a+a)v = 0 \cdot v = 0$.

Conversely, if R is a skew field, then $av = 0 \in V$ implies $v = 0 \in V$ or $a = 0 \in R$: If $a \neq 0$, then we have $0 = a^{-1} \cdot 0 = a^{-1}(av) = (a^{-1}a)v = 1 \cdot v = v$.

For example, R itself can be considered as an R-module, where scalar multiplication is given by **left multiplication** $R \times R \to R\colon [a, x] \mapsto ax$; and we have the **zero** R-module $\{0\} \subseteq R$.

b) Let $\mathcal{I}$ be a set. Then $\operatorname{Maps}(\mathcal{I}, V)$ is an R-module with **pointwise** addition $f + g\colon \mathcal{I} \to V\colon x \mapsto f(x) + g(x)$ and scalar multiplication $af\colon \mathcal{I} \to V\colon x \mapsto af(x)$, for all $f, g \in \operatorname{Maps}(\mathcal{I}, V)$ and $a \in R$; writing f as a **sequence** $[f(i) \in V; i \in \mathcal{I}]$, we have $f + g = [f(i) + g(i) \in V; i \in \mathcal{I}]$ and $af = [af(i) \in V; i \in \mathcal{I}]$. Note that for $\mathcal{I} = \emptyset$ the set $\operatorname{Maps}(\mathcal{I}, V)$ can be identified with the zero R-module.

For example, letting $V = \mathbb{R}$, for $\mathcal{I} = \mathbb{R}$ and $\mathcal{I} = \mathbb{N}$ we get the $\mathbb{R}$-vector spaces $\operatorname{Maps}(\mathbb{R}, \mathbb{R})$ of all real-valued functions on the real numbers, and $\operatorname{Maps}(\mathbb{N}, \mathbb{R})$ of all sequences of real numbers, respectively.

If $\mathcal{I} = \{1, \ldots, n\}$ for some $n \in \mathbb{N}$, the bijection $\operatorname{Maps}(\{1, \ldots, n\}, V) \to V^n\colon f \mapsto [f(1), \ldots, f(n)]$ shows that V^n is an R-module by **componentwise** addition $[v_1, \ldots, v_n] + [w_1, \ldots, w_n] := [v_1 + w_1, \ldots, v_n + w_n]$ and scalar multiplication $a \cdot [v_1, \ldots, v_n] := [av_1, \ldots, av_n]$, for all $[v_1, \ldots, v_n], [w_1, \ldots, w_n] \in V^n$ and $a \in R$. In particular, for $V = R$ we get the **row** R-module $R^n = \{[a_1, \ldots, a_n]; a_i \in R \text{ for all } i \in \{1, \ldots, n\}\}$ of n-tuples with entries in R; for $n = 0$ we let $R^0 := \{0\}$. For example, the $\mathbb{R}$-vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ are underlying Euclidean geometry.

If $\mathcal{I} = \{1, \ldots, m\} \times \{1, \ldots, n\}$ for some $m, n \in \mathbb{N}$, we get the R-module $R^{m \times n} = \{[a_{ij}]_{ij}; a_{ij} \in R \text{ for all } i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ of $(m \times n)$**-matrices** with **entries** a_{ij}. We write $A \in R^{m \times n}$ as a **rectangular** scheme

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \in R^{m \times n}$$

with m **rows** and n **columns**; if $m = n$ then A is called **quadratic**, and if $m = 0$ or $n = 0$ we let $R^{m \times n} := \{0\}$. For $m = 1$ there is the bijection

$R^{1\times n} \to R^n \colon [a_{1j}]_j \mapsto [a_1, \ldots, a_n]$, and for $n = 1$ we get the **column** R-module $R^{n\times 1} = \{[a_1, \ldots, a_n]^{\mathrm{tr}}; a_i \in R \text{ for all } i \in \{1, \ldots, n\}\}$.

(3.2) Submodules. a) Let V be an R-module. An additive submonoid $U \subseteq V$ is called an **(left)** R**-submodule** if the scalar multiplication restricts to a map $\cdot\colon R\times U \to U$; if R is a skew field, then an R-submodule is called an R**-subspace**.

From $0 \cdot v = 0 \in V$, for all $v \in V$, we infer that it suffices to assume that $U \neq \emptyset$ and that addition restricts to a map $+\colon U \times U \to U$. Moreover, from $-v = -1 \cdot v \in V$, for all $v \in V$, we conclude that U is closed with respect to taking additive inverses, hence U again is an additive group, and thus again is an R-module; we write $U \leq_R V$ or just $U \leq V$.

For example, we have $\{0\} \leq V$ and $V \leq V$. If $S \subseteq R$ is a subring, then R becomes an S-module via restricting the left multiplication to S. For example, we have the $\mathbb{Z}$-modules $\mathbb{Z} \leq_{\mathbb{Z}} \mathbb{Q} \leq_{\mathbb{Z}} \mathbb{R} \leq_{\mathbb{Z}} \mathbb{Q} \leq_{\mathbb{Z}} \mathbb{C}$, the $\mathbb{Q}$-vector spaces $\mathbb{Q} \leq_{\mathbb{Q}} \mathbb{R} \leq_{\mathbb{Q}} \mathbb{C}$, and the $\mathbb{R}$-vector spaces $\mathbb{R} \leq_{\mathbb{R}} \mathbb{C}$.

b) The prototypical example is as follows: Given $A = [a_{ij}]_{ij} \in R^{m\times n}$, where R is commutative and $m, n \in \mathbb{N}_0$, we consider the **homogeneous system of linear equations** in the **indeterminates** $x_1, \ldots, x_n \in R$ given by $\sum_{j=1}^n a_{ij}x_j = 0$, for all $i \in \{1, \ldots, m\}$. Then the set $\mathcal{L}(A) := \{[x_1, \ldots, x_n]^{\mathrm{tr}} \in R^{n\times 1}; \sum_{j=1}^n a_{ij}x_j = 0 \text{ for all } i \in \{1, \ldots, m\}\}$ of **solutions** is an R-submodule of $R^{n\times 1}$: We have $0 \in \mathcal{L}(A)$, and for $[x_1, \ldots, x_n]^{\mathrm{tr}}, [x'_1, \ldots, x'_n]^{\mathrm{tr}} \in \mathcal{L}(A)$ we have $\sum_{j=1}^n a_{ij}(x_j + x'_j) = 0$ and $\sum_{j=1}^n a_{ij}(ax_j) = 0$, for all $a \in R$ and $i \in \{1, \ldots, m\}$.

For example, we consider the $\mathbb{R}$-vector space $V := \mathrm{Maps}(\mathbb{R}, \mathbb{R})$. Here are a few subsets, which we check for being $\mathbb{R}$-subspaces: **i)** Neither $\{f \in V; f(0) = 1\}$ nor $\{f \in V; f(1) = 1\}$ are subspaces, but both $\{f \in V; f(0) = 0\}$ and $\{f \in V; f(1) = 0\}$ are. **ii)** Neither $\{f \in V; f(x) \in \mathbb{Q} \text{ for } x \in \mathbb{R}\}$ nor $\{f \in V; f(x) \leq f(y) \text{ for } x \leq y \in \mathbb{R}\}$ are subspaces, but $\{f \in V; f(x+y) = f(x)+f(y) \text{ for } x, y \in \mathbb{R}\}$ is. **iii)** The set $\{f \in V; |f(x)| \leq c\}$, where $c \in \mathbb{R}$, is a subspace if and only if $c = 0$, but all of $\{f \in V; f \text{ bounded}\}$ and $\{f \in V; f \text{ continuous}\}$ and $\{f \in V; f \text{ differentiable}\}$ and $\{f \in V; f \text{ smooth}\}$ and $\{f \in V; f \text{ integrable}\}$ are.

(3.3) Example: The ISBN [1968, 2007]. To detect typing errors **parity check systems** are used, where typical errors and their frequencies are given in Table 2; an example of a phonetic error is replacing 'thirty' by 'thirteen'. The **International Standard Book Number** is used to identify books, where up to 2006 the standard was **ISBN-10**, which from 2007 on has been replaced by **ISBN-13**. The ISBN-10 is formed as follows:

The **alphabet** is the field $\mathbb{Z}_{11}$, where $10 \in \mathbb{Z}_{11}$ is replaced by the roman **letter** X, and **words** $[a_1; a_2, \ldots, a_6; a_7, \ldots, a_9; a_{10}] \in \mathbb{Z}_{10}^9 \times \mathbb{Z}_{11} \subseteq \mathbb{Z}_{11}^{10}$ have **length** 10, where X might possibly occur only as a last letter. Here $a_1, \ldots, a_9$ are **information symbols**, where a_1 is the group code, $a_1 \in \{0, 1\}$ referring to English, $a_1 = 2$ referring to French, and $a_1 = 3$ referring to German, $[a_2, \ldots, a_6]$

Table 2: Typing errors.

error		frequency
single	$a \longrightarrow b$	79.0%
adjacent transposition	$ab \longrightarrow ba$	10.2%
jump transposition	$abc \longrightarrow cba$	0.8%
twin	$aa \longrightarrow bb$	0.6%
jump twin	$aca \longrightarrow bcb$	0.3%
phonetic	$a0 \longleftrightarrow 1a$	0.5%
random		8.6%

is the publisher code, $[a_7, \ldots, a_9]$ is the title code, and a_{10} is a **check symbol** such that $a_{10} = \sum_{i=1}^{9} ia_i \in \mathbb{Z}_{11}$.

For example, a valid ISBN-10 is '1-58488-508-4': We have $1 \cdot 1 + 2 \cdot 5 + 3 \cdot 8 + 4 \cdot 4 + 5 \cdot 8 + 6 \cdot 8 + 7 \cdot 5 + 8 \cdot 0 + 9 \cdot 8 = 246 = 4 \in \mathbb{Z}_{11}$. The corresponding ISBN-13 is '978-1-58488-508-5'. The ISBN-13 is formed from an ISBN-10 as follows: After a 3-letter prefix, being a country code '978' or '979' referring to 'bookland', the first 9 letters of the ISBN-10 are taken, and then a check symbol is added fulfilling the standard of the **International Article Number (EAN)** [1977], formerly **European Article Number**.

Hence a valid ISBN-10 is an element of the $\mathbb{Z}_{11}$-subspace $\mathcal{C} := \{[a_1, \ldots, a_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} ia_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$. Then **single** and **transposition errors** can be **detected**, although not **corrected**: Let $v = [a_1, \ldots, a_{10}] \in \mathcal{C}$.

For $i \in \{1, \ldots, 10\}$ let $a_i \neq a_i' \in \mathbb{Z}_{11}$ and $v' := [a_1, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_{10}] \in \mathbb{Z}_{11}^{10}$. Assume that $v' \in \mathcal{C}$, then we have $[0, \ldots, 0, a_i - a_i', 0, \ldots, 0] = v - v' \in \mathcal{C}$, implying that $i(a_i - a_i') = 0 \in \mathbb{Z}_{11}$, which since $0 \neq i \in \mathbb{Z}_{11}$ and $a_i \neq a_i' \in \mathbb{Z}_{11}$ is a contradiction. Thus v' is not a valid ISBN-10.

For some $i, j \in \{1, \ldots, 10\}$ such that $i < j$ and $a_i \neq a_j \in \mathbb{Z}_{11}$ let $v'' := [a_1, \ldots, a_{i-1}, a_j, a_{i+1}, \ldots, a_{j-1}, a_i, a_{j+1}, \ldots, a_{10}] \in \mathbb{Z}_{11}^{10}$. Assume that $v'' \in \mathcal{C}$, then we have $[0, \ldots, 0, a_i - a_j, 0, \ldots, 0, a_j - a_i, 0, \ldots, 0] = v - v'' \in \mathcal{C}$, implying that $(i - j)(a_i - a_j) = 0 \in \mathbb{Z}_{11}$, which since $0 \neq i - j \in \mathbb{Z}_{11}$ and $a_i \neq a_j \in \mathbb{Z}_{11}$ is a contradiction. Thus v'' is not a valid ISBN-10. ♯

(3.4) Linear combinations. a) Let V be an R-module. If $\mathcal{I}$ is a set and $U_i \leq V$ for all $i \in \mathcal{I}$, then $\bigcap_{i \in \mathcal{I}} U_i \leq V$ is an R-submodule of V again, the empty intersection being defined as V.

If $U, U' \leq V$ then $U \cup U' \leq V$ if and only if $U \leq U'$ or $U' \leq U$ holds: For $U \leq U'$ w have $U \cup U' = U' \leq V$, while for $U' \leq U$ we have $U \cup U' = U \leq V$. Assume we have $U \cup U' \leq V$, while $U \not\leq U'$ and $U' \not\leq U$. Thus there are $v \in U \setminus U'$ and $w \in U' \setminus U$, and since $v, w \in U \cup U'$ we also have $v + w \in U \cup U'$. We may assume that $v + w \in U$, and thus $w = (v + w) - v \in U$, a contradiction.

b) For any subset $S \subseteq V$ let $\langle S \rangle_R := \{\sum_{i=1}^k a_i v_i \in V; k \in \mathbb{N}_0, a_i \in R, v_i \in S \text{ for all } i \in \{1, \ldots, k\}\}$. The finite sums $\sum_{i=1}^k a_i v_i \in V$ are called ***R*-linear combinations** of S; for $k = 0$ the empty sum is defined as $0 \in V$. Whenever $S = \{v_1, \ldots, v_n\}$ is finite we also write $\langle S \rangle_R = \langle v_1, \ldots, v_n \rangle_R = \{\sum_{i=1}^n a_i v_i \in V; a_i \in R \text{ for all } i \in \{1, \ldots, n\}\}$; in particular, for $S = \emptyset$ and $S = \{v\}$ we have $\langle \emptyset \rangle_R = \{0\}$ and $\langle v \rangle_R = \{av \in V; a \in R\}$, respectively.

Then we have $S \subseteq \langle S \rangle_R \leq V$, and $\langle S \rangle_R$ is called the R-submodule of V **generated** or **spanned** by S; note that $\langle \emptyset \rangle_R = \{0\} \leq V$. If $\langle S \rangle_R = V$, then $S \subseteq V$ is called an ***R*-generating set**. If V has a finite R-generating set then V is called a **finitely generated** R-module.

We have $\langle S \rangle_R = \bigcap\{U \leq V; S \subseteq U\}$, that is $\langle S \rangle_R$ is the smallest R-submodule of V containing S, with respect to the partial order on all R-submodules of V given by inclusion: Since $S \subseteq \langle S \rangle_R \leq V$ we have $\bigcap\{U \leq V; S \subseteq U\} \leq \langle S \rangle_R$. Conversely, since for all $v, w \in S \subseteq U$ and $a \in R$ we have $v+w \in U$ and $av \in U$, we also have $\langle S \rangle_R \leq U$, and hence $\langle S \rangle_R \leq \bigcap\{U \leq V; S \subseteq U\}$.

In particular, we have $S = \langle S \rangle_R$ if and only if $S \leq V$, and hence $\langle \langle S \rangle_R \rangle_R = \langle S \rangle_R$ for all $S \subseteq V$. For example, let $V = R^n$, where $n \in \mathbb{N}$, and for $i \in \{1, \ldots, n\}$ let $e_i := [0, \ldots, 0, 1, 0, \ldots, 0] \in R^n$ be the **unit tuple** with i-th entry non-zero. Then we have $[a_1, \ldots, a_n] = \sum_{i=1}^n a_i e_i \in R^n$, and thus $\langle e_1, \ldots, e_n \rangle_R = R^n$.

c) The problem, that for $U, U' \leq V$ the union $U \cup U' \subseteq V$ in general is not an R-submodule again, can now be remedied by going over to the R-submodule ov V generated by $U \cup U'$: More generally, if $\mathcal{I}$ is a set and $U_i \leq V$ for all $i \in \mathcal{I}$, then $\sum_{i \in \mathcal{I}} U_i := \langle \bigcup_{i \in \mathcal{I}} U_i \rangle_R = \{\sum_{j \in \mathcal{J}} v_j \in V; v_j \in U_j, \mathcal{J} \subseteq \mathcal{I} \text{ finite}\}$ is called the **sum** of the U_i; if $\mathcal{I} = \{1, \ldots, n\}$, for some $n \in \mathbb{N}$, we also write $\sum_{i \in \mathcal{I}} U_i = \sum_{i=1}^n U_i = U_1 + \cdots + U_n$, where for $n = 0$ the empty sum is defined to be $\{0\} \leq V$. Thus for subsets $S_i \subseteq V$, where $i \in \mathcal{I}$, we have $\langle \bigcup_{i \in \mathcal{I}} S_i \rangle_R = \sum_{i \in \mathcal{I}} \langle S_i \rangle_R$; in particular, a finite sum of finitely generated R-submodules is finitely generated again.

d) As far as distributivity is concerned, for $U_1, U_2, U_3 \leq V$ we always have $(U_1 \cap U_3) + (U_2 \cap U_3) \subseteq (U_1 + U_2) \cap U_3$ and $(U_1 \cap U_2) + U_3 \subseteq (U_1 + U_3) \cap (U_2 + U_3)$, but in general we do not have equality:

For example, let $V := R^2$, where $R \neq \{0\}$, and $U_1 := \{[a, 0] \in V; a \in R\} \leq V$ and $U_2 := \{[0, b] \in V; b \in R\} \leq V$ and $U_3 := \{[c, c] \in V; c \in R\} \leq V$. Then we have $U_1 + U_2 = U_1 + U_3 = U_2 + U_3 = V$, but $(U_1 \cap U_3) + (U_2 \cap U_3) = \{0\} + \{0\} = \{0\}$ and $(U_1 + U_2) \cap U_3 = V \cap U_3 = U_3 \neq \{0\}$, as well as $(U_1 \cap U_2) + U_3 = \{0\} + U_3 = U_3 \neq V$ and $(U_1 + U_3) \cap (U_2 + U_3) = V \cap V = V$.

If additionally $U_1 \leq U_3$, then distributivity is replaced by the **Dedekind identity** $(U_1 + U_2) \cap U_3 = (U_1 \cap U_3) + (U_2 \cap U_3) = U_1 + (U_2 \cap U_3)$: Let $v = u_1 + u_2 \in (U_1 + U_2) \cap U_3$, where $u_i \in U_i$. Since $u_1 \in U_1 \leq U_3$ we have $u_2 = v - u_1 \in U_2 \cap U_3$, hence $v = u_1 + u_2 \in U_1 + (U_2 \cap U_3)$.

(3.5) Linear independence. Let $R \neq \{0\}$ and let V be an R-module. Then a sequence $\mathcal{S} := [v_i \in V; i \in \mathcal{I}]$, where $\mathcal{I}$ is a set, is called ***R*-linearly independent**, if for all finite subsets $\mathcal{J} \subseteq \mathcal{I}$ and for all $[a_j \in R; j \in \mathcal{J}]$ we have $\sum_{j \in \mathcal{J}} a_j v_j = 0$ if and only if $a_j = 0$ for all $j \in \mathcal{J}$; otherwise $\mathcal{S}$ is called ***R*-linearly dependent**. A subset $S \subseteq V$ is called ***R*-linearly independent**, if $[f(i) \in V; i \in \mathcal{I}]$ is R-linearly independent for some bijection $f \colon \mathcal{I} \rightarrow S$.

If $\mathcal{S}$ is R-linearly independent then $[v_j \in V; j \in \mathcal{J}]$ is as well, for any $\mathcal{J} \subseteq \mathcal{I}$. If $\mathcal{I} = \emptyset$ then $\mathcal{S} = []$ is R-linearly independent. If $v_i = 0$ for some $i \in \mathcal{I}$, or if $v_i = v_j$ for some $i, j \in \mathcal{I}$ such that $i \neq j$, then $\mathcal{S}$ is R-linearly dependent. If R is a skew field, then since $av = 0$ implies $a = 0$ or $v = 0$, for $v \in V$ and $a \in R$, we infer that the singleton set $S = \{v\}$, where $v \neq 0$, is R-linearly independent. For example, $\{e_1, \ldots, e_n\} \subseteq R^n$, where $n \in \mathbb{N}$, is R-linearly independent: From $\sum_{i=1}^{n} a_i e_i = [a_1, \ldots, a_n] = 0 \in R^n$ we conclude $a_i = 0 \in R$ for all $i \in \{1, \ldots, n\}$.

b) For example, the sequence $[f_k \in C^\infty(\mathbb{R}); k \in \mathbb{N}]$, where $f_k \colon \mathbb{R} \rightarrow \mathbb{R} \colon x \mapsto \exp(kx)$ and $C^\infty(\mathbb{R}) := \{f \colon \mathbb{R} \rightarrow \mathbb{R}; f \text{ smooth}\} \leq \operatorname{Maps}(\mathbb{R}, \mathbb{R})$, is $\mathbb{R}$-linearly independent, see also (8.2):

We proceed by induction on $n \in \mathbb{N}_0$: The case $n = 0$ being trivial, for $n \in \mathbb{N}$ let $a_1, \ldots, a_n \in \mathbb{R}$ such that $\sum_{k=1}^{n} a_k f_k = 0$. Thus from $\sum_{k=1}^{n} a_k \exp(kx) = 0$, for all $x \in \mathbb{R}$, differentiation $\frac{\partial}{\partial x}$ yields $\sum_{k=1}^{n} k a_k \exp(kx) = 0$, hence we get $0 = n \cdot \sum_{k=1}^{n} a_k \exp(kx) - \sum_{k=1}^{n} k a_k \exp(kx) = \sum_{k=1}^{n} (n-k) a_k \exp(kx) = \sum_{k=1}^{n-1} (n-k) a_k \exp(kx)$. Hence we conclude $(n-k)a_k = 0$, implying $a_k = 0$, for all $k \in \{1, \ldots, n-1\}$. This yields $a_n \exp(nx) = 0$ for all $x \in \mathbb{R}$, hence $a_n = 0$. ♯

(3.6) Bases. a) Let $R \neq \{0\}$ and let V be an R-module. An R-linearly independent R-generating set $B \subseteq V$ is called an ***R*-basis** of V.

Given an R-basis $B = [v_i \in V; i \in \mathcal{I}]$ of V, where $\mathcal{I}$ is a set, we have the principle of **comparison of coefficients**: Since B is an R-generating set of V, any $v \in V$ is an R-linear combination of B. Let $\mathcal{J}', \mathcal{J}'' \subseteq \mathcal{I}$ be finite such that $v = \sum_{j \in \mathcal{J}'} a_j' v_j = \sum_{j \in \mathcal{J}''} a_j'' v_j$, where $0 \neq a_j' \in R$ for all $j \in \mathcal{J}'$ and $0 \neq a_j'' \in R$ for all $j \in \mathcal{J}''$. Then we have $\mathcal{J}' = \mathcal{J}''$ and $a_j' = a_j''$ for all $j \in \mathcal{J}' = \mathcal{J}''$: Letting $a_j' := 0$ for $j \in \mathcal{J}'' \setminus \mathcal{J}'$ and $a_j'' := 0$ for $j \in \mathcal{J}' \setminus \mathcal{J}''$, we have $0 = v - v = \sum_{j \in \mathcal{J}' \cup \mathcal{J}''} (a_j' - a_j'') v_j$, and $[v_j \in V; j \in \mathcal{J}' \cup \mathcal{J}'']$ being R-linearly dependent yields $a_j' = a_j''$ for all $j \in \mathcal{J}' \cup \mathcal{J}''$.

Thus, given $v \in V$, any representation $v = \sum_{j \in \mathcal{J}} a_j v_j$, where $\mathcal{J} \subseteq \mathcal{I}$ is finite and the $[a_j \in R; j \in \mathcal{J}]$ are called the associaed **coefficients** or **coordinates**, is **essentially unique**, that is $\mathcal{K} := \{j \in \mathcal{J}; a_j \neq 0\} \subseteq \mathcal{I}$ and $[a_j \in R \setminus \{0\}; j \in \mathcal{K}]$ are independent from the choice of $\mathcal{J}$.

In particular, if $B = [v_1, \ldots, v_n]$ is finite, where $n \in \mathbb{N}_0$, then taking $\mathcal{J} = \mathcal{J}' = \mathcal{J}'' = \{1, \ldots, n\}$ in the above computation shows that any $v \in V$ has a unique representation $v = \sum_{i=1}^{n} a_i v_i$, where $a_1, \ldots, a_n \in R$, giving rise to the **coordinate tuples** $v_B := [a_1, \ldots, a_n] \in R^n$ and ${}_B v := [a_1, \ldots, a_n]^{\text{tr}} \in R^{n \times 1}$.

b) Let $\mathcal{I}$ be a set, and let $U_i \leq V$ for all $i \in \mathcal{I}$. Then V is called the **direct**

sum of the U_i, if $V = \sum_{i\in\mathcal{I}} U_i$ and any sequence $[v_i \in U_i \setminus \{0\}; i \in \mathcal{I}, U_i \neq \{0\}]$ is R-linearly independent; we write $V = \bigoplus_{i\in\mathcal{I}} U_i$. Thus any $v \in V$ can be essentially uniquely written as an R-linear combination $v = \sum_{j\in\mathcal{J}} a_j v_j$, where $\mathcal{J} \subseteq \mathcal{I}$ is finite, and $v_j \in U_j$ for all $j \in \mathcal{J}$. In particular, if $[v_i \in V; i \in \mathcal{I}]$ is an R-basis of V, then $V = \bigoplus_{i\in\mathcal{I}} \langle v_i \rangle_R$.

In particular, given $U \leq V$ then $U' \leq V$ is called a **complement** of U in V if $V = U \oplus U'$. Note that complements are not necessarily unique: For example, if $V := R^2$ and $U := \{[x, 0] \in R^2; x \in R\} \leq R^2$, then any $U_a := \{[ay, y] \in R^2; y \in R\} \leq R^2$, where $a \in R$, is a complement of U in V.

c) For example, considering the maps $f_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \exp(kx)$ again, since all non-vanishing $\mathbb{R}$-linear combinations of $\{f_k \in C^\infty(\mathbb{R}); k \in \mathbb{N}\}$ are unbounded, this is not an $\mathbb{R}$-basis of $C^\infty(\mathbb{R})$.

For example, $S := \{e_1, \ldots, e_n\} \subseteq R^n$, for $n \in \mathbb{N}$, is called the **standard** R-basis of R^n, while $\emptyset$ is the only R-basis of $R^0 = \{0\}$; note that $[a_1, \ldots, a_n] = \sum_{i=1}^n a_i e_i \in R^n$ shows that any element of R^n coincides with its coordinate tuple with respect to the standard basis, that is we have $[a_1, \ldots, a_n]_S = [a_1, \ldots, a_n]$.

For example, we consider $V = \mathbb{R}^{2\times 1}$. Then $B := \{[1, 1]^{\text{tr}}, [1, -1]^{\text{tr}}\} \subseteq \mathbb{R}^{2\times 1}$ is an $\mathbb{R}$-basis: Letting $A := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ we get $\mathcal{L}(A, [x, y]^{\text{tr}}) = \{[\frac{x+y}{2}, \frac{x-y}{2}]^{\text{tr}}\}$. This shows $[x, y]^{\text{tr}} = \frac{x+y}{2} \cdot [1, 1]^{\text{tr}} + \frac{x-y}{2} \cdot [1, -1]^{\text{tr}} \in \langle B \rangle_{\mathbb{R}}$, for all $[x, y]^{\text{tr}} \in \mathbb{R}^{2\times 1}$, hence B is an $\mathbb{R}$-generating set. For the associated homogeneous system we get $\mathcal{L}(A) = \{0\}$, hence B is $\mathbb{R}$-linearly independent. Thus the **coordinate vector** of $[x, y]^{\text{tr}} \in \mathbb{R}^{2\times 1}$ with respect to B is given as ${}_B([x, y]^{\text{tr}}) = [\frac{x+y}{2}, \frac{x-y}{2}]^{\text{tr}} \in \mathbb{R}^{2\times 1}$.

(3.7) Theorem: Characterisation of bases. Let K be a skew field and let V be a K-vector space. Then for any subset $B \subseteq V$ the following are equivalent:
i) B is a K-basis of V.
ii) B is a maximal K-linearly independent subset, that is B is K-linearly independent and any $B \subset S \subseteq V$ is K-linearly dependent.
iii) B is a minimal K-generating set, that is B is a K-generating set and for any $T \subset B$ we have $\langle T \rangle_K < V$.

Proof. i)⇒ii): Let $v \in S \backslash B$. Then there are $v_1, \ldots, v_n \in B$ and $a_1, \ldots, a_n \in K$ such that $v = \sum_{i=1}^n a_i v_i$, for some $n \in \mathbb{N}_0$. Thus S is K-linearly dependent.

ii)⇒iii): We show that B is a K-generating set: Let $v \in V$, and since $B \subseteq \langle B \rangle_K$ we may assume that $v \notin B$. Then $B \,\dot{\cup}\, \{v\} \subseteq V$ is K-linearly dependent, thus there are $v_1, \ldots, v_n \in B$ and $a, a_1, \ldots, a_n \in K$, for some $n \in \mathbb{N}_0$, such that $av + \sum_{i=1}^n a_i v_i = 0$, where $a \neq 0$ or $[a_1, \ldots, a_n] \neq 0$. Assume that $a = 0$, then $[a_1, \ldots, a_n] \neq 0$ implies that B is K-linearly dependent, a contradiction. Thus we have $a \neq 0$, and hence $v = a^{-1}(av) = -a^{-1} \cdot \sum_{i=1}^n a_i v_i \in \langle B \rangle_K$.

To show minimality, let $T \subset B$, and assume that $\langle T \rangle_K = V$. Then for $v \in B \setminus T$ there are $v_1, \ldots, v_n \in T$ and $a_1, \ldots, a_n \in K$, for some $n \in \mathbb{N}_0$, such that $v = \sum_{i=1}^n a_i v_i$, hence B is K-linearly dependent, a contradiction.

iii)⇒i): We show that B is K-linearly independent: Assume that there are $v_1, \ldots, v_n \in B$ and $a_1, \ldots, a_n \in K$, for some $n \in \mathbb{N}$, such that $[a_1, \ldots, a_n] \neq 0$ and $\sum_{i=1}^n a_i v_i = 0$. We may assume that $a_1 \neq 0$, hence $v_1 = -a_1^{-1} \cdot \sum_{i=2}^n a_i v_i$. Thus for any $0 \neq v \in V$ there are $v_{n+1}, \ldots, v_m \in B \setminus \{v_1, \ldots, v_n\}$, for some $m \geq n$, and $b_1, \ldots, b_m \in K$ such that $v = \sum_{i=1}^m b_i v_i = -a_1^{-1} b_1 \cdot \sum_{i=2}^n a_i v_i + \sum_{i=2}^m b_i v_i$, thus $\langle B \setminus \{v_1\} \rangle_K = V$, a contradiction. ♯

(3.8) Theorem: Steinitz's Base Change Theorem. Let K be a skew field, let V be a K-vector space having a finite K-basis B. If $S \subseteq V$ is K-linearly independent, then we have $|S| \leq |B|$, and there is $T \subseteq B$ such that $|S| + |T| = |B|$ and $S \,\dot{\cup}\, T$ is a K-basis of V.

Proof. Let $B = \{v_1, \ldots, v_n\}$ for some $n \in \mathbb{N}_0$, and we may assume that S is finite, hence $S = \{w_1, \ldots, w_m\}$ for some $m \in \mathbb{N}_0$. We proceed by induction on $m \in \mathbb{N}$, where the assertion is trivial for $m = 0$, hence let $m \geq 1$.

The set $\{w_1, \ldots, w_{m-1}\}$ is K-linearly independent as well, thus we may assume that $B' := \{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$ is a K-basis of V. Hence there are $a_1, \ldots, a_n \in K$ such that $w_m = \sum_{i=1}^{m-1} a_i w_i + \sum_{i=m}^n a_i v_i$. Assume that $[a_m, \ldots, a_n] = 0$, then $w_m = \sum_{i=1}^{m-1} a_i w_i$, hence $\{w_1, \ldots, w_m\}$ is K-linearly dependent, a contradiction. Hence we may assume that $a_m \neq 0$. Then $B'' := (B' \setminus \{v_m\}) \,\dot{\cup}\, \{w_m\} = \{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ is a K-basis of V:

We have $v_m = a_m^{-1}(w_m - \sum_{i=1}^{m-1} a_i w_i - \sum_{i=m+1}^n a_i v_i) \in \langle B'' \rangle_K$, and hence from $B' \setminus \{v_m\} \subseteq B''$ we conclude $\langle B'' \rangle_K = V$, that is B'' is a K-generating set. Let $b_1, \ldots, b_n \in K$ such that $\sum_{i=1}^m b_i w_i + \sum_{i=m+1}^n b_i v_i = 0$, hence we have $\sum_{i=1}^{m-1} (b_m a_i + b_i) w_i + b_m a_m v_m + \sum_{i=m+1}^n (b_m a_i + b_i) v_i = 0$. Since B' is K-linearly independent, we get $b_m a_m = 0$, thus $b_m = 0$, which implies $0 = b_m a_i + b_i = b_i$ for all $m \neq i \in \{1, \ldots, n\}$, showing that B'' is K-linearly independent. ♯

For example, let $V = \mathbb{R}^{2 \times 1}$ with standard $\mathbb{R}$-basis $\{e_1, e_2\}$. Then $\{[1, 1]^{\text{tr}}\}$ is $\mathbb{R}$-linearly independent, and letting $A_1 := \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ and $A_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ we get $\mathcal{L}(A_1) = \{0\} = \mathcal{L}(A_2)$, implying that both $\{[1, 1]^{\text{tr}}, [1, 0]^{\text{tr}}\} \subseteq V$ and $\{[1, 1]^{\text{tr}}, [0, 1]^{\text{tr}}\} \subseteq V$ are $\mathbb{R}$-linearly independent, and thus are $\mathbb{R}$-bases.

(3.9) Dimension. a) Let K be a skew field, and let V be a finitely generated K-vector space. Then, by (3.7) and (3.8), any finite K-generating set of V contains a K-basis of V, any K-linearly independent subset of V can be extended to a K-basis of V, and all K-bases of V are finite of the same cardinality,

The cardinality of any K-basis of V is called the K**-dimension** $\dim_K(V) \in \mathbb{N}_0$ of V; if V is not finitely generated then we write $\dim_K(V) = \infty$. For example, for $m, n \in \mathbb{N}_0$ we have $\dim_K(K^n) = n$ and $\dim_K(K^{m \times n}) = m \cdot n$, while $\dim_{\mathbb{R}}(C^\infty(\mathbb{R})) = \infty$.

b) We have the following further **numerical characterisation of bases**: For any subset $B \subseteq V$ the following are equivalent:
i) B is a K-basis of V.
ii) B is a K-linearly independent subset of maximal cardinality.
ii') B is a K-linearly independent subset such that $|B| = \dim_K(V)$.
iii) B is a K-generating set of minimal cardinality.
iii') B is a K-generating set such that $|B| = \dim_K(V)$.

(3.10) Theorem. Let K be a skew field, let V be a finitely generated K-vector space. Then any $U \leq V$ is finitely generated, and we have $\dim_K(U) \leq \dim_K(V)$, with equality if and only if $U = V$. Moreover, U has a complement in V.

Note that the (in)equaity is the analogue for finite-dimensional K-vector spaces of a property of finite sets, see (1.7).

Proof. Since any K-linearly independent subset of $U \subseteq V$ has cardinality at most $\dim_K(V) \in \mathbb{N}_0$, there is a maximal K-linearly independent subset $B \subseteq U$. Hence B is a K-basis of U, and thus $\dim_K(U) = |B| \leq \dim_K(V)$. If $U < V$ then B is K-linearly independent, but is not a K-generating set of V, hence is properly contained in a K-basis of V, implying $|B| < \dim_K(V)$.

Let $B := \{v_1, \ldots, v_m\} \subseteq U$ be a K-basis of U, where $m := \dim_K(U) \in \mathbb{N}_0$, let $B' := \{w_{m+1}, \ldots, w_n\} \subseteq V$ such that $B \mathbin{\dot\cup} B'$ is a K-basis of V, where $n := \dim_K(V) \in \mathbb{N}_0$, and let $U' := \langle B' \rangle_K \leq V$. Thus we have $U + U' = \langle B \mathbin{\dot\cup} B' \rangle_K = V$. Moreover, whenever $a_1, \ldots, a_n \in K$ such that $\sum_{i=1}^m a_i v_i + \sum_{i=m+1}^n a_i w_i = 0$, since $B \mathbin{\dot\cup} B'$ is K-linearly independent we get $[a_1, \ldots, a_n] = 0$. Hence any $[u, u'] \in (U \setminus \{0\}) \times (U' \setminus \{0\})$ is K-linearly independent. ♯

(3.11) Theorem: Dimension formula for subspaces. Let K be a skew field, let V be a K-vector space, and let $U, U' \leq V$ be finitely generated. Then we have $\dim_K(U) + \dim_K(U') = \dim_K(U + U') + \dim_K(U \cap U')$; see (4.6).

Proof. Let $m := \dim_K(U) \in \mathbb{N}_0$ and $l := \dim_K(U') \in \mathbb{N}_0$. Then $U + U'$ is finitely generated, hence both $n := \dim_K(U + U') \in \mathbb{N}_0$ and $k := \dim_K(U \cap U') \in \mathbb{N}_0$ are well-defined. Let $C := \{v_1, \ldots, v_k\}$ be a K-basis of $U \cap U'$, and let $B := \{w_1, \ldots, w_{m-k}\} \subseteq U$ and $B' := \{w'_1, \ldots, w'_{l-k}\} \subseteq U'$ such that $C \mathbin{\dot\cup} B$ and $C \mathbin{\dot\cup} B'$ are K-bases of U and U', respectively. Hence we have $\langle C \cup B \cup B' \rangle_K = U + U'$.

Let $a_1, \ldots, a_k, b_1, \ldots, b_{m-k}, b'_1, \ldots, b'_{l-k} \in K$ such that $\sum_{i=1}^k a_i v_i + \sum_{i=1}^{m-k} b_i w_i + \sum_{i=1}^{l-k} b'_i w'_i = 0$. Then we have $\sum_{i=1}^k a_i v_i + \sum_{i=1}^{m-k} b_i w_i = -\sum_{i=1}^{l-k} b'_i w'_i \in U \cap U' = \langle C \rangle_K$, which since $C \mathbin{\dot\cup} B$ is K-linearly independent implies $[b_1, \ldots, b_{m-k}] = 0$. Similarly we infer that $[b'_1, \ldots, b'_{l-k}] = 0$, which yields $\sum_{i=1}^k a_i v_i = 0$ and thus $[a_1, \ldots, a_k] = 0$. This shows that $[v_1, \ldots, v_k, w_1, \ldots, w_{m-k}, w'_1, \ldots, w'_{l-k}]$ is K-linearly independent. Thus we have $B \cap B' = \emptyset$, and hence $C \mathbin{\dot\cup} B \mathbin{\dot\cup} B'$ is a K-basis of $U + U'$, where $|C \mathbin{\dot\cup} B \mathbin{\dot\cup} B'| = k + (m - k) + (l - k) = m + l - k$. ♯

(3.12) Example: Linear recurrent sequences. Let K be a field, let $V := \mathrm{Maps}(\mathbb{N}_0, K)$, where we write $x = [x_i \in K; i \in \mathbb{N}_0] \in V$, and let $a_0, \ldots, a_{n-1} \in K$, where $n \in \mathbb{N}$. Then the set of **linear recurrent sequences** $\mathcal{L} := \{x \in V; x_{i+n} = \sum_{j=0}^{n-1} a_j x_{i+j} \text{ for all } i \in \mathbb{N}_0\}$ of **degree** n is a K-subspace of V.

Moreover, $\mathcal{L}$ is finitely generated and we have $\dim_K(\mathcal{L}) = n$: Starting with any $[x_0, \ldots, x_{n-1}] \in K^n$, the recurrence shows that there is a unique $x \in \mathcal{L}$ whose first n entries are $[x_0, \ldots, x_{n-1}]$; in particular $0 \in K^n$ yields $0 \in \mathcal{L}$. For $k \in \{0, \ldots, n-1\}$ let $v_k := [0, \ldots, 0, 1, 0, \ldots, 0; *, \ldots] \in \mathcal{L}$, where amongst the first n entries only the k-th one is non-zero. Hence $\{v_0, \ldots, v_{n-1}\}$ is K-linearly independent, and given $x \in \mathcal{L}$ then $x - \sum_{k=0}^{n-1} x_k v_k = [0, \ldots, 0; *, \ldots] \in \mathcal{L}$ shows that $x = \sum_{k=0}^{n-1} x_k v_k \in \langle v_0, \ldots, v_{n-1}\rangle_K$, thus $\{v_0, \ldots, v_{n-1}\}$ is a K-basis of $\mathcal{L}$. We proceed to exhibit another K-basis of $\mathcal{L}$:

To this end, for $b \in K$ let $w_n(b) := [b^0, \ldots, b^{n-1}] \in K^n$. Then for pairwise different $b_1, \ldots, b_n \in K$ the set $\{w_n(b_1), \ldots, w_n(b_n)\}$ is K-linearly independent; see (6.8): We proceed by induction on $n \in \mathbb{N}_0$, the case $n = 0$ being trivial. Hence for $n \in \mathbb{N}$ let $a_1, \ldots, a_n \in K$ such that $\sum_{k=1}^{n} a_k w_n(b_k) = 0$. Thus we have $\sum_{k=1}^{n} a_k b_k^i = 0$ for all $i \in \{0, \ldots, n-1\}$, and hence $0 = b_n \cdot \sum_{k=1}^{n} a_k b_k^{i-1} - \sum_{k=1}^{n} a_k b_k^i = \sum_{k=1}^{n-1} a_k(b_n - b_k) b_k^{i-1}$ for all $i \in \{1, \ldots, n-1\}$. Since $\{w_{n-1}(b_1), \ldots, w_{n-1}(b_{n-1})\} \subseteq K^{n-1}$ is K-linearly independent, we get $a_k(b_n - b_k) = 0$, which since $b_n \neq b_k$ implies $a_k = 0$, for all $k \in \{1, \ldots, n-1\}$. Thus from $w_n(b_n) \neq 0 \in K^n$ we get $a_n = 0$ as well.

Letting $w(b) := \sum_{k=0}^{n-1} b^k v_k \in \mathcal{L}$, we infer that $\{w(b_1), \ldots, w(b_n)\}$ is K-linearly independent, and thus is a K-basis of $\mathcal{L}$. If $b \in K$ is a **root** of the **polynomial** $f := X^n - \sum_{i=0}^{n-1} a_i X^i \in K[X]$, see (7.9), that is $f(b) := b^n - \sum_{i=0}^{n-1} a_i b^i = 0 \in K$, then the recurrence yields $w(b) = [b^i; i \in \mathbb{N}_0]$. Thus if f has n pairwise different roots $b_1, \ldots, b_n \in K$ then $\{[b_k^i; i \in \mathbb{N}_0]; k \in \{1, \ldots, n\}\}$ is a K-basis of $\mathcal{L}$.

a) Let $K := \mathbb{R}$, and $n = 2$, and $a_0 = a_1 = 1$, that is the recurrence reads $x_{i+2} = x_i + x_{i+1}$ for all $i \in \mathbb{N}_0$. Starting with $x_0 = 0$ and $x_1 = 1$ we get the sequence $[0, 1; 1, 2, 3, 5, 8, 13, \ldots] \in \mathcal{L}$ of **Fibonacci numbers [1202]**; see (8.5).

Now $f = X^2 - X - 1 \in \mathbb{R}[X]$ has roots $\rho := \frac{1}{2}(1 + \sqrt{5}) \in \mathbb{R}$, called the **golden ratio**, and $\rho' := \frac{1}{2}(1 - \sqrt{5}) \in \mathbb{R}$. Hence $\{w(\rho), w(\rho')\}$ is an $\mathbb{R}$-basis of $\mathcal{L}$, and from $\frac{1}{\sqrt{5}} \cdot ([1, \rho] - [1, \rho']) = [0, 1] \in \mathbb{R}^2$ we conclude $x_i = \frac{1}{\sqrt{5}} \cdot (\rho^i - \rho'^i) = \frac{1}{\sqrt{5}} \cdot ((\frac{1+\sqrt{5}}{2})^i - (\frac{1-\sqrt{5}}{2})^i) \in \mathbb{N}_0$ for all $i \in \mathbb{N}_0$; in particular we have $\lim_{i \to \infty} \frac{x_i \sqrt{5}}{\rho^i} = 1$.

b) A **stream cipher** is built as follows: The **plain text** and **cipher text** spaces are $V := \mathrm{Maps}(\mathbb{N}_0, \mathbb{Z}_2)$, and a plain text $x \in V$ is **encrypted** by adding a fixed **key** $y \in V$, yielding the cipher text $x' := x + y \in V$, which in turn is **decrypted** by adding the key y again, giving back the plain text $x = x' + y$.

To facilitate the exchange of the key y, the latter is chosen as a recurrent sequence, so that only the recurrence relation and the initial values $y_0, \ldots, y_{n-1} \in \mathbb{Z}_2$, where n is the degree of the recurrence, have to be agreed upon. Then the whole of $y \in V$ is generated using **linear feedback shift registers (LFSR)**, where identifying the elements 0 and 1 of $\mathbb{Z}_2$ with false and true, respectively, addition and multiplication in $\mathbb{Z}_2$ become the logical xor and and operations. To analyse the behaviour of LFSRs, the recurrence polynomial is used.

4 Linear maps

(4.1) Linear maps. a) Let R be a ring, and let V and W be R-modules. Then a map $\varphi\colon V \to W$ fulfilling the following conditions is called **R-linear** or an **R-homomorphism**: We have **additivity** $\varphi(v + v') = \varphi(v) + \varphi(v')$ and **proportionality** $\varphi(av) = a\varphi(v)$, for all $v, v' \in V$ and $a \in R$,

An R-linear map $\varphi\colon V \to W$ is called an **R-epimorphism**, an **R-monomorphism**, or an **R-isomorphism**, if φ is surjective, injective, or bijective, respectively; if there is an R-isomorphism $V \to W$, then we write $V \cong W$. An R-linear map $V \to V$ is called an **R-endomorphism**, and a bijective R-endomorphism is called an **R-automorphism** or **regular**; a non-bijective R-endomorphism is called **singular**.

For example, we consider the column K-spaces $K^{n\times 1}$ and $K^{m\times 1}$, where $m, n \in \mathbb{N}_0$: Given $A = [a_{ij}]_{ij} \in K^{m\times n}$, the matrix product yields the K-linear map $\varphi_A\colon K^{n\times 1} \to K^{m\times 1}\colon x \mapsto Ax$: It is immediate from the explicit description $\varphi_A\colon [x_1, \ldots, x_n]^{\mathrm{tr}} \mapsto [\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j]^{\mathrm{tr}}$ that we have $A(x + y) = Ax + Ay$ and $A(ax) = a \cdot Ax$, for all $x, y \in K^{n\times 1}$ and $a \in K$.

For example, for $C^\infty(\mathbb{R}) := \{f\colon \mathbb{R} \to \mathbb{R}; f \text{ smooth}\} \le \mathrm{Maps}(\mathbb{R}, \mathbb{R})$, differentiation $\frac{\partial}{\partial x}\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ is an $\mathbb{R}$-endomorphism, which since $\frac{\partial}{\partial x}(1) = 0$ is not injective, but using integration is seen to be surjective. For $C^0([0, 1]) := \{f\colon [0, 1] \to \mathbb{R}; f \text{ continuous}\} \le \mathrm{Maps}([0, 1], \mathbb{R})$, by letting $\int_0 f\colon [0, 1] \to \mathbb{R}\colon x \mapsto \int_0^x f$ we get an $\mathbb{R}$-endomorphism $\int_0$ of $C^0([0, 1])$, which since $f(0) = 0$ for all $f \in \mathrm{im}(\int_0)$ is not surjective, but using differentiation is seen to be injective.

b) For an R-linear map $\varphi\colon V \to W$ we have $\varphi(0) = \varphi(0+0) = \varphi(0)+\varphi(0)$, hence $\varphi(0) = 0$, and for $v \in V$ we have $\varphi(-v) = \varphi((-1) \cdot v) = (-1) \cdot \varphi(v) = -\varphi(v)$. Hence we have $0 \in \mathrm{im}(\varphi)$, and for $w, w' \in \mathrm{im}(\varphi)$ and $a \in R$, letting $v, v' \in V$ such that $\varphi(v) = w$ and $\varphi(v') = w'$, we have $w+w' = \varphi(v)+\varphi(v') = \varphi(v+v') \in \mathrm{im}(\varphi)$ and $aw = a\varphi(v) = \varphi(av) \in \mathrm{im}(\varphi)$, hence $\mathrm{im}(\varphi) \le W$. If $R = K$ is a skew field, then $\mathrm{rk}(\varphi) := \dim_K(\mathrm{im}(\varphi)) \in \mathbb{N}_0 \,\dot{\cup}\, \{\infty\}$ is called the **rank** of φ.

Let $\ker(\varphi) := \{v \in V; \varphi(v) = 0\}$ be the **kernel** of φ. Hence we have $0 \in \ker(\varphi)$, and for $v, v' \in \ker(\varphi)$ and $a \in R$ we have $\varphi(v + v') = \varphi(v) + \varphi(v') = 0$ and $\varphi(av) = a\varphi(v) = 0$, hence $v + v' \in \ker(\varphi)$ and $av \in \ker(\varphi)$, thus $\ker(\varphi) \le V$. Moreover, for all $v, v' \in V$ we have $\varphi(v) = \varphi(v')$ if and only if $\varphi(v - v') = 0$, that is $v - v' \in \ker(\varphi)$. Thus φ is injective if and only if $\ker(\varphi) = \{0\}$.

If $\varphi\colon V \to W$ is an R-isomorphism, then $\varphi^{-1}\colon W \to V$ is a R-linear as well, showing that V and W, together with addition and scalar multiplication, can be identified via φ: For $w, w' \in W$ and $a \in R$, letting $v := \varphi^{-1}(w)$ and $v' := \varphi^{-1}(w')$, we have $\varphi(v+v') = \varphi(v)+\varphi(v') = w+w'$, thus $\varphi^{-1}(w+w') = v+v' = \varphi^{-1}(w) + \varphi^{-1}(w')$, and $\varphi(av) = a\varphi(v) = aw$, thus $\varphi^{-1}(aw) = av = a\varphi^{-1}(w)$.

c) Let $\mathrm{Hom}_R(V,W) := \{\varphi\colon V \to W; \varphi\ R\text{-linear}\}$. Then for $\varphi, \psi \in \mathrm{Hom}_R(V,W)$ and $a \in R$ we have $\varphi + \psi \in \mathrm{Hom}_R(V,W)$ and $a\varphi \in \mathrm{Hom}_R(V,W)$, hence we conclude that $\mathrm{Hom}_R(V,W) \leq \mathrm{Maps}(V,W)$.

If $\varphi \in \mathrm{Hom}_R(V,W)$ and $\psi \in \mathrm{Hom}_R(U,V)$, where U is an R-module, then for all $u, u' \in U$ and $a \in R$ we have $\varphi\psi(u+u') = \varphi(\psi(u) + \psi(u')) = \varphi\psi(u) + \varphi\psi(u')$ and $\varphi\psi(au) = \varphi(a\psi(u)) = a\varphi\psi(u)$, hence $\varphi\psi \in \mathrm{Hom}_R(U,W)$.

Hence $\mathrm{End}_R(V) := \mathrm{Hom}_R(V,V)$ is a ring, called the ***R*-endomorphism ring** of V, with pointwise addition and composition as multiplication, the multiplicative neutral element being the identity map: We have $(\varphi + \varphi')\psi = \varphi\psi + \varphi'\psi \in \mathrm{End}_R(V)$ and $\psi(\varphi + \varphi') = \psi\varphi + \psi\varphi' \in \mathrm{End}_R(V)$, for all $\varphi, \varphi', \psi \in \mathrm{End}_R(V)$. Note that $\mathrm{End}_R(V)$ is in general non-commutative.

The unit group $\mathrm{GL}(V) := \mathrm{End}_R(V)^*$ is called the **general linear group** on V. Since an R-endomorphism is invertible in $\mathrm{End}_R(V)$ if and only if it is invertible in $\mathrm{Maps}(V,V)$, we have $\mathrm{GL}(V) = \{\varphi \in \mathrm{End}_R(V); \varphi \text{ bijective}\}$, an in general non-commutative group, with neutral element the identity map $\mathrm{id}_V \in \mathrm{GL}(V)$ and inverses given by $\varphi^{-1} \in \mathrm{GL}(V)$, for $\varphi \in \mathrm{GL}(V)$.

(4.2) Theorem: Linear maps and bases. Let $R \neq \{0\}$ be a ring, let V and W be R-modules, let $B := \{v_i \in V; i \in \mathcal{I}\}$ be an R-generating set of V, where $\mathcal{I}$ is a set, and let $C := [w_i \in W; i \in \mathcal{I}]$.
a) There is at most one R-linear map $\varphi\colon V \to W$ such that $\varphi(v_i) = w_i$, for all $i \in \mathcal{I}$. If φ exists then we have $\mathrm{im}(\varphi) = \langle C \rangle_R \leq W$; in particular, in this case φ is surjective if and only if C is an R-generating set of W.
b) If B is an R-basis, then such a map φ exists; in other words, R-linear maps can be defined, and then are uniquely determined, by prescribing arbitrarily the images of the elements of any chosen R-basis. Moreover, φ is injective if and only if C is R-linearly independent; in particular, φ is bijective if and only if $C \subseteq W$ is an R-basis.
c) If $B := [v_1, \ldots, v_n]$ is an R-basis of V, where $n \in \mathbb{N}_0$, then the R-linear map $V \to R^n\colon v \mapsto v_B$ is an R-isomorphism. In particular, if $R = K$ is a skew field and V is finitely generated, then we have $V \cong K^n$ if and only if $n = \dim_K(V)$; in other words, any K-vector space of K-dimension $n \in \mathbb{N}_0$ can be identified, via chosing a K-basis, with the row space K^n.

Proof. a) Since $B := \{v_i \in V; i \in \mathcal{I}\}$ is an R-generating set, for all $v \in V$ there is a finite subset $\mathcal{J} \subseteq \mathcal{I}$ and $[a_j \in R; j \in \mathcal{J}]$ such that $v = \sum_{j \in \mathcal{J}} a_j v_j \in V$. Hence if φ is as asserted, then we have $\varphi(v) = \sum_{j \in \mathcal{J}} a_j \varphi(v_j) = \sum_{j \in \mathcal{J}} a_j w_j \in W$, thus φ is uniquely determined.

b) Since B is R-linearly independent, the above representation $v = \sum_{j \in \mathcal{J}} a_j v_j$ is essentially unique. Hence there is a well-defined map $\varphi\colon V \to W$ given by letting $\varphi(v) := \sum_{j \in \mathcal{J}} a_j w_j \in W$. We show that φ is R-linear: We may assume that $v' = \sum_{j \in \mathcal{J}} a'_j v_j$ where $[a'_j \in R; i \in \mathcal{J}]$. Then $v + v' = \sum_{j \in \mathcal{J}} (a_j + a'_j) v_j$ and $av = \sum_{j \in \mathcal{J}} a a_j v_j$, for all $a \in R$, shows $\varphi(v + v') = \sum_{j \in \mathcal{J}} (a_j + a'_j) w_j = \sum_{j \in \mathcal{J}} a_j w_j + \sum_{j \in \mathcal{J}} a'_j w_j = \varphi(v) + \varphi(v')$, and $\varphi(av) = \sum_{j \in \mathcal{J}} a a_j w_j = a\varphi(v)$.

If C is R-linearly independent, then for $v = \sum_{j\in\mathcal{J}} a_j v_j \in \ker(\varphi)$ we have $\varphi(v) = \sum_{j\in\mathcal{J}} a_j w_j = 0$, implying $a_j = 0$ for all $j \in \mathcal{J}$, showing that $\ker(\varphi) = \{0\}$. Conversely, if $\ker(\varphi) = \{0\}$, then for $\sum_{j\in\mathcal{J}} a_j w_j = 0 \in W$, where $\mathcal{J} \subseteq \mathcal{I}$ finite and $[a_j \in R; j \in \mathcal{J}]$, we have $\sum_{j\in\mathcal{J}} a_j v_j \in \ker(\varphi) = \{0\}$, implying $a_j = 0$ for all $j \in \mathcal{J}$, showing that C is R-linearly independent.

c) Let $[e_1, \ldots, e_n]$ be the standard R-basis of R^n, and let $\beta\colon V \to R^n$ be the R-linear map defined by $v_i \mapsto e_i$, for $i \in \{1, \ldots, n\}$. Then for $v = \sum_{i=1}^n a_i v_i \in V$, where $a_1, \ldots, a_n \in R$, we have $\beta(v) = \sum_{i=1}^n a_i e_i = [a_1, \ldots, a_n] = v_B \in R^n$. ♯

(4.3) Theorem: Dimension formula. Let K be a skew field, V be a finitely generated K-vector space, W be a K-vector space, and $\varphi\colon V \to W$ be a K-linear map. Then we have $\dim_K(V) = \dim_K(\ker(\varphi)) + \mathrm{rk}(\varphi) \in \mathbb{N}_0$; see (4.6).

Proof. Since $\mathrm{im}(\varphi) \leq W$ is finitely generated, we conclude that $r := \mathrm{rk}(\varphi) = \dim_K(\mathrm{im}(\varphi)) \in \mathbb{N}_0$. Let $C := \{w_1, \ldots, w_r\} \subseteq \mathrm{im}(\varphi)$ be a K-basis, let $v_j \in V$ such that $\varphi(v_j) = w_j$ for all $j \in \{1, \ldots, r\}$, and let $B := \{v_1, \ldots, v_r\} \subseteq V$. Moreover, let $B' := \{v'_1, \ldots, v'_k\} \subseteq \ker(\varphi)$ be a K-basis, where $k := \dim_K(\ker(\varphi)) \leq \dim_K(V) \in \mathbb{N}_0$. We show that the sequence $[B', B] := [v'_1, \ldots, v'_k, v_1, \ldots, v_r] \subseteq V$ is a K-basis, implying $\dim_K(V) = k + r$:

Since $C \subseteq \mathrm{im}(\varphi)$ is a K-generating set, for $v \in V$ we have $\varphi(v) = \sum_{j=1}^r a_j w_j \in \mathrm{im}(\varphi)$, for some $a_1, \ldots, a_r \in K$. Hence we have $v - \sum_{j=1}^r a_j v_j \in \ker(\varphi)$, thus since $B' \subseteq \ker(\varphi)$ is a K-generating set there are $a'_1, \ldots, a'_k \in K$ such that $v = \sum_{i=1}^k a'_i v'_i + \sum_{j=1}^r a_j v_j \in V$, thus $[B', B] \subseteq V$ is a K-generating set.

Let $a'_1, \ldots, a'_k, a_1, \ldots, a_r \in K$ such that $\sum_{i=1}^k a'_i v'_i + \sum_{j=1}^r a_j v_j = 0 \in V$, thus $0 = \varphi(\sum_{i=1}^k a'_i v'_i) + \varphi(\sum_{j=1}^r a_j v_j) = \sum_{j=1}^r a_j w_j \in \mathrm{im}(\varphi)$. Since $C \subseteq \mathrm{im}(\varphi)$ is K-linearly independent, we conclude $a_1 = \cdots = a_r = 0$, from which we infer $\sum_{i=1}^k a'_i v'_i = 0 \in V$, and since $B' \subseteq \ker(\varphi)$ is K-linearly independent we get $a'_1 = \cdots = a'_k = 0$, showing that $[B', B] \subseteq V$ is K-linearly independent. ♯

(4.4) Corollary. Let K be a skew field, let V be a finitely generated K-vector space, and let $\varphi \in \mathrm{End}_K(V)$. Then the following are equivalent:
i) φ is a K-automorphism, that is φ is invertible, that is φ is bijective.
ii) φ is a K-monomorphism, that is φ is injective.
ii') $\varphi \in \mathrm{End}_K(V)$ is left invertible.
ii") For $\psi \in \mathrm{End}_K(V)$ we have $\varphi\psi = 0$ if and only if $\psi = 0$.
iii) φ is a K-epimorphism, that is φ is surjective.
iii') $\varphi \in \mathrm{End}_K(V)$ is right invertible.
iii") For $\psi \in \mathrm{End}_K(V)$ we have $\psi\varphi = 0$ if and only if $\psi = 0$.

Note that this is the analogue for K-vector spaces of the equivalence of injectivity and surjectivity of maps from a finite set to itself, see (1.7) and (1.10).

Proof. i)⇔ii)⇔iii): We have $\dim_K(V) = \dim_K(\ker(\varphi)) + \mathrm{rk}(\varphi)$, hence we have $\dim_K(\ker(\varphi)) = 0$ if and only if $\mathrm{rk}(\varphi) = \dim_K(V)$, in other words we have $\ker(\varphi) = \{0\}$ if and only if $\mathrm{im}(\varphi) = V$, which says that φ is injective if and only if φ is surjective.

i)⇒ii') and **i)⇒iii')** are clear. **ii')⇒ii)** and **iii')⇒iii)**: If $\varphi'' \in \mathrm{End}_K(V)$ is a left inverse of φ, then $\varphi''\varphi = \mathrm{id}_V$ implies that φ is injective. If $\varphi' \in \mathrm{End}_K(V)$ is a right inverse of φ, then $\varphi\varphi' = \mathrm{id}_V$ implies that φ is surjective.

i)⇒ii") and **i)⇒iii")**: From $\varphi\psi = 0$ we get $\psi = \varphi^{-1}\varphi\psi = 0$, and from $\psi\varphi = 0$ we get $\psi = \psi\varphi\varphi^{-1} = 0$.

ii")⇒ii): Let $B = [v_1, \ldots, v_n]$ be a K-basis of V, where $n := \dim_K(V) \in \mathbb{N}_0$ and $[v_1, \ldots, v_m]$ is a K-basis of $\ker(\varphi) \leq V$, for some $m \leq n$, and let $\psi \in \mathrm{End}_K(V)$ be given by $\psi(v_i) = v_i$ for $i \in \{1, \ldots, m\}$, and $\psi(v_i) = 0$ for $i \in \{m+1, \ldots, n\}$. Then we have $\varphi\psi = 0$, implying $\psi = 0$, and thus $m = 0$, that is $\ker(\varphi) = \{0\}$.

iii")⇒iii): Let $B = [v_1, \ldots, v_n]$ be a K-basis of V, where $n := \dim_K(V)$ and $[v_1, \ldots, v_m]$ is a K-basis of $\mathrm{im}(\varphi) \leq V$, for some $m \leq n$, and let $\psi \in \mathrm{End}_K(V)$ be given by $\psi(v_i) = 0$ for $i \in \{1, \ldots, m\}$, and $\psi(v_i) = v_i$ for $i \in \{m+1, \ldots, n\}$. Then we have $\psi\varphi = 0$, implying $\psi = 0$, and thus $m = n$, that is $\mathrm{im}(\varphi) = V$. ♯

(4.5) Theorem. Let R be a ring, let V be an R-module and let $U \leq V$.
a) Then the relation on V defined by $v \sim_U w$ if $v - w \in U$, for all $v, w \in V$, is an equivalence relation.

Let $V/U := \{v + U; v \in V\}$ be the associated set of equivalence classes, called **cosets** of U in V, where we write $v + U := [v]_{\sim_U} = \{w \in V; w - v \in U\} = \{v + u \in V; u \in U\}$; for example, if $U = \{0\}$ then $v + U = \{v\}$ for all $v \in V$, and if $U = V$ then $V/U = \{0 + U\}$.

b) Then V/U is an R-module, called the **quotient R-module** of V with respect to U, with addition $(v+U) + (w+U) := (v+w) + U$, for all $v, w \in V$, additive neutral element $0 + U$, the additive inverse of $v + U$ being $(-v) + U$, and scalar multiplication $a \cdot (v + U) := av + U$, for all $a \in R$.

The **natural map** $\nu_U \colon V \to V/U \colon v \mapsto v + U$ is an R-epimorphism such that $\ker(\nu_U) = U$; in particular, any subset of V is an R-submodule if and only if it is the kernel of a suitably chosen R-homomorphism.

c) Let $R \neq \{0\}$, let $\{v_i \in U; i \in \mathcal{I}\}$ be an R-basis of U, and let $\{v_j \in V; j \in \mathcal{J}\}$, where $\mathcal{I} \cap \mathcal{J} = \emptyset$. Then $\{v_j + U \in V/U; j \in \mathcal{J}\}$ is an R-basis of V/U if and only if $\{v_k \in V; k \in \mathcal{I} \,\dot\cup\, \mathcal{J}\}$ is an R-basis of V.

In particular, if K is a skew field and V is finitely generated, then V/U is finitely generated such that $\dim_K(V) = \dim_K(U) + \dim_K(V/U)$.

Proof. a) We have $v - v = 0 \in U$, hence $\sim_U$ is reflexive. If $v - w \in U$ then $w - v = -(v - w) \in U$ as well, hence $\sim_U$ is symmetric. If $v - w \in U$ and $w - x \in U$ then we have $v - x = (v - w) + (w - x) \in U$, hence $\sim_U$ is transitive.

b) We only have to show that addition and scalar multiplication are independent from the choice of representatives of the equivalence classes: For $v - v' \in U$ and $w - w' \in U$ we have $(v + w) - (v' + w') = (v - v') + (w - w') \in U$, hence $(v + w) + U = (v' + w') + U$ and $(av) - (av') = a(v - v') \in U$, for $a \in R$.

We have $\nu_U(v + v') = (v + v') + U = (v + U) + (v' + U) = \nu_U(v) + \nu_U(v')$ and $\nu_U(av) = av+U = a(v+U) = a\nu_U(v)$, for all $v, v' \in V$ and $a \in R$. Moreover, we have $\nu_U(v) = v+U = 0+U \in V/U$ if and only if $v \in U$, hence $U = \ker(\nu_U) \leq V$.

c) Let $\{v_k; k \in \mathcal{I} \,\dot\cup\, \mathcal{J}\}$ be an R-basis of V. For $v \in V$ there are $\mathcal{I}' \subseteq \mathcal{I}$ and $\mathcal{J}' \subseteq \mathcal{J}$ finite and $[a_k \in R; k \in \mathcal{I}' \,\dot\cup\, \mathcal{J}']$ such that $v = \sum_{k\in\mathcal{I}'\dot\cup\mathcal{J}'} a_k v_k \in V$, thus $v+U = \sum_{k\in\mathcal{I}'\dot\cup\mathcal{J}'} a_k(v_k+U) = \sum_{j\in\mathcal{J}'} a_j(v_j+U) \in V/U$, hence $\{v_j+U; j \in \mathcal{J}\}$ is an R-generating set of V/U. Let $\mathcal{J}' \subseteq \mathcal{J}$ be finite and $[a_j \in R; j \in \mathcal{J}']$ such that $\sum_{j\in\mathcal{J}'} a_j(v_j + U) = 0 \in V/U$, then there are $\mathcal{I}' \subseteq \mathcal{I}$ finite and $[a_i \in R; i \in \mathcal{I}']$ such that $\sum_{j\in\mathcal{J}'} a_j v_j = \sum_{i\in\mathcal{I}'} a_i v_i \in U$, implying that $a_k = 0$ for all $k \in \mathcal{I}' \,\dot\cup\, \mathcal{J}'$, hence $\{v_j + U; j \in \mathcal{J}\}$ is R-linearly independent.

Let conversely $\{v_j + U; j \in \mathcal{J}\}$ be an R-basis of V/U. For $v \in V$ there are $\mathcal{J}' \subseteq \mathcal{J}$ finite and $[a_j \in R; j \in \mathcal{J}']$ such that $v + U = \sum_{j\in\mathcal{J}'} a_j(v_j + U) \in V/U$, then there are $\mathcal{I}' \subseteq \mathcal{I}$ finite and $[a_i \in R; i \in \mathcal{I}']$ such that $v - \sum_{j\in\mathcal{J}'} a_j v_j = \sum_{i\in\mathcal{I}'} a_i v_i \in U$, hence $\{v_k; k \in \mathcal{I} \,\dot\cup\, \mathcal{J}\}$ is an R-generating set of V. Let $\mathcal{I}' \subseteq \mathcal{I}$ and $\mathcal{J}' \subseteq \mathcal{J}$ be finite, and let $[a_i \in R; i \in \mathcal{I}' \,\dot\cup\, \mathcal{J}']$ such that $\sum_{k\in\mathcal{I}'\dot\cup\mathcal{J}'} a_k v_k = 0 \in V$, then $\sum_{k\in\mathcal{J}'} a_j(v_j + U) = 0 \in V/U$ implies $a_j = 0$ for all $j \in \mathcal{J}'$, and from that $\sum_{i\in\mathcal{I}'} a_i v_i = 0 \in U$ implies $a_i = 0$ for all $i \in \mathcal{I}'$ as well, hence $\{v_k; k \in \mathcal{I} \,\dot\cup\, \mathcal{J}\}$ is R-linearly independent. ♯

(4.6) Theorem. Let R be a ring and let V be an R-module.
a) Let W be an R-module and $\varphi \in \mathrm{Hom}_R(V, W)$. Then we have the **homomorphism principle**: The map $\overline{\varphi}\colon V/\ker(\varphi) \to \mathrm{im}(\varphi)\colon v + \ker(\varphi) \mapsto \varphi(v)$ is an R-isomorphism such that $\varphi = \overline{\varphi}\nu_{\ker(\varphi)}$.

In particular, if $R = K$ is a skew field and V is finitely generated, then we have $\dim_K(V) = \dim_K(\ker(\varphi)) + \dim_K(\mathrm{im}(\varphi)) = \dim_K(\ker(\varphi)) + \mathrm{rk}(\varphi)$; thus we recover (4.3) as a special case.

b) If $U, W \leq V$ then we have the **isomorphism theorems** $U/(U \cap W) \cong (U + W)/W$, and if additionally $U \leq W \leq V$ then $(V/U)/(W/U) \cong V/W$.

In particular, if K is a skew field and $U, W \leq V$ are finitely generated, then we have $\dim_K(U) + \dim_K(W) = \dim_K(U + W) + \dim_K(U \cap W)$; thus we recover (3.11) as a special case.

Proof. a) Let $U := \ker(\varphi) \leq V$. Since for $v-v' \in U$ we have $\varphi(v) = \varphi(v') \in W$, in the definition of $\overline{\varphi}$ we have independence from the choice of representatives of the cosets. We have $\overline{\varphi}((v+v')+U) = \varphi(v+v') = \varphi(v)+\varphi(v') = \overline{\varphi}(v+U)+\overline{\varphi}(v'+U)$ and $\overline{\varphi}(avU) = \varphi(av) = a\varphi(v) = a\overline{\varphi}(v+U)$, for all $v, v' \in V$ and $a \in R$, hence $\overline{\varphi}$ is R-linear. Moreover, $\overline{\varphi}$ is surjective, and we have $v + U \in \ker(\overline{\varphi}) \in V/U$ if and only if $\overline{\varphi}(v + U) = \varphi(v) = 0$, that is $v \in \ker(\varphi) = U$, which in turn is equivalent to $v + U = 0 + U \in V/U$, implying that $\overline{\varphi}$ is injective. Finally we have $\overline{\varphi}\nu_U(v) = \overline{\varphi}(v + U) = \varphi(v)$, for all $v \in V$.

b) Let $U, W \leq V$ and $\varphi\colon U \to (U + W)/W\colon u \mapsto u + W$. Then φ is an R-epimorphism, and we have $u \in \ker(\varphi)$ if and only if $u + W = 0 + W \in (U + W)/W \leq V/W$, that is $u \in W$, implying that $\ker(\varphi) = U \cap W$.

Let $U \leq W \leq V$ and $\varphi\colon V/U \to V/W\colon v + U \mapsto v + W$. Since $v - v' \in U$ implies $v - v' \in W$, in the definition of φ we have independence from the choice of representatives of the cosets. Moreover, φ is an R-epimorphism, and we have $v + U \in \ker(\varphi)$ if and only if $v + W = 0 + W \in V/W$, that is $v \in W$, implying that $\ker(\varphi) = W/U$. $\sharp$

5 Matrices

(5.1) Linear maps and matrices. Let $R \neq \{0\}$ be a commutative ring, and let V and W be R-modules having R-bases $B := [v_1, \ldots, v_n]$ and $C := [w_1, \ldots, w_m]$, respectively, where $m, n \in \mathbb{N}_0$. Given $\varphi \in \operatorname{Hom}_R(V, W)$, let $a_{ij} \in R$, for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, such that $\varphi(v_j) = \sum_{i=1}^m a_{ij}w_i$. Thus for $v = \sum_{j=1}^n b_jv_j \in V$, where $b_1, \ldots, b_n \in R$, we obtain $\varphi(v) = \sum_{j=1}^n b_j\varphi(v_j) = \sum_{j=1}^n b_j(\sum_{i=1}^m a_{ij}w_i) = \sum_{i=1}^m(\sum_{j=1}^n a_{ij}b_j)w_j \in W$.

Hence identifying $V \to R^{n\times 1}\colon v \mapsto {}_Bv$ and $W \to R^{m\times 1}\colon w \mapsto {}_Cw$, the map φ translates into the map ${}_Bv = [b_1, \ldots, b_n]^{\text{tr}} \mapsto {}_C(\varphi(v)) = [c_1, \ldots, c_m]^{\text{tr}}$, where $c_i := \sum_{j=1}^n a_{ij}b_j \in R$. In other words, letting ${}_C\varphi_B = [a_{ij}]_{ij} \in R^{m\times n}$ be the **matrix** of φ with respect to the R-bases B and C, we get ${}_Bv \mapsto {}_C\varphi_B \cdot {}_Bv$, that is φ translates into $\varphi_A \in \operatorname{Hom}_R(R^{n\times 1}, R^{m\times 1})$, where $A := {}_C\varphi_B \in R^{m\times n}$.

Then the map ${}_C\Phi_B\colon \operatorname{Hom}_R(V, W) \to R^{m\times n}\colon \varphi \mapsto {}_C\varphi_B$ is an R-isomorphism:

Since for $\varphi, \varphi' \in \operatorname{Hom}_R(V, W)$ with matrices ${}_C\varphi_B = [a_{ij}]_{ij}$ and ${}_C\varphi'_B = [a'_{ij}]_{ij}$, respectively, and $a \in R$ we have $(\varphi+\varphi')(v_j) = \sum_{i=1}^m(a_{ij}+a'_{ij})w_i$ and $(a\varphi)(v_j) = \sum_{i=1}^m a \cdot a_{ij}w_i$, for all $j \in \{1, \ldots, n\}$, we conclude that ${}_C\Phi_B$ is R-linear. Moreover, $\varphi \in \operatorname{Hom}_R(V, W)$ being uniquely determined by ${}_C\varphi_B \in R^{m\times n}$ shows that ${}_C\Phi_B$ is injective. Finally, given any $A = [a_{ij}]_{ij} \in R^{m\times n}$, there is $\varphi \in \operatorname{Hom}_R(V, W)$ defined by $\varphi(v_j) := \sum_{i=1}^m a_{ij}w_i \in W$, for all $j \in \{1, \ldots, n\}$, thus ${}_C\varphi_B = A$, and hence ${}_C\Phi_B$ is surjective as well. $\sharp$

Letting the **matrix unit** $E_{ij} = [a_{kl}]_{kl} \in R^{m\times n}$ be defined by $a_{kl} := 1$ if $[k, l] = [i, j]$, and $a_{kl} := 0$ if $[k, l] \neq [i, j]$, where $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Then $\{E_{ij} \in R^{m\times n}; i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ is an R-basis of $R^{m\times n}$, called the **standard** R-basis, where E_{ij} describes the R-linear map sending $v_j \mapsto w_i$ and $v_k \mapsto 0$, whenever $j \neq k \in \{1, \ldots, n\}$. In particular, if $R = K$ is a field and V and W are finitely generated, then we have $\dim_K(\operatorname{Hom}_K(V, W)) = \dim_K(K^{m\times n}) = \dim_K(V) \cdot \dim_K(W)$.

Letting $\operatorname{diag}[a_1, \ldots, a_s] := \sum_{i=1}^s a_iE_{ii} \in R^{m\times n}$, where $s := \min\{m, n\}$, be the **diagonal matrix** associated with $a_1, \ldots, a_s \in R$, the identity map id_V is translated into the **identity matrix** ${}_B(\operatorname{id}_V)_B = E_n = \operatorname{diag}[1, \ldots, 1] = \sum_{i=1}^n E_{ii} =$

$[a_{ij}]_{ij} \in R^{n \times n}$, where $a_{ij} := 1$ if $i = j$, and $a_{ij} := 0$ if $i \neq j$, for $i, j \in \{1, \ldots, n\}$.

For example, with respect to the standard $\mathbb{R}$-basis of $\mathbb{R}^{2 \times 1}$, the **reflections** at the **hyperplanes perpendicular** to $[1,0]^{\mathrm{tr}}$ and $[0,1]^{\mathrm{tr}}$ are given by $\begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ and $\begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, respectively, the **rotation** by an **angle** of $\omega \in \mathbb{R}$ is given by $\begin{bmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, see (10.10), and the **rotation-dilatation** by an angle of $\frac{\pi}{4}$ and scaling factor of $\sqrt{2}$ is given by $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$.

(5.2) Matrix products. Let $R \neq \{0\}$ be a commutative ring, let U, V, and W be R-modules having R-bases $Q := [u_1, \ldots, u_l]$, $S := [v_1, \ldots, v_n]$, and $T := [w_1, \ldots, w_m]$, respectively, where $l, m, n \in \mathbb{N}_0$. Moreover, let $\varphi \in \mathrm{Hom}_R(V, W)$ and $\psi \in \mathrm{Hom}_R(U, V)$, with associated matrices ${}_T\varphi_S = [a_{ij}]_{ij} \in R^{m \times n}$ and ${}_S\psi_Q = [b_{ij}]_{ij} \in R^{n \times l}$. Then for all $k \in \{1, \ldots, l\}$ we have $\varphi\psi(u_k) = \varphi(\sum_{j=1}^n b_{jk} v_j) = \sum_{j=1}^n \sum_{i=1}^m b_{jk} a_{ij} w_i = \sum_{i=1}^m (\sum_{j=1}^n a_{ij} b_{jk}) w_i$, thus the composition $\varphi\psi \in \mathrm{Hom}_R(U, W)$ is described by ${}_T(\varphi\psi)_Q = [\sum_{j=1}^n a_{ij} b_{jk}]_{ik} \in R^{m \times l}$. Then we have ${}_T(\varphi\psi)_Q = {}_T\varphi_S \cdot {}_S\psi_Q$, as soon as we adopt the following definition:

For $A = [a_{ij}]_{ij} \in R^{m \times n}$ and $B = [b_{ij}]_{ij} \in R^{n \times l}$ we define the **matrix product** $A \cdot B := [\sum_{j=1}^n a_{ij} b_{jk}]_{ik} \in R^{m \times l}$. Moreover, for $l = 1$ the elements of $R^{n \times 1}$ are just columns, and for $[x_1, \ldots, x_n]^{\mathrm{tr}} \in R^{n \times 1}$ we recover $A \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = [\sum_{j=1}^n a_{ij} x_j]_i^{\mathrm{tr}} \in R^{m \times 1}$, coinciding with the definition made in (5.1).

Identifying $A \in R^{m \times n}$ with $\varphi_A \in \mathrm{Hom}_R(R^{n \times 1}, R^{m \times 1})$, from the associativity of maps we infer $A(BC) = (AB)C \in R^{m \times k}$, whenever $C \in R^{l \times k}$ for some $k \in \mathbb{N}_0$, hence associativity holds for the matrix product as well. Similarly, letting $A' \in R^{m \times n}$ and $B' \in R^{n \times l}$ we get $(A + A')B = AB + A'B \in R^{m \times l}$ and $A(B + B') = AB + AB' \in R^{m \times l}$.

In particular, $R^{n \times n}$ is a ring, called the **matrix ring** of degree n over R, with pointwise addition and matrix multiplication, the multiplicative neutral element being the identity matrix E_n; then $R^{n \times n}$ can be identified with $\mathrm{End}_R(R^{n \times 1})$, and in general is non-commutative.

Its unit group $\mathrm{GL}_n(R) := (R^{n \times n})^*$, having neutral element E_n, is called the **general linear group** of degree n over R; it in general is non-commutative, and its elements are called **invertible** matrices, where the inverse $A^{-1} \in \mathrm{GL}_n(R)$ of $A \in \mathrm{GL}_n(R)$ is given by the property $AA^{-1} = E_n = A^{-1}A \in R^{n \times n}$. Moreover, $\mathrm{GL}_n(R)$ can be identified with $\mathrm{GL}(R^{n \times 1})$, yielding the alternative description $\mathrm{GL}_n(R) = \{A \in R^{n \times n}; \varphi_A \in \mathrm{End}_R(R^{n \times 1}) \text{ bijective}\}$, where the description of inverses translates into $\varphi_{A^{-1}} = (\varphi_A)^{-1} \in \mathrm{End}_R(R^{n \times 1})$.

For example we have

$$\begin{bmatrix} . & -1 \\ -1 & . \end{bmatrix} \cdot \begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix} = \begin{bmatrix} . & -1 \\ 1 & . \end{bmatrix} \neq \begin{bmatrix} . & 1 \\ -1 & . \end{bmatrix} = \begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix} \cdot \begin{bmatrix} . & -1 \\ -1 & . \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R});$$

note that reflections indeed are **self-inverse** $\begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix}^2 = E_2 = \begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix}^2$.

(5.3) Base change. a) Let $R \neq \{0\}$ be a commutative ring, and let V be an R-module having R-bases $B := [v_1, \ldots, v_n]$ and $B' := [v'_1, \ldots, v'_n]$, where $n \in \mathbb{N}_0$. Then ${}_B\mathrm{id}_{B'} = [b_{ij}]_{ij} \in R^{n \times n}$ is called the associated **base change matrix**, that is we have $v'_j = \sum_{i=1}^n b_{ij} v_i$, for $j \in \{1, \ldots, n\}$. Hence we have ${}_B\mathrm{id}_{B'} \cdot {}_{B'}\mathrm{id}_B = {}_B\mathrm{id}_B = E_n$ and ${}_{B'}\mathrm{id}_B \cdot {}_B\mathrm{id}_{B'} = {}_{B'}\mathrm{id}_{B'} = E_n$, thus ${}_B\mathrm{id}_{B'} \in \mathrm{GL}_n(R)$ with inverse $({}_B\mathrm{id}_{B'})^{-1} = {}_{B'}\mathrm{id}_B$.

Letting W be an R-module having R-bases C and C', where $|C| = |C'| \in \mathbb{N}_0$, we get the **base change formula** ${}_{C'}\varphi_{B'} = {}_{C'}\mathrm{id}_C \cdot {}_C\varphi_B \cdot {}_B\mathrm{id}_{B'} = ({}_C\mathrm{id}_{C'})^{-1} \cdot {}_C\varphi_B \cdot {}_B\mathrm{id}_{B'}$.

b) Let K be a field, let V and W be K-vector spaces such that $n := \dim_K(V) \in \mathbb{N}_0$ and $m := \dim_K(W) \in \mathbb{N}_0$, and let $\varphi \in \mathrm{Hom}_K(V, W)$. Then there are K-bases B and C such that ${}_C\varphi_B = \mathrm{diag}[1, \ldots, 1, 0, \ldots, 0] = \sum_{i=1}^r E_{ii} \in K^{m \times n}$ with $r := \mathrm{rk}(\varphi) \in \{0, \ldots, \min\{m, n\}\}$ non-zero entries:

We have $\dim_K(\ker(\varphi)) = \dim_K(V) - \mathrm{rk}(\varphi) = n - r \geq 0$, hence $r \leq \min\{m, n\}$. Let $B := [v_1, \ldots, v_n]$ be a K-basis of V such that $[v_{r+1}, \ldots, v_n]$ is a K-basis of $\ker(\varphi) \leq V$. Since $\varphi(v_j) = 0$ for all $j \in \{r+1, \ldots, n\}$, we have $\mathrm{im}(\varphi) = \langle \varphi(v_1), \ldots, \varphi(v_r) \rangle_K$, thus $[\varphi(v_1), \ldots, \varphi(v_r)]$ is a K-basis of $\mathrm{im}(\varphi) \leq W$. Extending the latter to a K-basis $C := [\varphi(v_1), \ldots, \varphi(v_r), w_{r+1}, \ldots, w_m] \subseteq W$ yields ${}_C\varphi_B$ as desired. Note that this is reminiscent of the proof of (4.3). $\sharp$

c) We present an example for the base change mechanism: Let $B \subseteq \mathbb{R}^{2 \times 1}$ be the standard $\mathbb{R}$-basis and $C := [v_1, v_2] \subseteq \mathbb{R}^{2 \times 1}$ be the $\mathbb{R}$-basis given by $v_1 := [1, 1]^{\mathrm{tr}}$ and $v_2 := [-1, 1]^{\mathrm{tr}}$. Thus we have ${}_B\mathrm{id}_C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$, writing B as $\mathbb{R}$-linear combinations in C we get ${}_C\mathrm{id}_B = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$; indeed we have ${}_B\mathrm{id}_C \cdot {}_C\mathrm{id}_B = E_2 = {}_C\mathrm{id}_B \cdot {}_B\mathrm{id}_C$, that is ${}_C\mathrm{id}_B = ({}_B\mathrm{id}_C)^{-1}$.

For the reflection ρ at the hyperplane perpendicular to $[-1, 1]^{\mathrm{tr}}$, with respect to the $\mathbb{R}$-basis B we have ${}_B\rho_B = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{R}^{2 \times 2}$. There are various (equivalent) ways to find the matrix of ρ with respect to the $\mathbb{R}$-basis C:

Geometrically, we have $\rho(v_1) = v_1$ and $\rho(v_2) = -v_2$. In terms of matrices with respect to the $\mathbb{R}$-basis B this reads ${}_B(\rho(v_1)) = {}_B\rho_B \cdot {}_B(v_1) = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = {}_B(v_1)$ and ${}_B(\rho(v_2)) = {}_B\rho_B \cdot {}_B(v_2) = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = {}_B(-v_2)$. Anyway, we thus get ${}_C\rho_C = \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$; note that the latter is a diagonal matrix. Alternatively, using the base change formula we obtain

$$ {}_C\rho_C = {}_C\mathrm{id}_B \cdot {}_B\rho_B \cdot {}_B\mathrm{id}_C = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix}. $$

(5.4) Matrix rank. a) Let $R \neq \{0\}$ be a commutative ring, and let $A = [a_{ij}]_{ij} \in R^{m \times n}$, where $m, n \in \mathbb{N}_0$. Then $A^{\mathrm{tr}} := [a_{ji}]_{ij} \in R^{n \times m}$ is called the **transpose** of A. Then the map $\mathrm{tr} \colon R^{m \times n} \to R^{n \times m} \colon A \mapsto A^{\mathrm{tr}}$ is an R-isomorphism, and we have $\mathrm{tr} \cdot \mathrm{tr} = \mathrm{id}_{R^{m \times n}}$. In particular, we recover $\mathrm{tr} \colon R^n \to R^{n \times 1}$ and $\mathrm{tr} \colon R^{n \times 1} \to R^n$, mapping rows to columns and vice versa.

For $B = [b_{ij}]_{ij} \in R^{n \times l}$, where $l \in \mathbb{N}_0$, we have $(AB)^{\mathrm{tr}} = ([\sum_{j=1}^n a_{ij} b_{jk}]_{ik})^{\mathrm{tr}} = [\sum_{j=1}^n b_{ji} a_{kj}]_{ik} = [b_{ji}]_{ij} \cdot [a_{kj}]_{jk} = B^{\mathrm{tr}} A^{\mathrm{tr}}$. Thus for $A \in \mathrm{GL}_n(R)$ we have $A^{\mathrm{tr}} (A^{-1})^{\mathrm{tr}} = (A^{-1} A)^{\mathrm{tr}} = E_n$ and $(A^{-1})^{\mathrm{tr}} A^{\mathrm{tr}} = (A A^{-1})^{\mathrm{tr}} = E_n$, hence $A^{\mathrm{tr}} \in \mathrm{GL}_n(R)$ and $(A^{\mathrm{tr}})^{-1} = (A^{-1})^{\mathrm{tr}} =: A^{-\mathrm{tr}}$, called the **inverse transpose** of A.

b) Identifying $A \in R^{m \times n}$ with the R-linear map $\varphi_A \colon R^{n \times 1} \to R^{m \times 1} \colon A \mapsto Av$, let $\ker(A) := \ker(\varphi_A) = \{v \in R^{n \times 1}; Av = 0\} \leq R^{n \times 1}$ be the **(column) kernel** of A, and $\mathrm{im}(A) := \mathrm{im}(\varphi_A) = \langle w_1, \ldots, w_n \rangle_R \leq R^{m \times 1}$ be the **image** or the **column R-module** of A, where $w_j := [a_{1j}, \ldots, a_{mj}]^{\mathrm{tr}} \in R^{m \times 1}$, for $j \in \{1, \ldots, n\}$, are the **columns** of A.

For $Q \in R^{n \times n}$ we have $\mathrm{im}(AQ) = \mathrm{im}(\varphi_{AQ}) = \mathrm{im}(\varphi_A \varphi_Q) \leq \mathrm{im}(\varphi_A) = \mathrm{im}(A)$, hence for $Q \in \mathrm{GL}_n(R)$ we have $\mathrm{im}(A) = \mathrm{im}(AQQ^{-1}) \leq \mathrm{im}(AQ) \leq \mathrm{im}(A)$, implying $\mathrm{im}(AQ) = \mathrm{im}(A)$, that is the column R-modules of AQ and A coincide.

If $R = K$ is a field, then $\mathrm{rk}(A) := \mathrm{rk}(\varphi_A) = \dim_K(\mathrm{im}(\varphi_A)) = \dim_K(\mathrm{im}(A)) = \dim_K(\langle w_1, \ldots, w_n \rangle_K) \in \mathbb{N}_0$ is called the **column rank** of A; thus we have $n = \dim_K(\ker(A)) + \mathrm{rk}(A)$. In particular, for $Q \in K^{n \times n}$ we have $\mathrm{rk}(AQ) \leq \mathrm{rk}(A)$, where for $Q \in \mathrm{GL}_n(K)$ we get equality $\mathrm{rk}(AQ) = \mathrm{rk}(A)$.

Moreover, still assuming that $R = K$ is a field, for $m = n$ this yields a criterion when a matrix $A \in K^{n \times n}$ is invertible: We have $A \in \mathrm{GL}_n(K)$, that is φ_A is bijective, if and only if φ_A is surjective, that is $\mathrm{rk}(A) = n$, if and only if φ_A is injective, that is $\ker(A) = \{0\}$. Hence to infer that $A \in K^{n \times n}$ is invertible, it suffices to exhibit $B \in K^{n \times n}$ such that $AB = E_n$, since then φ_A is surjective and φ_B is injective, and thus $B = A^{-1} \in \mathrm{GL}_n(K)$; likewise it suffices to exhibit $C \in K^{n \times n}$ such that $CA = E_n$, since then φ_A is injective and φ_C is surjective, and thus $C = A^{-1} \in \mathrm{GL}_n(K)$.

c) We may also identify $A \in R^{m \times n}$ with the R-linear map $\varphi_A^{\mathrm{tr}} \colon R^m \to R^n \colon v \mapsto vA = (A^{\mathrm{tr}} v^{\mathrm{tr}})^{\mathrm{tr}}$. Then $\ker(\varphi_A^{\mathrm{tr}}) = \ker(A^{\mathrm{tr}})^{\mathrm{tr}} = \{v \in R^m; vA = 0\} \leq R^m$ is called the **row kernel** of A, and $\mathrm{im}(\varphi_A^{\mathrm{tr}}) = \mathrm{im}(A^{\mathrm{tr}})^{\mathrm{tr}} = \langle v_1, \ldots, v_m \rangle_R \leq R^n$ is called the **row image** or **row R-module** of A, where $v_i := [a_{i1}, \ldots, a_{in}] \in R^n$, for $i \in \{1, \ldots, m\}$, are the **rows** of A.

We have $\ker(A^{\mathrm{tr}}) \leq R^{m \times 1}$ and $\mathrm{im}(A^{\mathrm{tr}}) \leq R^{n \times 1}$, thus for $P \in R^{m \times m}$ we have $\mathrm{im}((PA)^{\mathrm{tr}}) = \mathrm{im}(A^{\mathrm{tr}} P^{\mathrm{tr}}) \leq \mathrm{im}(A^{\mathrm{tr}})$, that is $\mathrm{im}(\varphi_{PA}^{\mathrm{tr}}) \leq \mathrm{im}(\varphi_A^{\mathrm{tr}})$. Hence for $P \in \mathrm{GL}_m(R)$ we get $\mathrm{im}((P^{-1}PA)^{\mathrm{tr}}) \leq \mathrm{im}((PA)^{\mathrm{tr}}) \leq \mathrm{im}(A^{\mathrm{tr}})$, implying

$\mathrm{im}((PA)^{\mathrm{tr}}) = \mathrm{im}(A^{\mathrm{tr}})$, that is $\mathrm{im}(\varphi^{\mathrm{tr}}_{PA}) = \mathrm{im}(\varphi^{\mathrm{tr}}_{A})$, in other words the row R-modules of PA and A coincide.

If $R = K$ is a field, then $\mathrm{rk}(A^{\mathrm{tr}}) = \dim_K(\mathrm{im}(A^{\mathrm{tr}})) = \dim_K(\mathrm{im}(A^{\mathrm{tr}})^{\mathrm{tr}}) = \dim_K(\mathrm{im}(\varphi^{\mathrm{tr}}_{A})) = \dim_K(\langle v_1, \ldots, v_m\rangle_K) \in \mathbb{N}_0$ is called the **row rank** of A. In particular, for $P \in K^{m\times m}$ we have $\mathrm{rk}((PA)^{\mathrm{tr}}) \leq \mathrm{rk}(A^{\mathrm{tr}})$, where for $P \in \mathrm{GL}_m(K)$ we get equality $\mathrm{rk}((PA)^{\mathrm{tr}}) = \mathrm{rk}(A^{\mathrm{tr}})$.

For example, let K be any field and

$$A := \begin{bmatrix} 1 & 1 & . \\ 1 & . & 1 \\ . & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \in K^{4\times 3}, \quad \text{thus} \quad A^{\mathrm{tr}} = \begin{bmatrix} 1 & 1 & . & 2 \\ 1 & . & 1 & 1 \\ . & 1 & -1 & 1 \end{bmatrix} \in K^{3\times 4}.$$

For the columns $w_1, \ldots, w_3 \in K^{4\times 1}$ of A we have $w_1 = w_2 + w_3$, hence $\{w_2, w_3\} \subseteq K^{4\times 1}$ being K-linearly independent is a K-basis of $\mathrm{im}(A) \leq K^{4\times 1}$; for the rows $v_1, \ldots, v_4 \in K^3$ of A we have $v_3 = v_1 - v_2$ and $v_4 = v_1 + v_2$, hence $\{v_1, v_2\} \subseteq K^3$ being K-linearly independent is a K-basis of $\mathrm{im}(A^{\mathrm{tr}})^{\mathrm{tr}} \leq K^3$.

(5.5) Theorem. If $R = K$ is a field, then we have $\mathrm{rk}(A^{\mathrm{tr}}) = \mathrm{rk}(A)$, that is column rank and row rank of A coincide, just being called the **rank** of A.

Proof. For any matrix $P \in K^{m\times m}$ we have $\mathrm{rk}(PA) = \dim_K(\mathrm{im}(PA)) = \dim_K(\mathrm{im}(\varphi_P\varphi_A)) \leq \dim_K(\mathrm{im}(\varphi_A)) = \dim_K(\mathrm{im}(A)) = \mathrm{rk}(A)$, hence for $P \in \mathrm{GL}_m(K)$ we have $\mathrm{rk}(A) = \mathrm{rk}(P^{-1}PA) \leq \mathrm{rk}(PA) \leq \mathrm{rk}(A)$, implying $\mathrm{rk}(PA) = \mathrm{rk}(A)$; note that we only get a statement on column ranks here, while the associated column spaces in general are not included in each other.

Let $B \subseteq K^{n\times 1}$ and $C \subseteq K^{m\times 1}$ be K-bases such that $D := {}_C(\varphi_A)_B = \mathrm{diag}[1, \ldots, 1, 0, \ldots, 0] \in K^{m\times n}$, with $r := \mathrm{rk}(A) \in \{0, \ldots, \min\{m, n\}\}$ non-zero entries. Let $S \subseteq K^{n\times 1}$ and $T \subseteq K^{m\times 1}$ be the associated standard K-bases, and let $P := {}_T\mathrm{id}_C \in \mathrm{GL}_m(K)$ and $Q := {}_S\mathrm{id}_B \in \mathrm{GL}_n(K)$. Then we have $P^{-1}AQ = {}_C\mathrm{id}_T \cdot {}_T(\varphi_A)_S \cdot {}_S\mathrm{id}_B = {}_C(\varphi_A)_B = D$, thus $\mathrm{rk}(A) = \mathrm{rk}(PAQ^{-1}) = \mathrm{rk}(D) = r = \mathrm{rk}(D^{\mathrm{tr}}) = \mathrm{rk}((PAQ^{-1})^{\mathrm{tr}}) = \mathrm{rk}(Q^{-\mathrm{tr}}A^{\mathrm{tr}}P^{\mathrm{tr}}) = \mathrm{rk}(A^{\mathrm{tr}})$. $\sharp$

(5.6) Row operations. Let $R \neq \{0\}$ be a commutative ring, and let $A \in R^{m\times n}$, where $m, n \in \mathbb{N}_0$, have rows $v_1, \ldots, v_m \in R^n$. The following (elementary) **row operations** on A yield $A' \in R^{m\times n}$ having rows $v'_1, \ldots, v'_m \in R^n$:
i) Multiplying row i with a, that is $v'_i := av_i$ where $i \in \{1, \ldots, m\}$ and $a \in R^*$; hence $A' = E_{i,a}A$ where $E_{i,a} := \mathrm{diag}[1, \ldots, 1, a, 1, \ldots, 1] \in R^{m\times m}$.
ii) Adding the a-fold of row j to row i, that is $v'_i := v_i + av_j$ where $i \neq j \in \{1, \ldots, m\}$ and $a \in R$; hence $A' = E_{i,j,a}A$ where $E_{i,j,a} := E_m + aE_{ij} \in R^{m\times m}$.

The above **elementary** matrices fulfill $E_i(a)E_i(a^{-1}) = E_m = E_i(a^{-1})E_i(a)$ and $E_{i,j}(a)E_{i,j}(-a) = E_m = E_{i,j}(-a)E_{i,j}(a)$, implying that $E_i(a), E_{i,j}(a) \in \mathrm{GL}_m(R)$. Thus, if A is transformed into A' by row operations, then we have

$A' = PA$ for some $P \in \mathrm{GL}_m(R)$, and the row R-modules of A and A' coincide; in particular, if $R = K$ is a field then $\mathrm{rk}(A) = \mathrm{rk}(A')$.

In particular, **interchanging** row i and row j, that is $v_i' := v_j$ and $v_j' := v_i$ where $i \neq j \in \{1, \ldots, m\}$, is a row operation, where $A' = E_{i,j}A$ and $E_{i,j} := E_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} \in R^{m \times m}$: Considering only the indices $i < j$ yields

$$E_{j,-1}E_{i,j,1}E_{j,i,-1}E_{i,j,1} = \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ . & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & . \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ . & 1 \end{bmatrix} = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} = E_{i,j}.$$

Similarly, (elementary) **column operations** are defined; if A is transformed into A', then $A' = AQ$ for some $Q \in \mathrm{GL}_n(R)$, and the column R-modules of A and A' coincide; in particular, if $R = K$ is a field then $\mathrm{rk}(A) = \mathrm{rk}(A')$.

(5.7) Theorem: Gauß algorithm. Let K be a field and let $A \in K^{m \times n}$, where $m, n \in \mathbb{N}_0$. Then using row operations A can be transformed into **Gaussian normal form**, having $r = \mathrm{rk}(A) \in \{0, \ldots, \min\{m, n\}\}$ non-zero rows,

$$A' = \begin{bmatrix} \ldots & 1 & *** & . & *** & . & *** & . & *** & . & *** \\ \ldots & . & \ldots & 1 & *** & . & *** & . & *** & . & *** \\ \ldots & . & \ldots & . & \ldots & 1 & *** & . & *** & . & *** \\ \ldots & . & \ldots & . & \ldots & . & \ldots & 1 & *** & . & *** \\ \vdots & & & & & & & & & & \vdots \\ \ldots & . & \ldots & . & \ldots & . & \ldots & . & \ldots & 1 & *** \\ \ldots & . & \ldots & . & \ldots & . & \ldots & . & \ldots & . & \ldots \\ \vdots & & & & & & & & & & \vdots \end{bmatrix} \in K^{m \times n},$$

where the unit vectors occur in the uniquely determined **pivot columns** $1 \leq j_1 < j_2 < \cdots < j_r \leq n$; hence the Gaussian normal form is uniquely determined.

Proof. We proceed row by row: Looking at row $k \in \mathbb{N}$, by induction we may assume that $k - 1 \in \mathbb{N}_0$ rows have already been processed; let $j_0 := 0$. We consider the **submatrix** $\widetilde{A}$ consisting of rows $[k, \ldots, m]$ and columns $[j_{k-1} + 1, \ldots, n]$ of A. Then if $\widetilde{A} = 0$ we are done. Hence let $\widetilde{A} \neq 0$, and let $j_k \in \{j_{k-1} + 1, \ldots, n\}$ be the first column containing an entry $a := a_{i,j_k} \neq 0$, where $i \in \{k, \ldots, m\}$. Then interchanging rows k and i, and multiplying row k by a^{-1}, yields a matrix such that $a_{k,j_k} = 1$. Adding the $(-a_{i,j_k})$-fold of row k to row i, for all $i \in \{1, \ldots, m\}$, **cleans up** all of column j_k, yielding a matrix such that $a_{i,j_k} = 0$ for all $k \neq i \in \{1, \ldots, m\}$.

Thus we end up with a Gaussian normal form A' of A, having rows $v_1', \ldots, v_m' \in K^n$, where $r \in \{0, \ldots, \min\{m, n\}\}$ is the number of non-zero rows. Letting $a_1, \ldots, a_r \in K$ such that $\sum_{i=1}^{r} a_i v_i' = 0 \in K^n$, considering the pivot columns $[j_1, \ldots, j_r]$ shows that $a_i = 0 \in K$, for all $i \in \{1, \ldots, r\}$, saying that $[v_1', \ldots, v_r']$ is K-linearly independent. Since the row spaces of A and A' coincide, we conclude that $\{v_1', \ldots, v_r'\}$ is a K-basis of the row space of A, hence we have

$r = \mathrm{rk}(A)$. Note that any $v = [b_1, \ldots, b_n] \in \langle v_1', \ldots, v_r' \rangle_K \leq K^n$ decomposes as $v = \sum_{k=1}^{r} b_{j_k} v_k'$, that is the coordinates are just the pivot entries of v.

Let $\widetilde{A} \in K^{m \times n}$ be a Gaussian normal form of A, with non-zero rows $\widetilde{v}_1, \ldots, \widetilde{v}_r \in K^n$ and pivot columns $[\widetilde{j}_1, \ldots, \widetilde{j}_r]$; we may assume that $r \geq 1$. Assume that $[j_1, \ldots, j_r] \neq [\widetilde{j}_1, \ldots, \widetilde{j}_r]$, then there is $k \in \{1, \ldots, r\}$ minimal such that $j_k < \widetilde{j}_k$, and $j_l = \widetilde{j}_l$ for $l \in \{1, \ldots, k-1\}$; thus we have $\langle \widetilde{v}_k, \ldots, \widetilde{v}_r \rangle_K \leq \langle v_{k+1}, \ldots, v_r \rangle_K$, which since $\dim_K(\langle \widetilde{v}_k, \ldots, \widetilde{v}_r \rangle_K) = r - k + 1 > r - k = \dim_K(\langle v_{k+1}, \ldots, v_r \rangle_K)$ is a contradiction. Thus we have $[j_1, \ldots, j_r] = [\widetilde{j}_1, \ldots, \widetilde{j}_r]$, then decomposing $\widetilde{v}_i$ into $\{v_1', \ldots, v_r'\}$ shows that $\widetilde{v}_i = v_i'$, for all $i \in \{1, \ldots, r\}$. ♯

(5.8) Applications of the Gaussian normal form. a) Let K be a field, let $A \in K^{m \times n}$, where $m, n \in \mathbb{N}_0$, let $A' \in K^{m \times n}$ be its Gaussian normal form with pivot columns $1 \leq j_1 < j_2 < \cdots < j_r \leq n$, where $r := \mathrm{rk}(A) \in \mathbb{N}_0$. Let $P \in \mathrm{GL}_m(K)$ such that $PA = A'$.

Then rows $[1, \ldots, r]$ of P contain K-linear combinations of the rows of A yielding the K-basis $\{v_1', \ldots, v_r'\}$ of the row space of A. Since the row kernel of A has K-dimension $m - r$, we conclude that rows $[r+1, \ldots, m]$ of P consist of a K-basis of the row kernel of A.

Letting $w_1, \ldots, w_n \in K^{m \times 1}$ be the columns of A, from $\mathrm{rk}(P \cdot [w_{j_1}, \ldots, w_{j_r}]) = r$ we infer that $\mathrm{rk}([w_{j_1}, \ldots, w_{j_r}]) = r$, thus $\{w_{j_1}, \ldots, w_{j_r}\} \subseteq K^{m \times 1}$ is a K-basis of the image of A; the kernel of A is discussed in (5.9).

b) To find a matrix $P \in \mathrm{GL}_m(K)$ such that $PA = A'$ in the first place, the row operations used in the Gauß algorithm are kept track of by simultaneously applying them to the identity matrix E_m, that is to the **extended matrix** $[A|E_m] \in K^{m \times (n+m)}$, whose columns are the concatenation of the columns of A and those of E_m. Then we end up with the matrix $[A'|P] \in K^{m \times (n+m)}$, displaying the Gaussian normal form A' and a transforming matrix P.

In particular, for $m = n$ we have $A \in \mathrm{GL}_n(K)$, that is $\mathrm{rk}(A) = n$, if and only if $A' = E_n \in K^{n \times n}$. In this case, we have $PA = E_n$, that is $P = A^{-1} \in \mathrm{GL}_n(K)$, and the extended matrix after running the Gauß algorithm becomes $[E_n|A^{-1}] \in K^{n \times 2n}$, hence this is an algorithm for **matrix inversion**. Note that this also shows that any $A \in \mathrm{GL}_n(K)$ is a product of elementary matrices. Here is a couple of examples:

i) Letting $A := \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ we get

$$[A|E_2] \mapsto \left[\begin{array}{cc|cc} 1 & 2 & 1 & . \\ . & -1 & -3 & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & 2 & 1 & . \\ . & 1 & 3 & -1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & . & -5 & 2 \\ . & 1 & 3 & -1 \end{array}\right],$$

implying that $A \in \mathrm{GL}_2(\mathbb{R})$, where $A^{-1} := \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$.

ii) Letting $A := \begin{bmatrix} 1 & -4 & 1 \\ 2 & -3 & 2 \\ 5 & 1 & . \end{bmatrix} \in \mathbb{Q}^{3\times 3}$ we get:

$$\left[\begin{array}{ccc|ccc} 1 & -4 & 1 & 1 & . & . \\ 2 & -3 & 2 & . & 1 & . \\ 5 & 1 & . & . & . & 1 \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & -4 & 1 & 1 & . & . \\ . & 5 & . & -2 & 1 & . \\ . & 21 & -5 & -5 & . & 1 \end{array}\right]$$

$$\mapsto \left[\begin{array}{ccc|ccc} 1 & -4 & 1 & 1 & . & . \\ . & 1 & . & \frac{-2}{5} & \frac{1}{5} & . \\ . & 21 & -5 & -5 & . & 1 \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & . & 1 & \frac{-3}{5} & \frac{4}{5} & . \\ . & 1 & . & \frac{-2}{5} & \frac{1}{5} & . \\ . & . & -5 & \frac{17}{5} & \frac{-21}{5} & 1 \end{array}\right]$$

$$\mapsto \left[\begin{array}{ccc|ccc} 1 & . & 1 & \frac{-3}{5} & \frac{4}{5} & . \\ . & 1 & . & \frac{-2}{5} & \frac{1}{5} & . \\ . & . & 1 & \frac{-17}{25} & \frac{21}{25} & \frac{-1}{5} \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & . & . & \frac{2}{25} & \frac{-1}{25} & \frac{1}{5} \\ . & 1 & . & \frac{-2}{5} & \frac{1}{5} & . \\ . & . & 1 & \frac{-17}{25} & \frac{21}{25} & \frac{-1}{5} \end{array}\right]$$

This shows that $A^{-1} = \begin{bmatrix} \frac{2}{25} & \frac{-1}{25} & \frac{1}{5} \\ \frac{-2}{5} & \frac{1}{5} & . \\ \frac{-17}{25} & \frac{21}{25} & \frac{-1}{5} \end{bmatrix} = \frac{1}{25} \cdot \begin{bmatrix} 2 & -1 & 5 \\ -10 & 5 & . \\ -17 & 21 & -5 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$.

(5.9) Linear equations. a) Let $R \neq \{0\}$ be a commutative ring, let $A = [a_{ij}]_{ij} \in R^{m\times n}$, where $m, n \in \mathbb{N}_0$, and $w := [y_1, \ldots, y_m]^{\text{tr}} \in R^{m\times 1}$. We consider the **system of linear equations** in the **indeterminates** $x_1, \ldots, x_n \in R$ given by $\sum_{j=1}^n a_{ij}x_j = b_i$, for all $i \in \{1, \ldots, m\}$, which if $w = 0$ is called **homogeneous**, otherwise **inhomogeneous**.

Writing $v := [x_1, \ldots, x_n]^{\text{tr}} \in R^{n\times 1}$, the set of **solutions** becomes $\mathcal{L}(A, w) := \{v \in R^{n\times 1}; Av = w\} = \{v \in R^{n\times 1}; \varphi_A(v) = w\}$. Hence we have $\mathcal{L}(A, w) \neq \emptyset$ if and only if $w \in \operatorname{im}(A)$; in this case the system is called **solvable**. In particular, we have $\mathcal{L}(A, w) \neq \emptyset$ for all $w \in R^{m\times 1}$ if and only if $\operatorname{im}(A) = R^{m\times 1}$, that is φ_A is surjective.

For the homogeneous system associated with A we have $\mathcal{L}(A) := \mathcal{L}(A, 0) = \ker(A) \leq R^{n\times 1}$, thus any homogeneous system is solvable, where there is a unique solution if and only if $\ker(A) = \{0\}$, that is φ_A is injective.

Moreover, if $v, v' \in \mathcal{L}(A, w)$ then we have $A(v - v') = w - w = 0$, that is $v - v' \in \ker(A)$; conversely, for all $u \in \ker(A)$ we have $A(v + u) = w + 0 = w$, that is $v + u \in \mathcal{L}(A, w)$. Hence if $\mathcal{L}(A, w) \neq \emptyset$, then we have $\mathcal{L}(A, w) = v_0 + \ker(A) \subseteq R^{n\times 1}$, where $v_0 \in \mathcal{L}(A, w)$ is a fixed **particular solution**; note that whenever $w \neq 0$, then $0 \notin \mathcal{L}(A, w)$, thus the latter is not an R-submodule.

In terms of the homomorphism principle, the above observation reads as follows: Let $\overline{\varphi}_A \colon R^{n\times 1}/\ker(A) \to R^{m\times 1}$ be the R-monomorphism such that $\varphi_A = \overline{\varphi}_A \nu_{\ker(A)}$, where $\nu_{\ker(A)} \colon R^{n\times 1} \to R^{n\times 1}/\ker(A)$ is the natural map. Hence we have $\mathcal{L}(A, w) \neq \emptyset$ if and only if $w \in \operatorname{im}(A) = \operatorname{im}(\varphi_A) = \operatorname{im}(\overline{\varphi}_A)$. In this case there is a unique $v_0 + \ker(A) \in R^{n\times 1}/\ker(A)$ such that $w = \overline{\varphi}_A(v_0 + \ker(A)) = \varphi_A(v_0) = Av_0$, thus we recover the description $\mathcal{L}(A, w) = v_0 + \ker(A) \subseteq R^{n\times 1}$ with respect to the particular solution $v_0 \in R^{n\times 1}$.

Hence solving the system amounts to prove solvability, and in this case to find a particular solution and the solutions $\mathcal{L}(A, 0) = \ker(A)$ of the associated homogeneous system. To decide solvability we consider the **extended matrix** $[A|w] \in R^{m\times(n+1)}$, obtained by concatenating the columns of A with the column w: We have $\mathcal{L}(A, w) \neq \emptyset$, that is $w \in \operatorname{im}(A)$, if and only if $\operatorname{im}([A|w]) = \operatorname{im}(A)$,

b) Let $R = K$ be a field. Hence $\mathcal{L}(A, w) \neq \emptyset$ if and only if $\operatorname{rk}(A) = \operatorname{rk}([A|w])$.

We have $\mathcal{L}(A, w) \neq \emptyset$ for all $w \in K^{m\times 1}$ if and only if $\operatorname{rk}(A) = m$; in this case, since $\operatorname{rk}(A) \leq n$, we necesarily have $m \leq n$. Moreover, if $\mathcal{L}(A, w) \neq \emptyset$ for some $w \in K^{m\times 1}$, then we have $|\mathcal{L}(A, w)| = 1$ if and only if $\dim_K(\ker(A)) = 0$; in this case, since $\dim_K(\ker(A)) = n - \operatorname{rk}(A) \geq n - m$, we necessarily have $m \geq n$.

Thus the case $m = n$ is particularly interesting, where it turns out that the above cases are equivalent: Indeed, for $A \in K^{n\times n}$ the map φ_A is injective, that is $\ker(A) = \{0\}$, if and only if φ_A is surjective, that is $\operatorname{rk}(A) = n$. Hence the following are equivalent:
i) We have $A \in \mathrm{GL}_n(K)$.
ii) For all $w \in K^{n\times 1}$ we have $\mathcal{L}(A, w) \neq \emptyset$.
iii) For all $w \in K^{n\times 1}$ we have $|\mathcal{L}(A, w)| = 1$.
iv) There is $w \in K^{n\times 1}$ such that $|\mathcal{L}(A, w)| = 1$.
v) We have $|\mathcal{L}(A)| = 1$.

c) Let still $R = K$ be a field. Applying the Gauß algorithm to the extended matrix shows that $\operatorname{rk}(A) = \operatorname{rk}([A|w])$ if and only if $n + 1$ is not a pivot column. In this case the solutions $v = [x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathcal{L}(A, w)$ are described as follows:

Let $[A|w]$ have Gaussian normal form $[A'|w'] \in K^{m\times(n+1)}$ with pivot columns $1 \leq j_1 < \cdots < j_r \leq n$, where $r := \operatorname{rk}(A) \in \mathbb{N}_0$. Let $P \in \mathrm{GL}_m(K)$ such that $A' = [a'_{ij}] = PA \in K^{m\times n}$, writing $w' = [y'_1, \ldots, y'_m]^{\mathrm{tr}} := Pw \in K^{m\times 1}$ we have $\mathcal{L}(A, w) = \{v \in K^{n\times 1}; Av = w\} = \{v \in K^{n\times 1}; PAv = Pw\} = \mathcal{L}(A', w')$.

The $n - r$ entries $x_j \in K$, where $j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$ are the non-pivot columns, can be chosen arbitrarily, thus are considered as **parameters**. Then the r entries $x_{j_k} \in K$, for the pivot columns $\{j_1, \ldots, j_r\}$, are uniquely determined by $x_{j_k} := b'_k - \sum_{j\in\{j_k+1,\ldots,n\}\setminus\{j_{k+1},\ldots,j_r\}} a'_{kj}x_j$, for $k \in \{1, \ldots, r\}$; the particular solution given by letting $x_j := 0$, for all $j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$, equals $\sum_{k=1}^{r} b'_k e_{j_k} \in K^{n\times 1}$. Finally, considering the associated homgeneous system, a K-basis $\{v_j \in K^{n\times 1}; j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}\} \subseteq \ker(A)$ is found by specifying a K-basis of $\mathrm{Maps}(\{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}, K) \cong K^{n-r}$, thus with respect to the standard K-basis of K^{n-r} we get $v_j := e_j - \sum_{k\in\{1,\ldots,r\},j_k<j} a'_{kj}e_{j_k}$.

For example, the Gauß algorithm yields for $w := [a, b, c]^{\text{tr}} \in \mathbb{Q}^{3 \times 1}$:

$$[A|w] := \left[\begin{array}{ccc|c} 1 & 1 & 5 & a \\ 6 & -1 & 16 & b \\ 3 & 1 & 11 & c \end{array}\right] \qquad \mapsto \left[\begin{array}{ccc|c} 1 & 1 & 5 & a \\ . & -7 & -14 & -6a+b \\ . & -2 & -4 & -3a+c \end{array}\right]$$

$$\mapsto \left[\begin{array}{ccc|c} 1 & 1 & 5 & a \\ . & 1 & 2 & \frac{6a-b}{7} \\ . & -2 & -4 & -3a+c \end{array}\right] \mapsto \left[\begin{array}{ccc|c} 1 & . & 3 & \frac{a+b}{7} \\ . & 1 & 2 & \frac{6a-b}{7} \\ . & . & . & \frac{-9a-2b+7c}{7} \end{array}\right]$$

Hence the system $A \cdot [x_1, x_2, x_3]^{\text{tr}} = [a, b, c]^{\text{tr}}$ is solvable if and only if $9a + 2b - 7c = 0$. For the associated homogeneous system we get $\mathcal{L}(A) = \langle [-3, -2, 1]^{\text{tr}} \rangle_{\mathbb{Q}}$, hence if $c = \frac{9a+2b}{7}$ we have $\mathcal{L}(A, [a, b, c]^{\text{tr}}) = [\frac{a+b}{7}, \frac{6a-b}{7}, 0]^{\text{tr}} + \langle [-3, -2, 1]^{\text{tr}} \rangle_{\mathbb{Q}}$; for example, we have $\mathcal{L}(A, [3, 11, 7]^{\text{tr}}) = \{[2 - 3x, 1 - 2x, x]^{\text{tr}}; x \in \mathbb{Q}\}$.

(5.10) Example: Geometric interpretation. For $K := \mathbb{R}$ and $n \in \{2, 3\}$, we discuss $\mathcal{L} := \mathcal{L}(A, w)$, where $A \in \mathbb{R}^{m \times n}$ and $w \in \mathbb{R}^{m \times 1}$, in dependence of $r := \text{rk}(A) \in \{0, \ldots, n\}$. We may assume that the system $[A|w] \in \mathbb{R}^{m \times (n+1)}$ is in Gaussian normal form, hence we may also assume that $m = r + 1$.

If $r = 0$, then we get $\left[\begin{array}{c|c} . & \epsilon \end{array}\right]$, where $\epsilon \in \{0, 1\}$, hence we have $\mathcal{L} \neq \emptyset$ if and only if $\epsilon = 0$, and in this case we have $\mathcal{L} = \ker(A) = \mathbb{R}^n$. If $r = n$, then $\mathcal{L}$ has at most one element. The intermediate cases are more interesting:

a) If $n = 2$ and $r = 1$, then for some $s, a \in \mathbb{R}$ and $\epsilon \in \{0, 1\}$ we get

$$\left[\begin{array}{cc|c} 1 & s & a \\ . & . & \epsilon \end{array}\right] \quad \text{or} \quad \left[\begin{array}{cc|c} . & 1 & a \\ . & . & \epsilon \end{array}\right].$$

Hence we have $\mathcal{L} \neq \emptyset$ if and only if $\epsilon = 0$. In this case, in the first case we have $\mathcal{L} = \{[a - sy, y]^{\text{tr}} \in \mathbb{R}^{2 \times 1}; y \in \mathbb{R}\}$, a line in $\mathbb{R}^{2 \times 1}$; it intersects the first coordinate axis in $[a, 0]^{\text{tr}}$, and the second coordinate axis in $[0, \frac{a}{s}]^{\text{tr}}$ if $s \neq 0$, while it is parallel to it if $s = 0$, where it coincides with it if and only if $a = 0$. In the second case we have $\mathcal{L} = \{[x, a]^{\text{tr}} \in \mathbb{R}^{2 \times 1}; x \in \mathbb{R}\}$, a line in $\mathbb{R}^{2 \times 1}$; it intersects the second coordinate axis in $[0, a]^{\text{tr}}$, and is parallel to the first coordinate axis, where it coincides with it if and only if $a = 0$.

b) If $n = 3$ and $r = 2$, then for some $s, t, a, b \in \mathbb{R}$ and $\epsilon \in \{0, 1\}$ we get

$$\left[\begin{array}{ccc|c} 1 & . & s & a \\ . & 1 & t & b \\ . & . & . & \epsilon \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} 1 & s & . & a \\ . & . & 1 & b \\ . & . & . & \epsilon \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} . & 1 & . & a \\ . & . & 1 & b \\ . & . & . & \epsilon \end{array}\right].$$

Hence we have $\mathcal{L} \neq \emptyset$ if and only if $\epsilon = 0$. In this case, we have $\mathcal{L} = \{[a - sz, b - tz, z]^{\text{tr}} \in \mathbb{R}^{3 \times 1}; z \in \mathbb{R}\}$, and $\mathcal{L} = \{[a - sy, y, b]^{\text{tr}} \in \mathbb{R}^{3 \times 1}; y \in \mathbb{R}\}$, and $\mathcal{L} = \{[x, a, b]^{\text{tr}} \in \mathbb{R}^{3 \times 1}; x \in \mathbb{R}\}$, respectively, lines in $\mathbb{R}^{3 \times 1}$.

c) If $n = 3$ and $r = 1$, then for some $s, t, a \in \mathbb{R}$ and $\epsilon \in \{0, 1\}$ we get

$$\left[\begin{array}{ccc|c} 1 & s & t & a \\ . & . & . & \epsilon \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} . & 1 & s & a \\ . & . & . & \epsilon \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} . & . & 1 & a \\ . & . & . & \epsilon \end{array}\right].$$

Hence we have $\mathcal{L} \neq \emptyset$ if and only if $\epsilon = 0$. In this case, we have $\mathcal{L} = \{[a - sy - tz, y, z]^{\text{tr}} \in \mathbb{R}^{3\times 1}; y, z \in \mathbb{R}\}$, and $\mathcal{L} = \{[x, a - zs, z]^{\text{tr}} \in \mathbb{R}^{3\times 1}; x, z \in \mathbb{R}\}$, and $\mathcal{L} = \{[x, y, a]^{\text{tr}} \in \mathbb{R}^{3\times 1}; x, y \in \mathbb{R}\}$, respectively, hyperplanes in $\mathbb{R}^{3\times 1}$.

(5.11) Example: Block designs. A 2**-design** is a set $P = [p_1, \ldots, p_v]$ of $v \geq 2$ **points**, together with $b \in \mathbb{N}$ **blocks** $\emptyset \neq B_1, \ldots, B_b \subseteq P$ each containing $k \leq v$ points, such that any point is an element of precisely $r \leq b$ blocks, and any pair of distinct points is contained in precisely $\lambda \leq r$ blocks. If $\lambda = r$, then we may assume that $\{p_1, p_2\} \subseteq B_1, \ldots, B_\lambda$, implying that $p_1 \notin B_{\lambda+1}, \ldots, B_b$, hence $\{p_1, p_i\} \subseteq B_1$ for all $i \in \{2, \ldots, v\}$, showing that $P \subseteq B_1$, and thus we have the **complete** design $P = B_j$, for all $j \in \{1, \ldots, b\}$, hence $v = k$ and $\lambda = r = b$; more interesting are the **incomplete** designs fulfilling $\lambda < r$.

Let $A = [a_{ij}]_{ij} \in \mathbb{Q}^{v\times b}$ be the associated **incidence matrix** defined by $a_{ij} := 1$ if $p_i \in B_j$, and $a_{ij} := 0$ if $p_i \notin B_j$, for $i \in \{1, \ldots, v\}$ and $j \in \{1, \ldots, b\}$. Then **double counting**, taking column sums and row sums first, respectively, yields $bk = [k, \ldots, k] \cdot [1, \ldots, 1]^{\text{tr}} = ([1, \ldots, 1] \cdot A) \cdot [1, \ldots, 1]^{\text{tr}} = [1, \ldots, 1] \cdot (A \cdot [1, \ldots, 1]^{\text{tr}}) = [1, \ldots, 1] \cdot [r, \ldots, r]^{\text{tr}} = vr$. Moreover, we have

$$AA^{\text{tr}} = \begin{bmatrix} r & \lambda & \lambda & \ldots & \lambda & \lambda \\ \lambda & r & \lambda & \ldots & \lambda & \lambda \\ \lambda & \lambda & r & \ldots & \lambda & \lambda \\ \vdots & & & & & \vdots \\ \lambda & \lambda & \lambda & \ldots & r & \lambda \\ \lambda & \lambda & \lambda & \ldots & \lambda & r \end{bmatrix} \in \mathbb{Q}^{v\times v},$$

hence we have $v(r + \lambda(v-1)) = [1, \ldots, 1] \cdot AA^{\text{tr}} \cdot [1, \ldots, 1]^{\text{tr}} = ([1, \ldots, 1] \cdot A) \cdot ([1, \ldots, 1] \cdot A)^{\text{tr}} = [k, \ldots, k] \cdot [k, \ldots, k]^{\text{tr}} = bk^2 = vrk$, thus $\lambda(v-1) = r(k-1)$.

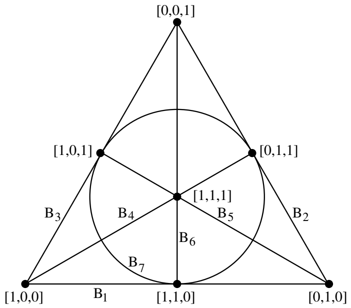
If $\lambda < r$, then we have $\text{rk}(AA^{\text{tr}}) = v$, implying **Fisher's inequality [1935]** $v = \text{rk}(AA^{\text{tr}}) \leq \min\{\text{rk}(A), \text{rk}(A^{\text{tr}})\} \leq \min\{v, b\} \leq b$: Starting off with $AA^{\text{tr}} \in \mathbb{Q}^{v\times v}$, subtracting row 1 from row i, for all $i \in \{2, \ldots, v\}$, yields

$$\begin{bmatrix} r & \lambda & \lambda & \ldots & \lambda & \lambda \\ \lambda - r & r - \lambda & . & \ldots & . & . \\ \lambda - r & . & r - \lambda & \ldots & . & . \\ \vdots & & & & & \vdots \\ \lambda - r & . & . & \ldots & r - \lambda & . \\ \lambda - r & . & . & \ldots & . & r - \lambda \end{bmatrix} \in \mathbb{Q}^{v\times v},$$

and adding columns $2, \ldots, v$ to column 1 yields

$$\begin{bmatrix} r + (v-1)\lambda & \lambda & \lambda & \ldots & \lambda & \lambda \\ . & r - \lambda & . & \ldots & . & . \\ . & . & r - \lambda & \ldots & . & . \\ \vdots & & & & & \vdots \\ . & . & . & \ldots & r - \lambda & . \\ . & . & . & \ldots & . & r - \lambda \end{bmatrix} \in \mathbb{Q}^{v\times v}.$$

Table 3: The Fano plane.



$$\begin{bmatrix} . & 1 & 1 & . & . & 1 & . \\ 1 & 1 & . & . & 1 & . & . \\ . & 1 & . & 1 & . & . & 1 \\ 1 & . & 1 & 1 & . & . & . \\ . & . & 1 & . & 1 & . & 1 \\ 1 & . & . & . & . & 1 & 1 \\ . & . & . & 1 & 1 & 1 & . \end{bmatrix} \in \mathbb{Q}^{7\times 7}$$

a) Design of experiments. For example, in a wine contest, v brands of wine have to be judged by b referees, such that any brand is tasted by precisely r referees, and any pair of distinct brands is tasted by precisely λ referees.

b) The Fano plane. Let $K := \mathbb{Z}_2$ and $V := K^3$, let P be the set of 1-dimensional K-subspaces of V, and let B be the set of 2-dimensional K-subspaces of V; see Table 3, where also the incidence matrix with respect to the lexicographical ordering of P is indicated. Hence we have $v = 7$ and $b = 7$, as well as $k = 3$, and $r = 3$, and $\lambda = 1$; the Fano plane is a **symmetric** design inasmuch $b = v$.

(5.12) Example: Ternary codes. Identifying the possible outcomes 'draw', 'home team wins', and 'guest team wins' of a soccer match with the elements of the finite field $\mathbb{Z}_3 = \{0, 1, 2\}$, the combined outcome of 11 matches is an element of the $\mathbb{Z}_3$-vector space $\mathbb{Z}_3^{11}$. The task is to bet on the outcome of these matches, and the more guesses are correct the higher the reward is; see (0.6). To launch a systematic attack, we proceed as follows:

Let F be a finite field and $n \in \mathbb{N}_0$. For $v = [a_1, \ldots, a_n] \in F^n$ and $w = [b_1, \ldots, b_n] \in F^n$ let $d(v, w) := |\{i \in \{1, \ldots, n\}; a_i \neq b_i\}| \in \mathbb{N}_0$ be the **Hamming distance** of v and w. This defines a **metric** on F^n, that is we have **positive definiteness** $d(v, w) \in \mathbb{R}_{\geq 0}$ where $d(v, w) = 0$ if and only if $v = w$, **symmetry** $d(v, w) = d(w, v)$, and the **triangle inequality** $d(u, v) + d(v, w) \geq d(u, w)$, for all $u, v, w \in F^n$; moreover we have $d(v, w) = d(v - w, 0)$. Let $B_r(v) := \{w \in F^n; d(v, w) \leq 2\}$ be the **sphere** with **centre** $v \in F^n$ and **radius** $r \in \mathbb{N}_0$. Thus we have $|B_r(v)| = \sum_{i=0}^{r} |\{w \in F^n; d(v, w) = i\}| = \sum_{i=0}^{r} \binom{n}{i} \cdot (|F| - 1)^i$. The aim is to cover F^n **perfectly**, that is completely by pairwise disjoint spheres of a fixed radius. Thus for a perfect cover to exist we necessary have $|B_r(v)| \mid |F^n| = |F|^n$.

For the given parameters $F := \mathbb{Z}_3$ and $n := 11$ we get the following cardinalities:

r	0	1	2	3	4	5	6
$\lvert B_r(v)\rvert$	1	23	**243**	1563	6843	21627	51195

r	7	8	9	10	11
$\vert B_r(v)\vert$	93435	135675	163835	175099	177147

Hence precisely for $r \in \{0, 2, 11\}$ we have $|B_r(v)| \mid 3^{11} = 177147$, in particular for $e := 2$ we have $|B_2(v)| = 243 = 3^5$. Thus we look for $\mathcal{C} \subseteq \mathbb{Z}_3^{11}$ such that $|\mathcal{C}| = \frac{3^{11}}{3^5} = 3^6 = 729$ and $\mathbb{Z}_3^{11} = \coprod_{v\in\mathcal{C}} B_2(v)$, that is any element of $\mathbb{Z}_3^{11}$ coincides in at least $n - e = 9$ positions with some and hence unique element of $\mathcal{C}$, or equivalently $\mathcal{C}$ has **minimum distance** $d := 2e + 1 = 5$, that is any two distinct elements of $\mathcal{C}$ differ in at least d positions. At best, we would like to choose $\mathcal{C}$ as a $\mathbb{Z}_3$-subspace of $\mathbb{Z}_3^{11}$, then necessarily $k := \dim_{\mathbb{Z}_3}(\mathcal{C}) = 6$.

We give a $\mathbb{Z}_3$-basis of the **ternary Golay code [1949]** $\mathcal{C} \leq \mathbb{Z}_3^{11}$ in terms of the **generator matrix** $G \in \mathbb{Z}_3^{6\times 11}$ shown in Table 4. We have to show that $\mathcal{C}$ has minimum distance $d = 5$: Since $\mathcal{C} \leq \mathbb{Z}_3^{11}$ is a $\mathbb{Z}_3$-subspace, we may to the contrary assume that there is $0 \neq v = [a_1, \ldots, a_{11}] \in \mathcal{C}$ such that $d(v, 0) \leq 4$, hence there is $\mathcal{I} \subseteq \{1, \ldots, 11\}$ such that $1 \leq |\mathcal{I}| \leq 4$ and $a_j = 0$ for all $j \notin \mathcal{I}$. We have $\mathrm{rk}(G) = 6$, hence $\dim_{\mathbb{Z}_3}(\ker(G)) = 5$, and a $\mathbb{Z}_3$-basis of $\ker(G) \leq \mathbb{Z}_3^{11}$ is given in terms of the **parity check matrix** $H \in \mathbb{Z}_3^{11\times 5}$ also shown in Table 4; thus we have $G \cdot H = 0 \in \mathbb{Z}_3^{6\times 5}$. Now $v \in \mathcal{C}$, the row $\mathbb{Z}_3$-space of A, implies that the rows $\mathcal{I}$ of H are $\mathbb{Z}_3$-linearly dependent. But by inspection we find that any 4 rows of H actually are $\mathbb{Z}_3$-linearly independent, a contradiction.

Similarly, for $n' := 13$ we get

r	0	1	2	3	4	5	6	7
$\vert B_r(v)\vert$	1	**27**	339	2627	14067	55251	165075	384723

r	8	9	10	11	12	13
$\vert B_r(v)\vert$	714195	1080275	1373139	1532883	1586131	1594323

Hence precisely for $r \in \{0, 1, 13\}$ we have $|B_r(v)| \mid 3^{13} = 1594323$, in particular for $e' := 1$ we have $|B_2(v)| = 27 = 3^3$. Thus we look for $\mathcal{C}' \leq \mathbb{Z}_3^{13}$ such that $k' := \dim_{\mathbb{Z}_3}(\mathcal{C}') = n' - 3 = 10$, that is $|\mathcal{C}'| = 3^{10} = 59049$, and minimum distance $d' := 2e' + 1 = 3$: We give a $\mathbb{Z}_3$-basis of a **ternary Hamming code [1950]** $\mathcal{C}' \leq \mathbb{Z}_3^{13}$ in terms of the generator matrix $G' \in \mathbb{Z}_3^{10\times 13}$ shown in Table 5; considering any choice of 2 rows of the parity check matrix $H' \in \mathbb{Z}_3^{13\times 3}$ given there as well shows that indeed $d' = 3$.

6 Determinants

(6.1) Oriented volumes. a) Let R be a ring, V be an R-module, and $n \in \mathbb{N}_0$. A map $\delta \colon V^n \to R$ is called a **determinant form** or **oriented volume** of **degree** n, if it is R-**multilinear**, that is for all $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$, where $i \in \{1, \ldots, n\}$, the map $V \to R \colon v \mapsto \delta(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$ is R-linear, and **alternating**, that is for $i < j$ and $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in V$ and $v \in V$ we have $\delta(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n) = 0$.

The set of determinant forms is an R-submodule of $\mathrm{Maps}(V^n, R)$. For $i \neq j \in \{1, \ldots, n\}$ and $v_1, \ldots, v_n \in V$ and $a \in R$ we have $\delta(\ldots, v_i + av_j, \ldots, v_j, \ldots) =$

Table 4: Generator and parity check matrices of the ternary Golay code.

$$G := \begin{bmatrix} 2 & . & 1 & 2 & 1 & 1 & . & . & . & . & . \\ . & 2 & . & 1 & 2 & 1 & 1 & . & . & . & . \\ . & . & 2 & . & 1 & 2 & 1 & 1 & . & . & . \\ . & . & . & 2 & . & 1 & 2 & 1 & 1 & . & . \\ . & . & . & . & 2 & . & 1 & 2 & 1 & 1 & . \\ . & . & . & . & . & 2 & . & 1 & 2 & 1 & 1 \end{bmatrix} \qquad H := \begin{bmatrix} 1 & . & . & . & . \\ 2 & 1 & . & . & . \\ 2 & 2 & 1 & . & . \\ 2 & 2 & 2 & 1 & . \\ 1 & 2 & 2 & 2 & 1 \\ . & 1 & 2 & 2 & 2 \\ 1 & . & 1 & 2 & 2 \\ . & 1 & . & 1 & 2 \\ . & . & 1 & . & 1 \\ . & . & . & 1 & . \\ . & . & . & . & 1 \end{bmatrix}$$

Table 5: Generator and parity check matrices of a ternary Hamming code.

$$G' := \begin{bmatrix} 2 & 2 & 1 & . & . & . & . & . & . & . & . & . & . \\ 1 & 2 & . & 1 & . & . & . & . & . & . & . & . & . \\ 2 & . & . & . & 2 & 1 & . & . & . & . & . & . & . \\ 1 & . & . & . & 2 & . & 1 & . & . & . & . & . & . \\ . & 2 & . & . & 2 & . & . & 1 & . & . & . & . & . \\ 2 & 2 & . & . & 2 & . & . & . & 1 & . & . & . & . \\ 1 & 2 & . & . & 2 & . & . & . & . & 1 & . & . & . \\ . & 1 & . & . & 2 & . & . & . & . & . & 1 & . & . \\ 2 & 1 & . & . & 2 & . & . & . & . & . & . & 1 & . \\ 1 & 1 & . & . & 2 & . & . & . & . & . & . & . & 1 \end{bmatrix} \qquad H' := \begin{bmatrix} . & 2 & 1 \\ 2 & . & 1 \\ 2 & 2 & 2 \\ 2 & 1 & . \\ 2 & 2 & . \\ 2 & 1 & 1 \\ 2 & . & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & . & . \\ . & 2 & 2 \\ . & 1 & . \\ . & . & 1 \end{bmatrix}$$

$\delta(\ldots, v_i, \ldots, v_j, \ldots) + a\delta(\ldots, v_j, \ldots, v_j, \ldots) = \delta(v_1, \ldots, v_n) \in R$. Moreover, for $i < j \in \{1, \ldots, n\}$ and $v_1, \ldots, v_n \in V$ we have $0 = \delta(\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) = \delta(\ldots, v_i, \ldots, v_i, \ldots) + \delta(\ldots, v_i, \ldots, v_j, \ldots) + \delta(\ldots, v_j, \ldots, v_i, \ldots) + \delta(\ldots, v_j, \ldots, v_j \ldots) = \delta(\ldots, v_i, \ldots, v_j, \ldots) + \delta(\ldots, v_j, \ldots, v_i, \ldots)$, which implies $\delta(\ldots, v_i, \ldots, v_j, \ldots) = -\delta(\ldots, v_j, \ldots, v_i, \ldots) \in R$; in other words, swapping two of the arguments of δ results in a sign change.

Actually, the latter condition is slightly weaker than the property of being alternating: Indeed, from $\delta(\ldots, v_i, \ldots, v_j, \ldots) + \delta(\ldots, v_j, \ldots, v_i, \ldots) = 0$, for $i < j \in \{1, \ldots, n\}$ and $v_1, \ldots, v_n \in V$, we get $2\delta(\ldots, v, \ldots, v, \ldots) = 0$ for $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in V$ and $v \in V$, which only under additional assumptions implies that δ is alternating, for example if $R = K$ is a field such that $2 \neq 0 \in K$.

b) We have the following geometric interpretation: Given $v_1, \ldots, v_n \in \mathbb{R}^n$ we consider the **parallelotope** $\{\sum_{i=1}^n a_i v_i \in \mathbb{R}^n; 0 \leq a_i \leq 1 \text{ for all } i \in \{1, \ldots, n\}\}$. The aim is to associate an **oriented volume** to this parallelotope, having the following properties: In either argument, the volume is additive and proportional with respect to positive scalars; reverting an argument, or exchanging two of them by **handedness** negates the volume; shearing does not change the volume; if the parallelotope is contained in a hyperplane, the volume vanishes; and as normalisation condition the **unit cube** $\{\sum_{i=1}^n a_i e_i \in \mathbb{R}^n; 0 \leq a_i \leq 1 \text{ for all } i \in \{1, \ldots, n\}\}$, which is associated with the standard $\mathbb{R}$-basis, has volume 1.

For $n = 2$, identifying $\mathbb{R}^{2 \times 1}$ with $\mathbb{C}$, let $z := x + iy = r\exp(i\varphi) \in \mathbb{C}$ and $z' := x' + iy' = r'\exp(i\varphi') \in \mathbb{C}$, where $x, y, r, \varphi, x', y', r', \varphi' \in \mathbb{R}$. Then the area of the parallelotope associated with $[z, z']$, which can be identified with $\begin{bmatrix} x & x' \\ y & y' \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, is given by $rr'\sin(\varphi' - \varphi) = rr' \cdot \frac{r}{r'} \cdot \operatorname{im}(\frac{z'}{z}) = \frac{x^2+y^2}{2i}(\frac{x'+iy'}{x+iy} - \frac{x'-iy'}{x-iy}) = \frac{x^2+y^2}{2i} \cdot \frac{2i(xy'-yx')}{x^2+y^2} = xy' - yx' \in \mathbb{R}$.

(6.2) The sign map. For $n \in \mathbb{N}_0$ let $\mathcal{S}_n$ be the symmetric group on the set $\{1, \ldots, n\}$. Any permutation $\pi \in \mathcal{S}_n$ can be written as the tuple $\pi = [\pi(1), \ldots, \pi(n)]$, or alternatively in cycle notation.

a) Let $\tau_{ij} := [1, \ldots, i-1, j, i+1, \ldots, j-1, i, j+1, \ldots, n] \in \mathcal{S}_n$ be a **transposition**, where $i < j \in \{1, \ldots, n\}$, and let $\tau_i := \tau_{i,i+1} = [1, \ldots, i-1, i+1, i, i+2, \ldots, n]$ be an **adjacent transposition**. We show that any $\pi \in \mathcal{S}_n$ can be written as a (possibly empty) product of adjacent transpositions:

We proceed by induction on $n \in \mathbb{N}_0$, the cases $n \in \{0, 1\}$ being trivial, we may assume $n \geq 2$: Let $m := \pi(n) \in \{1, \ldots, n\}$ and $\pi' := \tau_{n-1} \cdots \tau_m \pi \in \mathcal{S}_n$; then we have $\pi'(n) = n$, thus we get $\pi'|_{\{1,\ldots,n-1\}} \in \mathcal{S}_{n-1}$, which by induction is a product of adjacent transpositions $\tau_i \in \mathcal{S}_{n-1} \subseteq \mathcal{S}_n$, where $i \in \{1, \ldots, n-2\}$. $\sharp$

Letting $l(\pi) := |\{\{i, j\}; i < j, \pi(i) > \pi(j)\}| \in \mathbb{N}_0$ be the **inversion number** of $\pi \in \mathcal{S}_n$, we infer that $l(\pi)$ coincides with the minimum number of factors needed to write π as a product of adjacent transpositions. Hence we have $l(\pi) = l(\pi^{-1})$, as well as $l(\mathrm{id}) = 0$ and $l(\tau_{ij}) = 2(j-i) - 1$, in particular $l(\tau_i) = 1$.

Let the **sign** map $\mathrm{sgn}\colon \mathcal{S}_n \to \mathbb{Q}$ be defined by $\mathrm{sgn}(\pi) := \prod_{1\leq i<j\leq n} \frac{\pi(j)-\pi(i)}{j-i}$.

Since π induces a bijection on the set of all 2-element subsets of $\{1,\ldots,n\}$, with $\{i,j\}$ also $\{\pi(i),\pi(j)\}$ runs through these subsets. Thus this implies $\prod_{1\leq i<j\leq n} |\pi(j)-\pi(i)| = \prod_{1\leq i<j\leq n}(j-i)$, which equals $\prod_{k\in\{1,\ldots,n-1\}}(n-k)!$. Hence we actually have $\mathrm{sgn}\colon \mathcal{S}_n \to \{\pm 1\}\colon \pi \mapsto (-1)^{l(\pi)}$. For example, we have $\mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})$, as well as $\mathrm{sgn}(\mathrm{id}) = 1$ and $\mathrm{sgn}(\tau_i) = \mathrm{sgn}(\tau_{ij}) = -1$; in particular, we infer $\mathrm{im}(\mathrm{sgn}) = \{\pm 1\}$ for $n \geq 2$.

For $\pi,\rho \in \mathcal{S}_n$ we have **multiplicativity** $\mathrm{sgn}(\pi\rho) = \mathrm{sgn}(\pi)\cdot\mathrm{sgn}(\rho)$:

We have $\mathrm{sgn}(\pi\rho) = \prod_{1\leq i<j\leq n} \frac{\pi\rho(j)-\pi\rho(i)}{j-i} = \prod_{1\leq i<j\leq n}(\frac{\pi\rho(j)-\pi\rho(i)}{\rho(j)-\rho(i)}\cdot\frac{\rho(j)-\rho(i)}{j-i}) = (\prod_{1\leq i<j\leq n} \frac{\pi(\rho(j))-\pi(\rho(i))}{\rho(j)-\rho(i)})\cdot(\prod_{1\leq i<j\leq n}\frac{\rho(j)-\rho(i)}{j-i})$. Since $\{\rho(i),\rho(j)\}$ runs through the 2-element subsets of $\{1,\ldots,n\}$ if $\{i,j\}$ does so, from this we get $\mathrm{sgn}(\pi\rho) = (\prod_{1\leq i<j\leq n}\frac{\pi(j)-\pi(i)}{j-i})\cdot(\prod_{1\leq i<j\leq n}\frac{\rho(j)-\rho(i)}{j-i}) = \mathrm{sgn}(\pi)\cdot\mathrm{sgn}(\rho)$.

Alternatively, we show that $l(\pi\rho) \equiv l(\pi)+l(\rho) \pmod 2$: We may assume that $\rho = \tau_i$ for some $i \in \{1,\ldots,n-1\}$; hence $l(\tau_i) = 1$. Then we have $\pi\tau_i = [\pi(1),\ldots,\pi(i-1),\pi(i+1),\pi(i),\pi(i+2),\ldots,\pi(n)] \in \mathcal{S}_n$ showing that $l(\pi\tau_i) = l(\pi)+1$ if $\pi(i) < \pi(i+1)$, and $l(\pi\tau_i) = l(\pi)-1$ if $\pi(i) > \pi(i+1)$. ♯

b) An alternative approach is as follows: Recall that any permutation $\pi \in \mathcal{S}_n$ can be uniquely written as a product of $r \in \mathbb{N}_0$ disjoint cycles, where the order of the factors does not matter. To show that π can be written as a product of not necessarily disjoint transpositions, that is cycles of length 2, we may assume that π is a cycle of length $k \geq 2$, and even that $\pi = (1,\ldots,k)$. Then we have $\pi = (1,\ldots,k) = (1,2)(2,3)\cdots(k-1,k) = (k-1,k)(k-2,k)\cdots(1,k) \in \mathcal{S}_n$. Here the order of the transpositions does matter, and the factorisation in general is not unique, not even the number of transpositions occurring is; for example, we have $(1,2,3) = (1,2)(2,3) = (2,3)(1,3) = (1,2)(1,3)(2,3)(1,2) \in \mathcal{S}_3$.

Let $\pi = \sigma_1\cdots\sigma_s \in \mathcal{S}_n$ with transpositions $\sigma_i \in \mathcal{S}_n$, where $s \in \mathbb{N}_0$, be a product of $r \in \mathbb{N}_0$ disjoint cycles. Then we have $s \equiv n-r \pmod 2$:

We proceed by induction on $s \in \mathbb{N}_0$, where for $s = 0$ we have $\pi = ()$ and $r = n$. Hence let $s \geq 1$, let $\sigma_1 = (i,j) \in \mathcal{S}_n$ and $\pi' := \sigma_2\cdots\sigma_s \in \mathcal{S}_n$. If π' is a product of $r' \in \mathbb{N}_0$ disjoint cycles, we by induction have $s-1 \equiv n-r' \pmod 2$. If i and j are in the same cycle of π', then we have $\pi = \sigma_1\pi' = (i,j)(i,\ldots,k,j,l,\ldots,m)(\ldots)\cdots(\ldots) = (i,\ldots,k)(j,l,\ldots,m)(\ldots)\cdots(\ldots)$, thus π is a product of $r = r'+1$ disjoint cycles, hence $n-r \equiv n-r'-1 \equiv s-2 \equiv s \pmod 2$. If i and j are in different cycles of π', then $\pi = \sigma_1\pi' = (i,j)(i\ldots k)(j,\ldots,l)(\ldots)\cdots(\ldots) = (i,\ldots,k,j,\ldots,l)(\ldots)\cdots(\ldots)$, thus π is a product of $r = r'-1$ disjoint cycles, hence $n-r \equiv n-r'+1 \equiv s \pmod 2$. ♯

Thus the map $\mathcal{S}_n \to \{\pm 1\}\colon \pi = \sigma_1\cdots\sigma_s \mapsto (-1)^s = (-1)^{n-r}$ is well-defined, inasmuch the right hand side does not depend on the particular factorisation of π as a product of transpositions, and using shortest possible factorisations into adjacent transpositions shows that it coincides with the sign map defined above. Moreover, from the new description we independently infer that sgn is surjective for $n \geq 2$, that $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$, and that sgn is multiplicative.

c) The elements of $\mathcal{A}_n := \{\pi \in \mathcal{S}_n; \operatorname{sgn}(\pi) = 1\}$ and $\mathcal{S}_n \backslash \mathcal{A}_n = \{\pi \in \mathcal{S}_n; \operatorname{sgn}(\pi) = -1\}$ are called **even** and **odd** permutations, respectively. For example, a cycle of length $k \in \mathbb{N}$ is even if and only if k is odd, in particular transpositions are odd. Moreover, from multiplicativity we infer that where $\mathcal{A}_n \leq \mathcal{S}_n$ is a subgroup, being the **alternating group**, and we have $\mathcal{S}_n = \mathcal{A}_n \,\dot{\cup}\, \mathcal{A}_n\pi$, whenever $n \geq 2$ and $\pi \in \mathcal{S}_n \setminus \mathcal{A}_n$, for example if π is a transposition.

This applies to determinant forms as follows: Let R be a ring, V an R-module, and $\delta\colon V^n \to R$ a determinant form of degree n, writing $\pi \in \mathcal{S}_n$ as a product of transpositions yields $\delta(v_{\pi^{-1}(1)}, \ldots, v_{\pi^{-1}(n)}) = \operatorname{sgn}(\pi) \cdot \delta(v_1, \ldots, v_n) \in R$, that is applying $\pi \in \mathcal{S}_n$ to a tuple in V^n changes the value of δ by a factor $\operatorname{sgn}(\pi)$.

(6.3) Determinants. a) Let $R \neq \{0\}$ be a commutative ring. Then the **determinant** of a square matrix $A := [a_{ij}]_{ij} \in R^{n \times n}$, where $n \in \mathbb{N}_0$, is defined as $\det(A) := \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi) \cdot \prod_{j=1}^n a_{\pi(j),j} \in R$.

For example, for an **upper triangular** matrix $A \in R^{n \times n}$, that is $a_{ij} = 0$ for all $i > j \in \{1, \ldots, n\}$, all summands in the defining sum of $\det(A)$ vanish except for $\pi = \mathrm{id} \in \mathcal{S}_n$, yielding $\det(A) = \prod_{j=1}^n a_{jj} \in R$; in particular, for the identity matrix $E_n \in R^{n \times n}$ we get $\det(E_n) = 1$.

For $n = 0$ we have $\mathcal{S}_0 = \{[]\} = \{()\}$, and hence $\det([]) = 1$; for $n = 1$ we have $\mathcal{S}_1 = \{[1]\} = \{()\}$, and hence $\det([a]) = a$; for $n = 2$ we have $\mathcal{S}_2 = \{[1,2],[2,1]\} = \{(),(1,2)\}$, and hence $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} \in R$. For $n = 3$ we have $\mathcal{S}_3 = \{[1,2,3],[2,3,1],[3,1,2];[3,2,1],[2,1,3],[1,3,2] = \{(),(1,2,3),(1,3,2),(1,3),(1,2),(2,3)\}$, where $\mathcal{A}_3 = \{1,2,3],[2,3,1],[3,1,2]\} = \{(),(1,2,3),(1,3,2)\}$, thus $\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32}) \in R$, called the **Sarrus rule**.

b) We show that $\det\colon (R^{n \times 1})^n \to R\colon [v_1, \ldots, v_n] \mapsto \det([v_1, \ldots, v_n])$ actually is a determinant form of degree n:

We may assume that $n \geq 2$. Since each summand in the defining sum of det is R-multilinear, the map det is as well. In order to show that det is alternating, we consider the transposition $\tau_{ij} \in \mathcal{S}_n \setminus \mathcal{A}_n$, where $i < j \in \{1, \ldots, n\}$. Writing the columns of A as $v_j = [a_{1j}, \ldots, a_{nj}]^{\mathrm{tr}} \in R^{n \times 1}$, for all $j \in \{1, \ldots, n\}$, we get $\det([v_1, \ldots, v_n]) = \sum_{\pi \in \mathcal{A}_n} (\prod_{k=1}^n a_{\pi(k),k} - \prod_{k=1}^n a_{\pi\tau(k),k})$ $= \sum_{\pi \in \mathcal{A}_n} (\prod_{k \neq i,j} a_{\pi(k),k})(a_{\pi(i),i}a_{\pi(j),j} - a_{\pi(j),i}a_{\pi(i),j})$. Now, if $v_i = v_j$ then we have $a_{\pi(i),i} = a_{\pi(i),j}$ and $a_{\pi(j),i} = a_{\pi(j),j}$, thus $\det([v_1, \ldots, v_n]) = 0$. $\sharp$

c) From $A^{\mathrm{tr}} = [a_{ji}]_{ij} \in K^{n \times n}$, using $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$ for all $\pi \in \mathcal{S}_n$, we get $\det(A^{\mathrm{tr}}) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)} = \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi^{-1}) \cdot \prod_{j=1}^n a_{\pi^{-1}(j),j} = \det(A)$. Hence det is row R-multilinear and row alternating as well.

Thus, performing row operations on $A \in R^{n \times n}$ yielding $A' \in R^{n \times n}$, changes

$\det(A) \in R$ into $\det(A') \in R$ as follows: Multiplying row i with a scalar $a \in R$, where $i \in \{1, \dots, n\}$, yields $\det(A') = a \cdot \det(A)$; adding a multiple of row j to row i, where $j \neq i \in \{1, \dots, n\}$, yields $\det(A') = \det(A)$; interchanging row i and row j yields $\det(A') = -\det(A)$.

Hence this allows to compute the determinant of A by applying the Gauß algorithm, keeping track of the row operations made, and to read off the determinant of its Gaussian normal form A', which is an upper triangular matrix.

(6.4) Theorem. Let $R \neq \{0\}$ be a commutative ring, let V be an R-module having an R-basis C such that $n := |C| \in \mathbb{N}_0$.

Then $\det_C \colon V^n \to R \colon [v_1, \dots, v_n] \mapsto \det([{}_Cv_1, \dots, {}_Cv_n])$ is a determinant form of degree n, such that $\det_C(C) = 1$. Moreover, if $\delta \colon V^n \to R$ is any determinant form of degree n, then we have $\delta = \delta(C) \cdot \det_C \in \mathrm{Maps}(V^n, R)$, that is $\{\det_C\}$ generates the R-submodule of determinant forms of degree n.

Proof. Since $V \to R^{n\times 1} \colon v \mapsto {}_Cv$ is R-linear and $\det$ is a determinant form, $\det_C$ is as well, and we have $\det_C(C) = \det(E_n) = 1$.

For $v_1, \dots, v_n \in V$, letting $[{}_Cv_1, \dots, {}_Cv_n] = [a_{ij}]_{ij} \in R^{n\times n}$, we have $v_j = \sum_{i=1}^n a_{ij} w_i$, for all $j \in \{1, \dots, n\}$, where $C = [w_1, \dots, w_n]$. By R-multilinearity this yields $\delta(v_1, \dots, v_n) = \sum_{[i_1,\dots,i_n]\in\{1,\dots,n\}^n} (\prod_{j=1}^n a_{i_j,j}) \delta(w_{i_1}, \dots, w_{i_n})$. The summands for tuples $[i_1, \dots, i_n]$ having two identical entries vanish, hence we obtain $\delta(v_1, \dots, v_n) = \sum_{\pi\in\mathcal{S}_n} (\prod_{j=1}^n a_{\pi(j),j}) \delta(w_{\pi(1)}, \dots, w_{\pi(n)})$. Since applying π to the arguments of δ changes the value of δ by a factor $\mathrm{sgn}(\pi)$, we have $\delta(w_{\pi(1)}, \dots, w_{\pi(n)}) = \mathrm{sgn}(\pi^{-1}) \cdot \delta(w_1, \dots, w_n)$. Hence we get $\delta(v_1, \dots, v_n) = (\sum_{\pi\in\mathcal{S}_n} \mathrm{sgn}(\pi) \cdot \prod_{j=1}^n a_{\pi(j),j}) \cdot \delta(w_1, \dots, w_n) = \det([{}_Cv_1, \dots, {}_Cv_n]) \cdot \delta(C)$. $\sharp$

For example, let $V := \mathbb{R}^{2\times 1}$ with standard $\mathbb{R}$-basis B and $\mathbb{R}$-basis C given by ${}_B\mathrm{id}_C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$; hence ${}_C\mathrm{id}_B = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$. Then for $[v_1, v_2] \in (\mathbb{R}^{2\times 1})^2$ given by $[{}_Bv_1, {}_Bv_2] = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$, with respect to the standard $\mathbb{R}$-basis we get $\det_B([v_1, v_2]) = \det([{}_Bv_1, {}_Bv_2]) = ad - bc \in \mathbb{R}$, while using $[{}_Cv_1, {}_Cv_2] = {}_C\mathrm{id}_B \cdot [{}_Bv_1, {}_Bv_2] = \frac{1}{2} \cdot \begin{bmatrix} a+b & c+d \\ b-a & d-c \end{bmatrix} \in \mathbb{R}^{2\times 2}$, with respect to the $\mathbb{R}$-basis C we similarly get $\det_C([v_1, v_2]) = \det([{}_Cv_1, {}_Cv_2]) = \frac{1}{4} \cdot ((a+b)(d-c) - (c+d)(b-a)) = \frac{1}{2} \cdot (ad - bc) = \frac{1}{2} \cdot \det_B([v_1, v_2]) \in \mathbb{R}$, where indeed we have $\det_C(B) = \det({}_C\mathrm{id}_B) = \frac{1}{2}$; note that multiplicativity to be shown below also yields $\det([{}_Cv_1, {}_Cv_2]) = \det({}_C\mathrm{id}_B \cdot [{}_Bv_1, {}_Bv_2]) = \det({}_C\mathrm{id}_B) \cdot \det([{}_Bv_1, {}_Bv_2])$.

(6.5) Corollary. a) For all $A, B \in R^{n\times n}$ we have **multiplicativity** $\det(AB) = \det(A) \cdot \det(B)$. Hence if $A \in \mathrm{GL}_n(R)$ then we have $\det(A^{-1}) = \det(A)^{-1} \in R^*$,

Moreover, $\mathrm{SL}_n(R) := \{A \in \mathrm{GL}_n(R); \det(A) = 1\} \leq \mathrm{GL}_n(R)$ is a subgroup, being called the **special linear group** of degree n over R

b) If $\varphi \in \mathrm{End}_R(V)$, then $\det(\varphi) := \det({}_C\varphi_C) \in R$ is independent from the R-basis chosen, and called the **determinant of** φ.

In particular, in view of the geometric interpretation of determinants for $R = K = \mathbb{R}$ being the real number field, the determinant of $\varphi \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$, computed using the standard $\mathbb{R}$-basis of $\mathbb{R}^n$, is the oriented volume of the image of the unit cube under φ, thus describes the change in volume application of φ entails.

Proof. a) Letting $S \subseteq R^{n\times 1}$ be the standard R-basis, we have $\det = \det_S \in \mathrm{Maps}((R^{n\times 1})^n, R)$. Since $R^{n\times 1} \to R^{n\times 1} \colon v \mapsto Av$ is R-linear and det is a determinant form of degree n, we conclude that $\delta_A \colon (R^{n\times 1})^n \to R \colon [v_1, \ldots, v_n] \mapsto \det([Av_1, \ldots, Av_n])$ is a determinant form of degree n as well. Thus from $\delta_A(S) = \delta_A(E_n) = \det(A)$ we get $\delta_A = \delta_A(S) \cdot \det_S = \det(A) \cdot \det$. Hence, letting $v_1, \ldots, v_n \in R^{n\times 1}$ be the columns of B, this yields $\delta_A(B) = \det(A) \cdot \det([v_1, \ldots, v_n]) = \det(A) \cdot \det(B)$, while directly from the definition we get $\delta_A(B) = \delta_A([v_1, \ldots, v_n]) = \det([Av_1, \ldots, Av_n]) = \det(AB)$.

In particular, for $A \in \mathrm{GL}_n(R)$ we have $\det(A) \cdot \det(A^{-1}) = \det(AA^{-1}) = \det(E_n) = 1$, entailing $\det(A^{-1}) = \det(A)^{-1} \in R^*$.

b) If $D \subseteq V$ is a R-basis, then we have ${}_D\mathrm{id}_C = ({}_C\mathrm{id}_D)^{-1} \in \mathrm{GL}_n(R)$, thus we get $\det({}_D\varphi_D) = \det({}_D\mathrm{id}_C \cdot {}_C\varphi_C \cdot {}_C\mathrm{id}_D) = \det({}_C\mathrm{id}_D)^{-1} \cdot \det({}_C\varphi_C) \cdot \det({}_C\mathrm{id}_D) = \det({}_C\varphi_C) \in R$. ♯

For example, let still $V := \mathbb{R}^{2\times 1}$ with standard $\mathbb{R}$-basis B and $\mathbb{R}$-basis C given by ${}_B\mathrm{id}_C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$. Then for the reflection σ at the hyperplane perpendicular to $[-1, 1]^{\mathrm{tr}}$ we get $\det(\sigma) = \det({}_B\sigma_B) = \det\left(\begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix}\right) = -1$, or alternatively $\det(\sigma) = \det({}_C\sigma_C) = \left(\begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix}\right) = -1$; for the rotation ρ with angle $\omega \in \mathbb{R}$ we get $\det(\rho) = \det({}_B\rho_B) = \det\left(\begin{bmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{bmatrix}\right) = \cos^2(\omega) + \sin^2(\omega) = 1$; and for the rotation-dilatation τ with angle $\frac{\pi}{4}$ and scaling factor $\sqrt{2}$ we get $\det(\tau) = \det({}_B\tau_B) = \det\left(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right) = 2$. Indeed, reflections do not change the absolute value of the volume but invert the orientation, rotations leave the oriented volume invariant, while dilatations change the absolute value of the volume but keep the orientation.

(6.6) Theorem: Laplace expansion. Let $R \neq \{0\}$ be a commutative ring, let $A = [a_{ij}]_{ij} \in R^{n \times n}$ where $n \in \mathbb{N}$, and for $i, j \in \{1, \ldots, n\}$ let

$$A_{ij} := \begin{bmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix} \in R^{(n-1)\times(n-1)}$$

be the matrix obtained from A by deleting row i and column j, where $\det(A_{ij}) \in R$ is called the (i, j)-th $(n-1)$**-minor** of A.

Then we have the **column expansion** formula $\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$, for all $j \in \{1, \ldots, n\}$, as well as the **row expansion** formula $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$, for all $i \in \{1, \ldots, n\}$.

Proof. Letting $w_1, \ldots, w_n \in R^{n \times 1}$ be the columns of A, for $j \in \{1, \ldots, n\}$, and $w \in R^{n \times 1}$, let $A_j(w) := [w_1, \ldots, w_{j-1}, w, w_{j+1}, \ldots, w_n] \in R^{n \times n}$ be the matrix obtained from A by replacing column j by w. Now we consider $A_j(e_i) \in R^{n \times n}$, where $e_i \in R^{n \times 1}$ is the i-th unit vector, for $i, j \in \{1, \ldots, n\}$:

We apply the permutation $\alpha := [1, \ldots, i-1, n, i, i+1, \ldots, n-2, n-1] = (n, n-1, \ldots, i) \in \mathcal{S}_n$ to the rows, and the permutation $\beta := [1, \ldots, j-1, n, j, j+1, \ldots, n-2, n-1] = (n, n-1, \ldots, j) \in \mathcal{S}_n$ to the columns of $A_j(e_i)$, which yields $\left[\begin{array}{ccccc|c} \multicolumn{5}{c|}{A_{ij}} & . \\ \hline a_{i1} & \ldots & a_{i,j-1} & a_{i,j+1} & \ldots & a_{in} & 1 \end{array}\right] \in R^{n \times n}$. All summands in the defining sum of the determinant of the latter matrix vanish, except for $\pi \in \mathcal{S}_n$ such that $\pi(n) = n$, hence the remaining sum runs over $\mathcal{S}_{n-1} \subseteq \mathcal{S}_n$, thus $\operatorname{sgn}(\alpha) = (-1)^{n-i}$ and $\operatorname{sgn}(\beta) = (-1)^{n-j}$ yield $\det(A_j(e_i)) = (-1)^{i+j} \cdot \det(A_{ij})$.

Thus $\det(A) = \det(A_j(w_j)) = \det(A_j(\sum_{i=1}^{n} a_{ij} e_i)) = \sum_{i=1}^{n} a_{ij} \cdot \det(A_j(e_i)) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$, as well as $\det(A) = \det(A^{\mathrm{tr}}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ji} \cdot \det((A^{\mathrm{tr}})_{ij}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ji} \cdot \det((A_{ji})^{\mathrm{tr}}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ji} \cdot \det(A_{ji})$. ♯

(6.7) Corollary. a) Let $\operatorname{adj}(A) := [(-1)^{i+j} \cdot \det(A_{ji})]_{ij} \in R^{n \times n}$ be the **adjoint matrix** of A. Then we have $A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot E_n \in R^{n \times n}$.

Hence we have $A \in \mathrm{GL}_n(R)$ if and only if $\det(A) \in R^*$, and in this case $A^{-1} = \det(A)^{-1} \cdot \operatorname{adj}(A) \in \mathrm{GL}_n(R)$; in particular, if $R = K$ is a field then A is invertible if and only if $\det(A) \neq 0$.

b) For $A \in \mathrm{GL}_n(R)$ and $w \in R^{n \times 1}$, the unique solution $v = [x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathcal{L}(A, w)$ of the system of linear equations $Av = w$ is by **Cramer's rule** given as $x_i := \det(A)^{-1} \cdot \det(A_i(w)) \in R$, for all $i \in \{1, \ldots, n\}$, where still $A_i(w) \in R^{n \times n}$ is the matrix obtained from A by replacing column i by w.

Proof. a) For $i, k \in \{1, \ldots, n\}$ we get $(\mathrm{adj}(A) \cdot A)_{ik} = \sum_{j=1}^{n}(-1)^{i+j} \cdot \det(A_{ji}) \cdot a_{jk} = \sum_{j=1}^{n} \det(A_i(e_j)) \cdot a_{jk} = \det(A_i(\sum_{j=1}^{n} a_{jk}e_j)) = \det(A_i(w_k))$, where $w_k \in R^{n\times 1}$ denotes column k of A. Hence for $i \neq k$ columns i and k of $A_i(w_k)$ coincide, implying $\det(A_i(w_k)) = 0$, while for $i = k$ we have $A_i(w_i)) = A$, hence $\det(A_i(w_i)) = \det(A)$. This shows that $\mathrm{adj}(A) \cdot A = \det(A) \cdot E_n$.

Moreover, we have $(A_{ij})^{\mathrm{tr}} = (A^{\mathrm{tr}})_{ji}$, hence $\mathrm{adj}(A)^{\mathrm{tr}} = [(-1)^{i+j} \cdot \det(A_{ij})]_{ij} = [(-1)^{i+j} \cdot \det((A_{ij})^{\mathrm{tr}})]_{ij} = [(-1)^{i+j} \cdot \det((A^{\mathrm{tr}})_{ji})]_{ij} = \mathrm{adj}(A^{\mathrm{tr}})$, from which we infer $(A \cdot \mathrm{adj}(A))^{\mathrm{tr}} = \mathrm{adj}(A)^{\mathrm{tr}} \cdot A^{\mathrm{tr}} = \mathrm{adj}(A^{\mathrm{tr}}) \cdot A^{\mathrm{tr}} = \det(A^{\mathrm{tr}}) \cdot E_n = \det(A) \cdot E_n$, hence $A \cdot \mathrm{adj}(A) = \det(A) \cdot E_n$ as well. Note that we can spare the second half of this argument in case $\det(A) \neq 0$, but a priorily not in case $\det(A) = 0$.

b) Letting $w = [y_1, \ldots, y_n]^{\mathrm{tr}} \in R^{n\times 1}$, from $Av = w$ we get $v = A^{-1}w = \det(A)^{-1} \cdot \mathrm{adj}(A)w$, hence we obtain $\det(A) \cdot x_i = \sum_{j=1}^{n}(-1)^{i+j} \cdot \det(A_{ji}) \cdot y_j = \sum_{j=1}^{n} \det(A_i(e_j)) \cdot y_j = \det(A_i(\sum_{j=1}^{n} y_j e_j)) = \det(A_i(w))$. ♯

Here is a couple of examples, see (5.8):

i) Let $A := \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2}$. We explicitly get $\det(A) = 1 \cdot 5 - 2 \cdot 3 = -1$, expansion with respect to the first row yields $\det(A) = 1 \cdot \det([5]) - 2 \cdot \det([3]) = 1 \cdot 5 - 2 \cdot 3 = -1$, and expansion with respect to the first column yields $\det(A) = 1 \cdot \det([5]) - 3 \cdot \det([2]) = 1 \cdot 5 - 3 \cdot 2 = -1$; in particular we have $A \in \mathrm{GL}_2(\mathbb{R})$. Moreover, we have $\mathrm{adj}(A) = \begin{bmatrix} +(\det([5])) & -(\det([2])) \\ -(\det([3])) & +(\det([1])) \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -3 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, thus we get $A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A) = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$.

Letting $a, b \in \mathbb{R}$, the unique solution $v = [x_1, x_2]^{\mathrm{tr}} \in \mathbb{R}^{2\times 1}$ of $Av = [a, b]^{\mathrm{tr}}$ is given as $v = A^{-1} \cdot [a, b]^{\mathrm{tr}} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \cdot [a, b]^{\mathrm{tr}} = [-5a + 2b, 3a - b]^{\mathrm{tr}}$, which is also obtained from Cramer's rule by $x_1 = \det(A)^{-1} \cdot \det\left(\begin{bmatrix} a & 2 \\ b & 5 \end{bmatrix}\right) = -(5a-2b)$ and $x_2 = \det(A)^{-1} \cdot \det\left(\begin{bmatrix} 1 & a \\ 3 & b \end{bmatrix}\right) = -(b - 3a)$.

ii) Let $A := \begin{bmatrix} 1 & -4 & 1 \\ 2 & -3 & 2 \\ 5 & 1 & . \end{bmatrix} \in \mathbb{Q}^{3\times 3}$. The Sarrus rule yields $\det(A) = (1 \cdot (-3) \cdot 0 + (-4) \cdot 2 \cdot 5 + 1 \cdot 2 \cdot 1) - (1 \cdot (-3) \cdot 5 + (-4) \cdot 2 \cdot 0 + 1 \cdot 2 \cdot 1) = -25$; in particular we have $A \in \mathrm{GL}_3(\mathbb{R})$. Moreover, we have $\mathrm{adj}(A) = \begin{bmatrix} +(-2) & -(-1) & +(-5) \\ -(-10) & +(-5) & -(0) \\ +(17) & -(21) & +(5) \end{bmatrix} \in \mathbb{Q}^{3\times 3}$, and hence we get $A^{-1} = \frac{1}{25} \cdot \begin{bmatrix} 2 & -1 & 5 \\ -10 & 5 & . \\ -17 & 21 & -5 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$.

Letting $a, b, c \in \mathbb{R}$, the unique solution $v = [x_1, x_2, x_3]^{\mathrm{tr}} \in \mathbb{R}^{3\times 1}$ of $Av = [a, b, c]^{\mathrm{tr}}$

is given as $v = A^{-1}\cdot[a,b,c]^{\mathrm{tr}} = \frac{1}{25}\cdot[2a-b+5c, -10a+5b, -17a+21b-5c]^{\mathrm{tr}}$, which is also obtained from Cramer's rule from $\det\left(\begin{bmatrix} a & -4 & 1 \\ b & -3 & 2 \\ c & 1 & . \end{bmatrix}\right) = -2a+b-5c$ and $\det\left(\begin{bmatrix} 1 & a & 1 \\ 2 & b & 2 \\ 5 & c & . \end{bmatrix}\right) = 10a-5b$ and $\det\left(\begin{bmatrix} 1 & -4 & a \\ 2 & -3 & b \\ 5 & 1 & c \end{bmatrix}\right) = 17a-21b+5c$.

(6.8) Example: Vandermonde matrix. Let $R \neq \{0\}$ be a commutative ring. For $a_1, \ldots, a_n \in R$, where $n \in \mathbb{N}_0$, let the **Vandermonde matrix** be

$$A := [a_i^{j-1}]_{ij} = \begin{bmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{bmatrix} \in R^{n\times n}.$$

Then we have $\det(A) = \Delta(a_1, \ldots, a_n) := \prod_{1\leq i<j\leq n}(a_j - a_i) \in R$; in particular, if $R = K$ is a field then we have $A \in \mathrm{GL}_n(K)$ if and only if the $a_1, \ldots, a_n \in K$ are pairwise distinct:

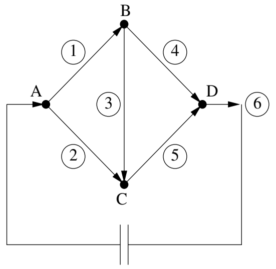
We proceed by induction; see (3.12): The cases $n \leq 1$ are trivial, hence let $n \geq 2$. Adding the $(-a_1)$-fold of column $n-1$ to column n, then adding the $(-a_1)$-fold of column $n-2$ to column $n-1$, and so on, until finally adding the $(-a_1)$-fold of column 1 to column 2, we get

$$\begin{aligned} \Delta(a_1, \ldots, a_n) &= \det \begin{bmatrix} 1 & . & . & \ldots & . \\ 1 & a_2 - a_1 & (a_2-a_1)a_2 & \ldots & (a_2-a_1)a_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_1 & (a_n-a_1)a_n & \ldots & (a_n-a_1)a_n^{n-2} \end{bmatrix} \\ &= \det \begin{bmatrix} a_2 - a_1 & (a_2-a_1)a_2 & \ldots & (a_2-a_1)a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ a_n - a_1 & (a_n-a_1)a_n & \ldots & (a_n-a_1)a_n^{n-2} \end{bmatrix} \\ &= \Delta(a_2, \ldots, a_n) \cdot \prod_{j\in\{2,\ldots,n\}}(a_j - a_1). \end{aligned}$$

$\sharp$

(6.9) Example: Direct current networks. We consider the **Wheatstone bridge** as depicted in Table 6: We have electrical connections between the vertices (A,B), (A,C), (B,C), (B,D), (C,D), and (D,A), whose internal resistances are given as $r := [r_1, \ldots, r_6] \in \mathbb{R}^6$, respectively, where $r_j > 0$. Voltage $v \in \mathbb{R}$ is fed into (D,A), and the task is to determine the currents $c := [c_1, \ldots, c_6]^{\mathrm{tr}} \in \mathbb{R}^{6\times 1}$ in the connections. In particular, we wonder whether it is possible to adjust the internal resistances such that the current c_3 through the bridge (B,C) vanishes.

Table 6: The Wheatstone bridge.



By **Kirchhoff's laws**, incoming and outgoing currents cancel out at each of the vertices A, B, C, D, leading to the first four of the following equations, moreover the voltage between two vertices is given as the product of the internal resistance and the current, and the voltages cancel out along all closed circuits in the network without source or sink, using the circuits (A, B, C) and (B, C, D) this leads to the next two of the following equations, while using the circuit (A, B, D) the last one is due to the voltage v fed into the network. Hence we have the overdetermined system $A'X^{\mathrm{tr}} = w'$, where

$$[A'|w'] := \left[\begin{array}{cccccc|c} -1 & -1 & . & . & . & 1 & . \\ 1 & . & -1 & -1 & . & . & . \\ . & 1 & 1 & . & -1 & . & . \\ . & . & . & 1 & 1 & -1 & . \\ \hline r_1 & -r_2 & r_3 & . & . & . & . \\ . & . & r_3 & -r_4 & r_5 & . & . \\ r_1 & . & . & r_4 & . & r_6 & v \end{array}\right] \in \mathbb{R}^{7\times(6+1)}.$$

Since the currents are accounted for with opposite signs at their respective end vertices, the column sums of the equations coming from the balance of currents all vanish. Thus summing up the first four rows of A' yields a zero row, and we may leave out row 4 and look at the system $AX^{\mathrm{tr}} = w$, where

$$[A|w] := \left[\begin{array}{cccccc|c} -1 & -1 & . & . & . & 1 & . \\ 1 & . & -1 & -1 & . & . & . \\ . & 1 & 1 & . & -1 & . & . \\ r_1 & -r_2 & r_3 & . & . & . & . \\ . & . & r_3 & -r_4 & r_5 & . & . \\ r_1 & . & . & r_4 & . & r_6 & v \end{array}\right] \in \mathbb{R}^{6\times(6+1)}.$$

If Kirchhoff's laws describe direct current networks completely, the above system should have a unique solution. Thus we check that $A \in \mathbb{R}^{6\times 6}$ is invertible:

Adding column 6 to columns 1 and 2, and using row expansion with respect to row 1 we get

$$\det(A) = -\det \begin{bmatrix} 1 & . & -1 & -1 & . \\ . & 1 & 1 & . & -1 \\ r_1 & -r_2 & r_3 & . & . \\ . & . & r_3 & -r_4 & r_5 \\ r_1+r_6 & r_6 & . & r_4 & . \end{bmatrix}.$$

Adding the r_5-fold of row 2 to row 4, and using column expansion with respect to column 5; and adding column 1 to columns 3 and 4, and using row expansion with respect to row 1, the right hand side equals

$$-\det \begin{bmatrix} 1 & . & -1 & -1 \\ r_1 & -r_2 & r_3 & . \\ . & r_5 & r_3+r_5 & -r_4 \\ r_1+r_6 & r_6 & . & r_4 \end{bmatrix} = -\det \begin{bmatrix} -r_2 & r_1+r_3 & r_1 \\ r_5 & r_3+r_5 & -r_4 \\ r_6 & r_1+r_6 & r_1+r_4+r_6 \end{bmatrix}.$$

The Sarrus rule implies $\det(A) = r_2(r_3+r_5)(r_1+r_4+r_6) + (r_1+r_3)r_4r_6 - r_1r_5(r_1+r_6) + r_1(r_3+r_5)r_6 + (r_1+r_3)r_5(r_1+r_4+r_6) + r_2r_4(r_1+r_6) = r_1r_2r_3 + r_1r_2r_4 + r_1r_2r_5 + r_1r_3r_5 + r_1r_3r_6 + r_1r_4r_5 + r_1r_4r_6 + r_1r_5r_6 + r_2r_3r_4 + r_2r_3r_6 + r_2r_4r_5 + r_2r_4r_6 + r_2r_5r_6 + r_3r_4r_5 + r_3r_4r_6 + r_3r_5r_6 > 0$, where the only summand with negative sign cancels out. Sorting terms in view of the symmetries of the physical system yields $\det(A) = (r_1+r_4)(r_2+r_5)r_3 + (r_1+r_2)(r_4+r_5)r_6 + (r_1+r_2+r_4+r_5)r_3r_6 + r_1r_2r_4r_5(\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_4} + \frac{1}{r_5})$, which is not too enlightening. $\sharp$

Hence we have $A \in \mathrm{GL}_6(\mathbb{R})$, and the system $AX^{\mathrm{tr}} = w$ has a unique solution $c = [c_1, \ldots, c_6]^{\mathrm{tr}} \in \mathbb{R}^{6\times 1}$. By Cramer's rule we have

$$c_3 \cdot \det(A) = \det \begin{bmatrix} -1 & -1 & . & . & . & 1 \\ 1 & . & . & -1 & . & . \\ . & 1 & . & . & -1 & . \\ r_1 & -r_2 & . & . & . & . \\ . & . & . & -r_4 & r_5 & . \\ r_1 & . & v & r_4 & . & r_6 \end{bmatrix}.$$

Using column expansion with respect to column 3, and column expansion with respect to column 5, the right hand side equals

$$-v \cdot \det \begin{bmatrix} -1 & -1 & . & . & 1 \\ 1 & . & -1 & . & . \\ . & 1 & . & -1 & . \\ r_1 & -r_2 & . & . & . \\ . & . & -r_4 & r_5 & . \end{bmatrix} = -v \cdot \det \begin{bmatrix} 1 & . & -1 & . \\ . & 1 & . & -1 \\ r_1 & -r_2 & . & . \\ . & . & -r_4 & r_5 \end{bmatrix}.$$

Adding column 1 to column 3, adding column 2 to column 4, and using row expansion with respect to rows 1 and 2, this in turn equals

$$-v \cdot \det \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ r_1 & -r_2 & r_1 & -r_2 \\ . & . & -r_4 & r_5 \end{bmatrix} = -v \cdot \det \begin{bmatrix} r_1 & -r_2 \\ -r_4 & r_5 \end{bmatrix}.$$

Hence we have $c_3 \cdot \det(A) = v \cdot (r_2 r_4 - r_1 r_5)$. Thus for $v \neq 0$ the current c_3 vanishes if and only if the internal resistances fufill $r_2 r_4 = r_1 r_5$, in other words if and only if we have $\frac{r_1}{r_2} = \frac{r_4}{r_5}$. $\sharp$

The physical interpretation is as follows: The voltage v applied to vertex A is distributed to vertices B and C according to the quotient $\frac{r_1}{r_2}$, similarly the voltage $-v$ applied to vertex D is distributed to vertices B and C according to the quotient $\frac{r_4}{r_5}$. There is no current through the bridge (B, C) if and only if B and C are on the same potential, thus if and only if $\frac{r_1}{r_2} = \frac{r_4}{r_5}$.

II Linear algebra II

7 Domains

(7.1) Ideals. a) Let R and S be rings. A map $\varphi\colon R \to S$ is called a **ring homomorphism**, if $\varphi(1_R) = 1_S$ and $\varphi(a+b) = \varphi(a) + \varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$, for all $a, b \in R$.

From $\varphi(0) = \varphi(0+0) = \varphi(0)+\varphi(0)$ we get $\varphi(0) = 0$. Hence from $\varphi(1)+\varphi(-1) = \varphi(1-1) = \varphi(0) = 0$ we get $\varphi(-1) = -\varphi(1)$, and thus $\varphi(-a) = \varphi((-1)\cdot a) = (-1)\cdot\varphi(a) = -\varphi(a)$, for all $a \in R$. This implies that $\mathrm{im}(\varphi) \leq S$ is an additive subgroup, and being a multiplicative submonoid, $\mathrm{im}(\varphi) \subseteq S$ is a subring. Let $\ker(\varphi) := \varphi^{-1}(\{0\}) = \{a \in R; \varphi(a) = 0\} \subseteq R$ be the **kernel** of φ.

b) Recall that R can be considered as an R-module via left multiplication. An R-submodule $I \leq R$ is called an **ideal** of R, if we have $IR := \{ab \in R; a \in I, b \in R\} \subseteq I$; we write $I \trianglelefteq R$. Equivalently, an additive submonoid $I \subseteq R$ is an ideal if and only if $RIR := \{bac \in R; a \in I, b, c \in R\} \subseteq I$. If R is commutative, then we have $RIR = IR = RI := \{ba \in R; a \in I, b \in R\}$, thus the ideals of R coincide with the R-submodules of R.

For example, $\{0\} \trianglelefteq R$ and $R \trianglelefteq R$ are the **trivial** ideals, and for any ring homomorphism $\varphi\colon R \to S$ we have $\ker(\varphi) \trianglelefteq R$: The latter is an additive submonoid, and for all $a \in \ker(\varphi)$ and $b, c \in R$ we have $\varphi(bac) = \varphi(b)\varphi(a)\varphi(c) = \varphi(b)\cdot 0\cdot\varphi(c) = 0$.

c) If $\mathcal{J}$ is a set and $I_j \trianglelefteq R$ for all $j \in \mathcal{J}$, then we have $\bigcap_{j\in\mathcal{J}} I_j \trianglelefteq R$, the empty intersection being defined as R.

For any subset $M \subseteq R$ let $\langle M\rangle := \{\sum_{i=1}^k b_i a_i c_i \in R; k \in \mathbb{N}_0, a_i \in M, b_i, c_i \in R \text{ for all } i \in \{1, \ldots, k\}\}$; for $k = 0$ the empty sum is defined as $0 \in R$, and for $M = \{a_1, \ldots, a_n\}$ we write $\langle M\rangle = \langle a_1, \ldots, a_n\rangle$. We have $M \subseteq \langle M\rangle \trianglelefteq R$, being called the ideal of R **generated** by M. If $\langle M\rangle = I \trianglelefteq R$, then M is called a **generating set** of I, and if $I \trianglelefteq R$ has a finite generating set then it is called **finitely generated**. For $a \in R$ the ideal $\langle a\rangle = RaR := \{bac; b, c \in R\} \trianglelefteq R$ is called **principal**; for example we have $\langle\emptyset\rangle = \langle 0\rangle = \{0\}$ and $\langle 1\rangle = R$.

We have $\langle M\rangle = \bigcap\{I \trianglelefteq R; M \subseteq I\}$, that is $\langle M\rangle$ is the smallest ideal of R containing $\mathcal{S}$, with respect to the partial order on all ideals of R given by inclusion: Since $M \subseteq \langle M\rangle \trianglelefteq R$ the intersection on the right hand side is well-defined and we have $\bigcap\{I \trianglelefteq R; M \subseteq I\} \leq \langle M\rangle$. Conversely, since for all $a \in M \subseteq I$ and $b, c \in R$ we have $bac \in I$, we also have $\langle M\rangle \leq I$, and hence $\langle M\rangle \leq \bigcap\{I \trianglelefteq R; M \subseteq I\}$.

If $\mathcal{J}$ is a set and $I_j \trianglelefteq R$ for all $j \in \mathcal{J}$, then $\sum_{j\in\mathcal{J}} I_j := \langle\bigcup_{j\in\mathcal{J}} I_j\rangle$ is called the **sum** of the I_j; if $\mathcal{J} = \{1, \ldots, n\}$, for some $n \in \mathbb{N}$, we also write $\sum_{j\in\mathcal{J}} I_j = \sum_{j=1}^n I_j = I_1 + \cdots + I_n$, where for $n = 0$ the empty sum is defined to be $\{0\} \trianglelefteq R$. Thus for subsets $M_j \subseteq V$, where $j \in \mathcal{J}$, we have $\langle\bigcup_{j\in\mathcal{J}} M_j\rangle = \sum_{j\in\mathcal{J}}\langle M_j\rangle$.

d) Let $I \trianglelefteq R$. If $I \cap R^* \neq \emptyset$, then we have $R \subseteq IR \subseteq I \subseteq R$, implying $I = R$. Thus I is proper if and only if $I \cap R^* = \emptyset$; in particular, if K is a skew field then $\{0\} \triangleleft K$ and $K \trianglelefteq K$ are the only ideals of K.

Conversely, any commutative ring $R \neq \{0\}$ having only the trivial ideals $\{0\} \lhd R$ and $R \unlhd R$ is a field: For any $0 \neq a \in R$ we have $R = \langle a \rangle = aR := \{ab \in R; b \in R\}$, hence there is $b \in R$ such that $ab = 1$, thus $a \in R^*$.

For example, for $n \in \mathbb{Z}$ we have $\langle n \rangle = n\mathbb{Z} := \{nk \in \mathbb{Z}; k \in \mathbb{Z}\} = \{x \in \mathbb{Z}; n \mid x\} \unlhd \mathbb{Z}$. In particular, $n\mathbb{Z} \leq \mathbb{Z}$ is a $\mathbb{Z}$-submodule, and we have the quotient $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$, such that the natural map $\nu_n \colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \colon a \mapsto a + n\mathbb{Z}$ is a $\mathbb{Z}$-epimorphism. Actually, ν_n is a ring epimorphism.

(7.2) Theorem. Let R be a ring.
a) Let $I \unlhd R$. Then the set $R/I := \{a + I; a \in R\}$ of cosets of I in R is a ring, called the **quotient ring** or **residue class ring** of R with respect to I, with addition $(a + I) + (b + I) := (a + b) + I$ and multiplication and $(a + I)(b + I) := ab + I$, for all $a, b \in R$, additive and multiplicative neutral element $0+I$ and $1+I$, respectively, the additive inverse of $a+U$ being $(-a)+U$.

The **natural map** $\nu_I \colon R \to R/I \colon a \mapsto a + I$ is a ring epimorphism such that $\ker(\nu_I) = I$; in particular, any subset of R is an ideal if and only if it is the kernel of a suitably chosen ring homomorphism.

b) Let S be a ring, and let $\varphi \colon R \to S$ be a ring homomorphism. Then we have the **homomorphism principle**: The map $\overline{\varphi} \colon R/\ker(\varphi) \to \mathrm{im}(\varphi) \colon a+\ker(\varphi) \mapsto \varphi(a)$ is a ring isomorphism such that $\varphi = \overline{\varphi}\nu_{\ker(\varphi)}$.

Proof. a) We only have to show that multiplication is independent from the choice of representatives of the equivalence classes: For $a' \in a+I$ and $b' \in b+I$ there are $x, y \in I$ such that $a' = a+x$ and $b' = b+y$, thus $a'b'+I = (a+x)(b+y)+I = ab+ay+xb+xy+I = ab+I$. From $\nu_I(ab) = ab+I = (a+I)(b+I) = \nu_I(a)\nu_I(b)$ and $\nu_I(1) = 1 + I$ we get that ν_I is a ring homomorphism.
b) From $\overline{\varphi}(ab + I) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(a + I)\overline{\varphi}(b + I)$ and $\overline{\varphi}(1 + I) = \varphi(1) = 1$ we conclude that $\overline{\varphi}$ is a ring homomorphism. ♯

(7.3) Factorial domains. a) Let R be an integral domain. Then $a \in R$ is called a **divisor** of $b \in R$, and b is called a **multiple** of a, if there is $c \in R$ such that $ac = b$; we write $a \mid b$. Hence we have $a \mid b$ if and only if $bR \subseteq aR \unlhd R$.

Elements $a, b \in R$ are called **associate**, if there is $u \in R^*$ such that $b = au \in R$; we write $a \sim b$. We have $a \sim b$ if and only if $aR = bR \unlhd R$, hence $\sim$ is an equivalence relation on R: From $a \sim b$ we have $a \mid b$ and $b \mid a$, and conversely if $a, b \in R$ such that $a \mid b$ and $b \mid a$, there are $u, v \in R$ such that $b = au$ and $a = bv$, thus $a = auv$, implying $a(1 - uv) = 0$, hence $a = 0$ or $uv = 1$, where in the first case $a = b = 0$, and in the second case $u, v \in R^*$ and $a \sim b$.

Let $\emptyset \neq M \subseteq R$ be a subset. Then $d \in R$ such that $d \mid a$ for all $a \in M$ is called a **common divisor** of M; any $u \in R^*$ always is a common divisor of M. If for all common divisors $c \in R$ of M we have $c \mid d$, then $d \in R$ is called a **greatest common divisor** of M. Let $\gcd(M) \subseteq R$ be the set of all greatest common divisors of M. In general greatest common divisors do not exist; if $\gcd(M) \neq \emptyset$

then it consists of an associate class: If $d, d' \in \gcd(M)$, then $d \mid d'$ and $d' \mid d$, hence $d \sim d'$. For $a \in R$ we have $a \in \gcd(a) = \gcd(0, a)$; elements $a, b \in R$ such that $\gcd(a, b) = R^*$ are called **coprime**. Similarly, we get the notion and basic properties of **lowest common multiples** $\operatorname{lcm}(M) \subseteq R$.

b) Let $0 \neq c \in R \setminus R^*$. Then c is called **irreducible** or **indecomposable**, if $c = ab$ implies $a \in R^*$ or $b \in R^*$ for all $a, b \in R$; otherwise c is called **reducible** or **decomposable**; hence if c is irreducible then all its associates also are. The element c is called a **prime**, if $c \mid ab$ implies $c \mid a$ or $c \mid b$, for all $a, b \in R$; hence if c is a prime then all its associates also are.

Then c is irreducible if and only if $cR \lhd R$ is a maximal proper principal ideal: If c is irreducible and $cR \subseteq aR \lhd R$, where $a \in R$, then we have $c = ab$ for some $b \in R$, and since $a \notin R^*$ we conclude $b \in R^*$, thus $cR = aR$. Conversely, if $cR \lhd R$ fulfils the maximality condition and $c = ab$, where $a \in R \setminus R^*$ and $b \in R$, then $cR \subseteq aR \lhd R$, hence $cR = aR$, implying $c \sim a$ and $b \in R^*$.

If c is a prime, then c is irreducible: Let $c = ab$ for some $a, b \in R$, hence we may assume that $c \mid a$, thus there is $d \in R$ such that $cd = a$, hence $c = cdb$, yielding $c(1 - db) = 0$, implying $b \in R^*$. The converse does not hold, that is an irreducible element in general is not a prime.

c) R is called **factorial** or a **Gaussian domain**, if any element $0 \neq a \in R$ can be written uniquely, up to reordering and taking associates, in the form $a = u \cdot \prod_{i=1}^n p_i \in R$, where the $p_i \in R$ are irreducible, $n \in \mathbb{N}_0$ and $u \in R^*$.

Let $\mathcal{P} \subseteq R$ be a set of representatives of the associate classes of irreducible elements of R; these exist by the Axiom of Choice. If R is factorial, then any $0 \neq a \in R$ has a unique **factorisation** $a = u_a \cdot \prod_{p \in \mathcal{P}} p^{\nu_p(a)}$, where $u_a \in R^*$ and $\nu_p(a) \in \mathbb{N}_0$ is called the associated **multiplicity**; we have $\nu_p(a) = 0$ for almost all $p \in \mathcal{P}$, and $\sum_{p \in \mathcal{P}} \nu_p(a) \in \mathbb{N}_0$ is called the **length** of the factorisation, and a is called **squarefree** if $\nu_p(a) \leq 1$ for all $p \in \mathcal{P}$. For any subset $\emptyset \neq M \subseteq R \setminus \{0\}$ we have $\prod_{p \in \mathcal{P}} p^{\min\{\nu_p(a); a \in M\}} \in \gcd(M)$ and $\prod_{p \in \mathcal{P}} p^{\max\{\nu_p(a); a \in M\}} \in \operatorname{lcm}(M)$.

If R is factorial, then any irreducible element $a \in R$ is a prime: Let $0 \neq b, c \in R$ such that $a \mid bc$. Hence there is $d \in R$ such that $ad = bc = u \cdot \prod_{p \in \mathcal{P}} p^{\nu_p(b) + \nu_p(c)}$, where $u \in R^*$. Since a is irreducible, uniqueness of factorisation implies $a \sim p$ for some $p \in \mathcal{P}$ such that $\nu_p(b) + \nu_p(c) > 0$, hence $a \mid b$ or $a \mid c$.

(7.4) Theorem: Bézout. Let R be a **principal ideal domain**, that is R is an integral domain all of whose ideals are principal.
a) Then R is factorial.
b) Let $\emptyset \neq M \subseteq R$ be a subset. Then for $d \in R$ we have $d \in \gcd(M)$ if and only if $dR = \langle M \rangle \trianglelefteq R$, that is there are $a_1, \ldots, a_n \in M$ and **Bézout coefficients** $c_1, \ldots, c_n \in R$, for some $n \in \mathbb{N}$, such that $d = \sum_{i=1}^n c_i a_i \in R$.

Proof. a) For existence of factorisations, assume to the contrary that there is $0 \neq a_1 \in R \setminus R^*$ which is not a product of irreducible elements. Then we have $a_1 R \lhd R$, and since a_1 is reducible there are $a_1 \not\sim a_2, a_2' \in R \setminus R^*$ such that

$a_1 = a_2 a_2'$. We may assume that a_2 is not a product of irreducible elements, hence $a_1 R \subset a_2 R \lhd R$. Thus by induction we get a strictly increasing chain of ideals $a_1 R \subset a_2 R \subset \cdots \subset a_i R \subset \cdots \lhd R$, where $i \in \mathbb{N}$. Hence there is $a \in R$ such that $\bigcup_{i \in \mathbb{N}} a_i R = aR \unlhd R$. Since there is $j \in \mathbb{N}$ such that $a \in a_j R$, we deduce $a_{j+i} R = aR$ for all $i \in \mathbb{N}_0$, a contradiction.

We show that any irreducible element $a \in R$ is a prime: The ideal $aR \lhd R$ is a maximal proper principal ideal, hence is a maximal proper ideal. Thus $R/aR \neq \{0\}$ has only the trivial ideals, thus is a field, hence is an integral domain. Let $a \mid bc$ where $b, c \in R$, then the natural epimorphism $R \rightarrow R/aR \colon x \mapsto \overline{x}$ yields $\overline{b} \cdot \overline{c} = 0$, implying $\overline{b} = 0$ or $\overline{c} = 0$, that is $a \mid b$ or $a \mid c$.

For uniqueness of factorisations, let $\prod_{i=1}^{n} p_i = u \cdot \prod_{j=1}^{m} q_j \in R$, where $u \in R^*$, the p_i and q_j are irreducible, and $n, m \in \mathbb{N}$ such that $n \leq m$. Since p_n is a prime we may assume that $p_n \mid q_m$, and since q_m is irreducible we have $p_n \sim q_m$. Thus we have $\prod_{i=1}^{n-1} p_i = u' \cdot \prod_{j=1}^{m-1} q_j$, where $u' \in R^*$. Hence by induction we get $1 = \widetilde{u} \cdot \prod_{j=1}^{m-n} q_j$, where $\widetilde{u} \in R^*$, showing that $m = n$.

b) Let $\langle M \rangle = bR$ for some $b \in R$. Then for any $c \in R$, such that $c \mid a$ for all $a \in M$, we have $c \mid b$ as well. Thus from $b \mid a$, for all $a \in M$, we conclude $b \in \gcd(M)$. Thus for $d \in R$ we have $d \in \gcd(M)$ if and only if $d \sim b$, which holds if and only if $dR = bR = \langle M \rangle$. $\sharp$

(7.5) Euclidean domains. a) An integral domain R is called **Euclidean**, if R has a **degree map** $\delta \colon R \setminus \{0\} \rightarrow \mathbb{N}_0$ having the following property: For all $a, b \in R$ such that $b \neq 0$ there are $q, r \in R$, called **quotient** and **remainder** respectively, such that $a = qb + r$ where $r = 0$ or $\delta(r) < \delta(b)$; note that no uniqueness assumption is made here.

If R is Euclidean, then it is a principal ideal domain: Letting $\{0\} \neq I \unlhd R$, we from $\delta(I \setminus \{0\}) \neq \emptyset$ infer that there is $0 \neq b \in I$ such that $\delta(b) \in \delta(I \setminus \{0\}) \subseteq \mathbb{N}_0$ is minimal. Hence, for any $a \in I$ let $q, r \in R$ such that $a = qb + r$ where $r = 0$ or $\delta(r) < \delta(b)$. From $r = a - qb \in I$ we by the minimality of $\delta(b)$ conclude $r = 0$, implying $a \in bR$ and thus $I = bR \unlhd R$.

For example, any field K is Euclidean with respect to the degree map $\delta \colon K^* \rightarrow \mathbb{N}_0 \colon x \mapsto 0$, and $\mathbb{Z}$ is Euclidean with respect to the degree map $\delta \colon \mathbb{Z} \setminus \{0\} \rightarrow \mathbb{N}_0 \colon z \mapsto |z|$. Note that these degree maps additionally fulfil $\delta(a) \leq \delta(ab)$, for all $0 \neq a, b \in R$, and if moreover $b \notin R^*$ then even $\delta(a) < \delta(ab)$; this also holds for the degree maps of Euclidean polynomial rings, see (7.7).

b) If R is Euclidean, given $a, b \in R$, a greatest common divisor $r \in R$ and Bézout coefficients $s, t \in R$ such that $r = sa + tb \in R$ can be computed by the **extended Euclidean algorithm**; leaving out the steps indicated by ∘, needed to compute the $s_i, t_i \in R$, just yields a greatest common divisor:

```
• $r_0 \leftarrow a$, $r_1 \leftarrow b$, $i \leftarrow 1$
∘ $s_0 \leftarrow 1$, $t_0 \leftarrow 0$, $s_1 \leftarrow 0$, $t_1 \leftarrow 1$
• while $r_i \neq 0$ do
```

Table 7: Extended Euclidean algorithm in $\mathbb{Z}$.

i	q_i	r_i	s_i	t_i
0		126	1	0
1	3	35	0	1
2	1	21	1	-3
3	1	14	-1	4
4	2	7	2	-7
5		0	-5	18

```
    • $[q_i, r_{i+1}] \leftarrow$ QuotRem$(r_{i-1}, r_i)$    # quotient and remainder
    # $q_i, r_{i+1} \in R$ such that $r_{i+1} = r_{i-1} - q_i r_i$ where $r_{i+1} = 0$ or $\delta(r_{i+1}) < \delta(r_i)$
    ∘ $s_{i+1} \leftarrow s_{i-1} - q_i s_i$, $t_{i+1} \leftarrow t_{i-1} - q_i t_i$
    • $i \leftarrow i + 1$
• return $[r; s, t] \leftarrow [r_{i-1}; s_{i-1}, t_{i-1}]$
```

Since $\delta(r_i) > \delta(r_{i+1}) \geq 0$ for $i \in \mathbb{N}$, there is $l \in \mathbb{N}_0$ such that $r_l \neq 0$ and $r_{l+1} = 0$, hence the algorithm terminates. We have $r_i = s_i a + t_i b$ for all $i \in \{0, \ldots, l+1\}$, hence $r = r_l = sa + tb$. From $r_{i+1} = r_{i-1} - q_i r_i$, for all $i \in \{1, \ldots, l\}$, we get $r = r_l \in \gcd(r_l, 0) = \gcd(r_l, r_{l+1}) = \gcd(r_i, r_{i+1}) = \gcd(r_0, r_1) = \gcd(a, b)$. ♯

Note that in terms of matrices for all $i \in \{1, \ldots, l\}$ we have

$$\begin{bmatrix} 0 & 1 \\ 1 & -q_i \end{bmatrix} \cdot \begin{bmatrix} r_{i-1} & s_{i-1} & t_{i-1} \\ r_i & s_i & t_i \end{bmatrix} = \begin{bmatrix} r_i & s_i & t_i \\ r_{i+1} & s_{i+1} & t_{i+1} \end{bmatrix} \in R^{2 \times 3},$$

where $\begin{bmatrix} r_0 & s_0 & t_0 \\ r_1 & s_1 & t_1 \end{bmatrix} = \begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \end{bmatrix}$. For example, let $R := \mathbb{Z}$ and $a := 2 \cdot 3^2 \cdot 7 = 126$ and $b := 5 \cdot 7 = 35$, then Table 7 shows that $d := 7 = 2a - 7b \in \gcd(a, b)$.

(7.6) Polynomial rings. a) Let $R \neq \{0\}$ be a commutative ring, and let M be a multiplicative monoid. Then let $R[M] := \{a \in \mathrm{Maps}(M, R); a(x) = 0 \text{ for almost all } x \in M\}$. Hence $a \in R[M]$ can be written essentially uniquely as $\sum_{x \in M} a(x)x = \sum_{x \in M} a_x x$. Considering $M \subseteq R[M]$ via the injective map $M \to R[M] \colon x \mapsto 1 \cdot x$, the set $R[M]$ is an R-module with R-basis M.

We define **convolutional** multiplication on $R[M]$ by letting $(\sum_{x \in M} a_x x) \cdot (\sum_{y \in M} b_y y) := \sum_{x,y \in M} a_x b_y xy$. Then we have $((\sum_{x \in M} a_x x)(\sum_{y \in M} b_y y)) \cdot (\sum_{z \in M} c_z z) = \sum_{x,y,z \in M} a_x b_y c_z (xy)z = \sum_{x,y,z \in M} a_x b_y c_z x(yz) = (\sum_{x \in M} a_x x) \cdot ((\sum_{y \in M} b_y y)(\sum_{z \in M} c_z z))$, showing associativity. Letting $1 := 1_R \cdot 1_M \in R[M]$, we get $1 \cdot (\sum_{x \in M} a_x x) = \sum_{x \in M} a_x x = (\sum_{x \in M} a_x x) \cdot 1$, hence $R[M]$ is a monoid with neutral element 1. We have $((\sum_{x \in M} a_x x) + (\sum_{x \in M} b_x x)) \cdot (\sum_{z \in M} c_z z) = (\sum_{x \in M} (a_x + b_x) x)(\sum_{z \in M} c_z z) = \sum_{x,z \in M} (a_x + b_x) c_z xz = (\sum_{x,z \in M} a_x c_z xz) + (\sum_{x,z \in M} b_x c_z xz) = (\sum_{x \in M} a_x x)(\sum_{z \in M} c_z z) + (\sum_{x \in M} b_x x)(\sum_{z \in M} c_z z)$, similarly $(\sum_{z \in M} c_z z) \cdot ((\sum_{x \in M} a_x x) + (\sum_{x \in M} b_x x) = (\sum_{z \in M} c_z z)(\sum_{x \in M} a_x x) + (\sum_{z \in M} c_z z)(\sum_{x \in M} b_x x)$, showing distributivity.

Thus $R[M]$ is a ring, called the **monoid ring** of M over R, which is commutative if and only if M is a commutative monoid. Moreover, $M \subseteq R[M]$ is a submonoid, and we may consider $R \subseteq R[M]$ as a subring via the injective ring homomorphism $R \to R[M] \colon r \mapsto r \cdot 1$.

b) Letting X be a **symbol** or **indeterminate**, let $\mathcal{W}(X) := \{X^i; i \in \mathbb{N}_0\}$ be the set of **words** in X, where commutative multiplication is given by **concatenation**, that is we may identify the additive monoid $\mathbb{N}_0$ with $\mathcal{W}(X)$ via $\mathbb{N}_0 \to \mathcal{W}(X) \colon i \mapsto X^i$. The associated monoid ring $R[X] := R[\mathcal{W}(X)]$ is called the **(univariate) polynomial ring** in X over R.

The elements $f = \sum_{i \geq 0} a_i X^i \in R[X]$ are called **polynomials**, where $a_i \in R$ is called the i-th **coefficient** of f. If $f \neq 0$ then $\deg(f) := \max\{i \in \mathbb{N}_0; a_i \neq 0\} \in \mathbb{N}_0$ is called its **degree**, where polynomials of degree $0, \ldots, 3$ are called **constant**, **linear**, **quadratic**, and **cubic**, respectively, and $\mathrm{lc}(f) := a_{\deg(f)} \in R$ is called its **leading coefficient**; if $\mathrm{lc}(f) = 1$ then f is called **monic**.

For $g = \sum_{j \geq 0} b_j X^j \in R[X]$ we have $fg = \sum_{k \geq 0} (\sum_{l=0}^{k} a_l b_{k-l}) X^k \in R[X]$, Hence for the degree map $\deg \colon R[X] \setminus \{0\} \to \mathbb{N}_0$ we have either $fg = 0$ or $\deg(fg) \leq \deg(f) + \deg(g)$. If $\mathrm{lc}(f)\mathrm{lc}(g) \neq 0$, then we have $fg \neq 0$ where $\mathrm{lc}(fg) = \mathrm{lc}(f)\mathrm{lc}(g)$ and $\deg(fg) = \deg(f) + \deg(g)$.

Hence for any $0 \neq f \in R[X]$ such that $\mathrm{lc}(f) \in R$ is not a zero-divisor, $f \in R[X]$ is not a zero-divisor either. Thus, since $R \subseteq R[X]$ is a subring, $R[X]$ is an integral domain if and only if R is; in this case we have $R[X]^* = R^*$, and $R(X) := \mathrm{Q}(R[X])$ is called the **field of (univariate) rational functions** in X over R, where from $R \subseteq R[X] \subseteq \mathrm{Q}(R)[X] \subseteq R(X)$ we get $R(X) = \mathrm{Q}(R)(X)$.

(7.7) Theorem: Polynomial division. Let $R \neq \{0\}$ be a commutative ring, let $f \in R[X]$ and let $0 \neq g \in R[X]$ such that $\mathrm{lc}(g) \in R^*$. Then there are uniquely determined $q, r \in R[X]$, called **quotient** and **remainder**, respectively, such that $f = qg + r$ where $r = 0$ or $\deg(r) < \deg(g)$.

Proof. Let $qg + r = f = q'g + r'$ where $q, q', r, r' \in R[X]$ such that $r = 0$ or $\deg(r) < \deg(g)$, and $r' = 0$ or $\deg(r') < \deg(g)$. Then we have $(q-q')g = r'-r$, where $r' - r = 0$ or $\deg(r' - r) < \deg(g)$, and where $(q - q')g = 0$ or since $\mathrm{lc}(g) \in R^*$ we have $\deg((q - q')g) = \deg(g) + \deg(q - q') \geq \deg(g)$. Hence we have $r' = r$ and $(q - q')g = 0$, implying $q = q'$, showing uniqueness.

To show existence, we may assume that $f \neq 0$ and $m := \deg(f) \geq \deg(g) := n$. We proceed by induction on $m \in \mathbb{N}_0$: Letting $f' := f - \mathrm{lc}(f)\mathrm{lc}(g)^{-1} g X^{m-n} \in R[X]$, the m-th coefficient of f' shows that $f' = 0$ or $\deg(f') < m$. By induction there are $q', r' \in R[X]$ such that $f' = q'g + r'$, where $r' = 0$ or $\deg(r') < \deg(g)$, hence $f = (q'g + r') + \mathrm{lc}(f)\mathrm{lc}(g)^{-1} g X^{m-n} = (q' + \mathrm{lc}(f)\mathrm{lc}(g)^{-1} X^{m-n})g + r'$. $\sharp$

(7.8) Corollary. If K is a field then $K[X]$ is Euclidean with respect to the degree map deg. Hence any $0 \neq f \in K[X]$ can be written uniquely as $f = \mathrm{lc}(f) \cdot \prod_{p \in \mathcal{P}} p^{\nu_p(f)}$, where $\nu_p(f) \in \mathbb{N}_0$ and $\mathcal{P} \subseteq K[X]$ is the set of irreducible

Table 8: Extended Euclidean algorithm in $\mathbb{Q}[X]$.

i	q_i	r_i	s_i	t_i
0		$X^5 - X^3 + 2X^2 - 2$	1	0
1	$X^2 - 2X + 1$	$X^3 + 2X^2 + 2X + 1$	0	1
2	$\frac{1}{3}(X+2)$	$3X^2 - 3$	1	$-X^2 + 2X - 1$
3	$X - 1$	$3X + 3$	$\frac{-1}{3}(X+2)$	$\frac{1}{3}(X^3 - 3X + 5)$
4		0	$\frac{1}{3}(X^2 + X + 1)$	$\frac{-1}{3}(X^4 - X^3 + 2X - 2)$

monic polynomials, being a set of representatives of the associate classes of irreducible polynomials; we have $\deg(f) = \sum_{p \in \mathcal{P}} \nu_p(f) \deg(p) \in \mathbb{N}_0$.

For example, for $f := (X^3+2)(X+1)(X-1) = X^5-X^3+2X^2-2 \in \mathbb{Z}[X] \subseteq \mathbb{Q}[X]$ and $g := (X^2+X+1)(X+1) = X^3+2X^2+2X+1 \in \mathbb{Z}[X] \subseteq \mathbb{Q}[X]$ we get $f = qg + r$ where $q := X^2 - 2X + 1 \in \mathbb{Z}[X]$ and $r := 3X^2 - 3 \in \mathbb{Z}[X]$. Moreover, Table 8 shows that $d := X + 1 \in \gcd(f, g) \subseteq \mathbb{Q}[X]$ and $d = \frac{-1}{9}(X+2) \cdot f + \frac{-1}{9}(X^4 - X^3 + 2X - 2) \cdot g \in \mathbb{Q}[X]$.

(7.9) Evaluation. a) Let $R \neq \{0\}$ be a commutative ring, let $\varphi\colon R \to S$ be a ring homomorphism into a ring S such that $\varphi(a)z = z\varphi(a)$, for all $a \in R$ and $z \in S$. For $z \in S$ we have the **evaluation map** $\varphi_z\colon R[X] \to S\colon f = \sum_{i\geq 0} a_i X^i \mapsto \sum_{i \geq 0} \varphi(a_i) z^i =: f_\varphi(z)$. We have $\varphi_z(f+g) = \varphi_z(\sum_{i\geq 0}(a_i + b_i)X^i) = \sum_{i \geq 0} \varphi(a_i + b_i) z^i = \sum_{i\geq 0} \varphi(a_i) z^i + \sum_{i \geq 0} \varphi(b_i) z^i = \varphi_z(f) + \varphi_z(g)$ and $\varphi_z(fg) = \varphi_z(\sum_{i \geq 0}(\sum_{j=0}^i a_j b_{i-j})X^i) = \sum_{i\geq 0}(\sum_{j=0}^i \varphi(a_j)\varphi(b_{i-j}))z^i = (\sum_{i \geq 0} \varphi(a_i) z^i) \cdot (\sum_{i\geq 0} \varphi(b_i) z^i) = \varphi_z(f)\varphi_z(g)$ and $\varphi_z(1) = \varphi(1) = 1$, for $g = \sum_{i \geq 0} b_i X^i \in R[X]$, thus φ_z is a ring homomorphism.

If $S \neq \{0\}$ is commutative, $S \subseteq S[X]$ yields the evaluation homomorphism $\varphi_X\colon R[X] \to S[X]\colon \sum_{i \geq 0} a_i X^i \mapsto \sum_{i \geq 0} \varphi(a_i) X^i$, which is injective, respectively surjective, if and only if $\varphi\colon R \to S$ is injective, respectively surjective.

b) For $f \in R[X]$ we get the **polynomial map** $\widehat{f}_\varphi\colon S \to S\colon z \mapsto f_\varphi(z)$. The set $\mathrm{Maps}(S, S)$ is a ring with pointwise addition $f+g\colon S \to S\colon z \mapsto f(z)+g(z)$ and multiplication $fg\colon S \to S\colon z \mapsto f(z)g(z)$, neutral elements being the constant maps $S \to S\colon z \mapsto 0$ and $S \to S\colon z \mapsto 1$, respectively.

Then $\widehat{\varphi}\colon R[X] \to \mathrm{Maps}(S,S)\colon f \mapsto \widehat{f}_\varphi$ is a ring homomorphism: For $f, g \in R[X]$ we have $\widehat{\varphi}(f+g) = (z \mapsto (f+g)_\varphi(z)) = (z \mapsto f_\varphi(z) + g_\varphi(z)) = \widehat{\varphi}(f) + \widehat{\varphi}(g)$ and $\widehat{\varphi}(fg) = (z \mapsto (fg)_\varphi(z)) = (z \mapsto f_\varphi(z) g_\varphi(z)) = \widehat{\varphi}(f)\widehat{\varphi}(g)$ and $\widehat{\varphi}(1) = (z \mapsto (1)_\varphi(z)) = (z \mapsto 1) = 1$. Note that $\widehat{\varphi}$ in general is not injective; for example for $R = S := \mathbb{Z}/2\mathbb{Z}$ and $\varphi := \mathrm{id}_{\mathbb{Z}/2\mathbb{Z}}$ we have $0 \neq f = X^2 + X \in \mathbb{Z}/2\mathbb{Z}[X]$, but $\widehat{f}(0) = \widehat{f}(1) = 0 \in \mathbb{Z}/2\mathbb{Z}$ implies $\widehat{f} = 0 \in \mathrm{Maps}(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/2\mathbb{Z})$.

c) If $f_\varphi(z) = 0$ then $z \in S$ is called a **root** or **zero** of f in S.

In particular, considering the case $S := R$ and $\varphi := \mathrm{id}_R$, an element $a \in R$ is a root of f if and only if $(X-a) \mid f \in R[X]$: There are $q, r \in R[X]$ such that $f = q \cdot (X-a) + r$, where $r = 0$ or $\deg(r) < \deg(X-a) = 1$, that is $r \in R$, showing that $r = f(a) - q(a) \cdot (a-a) = f(a)$.

If $R = K$ is a field, then $a \in K$ is called a root of $0 \neq f \in K[X]$ of **multiplicity** $\nu_a(f) := \nu_{X-a}(f) \in \mathbb{N}_0$; note that $X - a \in \mathcal{P}$. From $\sum_{a \in K} \nu_a(f) \leq \deg(f)$ we conclude that f has at most $\deg(f) \in \mathbb{N}_0$ roots in K, counted with multiplicity. If $\deg(f) \in \{2, 3\}$ then considering factorisations shows that f is irreducible if and only if f does not have any root in K. The field K is called **algebraically closed** if any polynomial in $K[X] \setminus K$ has a root in K, or equivalently if $\mathcal{P} = \{X - a \in K[X]; a \in K\}$; for example, by the **Fundamental Theorem of Algebra [Gauß, 1801]** the field $\mathbb{C}$ is algebraically closed.

8 Eigenvalues

(8.1) Similarity. Let K be a field. Matrices $A, D \in K^{n \times n}$, where $n \in \mathbb{N}_0$, are called **similar**, if there is $P \in \mathrm{GL}_n(K)$ such that $D = P^{-1}AP$. Similarity is an equivalence relation, the equivalence classes are called **similarity classes**.

The matrix A is called **diagonalisable**, if it is similar to a diagonal matrix. The matrix A is called **triangularisable**, if it is similar to a **(lower) triangular matrix**, that is a matrix $M := [b_{ij}]_{ij} \in K^{n \times n}$ such that $b_{ij} = 0$ for all $j > i \in \{1, \dots, n\}$; in particular a diagonalisable matrix is triangularisable. Note that triangularisability is equivalent to requiring that A is similar to an **upper triangular matrix**, that is a matrix $N := [c_{ij}]_{ij} \in K^{n \times n}$ such that $c_{ij} = 0$ for all $i > j \in \{1, \dots, n\}$: Letting $P := [a_{ij}]_{ij} \in K^{n \times n}$ where $a_{ij} := 1$ if and only if $i + j = n + 1$, and $a_{ij} := 0$ elsewise, for any lower triangular matrix $M \in K^{n \times n}$ the matrix $P^{-1}NP \in K^{n \times n}$ is an upper triangular.

Let V be a K-vector space such that $\dim_K(V) = n$. Then $\varphi, \psi \in \mathrm{End}_K(V)$ are called **similar**, if there are K-bases B and C of V such that ${}_B\varphi_B = {}_C\psi_C \in K^{n \times n}$; note that since $P := {}_B\mathrm{id}_C \in \mathrm{GL}_n(K)$ and ${}_C\psi_C = P^{-1} \cdot {}_B\psi_B \cdot P$ this is equivalent to saying that ${}_B\varphi_B$ and ${}_B\psi_B$ are similar. Moreover, φ is called **diagonalisable** or **triangularisable**, if ${}_B\varphi_B$ is diagonalisable or triangularisable, respectively, for some and hence any K-basis $B \subseteq V$.

For example, let $A := \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ and $D := \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$. Then A is similar to D, hence is diagonalisable: Letting $P := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ we have $P^{-1} := \frac{1}{2} \cdot P$ and $P^{-1}AP = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} = D$.

(8.2) Eigenvalues. a) Let K be a field, let V be a K-vector space, and let $\varphi \in \mathrm{End}_K(V)$. Then $a \in K$ is called an **eigenvalue** of φ, if there is an **eigenvector** $0 \neq v \in V$ such that $\varphi(v) = av$.

Given $a \in K$, we have $\varphi - a \cdot \mathrm{id}_V \in \mathrm{End}_K(V)$ as well, hence we have $T_a(\varphi) := \ker(\varphi - a \cdot \mathrm{id}_V) = \{v \in V; \varphi(v) = av\} \leq V$, being called the **eigenspace** of φ with respect to a. Hence $T_a(\varphi) \setminus \{0\}$ is the set of eigenvectors of φ with eigenvalue a. Thus letting $\gamma_a(\varphi) := \dim_K(T_a(\varphi)) \in \mathbb{N}_0 \stackrel{.}{\cup} \{\infty\}$ be its **geometric multiplicity**, a is an eigenvalue of φ if and only if $\gamma_a(\varphi) \geq 1$. In particular, from $\ker(\varphi) = T_0(\varphi)$ we infer that φ is injective if and only if 0 is not an eigenvalue.

Let $V' := \sum_{a \in K} T_a(\varphi) \leq V$. Then we have $V' := \bigoplus_{a \in K, \gamma_a(\varphi) \geq 1} T_a(\varphi)$: We show that any sequence $[v_i \in T_{a_i}(\varphi) \setminus \{0\}; i \in \mathcal{I}]$, where $\mathcal{I}$ is a set and the $a_i \in K$ are pairwise different, is K-linearly independent: Let $\mathcal{J} \subseteq \mathcal{I}$ be finite, where we may assume that $\mathcal{J} = \{1, \ldots, n\}$ for some $n \in \mathbb{N}_0$. We proceed by induction, the case $n = 0$ being trivial: Let $b_1, \ldots, b_n \in K$ such that $\sum_{i=1}^n b_i v_i = 0$. Hence we have $0 = \varphi(\sum_{i=1}^n b_i v_i) = \sum_{i=1}^n a_i b_i v_i$, and thus $0 = a_n \cdot \sum_{i=1}^n b_i v_i - \sum_{i=1}^n a_i b_i v_i = \sum_{i=1}^{n-1} (a_n - a_i) b_i v_i$. By induction we get $(a_n - a_i) b_i = 0$, and $a_n - a_i \neq 0$ implies $b_i = 0$, for all $i \in \{1, \ldots, n-1\}$, thus $v_n \neq 0$ implies $b_n = 0$.

For example, let $\frac{\partial}{\partial x} \in \mathrm{End}_{\mathbb{R}}(C^\infty(\mathbb{R}))$. For $\epsilon_a \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \exp(ax)$, where $a \in \mathbb{R}$, we have $\frac{\partial}{\partial x}(\epsilon_a) = a\epsilon_a$. Hence a is an eigenvalue of $\frac{\partial}{\partial x}$, having $\epsilon_a \in C^\infty(\mathbb{R})$ as an eigenvector, and $[\epsilon_a \in C^\infty(\mathbb{R}); a \in \mathbb{R}]$ is $\mathbb{R}$-linearly independent; see (3.5).

b) If V is finitely generated such that $\dim_K(V) = n \in \mathbb{N}_0$, then choosing a K-basis $B \subseteq V$ and identifying $V \to K^{n \times 1} \colon v \mapsto {}_B v$ translates eigenvalues and eigenvectors of $\varphi \in \mathrm{End}_K(V)$ into those of ${}_B\varphi_B \in K^{n \times n}$:

Letting $V := K^{n \times 1}$, the eigenvalues and eigenvectors of $A \in K^{n \times n}$ are defined to be those of $\varphi_A \colon V \to V \colon v \mapsto Av$. Hence $a \in K$ is an eigenvalue of A if and only if $T_a(A) := \ker(A - aE_n) \neq \{0\}$, which holds if and only if $\det(aE_n - A) = \det(-(A - aE_n)) = (-1)^n \cdot \det(A - aE_n) = 0$. For $P \in \mathrm{GL}_n(K)$ we have $n - \gamma_a(P^{-1}AP) = \mathrm{rk}(P^{-1}AP - aE_n) = \mathrm{rk}(P^{-1}(A - aE_n)P) = \mathrm{rk}(A - aE_n) = n - \gamma_a(A)$, thus geometric multiplicities only depend on similarity classes.

The matrix A is diagonalisable if and only if there is a K-basis $\{v_1, \ldots, v_n\} \subseteq V$ consisting of eigenvectors of A. In this case, for $P := [v_1, \ldots, v_n] \in \mathrm{GL}_n(K)$ we have $P^{-1}AP = D := \mathrm{diag}[a_1, \ldots, a_n] \in K^{n \times n}$, where $\{a_1, \ldots, a_n\}$ are the eigenvalues of A; since $n - \gamma_a(A) = \mathrm{rk}(D - aE_n) = |\{i \in \{1, \ldots, n\}; a_i \neq a\}|$, for all $a \in K$, any eigenvalue occurs with its geometric multiplicity, and the eigenvalues together with their geometric multiplicities are called the **spectrum** of A. In particular, A has at most n pairwise different eigenvalues; in this case A is diagonalisable, and we have $\gamma_a(A) \leq 1$ for all $a \in K$.

(8.3) Characteristic polynomials. Let K be a field and let $A \in K^{n \times n}$, where $n \in \mathbb{N}_0$. Then $XE_n - A \in K[X]^{n \times n}$ is called the **characteristic matrix** associated with A, and let $\chi_A := \det(XE_n - A) \in K[X]$ be the monic **characteristic polynomial** of A. For example, for a diagonal matrix $D := \mathrm{diag}[a_1, \ldots, a_n] \in K^{n \times n}$ we have $\chi_D = \det(XE_n - D) = \prod_{i=1}^n (X - a_i) \in K[X]$.

We have $\deg(\chi_A) = n$ and $\chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)$. Moreover, χ_A only depends on the similarity class of A: For $P \in \mathrm{GL}_n(K)$ we have $\chi_{P^{-1}AP} = \det(XE_n - P^{-1}AP) = \det(P^{-1}(XE_n - A)P) = \det(XE_n - A) = \chi_A \in K[X]$.

Thus, if V is a finitely generated K-vector space and $\varphi \in \mathrm{End}_K(V)$, choosing a K-basis $B \subseteq V$ yields the **characteristic polynomial** $\chi_\varphi := \chi_{{}_B\varphi_B} \in K[X]$.

Given $a \in K$, the multiplicity $\nu_a(A) := \nu_a(\chi_A) = \nu_{X-a}(\chi_A) \in \mathbb{N}_0$ is called the associated **algebraic multiplicity**; in particular, algebraic multiplicities only depend on the similarity class of A. Hence a is an eigenvalue of A if and only if $\chi_A(a) = 0$, that is if and only if $\nu_a(A) \geq 1$; in particular, if K is algebraically closed then A has an eigenvalue. Moreover, this again shows that A has at most n pairwise different eigenvalues; in this case we have $\nu_a(A) \leq 1$ for all $a \in K$.

For any $a \in K$ we have $\nu_a(A) \geq \gamma_a(A)$: Let $P := [v_1, \ldots, v_n] \in \mathrm{GL}_n(K)$ be a K-basis of $V := K^{n\times 1}$ such that $[v_1, \ldots, v_m]$ is a K-basis of $T_a(A) \leq V$, where $m := \gamma_a(A)$. $P^{-1}AP = \left[\begin{array}{c|c} D & * \\ \hline 0 & A' \end{array}\right]$, where $D = \mathrm{diag}[a, \ldots, a] \in K^{m\times m}$ and $A' \in K^{(n-m)\times(n-m)}$, implies $\chi_A = \det \left[\begin{array}{c|c} XE_m - D & * \\ \hline 0 & XE_{n-m} - A' \end{array}\right] = \det(XE_m - D) \cdot \det(XE_{n-m} - A') = \chi_D \cdot \chi_{A'} = (X-a)^m \cdot \chi_{A'} \in K[X]$, thus $\nu_a(A) \geq m$. Recall that we have $\gamma_a(A) = 0$ if and only if $\nu_a(A) = 0$, and thus $\nu_a(A) = 1$ implies $\gamma_a(A) = 1$.

(8.4) Diagonalisability. Let K be a field and let $A \in K^{n\times n}$, where $n \in \mathbb{N}_0$. Then A is diagonalisable if and only if $\chi_A \in K[X]$ splits into linear factors and for all $a \in K$ we have $\nu_a(A) = \gamma_a(A)$:

If $A = \mathrm{diag}[a_1, \ldots, a_n] \in K^{n\times n}$, then we have $\chi_A = \prod_{i=1}^n (X - a_i) \in K[X]$, where $\nu_a(A) = |\{i \in \{1, \ldots, n\}; a_i = a\}| = n - |\{i \in \{1, \ldots, n\}; a_i \neq a\}| = \gamma_a(A)$ for all $a \in K$. Conversely, if $\chi_A = \prod_{i=1}^s (X - a_i)^{\nu_{a_i}(A)} \in K[X]$, where $\{a_1, \ldots, a_s\} \subseteq K$ are the eigenvalues of A with multipliticities $\nu_{a_i}(A) = \gamma_{a_i}(A) \in \mathbb{N}$, for some $s \in \mathbb{N}_0$, then let $B_i \subseteq T_{a_i}(A)$ be a K-basis for $i \in \{1, \ldots, s\}$. Since $\bigoplus_{i=1}^s T_{a_i}(A)$ is a direct sum, we conclude that $B := [B_1, \ldots, B_s]$ is K-linearly independent, and $\sum_{i=1}^s |B_i| = \sum_{i=1}^s \nu_{a_i}(A) = n$ implies that $B \subseteq V$ is a K-basis consisting of eigenvectors of A, hence A is diagonalisable. ♯

For example: **i)** Let $A := \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{R}^{2\times 2}$, which is a reflection in $\mathbb{R}^{2\times 1}$. Then we have $XE_n - A = \begin{bmatrix} X & -1 \\ -1 & X \end{bmatrix} \in \mathbb{R}[X]^{2\times 2}$, thus $\chi_A = X^2 - 1 = (X-1)(X+1) \in \mathbb{R}[X]$. Hence A has precisely the eigenvalues $\pm 1 \in \mathbb{R}$, where $\nu_{\pm 1}(A) = 1$. We have $\ker(A - E_2) = \langle [1,1]^{\mathrm{tr}} \rangle_{\mathbb{R}}$ and $\ker(A + E_2) = \langle [1,-1]^{\mathrm{tr}} \rangle_{\mathbb{R}}$. Hence letting $P := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ we have $P^{-1}AP = \mathrm{diag}[1,-1] \in \mathbb{R}^{2\times 2}$.

ii) Let $A = \begin{bmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{bmatrix} \in \mathbb{R}^{2\times 2}$, which is the rotation with respect to the angle $\omega \in \mathbb{R}$ in $\mathbb{R}^{2\times 1}$. Then we have $XE_n - A = \begin{bmatrix} X - \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & X - \cos(\omega) \end{bmatrix} \in \mathbb{R}[X]^{2\times 2} \subseteq \mathbb{C}[X]^{2\times 2}$, from which we get $\chi_A = X^2 - 2\cos(\omega)X + 1 \in \mathbb{R}[X] \subseteq \mathbb{C}[X]$, having roots $a_\pm := \cos(\omega) \pm i \cdot \sin(\omega) = \exp(\pm i\omega) \in \mathbb{C}$. Hence we have

$a_\pm \in \mathbb{R}$ if and only if $\omega = k\pi$, where $k \in \mathbb{Z}$; in this case we have $A = (-1)^k \cdot E_2$, which already is a diagonal matrix, and $\chi_A = (X - (-1)^k)^2$, thus $\nu_{(-1)^k}(A) = \gamma_{(-1)^k}(A) = 2$.

If $\omega \notin \pi\mathbb{Z}$ then $a_\pm \in \mathbb{C} \setminus \mathbb{R}$. Thus $\chi_A \in \mathbb{R}[X]$ is irreducible, and A does not have any eigenvalues in $\mathbb{R}$, in particular A is not diagonalisable over $\mathbb{R}$. But from $\chi_A = (X - a_+)(X - a_-) \in \mathbb{C}[X]$, where $a_+ \neq a_-$, we infer that A has precisely the eigenvalues $a_\pm \in \mathbb{C}$, where $\nu_{a_\pm}(A) = \gamma_{a_\pm}(A) = 1$, hence A is diagonalisable over $\mathbb{C}$, being similar to $\mathrm{diag}[a_+, a_-] \in \mathbb{C}^{2\times 2}$. More precisely, we have $\ker(A - a_+E_2) = \ker\left(\begin{bmatrix} -i\sin(\omega) & -\sin(\omega) \\ \sin(\omega) & -i\sin(\omega) \end{bmatrix}\right) = \ker\left(\begin{bmatrix} i & 1 \\ i & 1 \end{bmatrix}\right) = \langle [i, 1]^{\mathrm{tr}} \rangle_{\mathbb{C}}$ and $\ker(A - a_-E_2) = \ker\left(\begin{bmatrix} i\sin(\omega) & -\sin(\omega) \\ \sin(\omega) & i\sin(\omega) \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix}\right) = \langle [1, i]^{\mathrm{tr}} \rangle_{\mathbb{C}}$; thus picking the vectors indicated we get the $\mathbb{C}$-basis given by $P := \begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C})$ and $P^{-1}AP = \mathrm{diag}[a_+, a_-] \in \mathbb{C}^{2\times 2}$. Note that the latter statement also holds for $\omega \in \pi\mathbb{Z}$.

iii) Let $A := \begin{bmatrix} 1 & . \\ 1 & 1 \end{bmatrix} \in \mathbb{C}^{2\times 2}$. Then we have $XE_n - A = \begin{bmatrix} X-1 & . \\ -1 & X-1 \end{bmatrix} \in \mathbb{C}[X]^{2\times 2}$, thus $\chi_A = (X-1)^2 \in \mathbb{C}[X]$. Hence A has precisely the eigenvalue $1 \in \mathbb{C}$, where $\nu_1(A) = 2$. We have $\ker(A - E_2) = \langle [0, 1]^{\mathrm{tr}} \rangle_{\mathbb{C}}$, thus $\gamma_1(A) = 1$, implying that A is not diagonalisable.

(8.5) Example: Fibonacci numbers. Let $[x_i \in \mathbb{R}; i \in \mathbb{N}_0]$ be a linear recurrent sequence of degree 2 given by $x_{i+2} := x_i + x_{i+1}$; see (3.12). Starting with $x_0 := 0$ and $x_1 := 1$ we obtain the sequence $[0, 1; 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots]$ of **Fibonacci numbers [1202]**, and starting with $x'_0 := 2$ and $x'_1 := 1$ we obtain the sequence $[2, 1; 3, 4, 7, 11, 18, 29, 47, 76, 123, \ldots]$ of **Lucas numbers**. To find a closed formula for the x_i we proceed as follows:

Letting $A := \begin{bmatrix} . & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ we have $A \cdot [x_i, x_{i+1}]^{\mathrm{tr}} = [x_{i+1}, x_{i+2}]^{\mathrm{tr}} \in \mathbb{R}^{2\times 1}$, thus $[x_i, x_{i+1}]^{\mathrm{tr}} = A^i \cdot [x_0, x_1]^{\mathrm{tr}}$ for all $i \in \mathbb{N}_0$. Assume that A is diagonalisable, that is there is $P \in \mathrm{GL}_2(\mathbb{R})$ such that $P^{-1}AP = \mathrm{diag}[\rho, \rho']$. Then we have $P^{-1}A^iP = (P^{-1}AP)^i = (\mathrm{diag}[\rho, \rho'])^i = \mathrm{diag}[\rho^i, \rho'^i]$, and thus $A^i = P \cdot \mathrm{diag}[\rho^i, \rho'^i] \cdot P^{-1}$.

Now we have $XE_n - A = \begin{bmatrix} X & -1 \\ -1 & X-1 \end{bmatrix} \in \mathbb{R}[X]^{2\times 2}$, thus $\chi_A = X^2 - X - 1 = (X - \rho)(X - \rho') \in \mathbb{R}[X]$, where $\rho := \frac{1}{2}(1 + \sqrt{5}) \in \mathbb{R}$ and $\rho' := \frac{1}{2}(1 - \sqrt{5}) \in \mathbb{R}$; we have $\rho + \rho' = 1$ and $\rho\rho' = -1$. Hence A has precisely the eigenvalues $\rho, \rho' \in \mathbb{R}$, where $\nu_\rho(A) = \nu_{\rho'}(A) = 1$. We have $\ker(A - \rho E_2) = \langle [1, \rho]^{\mathrm{tr}} \rangle_{\mathbb{R}}$ and $\ker(A - \rho' E_2) = \langle [1, \rho']^{\mathrm{tr}} \rangle_{\mathbb{R}}$, thus $\gamma_\rho(A) = \gamma_{\rho'}(A) = 1$. Letting $P := \begin{bmatrix} 1 & 1 \\ \rho & \rho' \end{bmatrix} \in$

$\mathbb{R}^{2\times 2}$ yields $P^{-1}AP = \mathrm{diag}[\rho, \rho']$, and $P^{-1} := \frac{1}{\rho'-\rho} \cdot \begin{bmatrix} \rho' & -1 \\ -\rho & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ yields

$$A^i = P \cdot \mathrm{diag}[\rho^i, \rho'^i] \cdot P^{-1} = \frac{1}{\rho' - \rho} \cdot \begin{bmatrix} \rho^i\rho' - \rho\rho'^i & \rho'^i - \rho^i \\ \rho^{i+1}\rho' - \rho\rho'^{i+1} & \rho'^{i+1} - \rho^{i+1} \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

implying that the i-th Fibonacci and Lucas numbers are given as $x_i = \frac{\rho^i - \rho'^i}{\rho - \rho'} = \frac{1}{\sqrt{5}} \cdot (\rho^i - \rho'^i)$ and $x'_i = \frac{2(\rho^i\rho' - \rho\rho'^i) + (\rho'^i - \rho^i)}{\rho - \rho'} = \rho^i + \rho'^i$, respectively.

(8.6) Minimum polynomials. a) Let K be a field and $A \in K^{n\times n}$, where $n \in \mathbb{N}_0$. For the ring homomorphism $\sigma\colon K \to K^{n\times n}\colon a \mapsto aE_n$ we have $aE_n \cdot A = A \cdot aE_n$, hence there is an evaluation map $\sigma_A\colon K[X] \to K^{n\times n}\colon f \mapsto f(A)$.

For $f \in K[X]$ let $T_f(A) := \ker(f(A)) := \{v \in K^{n\times 1}; f(A)v = 0\} \leq K^{n\times 1}$ be the **generalised eigenspace** of A with respect to f; note that $T_a(A) = T_{X-a}(A)$. Then $\dim_K(T_f(A)) \in \mathbb{N}_0$ only depends on the similarity class of A: For $P \in \mathrm{GL}_n(K)$ and $v \in T_f(A)$ we have $f(P^{-1}AP) \cdot P^{-1}v = P^{-1}f(A)P \cdot P^{-1}v = P^{-1}f(A)v = 0$, thus $P^{-1}T_f(A) \leq T_f(P^{-1}AP)$, hence replacing A by $P^{-1}AP$ yields $PT_f(P^{-1}AP) \leq T_f(A)$, thus we infer $P^{-1}T_f(A) = T_f(P^{-1}AP)$, in particular implying $\dim_K(T_f(A)) = \dim_K(T_f(P^{-1}AP))$.

For $v \in T_f(A)$ we have $f(A)Av = Af(A)v = 0$, thus $T_f(A)$ is **A-invariant**, that is we have $A \cdot T_f(A) \leq T_f(A)$. Moreover, if $f = gh \in K[X]$, then for $v \in T_g(A)$ we have $f(A)v = h(A)g(A)v = 0$, thus $T_g(A) \leq T_f(A)$; in particular if $f \sim g \in K[X]$ then we have $T_f(A) = T_g(A)$.

b) Let $I_A := \ker(\sigma_A) = \{f \in K[X]; f(A) = 0 \in K^{n\times n}\} \trianglelefteq K[X]$ be the **order ideal** of A; note that $I_A = K[X]$ if and only if $n = 0$. Then we have $I_A \neq \{0\}$: The sequence $[A^i \in K^{n\times n}; i \in \mathbb{N}_0]$ is K-linearly dependent. Hence let $k \in \mathbb{N}_0$ be minimal such that $[A^i \in K^{n\times n}; i \in \{0,\ldots,k\}]$ is K-linearly dependent; we have $k \leq \dim_K(K^{n\times n}) = n^2$. Then there are $c_0,\ldots,c_{k-1} \in K$ such that $A^k + \sum_{i=0}^{k-1} c_iA^i = 0$, hence $0 \neq \mu := X^k + \sum_{i=0}^{k-1} c_iX^i \in I_A \trianglelefteq K[X]$.

Since $I_A \trianglelefteq K[X]$ is principal, there is a unique monic polynomial $0 \neq \mu_A \in K[X]$ such that $\langle \mu_A \rangle = I_A \trianglelefteq K[X]$, being called the **minimum polynomial** of A. That is we have $\mu_A \in \gcd(I_A)$, or equivalently μ_A is the unique monic non-zero polynomial of smallest degree such that $\mu_A(A) = 0$. Indeed, letting $k \in \mathbb{N}_0$ and $\mu \in K[X]$ be as above, since $[A^i \in K^{n\times n}; i \in \{0,\ldots,k-1\}]$ is K-linearly independent we conclude that $\mu_A = \mu$; in particular we have $\deg(\mu_A) = k \leq n^2$. For example, we have $\deg(\mu_A) = 0$ if and only if $n = 0$, in which case we have $\mu_A = 1 \in K[X]$; for $n \geq 1$ and $a \in K$ we have $\mu_{aE_n} = X - a \in K[X]$.

Then $\mu_A \in K[X]$ only depends on the similarity class of A: For $P \in \mathrm{GL}_n(K)$ we have $\mu_A(P^{-1}AP) = P^{-1}\mu_A(A)P = 0$, thus we have $\mu_{P^{-1}AP} \mid \mu_A \in K[X]$, hence replacing A by $P^{-1}AP$ yields $\mu_A \mid \mu_{P^{-1}AP}$, that is $\mu_{P^{-1}AP} = \mu_A$. Thus, if V is a finitely generated K-vector space and $\varphi \in \mathrm{End}_K(V)$, choosing a K-basis $B \subseteq V$ yields the **minimum polynomial** $\mu_\varphi := \mu_{{}_B\varphi_B} \in K[X]$.

(8.7) Theorem: Cayley-Hamilton. Let K be a field and let $A \in K^{n \times n}$, where $n \in \mathbb{N}_0$. Then we have $\chi_A \in I_A$, that is we have $\mu_A \mid \chi_A \in K[X]$; in particular we have $\deg \mu_A \leq n$.

Proof. For $n = 0$ we have $\mu_A = \chi_A = 1 \in K[X]$, hence we may assume $n \geq 1$. The entries of the adjoint matrix $\mathrm{adj}(XE_n - A) \in K[X]^{n \times n}$ consist of $(n-1)$-minors of the characteristic matrix $XE_n - A \in K[X]^{n \times n}$, hence are 0 or have degree at most $n-1$. Thus there are $A_0, \ldots, A_{n-1} \in K^{n \times n}$ such that $\mathrm{adj}(XE_n - A) = \sum_{i=0}^{n-1} X^i A_i \in K[X]^{n \times n}$. Letting $\chi_A = \sum_{i=0}^{n} b_i X^i \in K[X]$ we get $\sum_{i=0}^{n} b_i X^i E_n = \chi_A E_n = (XE_n - A) \cdot \mathrm{adj}(XE_n - A) = (XE_n - A) \cdot \sum_{i=0}^{n-1} X^i A_i = X^n A_{n-1} - AA_0 + \sum_{i=1}^{n-1} X^i (A_{i-1} - AA_i) \in K[X]^{n \times n}$, thus a comparison of coefficients yields $b_n E_n = A_{n-1}$ and $b_0 E_n = -AA_0$, as well as $b_i E_n = A_{i-1} - AA_i$ for $i \in \{1, \ldots, n-1\}$. Hence we obtain $\chi_A(A) = \sum_{i=0}^{n} b_i A^i = \sum_{i=0}^{n} A^i (b_i E_n) = A^n A_{n-1} - AA_0 + \sum_{i=1}^{n-1} (A^i A_{i-1} - A^{i+1} A_i) = 0$. ♯

(8.8) Elementary components. a) Let K be a field, let $V := K^{n \times 1}$ and let $A \in K^{n \times n}$, where $n \in \mathbb{N}_0$. Let $f = gh \in K[X]$, where g and h are coprime, such that $\mu_A \mid f$, that is $f(A) = 0$. Then we have $V = T_g(A) \oplus T_h(A)$, where moreover $T_g(A) = \mathrm{im}(h(A))$ and $T_h(A) = \mathrm{im}(g(A))$:

Since $1 \in \gcd(g, h) \subseteq K[X]$, there are $g', h' \in K[X]$ such that $1 = gg' + hh' \in K[X]$. For $v = g(A)w \in \mathrm{im}(g(A))$, where $w \in V$, we get $h(A)v = h(A)g(A)w = f(A)w = 0$, implying $\mathrm{im}(g(A)) \leq T_h(A)$, and similarly $\mathrm{im}(h(A)) \leq T_g(A)$. For $v \in T_g(A)$ we have $v = E_n v = E_n v - g'(A)g(A)v = h(A)h'(A)v \in \mathrm{im}(h(A))$, implying $T_g(A) \leq \mathrm{im}(h(A))$, and similarly $T_h(A) \leq \mathrm{im}(g(A))$. Hence we have $T_g(A) = \mathrm{im}(h(A))$ and $T_h(A) = \mathrm{im}(g(A))$.

For $v \in V$ we have $v = E_n v = g(A)g'(A)v + h(A)h'(A)v$, hence we have $V = \mathrm{im}(g(A)) + \mathrm{im}(h(A)) = T_g(A) + T_h(A)$. Finally, assume that $0 \neq v \in T_g(A)$ and $0 \neq w \in T_h(A)$ are K-linearly dependent, then we have $0 \neq v \in T_g(A) \cap T_h(A)$, hence $v = E_n v = g'(A)g(A)v + h'(A)h(A)v = 0$, a contradiction. This shows that $V = T_g(A) \oplus T_h(A)$. ♯

Since $T_g(A) \leq V$ and $T_h(A) \leq V$ are A-invariant, choosing K-bases $B \subseteq T_g(A)$ and $C \subseteq T_h(A)$ we get $A_g := {}_B(\varphi_A|_{T_g(A)})_B \in K^{l \times l}$ and $A_h := {}_C(\varphi_A|_{T_h(A)})_C \in K^{m \times m}$, where $l := \dim_K(T_g(A)) \in \mathbb{N}_0$ and $m := \dim_K(T_h(A)) \in \mathbb{N}_0$. Hence $P := [B, C] \in \mathrm{GL}_n(K)$ is a K-basis of V, and A is similar to the **block diagonal matrix** $P^{-1}AP = A_g \oplus A_h \in K^{n \times n}$.

We have $\mu_{A_g} \mid g \in K[X]$ and $\mu_{A_h} \mid h \in K[X]$, as well as $\mu_A \in \mathrm{lcm}(\mu_{A_g}, \mu_{A_h}) \subseteq K[X]$, hence μ_{A_h} and μ_{A_h} being coprime we infer $\mu_A = \mu_{A_g}\mu_{A_h}$. In particular, since $\mu_{A_g} \mid \gcd(g, \mu_A)$, we infer that if $1 \in \gcd(g, \mu_A)$ then we have $\mu_{A_g} = 1$, in other words $T_g(A) = \{0\}$.

If $\mu_A \sim f$ then we have $\mu_{A_g} \sim g$ and $\mu_{A_h} \sim h$, entailing $\deg(g) = \deg(\mu_{A_g}) \leq \deg(\chi_{A_g}) = \dim_K(T_g(A))$, and similarly $\deg(h) \leq \dim_K(T_h(A))$; in particular, if g is non-constant then we have $T_g(A) \neq \{0\}$. Hence, if $\mu_A = \prod_{p \in \mathcal{P}}^{s} p^{\nu_p} \in K[X]$, where $\nu_p \in \mathbb{N}_0$ and $\mathcal{P} \subseteq K[X]$ is the set of monic irreducible polynomials, then we have the direct sum decomposition $V = \bigoplus_{p \in \mathcal{P}} T_{p^{\nu_p}}(A)$ into **elementary components** $T_{p^{\nu_p}}(A) \leq V$, where $T_{p^{\nu_p}}(A) = \{0\}$ if and only if $\nu_p = 0$.

For example, let $A := \begin{bmatrix} . & . & 1 \\ 1 & . & . \\ . & 1 & . \end{bmatrix} \in \mathbb{Q}^{3 \times 3}$. Then we have $\chi_A = X^3 - 1 = (X - 1)(X^2 + X + 1) \in \mathbb{Q}[X]$, where the factors are irreducible and hence coprime. We have $\ker(A - E_3) = \langle [1,1,1]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$ and $\ker(A^2 + A + E_3) = \langle [1,-1,0]^{\mathrm{tr}}, [0,1,-1]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$. Hence letting $P := \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{Q})$ we get $P^{-1}AP = \left[\begin{array}{c|cc} 1 & . & . \\ \hline . & . & -1 \\ . & 1 & -1 \end{array}\right]$.

b) The matrix A is diagonalisable if and only if μ_A splits into pairwise non-associate linear factors:

If A is diagonal, then we have $\chi_A = \prod_{i=1}^{s} (X - a_i)^{\nu_{a_i}(A)} \in K[X]$, for some $s \in \mathbb{N}_0$, where $\{a_1, \ldots, a_s\} \subseteq K$ are the eigenvalues of A, each occurring with multiplicity $\nu_{a_i}(A) = \gamma_{a_i}(A) \in \mathbb{N}$. Then we have $\mu_A = \prod_{i=1}^{s} (X - a_i) \in K[X]$: Letting $f := \prod_{i=1}^{s} (X - a_i)$, we have $f(A) = \prod_{i=1}^{s} (A - a_i E_n) = 0$, hence $\mu_A \mid f$; the maximal proper divisors of f being $f_j := \prod_{i \neq j} (X - a_i) \in K[X]$, where $j \in \{1, \ldots, s\}$, we from $\mathrm{rk}(f_j(A)) = \nu_{a_j}(A) \geq 1$ infer that $\mu_A \nmid f_j$.

Conversely, let $\mu_A = \prod_{i=1}^{s} (X - a_i) \in K[X]$, for some $s \in \mathbb{N}_0$, where the $a_i \in K$ are pairwise different. Then we have $V = \bigoplus_{i=1}^{s} T_{a_i}(A)$, that is V is the direct sum of the eigenspaces with respect to the a_i, thus A is diagonalisable. $\sharp$

(8.9) Jordan normal form. a) Let K be a field, let $V := K^{n \times 1}$ and let $A \in K^{n \times n}$, where $n \in \mathbb{N}_0$. Let $p := X - a \in K[X]$, and let $\mu_A = p^l$ for some $l \leq n$. For $i \in \mathbb{N}_0$ let $V_i := T_{p^i}(A) = \ker((A - aE_n)^i) \leq V$. Thus we have $V_{i-1} \leq V_i$ for $i \in \mathbb{N}$, where $V_0 = \{0\}$ and $V_{l-1} < V_l = V$. Letting $n_i := \dim_K(V_i / V_{i-1}) = \dim_K(V_i) - \dim_K(V_{i-1})$ for $i \in \mathbb{N}$, we have $n_i = 0$ for $i > l$, while $n_l > 0$, and $\sum_{i=1}^{l} n_i = n$.

Then there is a K-basis $[v_{l1}, \ldots, v_{ln_l}; v_{l-1,1}, \ldots, v_{l-1,n_{l-1}}; \ldots; v_{11}, \ldots, v_{1n_1}]$ of V, such that $[v_{i1}, \ldots, v_{in_i}; \ldots; v_{11}, \ldots, v_{1n_1}]$ is a K-basis of V_i for all $i \in \{1, \ldots, l\}$, and $v_{i-1,j} = p(A)v_{ij}$ for all $i \in \{2, \ldots, l\}$ and $j \in \{1, \ldots, n_i\}$; thus in particular we have $n_{i-1} \geq n_i$ for all $i \in \mathbb{N}$:

We proceed by induction on $l \in \mathbb{N}_0$; the cases $l \leq 1$ being trivial, let $l \geq 2$: Let $k := n_l$ and let $v_1, \ldots, v_k \in V$ such that $\{v_1 + V_{l-1}, \ldots, v_k + V_{l-1}\} \subseteq V/V_{l-1}$ is a K-basis. Letting $w_j := p(A)v_j = (A - aE_n)v_j$ for $j \in \{1, \ldots, k\}$, we have $p^{l-1}(A)w_j = p^l(A)v_j = 0$, that is $w_j \in V_{l-1}$. Moreover, $[w_1 + V_{l-2}, \ldots, w_k + V_{l-2}]$ is K-linearly independent: Let $a_1, \ldots, a_k \in K$ such that $\sum_{j=1}^{k} a_j w_j \in V_{l-2}$, then $p(A)^{l-1}(\sum_{j=1}^{k} a_j v_j) = p(A)^{l-2}(\sum_{j=1}^{k} a_j w_j) = 0$ implies $\sum_{j=1}^{k} a_j v_j \in V_{l-1}$, hence $\{v_1 + V_{l-1}, \ldots, v_k + V_{l-1}\} \subseteq V/V_{l-1}$ being K-linearly independent yields $a_j = 0$ for all $j \in \{1, \ldots, k\}$. Extending $\{w_1 + V_{l-2}, \ldots, w_k + V_{l-2}\}$ to a K-basis of V_{l-1}/V_{l-2}, we are done by induction. $\sharp$

Reordering the above K-basis we obtain

$$P = [P_{l1}, \ldots, P_{ln_l}; P_{l-1,n_l+1}, \ldots, P_{l-1,n_{l-1}}; \ldots; P_{1,n_2+1}, \ldots, P_{1,n_1}] \in \mathrm{GL}_n(K),$$

where $P_{ij} := [v_{ij}, v_{i-1,j}, \ldots, v_{1j}]$ for $i \in \{1, \ldots, l\}$ and $j \in \{n_{i+1}+1, \ldots, n_i\}$. Then $Av_{ij} = p(A)v_{ij} + av_{ij} = v_{i-1,j} + av_{ij}$ implies that $\mathrm{im}(P_{ij}) \leq V$ is A-invariant. Hence letting $m_i := n_i - n_{i+1} \in \mathbb{N}_0$, the matrix A is similar to the block diagonal matrix $P^{-1}AP = \bigoplus_{i=1}^{l} \bigoplus_{j=1}^{m_i} C_i(a)$, whose diagonal entries are **Jordan matrices**

$$C_i(a) = C_i(X-a) := \begin{bmatrix} a & . & . & . & \ldots & . \\ 1 & a & . & . & \ldots & . \\ . & 1 & a & . & \ldots & . \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ . & \ldots & . & 1 & a & . \\ . & \ldots & . & . & 1 & a \end{bmatrix} \in K^{i \times i}.$$

b) The multiplicities m_i are uniquely determined by A: For a Jordan matrix $C := C_n(a) \in K^{n \times n}$ we have $\chi_C = \det(XE_n - C) = p^n \in K[X]$, that is $\nu_a(C) = n$. Moreover, we have $\mathrm{rk}(p^i(C)) = \mathrm{rk}((C - aE_n)^i) = n - i$, that is $\dim_K(T_{p^i}(C)) = i$ for $i \in \{0, \ldots, n\}$. Thus we have $\mu_C = \chi_C = p^n \in K[X]$, and $\dim_K(T_{p^i}(C)/T_{p^{i-1}}(C)) = 1$ for $i \in \{1, \ldots, n\}$, while $\dim_K(T_{p^i}(C)/T_{p^{i-1}}(C)) = 0$ for $i > n$; in particular, we have $\gamma_a(C) = \dim_K(T_a(C)) = \dim_K(T_p(C)) = 1$.

Hence for any matrix $A = \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{m_i} C_i(a) \in K^{n \times n}$, where $m_i \in \mathbb{N}_0$ such that $\sum_{i=1}^{n} im_i = n$, we have $n_i := \dim_K(T_{p^i}(A)/T_{p^{i-1}}(A)) = \sum_{j=i}^{n} m_j$, implying that $m_i = n_i - n_{i+1}$ for all $i \in \{1, \ldots, n\}$.

c) Thus, in practice **Jordan normal forms** can be computed combinatorially, without specifying a Jordan K-basis, as soon as the K-dimensions of the K-subspaces $V_i = T_{p^i}(A) \leq V$ are known, for all $i \in \mathbb{N}_0$.

This is best explained by way of an example: Let $n = 13$ and $[\dim_K(V_i) \in \mathbb{N}_0; i \in \mathbb{N}_0] = [0, 5, 8, 10, 12, 13, 13, \ldots]$, hence we have $l = 5$ and the numbers $n_i = \dim_K(V_i) - \dim_K(V_{i-1}) \in \mathbb{N}_0$, for $i \in \mathbb{N}$, are given as $[n_i \in \mathbb{N}_0; i \in \mathbb{N}] = [5, 3, 2, 2, 1, 0, \ldots]$. We depict the numbers $n_i \in \mathbb{N}$, for $i \in \{1, \ldots, l\}$, as the rows of the following diagram, from bottom to top:

$\mathbf{v_{51}}$				
v_{41}	$\mathbf{v_{42}}$			
v_{31}	v_{32}			
v_{21}	v_{22}	$\mathbf{v_{23}}$		
v_{11}	v_{12}	v_{13}	$\mathbf{v_{14}}$	$\mathbf{v_{15}}$

Then the multiplicity $m_i = n_i - n_{i+1} \in \mathbb{N}_0$ can be read off from the diagram, as the number of columns of height $i \in \mathbb{N}$; of course it suffices to consider $i \in \{1, \ldots, l\}$. Here we obtain the column heights $[5, 4, 2, 1, 1, 0, \ldots]$, and therefrom

$[m_i \in \mathbb{N}_0; i \in \mathbb{N}] = [2, 1, 0, 1, 1, 0, \ldots]$, thus the Jordan normal form of the matrix A in question is $C_5(a) \oplus C_4(a) \oplus C_2(a) \oplus C_1(a) \oplus C_1(a) \in K^{13\times 13}$.

Moreover, the vectors v_{ij} constituting the Jordan K-basis $P \subseteq V$ can be filled into the diagram as indicated above. Then the subset $P_{ij} \subseteq P$ coincides with the vectors in column i, in other words the K-subspaces generated by tbe vectors in either column are A-invariant. Now, the construction of P can be described as follows, again by way of the above example; the vectors we are free to choose are depicted in bold face in the above diagram:

We choose $v_{51} \in V_5$, being placed on top of column 1, extending any K-basis of V_4 to a K-basis of V_5; then successively working down column 1 we get $v_{5-i,1} = p^i(A)(v_{51}) \in V_{5-i} \setminus V_{4-i}$ for $i \in \{1, \ldots, 4\}$. Then we chosse $v_{42} \in V_4$, being placed on top of column 2, so that $\{v_{41}, v_{42}\}$ extends any K-basis of V_3 to a K-basis of V_4; then successively working down column 2 we get $v_{4-i,2} = p^i(A)(v_{42}) \in V_{4-i} \setminus V_{3-i}$ for $i \in \{1, \ldots, 3\}$. Next we observe that $\{v_{31}, v_{32}\}$ already extends any K-basis of V_2 to a K-basis of V_3, so we are done for V_3. Proceeding further, we chosse $v_{23} \in V_2$, being placed on top of column 3, so that $\{v_{21}, v_{22}, v_{23}\}$ extends any K-basis of V_1 to a K-basis of V_2; then working down column 3 we get $v_{13} = p(A)(v_{23})$. Finally, we chosse $v_{14}, v_{15} \in V_1$, being placed in columns 4 and 5, extending $\{v_{11}, v_{12}, v_{13}\}$ to a K-basis of V_2; recall that we have $V_0 = \{0\}$ which has an empty K-basis. ♯

In order to present an explicit example, let $A := \begin{bmatrix} 1 & -1 & 1 \\ 3 & 5 & -3 \\ 2 & 2 & 0 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$. Then we have $\chi_A = X^3 - 6X^2 + 12X - 8 = (X-2)^3 \in \mathbb{Q}[X]$. We have $\dim_{\mathbb{Q}}(\ker(A - 2E_3)) = 3 - \mathrm{rk}(A - 2E_3) = 2$, which implies $\dim_{\mathbb{Q}}(\ker((A - 2E_3)^2)) = 3$, thus the Jordan normal form of A is $C_2(2) \oplus C_1(2) = \left[\begin{array}{cc|c} 2 & . & . \\ 1 & 2 & . \\ \hline . & . & 2 \end{array}\right]$. We have $V_1 := \ker(A - 2E_3) = \langle [1, -1, 0]^{\mathrm{tr}}, [0, 1, 1]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$, hence for $v_{21} := [1, 0, 0]^{\mathrm{tr}} \in V$ we infer that $\{v_{21} + V_1\}$ is a $\mathbb{Q}$-basis of V/V_1, thus letting $v_{11} := (A - 2E_3)v_{21} = [-1, 3, 2]^{\mathrm{tr}} \in V_1$, and extending by $v_{12} := [1, -1, 0]^{\mathrm{tr}} \in V_1$ to the $\mathbb{Q}$-basis $\{v_{11}, v_{12}\} \subseteq V_1$, we get the $\mathbb{Q}$-basis $P := \begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & -1 \\ 0 & 2 & 0 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{Q})$ such that $P^{-1}AP = C_2(2) \oplus C_1(2)$.

(8.10) Triangularisability. a) Let K be a field, let $V := K^{n\times 1}$ and let $A \in K^{n\times n}$, where $n \in \mathbb{N}_0$. Then A is triangularisable if and only if $\chi_A \in K[X]$ splits into linear factors, or equivalently if and only if $\mu_A \in K[X]$ splits into linear factors; in particular, if K is algebraically closed then A is triangularisable:

If A is triangular, then $\chi_A = \prod_{i=1}^{s} (X - a_i)^{\nu_{a_i}(A)} \in K[X]$, for some $s \in \mathbb{N}_0$, where $\{a_1, \ldots, a_s\} \subseteq K$ are the diagonal entries of A, each occurring with multiplicity $\nu_{a_i}(A) \in \mathbb{N}$. Hence χ_A and thus μ_A split into linear factors.

Let $\mu_A \in K[X]$ split into linear factors, that is we have $\mu_A = \prod_{i=1}^{s} f_i \in K[X]$, where $s \in \mathbb{N}_0$ and $f_i = (X - a_i)^{\nu_i} \in K[X]$, where in turn the $a_i \in K$ are pairwise different and $\nu_i \in \mathbb{N}$. Then we have $V = \bigoplus_{i=1}^{s} T_{f_i}(A)$, and $\mu_{A_{f_i}} \mid f_i \in K[X]$ for $i \in \{1, \ldots, s\}$. Thus choosing Jordan K-bases $B_i \subseteq T_{f_i}(A)$, for all $i \in \{1, \ldots, s\}$, and letting $B := [B_1, \ldots, B_s] \in \mathrm{GL}_n(K)$, then A is similar to the block diagonal matrix $B^{-1}AB = \bigoplus_{i=1}^{s} B_i^{-1} A_{f_i} B_i \in K^{n \times n}$, where each $B_i^{-1} A_{f_i} B_i \in K^{n_i \times n_i}$, for $n_i := \dim_K(T_{f_i}(A)) \in \mathbb{N}_0$, is again a block diagonal matrix, consisting of Jordan matrices with respect to the eigenvalue a_i. ♯

b) Since $\mu_A \mid \chi_A \in K[X]$, the irreducible divisors of μ_A are amongst those of χ_A. Actually, we proceed to show that the linear factors of μ_A and of χ_A coincide: (Indeed, all the irreducible divisors of μ_A and of χ_A coincide, but we are not able to prove this here.)

Assume to the contrary that $X - a \mid \chi_A$, but $X - a \nmid \mu_A$; then we have $\chi_A(a) = 0$, saying that $a \in K$ is an eigenvalue of A, that is $T_{X-a}(A) \neq \{0\}$; but since $X - a$ and μ_A are coprime we have $T_{X-a}A = \{0\}$, a contradiction. ♯

Now, if $\mu_A = \prod_{i=1}^{s}(X - a_i)^{l_i} \in K[X]$ splits into linear factors, then by the above we have $\chi_A = \prod_{i=1}^{s}(X - a_i)^{n_i}$, that is the algebraic multiplicity of the eigenvalue a_i is given as $\nu_{a_i}(A) = n_i = \dim_K(T_{(X-a_i)^{l_i}}(A))$, for $i \in \{1, \ldots, s\}$.

For example, we have

$$A := \begin{bmatrix} 1 & -2 & -1 & 2 \\ 0 & -1 & -1 & 2 \\ 2 & -2 & -1 & 4 \\ 1 & -1 & 0 & 1 \end{bmatrix} \sim C_2(1) \oplus C_2(-1) = \left[\begin{array}{cc|cc} 1 & . & . & . \\ 1 & 1 & . & . \\ \hline . & . & -1 & . \\ . & . & 1 & -1 \end{array}\right] \in \mathbb{Q}^{4 \times 4} :$$

We have $\chi_A = X^4 - 2X^2 + 1 = (X-1)^2(X+1)^2 \in \mathbb{Q}[X]$. Thus we have the direct sum decomposition $V = \ker((A - E_3)^2) \oplus \ker((A + E_3)^2)$ into elementary components, where $\dim_{\mathbb{Q}}(\ker((A - E_3)^2)) = 2 = \dim_{\mathbb{Q}}(\ker((A + E_3)^2))$. We have $\dim_{\mathbb{Q}}(\ker(A - E_3)) = 1 = \dim_{\mathbb{Q}}(\ker(A + E_3)) = 1$, which already implies that the Jordan normal form of A is $C_2(1) \oplus C_2(-1)$.

To obtain $P \in \mathrm{GL}_4(\mathbb{Q})$ such that $P^{-1}AP = C_2(1) \oplus C_2(-1)$ we proceed as follows: We have

$$A - E_3 = \left(\begin{bmatrix} 0 & -2 & -1 & 2 \\ 0 & -2 & -1 & 2 \\ 2 & -2 & -2 & 4 \\ 1 & -1 & 0 & 0 \end{bmatrix}\right) \quad \text{and} \quad (A - E_3)^2 = \left(\begin{bmatrix} 0 & 4 & 4 & -8 \\ 0 & 4 & 4 & -8 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right),$$

$$A + E_3 = \left(\begin{bmatrix} 2 & -2 & -1 & 2 \\ 0 & 0 & -1 & 2 \\ 2 & -2 & 0 & 4 \\ 1 & -1 & 0 & 2 \end{bmatrix}\right) \quad \text{and} \quad (A - E_3)^2 = \left(\begin{bmatrix} 4 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 8 & -8 & 0 & 8 \\ 4 & -4 & 0 & 4 \end{bmatrix}\right),$$

from which we get $\ker(A - E_3) = \langle [0,0,2,1]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$ as well as $\ker((A - E_3)^2) = \langle [0,0,2,1]^{\mathrm{tr}}, [1,0,0,0]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$, and similarly $\ker(A + E_3) = \langle [1,1,0,0]^{\mathrm{tr}} \rangle_{\mathbb{Q}}$ as well as

$\ker((A+E_3)^2) = \langle [1,1,0,0]^{\text{tr}}, [0,0,1,0]^{\text{tr}} \rangle_{\mathbb{Q}}$. Thus from $(A-E_3)[1,0,0,0]^{\text{tr}} = [0,0,2,1]^{\text{tr}}$ and $(A+E_3)[0,0,1,0]^{\text{tr}} = [-1,-1,0,0]^{\text{tr}}$ we obtain the matrix $P := \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \in \mathrm{GL}_4(\mathbb{Q})$ such that $P^{-1}AP = C_2(1) \oplus C_2(-1)$.

(8.11) Example: Harmonic oscillator. Let again $C^\infty(\mathbb{R}) := \{R \to \mathbb{R} \colon t \mapsto x(t) \text{ smooth}\}$, where now we denote variables and maps by the letters 't' and 'x', respectively, being reminiscent of their forthcoming physical interpretation as time and place, respectively. We again use the differential operator $D := \frac{\partial}{\partial t} \in \mathrm{End}_{\mathbb{R}}(C^\infty(\mathbb{R}))$, where we abbreviate $\dot{x} := D(x) = \frac{\partial}{\partial t}(x)$.

We consider a **(single) body** of **(inert) mass** $m > 0$, being fixed to a spring. Pulling the body away from the point of equilibrium, and releasing it, it will start to oscillate. Letting $x = x(t) \in \mathbb{R}$ denote the **place** of the body at time $t \in \mathbb{R}$, its **velocity** and **acceleration** are given as $\dot{x} = \dot{x}(t) \in \mathbb{R}$ and $\ddot{x} = \ddot{x}(t) \in \mathbb{R}$, respectively. By **Newton's Law of Motion** the acceleration of the body is proportional to the force exerted to it, the proportionality factor just being its mass m. In turn, the pulling-back force exerted to the body by the spring is proportional to the distance of the place of the body to the point of equilibrium, the proportionality factor being the square of the **spring constant** $k > 0$. Assuming that the point of equilibrium is $x = 0$, we thus obtain the differential equation $m\ddot{x} = -k^2 x$ of the **(free) harmonic oscillator**.

Now we additionally allow for **friction**, which exerts a decelerating force to the body. The latter is proportional to its velocity, the proportionality factor being the **friction constant** $r \geq 0$; for $r = 0$ we recover the free harmonic oscillator. Hence the differential equation of the **damped** harmonic oscillator, describing the motion of the body in this **physical system**, is given as $m\ddot{x} = -r\dot{x} - k^2 x$, a **linear** differential equation of **degree** 2 with **constant coefficients**.

Hence we are looking for the $\mathbb{R}$-subspace $\mathcal{L} = \mathcal{L}_{\rho,\omega} \leq C^\infty(\mathbb{R})$ of solutions of the $\mathbb{R}$-endomorphism $D^2 + 2\rho D + \omega^2$ of $C^\infty(\mathbb{R})$, where $\rho := \frac{r}{2m} \geq 0$ and $\omega := \frac{k}{\sqrt{m}} > 0$. A consideration of **Taylor series** shows that $\dim_{\mathbb{R}}(\mathcal{L}) = 2$. More precisely, the motion of the body is uniquely described by imposing arbitrary **initial values** $x(0) \in \mathbb{R}$ and $\dot{x}(0) \in \mathbb{R}$ for the place and the velocity of the body at time $t = 0$. In particular, pulling the body away from the point of equilibrium and releasing it, amounts to letting $x(0) := 1$, say, and $\dot{x}(0) := 0$.

Since $D^2 = -2\rho D - \omega^2$ on $\mathcal{L}$, we conclude that $\mathcal{L}$ is D-invariant, and that the action of D on $\mathcal{L}$ has minimum polynomial $\mu_D \mid p = p_{\rho,\omega} := X^2 + 2\rho X + \omega^2 = (X+\rho)^2 - (\rho^2 - \omega^2) \in \mathbb{R}[X]$. We distinguish three cases with respect to the **discriminant** $\rho^2 - \omega^2 \in \mathbb{R}$ of p being positive, zero or negative, respectively:

i) Let $\rho > \omega > 0$; physically this is the 'large friction' case. Then we have $p = (X-a)(X-b) \in \mathbb{R}[X]$, where $\{a, b\} = \{-\rho \pm \sqrt{\rho^2 - \omega^2}\}$, in particular $a \neq b$ and both $a, b < 0$. We have $\mu_D \in \{X - a, X - b, p\}$, and depending on the case for μ_D we have $\chi_D \in \{(X-a)^2, (X-b)^2, p\}$. Anyway, μ_D splits into pairwise non-associate linear factors, that is D acts diagonalisably on $\mathcal{L}$.

The map $\epsilon_c\colon \mathbb{R} \to \mathbb{R}\colon t \mapsto \exp(ct)$ fulfills $\dot{\epsilon_c} = c\epsilon_c$, that is ϵ_c is an eigenvector of D on $C^\infty(\mathbb{R})$, with respect to the eigenvalue $c \in \mathbb{R}$, see (8.2). Moreover, a consideration of Taylor series shows that the corresponding eigenspace of D actually equals $\langle \epsilon_c \rangle_\mathbb{R}$. Hence we conclude that $T_{X-a}(D) = \langle \epsilon_a \rangle_\mathbb{R}$ and $T_{X-b}(D) = \langle \epsilon_b \rangle_\mathbb{R}$, thus we have the principal subspace decomposition $\mathcal{L} = T_{X-a}(D) \oplus T_{X-b}(D) = \langle \epsilon_a \rangle_\mathbb{R} \oplus \langle \epsilon_b \rangle_\mathbb{R}$; in particular, we have $\mu_D = p = \chi_D$.

Hence any solution is of the form $x(t) = \alpha \exp(at) + \beta \exp(bt)$, for all $t \in \mathbb{R}$, where $\alpha, \beta \in \mathbb{R}$, entailing $\dot{x}(t) = \alpha a \exp(at) + \beta b \exp(bt)$. Since both $a, b < 0$ we have $\lim_{t\to\infty} x(t) = 0$, saying that the body ultimately tends to the point of equilibrium. Since for any non-zero solution we may assume that $\beta \neq 0$, we have $\dot{x}(t) = 0$ if and only if $\exp((b-a)t) = -\frac{\alpha}{\beta} \cdot \frac{a}{b}$; hence this happens for at most one $t \in \mathbb{R}$, saying that the body changes direction at most once. In particular, letting $x(0) := 1$ and $\dot{x}(0) := 0$, we get $\alpha + \beta = x(0) = 1$ and $\alpha a + \beta b = \dot{x}(0) = 0$, yielding $\alpha = \frac{b}{b-a}$ and $\beta = \frac{a}{a-b}$, that is $x(t) = \frac{1}{b-a} \cdot (b \exp(at) - a \exp(bt))$.

ii) Let $\rho = \omega > 0$. Then we have $p = (X+\rho)^2 \in \mathbb{R}[X]$. We have $\mu_D \in \{X+\rho, p\}$, thus μ_D splits into linear factors anyway, that is D acts triangularisably on $\mathcal{L}$, and we have $\chi_D = p$. Moreover, we have $T_{X+\rho}(D) = \langle \epsilon_{-\rho} \rangle_\mathbb{R}$, implying $\mu_D = p = \chi_D$, that is $\mathcal{L} = T_p(D)$ consists of a single Jordan block. Letting $\widehat{\epsilon}_c\colon \mathbb{R} \to \mathbb{R}\colon t \mapsto t\exp(ct)$, where $c \in \mathbb{R}$, we have $\dot{\widehat{\epsilon}}_c(t) = \exp(ct) + ct\exp(ct)$, for all $t \in \mathbb{R}$, hence $\dot{\widehat{\epsilon}}_c = \epsilon_c + c\widehat{\epsilon}_c$. Thus we have $(D+\rho)(\widehat{\epsilon}_{-\rho}) = \epsilon_{-\rho}$, implying that $\{\widehat{\epsilon}_{-\rho}, \epsilon_{-\rho}\}$ indeed is a Jordan $\mathbb{R}$-basis of $\mathcal{L}$.

Hence any solution is of the form $x(t) = (\alpha + \beta t)\exp(-\rho t)$, for all $t \in \mathbb{R}$, where $\alpha, \beta \in \mathbb{R}$, entailing $\dot{x}(t) = (\alpha a + \beta + \beta a t)\exp(-\rho t)$. Since $\rho > 0$ we have $\lim_{t\to\infty} x(t) = 0$, saying that the body ultimately tends to the point of equilibrium. If x is a non-zero solution, then if $\beta = 0$ we have $\dot{x}(t) = -\alpha\rho\exp(-\rho t) \neq 0$ for all $t \in \mathbb{R}$, while if $\beta \neq 0$ we have $\dot{x}(t) = 0$ if and only if $t = -\frac{\alpha}{\beta} + \frac{1}{\rho}$; hence this happens for at most one $t \in \mathbb{R}$, saying that the body changes direction at most once. In particular, letting $x(0) := 1$ and $\dot{x}(0) := 0$, we get $\alpha = x(0) = 1$ and $-\alpha\rho + \beta = \dot{x}(0) = 0$, yielding $\beta = \rho$, that is $x(t) = (1 + \rho t)\exp(-\rho t)$.

iii) Let $\omega > \rho \geq 0$; physically this is the 'small friction' case. Then $p \in \mathbb{R}[X]$ is irreducible. Hence we have $\mu_D = p = \chi_D$; in particular, does not act triangularisably on $\mathcal{L}$. To describe $\mathcal{L}$ we use **complexification**:

We have $p = (X-a)(X-\overline{a}) \in \mathbb{C}[X]$, where $\{a, \overline{a}\} = \{-\rho \pm i\varphi\} \subseteq \mathbb{C} \setminus \mathbb{R}$ and $\varphi := \sqrt{\omega^2 - \rho^2} > 0$, and where $\overline{\ }\colon \mathbb{C} \to \mathbb{C}$ denotes complex conjugation. We consider the $\mathbb{C}$-vector space $C^\infty(\mathbb{R}, \mathbb{C}) := \{\mathbb{R} \to \mathbb{C}\colon t \mapsto z(t) = x(t) + iy(t);\ x, y \in C^\infty(\mathbb{R})\}$, and we are looking for the $\mathbb{C}$-subspace $\mathcal{L}_\mathbb{C} \leq C^\infty(\mathbb{R}, \mathbb{C})$ of solutions of the $\mathbb{C}$-endomorphism $D^2 + 2\rho D + \omega^2$ of $C^\infty(\mathbb{R}, \mathbb{C})$. Again a consideration of Taylor series shows that $\dim_\mathbb{C}(\mathcal{L}_\mathbb{C}) = 2$.

Similarly, for any $c \in \mathbb{C}$ the corresponding eigenspace of D on $C^\infty(\mathbb{R}, \mathbb{C})$ is seen to be equal to $\langle \epsilon_c \rangle_\mathbb{C}$, where $\epsilon_c\colon \mathbb{C} \to \mathbb{C}\colon t \mapsto \exp(ct)$. This yields the principal subspace decomposition $\mathcal{L}_\mathbb{C} = T_{X-a}(D) \oplus T_{X-\overline{a}}(D) = \langle \epsilon_a \rangle_\mathbb{C} \oplus \langle \epsilon_{\overline{a}} \rangle_\mathbb{C}$; in particular, D acts diagonalisably on $\mathcal{L}_\mathbb{C}$. Letting $a = -\rho + i\varphi$, we have $\epsilon_a(t) = \exp(at) = \exp(-\rho t) \cdot \big(\cos(\varphi t) + i\sin(\varphi t)\big)$ and $\epsilon_{\overline{a}}(t) = \exp(-\rho t) \cdot \big(\cos(\varphi t) - i\sin(\varphi t)\big) = \overline{\epsilon_a(t)}$, for all $t \in \mathbb{R}$. Hence with respect to the $\mathbb{C}$-basis $\{\epsilon_a, \epsilon_{\overline{a}}\} \subseteq \mathcal{L}_\mathbb{C}$ the map D is represented by $\operatorname{diag}[-\rho + i \cdot \sqrt{\omega^2 - \rho^2}, -\rho - i \cdot \sqrt{\omega^2 - \rho^2}] \in \mathbb{C}^{2\times 2}$.

Now we are looking for solutions in $\mathcal{L} \subseteq \mathcal{L}_\mathbb{C}$: Letting $\tau_a := \frac{1}{2}(\epsilon_a + \epsilon_{\overline{a}})$ and $\sigma_a := \frac{1}{2i}(\epsilon_a - \epsilon_{\overline{a}})$ we have $\tau_a(t) = \exp(-\rho t)\cos(\varphi t)$ and $\sigma_a(t) = \exp(-\rho t)\sin(\varphi t)$, for all $t \in \mathbb{R}$, hence $\tau_a, \sigma_a \in \mathcal{L} \subseteq \mathcal{L}_\mathbb{C}$. Since $\epsilon_a, \epsilon_{\overline{a}} \in \langle \tau_a, \sigma_a \rangle_\mathbb{C}$ we conclude that $\langle \tau_a, \sigma_a \rangle_\mathbb{C} = \mathcal{L}_\mathbb{C}$, hence $\{\tau_a, \sigma_a\}$ is $\mathbb{C}$-linearly independent, in particular is $\mathbb{R}$-linearly independent, and thus is an $\mathbb{R}$-basis of $\mathcal{L}$; alternatively, evaluating at $t = 0$ and $t = \frac{\pi}{2}$ shows directly that $\{\tau_a, \sigma_a\}$ is $\mathbb{R}$-linearly independent. Now we have $\dot{\tau}_a(t) = -\rho\exp(-\rho t)\cos(\varphi t) - \varphi\exp(-\rho t)\sin(\varphi t)$ and $\dot{\sigma}_a(t) = -\rho\exp(-\rho t)\sin(\varphi t) + \varphi\exp(-\rho t)\cos(\varphi t)$, for all $t \in \mathbb{R}$, that is $\dot{\tau}_a = -\rho\tau_a - \varphi\sigma_a$ and $\dot{\sigma}_a = \varphi\tau_a - \rho\sigma_a$, hence with respect to the $\mathbb{R}$-basis $\{\tau_a, \sigma_a\} \subseteq \mathcal{L}$ the map D is represented by $\begin{bmatrix} -\rho & \varphi \\ -\varphi & -\rho \end{bmatrix} = \begin{bmatrix} -\rho & \sqrt{\omega^2 - \rho^2} \\ -\sqrt{\omega^2 - \rho^2} & -\rho \end{bmatrix} \in \mathbb{R}^{2\times 2}$.

Hence any solution is of the form $x(t) = \exp(-\rho t) \cdot \big(\alpha\cos(\varphi t) + \beta\sin(\varphi t)\big)$, for all $t \in \mathbb{R}$, where $\alpha, \beta \in \mathbb{R}$, entailing $\dot{x}(t) = \exp(-\rho t) \cdot \big((-\alpha\rho + \beta\varphi)\cos(\varphi t) + (-\alpha\varphi - \beta\rho)\sin(\varphi t)\big)$. Thus, if x is a non-zero solution, then we have $\dot{x}(t) = 0$ whenever $t = \frac{2k\pi}{\varphi} \in \mathbb{R}$ for some $k \in \mathbb{Z}$, saying that the body changes direction infinitely often. Now, if $\rho > 0$ we have $\lim_{t\to\infty} x(t) = 0$, saying that the body ultimately tends to the point of equilibrium, in other words the body oscillates with decreasing **amplitude**; in contrast, if $\rho = 0$ the limit $\lim_{t\to\infty} x(t)$ does not exist, and the body oscillates with constant amplitude. In particular, letting $x(0) := 1$ and $\dot{x}(0) := 0$, we get $\alpha = x(0) = 1$ and $-\alpha\rho + \beta\varphi = \dot{x}(0) = 0$, yielding $\beta = \frac{\rho}{\varphi}$, that is $x(t) = \exp(-\rho t) \cdot \big(\cos(\varphi t) + \frac{\rho}{\varphi}\sin(\varphi t)\big)$, where $\varphi = \sqrt{\omega^2 - \rho^2}$; for $\rho = 0$ we get $\varphi = \omega$ and $x(t) = \cos(\omega t)$, saying that $\frac{\omega}{2\pi} > 0$ is the **frequency** of the free harmonic oscillator.

9 Normal forms

(9.1) Theorem: Smith algorithm. Let R be a principal ideal domain and let $A \in R^{m\times n}$, where $m, n \in \mathbb{N}_0$.

a) Then there are $S \in \mathrm{GL}_m(R)$ and $T \in \mathrm{GL}_n(R)$, and $d_1, \ldots, d_r \in R$, where $r := \min\{m, n\} \in \mathbb{N}_0$, such that $SAT = \operatorname{diag}[d_1, \ldots, d_r] \in R^{m\times n}$ is in **Smith normal form**, that is $d_i \mid d_{i+1}$ for all $i \in \{1, \ldots, r-1\}$.

b) The **invariant factors** $[d_1, \ldots, d_r]$ of A are unique up to taking associates.

Let $s := |\{i \in \{1, \ldots, r\}; d_i \neq 0\}|$ and $d_i \sim \prod_{p\in\mathcal{P}} p^{\nu_p(d_i)} \in R$ for $i \in \{1, \ldots, s\}$. Then $[p^{\nu_p(d_i)} \in R; i \in \{1, \ldots, s\}, p \in \mathcal{P}, \nu_p(d_i) \geq 1]$, and their associates, are called the **elementary divisors** of A.

Proof. a) We proceed by induction on $m + n \in \mathbb{N}_0$, where we assume that $m + n \geq 1$. Let $A = [a_{ij}]_{ij} \in R^{m\times n}$, where we may assume that $A \neq 0$.

i) If there are $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, n\}$ such that $a_{kl} \mid a_{ij}$ for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, then by interchanging rows 1 and k, and columns 1 and l we may assume that $0 \neq a_{11} \mid a_{ij}$ for all i and j. Hence adding suitable multiples of row 1 to the other rows, and suitable multiples of column 1 to the other columns, we may assume that $A = \left[\begin{array}{c|c} a_{11} & . \\ \hline . & A' \end{array}\right]$, where $A' = [a_{i+1,j+1}]_{ij} \in R^{(m-1)\times(n-1)}$ and $a_{11} \mid a_{ij}$ for all i and j. By induction A' can be transformed into Smith normal form, whose entries are divisible by a_{11}.

ii) Otherwise, let $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, n\}$ such that $0 \neq a_{kl} \in R$ has a factorisation of minimal length amongst all non-zero matrix entries. By applying suitable row and column operations we may assume that $k = l = 1$, hence we have $0 \neq a_{11} \in R \setminus R^*$.

We show that we may assume that $a_{11} \not| a_{21}$: If there is $i \in \{2, \ldots, m\}$ such that $a_{11} \not| a_{i1}$ this is achieved by interchanging rows 2 and i. If there is $j \in \{2, \ldots, n\}$ such that $a_{11} \not| a_{1j}$ this is achieved by interchanging columns 2 and j and going over to A^{tr}. If finally $a_{11} \mid a_{i1}, a_{1j}$ for all i and j, by applying suitable row and column operations we may assume that $A = \left[\begin{array}{c|c} a_{11} & . \\ \hline . & A' \end{array}\right]$, where $A' = [a_{i+1,j+1}]_{ij} \in R^{(m-1)\times(n-1)}$ has an entry a_{ij} such that $a_{11} \not| a_{ij}$, hence by applying suitable row and column operations we may assume that $a_{11} \not| a_{22}$, thus by adding column 2 to column 1 we may assume that $a_{11} \not| a_{21}$.

Let $a := a_{11}$ and $b := a_{21}$, and let $s, t \in R$ such that $d := sa + tb \in \gcd(a, b)$. Letting $a', b' \in R$ such that $a = a'd$ and $b = b'd$, we get $d = sa+tb = (sa'+tb')d$, and hence $sa' + tb' = 1$. Thus for $B := \left[\begin{array}{cc|c} s & t & . \\ -b' & a' & . \\ \hline . & . & E_{m-2} \end{array}\right] \in R^{m\times m}$ we have $\det(B) = sa' + tb' = 1$ and hence $B \in \mathrm{GL}_m(R)$. Moreover, the upper left entry of $BA \in R^{m\times n}$ is $sa + tb = d$. Since $0 \neq d \in R$ has a shorter factorisation than a has, after a finite number of steps we end up in case i).

b) For $\mathcal{I} \subseteq \{1, \ldots, m\}$ and $\mathcal{J} \subseteq \{1, \ldots, n\}$ such that $|\mathcal{I}| = |\mathcal{J}| = i$, where $i \in \{1, \ldots, r\}$, let $A_{\mathcal{I},\mathcal{J}} \in R^{i\times i}$ be the submatrix consisting of rows $\mathcal{I}$ and columns $\mathcal{J}$. Then $\det(A_{\mathcal{I},\mathcal{J}}) \in R$ is called an ***i*-minor** of A, and $\delta_i \in \gcd(\{\det(A_{\mathcal{I},\mathcal{J}}) \in R; |\mathcal{I}| = |\mathcal{J}| = i\})$ is called an i-th **determinantal divisor** of A.

Since the rows of SA are R-linear combinations of the rows of A, the i-minors of SA are R-linear combinations of the i-minors of A, hence $\delta_i(A)$ divides the i-minors of SA, thus $\delta_i(A) \mid \delta_i(SA)$, and from $S \in \mathrm{GL}_n(R)$ we infer $\delta_i(A) \sim \delta_i(SA)$. Similarly, by considering columns we have $\delta_i(A) \sim \delta_i(AT)$. Thus from $\delta_i(A) \sim \delta_i(SAT) \sim \prod_{j=1}^{i} d_j$, for all $i \in \{1, \ldots, r\}$, we get $\delta_1(A) \sim d_1$ and $\delta_i(A) \sim \delta_{i-1}(A) \cdot d_i$ for all $i \in \{2, \ldots, r\}$. ♯

(9.2) Corollary. If R is Euclidean, then A can be transformed into Smith normal form by row and column operations.

Proof. We only have to consider case ii), where we may choose a_{kl} to have minimal degree amongst all non-zero matrix entries. Again we may assume that $k = l = 1$ and that $a_{11} \not| \, a_{21}$. Then we have $a_{21} = qa_{11} + r$, where $q, r \in R$ such that $r \neq 0$ and for the degrees we have $\delta(r) < \delta(a_{11})$. Hence adding the $(-q)$-fold of row 1 to row 2, and interchanging rows 1 and 2, yields a matrix whose upper left entry r has a smaller degree than a_{11} has. Hence after a finite number of steps we end up in case i); note that if $\delta(r) = 0$ then $r \in R^*$. ♯

For example, $\begin{bmatrix} 3 & 4 & 5 \\ 3 & 2 & 1 \\ 7 & 6 & 7 \end{bmatrix} \in \mathbb{Z}^{3\times 3}$ has invariant factors $[d_1, d_2, d_3] \sim [1, 2, 6]$, its elementary divisors are $[2, 2; 3]$, and its determinantal divisors are $[1, 2, 12]$.

(9.3) Linear equations over principal ideal domains. Let R be a principal ideal domain, let $A \in R^{m\times n}$, where $m, n \in \mathbb{N}_0$, and let $w \in R^{m\times 1}$. Let $S \in \mathrm{GL}_m(R)$ and $T \in \mathrm{GL}_n(R)$ such that $SAT = D := \mathrm{diag}[d_1, \ldots, d_r] \in R^{m\times n}$ is in Smith normal form, where $r := \min\{m, n\} \in \mathbb{N}_0$. Then for $x \in R^{n\times 1}$ we have $Ax = w$ if and only if $D \cdot T^{-1}x = SATT^{-1}x = Sw$, hence we have $\mathcal{L}(A, w) = T \cdot \mathcal{L}(D, Sw) \subseteq R^{n\times 1}$. Letting $Sw = [c_1, \ldots, c_m]^{\mathrm{tr}} \in R^{m\times 1}$, we for $y = [y_1, \ldots, y_n]^{\mathrm{tr}} \in R^{n\times 1}$ have $y \in \mathcal{L}(D, Sw)$ if and only if $d_i y_i = c_i$ for all $i \in \{1, \ldots, r\}$; thus $\mathcal{L}(D, Sw) \neq \emptyset$ if and only if $d_i \mid c_i$ for all $i \in \{1, \ldots, r\}$.

For example: Five sailors are lost on an island, where they have to live from coconuts alone. They collect a number of them at the beach. But they do not trust each other. So at night the first of the sailors takes one fifth of the coconuts piled at the beach, buries them secretly, and throws one of the remaining ones into the sea. And so does successively the second, third, fourth and fifth of the sailors. In the morning there is number of coconuts divisible by five left on the beach. How many have there been at least the evening before?

Let $x_i \in \mathbb{N}_0$ be the number of coconuts the i-th sailors finds at the beach, for $i \in \{1, \ldots, 5\}$. Then we have $x_{i+1} = \frac{4}{5}x_i - 1$, that is $4x_i - 5x_{i+1} = 5$, for $i \in \{1, \ldots, 5\}$, and $x_6 = 5x_0$ for some $x_0 \in \mathbb{N}_0$. Then letting $x := [x_1, \ldots, x_6, x_0] \in \mathbb{Z}^7$ and $w := [5, 5, 5, 5, 5, 0] \in \mathbb{Z}^6$, we have to solve the system of linear equations $AX^{\mathrm{tr}} = w^{\mathrm{tr}}$ over $\mathbb{Z}$, where $X := [X_1, \ldots, X_6, X_0]$ and

$$A := \begin{bmatrix} 4 & -5 & . & . & . & . & . \\ . & 4 & -5 & . & . & . & . \\ . & . & 4 & -5 & . & . & . \\ . & . & . & 4 & -5 & . & . \\ . & . & . & . & 4 & -5 & . \\ . & . & . & . & . & 1 & -5 \end{bmatrix} \in \mathbb{Z}^{6\times 7}.$$

Then A has Smith normal form $SAT = D := \mathrm{diag}[1,1,1,1,1,1] \in \mathbb{Z}^{6\times 7}$, where

$$S := \begin{bmatrix} -205 & -51 & -269 & -131 & -369 & -1024 \\ -164 & -41 & -215 & -105 & -295 & -820 \\ -80 & -20 & -105 & -51 & -144 & -399 \\ -64 & -16 & -84 & -41 & -115 & -320 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -256 & -64 & -336 & -164 & -461 & -1281 \end{bmatrix} \in \mathbb{Z}^{6\times 6},$$

$$T := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -779 & 15625 \\ 1 & 1 & 0 & 0 & 0 & -624 & 12500 \\ 0 & 1 & 1 & 0 & 0 & -499 & 10000 \\ 0 & 0 & 1 & 1 & 0 & -399 & 8000 \\ 0 & 0 & 0 & 1 & 1 & -319 & 6400 \\ 0 & 0 & 0 & 0 & 1 & -255 & 5120 \\ 0 & 0 & 0 & 0 & 0 & -51 & 1024 \end{bmatrix} \in \mathbb{Z}^{7\times 7}.$$

Thus letting $y^{\mathrm{tr}} := T^{-1}x^{\mathrm{tr}} \in \mathbb{Z}^{1\times 7}$ we get $Dy^{\mathrm{tr}} = Sw^{\mathrm{tr}} \in \mathbb{Z}^{6\times 1}$, hence $y = [-5125, -4100, -2000, -1600, 0, -6405, y_0] \in \mathbb{Z}^7$, thus $x_1 = 4984370 + 15625y_0$, where $y_0 \in \mathbb{Z}$. Hence the smallest positive solution is given by $y_0 := -318$, yielding $x_1 := 15620$, that is $x = [15620, 12495, 9995, 7995, 6395, 5115, 1023]$. Thus there have been at least 15620 coconuts in the beginning. $\sharp$

(9.4) Theorem. Let R be a principal ideal domain, and let V be an R-module with an R-basis of cardinality $n \in \mathbb{N}_0$; then V is called **R-free**.
a) Any R-basis of V has cardinality $\mathrm{rk}_R(V) := n \in \mathbb{N}_0$, called the **$R$-rank** of V.
b) Any R-submodule of V has an R-basis of cardinality at most n.

Proof. a) Choosing R-bases of cardinality $n \in \mathbb{N}_0$ and $m \in \mathbb{N}_0$, respectively, we may assume that there is an R-isomorphism $\varphi\colon R^n \to R^m$. If R is a field, then $n = \dim_R(R^n) = \dim_R(R^m) = m$. Thus we assume that R is not a field.

Let $0 \neq p \in R \setminus R^*$ be irreducible. Then $pR \lhd R$ is a maximal proper ideal, thus $K := R/pR$ is a field. The natural map $\nu_p\colon R \to K$ yields R-epimorphisms $\nu_p^n\colon R^n \to K^n$ and $\nu_p^m\colon R^m \to K^m$ Since $(pR)^n \leq \ker(\nu_p^m\varphi)$ and $(pR)^m \leq \ker(\nu_p^n\varphi^{-1})$ there are induced R-linear, and thus K-linear, maps $\overline{\varphi}\colon K^n \to K^m$ and $\overline{\varphi^{-1}} = \overline{\varphi}^{-1}\colon K^m \to K^n$. This implies $n = \dim_K(K^n) = \dim_K(K^m) = m$.

b) Let $U \leq V$. We proceed by induction over $n \in \mathbb{N}_0$: We may assume that $V = R^n$ where $n \geq 1$. Then the standard R-basis of $R^{n-1} \cong \{[a_1, \ldots, a_{n-1}, 0] \in R^n; a_1, \ldots, a_{n-1} \in R\} \leq R^n$ has cardinality $n-1$, and thus by induction $U \cap R^{n-1} \leq R^{n-1}$ has an R-basis $\{v_1, \ldots, v_m\}$ where $m \leq n-1$. Hence we may assume that $U \not\leq R^{n-1}$, and let $I := \{a \in R; [a_1, \ldots, a_{n-1}, a] \in U \text{ for some } a_1, \ldots, a_{n-1} \in R\}$. Thus we have $\{0\} \neq I \unlhd R$, hence there is $0 \neq b \in R$ such that $I = bR$, and there are $b_1, \ldots, b_{n-1} \in R$ such that $v := [b_1, \ldots, b_{n-1}, b] \in U$. Then $\{v_1, \ldots, v_m, v\} \subseteq U$ is an R-basis, having cardinality $m + 1 \leq n$: For $w = [a_1, \ldots, a_n] \in U$ we have $a_n = bc$ for some $c \in R$, and thus $w - cv \in U \cap R^{n-1}$, thus $\langle v_1, \ldots, v_m, v\rangle_R = U$. If $av + \sum_{i=1}^m a_i v_i = 0$, where $a, a_i \in R$, then from $v_i \in R^{n-1}$ we get $av \in R^{n-1}$, thus $ab = 0$, implying $a = 0$, and hence $\sum_{i=1}^m a_i v_i = 0$ implies $a_i = 0$, for all $i \in \{1, \ldots, m\}$. $\sharp$

(9.5) Theorem: Structure of modules over principal ideal domains. Let R be a principal ideal domain, and let V be a finitely generated R-module.
a) There are $r, s \in \mathbb{N}_0$ and $0 \neq d_1, \ldots, d_s \in R \setminus R^*$, such that $d_i \mid d_{i+1}$ for all $i \in \{1, \ldots, s-1\}$ and $V \cong \bigoplus_{i=1}^s R/d_iR \oplus R^r$ as R-modules, where the **direct sum** denotes $R/d_1R \times \cdots \times R/d_sR \times R^r$ with componentwise operations. Moreover, if $d_i \sim \prod_{p \in \mathcal{P}} p^{\nu_p(d_i)} \in R$, then $R/d_iR \cong \bigoplus_{p \in \mathcal{P}, \nu_p(d_i) \geq 1} R/p^{\nu_p(d_i)}R$.
b) The numbers $r, s \in \mathbb{N}_0$ are uniquely determined, and the **torsion invariants** $[d_1, \ldots, d_s]$ are uniquely determined up to taking associates.

Proof. a) Let $\{v_1, \ldots, v_m\} \subseteq V$ be an R-generating set for some $m \in \mathbb{N}_0$. Letting $B := \{e_1, \ldots, e_m\} \subseteq R^m$ be the standard R-basis, we have the R-epimorphism $\varphi\colon R^m \to V\colon \sum_{i=1}^m a_i e_i \mapsto \sum_{i=1}^m a_i v_i$, where $\ker(\varphi) \leq R^m$ has an R-basis $C := \{w_1, \ldots, w_n\}$ for some $n \leq m$. Applying the Smith algorithm to ${}_B\mathrm{id}_C \in R^{m \times n}$ yields $S \in \mathrm{GL}_m(R)$ and $T \in \mathrm{GL}_n(R)$ and $d_1, \ldots, d_n \in R$ such that $d_i \mid d_{i+1}$ for all $i \in \{1, \ldots, n-1\}$ and $S \cdot {}_B\mathrm{id}_C \cdot T = \mathrm{diag}[d_1, \ldots, d_n] \in R^{m \times n}$.

Let $B' := \{u_1, \ldots, u_m\} \subseteq R^m$ be given by ${}_B\mathrm{id}_{B'} = S^{-1}$, then the surjectivity and injectivity of $\varphi_{S^{-1}} = \varphi_S^{-1} \in \mathrm{End}_R(R^{m \times 1})$ imply that $B' \subseteq R^m$ is an R-generating set and R-linearly independent, respectively, thus is an R-basis. Similarly, considering $\varphi_T \in \mathrm{End}_R(R^{n \times 1})$ shows that $C' := \{u'_1, \ldots, u'_n\} \subseteq \ker(\varphi)$ given by ${}_C\mathrm{id}_{C'} = T$ is an R-basis. We have ${}_{B'}\mathrm{id}_{C'} = {}_{B'}\mathrm{id}_B \cdot {}_B\mathrm{id}_C \cdot {}_C\mathrm{id}_{C'} = S \cdot {}_B\mathrm{id}_C \cdot T = \mathrm{diag}[d_1, \ldots, d_n] \in R^{m \times n}$, hence $u'_i = d_i u_i \in R^m$, implying $d_i \neq 0$ for all $i \in \{1, \ldots, n\}$, and thus $R^m / \ker(\varphi) = \langle u_1, \ldots, u_m\rangle_R / \langle d_1 u_1, \ldots, d_n u_n\rangle_R$.

For the R-epimorphism $\psi\colon R^m \to \bigoplus_{i=1}^n R/d_iR \oplus R^{m-n}\colon \sum_{i=1}^m a_i e_i \mapsto [a_1 + d_1R, \ldots, a_n + d_nR, a_{n+1}, \ldots, a_m]$ we have $\ker(\psi) = \langle d_1 e_1, \ldots, d_n e_n\rangle = \ker(\varphi)$. Since $R/d_iR \cong \{0\}$ if and only if $d_i \in R^*$, letting $s := |\{i \in \{1, \ldots, n\}; d_i \notin R^*\}|$ and $r := m - n$ we get an induced R-isomorphism $\overline{\psi}\overline{\varphi}^{-1}\colon V \to R^m/\ker(\varphi) = R^m/\ker(\psi) \to \bigoplus_{i=1}^s R/d_{n-s+i}R \oplus R^r$.

Finally, it is sufficient to show that $R/abR \cong R/aR \oplus R/bR$, where $0 \neq a, b \in R$ are coprime: The determinantal divisors of $\mathrm{diag}[a, b] \in R^{2 \times 2}$ are $\delta_1 \in \gcd(a, b) = R^*$ and $\delta_2 \sim \det(\mathrm{diag}[a, b]) = ab$, hence its invariant factors are $[1, ab]$.

b) For $0 \neq c \in R$ let $T_c(V) := \{v \in V; cv = 0\} \leq V$ be the **c-torsion** R-submodule, and let $T(V) := \langle T_c(V); 0 \neq c \in R\rangle = \{v \in V; cv = 0 \text{ for some } 0 \neq c \in R\} \leq V$ be the **torsion** R-submodule of V. Hence we have $T(R) = \{0\}$ and $T(R/dR) = R/dR$, where $0 \neq d \in R$.

For $V := \bigoplus_{i=1}^s R/d_iR \oplus R^r$ as above we have $T(V) = \bigoplus_{i=1}^s T(R/d_iR) \oplus \bigoplus_{i=1}^r T(R) \cong \bigoplus_{i=1}^s R/d_iR$. Moreover, we have the R-epimorphism $\pi\colon V \to R^r\colon [\overline{a_1}, \ldots, \overline{a_s}, a_{s+1}, \ldots, a_{s+r}] \mapsto [a_{s+1}, \ldots, a_{s+r}]$, where $\ker(\pi) = T(V)$. Hence $V/T(V) \cong R^r$ shows that $r = \mathrm{rk}_R(R^r) \in \mathbb{N}_0$ is determined by V. Thus we may assume that $V = T(V) := \bigoplus_{i=1}^s R/d_iR$ where $s \geq 1$; we proceed by induction on the length of the factorisation of $0 \neq d_s \in R \setminus R^*$:

If $p \in R$ is irreducible, then $pR \lhd R$ is a maximal proper ideal, thus $K := R/pR$ is a field and $T_p(V)$ is an K-vector space. For $0 \neq d \in R$ we have $T_p(R/dR) \neq \{0\}$ if and only if $p \mid d$: If $0 \neq \overline{a} \in T_p(R/dR)$, where $a \in R$, then $d \not\mid a$ and $d \mid pa$, implying $p \mid d$; if conversely $d = pd'$ for some $d' \in R$, then $d = pd' \mid pa$ if and only if $d' \mid a$, implying $T_p(R/dR) = d'R/dR \neq \{0\}$; in this case, the R-epimorphism $R \to d'R/dR \colon a \mapsto \overline{ad'}$, having kernel pR, shows $T_p(R/dR) \cong K$. Note that, letting $l := \nu_p(d) \in \mathbb{N}_0$ and $\widetilde{d} \in R$ such that $d = p^l\widetilde{d}$, we by induction from this get $T_{p^j}(R/dR) = p^{l-j}\widetilde{d}R/dR$, for $j \in \{0, \ldots, l\}$, and thus $T_{p^j}(R/dR)/T_{p^{j-1}}(R/dR) \cong p^{l-j}\widetilde{d}R/p^{l-j+1}\widetilde{d}R \cong R/pR = K$, for $j \in \{1, \ldots, l\}$,

Let now $p \mid d_s$. Then from $T_p(V) = \bigoplus_{i=1}^s T_p(R/d_iR)$, letting $t := \min\{i \in \{1, \ldots, s\}; p \mid d_i\} \in \mathbb{N}$ we get $s - t + 1 = \dim_K(T_p(V))$, showing that the number of d_i being divisible by p is determined by V. Moreover, letting $d_i = pd'_i$, where $d'_i \in R$ for $i \in \{t, \ldots, s\}$, from $(R/d_iR)/(d'_iR/d_iR) \cong R/d'_iR$ we get $V/T_p(V) = (\bigoplus_{i=1}^s R/d_iR)/(\bigoplus_{i=t}^s d'_iR/d_iR) \cong \bigoplus_{i=1}^{t-1} R/d_iR \oplus \bigoplus_{i=t}^s R/d'_iR$. Since d'_s has a shorter factorisation than d_s, the d_i for $i \in \{1, \ldots, t-1\}$ and the d'_i for $i \in \{t, \ldots, s\}$ are by induction determined by $V/T_p(V)$ and hence by V. $\sharp$

(9.6) $\mathbb{Z}$-modules. Let G be an additive abelian group. Then G is a $\mathbb{Z}$-module with scalar multiplication $\mathbb{Z} \times G \to G \colon [a, g] \mapsto ag$; conversely, any $\mathbb{Z}$-module is an additive abelian group. Thus any finitely generated additive abelian group is a finite direct sum of finite groups $\mathbb{Z}_d$, for various $1 \neq d \in \mathbb{N}$, and infinite groups $\mathbb{Z}$, and the **isomorphism classes** of these groups are given by the invariants specified in the above theorem.

For example, the isomorphism classes of abelian groups of cardinality $72 = 2^3 \cdot 3^2$ are given by the torsion invariants $[2, 2 \cdot 3, 2 \cdot 3]$, $[2 \cdot 3, 2^2 \cdot 3]$, $[3, 2^3 \cdot 3]$, $[2, 2, 2 \cdot 3^2]$, $[2, 2^2 \cdot 3^2]$, $[2^3 \cdot 3^2]$, yielding

$$\begin{array}{lcl}
\mathbb{Z}_2 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}_6 & \cong & (\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2) \oplus (\mathbb{Z}_3 \oplus \mathbb{Z}_3), \\
\mathbb{Z}_6 \oplus \mathbb{Z}_{12} & \cong & (\mathbb{Z}_2 \oplus \mathbb{Z}_4) \oplus (\mathbb{Z}_3 \oplus \mathbb{Z}_3), \\
\mathbb{Z}_3 \oplus \mathbb{Z}_{24} & \cong & \mathbb{Z}_8 \oplus (\mathbb{Z}_3 \oplus \mathbb{Z}_3), \\
\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_{18} & \cong & (\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2) \oplus \mathbb{Z}_9, \\
\mathbb{Z}_2 \oplus \mathbb{Z}_{36} & \cong & (\mathbb{Z}_2 \oplus \mathbb{Z}_4) \oplus \mathbb{Z}_9, \\
\mathbb{Z}_{72} & \cong & \mathbb{Z}_8 \oplus \mathbb{Z}_9.
\end{array}$$

(9.7) $K[X]$-modules. a) Let K be a field, let $V := K^{n \times 1}$ and let $A = [a_{ij}]_{ij} \in K^{n \times n}$, where $n \in \mathbb{N}_0$. Then V becomes a $K[X]$-module with scalar multiplication $K[X] \times V \to V \colon [f, v] \mapsto f(A)v$; we write V_A, the $K[X]$-module structure being given by $Xv := Av$ for all $v \in V$. The standard K-basis $\mathcal{B} := \{v_1, \ldots, v_n\} \subseteq V$ being a $K[X]$-generating set of V_A as well shows that V_A is a finitely generated $K[X]$-module.

If $A' \in K^{n \times n}$, then V also becomes a $K[X]$-module via $Xv := A'v$ for all $v \in V$. For any map $\varphi \colon V \to V$ we have $\varphi \in \mathrm{Hom}_{K[X]}(V_A, V_{A'})$ if and only if

$\varphi \in \mathrm{End}_K(V)$ and $\varphi\varphi_A = \varphi_{A'}\varphi$. In particular, we have $V_A \cong V_{A'}$ as $K[X]$-modules if and only if there is $P \in \mathrm{GL}_n(K)$ such that $PA = A'P$, that is if and only if A and A' are similar; thus the similarity classes of matrices in $K^{n\times n}$ coincide with the **isomorphism classes** of $K[X]$-module structures on V.

b) Let $B := \{e_1, \ldots, e_n\} \subseteq K[X]^{n\times 1}$ be the standard $K[X]$-basis, and let $\varphi\colon K[X]^{n\times 1} \to V_A$ be the $K[X]$-epimorphism given by $\varphi(e_i) := v_i$, for $i \in \{1, \ldots, n\}$. Then the $K[X]$-isomorphism $\overline{\varphi}\colon K[X]^{n\times 1}/\ker(\varphi) \to V_A$ is described by the characteristic matrix ${}_B\mathrm{id}_C = XE_n - A \in K[X]^{n\times n}$, that is $C := \{f_1, \ldots, f_n\}$ is a $K[X]$-basis of $\ker(\varphi) \le K[X]^{n\times 1}$, where $f_j := Xe_j - \sum_{i=1}^n a_{ij}e_i \in K[X]^n$, for $j \in \{1, \ldots, n\}$:

We have $\varphi(f_j) = Av_j - \sum_{i=1}^n a_{ij}v_i = 0 \in V_A$, for all $j \in \{1, \ldots, n\}$, hence $U := \langle C\rangle_{K[X]} \le \ker(\varphi)$. Conversely, let $g = \sum_{j=1}^n g_je_j \in \ker(\varphi)$, where $g_j \in K[X]$. Letting $\nu_U\colon K[X]^{n\times 1} \to K[X]^{n\times 1}/U$ be the natural $K[X]$-epimorphism, from $\nu_U(Xe_j) = \nu_U(\sum_{i=1}^n a_{ij}e_i) \in K[X]^{n\times 1}/U$ we get $\nu_U(g) = \nu_U(\sum_{j=1}^n g_je_j) = \nu_U(\sum_{i=1}^n a_ie_i) \in K[X]^{n\times 1}/U$, for suitable $a_i \in K$. Since $U \le \ker(\varphi)$ there is a $K[X]$-epimorphism $\widetilde{\varphi}\colon K[X]^{n\times 1}/U \to V_A$ such that $\varphi = \widetilde{\varphi}\nu_U$, implying $0 = \widetilde{\varphi}\nu_U(g) = \widetilde{\varphi}\nu_U(\sum_{i=1}^n a_ie_i) = \sum_{i=1}^n a_iv_i \in V_A$. Since $\mathcal{B}$ is K-linearly independent we get $a_i = 0$ for $i \in \{1, \ldots, n\}$, hence $\nu_U(g) = 0 \in K[X]^{n\times 1}/U$, that is $g \in U$.

Let $\sum_{j=1}^n g_jf_j = 0 \in K[X]^{n\times 1}$, where $g_j \in K[X]$. Hence we have $\sum_{j=1}^n Xg_je_j = \sum_{i=1}^n(\sum_{j=1}^n a_{ij}g_j)e_i \in K[X]^{n\times 1}$, thus since B is $K[X]$-linearly independent we get $Xg_j = \sum_{k=1}^n a_{jk}g_k \in K[X]$, for all $j \in \{1, \ldots, n\}$. Assume there is $g_j \neq 0$, chosen of maximum degree amongst the $\{g_k; k \in \{1, \ldots, n\}, g_k \neq 0\}$. Then since $a_{jk} \in K$, for all $k \in \{1, \ldots, n\}$, the j-th equation yields a contradiction. $\sharp$

Hence the matrices $A, A' \in K^{n\times n}$ are similar, if and only if the associated characteristic matrices $XE_n - A, XE_n - A' \in K[X]^{n\times n}$ have the same Smith normal form, which holds if and only if they are **equivalent**, that is there are $S, T \in \mathrm{GL}_n(K[X])$ such that $S(XE_n - A')T = XE_n - A$.

In this case, let $\psi\colon K[X]^{n\times 1} \to V_{A'}$ be the $K[X]$-epimorphism associated with A'. Hence we have $\ker(\varphi) = S \cdot \ker(\psi)$, thus $\varphi_S \in \mathrm{End}_{K[X]}(K[X]^{n\times 1})$ induces a $K[X]$-isomorphism $\overline{\varphi}_S\colon K[X]^{n\times 1}/\ker(\psi) \to K[X]^{n\times 1}/\ker(\varphi)$, and hence a $K[X]$-isomorphism $\overline{\varphi}\overline{\varphi}_S\overline{\psi}^{-1}\colon V_{A'} \to V_A$. Let $B' := S \cdot B \subseteq K[X]^{n\times 1}$, that is B' consists of the columns of S. Since $\psi(B) = \mathcal{B} \subseteq V_{A'}$ is a K-generating set, we infer that $\mathcal{B}' := \varphi(B') \subseteq V_A$ is a K-generating set, and thus a K-basis. Writing $S = \sum_{i\geq 0} X^iS_i \in K[X]^{n\times n}$, where $S_i \in K^{n\times n}$, shows that $\mathcal{B}'$ consists of the columns of $P := \sum_{i\geq 0} A^iS_i \in K^{n\times n}$, thus $P \in \mathrm{GL}_n(K)$ and $P^{-1}AP = A'$.

(9.8) Normal forms. a) We keep the setting of (9.7). Let A be equivalent to $D := \mathrm{diag}[1, \ldots, 1, d_1, \ldots, d_s] \in K[X]^{n\times n}$, where $s \in \mathbb{N}_0$ and the $d_i \in K[X] \setminus K$ are monic such that $d_i \mid d_{i+1}$ for all $i \in \{1, \ldots, s-1\}$. Hence we have $\prod_{i=1}^s d_i = \det(XE_n - A) = \chi_A \in K[X]$, thus $\sum_{i=1}^s \deg(d_i) = \deg(\chi_A) = n$.

Writing $d_i = \prod_{p\in\mathcal{P}} p^{\nu_p(d_i)} \in R$, then we have $\sum_{i=1}^s \sum_{p\in\mathcal{P}} \nu_p(d_i)\deg(p) = \sum_{i=1}^s \deg(d_i) = n$, thus the number of prime powers occurring is bounded above by n; note that it equals n if and only if all the d_i split into pairwise different linear factors, or equivalently d_s does so. Hence $D' := \operatorname{diag}[p^{\nu_p(d_i)}; i \in \{1, \ldots, s\}, p \in \mathcal{P}, \nu_p(d_i) \geq 1] \in K[X]^{n \times n}$ also has Smith normal form D, thus A is equivalent to D' as well. Hence as $K[X]$-modules we have

$$V_A \cong \bigoplus_{i=1}^{s} K[X]/d_iK[X] \cong \bigoplus_{i=1}^{s} \left(\bigoplus_{p \in \mathcal{P}, \nu_p(d_i) \geq 1} K[X]/p^{\nu_p(d_i)}K[X] \right).$$

For any $K[X]$-module $K[X]/fK[X]$, where $0 \neq f = X^n + \sum_{i=0}^{n-1} a_iX^i \in K[X]$ is monic, we have $\dim_K(K[X]/fK[X]) = n$, and scalar multiplication with X is given by the K-linear map $\lambda_X \colon \overline{g} \mapsto \overline{Xg}$, where for any $g \in K[X]$ we have $g(\lambda_X) = 0 \in \operatorname{End}_K(K[X]/fK[X])$ if and only if $f \mid g$, hence $\mu_{\lambda_X} = f$. Thus we infer that $\lambda_X \in \operatorname{End}_K(\bigoplus_{i=1}^{s} K[X]/d_iK[X])$ has minimum polynomial $\mu_{\lambda_X} \in \operatorname{lcm}(d_1, \ldots, d_s) \subseteq K[X]$, thus $\mu_{\lambda_X} = d_s$.

Hence we conclude that $\mu_A = d_s$ as well, and thus we recover the **Cayley-Hamilton-Frobenius Theorem**: We have $\mu_A \mid \chi_A \in K[X]$, and for any $p \in \mathcal{P}$ such that $p \mid \chi_A$ we also have $p \mid \mu_A$. Moreover, we recover the fact that A is diagonalisable if and only if μ_A splits into pairwise different linear factors.

The multiplicities $\nu_p(d_i) \in \mathbb{N}_0$, where $p \in \mathcal{P}$ and $i \in \{1, \ldots, s\}$, can also be found directly from V: For the field $L := K[X]/pK[X]$ we have $\dim_K(L) = \deg(p)$, thus for $j \in \mathbb{N}$ we have $\dim_K(T_{p^j}(A)/T_{p^{j-1}}(A)) = n_j(p)\deg(p)$, where $n_j(p) := \dim_L(T_{p^j}(A)/T_{p^{j-1}}(V)) = |\{i \in \{1, \ldots, s\}; \nu_p(d_i) \geq j\}|$.

b) A K-basis of $K[X]/fK[X]$ is given by $B(f) := \{\overline{X}^i \in K[X]/fK[X]; i \in \{0, \ldots, n-1\}\}$. From $X \cdot \overline{X}^{n-1} = \overline{X}^n = -\sum_{i=0}^{n-1} a_i\overline{X}^i$ we infer that $\lambda_X \in \operatorname{End}_K(K[X]/fK[X])$ is described by the **companion matrix**

$${}_{B(f)}(\lambda_X)_{B(f)} = C(f) = C_1(f) := \begin{bmatrix} . & . & . & \cdots & . & -a_0 \\ 1 & . & . & \cdots & . & -a_1 \\ . & 1 & . & \cdots & . & -a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ . & \cdots & . & 1 & . & -a_{n-2} \\ . & \cdots & . & . & 1 & -a_{n-1} \end{bmatrix} \in K^{n \times n}.$$

Moreover, for $l \in \mathbb{N}_0$ a K-basis of $K[X]/f^lK[X]$ is given by $B_l(f) := \{\overline{X}^i\overline{f}^j \in K[X]/f^lK[X]; i \in \{0, \ldots, n-1\}, j \in \{0, \ldots, l-1\}\}$; hence $B(f) = B_1(f)$. From $X \cdot \overline{X}^{n-1}\overline{f}^j = \overline{X}^n\overline{f}^j = (\overline{f} - \sum_{i=0}^{n-1} a_i\overline{X}^i)\overline{f}^j = \overline{f}^{j+1} - \sum_{i=0}^{n-1} a_i\overline{X}^i\overline{f}^j$, for $j \in \{0, \ldots, l-1\}$, we infer that $\lambda_X \in \operatorname{End}_K(K[X]/f^lK[X])$ is described by the

(generalised) Jordan matrix, using the matrix unit $E_{1n} \in K^{n \times n}$,

$$_{B_l(f)}(\lambda_X)_{B_l(f)} = C_l(f) := \begin{bmatrix} C(f) & . & . & . & \cdots & . \\ E_{1n} & C(f) & . & . & \cdots & . \\ . & E_{1n} & C(f) & . & \cdots & . \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ . & \cdots & . & E_{1n} & C(f) & . \\ . & \cdots & . & . & E_{1n} & C(f) \end{bmatrix} \in K^{nl \times nl}.$$

Thus A is similar to the block diagonal matrix $\bigoplus_{i=1}^s C(d_i) \in K^{n \times n}$, called the **Frobenius normal form** of A, as well as to the block diagonal matrix $\bigoplus_{i=1}^s \bigoplus_{p \in \mathcal{P}, \nu_p(d_i) \geq 1} C(p^{\nu_p(d_i)})$, called the **Weierstraß normal form** of A, and finally to the block diagonal matrix $\bigoplus_{i=1}^s \bigoplus_{p \in \mathcal{P}, \nu_p(d_i) \geq 1} C_{\nu_p(d_i)}(p)$, called the **(generalised) Jordan normal form** of A.

For example: Let $A \in \mathbb{Q}^{9 \times 9}$ such that $XE_9 - A \in \mathbb{Q}[X]^{9 \times 9}$ has non-constant invariant factors $d_1 := (X+1)(X^2+1) \in \mathbb{Q}[X]$ and $\mu_A = d_2 := (X+1)^2(X^2+1)^2 \in \mathbb{Q}[X]$, hence $\chi_A = d_1 d_2 = (X+1)^3(X^2+1)^3 \in \mathbb{Q}[X]$. The associated Frobenius, Weierstraß and Jordan normal forms are given as $C(d_1) \oplus C(d_2)$, and $\big(C(X+1) \oplus C(X^2+1)\big) \oplus \big(C((X+1)^2) \oplus C((X^2+1)^2)\big)$, as well as $\big(C_1(-1) \oplus C_1(X^2+1)\big) \oplus \big(C_2(-1) \oplus C_2(X^2+1)\big)$, respectively:

$$\left[\begin{array}{ccc|cccccc} . & . & -1 & . & . & . & . & . & . \\ 1 & . & -1 & . & . & . & . & . & . \\ . & 1 & -1 & . & . & . & . & . & . \\ \hline . & . & . & . & . & . & . & . & -1 \\ . & . & . & 1 & . & . & . & . & -2 \\ . & . & . & . & 1 & . & . & . & -3 \\ . & . & . & . & . & 1 & . & . & -4 \\ . & . & . & . & . & . & 1 & . & -3 \\ . & . & . & . & . & . & . & 1 & -2 \end{array}\right] \in \mathbb{Q}^{9 \times 9}$$

$$\left[\begin{array}{c|cc||cc|cccc} -1 & . & . & . & . & . & . & . & . \\ \hline . & . & -1 & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . & . \\ \hline\hline . & . & . & . & -1 & . & . & . & . \\ . & . & . & 1 & -2 & . & . & . & . \\ \hline . & . & . & . & . & . & . & . & -1 \\ . & . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & . & 1 & . & -2 \\ . & . & . & . & . & . & . & 1 & . \end{array}\right] \in \mathbb{Q}^{9 \times 9}$$

$$\left[\begin{array}{c|cc||cc|cc|cc}
-1 & . & . & . & . & . & . & . & . \\ \hline
. & . & -1 & . & . & . & . & . & . \\
. & 1 & . & . & . & . & . & . & . \\ \hline\hline
. & . & . & -1 & . & . & . & . & . \\
. & . & . & 1 & -1 & . & . & . & . \\ \hline
. & . & . & . & . & . & -1 & . & . \\
. & . & . & . & . & 1 & . & . & . \\ \hline
. & . & . & . & . & . & 1 & . & -1 \\
. & . & . & . & . & . & . & 1 & .
\end{array}\right] \in \mathbb{Q}^{9\times 9}$$

(9.9) Base changes. We keep the setting of (9.8). To find a K-basis reflecting Frobenius normal form, let $\psi\colon K[X]^{n\times 1} \to V_D$ be the $K[X]$-epimorphism associated with D, and let $S, T \in \mathrm{GL}_n(K[X])$ such that $XE_n - A = SDT \in K[X]^{n\times n}$. Hence we infer that $\varphi_S \in \mathrm{End}_{K[X]}(K[X]^{n\times 1})$ induces a $K[X]$-isomorphism $\overline{\varphi}\overline{\varphi_S}\overline{\psi}^{-1}\colon V_D \to K[X]^{n\times 1}/\ker(\psi) \to K[X]^{n\times 1}/\ker(\varphi) \to V_A$. Let $B_s := \{e_{n-s+1}, \ldots, e_n\} \subseteq B$ and $B'_s := S \cdot B_s \subseteq K[X]^{n\times 1}$, that is B'_s consists of columns $[n-s+1, \ldots, n]$ of S. Since $\psi(B_s) = \{v_{n-s+1}, \ldots, v_n\} \subseteq \mathcal{B} \subseteq V_D$ is a $K[X]$-generating set, $\mathcal{B}'_s = \{w_{n-s+1}, \ldots, w_n\} := \varphi(B'_s) \subseteq V_A$ is a $K[X]$-generating set as well. Writing $S = \sum_{i\geq 0} X^i S_i \in K[X]^{n\times n}$, where $S_i \in K^{n\times n}$, shows that $\mathcal{B}'_s$ consists of columns $[n-s+1, \ldots, n]$ of $\sum_{i\geq 0} A^i S_i \in K^{n\times n}$; since columns $[1, \ldots, n-s]$ of S are in $\ker(\varphi)$, these columns of $\sum_{i\geq 0} A^i S_i$ vanish. Moreover, $\langle \overline{e}_i\rangle_{K[X]} \leq K[X]^{n\times 1}/\ker(\psi)$ has K-basis $\overline{e}_i B(d_i)$, for $i \in \{n-s+1, \ldots, n\}$, thus $P := \coprod_{i=1}^{s} [A^j w_{n-s+i} \in V; j \in \{0, \ldots, \deg(d_i)-1\}] \in \mathrm{GL}_n(K)$ is such that $P^{-1}AP$ is in Frobenius normal form.

To find K-bases reflecting Weierstraß and Jordan normal forms, let $S', T' \in \mathrm{GL}_n(K[X])$ such that $XE_n - A = S'D'T' \in K[X]^{n\times n}$, write $S' = \sum_{i\geq 0} X^i S'_i \in K[X]^{n\times n}$, where $S'_i \in K^{n\times n}$, and let $\mathcal{B}' = \{w_{i,p} \in V; i \in \{1, \ldots, s\}, p \in \mathcal{P}, \nu_p(d_i) \geq 1\}$ consist of the non-vanishing columns of $\sum_{i\geq 0} A^i S'_i \in K^{n\times n}$. Then $P' := \coprod_{i=1}^{s} \coprod_{p\in\mathcal{P}, \nu_p(d_i)\geq 1} [A^j w_{i,p} \in V; j \in \{0, \ldots, \nu_p(d_i)\deg(p)-1\}] \in \mathrm{GL}_n(K)$ is such that $P'^{-1}AP'$ is in Weierstraß normal form, using the K-basis $B_{\nu_p(d_i)}(p)$ of $K[X]/p^{\nu_p(d_i)}K[X]$ instead yields $P'' := \coprod_{i=1}^{s} [A^j p(A)^k w_{n-s+i} \in V; j \in \{0, \ldots, \deg(p)-1\}, k \in \{0, \ldots, \nu_p(d_i)-1\}] \in \mathrm{GL}_n(K)$ such that $P''^{-1}AP''$ is in Jordan normal form.

For example: **i)** Let $A := \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{Q}^{2\times 2}$ and $A' := \mathrm{diag}[1, -1] \in \mathbb{Q}^{2\times 2}$; see (8.4). We have $S \cdot (XE_2 - A) \cdot T = \mathrm{diag}[1, X^2-1] = S' \cdot (XE_2 - A') \cdot T' \in \mathbb{Q}[X]^{2\times 2}$, where $S := \begin{bmatrix} . & 1 \\ 1 & X \end{bmatrix}$ and $T := \begin{bmatrix} -1 & X \\ . & 1 \end{bmatrix}$, and $S' := \begin{bmatrix} -1 & 1 \\ X+1 & -X+1 \end{bmatrix}$ and $T' := \frac{1}{2} \cdot \begin{bmatrix} 1 & X+1 \\ 1 & X-1 \end{bmatrix}$. Hence A and A' are similar, we have $\mu_A = \chi_A = d_1 = X^2 - 1 \in \mathbb{Q}[X]$, and the elementary divisors are $X-1$ and $X+1$.

A base change reflecting similarity is found as follows: From $S^{-1}S' = \begin{bmatrix} -X & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ X+1 & -X+1 \end{bmatrix} = \begin{bmatrix} 2X+1 & -2X+1 \\ -1 & 1 \end{bmatrix} = X \cdot \begin{bmatrix} 2 & -2 \\ . & . \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \in \mathbb{Q}[X]^{2\times 2}$ we get $P := A \cdot \begin{bmatrix} 2 & -2 \\ . & . \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Q})$ where $P^{-1}AP = A'$.

The matrix $A = C(X^2-1)$ already is in Frobenius normal form: From $S^{-1} = X \cdot \begin{bmatrix} -1 & . \\ . & . \end{bmatrix} + \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{Q}[X]^{2\times 2}$ we get $A \cdot \begin{bmatrix} -1 & . \\ . & . \end{bmatrix} + \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} = \begin{bmatrix} . & 1 \\ . & . \end{bmatrix} \in \mathbb{Q}^{2\times 2}$, hence $A \cdot [1,0]^{\mathrm{tr}} = [0,1]^{\mathrm{tr}}$ yields $P_{11} := E_2 \in \mathrm{GL}_2(\mathbb{Q})$ and $P_{11}^{-1}AP_{11} = A$.

The Frobenius normal form of A' is found as follows: $2 \cdot S'^{-1} = \begin{bmatrix} X-1 & 1 \\ X+1 & 1 \end{bmatrix} = X \cdot \begin{bmatrix} 1 & . \\ 1 & . \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{Q}[X]^{2\times 2}$ yields $A' \cdot \begin{bmatrix} 1 & . \\ 1 & . \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} . & 1 \\ . & 1 \end{bmatrix} \in \mathbb{Q}^{2\times 2}$, hence $A' \cdot [1,1]^{\mathrm{tr}} = [1,-1]^{\mathrm{tr}}$ yields $P_{21} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Q})$ and $P_{21}^{-1}A'P_{21} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \mathrm{diag}[1,-1] \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = A$.

The matrix $A' = C(X-1) \oplus C(X+1) = C_1(1) \oplus C_1(-1)$ already is in Weierstraß normal form, which coincides with its Jordan normal form: We already have $XE_2 - A' = \mathrm{diag}[X-1, X+1] \in \mathbb{Q}[X]^{2\times 2}$, thus the column transformation matrix being the identity matrix we get $P_{22} := E_2 \in \mathrm{GL}_2(\mathbb{Q})$ and $P_{22}^{-1}A'P_{22} = A'$. The Weierstraß normal form of A is found as follows: Since $XE_2 - A' = \mathrm{diag}[X-1, X+1]$ we use $S^{-1}S' \in \mathbb{Q}[X]^{2\times 2}$ again, yielding $P_{12} := P \in \mathrm{GL}_2(\mathbb{Q})$ and $P_{12}^{-1}AP_{12} = A'$.

ii) For example, let $A := \begin{bmatrix} 1 & -1 & 1 \\ 3 & 5 & -3 \\ 2 & 2 & 0 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$; see (8.9). Then $XE_3 - A \in \mathbb{Q}[X]^{3\times 3}$ has non-constant invariant factors $d_1 := X-2 \in \mathbb{Q}[X]$ and $\mu_A = d_2 := (X-2)^2 \in \mathbb{Q}[X]$, thus $\chi_A = d_1d_2 = (X-2)^3 \in \mathbb{Q}[X]$. Hence Frobenius, Weierstraß and Jordan normal forms are, respectively,

$$C(X-2) \oplus C((X-2)^2) = \left[\begin{array}{c|cc} 2 & . & . \\ \hline . & . & -4 \\ . & 1 & 4 \end{array}\right], \quad C_1(2) \oplus C_2(2) = \left[\begin{array}{c|cc} 2 & . & . \\ \hline . & 2 & . \\ . & 1 & 2 \end{array}\right].$$

Letting $S := \begin{bmatrix} 2 & 2 & X-5 \\ 2 & 4 & X-7 \\ 4 & 4 & 2X-8 \end{bmatrix}$ and $T := \frac{1}{4} \cdot \begin{bmatrix} 2 & -2 & -X^2+6X-9 \\ 0 & 2 & -X+5 \\ 0 & 0 & 2 \end{bmatrix}$ we have $S \cdot (XE_3 - A) \cdot T = D := \mathrm{diag}[1, X-2, (X-2)^2] \in \mathbb{Q}[X]^{3\times 3}$. Hence

$$4 \cdot S^{-1} = \begin{bmatrix} 2X-2 & -2 & -X+3 \\ -6 & 2 & 2 \\ -4 & 0 & 2 \end{bmatrix} = X \cdot \begin{bmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -2 & -2 & 3 \\ -6 & 2 & 2 \\ -4 & 0 & 2 \end{bmatrix}$$

yields $\begin{bmatrix} 1 & -1 & 1 \\ 3 & 5 & -3 \\ 2 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -2 & -2 & 3 \\ -6 & 2 & 2 \\ -4 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$,

thus from $A \cdot [2,-1,0]^{\mathrm{tr}} = [3,1,2]^{\mathrm{tr}}$ we get $P_1 := \begin{bmatrix} -2 & 2 & 3 \\ 2 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{Q})$ such that $P_1^{-1}AP_1$ is in Frobenius normal form, and from $(A-2E_3)\cdot[2,-1,0]^{\mathrm{tr}} = [-1,3,2]^{\mathrm{tr}}$ we get $P_2 := \begin{bmatrix} -2 & 2 & -1 \\ 2 & -1 & 3 \\ 0 & 0 & 2 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{Q})$ such that $P_2^{-1}AP_2$ is in Frobenius normal form.

10 Bilinear forms

(10.1) Adjoint matrices. Let K be a field, and let $\alpha\colon K \rightarrow K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$; for example, we may let $\alpha := \mathrm{id}_K$, and complex conjugation $\alpha := \bar{\ }\colon \mathbb{C} \rightarrow \mathbb{C}$ is a field automorphism such that $\alpha^2 = \mathrm{id}_{\mathbb{C}}$ and $\alpha \neq \mathrm{id}_{\mathbb{C}}$. Given K-vector spaces V and W, then V becomes a K-vector space V_α with respect to scalar multiplication given by $K \times V \rightarrow V\colon [a,v] \mapsto a^\alpha v$, and the maps in $\mathrm{Hom}_K(V_\alpha, W) = \mathrm{Hom}_K(V, W_\alpha)$ are called **α-semilinear**; note that we have $\dim_K(V) = \dim_K(V_\alpha) \in \mathbb{N}_0 \,\dot\cup\, \{\infty\}$.

For $m,n \in \mathbb{N}_0$ we have an α-semilinear map $K^{m\times n} \rightarrow K^{m\times n}\colon A = [a_{ij}]_{ij} \mapsto [a_{ij}^\alpha]_{ij} =: A^\alpha$, where A^α is called the associated **(α-)conjugate** matrix. Hence for $B \in K^{l\times m}$, where $l \in \mathbb{N}_0$, we have $(BA)^\alpha = B^\alpha A^\alpha \in K^{l\times n}$. For $A \in K^{n\times n}$ we have $\det(A^\alpha) = \det(A)^\alpha \in K$ and $\mathrm{adj}(A^\alpha) = \mathrm{adj}(A)^\alpha \in K^{n\times n}$, hence we have $\mathrm{rk}(A^\alpha) = \mathrm{rk}(A)$, in particular for $A \in \mathrm{GL}_n(K)$ we have $A^\alpha \in \mathrm{GL}_n(K)$, where $(A^\alpha)^{-1} = (A^{-1})^\alpha =: A^{-\alpha}$.

We have an α-semilinear map $K^{m\times n} \rightarrow K^{n\times m}\colon A \mapsto (A^\alpha)^{\mathrm{tr}} = (A^{\mathrm{tr}})^\alpha =: A^{\alpha\mathrm{tr}} = A^*$, where A^* is called the associated **(α-)adjoint** matrix. Hence for $B \in K^{l\times m}$ we have $(BA)^* = A^*B^* \in K^{n\times l}$. For $A \in K^{n\times n}$ we have $\det(A^*) = \det(A)^\alpha \in K$ and $\mathrm{adj}(A^*) = \mathrm{adj}(A)^* \in K^{n\times n}$, hence we have $\mathrm{rk}(A^*) = \mathrm{rk}(A)$, in particular for $A \in \mathrm{GL}_n(K)$ we have $A^* \in \mathrm{GL}_n(K)$, where $(A^*)^{-1} = (A^{-1})^* =: A^{-*} = A^{-\alpha\mathrm{tr}}$. Then $A \in K^{n\times n}$ is called **normal** if $AA^* = A^*A$, **hermitian** or **self-adjoint** if $A^* = A$, **skew-hermitian** if $A^* = -A$, and $A \in \mathrm{GL}_n(K)$ is called **unitary** if $A^* = A^{-1}$; if $\alpha = \mathrm{id}_K$ we in the latter cases have $A^{\mathrm{tr}} = A$ and $A^{\mathrm{tr}} = -A$ and $A^{\mathrm{tr}} = A^{-1}$, and A is called **symmetric** and **symplectic** and **orthogonal**, respectively.

(10.2) Sesquilinear forms. a) Let K be a field, and let $\alpha\colon K \rightarrow K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$. Given a K-vector space V, a map $\Phi\colon V \times V \rightarrow K\colon [v,w] \mapsto \langle v,w\rangle$ being K-linear in the second component, that is $\langle v, w+w'\rangle = \langle v,w\rangle + \langle v,w'\rangle$ and $\langle v, aw\rangle = a\langle v,w\rangle$, and α-semilinear in the first component, that is $\langle v+v', w\rangle = \langle v,w\rangle + \langle v',w\rangle$ and $\langle av, w\rangle = a^\alpha\langle v,w\rangle$, for all $v,v',w,w' \in V$ and $a \in K$, is called an **α-sesquilinear form** on V; if $\alpha = \mathrm{id}_K$

then Φ is K-linear in the first component as well and called a **K-bilinear form**.

An α-sesquilinear form Φ is called **hermitian** if $\langle w, v\rangle = \langle v, w\rangle^\alpha$ holds, and called **skew-hermitian** if $\langle w, v\rangle = -\langle v, w\rangle^\alpha$ holds, for all $v, w \in V$; if $\alpha = \mathrm{id}_K$ we in these cases have $\langle w, v\rangle = \langle v, w\rangle$ and $\langle w, v\rangle = -\langle v, w\rangle$, respectively, and Φ is called **symmetric** and **symplectic**, respectively.

b) Given $w \in V$, a vector $v \in V$ is called **right** and **left orthogonal** to w if $\langle w, v\rangle = 0$ and $\langle v, w\rangle = 0$, respectively; we write $w \perp v$ and $v \perp w$, respectively. Given $S \subseteq V$, then $S^\perp := \{v \in V; \langle w, v\rangle = 0 \text{ for all } w \in S\} \leq V$ and ${}^\perp S := \{v \in V; \langle v, w\rangle = 0 \text{ for all } w \in S\} \leq V$ are called the **right** and **left orthogonal spaces** of S, respectively; we have $S^\perp = (\langle S\rangle_K)^\perp$ and ${}^\perp S = {}^\perp(\langle S\rangle_K)$, and if Φ is (skew-)hermitian then we have $S^\perp = {}^\perp S$. In particular, $V^\perp$ and ${}^\perp V$ are called the **right** and **left radical** of Φ, respectively, and Φ is called **non-degenerate** if $V^\perp = \{0\} = {}^\perp V$.

For example, the **standard** α-sesquilinear form $\Gamma\colon K^{n\times 1} \times K^{n\times 1} \to K$ is given as $\langle [a_1, \ldots, a_n]^{\mathrm{tr}}, [b_1, \ldots, b_n]^{\mathrm{tr}}\rangle := [a_1, \ldots, a_n]^\alpha \cdot [b_1, \ldots, b_n]^{\mathrm{tr}} = \sum_{i=1}^n a_i^\alpha b_i$, for $n \in \mathbb{N}_0$. Then $\langle [b_1, \ldots, b_n]^{\mathrm{tr}}, [a_1, \ldots, a_n]^{\mathrm{tr}}\rangle = \sum_{i=1}^n b_i^\alpha a_i = \sum_{i=1}^n b_i^\alpha a_i^{\alpha^2} = (\sum_{i=1}^n a_i^\alpha b_i)^\alpha = \langle [a_1, \ldots, a_n]^{\mathrm{tr}}, [b_1, \ldots, b_n]^{\mathrm{tr}}\rangle^\alpha$ shows that Γ is hermitian, and since for $[a_1, \ldots, a_n]^{\mathrm{tr}} \in {}^\perp(K^{n\times 1})$ we get $0 = \langle e_i, [a_1, \ldots, a_n]^{\mathrm{tr}}\rangle = a_i$, for all $i \in \{1, \ldots, n\}$, hence Γ is non-degenerate. Finally, we have $\langle e_i, e_j\rangle = 0$ for all $i \neq j \in \{1, \ldots, n\}$, that is the standard K-basis is an **orthogonal** K-basis, and since $\langle e_i, e_i\rangle = 1$ for all $i \in \{1, \ldots, n\}$ it is even an **orthonormal** K-basis.

Let $\Gamma^{(n-1,1)}$ be the **Minkowski** α-sesquilinear form on $K^{n\times 1}$, for $n \in \mathbb{N}$, defined by $\langle [a_0, a_1, \ldots, a_{n-1}]^{\mathrm{tr}}, [b_0, b_1, \ldots, b_{n-1}]^{\mathrm{tr}}\rangle := -a_0^\alpha b_0 + \sum_{i=1}^{n-1} a_i^\alpha b_i$. Then we have $\langle [a_0, \ldots, a_{n-1}]^{\mathrm{tr}}, [b_0, \ldots, b_{n-1}]^{\mathrm{tr}}\rangle = \langle [b_0, \ldots, b_{n-1}]^{\mathrm{tr}}, [a_0, \ldots, a_{n-1}]^{\mathrm{tr}}\rangle^\alpha$, hence $\Gamma^{(n-1,1)}$ is hermitian, and from $[a_0, \ldots, a_{n-1}]^{\mathrm{tr}} \in {}^\perp\mathbb{R}^{n\times 1}$ we get $0 = \langle e_i, [a_0, \ldots, a_{n-1}]^{\mathrm{tr}}\rangle = a_i$, for $i \in \{1, \ldots, n-1\}$, and $0 = \langle e_0, [a_0, \ldots, a_{n-1}]^{\mathrm{tr}}\rangle = -a_0$, hence $\Gamma^{(n-1,1)}$ is non-degenerate. We have $\langle e_i, e_j\rangle = 0$ for $i \neq j$, hence the standard K-basis $\{e_0, \ldots, e_{n-1}\} \subseteq K^{n\times 1}$ is an orthogonal K-basis, where $\langle e_i, e_i\rangle = 1$ for $i \in \{1, \ldots, n\}$, but $\langle e_0, e_0\rangle = -1$. Thus for $n \geq 2$ there are isotropic vectors; for example $\langle [1, 0, \ldots, 0, 1]^{\mathrm{tr}}, [1, 0, \ldots, 0, 1]^{\mathrm{tr}}\rangle = -1 + 1 = 0$.

c) If $0 \neq v \in V$ such that $v \perp v$, that is $\langle v, v\rangle = 0$, or in other words $v \in \langle v\rangle_K^\perp \cap {}^\perp\langle v\rangle_K$, then v is called **isotropic**; if there are no isotropic vectors then Φ is called **anisotropic**. Note that any anisotropic form fulfills $V^\perp = V \cap V^\perp = \{0\} = V \cap {}^\perp V = {}^\perp V$, hence is non-degenerate.

We show that for any hermitian α-sesquilinear form $\Phi \neq 0$ there is a non-isotropic vector, unless $2 = 0 \in K$ and $\alpha = \mathrm{id}_K$: Assume that Φ is **totally isotropic**, that is $\langle v, v\rangle = 0$ for all $v \in V$. Since $V^\perp < V$, there are $v, w \in V$ such that $\langle v, w\rangle = 1$. Hence for all $a \in K$ we have $0 = \langle v + aw, v + aw\rangle = \langle v, v\rangle + a\langle v, w\rangle + a^\alpha\langle w, v\rangle + aa^\alpha\langle w, w\rangle = a + a^\alpha$, and thus $\alpha = -\mathrm{id}_K$, from which $1 = 1^\alpha = -1 \in K$ shows $2 = 0 \in K$ and $\alpha = \mathrm{id}_K$. For example, the symmetric $\mathbb{Z}_2$-bilinear form on $\mathbb{Z}_2^{2\times 1}$ given by $\langle [a, b]^{\mathrm{tr}}, [c, d]^{\mathrm{tr}}\rangle := ad + bc$ indeed is totally isotropic: We have $\langle [a, b]^{\mathrm{tr}}, [a, b]^{\mathrm{tr}}\rangle = ab + ba = 0$, for all $a, b \in \mathbb{Z}_2$.

(10.3) Gram matrices. a) Let K be a field, let $\alpha\colon K \to K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$, let V be finitely generated K-vector space with K-bases $B := [v_1, \ldots, v_n]$ and $C := [w_1, \ldots, w_n]$, where $n := \dim_K(V) \in \mathbb{N}_0$, and let Φ be an α-sesquilinear form on V. Then for $v = \sum_{i=1}^n a_i v_i \in V$ and $w = \sum_{j=1}^n b_j w_j$, where $a_i, b_j \in K$, we have $\langle v, w\rangle = \langle \sum_{i=1}^n a_i v_i, \sum_{j=1}^n b_j w_j\rangle = \sum_{i=1}^n \sum_{j=1}^n a_i^\alpha b_j \langle v_i, w_j\rangle \in K$. Thus letting ${}_B\Phi_C := [\langle v_i, w_j\rangle]_{ij} \in K^{n\times n}$ be the **Gram matrix** of Φ with respect to the K-bases B and C, using the coordinate tuples ${}_Bv = [a_1, \ldots, a_n]^{\mathrm{tr}} \in K^{n\times 1}$ and ${}_Cw = [b_1, \ldots, b_n]^{\mathrm{tr}} \in K^{n\times 1}$ we get $\langle v, w\rangle = ({}_Bv)^* \cdot {}_B\Phi_C \cdot {}_Cw \in K$.

For example, for the standard α-sesquilinear form Γ on $K^{n\times 1}$ with respect to the standard K-basis $B \subseteq K^{n\times 1}$ we get ${}_B\Gamma_B = E_n$; and for the Minkowski α-sesquilinear form $\Gamma^{(n-1,1)}$ on $K^{n\times 1}$ with respect to the standard K-basis $B \subseteq K^{n\times 1}$ we get ${}_B(\Gamma^{(n-1,1)})_B = \mathrm{diag}[-1, 1, \ldots, 1] \in K^{n\times n}$.

Hence Φ is uniquely determined by ${}_B\Phi_C$. Conversely, for any $A \in K^{n\times n}$ letting $\langle v, w\rangle := ({}_Bv)^* \cdot A \cdot {}_Cw \in K$ defines an α-sesquilinear form on V. Thus the set of all α-sesquilinear forms on V, being a K-vector space with respect to pointwise addition and scalar multiplication, is K-isomorphic to $K^{n\times n}$ via $\Phi \mapsto {}_B\Phi_C$.

If $B' := \{v'_1, \ldots, v'_n\}$ and $C' := \{w'_1, \ldots, w'_n\}$ are also K-bases of V, then for $i, j \in \{1, \ldots, n\}$ we have $\langle v'_i, w'_j\rangle = ({}_B(v'_i))^* \cdot {}_B\Phi_C \cdot {}_C(w'_j) \in K$, where ${}_C(w'_j) \in K^{n\times 1}$ is column j of ${}_C\mathrm{id}_{C'} \in K^{n\times n}$, and ${}_B(v'_i) \in K^{n\times 1}$ is column i of ${}_B\mathrm{id}_{B'} \in K^{n\times n}$, thus ${}_{B'}\Phi_{C'} := [\langle v'_i, w'_j\rangle]_{ij} = ({}_B\mathrm{id}_{B'})^* \cdot {}_B\Phi_C \cdot {}_C\mathrm{id}_{C'} \in K^{n\times n}$.

In particular, Φ is hermitian if and only if ${}_B\Phi_B = [\langle v_i, v_j\rangle]_{ij} = [\langle v_j, v_i\rangle^\alpha]_{ij} = [\langle v_i, v_j\rangle]_{ji}^\alpha = ({}_B\Phi_B)^* \in K^{n\times n}$, that is ${}_B\Phi_B$ is hermitian; similarly Φ is skew-hermitian if and only if ${}_B\Phi_B$ is skew-hermitian. Moreover, if B is an orthonormal K-basis, that is ${}_B\Phi_B = E_n$, and C is any K-basis, then C is orthonormal if and only if $E_n = {}_C\Phi_C = ({}_B\mathrm{id}_C)^* \cdot {}_B\Phi_B \cdot {}_B\mathrm{id}_C = ({}_B\mathrm{id}_C)^* \cdot {}_B\mathrm{id}_C$, thus if and only if $({}_B\mathrm{id}_C)^* = ({}_B\mathrm{id}_C)^{-1} \in \mathrm{GL}_n(K)$, that is ${}_B\mathrm{id}_C$ is unitary.

Moreover, this leads to the following notion: If Φ' also is an α-sesquilinear form on V, then Φ and Φ' are called **equivalent**, if there is a K-basis $B' \subseteq V$ such that ${}_{B'}\Phi'_{B'} = {}_B\Phi_B$, in other words if and only if there is $P \in \mathrm{GL}_n(K)$ such that ${}_B\Phi'_B = P^* \cdot {}_B\Phi_B \cdot P$; note that this is an equivalence relation on $K^{n\times n}$.

b) Let $B \subseteq V$ be a K-basis; then we may identify $V \to K^{n\times 1}\colon v \mapsto {}_Bv$. Letting $G := {}_B\Phi_B \in K^{n\times n}$, we have $\langle v, w\rangle = v^*Gw \in K$, for all $v, w \in V$.

Let $U \leq V$ be given as the image of $P \in K^{n\times m}$, where $m := \dim_K(U) \in \mathbb{N}_0$, and let $U^\perp \leq V$ and ${}^\perp U \leq V$ be given as the image of $Q' \in K^{n\times m'}$ and $Q'' \in K^{n\times m''}$, respectively, where $m' := \dim_K(U^\perp) \in \mathbb{N}_0$ and $m'' := \dim_K({}^\perp U) \in \mathbb{N}_0$. Then $P^*GQ' = 0 \in K^{m\times m'}$ implies $U^\perp = \ker(P^*G)$, and $Q''^*GP = 0 \in K^{m''\times m}$ implies ${}^\perp U = \ker((GP)^*) = \ker(P^*G^*)$. In particular we have $V^\perp = \ker(G)$ and ${}^\perp V = \ker(G^*)$, thus $\dim_K(V^\perp) = n - \mathrm{rk}(G) = n - \mathrm{rk}(G^*) = \dim_K({}^\perp V) \in \mathbb{N}_0$; hence Φ is non-degenerate if and only if $V^\perp = \{0\}$, if and only if ${}^\perp V = \{0\}$, if and only if $G \in \mathrm{GL}_n(K)$.

Considering the induced maps $\varphi_G \in \mathrm{End}_K(V)$ and $\varphi_{G^*} \in \mathrm{End}_K(V)$, we have

$\ker(GP) = \ker(\varphi_G|_U) = U \cap V^\perp$ and $\ker(G^*P) = \ker(\varphi_{G^*}|_U) = U \cap {}^\perp V$. This yields $m' = \dim_K(\ker(P^*G)) = n - \mathrm{rk}(P^*G) = n - \mathrm{rk}(G^*P) = n - m + \dim_K(\ker(G^*P)) = n - m + \dim_K(U \cap {}^\perp V)$ and $m'' = \dim_K(\ker(P^*G^*)) = n - \mathrm{rk}(P^*G^*) = n - \mathrm{rk}(GP) = n - m + \dim_K(\ker(GP)) = n - m + \dim_K(U \cap V^\perp)$.

In particular, if Φ is non-degenerate then we get $m' = n - m = m''$, and thus from $U \leq {}^\perp(U^\perp)$ and $U \leq ({}^\perp U)^\perp$ we infer $U = {}^\perp(U^\perp) = ({}^\perp U)^\perp$, that is U is **saturated**. Moreover, if Φ is even anisotropic, then we have $U \cap U^\perp = \{0\} = U \cap {}^\perp U$, thus we have the direct sum decompositions $V = U \oplus U^\perp = U \oplus {}^\perp U$, that is U has both a **right orthogonal** and a **left orthogonal** complement.

c) As an example, we consider $V := \mathbb{R}^{2\times 1}$ equipped with the standard $\mathbb{R}$-bilinear form Γ, which is symmetric and anisotropic. With respect to the standard $\mathbb{R}$-basis $B \subseteq V$ the associated Gram matrix is given as $G := {}_B\Gamma_B = E_2 \in \mathbb{R}^{2\times 2}$, reflecting the orthonormality of B; moreover, from $G = G^{\mathrm{tr}}$ and $\mathrm{rk}(G) = 2$ we recover the facts that Γ is symmetric and non-degenerate.

Let $v := [1,1]^{\mathrm{tr}} \in V$ and $U := \langle v\rangle_{\mathbb{R}}$. Then from $V = U \oplus U^\perp$ we get $\dim_{\mathbb{R}}(U^\perp) = 1$ and $U \cap U^\perp = \{0\}$. Indeed, letting $P := [v] \in \mathbb{R}^{2\times 1}$ we have $U^\perp = \ker(P^*G) = \ker(P^{\mathrm{tr}}G) = \ker([[1,1]] \cdot E_2) = \ker([[1,1]]) = \langle w\rangle_{\mathbb{R}}$, where $w := [-1,1]^{\mathrm{tr}} \in V$. Thus we infer that $C := [v,w] \subseteq V$ is an orthogonal $\mathbb{R}$-basis.

Letting $Q := {}_B\mathrm{id}_C = [v,w] \in \mathbb{R}^{2\times 2}$ be the associated base change matrix, we get ${}_C\Gamma_C = Q^*GQ = Q^{\mathrm{tr}}GQ = Q^{\mathrm{tr}}Q = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & . \\ . & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, where the diagonality of the latter matrix reflects the orthogonality of C, and the diagonal entries say that $\langle v,v\rangle = 2 = \langle w,w\rangle$.

Going over to unit vectors $v' := \frac{1}{\sqrt{2}} \cdot v \in V$ and $w' := \frac{1}{\sqrt{2}} \cdot w \in V$ yields the orthonormal $\mathbb{R}$-basis $C' := [v',w'] \subseteq V$, with associated base change matrix $Q' := {}_B\mathrm{id}_{C'} = [v',w'] = \frac{1}{\sqrt{2}} \cdot Q \in \mathbb{R}^{2\times 2}$. From this we get ${}_{C'}\Gamma_{C'} = Q'^{\mathrm{tr}}GQ' = Q'^{\mathrm{tr}}Q' = \frac{1}{2} \cdot Q^{\mathrm{tr}}Q = E_2 \in \mathbb{R}^{2\times 2}$, saying again that C' is orthonormal, and that Q' indeed is an orthogonal matrix; note that in order to go over to unit vectors we have to extract square roots.

(10.4) Orthogonalisation. a) Let K be a field, let $\alpha\colon K \to K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$, let V be finitely generated K-vector space, and let Φ be a hermitian α-sesquilinear form on V, where if $\alpha = \mathrm{id}_K$ we additionally assume that $2 \neq 0 \in K$. Then V actually has an orthogonal K-basis; note that orthogonal K-bases possibly exist only if Φ is hermitian:

We proceed by induction on $n := \dim_K(V) \in \mathbb{N}_0$, where the case $n = 0$ is trivial. Hence we may assume that $n \geq 1$ and $\Phi \neq 0$. Hence there is $v \in V$ such that $\langle v,v\rangle \neq 0$. Letting $U := \langle v\rangle_K$, then $v \notin U^\perp$ shows $U \cap U^\perp = \{0\}$, thus we have $U \cap V^\perp = \{0\}$, implying $\dim_K(U^\perp) = n - \dim_K(U) = n - 1$, hence $V = U \oplus U^\perp$.

A K-basis reflecting the direct sum decomposition $V = U \oplus U^\perp$ is found as follows: Let $B := [v, v_2, \ldots, v_n] \subseteq V$ be any K-basis containing v, and for $i \in \{2,\ldots,n\}$ let $w_i := v_i - \frac{\langle v,v_i\rangle}{\langle v,v\rangle} \cdot v \in V$. Thus $C := [v, w_2, \ldots, w_n]$ is a K-basis of

V as well, where $\langle v, w_i\rangle = \langle v, v_i\rangle - \frac{\langle v,v_i\rangle}{\langle v,v\rangle} \cdot \langle v, v\rangle = 0$ shows that $C' := [w_2, \ldots, w_n]$ is a K-basis of $U^\perp$. Note that $P := {}_B\mathrm{id}_{B'} = E_n - \sum_{i=2}^n \frac{\langle v,v_i\rangle}{\langle v,v\rangle} E_{1i} \in K^{n\times n}$ is an upper triangular matrix, and ${}_C\Phi_C = P^* \cdot {}_B\Phi_B \cdot P = [\langle v, v\rangle] \oplus {}_{C'}(\Phi|_{U^\perp \times U^\perp})_{C'} \in K^{n\times n}$ is found from ${}_B\Phi_B$ by subtracting the $\frac{\langle v,v_i\rangle}{\langle v,v\rangle}$-fold of column 1 from column i, and subtracting the $\frac{\langle v,v_i\rangle}{\langle v,v\rangle}$-fold of row 1 from row i, for all $i \in \{2, \ldots, n\}$.

For example, Let $K := \mathbb{R}$ and $\alpha = \mathrm{id}$, and let Φ be given with respect to some $\mathbb{R}$-basis $B \subseteq \mathbb{R}^{3\times 1}$ by ${}_B\Phi_B := \begin{bmatrix} 0 & -2 & 4 \\ -2 & 1 & -1 \\ 4 & -1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$. Hence we may choose the second basis vector as a non-isotropic vector to begin with, and letting $P_1 := \begin{bmatrix} . & 1 & . \\ 1 & . & . \\ . & . & 1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$ we get $G_1 = P_1^{\mathrm{tr}} G P_1 = \begin{bmatrix} 1 & -2 & -1 \\ -2 & 0 & 4 \\ -1 & 4 & 0 \end{bmatrix}$. Then, letting $P_2 := \begin{bmatrix} 1 & 2 & 1 \\ . & 1 & . \\ . & . & 1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$ yields $G_2 = P_2^{\mathrm{tr}} G_1 P_2 = \begin{bmatrix} 1 & . & . \\ . & -4 & 2 \\ . & 2 & -1 \end{bmatrix}$. Next, letting $P_3 := \begin{bmatrix} 1 & . & . \\ . & 1 & \frac{1}{2} \\ . & . & 1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$ we get $G_3 = P_3^{\mathrm{tr}} G_2 P_3 = \begin{bmatrix} 1 & . & . \\ . & -4 & . \\ . & . & . \end{bmatrix}$. Finally, rescaling with $P_4 := \begin{bmatrix} 1 & . & . \\ . & \frac{1}{2} & . \\ . & . & 1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$ yields $G' = P_4^{\mathrm{tr}} G_3 P_4 = \begin{bmatrix} 1 & . & . \\ . & -1 & . \\ . & . & . \end{bmatrix}$. Hence we have ${}_C\Phi_C = G' = P^{\mathrm{tr}} G P \in \mathbb{R}^{3\times 3}$, where the $\mathbb{R}$-basis $C \subseteq V$ is given as ${}_B\mathrm{id}_C = P := P_1P_2P_3P_4 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{GL}_3(\mathbb{R})$.

b) For $[K, \alpha] \in \{[\mathbb{R}, \mathrm{id}], [\mathbb{C}, \overline{\ \ }]\}$ we have **Sylvester's Theorem of Inertia**: An orthogonal K-basis $C \subseteq V$ can be chosen such that ${}_C\Phi_C = E_k \oplus (-E_l) \oplus (0 \cdot E_{n-k-l})$, where $k, l \in \mathbb{N}_0$ are independent from the particular choice of C:

The existence of C follows by replacing the non-isotropic elements v of an orthogonal K-basis as above by $\frac{1}{\sqrt{|\langle v,v\rangle|}} \cdot v$. To show uniqueness, for $\epsilon \in \{0, \pm 1\}$ let $C_\epsilon := \{v \in C, \langle v, v\rangle = \epsilon\}$ and $V_\epsilon := \langle C_\epsilon\rangle_K \leq V$, thus we have $V = \bigoplus_{\epsilon\in\{0,\pm1\}} V_\epsilon$ where $k = \dim_K(V_1)$ and $l = \dim_K(V_{-1})$ and $V_0 = V^\perp$, implying that $m := n - k - l = \dim_K(V_0) = \dim_K(V^\perp) \in \mathbb{N}_0$ is uniquely determined. For $w = \sum_{v\in C_\epsilon} a_v v \in V_\epsilon$ we have $\langle w, w\rangle = \epsilon \cdot \sum_{v\in C_\epsilon} |a_v|^2$, thus $\langle w, w\rangle \leq 0$ for $w \in V_{-1} \oplus V_0$, and $\langle w, w\rangle > 0$ for $0 \neq w \in V_1$. Let $C' \subseteq V$ be a K-basis such that ${}_{C'}\Phi_{C'} = E_{k'} \oplus (-E_{l'}) \oplus (0 \cdot E_m)$ with associated K-subspaces V'_ϵ, then we have $V_1 \cap (V'_{-1} \oplus V'_0) = 0$, implying $k + l' + m = \dim_K(V_1 + V'_{-1} + V'_0) \leq n = k' + l' + m$, thus $k \leq k'$; similarly, interchanging the roles of B and C we get $k' \leq k$. ♯

The uniquely defined pair $[k, l]$ is called the **signature** of Φ. Hence the equivalence classes of hermitian α-sesquilinear forms on V are described by the various signatures $[k, l]$, where $k, l \geq 0$ such that $k + l \leq n = \dim_K(V)$. In particular, $(-\Phi)$ has signature $[l, k]$.

In particular, Φ has signature $[n, 0]$, in other words V has an orthonormal K-basis, if and only if Φ is equivalent to the standard α-sesquilinear form Γ; this is the genuinely geometric case discussed in some more detail below. Moreover, Φ has signature $[n-1, 1]$ if and only if Φ is equivalent to the Minkowski α-sesquilinear form $\Gamma^{(n-1,1)}$; in this case V is called a **Minkowski space**.

(10.5) Scalar products. a) Let $[K, \alpha] \in \{[\mathbb{R}, \mathrm{id}], [\mathbb{C}, \overline{\ }]\}$, let V be a finitely generated K-vector space, let Φ be a hermitian α-sesquilinear form on V, and let $q \colon V \to K \colon v \mapsto \langle v, v \rangle$ be the associated **quadratic form**; note that $\langle v, v \rangle = \langle v, v \rangle^\alpha \in K$ implies that in both cases q has values in $\mathbb{R}$. Hence for $a \in K$ and $v \in V$ we have $q(av) = |a|^2 q(v) \in \mathbb{R}$; note that $q(0) = 0$.

If $q(v) > 0$ for all $0 \neq v \in V$, then q is called **positive definite**; if $q(v) \geq 0$ for all $v \in V$, then q is called **positive semi-definite**; if $q(v) < 0$ for all $0 \neq v \in V$, then q is called **negative definite**; if $q(v) \leq 0$ for all $v \in V$, then q is called **negative semi-definite**; otherwise q is called **indefinite**. In particular, if q is positive or negative definite then Φ is anisotropic.

These notions are related to the signature $[k, l]$ of Φ as follows: If $B \subseteq V$ is an orthogonal K-basis as in Sylvester's Theorem, then we have $q(x_1, \ldots, x_n) = [x_1, \ldots, x_n]^\alpha \cdot {}_B\Phi_B \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = (\sum_{i=1}^k |x_i|^2) - (\sum_{j=1}^l |x_{k+j}|^2) \in \mathbb{R}$, where $[x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}$ is the coordinate tuple with respect to B. Hence q is positive definite if and only if $k = n$; and q is positive semi-definite if and only if $l = 0$; while q is negative definite if and only if $l = n$; and q is negative semi-definite if and only if $k = 0$; thus q is indefinite in all the cases $\{k, l\} \neq \{0, n\}$.

If q is positive definite, then Φ is called a **scalar product**, where V is called **Euclidean** if $K = \mathbb{R}$, and **unitary** if $K = \mathbb{C}$. In particular, for the standard α-sesquilinear form Γ on $K^{n \times 1}$, where $n \in \mathbb{N}_0$, we have $q([a_1, \ldots, a_n]^{\mathrm{tr}}) = \langle [a_1, \ldots, a_n]^{\mathrm{tr}}, [a_1, \ldots, a_n]^{\mathrm{tr}} \rangle = \sum_{i=1}^n |a_i|^2 > 0$, for all $0 \neq [a_1, \ldots, a_n]^{\mathrm{tr}} \in K^{n \times 1}$, thus Γ is also called the **standard** scalar product on $K^{n \times 1}$.

b) We give a characterisation of the quadratic form q being positive or negative definite in terms of the **leading principal minors** of the Gram matrix of Φ, being called the **Hurwitz-Sylvester criterion**:

To this end, let $B = [v_1, \ldots, v_n] \subseteq V$ be any K-basis, where $n := \dim_K(V) \in \mathbb{N}_0$, and let $G := {}_B\Phi_B \in K^{n \times n}$. Moreover, for $k \in \{0, \ldots, n\}$ let $B_k := [v_1, \ldots, v_k]$ and $V_k := \langle B_k \rangle_K \leq V$ and $G_k := {}_{B_k}(\Phi|_{V_k})_{B_k} \in K^{k \times k}$; hence we have $G_n = G$.

Let q be positive or negative definite, and let $\epsilon := 1$ and $\epsilon := -1$, respectively. Then letting $C \subseteq V$ be an orthogonal K-basis as in Sylvester's Theorem, and $P := {}_B\mathrm{id}_C \in \mathrm{GL}_n(K)$, we have $P^* G P = {}_C\Phi_C = \epsilon E_n \in K^{n \times n}$, implying that $|\det(P)|^2 \cdot \det(G) = \det({}_C\Phi_C) = \epsilon^n$, hence $\epsilon^n \cdot \det(G) > 0$. Moreover, since definiteness is inherited to K-subspaces, we infer that $\epsilon^k \cdot \det(G_k) > 0$, for all

$k \in \{0, \ldots, n\}$; note that $\det(G_k)$ is the k-th leading principal minor of G, and that since G is hermitian we have $\det(G_k) \in \mathbb{R}$ indeed.

Conversely, let $\epsilon \in \{\pm 1\}$, and assume that $\epsilon^k \cdot \det(G_k) > 0$ for all $k \in \{0, \ldots, n\}$. We proceed by induction on $k \in \mathbb{N}_0$, where for $k \geq 1$ we may assume that $\Phi|_{V_{k-1}}$ is positive or negative definite, respectively; the case $k = 0$ being trivial: Let $[w_1, \ldots, w_{k-1}] \subseteq V_{k-1}$ be a K-basis as in Sylvester's Theorem, that is $\langle w_j, w_j \rangle = \epsilon$ for $j \in \{1, \ldots, k-1\}$. Then letting $w := v_k - \epsilon \cdot \sum_{j=1}^{k-1} \langle w_j, v_k \rangle w_j \in V_k$, we have $\langle w_i, w \rangle = \langle w_i, v_k \rangle - \langle w_i, v_k \rangle = 0$, for $i \in \{1, \ldots, k-1\}$, hence $w \in V_k \cap V_{k-1}^{\perp}$. Thus $C := [w_1, \ldots, w_{k-1}, w] \subseteq V_k$ is an orthogonal K-basis such that ${}_C(\Phi|_{V_k})_C = \epsilon E_{k-1} \oplus [\langle w, w \rangle]$; note that $V_{k-1} \cap V_{k-1}^{\perp} = \{0\}$. Hence from $\epsilon \langle w, w \rangle = \epsilon^k \cdot \det({}_C(\Phi|_{V_k})_C) = |\det({}_{B_k}\mathrm{id}_C)|^2 \cdot \epsilon^k \cdot \det(G_k) > 0$ we infer that $\langle w, w \rangle = \epsilon$, that is $\Phi|_{V_k}$ is positive or negative definite, respectively. $\sharp$

c) We now give a characterisation of the quadratic form q associated with Φ being positive or negative semi-definite in terms of all **principal minors** of the Gram matrix of Φ: To this end, for $S \subseteq \{1, \ldots, n\}$ let $G_S \in K^{|S| \times |S|}$ be the submatrix of $G = {}_B\Phi_B$ consisting of the columns and rows in S; hence we have $G_{\{1,\ldots,k\}} = G_k$, for $k \in \{0, \ldots, n\}$.

Let q be positive or negative semi-definite, and let $\epsilon := 1$ and $\epsilon := -1$, respectively. Then letting $C \subseteq V$ be an orthogonal K-basis as in Sylvester's Theorem, and $P := {}_B\mathrm{id}_C \in \mathrm{GL}_n(K)$, we have $P^* G P = {}_C\Phi_C = \epsilon E_r \oplus (0 \cdot E_{n-r}) \in K^{n \times n}$, for some $r \in \{0, \ldots, n\}$, implying that $|\det(P)|^2 \cdot \det(G) = \det({}_C\Phi_C) \in \{\epsilon^n, 0\}$, hence $\epsilon^n \cdot \det(G) \geq 0$. Moreover, since semi-definiteness is inherited to K-subspaces, we infer $\epsilon^{|S|} \cdot \det(G_S) \geq 0$, for all $S \subseteq \{1, \ldots, n\}$; note that $\det(G_S)$ is a principal minor of G, and that since G is hermitian we have $\det(G_S) \in \mathbb{R}$.

Conversely, let $\epsilon \in \{\pm 1\}$, and assume that $\epsilon^{|S|} \cdot \det(G_S) \geq 0$, for all $S \subseteq \{1, \ldots, n\}$. We consider the hermitian α-sesquilinear form $\Phi + \epsilon\xi\Gamma$, where $\xi > 0$ and Γ denotes the standard α-sesquilinear form with respect to the K-basis $B \subseteq V$, whose Gram matrix is given as ${}_B(\Phi + \epsilon\xi\Gamma)_B = G + \epsilon\xi E_n \in K^{n \times n}$, and whose associated quadratic form is given as $q_\xi(v) = q(v) + \epsilon\xi\Gamma(v, v)$, for all $v \in V$: For $k \in \{0, \ldots, n\}$ the characteristic polynomial of G_k equals $\chi_{G_k} = \det(XE_k - G_k) = X^k + \sum_{j=1}^{k} (-1)^j \cdot \big(\sum_{S \subseteq \{1,\ldots,k\},\, |S| = j} \det(G_S)\big) \cdot X^{k-j} \in \mathbb{R}[X]$. This yields $\det(G_k + XE_k) = (-1)^k \cdot \det((-X)E_k - G_k) = X^k + \sum_{j=1}^{k} \big(\sum_{S \subseteq \{1,\ldots,k\},\, |S| = j} \det(G_S)\big) \cdot X^{k-j}$. Hence we get $\epsilon^k \cdot \det(G_k + \epsilon\xi E_k) = \xi^k + \sum_{j=1}^{k} \epsilon^j \cdot \big(\sum_{S \subseteq \{1,\ldots,k\},\, |S| = j} \det(G_S)\big) \cdot \xi^{k-j} > 0$. Thus q_ξ is positive or negative definite, respectively, and hence $\epsilon q(v) = \lim_{\xi \to 0^+} (\epsilon q_\xi(v)) \geq 0$, for all $0 \neq v \in V$, showing that q is positive or negative semi-definite, respectively. $\sharp$

Note that the straightforward generalisation of the definite case, namely that $\epsilon^k \cdot \det(G_k) \geq 0$, for all $k \subseteq \{0, \ldots, n\}$, already entails semi-definiteness, does not hold, as the example $G := \begin{bmatrix} . & . \\ . & -1 \end{bmatrix}$, for $\epsilon = 1$, shows.

(10.6) Orthonormalisation. **a)** Let $[K,\alpha]\in\{[\mathbb{R},\mathrm{id}],[\mathbb{C},\bar{\ }]\}$, let Φ be a scalar product on a K-vector space V, and let $B=[v_1,\ldots,v_n]\subseteq V$ be a K-basis, where $n:=\dim_K(V)\in\mathbb{N}_0$. Then V has a unique orthonormal K-basis C such that ${}_B\mathrm{id}_C\in\mathrm{GL}_n(K)$ is upper triangular having positive diagonal entries, called the **Gram-Schmidt** K-basis associated with B; recall that orthonormal K-bases possibly exist only if Φ is a scalar product:

The existence of C follows from the above orthogonalisation procedure using that Φ is anisotropic. To show uniqueness, let $B_k:=[v_1,\ldots,v_k]$ and $V_k:=\langle B_k\rangle_K\leq V$, for $k\in\{0,\ldots,n\}$. We proceed by induction on $k\in\mathbb{N}_0$, the case $k=0$ being trivial: For $k\geq 1$ let $C_{k-1}=[w_1,\ldots,w_{k-1}]$ be as desired; then we have $\langle C_{k-1}\rangle_K=V_{k-1}$. Then for any $v=av_k+\sum_{j=1}^{k-1}a_jw_j\in V_k\cap V_{k-1}^\perp$ we get $0=\langle w_i,v\rangle=a\langle w_i,v_k\rangle+\sum_{j=1}^{k-1}a_j\langle w_i,w_j\rangle=a\langle w_i,v_k\rangle+a_i$, for all $i\in\{1,\ldots,k-1\}$, implying $v=aw$ where $w:=v_k-\sum_{j=1}^{k-1}\langle w_j,v_k\rangle w_j\in V_k$. Then $\langle v,v\rangle=1$ and the positivity of the diagonal entries of ${}_B\mathrm{id}_C$ yields $a=\frac{1}{\sqrt{\langle w,w\rangle}}$.

We have the **Cholesky decomposition** ${}_B\Phi_B=({}_C\mathrm{id}_B)^*\cdot{}_C\mathrm{id}_B\in K^{n\times n}$, where ${}_C\mathrm{id}_B=({}_B\mathrm{id}_C)^{-1}\in\mathrm{GL}_n(K)$ is upper triangular. Moreover, any orthonormal subset of V can be extended to an orthonormal K-basis of V; recall that orthogonal sets consisting of non-isotropic vectors are K-linearly independent.

b) If we are given a Gram matrix $G={}_B\Phi_B$ of some hermitian α-sesquilinear form Φ with respect to some K-basis $B\subseteq V$, the question arises how we may decide whether Φ is a scalar product. This can be done in various ways:

Firstly, the quadratic form associated with Φ is $q(x_1,\ldots,x_n)=[x_1,\ldots,x_n]^\alpha\cdot G\cdot[x_1,\ldots,x_n]^{\mathrm{tr}}$, where $[x_1,\ldots,x_n]^{\mathrm{tr}}\in\mathbb{R}^{n\times 1}$ is the coordinate tuple with respect to B, and we may try and decide whether q is positive definite. Secondly, we may apply the Hurwitz-Sylvester criterion. Thirdly, we may run the orthogonalisation procedure, regardless of whether or not Φ is a scalar product, which yields an orthonormal K-basis if Φ is a scalar product, and otherwise at a certain stage necessarily produces a vector $0\neq v\in V$ such that $\Phi(v,v)\leq 0$. Fourthly, yet another criterion will be given in (10.14).

For example, let Γ be the standard scalar product on $V:=\mathbb{R}^{2\times 1}$, let $A\subseteq V$ be the standard $\mathbb{R}$-basis, and let the $\mathbb{R}$-basis $B\subseteq V$ be given as $Q={}_A\mathrm{id}_B:=\begin{bmatrix}1 & -\frac{1}{2}\\ 0 & \frac{\sqrt{3}}{2}\end{bmatrix}\in\mathrm{GL}_2(\mathbb{R})$, hence we get $G:={}_B\Gamma_B=Q^{\mathrm{tr}}Q=\begin{bmatrix}1 & -\frac{1}{2}\\ -\frac{1}{2} & 1\end{bmatrix}\in\mathbb{R}^{2\times 2}$.

By construction G is the Gram matrix of a scalar product. But if we are just given the matrix G then this information is lost. Still, the associated quadratic form is given as $q(x,y)=[x,y]\cdot G\cdot[x,y]^{\mathrm{tr}}=x^2-xy+y^2=(x-\frac{1}{2}y)^2+\frac{3}{4}y^2$, hence $q(x,y)>0$ for all $0\neq[x,y]\in\mathbb{R}^2$; alternatively, we have $\det([1])=1>0$ and $\det(G)=\frac{3}{4}>0$, hence the Hurwitz-Sylvester criterion implies that G describes a scalar product. We aim to find an orthonormal $\mathbb{R}$-basis of V from G:

Letting $P_1:=\begin{bmatrix}1 & \frac{1}{2}\\ 0 & 1\end{bmatrix}$ yields $P_1^{\mathrm{tr}}GP_1=\mathrm{diag}[1,\frac{3}{4}]$, thus letting $P_2:=\mathrm{diag}[1,\frac{2}{\sqrt{3}}]$

and $P := P_1 P_2 = \begin{bmatrix} 1 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{3}} \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$ yields $P^{\mathrm{tr}} G P = E_2$, hence we get the orthonormal $\mathbb{R}$-basis $C \subseteq V$ defined by ${}_B\mathrm{id}_C := P$; indeed we have ${}_A\mathrm{id}_C = {}_A\mathrm{id}_B \cdot {}_B\mathrm{id}_C = QP = E_2$, thus C is just the standard $\mathbb{R}$-basis. $\sharp$

(10.7) Euclidean and unitary geometry. a) Let $[K, \alpha] \in \{[\mathbb{R}, \mathrm{id}_\mathbb{R}], [\mathbb{C}, \overline{}]\}$, and Φ be a scalar product on a finitely generated K-vector space V, with associated quadratic form q. Letting $\|v\| := \sqrt{q(v)} = \sqrt{\langle v, v\rangle} \in \mathbb{R}_{\geq 0}$ be the **length** or **norm** of $v \in V$, we have $\|av\| = |a| \cdot \|v\|$, for $a \in K$, that is **linearity** with respect to absolute values, and $\|v\| = 0$ if and only if and only if $v = 0$, that is **definiteness**. Then $\frac{1}{\|v\|} \cdot v \in V$ is called the **unit** vector associated with $v \neq 0$.

For $v, w \in V$ we have the **Cauchy-Schwarz inequality** $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$, where equality holds if and only if $[v, w]$ is K-linearly dependent: We may assume that $v \neq 0$. For any $u := av + bw \in V$, where $a, b \in K$, we have $\langle u, u\rangle = |a|^2 \langle v, v\rangle + a^\alpha b \langle v, w\rangle + a b^\alpha \langle w, v\rangle + |b|^2 \langle w, w\rangle$. Thus letting $a := -\langle v, w\rangle$ and $b := \langle v, v\rangle$ we get $\langle u, u\rangle = |\langle v, w\rangle|^2 \langle v, v\rangle - \langle v, w\rangle^\alpha \langle v, v\rangle \langle v, w\rangle - \langle v, w\rangle \langle v, v\rangle^\alpha \langle w, v\rangle + |\langle v, v\rangle|^2 \langle w, w\rangle = \langle v, v\rangle(\langle v, v\rangle \langle w, w\rangle - |\langle v, w\rangle|^2)$. Since $\langle u, u\rangle \geq 0$ and $\langle v, v\rangle > 0$ we conclude $|\langle v, w\rangle|^2 \leq \langle v, v\rangle \cdot \langle w, w\rangle$. Moreover, if equality holds then $\langle u, u\rangle = 0$, that is $u = 0$, thus $b = \langle v, v\rangle \neq 0$ implies that $[v, w]$ is K-linearly dependent. Conversely, if $[v, w]$ is K-linearly dependent, then there is $a \in K$ such that $w = av$, and hence $|\langle v, w\rangle|^2 = |\langle v, av\rangle|^2 = |a|^2 |\langle v, v\rangle|^2 = \langle v, v\rangle \langle av, av\rangle = \langle v, v\rangle \langle w, w\rangle$.

This yields the **Minkowski** or **triangle inequality** $\|v+w\| \leq \|v\| + \|w\|$: We have $\|v+w\|^2 = \langle v+w, v+w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle v, w\rangle^\alpha + \langle w, w\rangle = \langle v, v\rangle + 2\mathrm{Re}(\langle v, w\rangle) + \langle w, w\rangle \leq \langle v, v\rangle + 2|\langle v, w\rangle| + \langle w, w\rangle \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$. Thus V together with the norm $\|\cdot\|$, where the latter fulfills linearity, definiteness and the triangle inequality, becomes a **normed vector space**; since the norm is induced by a scalar product, V even is a **(pre-)Hilbert space**.

For $K = \mathbb{R}$ we have the following geometric interpretation: For $0 \neq v, w \in V$ we have $-1 \leq \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leq 1$. Thus there is a unique $0 \leq \omega \leq \pi$ such that $\cos(\omega) = \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$, called the **angle** between the unit vectors $\frac{1}{\|v\|} \cdot v$ and $\frac{1}{\|w\|} \cdot w$, and the latter are **perpendicular**, that is we have $\omega = \frac{\pi}{2}$, if and only if $\cos(\omega) = 0$, which holds if and only if $\langle v, w\rangle = 0$, that is $v \perp w$.

b) Let $\{v_1, \ldots, v_n\} \subseteq V$ be an orthonormal K-basis, where $n := \dim_K(V) \in \mathbb{N}_0$. Then for any $v \in V$ we have the **Fourier expansion** $v = \sum_{i=1}^n \langle v_i, v\rangle v_i$: Letting $v' \in V$ denote the right hand side, we for all $j \in \{1, \ldots, n\}$ have $\langle v_j, v'\rangle = \sum_{i=1}^n \langle v_i, v\rangle \langle v_j, v_i\rangle = \langle v_j, v\rangle$, implying $v - v' \in V^\perp = \{0\}$. Moreover, we have **Pythagoras's Theorem** $\|v\|^2 = \langle v, v\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v\rangle \langle v_j, v\rangle^\alpha \langle v_i, v_j\rangle = \sum_{i=1}^n \langle v_i, v\rangle \langle v_i, v\rangle^\alpha = \sum_{i=1}^n |\langle v_i, v\rangle|^2$.

Let $U \leq V$ be a K-subspace, and let $\{u_1, \ldots, u_m\} \subseteq U$ be an orthonormal K-basis, where $m := \dim_K(U) \leq n$. Then for all $a_1, \ldots, a_m \in K$ we have $\|v - \sum_{i=1}^m a_i u_i\|^2 = \langle v, v\rangle - \sum_{i=1}^m (a_i \langle u_i, v\rangle^\alpha + a_i^\alpha \langle u_i, v\rangle) + \sum_{i=1}^m |a_i|^2 = \langle v, v\rangle - \sum_{i=1}^m |\langle u_i, v\rangle|^2 + \sum_{i=1}^m (a_i - \langle u_i, v\rangle)(a_i^\alpha - \langle u_i, v\rangle^\alpha) = \langle v, v\rangle - \sum_{i=1}^m |\langle u_i, v\rangle|^2 +$

$\sum_{i=1}^m |a_i - \langle u_i, v\rangle|^2$, implying the **Bessel inequality** $\min\{\|v-u\|^2; u \in U\} = \|v\|^2 - \sum_{i=1}^m |\langle u_i, v\rangle|^2 \geq 0$, where the minimum is attained precisely for the **best approximation** $u_0 := \sum_{i=1}^m \langle u_i, v\rangle u_i \in U$; we have $\langle v - u_0, u_j\rangle = \langle v, u_j\rangle - \sum_{i=1}^m \langle u_i, v\rangle\langle u_i, u_j\rangle = \langle v, u_j\rangle - \langle u_j, v\rangle = 0$, for all $j \in \{1, \ldots, m\}$, hence $v - u_0 \in U^\perp$, in other words we have $U \cap (v + U^\perp) = \{u_0\}$.

(10.8) Adjoint maps. Let K be a field, let $\alpha\colon K \to K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$, and let V be finitely generated K-vector space with a non-degenerate α-sesquilinear form Φ. For any $\varphi \in \mathrm{End}_K(V)$ there is a unique **adjoint map** $\varphi^* \in \mathrm{End}_K(V)$ such that $\langle v, \varphi(w)\rangle = \langle \varphi^*(v), w\rangle$ for all $v, w \in V$:

Let $B \subseteq V$ be a K-basis and $G := {}_B\Phi_B \in \mathrm{GL}_n(K)$, where $n := \dim_K(V) \in \mathbb{N}_0$. Let $A := {}_B\varphi_B \in K^{n\times n}$, let $A' := (GAG^{-1})^* \in K^{n \times n}$, that is $A'^*G = GA$, and let $\varphi^* \in \mathrm{End}_K(V)$ be defined by ${}_B(\varphi^*)_B := A'$. Then for the α-sesquilinear form $\Psi\colon V \times V \to K\colon [v, w] \mapsto \langle v, \varphi(w)\rangle$ we have ${}_B\Psi_B = GA = A'^*G$, that is $\Psi(v, w) = \langle \varphi^*(v), w\rangle$. If $\varphi' \in \mathrm{End}_K(V)$ such that $\langle \varphi'(v), w\rangle = \langle \varphi^*(v), w\rangle$, for all $v, w \in V$, then we have $(\varphi' - \varphi^*)(v) \in {}^\perp V = \{0\}$, that is $\varphi' = \varphi^*$. ♯

In particular, if B is orthonormal, that is $G = E_n$, then we have ${}_B(\varphi^*)_B = A^*$. In general, from $(G \cdot aA \cdot G^{-1})^* = a^\alpha (GAG^{-1})^*$, for all $a \in K$, we conclude that the map ${}^*\colon \mathrm{End}_K(V) \to \mathrm{End}_K(V)\colon \varphi \mapsto \varphi^*$ is α-semilinear. We have $\mathrm{id}_V^* = \mathrm{id}_V$, as well as $\det(\varphi^*) = \det((GAG^{-1})^*) = \det(A)^\alpha = \det(\varphi)^\alpha$, and $\mathrm{rk}(\varphi^*) = \mathrm{rk}((GAG^{-1})^*) = \mathrm{rk}(A) = \mathrm{rk}(\varphi)$. In particular, we have $\varphi \in \mathrm{GL}(V)$ if and only if $\varphi^* \in \mathrm{GL}(V)$, and in this case we from $((GAG^{-1})^*)^{-1} = (GA^{-1}G^{-1})^*$ get $(\varphi^*)^{-1} = (\varphi^{-1})^*$. Moreover, for $\varphi' \in \mathrm{End}_K(V)$, letting $A' := {}_B(\varphi')_B \in K^{n\times n}$, we get $(GA'AG^{-1})^* = (GAG^{-1})^*(GA'G^{-1})^*$, thus $(\varphi'\varphi)^* = \varphi^*\varphi'^*$.

We have $v \in \ker(\varphi^*)$ if and only if $0 = \langle \varphi^*(v), w\rangle = \langle v, \varphi(w)\rangle$ for all $w \in V$, that is $\ker(\varphi^*) = {}^\perp\mathrm{im}(\varphi)$, and $w \in \ker(\varphi)$ if and only if $0 = \langle v, \varphi(w)\rangle = \langle \varphi^*(v), w\rangle$ for all $v \in V$, that is $\ker(\varphi) = \mathrm{im}(\varphi^*)^\perp$. Moreover, if $U \leq V$ is φ-invariant, then from $\langle \varphi^*(v), w\rangle = \langle v, \varphi(w)\rangle = 0$ for all $v \in {}^\perp U$ and $w \in U$ we infer that ${}^\perp U$ is φ^*-invariant; and if $U \leq V$ is φ^*-invariant then from $\langle v, \varphi(w)\rangle = \langle \varphi^*(v), w\rangle = 0$ for all $v \in U$ and $w \in U^\perp$ we infer that $U^\perp$ is φ-invariant.

Finally, if Φ is hermitian then we have $\langle v, \varphi^*(w)\rangle = \langle \varphi^*(w), v\rangle^\alpha = \langle w, \varphi(v)\rangle^\alpha = \langle \varphi(v), w\rangle$, hence we get $\varphi^{**} = \varphi$; we argue similarly if Φ is skew-hermitian.

(10.9) Normal maps. a) Let K be a field, let $\alpha\colon K \to K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$, and let V be finitely generated K-vector space with a non-degenerate hermitian form Φ. A map $\varphi \in \mathrm{End}_K(V)$ is called **normal** if $\varphi\varphi^* = \varphi^*\varphi$. In particular, if $\varphi^* = \varphi$ then φ is called **hermitian** or **self-adjoint**; if $\varphi \in \mathrm{GL}(V)$ such that $\varphi^* = \varphi^{-1}$ then φ is called **unitary** or an **isometry**; if $\alpha = \mathrm{id}_K$ then in these cases φ is also called **symmetric** and **orthogonal**, respectively. Hence if $B \subseteq V$ is an orthonormal K-basis, then these properties are translated into the respective properties of the matrix ${}_B\varphi_B$.

Then φ is normal if and only if $\langle \varphi(v), \varphi(w)\rangle = \langle \varphi^*(v), \varphi^*(w)\rangle$ for all $v, w \in V$: If φ is normal, then we have $\langle \varphi(v), \varphi(w)\rangle = \langle \varphi^*\varphi(v), w\rangle = \langle \varphi\varphi^*(v), w\rangle =$

$\langle \varphi^*(v), \varphi^*(w)\rangle$; conversely, $\langle \varphi^*\varphi(v), w\rangle = \langle \varphi(v), \varphi(w)\rangle = \langle \varphi^*(v), \varphi^*(w)\rangle = \langle \varphi\varphi^*(v), w\rangle$ shows $\varphi^*\varphi = \varphi\varphi^*$.

b) Let V be Euclidean or unitary. Then φ is normal if and only if $\|\varphi(v)\| = \|\varphi^*(v)\|$ for all $v \in V$: From $\langle \varphi(v+aw), \varphi(v+aw)\rangle = \langle \varphi^*(v+aw), \varphi^*(v+aw)\rangle$ we get $a\langle \varphi(v), \varphi(w)\rangle + a^\alpha\langle \varphi(w), \varphi(v)\rangle = a\langle \varphi^*(v), \varphi^*(w)\rangle + a^\alpha\langle \varphi^*(w), \varphi^*(v)\rangle$, for all $v, w \in V$ and $a \in K$, that is $2\mathrm{Re}(a\langle \varphi(v), \varphi(w)\rangle) = 2\mathrm{Re}(a\langle \varphi^*(v), \varphi^*(w)\rangle)$, hence letting $a := 1$ and $a := i$ shows that $\langle \varphi(v), \varphi(w)\rangle = \langle \varphi^*(v), \varphi^*(w)\rangle$.

If φ is normal, then this implies $\ker(\varphi) = \ker(\varphi^*)$. Moreover, for any φ-invariant K-subspace $U \le V$ the orthogonal space $U^\perp \le V$ is φ-invariant as well: Let $B := [v_1, \ldots, v_n] \subseteq V$ be an orthonormal K-basis such that $\langle v_1, \ldots, v_m\rangle_K = U$ and $\langle v_{m+1}, \ldots, v_n\rangle_K = U^\perp$, where $n := \dim_K(V) \in \mathbb{N}_0$ and $m := \dim_K(U) \in \mathbb{N}_0$; recall that $V = U \oplus U^\perp$. Then $A := M_B^B(\varphi) \in K^{n\times n}$ is a matrix of shape $A = \left[\begin{array}{c|c} A' & C \\ \hline . & A'' \end{array}\right]$, where $A' \in K^{m\times m}$ and $A'' \in K^{(n-m)\times(n-m)}$ and $C = [c_{ij}]_{ij} \in K^{m\times(n-m)}$. We have $A^* = \left[\begin{array}{c|c} A'^* & . \\ \hline C^* & A''^* \end{array}\right]$, hence normality, that is $AA^* = A^*A$, implies $A'^*A' = A'A'^* + CC^*$. Since $\mathrm{Tr}(A'^*A') = \mathrm{Tr}(A'A'^*)$, this entails $0 = \mathrm{Tr}(CC^*) = \sum_{i=1}^m \sum_{j=1}^{n-m} c_{ij}c_{ij}^\alpha = \sum_{i=1}^m \sum_{j=1}^{n-m} |c_{ij}|^2$, thus $C = 0$.

(10.10) Unitary maps. a) Let K be a field, let $\alpha\colon K \to K$ be a field automorphism such that $\alpha^2 = \mathrm{id}_K$, let V be finitely generated K-vector space with a non-degenerate hermitian form Φ, and let $\varphi \in \mathrm{End}_K(V)$. Then φ is unitary if and only if $\langle \varphi(v), \varphi(w)\rangle = \langle v, w\rangle$ for all $v, w \in V$: If φ is unitary, then we have $\langle \varphi(v), \varphi(w)\rangle = \langle \varphi^*\varphi(v), w\rangle = \langle \varphi^{-1}\varphi(v), w\rangle = \langle v, w\rangle$; conversely, $\langle \varphi^*\varphi(v), w\rangle = \langle \varphi(v), \varphi(w)\rangle = \langle v, w\rangle$ shows $\varphi^*\varphi = \mathrm{id}_V$.

If φ is unitary, then we have $\det(\varphi)^{1+\alpha} = \det(\varphi\varphi^*) = \det(\mathrm{id}_V) = 1$; in particular, if $[K, \alpha] = [\mathbb{C}, \overline{\ }]$ then we have $|\det(\varphi)|^2 = 1$, and if $\alpha = \mathrm{id}_K$ then we have $\det(\varphi) \in \{\pm 1\}$, where orthogonal maps of determinant 1 are called **rotations**. The subgroups $\mathrm{GU}(V) := \{\varphi \in \mathrm{GL}(V); \varphi \text{ unitary}\} \le \mathrm{GL}(V)$ and $\mathrm{SU}(V) := \mathrm{GU}(V) \cap \mathrm{SL}(V) = \{\varphi \in \mathrm{GU}(V); \det(\varphi) = 1\} \le \mathrm{GL}(V)$ are called the **general** and **special unitary groups**, respectively; if $\alpha = \mathrm{id}_K$ these are also called the **general** and **special orthogonal groups**, denoted by $\mathrm{GO}(V)$ and $\mathrm{SO}(V)$, respectively.

b) Let V be Euclidean or unitary. Then φ is unitary if and only if $\|\varphi(v)\| = \|v\|$ for all $v \in V$; in particular, for any eigenvalue $a \in K$ of a unitary map φ we have $|a| = 1$: From $\langle \varphi(v+aw), \varphi(v+aw)\rangle = \langle v+aw, v+aw\rangle$ we get $a\langle \varphi(v), \varphi(w)\rangle + a^\alpha\langle \varphi(w), \varphi(v)\rangle = a\langle v, w\rangle + a^\alpha\langle w, v\rangle$, for all $v, w \in V$ and $a \in K$, that is $2\mathrm{Re}(a\langle \varphi(v), \varphi(w)\rangle) = 2\mathrm{Re}(a\langle v, w\rangle)$, hence $\langle \varphi(v), \varphi(w)\rangle = \langle v, w\rangle$.

Since unitary maps leave the length of vectors invariant, they are also called isometries. Moreover, unitary maps also leave the angle between vectors invariant: Indeed, for $0 \ne v, w \in V$ we have $\frac{\langle \varphi(v), \varphi(w)\rangle}{\|\varphi(v)\|\cdot\|\varphi(w)\|} = \frac{\langle v, w\rangle}{\|v\|\cdot\|w\|}$.

(10.11) Example: The group $\mathbf{GO}_2(\mathbb{R})$. For $K = \mathbb{R}$ and $\alpha = \mathrm{id}_{\mathbb{R}}$, the general orthogonal group $\mathrm{GO}_1(\mathbb{R}) \leq \mathrm{GL}_1(\mathbb{R}) \cong \mathbb{R}^*$, with respect to the standard scalar product on $\mathbb{R}$, that is the absolute value, is easily described: Since orthogonal maps are isometries, we have $\mathrm{GO}_1(\mathbb{R}) = \{\pm 1\}$. More substantially, the general orthogonal group $\mathrm{GO}_2(\mathbb{R}) \leq \mathrm{GL}_2(\mathbb{R})$, with respect to the standard scalar products on $\mathbb{R}^{2\times 1}$, is described as follows:

a) A matrix $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$ is orthogonal if and only if $\begin{bmatrix} a & c \\ b & d \end{bmatrix} = A^{\mathrm{tr}} = A^{-1} = \frac{1}{ad-bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Hence taking determinants yields $(ad-bc)^2 = \det(A)^2 = 1$, entailing $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \cdot \mathrm{diag}[1, \epsilon]$, where $a^2 + b^2 = 1$ and $\epsilon = \det(A) \in \{\pm 1\}$; conversely all matrices of this form are orthogonal. Note that, the standard $\mathbb{R}$-basis of $\mathbb{R}^{2\times 1}$ being orthonormal, the colums $\{[a,b]^{\mathrm{tr}}, \epsilon \cdot [-b,a]^{\mathrm{tr}}\}$ of A indeed form an orthonormal $\mathbb{R}$-basis of $\mathbb{R}^{2\times 1}$ as well.

For $\omega \in \mathbb{R}$ let still $A_\omega := \begin{bmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{bmatrix} \in \mathrm{SL}_2(\mathbb{R})$, see (8.4); recall that $\chi_{A_\omega} = p_\omega := X^2 - 2\cos(\omega)X + 1 \in \mathbb{R}[X]$ is irreducible if and only if $\omega \notin \pi\mathbb{Z}$. Taking the 2π-periodicity of $\cos\colon \mathbb{R} \to \mathbb{R}$ and $\sin\colon \mathbb{R} \to \mathbb{R}$ into account, we have $A_\omega = A_{\omega'}$ if and only if $\omega - \omega' = 2\pi\mathbb{Z}$. Given A as above, there is a unique $0 \leq \omega < 2\pi$ such that $a = \cos(\omega)$ and $b = \sin(\omega)$, hence we get $A = A_\omega \cdot \mathrm{diag}[1, \epsilon] \in \mathrm{GL}_2(\mathbb{R})$. Thus we have $\mathrm{SO}_2(\mathbb{R}) = \{A_\omega \in \mathrm{SL}_2(\mathbb{R}); 0 \leq \omega < 2\pi\}$ and $\mathrm{GO}_2(\mathbb{R}) = \mathrm{SO}_2(\mathbb{R}) \,\dot\cup\, \mathrm{SO}_2(\mathbb{R}) \cdot \mathrm{diag}[1, -1]$.

b) The algebraic structure of $\mathrm{SO}_2(\mathbb{R})$ is elucidated as follows: We have already seen in (8.4) that using $P := \begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C})$ we get $P^{-1}A_\omega P = \mathrm{diag}[\exp(i\omega), \exp(-i\omega)] \in \mathbb{C}^{2\times 2}$, for all $\omega \in \mathbb{R}$. Note that the eigenvectors found are mutually orthogonal, hence going over to unit vectors shows that $\mathbb{C}^{2\times 1}$ indeed has an orthonormal $\mathbb{C}$-basis consisting of eigenvectors of A_ω.

Hence we have $A_\omega A_{\omega'} = P \cdot P^{-1}A_\omega P \cdot P^{-1}A_{\omega'}P \cdot P^{-1} = P \cdot \mathrm{diag}[\exp(i(\omega + \omega')), \exp(-i(\omega + \omega'))] \cdot P^{-1} = A_{\omega+\omega'}$, for $\omega, \omega' \in \mathbb{R}$. Thus we have $A_\omega A_{\omega'} = A_{\omega+\omega'} = A_{\omega'}A_\omega$, saying that $\mathrm{SO}_2(\mathbb{R})$ is commutative, while $A_\omega \cdot \mathrm{diag}[1,-1] = \mathrm{diag}[1,-1] \cdot A_{-\omega} = \mathrm{diag}[1,-1] \cdot A_\omega^{-1}$ implies that $\mathrm{GO}_2(\mathbb{R})$ is not commutative.

c) We have the following geometric interpretation of orthogonal maps: Let $-\pi \leq \omega \leq \pi$. For both $i \in \{1,2\}$ we have $\langle e_i, A_\omega e_i \rangle = \cos(\omega) = \cos(|\omega|)$, saying that the angle between e_i and $A_\omega e_i$ equals $|\omega|$. Thus indeed A_ω is a **rotation** with respect to the angle ω. Hence A_ω maps the standard $\mathbb{R}$-basis to an orthonormal $\mathbb{R}$-basis keeping the **orientation**.

The map $\mathrm{diag}[1,-1]$ is a **reflection** with respect to the hyperplane $\langle e_2 \rangle_{\mathbb{R}}^{\perp} = \langle e_1 \rangle_{\mathbb{R}} \leq \mathbb{R}^{2\times 1}$, that is e_2 is mapped to its negative, while the orthogonal space $\langle e_2 \rangle_{\mathbb{R}}^{\perp}$ consists of **fixed points**. Hence $\mathrm{diag}[1,-1]$ maps the standard $\mathbb{R}$-basis $[e_1, e_2]$ to the orthonormal $\mathbb{R}$-basis $[e_1, -e_2]$, thus changing the orientation.

For $A = A_\omega \cdot \operatorname{diag}[1,-1] = \begin{bmatrix} \cos(\omega) & \sin(\omega) \\ \sin(\omega) & -\cos(\omega) \end{bmatrix}$, which changes orientation as well, we have $\langle e_1, A_\omega e_1 \rangle = \cos(\omega)$ and $\langle e_2, A_\omega e_2 \rangle = -\cos(\omega)$, hence for $\omega \notin \{\pm\frac{\pi}{2}\}$ the angles between e_1 and Ae_1, and e_2 and Ae_2 are different. For $\omega = \pm\frac{\pi}{2}$ we get $A = \pm \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix}$, the reflections with respect to the hyperplanes $\langle [-1,1]^{\mathrm{tr}} \rangle_{\mathbb{R}}^{\perp} = \langle [1,1]^{\mathrm{tr}} \rangle_{\mathbb{R}}$ and $\langle [1,1]^{\mathrm{tr}} \rangle_{\mathbb{R}}^{\perp} = \langle [-1,1]^{\mathrm{tr}} \rangle_{\mathbb{R}}$, respectively; for $\omega = \pi$ we get $A = \operatorname{diag}[-1,1]$, the reflection with respect to the hyperplane $\langle e_1 \rangle_{\mathbb{R}}^{\perp} = \langle e_2 \rangle_{\mathbb{R}}$.

We proceed to show that $A \in \mathrm{GO}_2(\mathbb{R}) \setminus \mathrm{SO}_2(\mathbb{R})$ always is a reflection: We have $\chi_A = X^2 - 1 = (X-1)(X+1) \in \mathbb{R}[X]$, hence A is diagonalisable. Letting $v_\delta \in \mathbb{R}^{2\times 1}$ be eigenvectors with respect to the eigenvalues $\delta \in \{\pm 1\}$, respectively, we conclude that $\langle v_1 \rangle_{\mathbb{R}} \leq \mathbb{R}^{2\times 1}$ is a hyperplane consisting of fixed points, while v_{-1} is mapped to its negative. Hence it remains to be shown that the eigenspaces are mutually orthogonal, where we may assume that $\omega \notin \{0, \pm\pi\}$: We have $T_\delta(A) = \ker\left(\begin{bmatrix} \cos(\omega) - \delta & \sin(\omega) \\ \sin(\omega) & -(\cos(\omega) + \delta) \end{bmatrix}\right) = \langle v_\delta \rangle_{\mathbb{R}}$, where $v_\delta := [\cos(\omega) + \delta, \sin(\omega)]^{\mathrm{tr}} \in \mathbb{R}^{2\times 1}$, which indeed entails $\langle v_1, v_{-1} \rangle = 0$; note that by assumption we have $|\cos(\omega)| < 1$, hence $\|v_\delta\|^2 = 2 + 2\delta\cos(\omega) \neq 0$.

Going over to unit vectors, this implies that $\mathbb{R}^{2\times 1}$ has an orthonormal $\mathbb{R}$-basis consisting of eigenvectors of A; note that A is symmetric. This statement also holds for the excluded cases $\omega \in \{0, \pm\pi\}$, as does the statement to follow.

The angle between the reflection hyperplane $\langle v_{-1} \rangle_{\mathbb{R}}^{\perp}$ of A, and the reflection hyperplane $\langle e_2 \rangle_{\mathbb{R}}^{\perp}$ of $\operatorname{diag}[1,-1]$, is described by $\frac{\langle v_{-1}, e_2 \rangle}{\|v_{-1}\| \cdot \|e_2\|} = \frac{\sin(\omega)}{\sqrt{2 - 2\cos(\omega)}} = \frac{\omega}{|\omega|} \cdot \cos(\frac{\omega}{2})$; note that the latter equality follows from the matrix equation $\begin{bmatrix} \cos(\omega)^2 - \sin(\omega)^2 & -2\cos(\omega)\sin(\omega) \\ 2\cos(\omega)\sin(\omega) & \cos(\omega)^2 - \sin(\omega)^2 \end{bmatrix} = A_\omega^2 = A_{2\omega} = \begin{bmatrix} \cos(2\omega) & -\sin(2\omega) \\ \sin(2\omega) & \cos(2\omega) \end{bmatrix}$, which entails $\frac{\sin(2\omega)}{\sqrt{2 - 2\cos(2\omega)}} = \frac{2\cos(\omega)\sin(\omega)}{\sqrt{2 - 2\cos(\omega)^2 + 2\sin(\omega)^2}} = \frac{2\cos(\omega)\sin(\omega)}{\sqrt{4\sin(\omega)^2}} = \frac{\omega}{|\omega|} \cdot \cos(\omega)$.

(10.12) Theorem: Spectral theorem. Let V be an Euclidean or unitary K-vector space with scalar product Φ, and $\varphi \in \mathrm{End}_K(V)$. Then there is an orthonormal K-basis of V consisting of eigenvectors of φ if and only if φ is normal and $\chi_\varphi \in K[X]$ splits into linear factors.

In particular, these conditions are fulfilled if **i)** V is unitary and φ is normal, or **ii)** φ is hermitian.

In particular, if $A \in K^{n\times n}$, where $n \in \mathbb{N}_0$, such that **i)** $K = \mathbb{C}$ and A is normal, or **ii)** A is hermitian, then there is a unitary matrix $P \in \mathrm{GL}_n(K)$, such that $P^{-1}AP = P^*AP \in K^{n\times n}$ is a diagonal matrix.

Proof. Let $B \subseteq V$ be an orthonormal K-basis consisting of eigenvectors of φ. Then φ is diagonalisable, hence $\chi_\varphi \in K[X]$ splits into linear factors. Moreover, $A := {}_B\varphi_B \in K^{n\times n}$, where $n := \dim_K(V) \in \mathbb{N}_0$, is a diagonal matrix, hence from ${}_B(\varphi^*)_B = A^*$ we infer $AA^* = A^*A$.

Conversely, we proceed by induction on $n \in \mathbb{N}_0$, the case $n = 0$ being trivial: Let $n \geq 1$, and since $\chi_\varphi \in K[X]$ splits into linear factors, let $v \in V$ be an eigenvector of φ, where we may assume that $\|v\| = 1$, and let $U := \langle v \rangle_K$. Hence we have $V = U \oplus U^\perp$, where $U^\perp$ is both φ- and φ^*-invariant. Since $\Phi|_{U^\perp \times U^\perp}$ is a scalar product, and $(\varphi|_{U^\perp})^* = \varphi^*|_{U^\perp}$ implies that $\varphi|_{U^\perp}$ is normal, we are done by induction.

The assertion in i) follows from $\mathbb{C}$ being algebraically closed. To show ii), let $A := {}_B\varphi_B \in K^{n\times n} \subseteq \mathbb{C}^{n\times n}$, where $B \subseteq V$ be an orthonormal K-basis. Then $A - aE_n \in \mathbb{C}^{n\times n}$ being normal, for all $a \in \mathbb{C}$, we have $T_a(A) = \ker(A - aE_n) = \ker((A - aE_n)^*) = \ker(A^* - \overline{a}E_n) = T_{\overline{a}}(A^*)$, hence $A = A^*$ implies that for all eigenvalues $a \in \mathbb{C}$ of A we have $\overline{a} = a$, that is $a \in \mathbb{R}$. Since $\chi_A \in \mathbb{C}[X]$ splits into linear factors, we conclude that $\chi_A \in \mathbb{R}[X]$ also splits into linear factors. ♯

(10.13) Corollary. a) If $K = \mathbb{C}$, then φ is unitary if and only if φ is normal and all its eigenvalues have absolute value 1.
b) If $K = \mathbb{R}$, then φ is orthogonal if and only if there is an orthonormal $\mathbb{R}$-basis $B \subseteq V$ such that ${}_B\varphi_B = E_k \oplus (-E_l) \oplus \bigoplus_{i=1}^m A_{\omega_i} \in \mathbb{R}^{n\times n}$ is in **orthogonal normal form**, for unique $k, l, m \in \mathbb{N}_0$ and $0 < \omega_1 \leq \cdots \leq \omega_m < \pi$.

Proof. a) Let $C := \{v_1, \ldots, v_n\} \subseteq V$ be an orthonormal $\mathbb{C}$-basis such that $\varphi(v_i) = a_i v_i$, where $|a_i| = 1$. Then for any $v = \sum_{i=1}^n b_i v_i \in V$ we have $\|\varphi(v)\|^2 = \sum_{i=1}^n \langle a_i b_i v_i, a_i b_i v_i \rangle = \sum_{i=1}^n \langle b_i v_i, b_i v_i \rangle = \|v\|^2$, hence φ is unitary.

b) We show that for any irreducible polynomial $p \in \mathbb{R}[X]$ we have $\deg(p) \leq 2$: We may assume that $\deg(p) \geq 2$. The field automorphism $\overline{}\colon \mathbb{C} \to \mathbb{C}$ induces a ring automorphism $\overline{}\colon \mathbb{C}[X] \to \mathbb{C}[X]$, where since $p \in \mathbb{R}[X]$ we have $\overline{p} = p$. Thus whenever $X - a \mid p \in \mathbb{C}[X]$, where necessarily $a \in \mathbb{C} \setminus \mathbb{R}$, we have $X - \overline{a} \mid \overline{p} = p \in \mathbb{C}[X]$. Hence since $X - a, X - \overline{a} \in \mathbb{C}[X]$ are coprime we have $(X - a)(X - \overline{a}) \mid p \in \mathbb{C}[X]$, thus $p \sim X^2 - 2\mathrm{Re}(a)X + |a|^2 \in \mathbb{R}[X]$.

Let $p \in \mathbb{R}[X]$ be monic and irreducible such that $p \mid \chi_\varphi \in \mathbb{R}[X]$. Then we have $T_p(\varphi) \neq \{0\}$, and for some $0 \neq v \in T_p(\varphi)$ let $U := \langle \varphi^i(v) \in V; i \in \mathbb{N}_0 \rangle_\mathbb{R} \leq V$ be the φ-invariant $\mathbb{R}$-subspace **generated** by v. Then we have $V = U \oplus U^\perp$, where both U and $U^\perp$ are φ- and φ^*-invariant, and we may proceed by induction.

Let $\psi := \varphi|_U$. If $\deg(p) = 1$ then $p = X \pm 1 \in \mathbb{R}[X]$, hence $U = \langle v \rangle_\mathbb{R}$ and $\psi = \mp\mathrm{id}$. If $\deg(p) = 2$ then $U = \langle v, \psi(v) \rangle_\mathbb{R}$, and from $p(\psi) = 0$ we infer that $\mu_\psi = p \in \mathbb{R}[X]$, thus $\dim_\mathbb{R}(U) = 2$ and $\chi_\psi = \mu_\psi = p$ is irreducible, implying that $\psi \in \mathrm{SO}_2(\mathbb{R})$ is the rotation by the angle $0 < \omega < \pi$ given by $p_\omega = p$.

Finally, we have $k = \nu_1(\varphi) \in \mathbb{N}_0$ and $l = \nu_{-1}(\varphi) \in \mathbb{N}_0$, and for $0 < \omega < \pi$ we have $\nu_{p_\omega}(\chi_\varphi) = |\{i \in \{1, \ldots, m\}; \omega_i = \omega\}| \in \mathbb{N}_0$, implying uniqueness. ♯

(10.14) Corollary. a) φ is hermitian if and only if there is an orthonormal K-basis of V consisting of eigenvectors of φ and all its eigenvalues are in $\mathbb{R}$.
b) Let Ψ be a hermitian α-sesquilinear form on V, and let $G := {}_B\Psi_B \in K^{n\times n}$, where $B \subseteq V$ is any K-basis. Then we have $\sum_{a\in\mathbb{R}} \nu_a(G) = n$, and Ψ has

signature $[\sum_{a>0}\nu_a(G), \sum_{a<0}\nu_a(G)]$. The 1-dimensional K-subspaces of the eigenspace $T_a(G) \leq V$ are called **principal axes** of Ψ with respect to $a \in \mathbb{R}$. In particular, if $\dim_K(T_a(G)) = 1$ then the latter are uniquely determined.

Proof. a) Let $C := \{v_1, \ldots, v_n\} \subseteq V$ be an orthonormal K-basis such that $\varphi(v_i) = a_i v_i$, where $a_i \in \mathbb{R}$. Then we have $({}_C\varphi_C)^* = {}_C\varphi_C$, hence φ is hermitian.
b) Since G is hermitian, there is a unitary matrix $P \in \mathrm{GL}_n(K)$ such that $G' := P^{-1}GP = P^*GP = \mathrm{diag}[a_1, \ldots, a_n] \in K^{n\times n}$, where $a_i \in \mathbb{R}$, for all $i \in \{1, \ldots, n\}$. Hence on the one hand we have $\nu_a(G) = \nu_a(G')$ for all $a \in K$, where $\nu_a(G) > 0$ only if $a \in \mathbb{R}$, and on the other hand $G' = G_C^C(\Psi)$ is the Gram matrix of Ψ with respect to the K-basis $C \subseteq V$, where $M_B^C(\mathrm{id}) = P$. Hence replacing all non-isotropic vectors $v \in C$ by scalar multiples $v' := \frac{1}{\sqrt{|\Psi(v,v)|}} \cdot v$ yields a K-basis $C' \subseteq V$ such that $G_{C'}^{C'}(\Psi) = E_k \oplus (-E_l) \oplus (0 \cdot E_m)$, where $k = \sum_{a>0}\nu_a(G)$ and $l = \sum_{a<0}\nu_a(G)$; note that $m = \dim_K(V^\perp) = \nu_0(G)$. ♯

For example, for $K := \mathbb{R}$ and $\alpha = \mathrm{id}$, let Φ be given with respect to some $\mathbb{R}$-basis $B \subseteq \mathbb{R}^{3\times 1}$ by $G = G_B^B(\Phi) := \begin{bmatrix} 0 & -2 & 4 \\ -2 & 1 & -1 \\ 4 & -1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$. We have $\chi_G = X^3 - X^2 - 21X = X(X - a_+)(X - a_-) \in \mathbb{R}[X]$, hence we get the eigenvalues $a_0 := 0$ and $a_\pm := \frac{1}{2}(1 \pm \sqrt{85}) \in \mathbb{R}$. Hence we conclude that Φ has signature $[1, 1]$, as we have already observed in (10.4); note that is suffices to observe that $a_+a_- = -21 < 0$ to conclude that $a_- < 0 < a_+$.

Moreover, we get the principal axes $T_0(G) = \langle v_0\rangle_{\mathbb{R}}$ and $T_{a_\pm}(G) = \langle v_\pm\rangle_{\mathbb{R}}$, where $v_0 := [1, 4, 2]^{\mathrm{tr}} \in \mathbb{R}^{3\times 1}$ and $v_\pm := [\pm 4\sqrt{85}, -17 \mp \sqrt{85}, 34]^{\mathrm{tr}} \in \mathbb{R}^{3\times 1}$. Letting $\Gamma = \langle\cdot,\cdot\rangle$ denote the standard scalar product on $\mathbb{R}^{3\times 1}$, we indeed have $\langle v_0, v_\pm\rangle = 0 = \langle v_+, v_-\rangle$, as well as $\|v_0\|^2 = 21$ and $\|v_\pm\|^2 = 2890 \pm 34\sqrt{85}$. Hence $P := [v_+, v_-, v_0] \cdot \mathrm{diag}[\frac{1}{\|v_+\|}, \frac{1}{\|v_-\|}, \frac{1}{\|v_0\|}] \in \mathrm{GL}_3(\mathbb{R})$ is orthogonal, that is fulfills $P^{-1} = P^{\mathrm{tr}}$, and thus we get $P^{\mathrm{tr}}GP = P^{-1}GP = \mathrm{diag}[a_+, a_-, 0] \in \mathbb{R}^{3\times 3}$.

(10.15) Example: Binary quadrics. Let $V := \mathbb{R}^{2\times 1}$ be equipped with the standard scalar product, and let $B \subseteq V$ be the orthonormal standard $\mathbb{R}$-basis. Moreover, let Φ be a symmetric $\mathbb{R}$-bilinear form on V, having Gram matrix $G := {}_B\Phi_B = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$. Let $q\colon V \to \mathbb{R}$ be the quadratic form associated with Φ, which hence is given by $q(x, y) = [x, y]\cdot G\cdot[x, y]^{\mathrm{tr}} = ax^2 + bxy + cy^2$, where $[x, y]^{\mathrm{tr}} \in \mathbb{R}^{2\times 1}$ is the coordinate tuple with respect to B. We aim to describe the associated **real binary quadric** $\mathcal{V}(q, \rho) := \{[x, y] \in \mathbb{R}^2; q(x, y) = \rho\}$, where $\rho \in \mathbb{R}$. In order to do so, we consider $\mathcal{V}(q, \rho)$ in terms of an orthonormal $\mathbb{R}$-basis $C \subseteq V$ given by the principal axes of Φ:

We have $\chi_G = X^2 - (a + c)X + (ac - \frac{b^2}{4}) \in \mathbb{R}[X]$, yielding the eigenvalues $\lambda_\pm = \frac{1}{2}(a + c \pm \sqrt{d}) \in \mathbb{R}$, where $d := (a - c)^2 + b^2$. Note that $d = 0$ if and only if $a = c$ and $b = 0$, that is $G = aE_2$; in this case we have $\lambda_\pm = a$, and the principal axes are not uniquely defined.

We are going to determine the signature of Φ: We have $\det(G) = ac - \frac{b^2}{4} = \lambda_+\lambda_-$ and $\mathrm{Tr}(G) = a + c = \lambda_+ + \lambda_-$. Hence, if $\det(G) > 0$ then $\lambda_\pm$ are both positive or both negative, and we can read off from $\mathrm{Tr}(G)$ which case occurs; if $\det(G) = 0$ then at least one of $\lambda_\pm$ vanishes, and again we can read off from $\mathrm{Tr}(G)$ which ones vanish; if $\det(G) < 0$ then $\lambda_\pm$ have different signs, hence we have $\lambda_- < 0 < \lambda_+$. Thus we have the following signatures:

$\det(G)$	$\mathrm{Tr}(G)$	λ_+	λ_-	signature	
> 0	> 0	> 0	> 0	$[2, 0]$	positive definite
	< 0	< 0	< 0	$[0, 2]$	negative definite
$= 0$	> 0	> 0	$= 0$	$[1, 0]$	positive semi-definite
	$= 0$	$= 0$	$= 0$	$[0, 0]$	zero form
	< 0	$= 0$	< 0	$[0, 1]$	negative semi-definite
< 0		> 0	< 0	$[1, 1]$	Minkowski

We proceed to find an $\mathbb{R}$-basis $C \subseteq V$ as described above: If $b = 0$, then G already is a diagonal matrix; hence letting C be given by $P := {}_B\mathrm{id}_C = E_2$ or $P := {}_B\mathrm{id}_C = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathrm{GO}_2(\mathbb{R})$, we may assume that $\lambda_+ = a \geq c = \lambda_-$.

Assuming $b \neq 0$, we get the uniquely defined principal axes

$$T_{\lambda_\pm}(G) = \ker\left(\begin{bmatrix} a - c \mp \sqrt{d} & b \\ b & c - a \mp \sqrt{d} \end{bmatrix}\right) = \langle v_\pm \rangle_{\mathbb{R}},$$

where $v_+ := [a - c + \sqrt{d}, b]^{\mathrm{tr}}$ and $v_- := [-b, a - c + \sqrt{d}]^{\mathrm{tr}}$; note that with respect to the standard scalar product we indeed have $\langle v_+, v_- \rangle = 0$. We have $\lambda := \|v_+\|^2 = \|v_-\|^2 = 2d + 2(a - c)\sqrt{d}$, hence in this case we let C be given by $P := {}_B\mathrm{id}_C = \frac{1}{\sqrt{\lambda}} \cdot [v_+, v_-] = \frac{1}{\sqrt{\lambda}} \cdot \begin{bmatrix} a - c + \sqrt{d} & -b \\ b & a - c + \sqrt{d} \end{bmatrix} \in \mathrm{GO}_2(\mathbb{R})$.

Thus in any case we get $G' := {}_C\Phi_C = \mathrm{diag}[\lambda_+, \lambda_-]$, and the quadratic form becomes $q(x', y') = [x', y'] \cdot G' \cdot [x', y']^{\mathrm{tr}} = \lambda_+ x'^2 + \lambda_- y'^2$, where $[x', y']^{\mathrm{tr}} \in \mathbb{R}^{2 \times 1}$ is the coordinate tuple with respect to C. We now describe $\mathcal{V}(q, \rho)$ in the various cases, where by going over to $-\Phi$ if necessary, and observing that $\mathcal{V}(-q, \rho) = \mathcal{V}(q, -\rho)$, it suffices to consider the signatures $[2, 0]$, $[1, 1]$, $[1, 0]$ and $[0, 0]$, which we do in turn, letting $\rho > 0$:

i) If $\lambda_+ > \lambda_- > 0$, we infer that $\mathcal{V}(q, \rho)$ is an **ellipsis**, with **half axes** $\sqrt{\frac{\rho}{\lambda_\pm}}$, where for $d > 0$ the latter are different, while for $d = 0$ we have $\lambda_\pm = a$, in which case the ellipsis has two equal half axes, hence becomes the **circle** with **radius** $\sqrt{\frac{\rho}{a}}$; moreover, $\mathcal{V}(q, 0) = \{0\}$ is a single point, and $\mathcal{V}(q, -\rho) = \emptyset$. For $\lambda_- \to 0^+$ the associated half axe tends towards ∞, hence the ellipsis degenerates to a pair of vertical lines; for $\rho \to 0^+$ the ellipsis degenerates to a point.

ii) If $\lambda_+ > 0 > \lambda_-$, we infer that both $\mathcal{V}(q, \rho)$ and $\mathcal{V}(q, -\rho)$ are a **hyperbola**, the former with **vertices** $[\pm\sqrt{\frac{\rho}{\lambda_+}}, 0]^{\mathrm{tr}}$, the latter with vertices $[0, \pm\sqrt{\frac{\rho}{\lambda_-}}]^{\mathrm{tr}}$.

Moreover, $\mathcal{V}(q,0) = \{[x', y'] \in \mathbb{R}^2; y' = \pm\sqrt{-\frac{\lambda_+}{\lambda_-}}\cdot x'\}$ is the union of two different intersecting lines, being called the **asymptotic lines** of the hyperbolas; the latter are orthogonal to each other if and only if $\lambda_+ = \lambda_-$, that is $d = 0$, in which case we have $\mathcal{V}(q,0) = \{[x', y'] \in \mathbb{R}^2; y' = \pm x'\}$. For $\lambda_- \to 0^-$ the hyperbolas degenerate to a pair of vertical lines and to the empty set, respectively; for $\lambda_+ \to 0^+$ they degenerate to the empty set and to a pair of horizontal lines, respectively; for $\rho \to 0^+$ the hyperbolas degenerate to the asymptotic lines.

iii) If $\lambda_+ > 0 = \lambda_-$ we infer that $\mathcal{V}(q,\rho) = \{\pm\sqrt{\frac{\rho}{\lambda_+}}\} \times \mathbb{R}$ is a union of two different vertical lines; moreover, $\mathcal{V}(q,0) = \{0\} \times \mathbb{R}$ is a vertical line, and $\mathcal{V}(q,-\rho) = \emptyset$. For $\lambda_+ \to 0^+$ the pair of vertical lines degenerates to an empty set; for $\rho \to 0^+$ the pair of vertical lines degenerates to a single vertical line.

iv) If $\lambda_+ = \lambda_- = 0$ we infer that $\mathcal{V}(q,0) = \mathbb{R}^2$, while $\mathcal{V}(q,\rho) = \emptyset = \mathcal{V}(q,-\rho)$.

In particular, for $a = c$ and $b \neq 0$ we have $d = b^2 > 0$, hence $\lambda_\pm = a \pm \frac{|b|}{2}$, implying $\det(G) = a^2 - \frac{b^2}{4}$ and $\mathrm{Tr}(G) = 2a$, and as base change matrix we may take $P = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. For example, $G := \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ has eigenvalues $[3, 1]$, $G := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ has eigenvalues $[3, -1]$, and $G := \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ has eigenvalues $[2, 0]$.

III Exercises and references

11 Exercises for Part I (in German)

(11.1) Aufgabe: Rechnen mit Mengen.
Es seien L, M und N Mengen. Man zeige:
a) Es gelten $(L \cap M) \cap N = L \cap (M \cap N)$ und $M \cap N = N \cap M$ sowie $M \cap M = M$ und $M \cap \emptyset = \emptyset$.
b) Es gelten $(L \cup M) \cup N = L \cup (M \cup N)$ und $M \cup N = N \cup M$ sowie $M \cup M = M$ und $M \cup \emptyset = M$.
c) Es gelten $M \cap (M \cup N) = M$ und $M \cup (M \cap N) = M$.
d) Ist $L \subseteq N$, so gilt die **Dedekind-Identität** $(L \cup M) \cap N = L \cup (M \cap N)$.

(11.2) Aufgabe: Äquivalenzrelationen.
Es sei R eine symmetrische transitive Relation auf der Menge M mit folgender Zusatzeigenschaft: Für jedes $x \in M$ gibt es ein $y \in M$ mit $[x, y] \in R$. Man zeige: R ist eine Äquivalenzrelation. Kann man auf die Voraussetzung der Zusatzeigenschaft verzichten?

(11.3) Aufgabe: Ordnungsrelationen.
a) Es sei $\leq$ eine reflexive transitive Relation auf der Menge M, die zudem **antisymmetrisch** sei, das heißt, für $x, y \in M$ mit $x \leq y$ und $y \leq x$ gilt stets bereits $x = y$. Solche Relationen werden als **partielle Ordnungen** bezeichnet; man schreibt $x < y$, falls $x \leq y$ und $x \neq y$ gilt.

Man zeige: Gilt $x \leq y$ oder $y \leq x$ für alle $x, y \in M$, so gilt für alle $x, y \in M$ genau eine der Beziehungen $x < y$ oder $x = y$ oder $y < x$. Solche Relationen werden als **Ordnungen** bezeichnet.

b) Man zeige: Die Potenzmenge $\mathcal{P}(M)$ einer Menge M ist durch $\subseteq$ partiell geordnet. Wann ist dies eine Ordnung?

c) Man zeige: Die Menge $\mathbb{N}$ ist sowohl durch $\{[x, y] \in \mathbb{N}^2; x \mid y\}$ als auch durch $\leq$ partiell geordnet. Sind dies Ordnungen? Was gilt für $\geq$?

(11.4) Aufgabe: Arithmetik auf $\mathbb{N}_0$.
a) Man zeige: Durch die durch $0 + 0 := 0$ und $a + (b + 1) := (a + b) + 1$ sowie $(a+1)+b := (a+b)+1$ für alle $a, b \in \mathbb{N}_0$ definierte Addition $+ \colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ wird $\mathbb{N}_0$ zu einem kommutativen Monoid. Wird $\mathbb{N}_0$ damit zu einer Gruppe?
b) Man zeige: Durch die durch $0 \cdot 0 := 0$ und $a \cdot (b + 1) := (a \cdot b) + a$ sowie $(a+1) \cdot b := (a \cdot b) + b$ für alle $a, b \in \mathbb{N}_0$ definierte Multiplikation $\cdot \colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ wird $\mathbb{N}_0$ zu einem kommutativen Monoid. Wird $\mathbb{N}_0$ damit zu einer Gruppe?
c) Man zeige die Distributivgesetze $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ und $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ für alle $a, b, c \in \mathbb{N}_0$. Wird $\mathbb{N}_0$ damit zu einem Ring?
d) Man zeige man die folgenden Kürzungsregeln: Für alle $a, b, c \in \mathbb{N}_0$ gilt $a = b$ genau dann, wenn $a + c = b + c$ gilt; und für alle $a, b \in \mathbb{N}_0$ und $c \in \mathbb{N}$ gilt $a = b$ genau dann, wenn $a \cdot c = b \cdot c$ gilt.

(11.5) Aufgabe: Vollständige Induktion.
a) Für $n \in \mathbb{N}$ zeige man: Es gilt $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$.
b) Für $n \in \mathbb{N}$ zeige man: Es gilt $1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$. Man gebe einen geschlossenen Ausdruck für die Summe der ersten n Kubikzahlen an.
c) Man gebe geschlossene Ausdrücke für die folgenden Summen an, wobei $n \in \mathbb{N}$:

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} \quad \text{und} \quad 1 + 3 + \ldots + (2n-1).$$

(11.6) Aufgabe: Potenzmenge.
Es seien M eine beliebige Menge und $\mathcal{P}(M)$ ihre Potenzmenge. Man zeige: Es gibt keine Bijektion $M \to \mathcal{P}(M)$.

Hinweis. Für $f \colon M \to \mathcal{P}(M)$ bijektiv betrachte man $\{x \in M; x \notin f(x)\}$.

(11.7) Aufgabe: Endliche Mengen.
Es seien M und N endliche Mengen sowie $m := |M| \in \mathbb{N}_0$ und $n := |N| \in \mathbb{N}_0$.
a) Man zeige: Es gelten $|M \times N| = mn$ und $|\mathrm{Abb}(M, N)| = n^m$ sowie außerdem $|\{f \in \mathrm{Abb}(M, N); f \text{ injektiv}\}| = n(n-1) \cdots (n-m+1) = \frac{n!}{(n-m)!}$.
b) Man zeige: Für $n \neq 0$ gilt $|\{U \subseteq N; |U| \text{ gerade}\}| = |\{U \subseteq N; |U| \text{ ungerade}\}|$. Wieviele Elemente haben die genannten Mengen also jeweils?
c) Für $k \in \{0, \ldots, n\}$ sei $\binom{n}{k} := \frac{n!}{k!(n-k)!} \in \mathbb{Q}$ der zugehörige **Binomialkoeffizient**. Für $\mathcal{N}_k := \{U \subseteq N; |U| = k\}$ zeige man: Es gilt $|\mathcal{N}_k| = \binom{n}{k}$, und folgere $\binom{n}{k} \in \mathbb{N}$. Welche Identitäten folgen aus $|\mathcal{P}(N)| = 2^n$ und der Aussage in b)?

(11.8) Aufgabe: Abbildungen.
Es seien L, M, N Mengen, sowie $f \colon M \to N$ und $g \colon L \to M$ Abbildungen.
a) Man gebe ein Beispiel für die folgende Situation an: Es sind f nicht injektiv, und g nicht surjektiv, aber dennoch ist fg bijektiv.
b) Nun seien sowohl f als auch g bijektiv. Man gebe die Umkehrabbildung von fg als Ausdruck in f^{-1} und g^{-1} an.

(11.9) Aufgabe: Symmetrische Gruppe $\mathcal{S}_3$.
Man stelle die Verknüpfungstafel der symmetrischen Gruppe $\mathcal{S}_3$ auf und gebe alle Untergruppen von $\mathcal{S}_3$ an.

(11.10) Aufgabe: Symmetrische Gruppe $\mathcal{S}_4$.
a) Man zeige: Die Menge $V_4 := \{(), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ ist eine Untergruppe von $\mathcal{S}_4$; sie heißt **Kleinsche Vierergruppe**. Ist V_4 kommutativ? Man stelle die Verknüpfungstafel auf und gebe alle Untergruppen von V_4 an.
b) Man bestimme die kleinste Untergruppe D_8 von $\mathcal{S}_4$, die V_4 als Teilmenge und das Element $(1,2,3,4)$ enthält; sie heißt **Diedergruppe**. Wieviele Elemente hat D_8? Ist D_8 kommutativ? Man stelle die Verknüpfungstafel auf und gebe alle Untergruppen von D_8 an.
c) Man bestimme möglichst viele verschiedene Untergruppen von $\mathcal{S}_4$. Wieviele Elemente haben sie jeweils? Welche von ihnen sind kommutativ?

(11.11) Aufgabe: Ringe.
Es sei R ein Ring. Man zeige:
a) Sind $a, b \in R$ mit $ab = ba$, so gilt $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ für alle $n \in \mathbb{N}_0$.
b) Die Menge R wird auch durch die Multiplikation $a \circ b := ba$, für alle $a, b \in R$, zu einem Ring, dem **oppositären** Ring R°. Wann ist $R = R^\circ$?

(11.12) Aufgabe: Ordnungen auf $\mathbb{Z}$.
a) Man zeige: Auf $\mathbb{Z}$ wird durch $[a, b]_\Delta \leq [a', b']_\Delta$ falls $a + b' \leq b + a'$, eine Ordnung definiert, die die natürliche Ordnung auf $\mathbb{N}_0$ fortsetzt. Man gebe die Mengen $\{x \in \mathbb{Z}; x > 0\}$ und $\{x \in \mathbb{Z}; x < 0\}$ an.
b) Man zeige, daß $\mathbb{Z}$ ein **geordneter Ring** ist, das heißt, für alle $x, y, z \in \mathbb{Z}$ gilt $x \leq y$ genau dann, wenn $x + z \leq y + z$; ist $z > 0$, so gilt $x \leq y$ genau dann, wenn $xz \leq yz$; ist $z < 0$, so gilt $x \leq y$ genau dann, wenn $xz \geq yz$.

(11.13) Aufgabe: Ordnungen auf $\mathbb{Q}$.
a) Man zeige: Auf $\mathbb{Q}$ wird durch $\frac{a}{b} \leq \frac{a'}{b'}$ falls $ab' \leq ba'$ und $bb' > 0$, oder $ab' \geq ba'$ und $bb' < 0$, eine Ordnung definiert, die die natürliche Ordnung auf $\mathbb{Z}$ fortsetzt. Man gebe die Mengen $\{x \in \mathbb{Q}; x > 0\}$ und $\{x \in \mathbb{Q}; x < 0\}$ an.
b) Man zeige, daß $\mathbb{Q}$ ein **geordneter Körper** ist, das heißt, für alle $x, y, z \in \mathbb{Q}$ gilt $x \leq y$ genau dann, wenn $x + z \leq y + z$; ist $z > 0$, so gilt $x \leq y$ genau dann, wenn $xz \leq yz$; ist $z < 0$, so gilt $x \leq y$ genau dann, wenn $xz \geq yz$.

(11.14) Aufgabe: Restklassenringe.
Für $n \in \{1, \ldots, 9\}$ stelle man die Verknüpfungstafeln der Addition und der Multiplikation in $\mathbb{Z}/n\mathbb{Z}$ auf. Man bestimme die Einheitengruppe $(\mathbb{Z}/n\mathbb{Z})^*$, und gebe zu allen Einheiten jeweils das multiplikativ inverse Element an.

(11.15) Aufgabe: Fermat-Zahlen [Landry, 1880].
Es sei $F_6 := 2^{2^6} + 1 = 18\,446\,744\,073\,709\,551\,617$ die sechste Fermat-Zahl. Man zeige: Es gilt $274\,177 \mid F_6$.

(11.16) Aufgabe: Endliche Primkörper.
Es sei $p \in \mathbb{N}$ eine Primzahl. Man zeige:
a) Es gilt der **Satz von Wilson** $(p-1)! \equiv -1 \pmod{p}$.
b) Ist p ungerade, so gilt $((\frac{p-1}{2})!)^2 \equiv (-1)^{\frac{p+1}{2}} \pmod{p}$.

Hinweis zu a). Man zeige, daß $X^2 = 1$ in $\mathbb{Z}_p$ genau zwei Lösungen hat.

(11.17) Aufgabe: Einheitswurzeln in $\mathbb{C}$.
a) Man zeige: Die Menge $U := \{z \in \mathbb{C}; |z| = 1\}$ und die Menge $U_n := \{z \in \mathbb{C}; z^n = 1\}$ der n-ten **Einheitswurzeln**, für $n \in \mathbb{N}$, sind Untergruppen von $\mathbb{C}^*$. Man stelle U in der Gauß-Ebene dar.
b) Für $n \in \{2, 3, 4, 6, 8, 12\}$ bestimme man die Menge U_n, und stelle sie in der Gauß-Ebene dar. Wieviele Elemente hat U_n jeweils?

(11.18) Aufgabe: Hamilton-Quaternionen.
Es seien $\mathbb{H} := \mathbb{R}^4$ die Menge der **Hamilton-Quaternionen**, und weiter $1 := [1,0,0,0] \in \mathbb{H}$ und $i := [0,1,0,0] \in \mathbb{H}$ sowie $j := [0,0,1,0] \in \mathbb{H}$ und $k := [0,0,0,1] \in \mathbb{H}$; dann hat jedes Element $a \in \mathbb{H}$ eine eindeutige Darstellung der Form $a = a_0 + ia_1 + ja_2 + ka_3 \in \mathbb{H}$, wobei $a_1, \ldots, a_4 \in \mathbb{R}$.

Man zeige: Durch komponentenweise Addition, und die durch $i^2 = j^2 = k^2 = -1$ sowie $ij = k$ und $jk = i$ und $ki = j$ festgelegte Multiplikation wird $\mathbb{H}$ zu einem nichtkommutativen Schiefkörper. Wie kann man $\mathbb{C}$ mit Teilkörpern von $\mathbb{H}$ identifizieren?

(11.19) Aufgabe: Symmetrische Differenz.
Die **symmetrische Differenz** der Mengen M und N ist definiert als $M \vartriangle N := (M \setminus N) \,\dot{\cup}\, (N \setminus M) = (M \cup N) \setminus (M \cap N)$. Man zeige: Ist U eine Menge, so kann $\mathcal{P}(U)$ zusammen mit der symmetrischen Differenz $\vartriangle$ als Addition als $\mathbb{Z}_2$-Vektorraum aufgefaßt werden.

(11.20) Aufgabe: International Standard Book Number (ISBN).
Es sei $\mathcal{C} := \{[a_1, \ldots, a_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} ia_i = 0 \in \mathbb{Z}_{11}\}$ der $\mathbb{Z}_{11}$-Vektorraum des ISBN-10-Standards. Kann man **Zwillingsfehler** der Form $aa \to bb$ und **Sprungzwillingsfehler** der Form $aca \to bcb$ immer erkennen?

(11.21) Aufgabe: Folgenräume.
Es sei $V := \mathrm{Abb}(\mathbb{N}, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller Zahlenfolgen $[x_j \in \mathbb{R}; j \in \mathbb{N}]$. Welche der folgenden Teilmengen von V sind $\mathbb{R}$-Teilräume?
a) $\{[x_j] \in V; \text{es gibt ein } c \geq 0 \text{ mit } |x_j| \leq c \text{ für alle } j \in \mathbb{N}\}$
b) $\{[x_j] \in V; x_j \neq 0 \text{ für nur endlich viele } j \in \mathbb{N}\}$
c) $\{[x_j] \in V; x_j \neq 0 \text{ für unendlich viele } j \in \mathbb{N}\}$
d) $\{[x_j] \in V; x_{j+1} = x_j + a \text{ für alle } j \in \mathbb{N}\}$, wobei $a \in \mathbb{R}$.
e) $\{[x_j] \in V; x_{j+2} = x_{j+1} + x_j \text{ für alle } j \in \mathbb{N}\}$

(11.22) Aufgabe: Vereinigung von Teilräumen.
a) Es seien K ein Körper, V ein K-Vektorraum, und $U_1, \ldots, U_n < V$, wobei $n \in \mathbb{N}$. Man zeige: Ist K unendlich oder gilt $|K| \geq n$, so ist $\bigcup_{i=1}^n U_i \subset V$.
b) Es seien K ein endlicher Körper und $V := K^n$ für ein $n \geq 2$. Weiter seien $U_a := \{[a_1, \ldots, a_{n-1}, aa_n] \in K^n; a_1, \ldots, a_n \in K\}$, für $a \in K$, und $U := \{[a_1, \ldots, a_{n-2}, 0, a_n] \in K^n; a_1, \ldots, a_{n-2}, a_n \in K\}$. Man zeige: Es gilt $U < V$ und $U_a < V$, für alle $a \in K$, sowie $(\bigcup_{a \in K} U_a) \cup U = V$. Was bedeutet das für die Aussage in a)?

Hinweis zu a). Induktion nach n und Schubfachprinzip.

(11.23) Aufgabe: Lineare Abhängigkeit.
Es seien K ein Körper, V ein K-Vektorraum und $\mathcal{I}$ eine Menge. Man zeige: Eine Folge $\mathcal{S} := [v_i \in V; i \in \mathcal{I}]$ von Vektoren ist genau dann K-linear abhängig, wenn es ein $i \in \mathcal{I}$ gibt, so daß $v_i \in \langle v_j; i \neq j \in \mathcal{I} \rangle_K$ gilt.

(11.24) Aufgabe: Lineare Unabhängigkeit.
Es seien $p_1, \ldots, p_n \in \mathbb{N}$ paarweise verschiedene Primzahlen, wobei $n \in \mathbb{N}_0$. Man zeige: $\{1, \sqrt{p_1}, \ldots, \sqrt{p_n}\} \subseteq \mathbb{R}$ ist $\mathbb{Q}$-linear unabhängig.

(11.25) Aufgabe: Lineare Abhängigkeit in Matrixräumen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und folgende Teilmengen von $K^{2\times 2}$ gegeben:

$$\begin{aligned} S_1 &:= \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & -3 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix} \right\} \\ S_2 &:= \left\{ \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} \right\} \end{aligned}$$

Ferner seien die folgenden Matrizen gegeben:

$$A_1 := \begin{bmatrix} 11 & 0 \\ -9 & 5 \end{bmatrix}, \quad A_2 := \begin{bmatrix} 11 & -7 \\ -9 & 5 \end{bmatrix}, \quad A_3 := \begin{bmatrix} -2 & 3 \\ 0 & -1 \end{bmatrix}$$

Man entscheide für $i \in \{1, 2\}$ und $j \in \{1, 2, 3\}$, ob A_j als K-Linearkombination der Elemente von S_i geschrieben werden kann, und gebe gegebenenfalls eine solche Linearkombination an. Wie hängen die Rechnungen von K ab?

(11.26) Aufgabe: Lineare Abhängigkeit mit Parameter.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Wie muß $\lambda \in K$ gewählt werden, damit die folgende Menge $T_\lambda \subseteq K^{2\times 2}$ K-linear abhängig ist? Hängen die Rechnungen von K ab?

$$T_\lambda := \left\{ \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & \lambda \end{bmatrix} \right\}$$

(11.27) Aufgabe: Lineare Unabhängigkeit in Funktionenräumen.
a) Es sei $C^{\text{pol}}(\mathbb{R}) \subseteq C^\infty(\mathbb{R})$ die Menge der **Polynomfunktionen**, und für $k \in \mathbb{N}_0$ seien $p_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto x^k$. Man zeige: Es gilt $C^{\text{pol}}(\mathbb{R}) \leq C^\infty(\mathbb{R})$ und $P := \{p_k \in C^{\text{pol}}(\mathbb{R}); k \in \mathbb{N}_0\}$ ist eine $\mathbb{R}$-Basis von $C^{\text{pol}}(\mathbb{R})$. Welche Teilmengen von P sind ebenfalls $\mathbb{R}$-Basen von $C^{\text{pol}}(\mathbb{R})$? Ist P eine $\mathbb{R}$-Basis von $C^\infty(\mathbb{R})$?
b) Für $k \in \mathbb{N}$ seien $s_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \sin(kx)$ und $c_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \cos(kx)$. Man zeige: Die Menge $\{s_k, c_k \in C^\infty(\mathbb{R}); k \in \mathbb{N}\}$ ist $\mathbb{R}$-linear unabhängig. Ist sie eine $\mathbb{R}$-Basis von $C^\infty(\mathbb{R})$?

(11.28) Aufgabe: Basen in Funktionenräumen.
Für einen Körper K sei $\text{Abb}'(\mathcal{I}, K) := \{f \in \text{Abb}(\mathcal{I}, K); f^{-1}(K \setminus \{0\}) \text{ endlich}\}$, für eine Menge $\mathcal{I}$. Man zeige: Es ist $\text{Abb}'(\mathcal{I}, K) \leq \text{Abb}(\mathcal{I}, K)$. Man gebe eine K-Basis von $\text{Abb}'(\mathcal{I}, K)$ an. Wann hat $\text{Abb}'(\mathcal{I}, K)$ eine endliche K-Basis?

(11.29) Aufgabe: Basisergänzung.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, sowie $S_1 := \{[4, 1, 2, -1], [2, 1, 2, -1]\} \subseteq K^4$ und $S_2 := \{[2, -1, 0, 0], [3, 0, 4, 0]\} \subseteq K^4$. Man zeige, daß S_1 und S_2 K-linear unabhängig sind, und ergänze sie mit Hilfe der Standardbasis von K^4 zu K-Basen von K^4.

(11.30) Aufgabe: Linear abhängige Teilmengen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $S := \{[1,2,3,4], [4,3,2,1], [1,1,1,1]\} \subseteq K^4$.
a) Man gebe die K-linear unabhängigen Teilmengen $T \subseteq S$ an. Welche sind maximal? Man bestimme $\dim_K(\langle S \rangle_K)$. Hängen die Rechnungen von K ab?
b) Man ergänze die obigen Teilmengen $T \subseteq S$ jeweils zu K-Basen von K^4, und stelle die Einheitsvektoren $e_1, \ldots, e_4 \in K^4$ als K-Linearkombinationen dar.

(11.31) Aufgabe: Basen in Matrixräumen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, und $B := \{v_1, \ldots, v_4\} \subseteq K^{2 \times 2}$, wobei

$$v_1 := \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad v_2 := \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}, \quad v_3 := \begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix}, \quad v_4 := \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}.$$

a) Man zeige, daß B eine K-Basis von $K^{2 \times 2}$ ist, und bestimme den Koordinatenvektor $A_B \in K^4$ für eine beliebige Matrix $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in K^{2 \times 2}$.
b) Es sei $U := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in K^{2 \times 2}; a - b + c - d = b + c - 3d = 0 \right\}$. Man zeige, daß U ein K-Teilraum von $K^{2 \times 2}$ ist, gebe eine K-Basis von U an, und ergänze diese mit Hilfe von B zu einer K-Basis von $K^{2 \times 2}$.

(11.32) Aufgabe: Teilräume von Matrixräumen.
Es seien

$$\begin{aligned} U_1 &:= \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in \mathbb{R}^{2 \times 3}; a + b + c = d + e + f = a + c + e = 0 \right\}, \\ U_2 &:= \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in \mathbb{R}^{2 \times 3}; a - 2d = 2b - e = c - 2f = 0 \right\}. \end{aligned}$$

a) Man zeige, daß U_1 und U_2 $\mathbb{R}$-Teilräume von $\mathbb{R}^{2 \times 3}$ sind, und bestimme $\mathbb{R}$-Basen für U_1 und U_2. Wie erhält man daraus eine $\mathbb{R}$-Basis für $U_1 + U_2$?
b) Außerdem bestimme man eine $\mathbb{R}$-Basis für $U_1 \cap U_2$, und ergänze sie zu $\mathbb{R}$-Basen für U_1 bzw. U_2.

(11.33) Aufgabe: Magische Quadrate.
Eine quadratische Anordnung

a	b	c
d	e	f
g	h	i

neun reeller Zahlen $a, \ldots, i \in \mathbb{R}$ heißt **magisches Quadrat**, wenn die Summe jeder Zeile, jeder Spalte sowie der beiden Diagonalen des Quadrats jeweils Null ist. Es sei

$$\mathcal{M} := \left\{ [a, \ldots, i] \in \mathbb{R}^9; \begin{array}{|c|c|c|} \hline a & b & c \\ \hline d & e & f \\ \hline g & h & i \\ \hline \end{array} \text{ ist magisches Quadrat} \right\}.$$

a) Man zeige, daß $\mathcal{M} \subseteq \mathbb{R}^9$ ein $\mathbb{R}$-Teilraum ist, und gebe eine $\mathbb{R}$-Basis B an.
b) Man bestimme den Koordinatenvektor A_B für das magische Quadrat

$$A := \begin{array}{|c|c|c|} \hline -1 & 4 & -3 \\ \hline -2 & 0 & 2 \\ \hline 3 & -4 & 1 \\ \hline \end{array}$$

(11.34) Aufgabe: Endliche Vektorräume.
Es seien K ein endlicher Körper mit $q := |K| \in \mathbb{N}$, und V ein K-Vektorraum mit $\dim_K(V) = n \in \mathbb{N}_0$. Man zeige:
a) Es gilt $|V| = q^n$, und V hat genau $b_n(q) := \prod_{i=0}^{n-1}(q^n - q^i)$ verschiedene K-Basisfolgen. Was folgt daraus für die Anzahl der Elemente von $\mathrm{GL}(V)$?
b) Für $m \in \{0, \ldots, n\}$ hat V genau

$$s_{n,m}(q) = \prod_{i=0}^{m-1} \frac{q^n - q^i}{q^m - q^i} = \frac{1}{q^{m(n-m)}} \cdot \frac{b_n(q)}{b_{n-m}(q) b_m(q)}$$

verschiedene m-dimensionalen K-Teilräume, und es gilt $s_{n,m}(q) = s_{n,n-m}(q)$.
c) Ein m-dimensionaler K-Teilraum hat genau $q^{m(n-m)}$ Komplemente in V.

(11.35) Aufgabe: Teilräume von K^n.
Es seien K ein Körper und $V := K^n$ für ein $n \in \mathbb{N}$.
a) Es seien $U := \{[a, \ldots, a] \in V; a \in K\}$ und $W := \{[a_1, \ldots, a_n] \in V; \sum_{i=1}^n a_i = 0\}$. Man zeige: Es gilt $U \leq V$ und $W \leq V$. Man berechne $U \cap W$ und $U + W$, und gebe die jeweiligen K-Dimensionen an. Wann gilt $V = U \oplus W$?
b) Für $i \in \{1, \ldots, n\}$ sei $v_i := [1, \ldots, 1, 0, 1, \ldots, 1] \in V$ mit 0 als i-tem Eintrag. Man bestimme $\dim_K(\langle v_1, \ldots, v_n \rangle_K)$. Hängen die Rechnungen von K ab?

(11.36) Aufgabe: Schnitte von Teilräumen.
Es seien K ein Körper und V ein K-Vektorraum mit $\dim_K(V) = n \in \mathbb{N}$.
a) Es sei $k \in \mathbb{N}$, und für $i \in \{1, \ldots, k\}$ sei $U_i < V$ mit $\dim_K(U_i) = n - 1$. Man zeige: Es gilt $\dim_K(\bigcap_{i=1}^k U_i) \geq n - k$.
b) Man zeige: Ist $W < V$ mit $\dim_K(W) = n - k$, für ein $k \in \{1, \ldots, n\}$, so gibt es K-Teilräume $U_1, \ldots, U_k < V$ mit $\dim_K(U_i) = n - 1$ und $W = \bigcap_{i=1}^k U_i$.

(11.37) Aufgabe: Direkte Summen.
Es seien K ein Körper, V ein K-Vektorraum, $\mathcal{I}$ eine Menge, sowie $U_i \leq V$ für $i \in \mathcal{I}$ mit $\sum_{i \in \mathcal{I}} U_i = V$. Man zeige: Es ist genau dann $V = \bigoplus_{i \in \mathcal{I}} U_i$, wenn $U_i \cap \sum_{i \neq j \in \mathcal{I}} U_i = \{0\}$ für alle $i \in \mathcal{I}$ gilt.

(11.38) Aufgabe: Lineare Abbildungen in Funktionenräumen.
Man zeige: Die Ableitung $\frac{\partial}{\partial x}$ und das Integral $\int_0 \colon f \mapsto (x \mapsto \int_0^x f)$ sind $\mathbb{R}$-Endomorphismen von $C^{\mathrm{pol}}(\mathbb{R})$. Man beschreibe $\frac{\partial}{\partial x}$ und $\int_0$ mittels der $\mathbb{R}$-Basis $\{p_k \in C^{\mathrm{pol}}(\mathbb{R}); k \in \mathbb{N}_0\}$ von $C^{\mathrm{pol}}(\mathbb{R})$, wobei $p_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto x^k$, bestimme jeweils Kern und Bild, und untersuche auf Injektivität und Surjektivität.

(11.39) Aufgabe: Lineare Abbildungen in Folgenräumen.
Es seien K ein Körper und $V := \mathrm{Abb}(\mathbb{N}, K)$ sowie $\sigma\colon V \to V\colon [a_1, a_2, \ldots] \mapsto [a_2, a_3, \ldots]$ und $\tau\colon V \to V\colon [a_1, a_2, \ldots] \mapsto [0, a_1, a_2, \ldots]$. Man zeige: Es sind $\sigma, \tau \in \mathrm{End}_K(V)$. Man bestimme jeweils Kern und Bild, und untersuche auf Injektivität und Surjektivität.

(11.40) Aufgabe: Existenz linearer Abbildungen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ sowie $v_1 := [1, 1, 0] \in K^3$ und $v_2 := [0, 1, 1] \in K^3$, sowie $w_1 := [1, 2, 3] \in K^3$ und $w_2 := [3, 1, 2] \in K^3$. Für **i)** $v := [1, 0, 1] \in K^3$ und **ii)** $v := [1, 0, -1] \in K^3$ gebe man jeweils diejenigen $w \in K^3$ an, für die es $\varphi \in \mathrm{Hom}_K(\langle v, v_1, v_2\rangle_K, K^3)$ gibt mit $\varphi\colon v \mapsto w, v_1 \mapsto w_1, v_2 \mapsto w_2$. Wie hängen die Rechnungen von K ab?

(11.41) Aufgabe: Lineare Abbildungen und Basen.
Es seien K ein Körper, V ein K-Vektorraum mit K-Basis $B = \{v_1, \ldots, v_5\}$, und W ein K-Vektorraum mit K-Basis $C = \{w_1, \ldots, w_4\}$. Es sei φ diejenige K-lineare Abbildung von V nach W, für die gilt:

$$\begin{array}{lll} \varphi(v_1) = w_2 - w_1, & \varphi(v_2) = w_4 - w_2, & \varphi(v_3) = w_4 - w_1, \\ \varphi(v_4) = w_3 - w_1, & \varphi(v_5) = w_4 - w_3. & \end{array}$$

a) Man bestimme die Abbildungsmatrix ${}_C\varphi_B$ der Abbildung φ bezüglich der K-Basen B von V und C von W.
b) Man gebe K-Basen von $\mathrm{Kern}(\varphi)$ und $\mathrm{Bild}(\varphi)$ an.

(11.42) Aufgabe: Lineare Abbildungen.
Es sei K ein Körper. Welche der folgenden Abbildungen sind linear? Gegebenenfalls gebe man die Matrix bezüglich der jeweiligen K-Standardbasen an, bestimme Kern und Bild, und untersuche auf Injektivität und Surjektivität:
a) $\varphi\colon K^3 \to K^2\colon [a, b, c] \mapsto [a, c]$ **b)** $\varphi\colon K^3 \to K^3\colon [a, b, c] \mapsto [a, b+1, c]$
c) $\varphi\colon K^3 \to K^3\colon [a, b, c] \mapsto [c, a, b]$ **d)** $\varphi\colon K^2 \to K^2\colon [a, b] \mapsto [a+b, b]$
e) $\varphi\colon K^2 \to K\colon [a, b] \mapsto ab$ **f)** $\varphi\colon K^3 \to K^3\colon [a, b, c] \mapsto [a, b-c, 0]$

(11.43) Aufgabe: Invertierbare lineare Abbildungen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $\varphi\colon K^{2\times 1} \to K^{2\times 1}\colon [x, y]^{\mathrm{tr}} \mapsto [2x + 5y, x + 3y^{\mathrm{tr}}]$. Man zeige, dass φ K-linear und bijektiv ist, und gebe die Umkehrabbildung an.

(11.44) Aufgabe: Geometrische Interpretation linearer Abbildungen.
a) Man zeige, daß die Abbildungen

$$\begin{array}{lll} \varphi\colon & \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}\colon & [x_1, x_2]^{\mathrm{tr}} \mapsto [-\frac{1}{2}x_1 - \frac{\sqrt{3}}{2}x_2, \frac{\sqrt{3}}{2}x_1 - \frac{1}{2}x_2]^{\mathrm{tr}} \\ \psi\colon & \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}\colon & [x_1, x_2]^{\mathrm{tr}} \mapsto [\frac{5}{13}x_1 - \frac{12}{13}x_2, -\frac{12}{13}x_1 - \frac{5}{13}x_2]^{\mathrm{tr}} \end{array}$$

$\mathbb{R}$-linear sind. Man gebe die Abbildungsmatrizen von φ und ψ bezüglich der Standardbasis von $\mathbb{R}^{2\times 1}$ an.

b) Man betrachte das Dreieck $\Delta \subseteq \mathbb{R}^{2\times 1}$ mit den Ecken $u := [1,1]^{\mathrm{tr}}$, $v := [\frac{1}{2}, 2]^{\mathrm{tr}}$ und $w := [3, \frac{3}{2}]^{\mathrm{tr}}$. Man berechne die Bilder dieser Punkte unter den Abbildungen φ und ψ, und trage die Ergebnisse in ein kartesisches Koordinatensystem ein. Welche geometrische Beschreibung haben die Abbildungen φ und ψ?

(11.45) Aufgabe: Abbildungsmatrizen.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $\varphi_A \colon K^{2\times 1} \to K^{3\times 1} \colon v \mapsto A \cdot v$ die K-lineare Abbildung mit

$$A := \begin{bmatrix} 4 & -6 \\ 3 & -2 \\ -1 & 5 \end{bmatrix} \in K^{3\times 2}.$$

a) Es seien S und T die Standardbasen von $K^{2\times 1}$ bzw. $K^{3\times 1}$. Man bestimme die Abbildungsmatrix ${}_T(\varphi_A)_S$.
b) Man zeige, daß $B := \{[3,2]^{\mathrm{tr}}, [7,5]^{\mathrm{tr}}\}$ eine K-Basis von $K^{2\times 1}$ ist, und daß $C := \{[-1,2,3]^{\mathrm{tr}}, [0,1,2]^{\mathrm{tr}}, [0,0,-1]^{\mathrm{tr}}\}$ eine K-Basis von $K^{3\times 1}$ ist.
c) Für einen beliebigen Vektor $v := [x,y]^{\mathrm{tr}} \in K^{2\times 1}$ bestimme man den Koordinatenvektor ${}_B v \in K^{2\times 1}$, und für einen beliebigen Vektor $w := [x,y,z]^{\mathrm{tr}} \in K^{3\times 1}$ bestimme man den Koordinatenvektor ${}_C w \in K^{3\times 1}$, und
d) Man bestimme die Abbildungsmatrix ${}_C(\varphi_A)_B \in K^{3\times 2}$. Wie hängen die Rechnungen von K ab?

(11.46) Aufgabe: Basiswechsel.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $\varphi \in \mathrm{End}_K(K^{4\times 1})$ bezüglich der K-Standardbasis B von $K^{4\times 1}$ gegeben durch

$${}_B\varphi_B = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & 1 & 1 & 1 \\ 3 & 2 & 0 & 1 \\ 4 & 1 & 1 & 1 \end{bmatrix} \in K^{4\times 4}.$$

a) Es sei $C := \{[1,0,0,2]^{\mathrm{tr}}, [2,1,1,0]^{\mathrm{tr}}, [1,2,0,4]^{\mathrm{tr}}, [3,2,1,5]^{\mathrm{tr}}\} \subseteq K^{4\times 1}$. Man zeige: C ist eine K-Basis von $K^{4\times 1}$. Man berechne die Matrizen ${}_B\varphi_C$, ${}_C\varphi_B$ und ${}_C\varphi_C$. Wie hängen die Rechnungen von K ab?
b) Es sei $D := \{[1,-1,0,0]^{\mathrm{tr}}, [0,1,1,0]^{\mathrm{tr}}, [0,0,1,1]^{\mathrm{tr}}, [1,0,0,1]^{\mathrm{tr}}\} \subseteq K^{4\times 1}$. Man zeige: D ist eine K-Basis von $K^{4\times 1}$. Man bestimme alle Basiswechselmatrizen zwischen B, C und D, und berechne damit erneut ${}_C\varphi_C$ sowie ${}_D\varphi_D$.

(11.47) Aufgabe: Rang linearer Abbildungen.
Es seien K ein Körper sowie V, W und U endlich erzeugte K-Vektorräume.
a) Es seien $\varphi, \psi \in \mathrm{Hom}_K(V,W)$. Man zeige: Es gilt $|\mathrm{Rang}(\varphi) - \mathrm{Rang}(\psi)| \leq \mathrm{Rang}(\varphi + \psi) \leq \mathrm{Rang}(\varphi) + \mathrm{Rang}(\psi)$.
b) Es seien $\varphi \in \mathrm{Hom}_K(V,W)$ und $\psi \in \mathrm{Hom}_K(W,U)$. Man zeige: Es gilt $\mathrm{Rang}(\varphi) + \mathrm{Rang}(\psi) - \dim_K(W) \leq \mathrm{Rang}(\psi\varphi) \leq \min\{\mathrm{Rang}(\psi), \mathrm{Rang}(\varphi)\}$.

(11.48) Aufgabe: Räume von Endomorphismen.
Es seien K ein Körper, V ein endlich erzeugter K-Vektorraum, und für $U \leq V$ seien $\mathcal{Q} := \{\varphi \in \mathrm{End}_K(V); \mathrm{Bild}(\varphi) \leq U\}$ und $\mathcal{R} := \{\varphi \in \mathrm{End}_K(V); \varphi|_U = 0\}$. Man zeige: Es sind $\mathcal{Q}, \mathcal{R} \leq \mathrm{End}_K(V)$. Man bestimme $\dim_K(\mathcal{Q})$, $\dim_K(\mathcal{R})$, $\dim_K(\mathcal{Q} \cap \mathcal{R})$ und $\dim_K(\mathcal{Q} + \mathcal{R})$, als Ausdrücke in $\dim_K(V)$ und $\dim_K(U)$.

(11.49) Aufgabe: Endomorphismen.
Es seien K ein Körper und V ein endlich erzeugter K-Vektorraum.
a) Für $\varphi \in \mathrm{End}_K(V)$ zeige man die Äquivalenz der folgenden Aussagen:
i) Es ist $\mathrm{Kern}(\varphi) \cap \mathrm{Bild}(\varphi) = \{0\}$. **ii)** Es ist $\mathrm{Kern}(\varphi) + \mathrm{Bild}(\varphi) = V$.
c) Man zeige: Ist $\varphi \in \mathrm{End}_K(V)$ mit $\varphi\varphi = \varphi$, so gilt $V = \mathrm{Kern}(\varphi) \oplus \mathrm{Bild}(\varphi)$.
d) Man gebe jeweils ein $\varphi \in \mathrm{End}_K(V)$ mit den folgenden Eigenschaften an:
i) Es ist $\{0\} \neq \mathrm{Kern}(\varphi) \leq \mathrm{Bild}(\varphi) \neq V$. **ii)** Es ist $\mathrm{Bild}(\varphi) \leq \mathrm{Kern}(\varphi) \neq V$.

(11.50) Aufgabe: Skalare Abbildungen.
Es seien K ein Körper, V ein K-Vektorraum und $\varphi \in \mathrm{End}_K(V)$ mit $\varphi(v) \in \langle v \rangle_K$ für alle $v \in V$. Man zeige: Es gibt ein $a \in K$ mit $\varphi(v) = av$ für alle $v \in V$.

(11.51) Aufgabe: Faktorräume.
a) Es seien K ein Körper, V ein K-Vektorraum mit endlicher K-Basis B und $U \leq V$. Man zeige: Es gibt $C \subseteq B$, so daß $\{v + U \in V/U; v \in C\}$ eine K-Basis des Quotienten V/U ist. Gilt eine analoge Aussage auch für den Teilraum U?
b) Es seien $V := K^n$, für ein $n \in \mathbb{N}$, sowie $U := \{[a, \ldots, a] \in V; a \in K\}$ und $W := \{[a_1, \ldots, a_n] \in V; \sum_{i=1}^n a_i = 0\}$. Man gebe Teilmengen B und C der K-Standardbasis von V an, so daß $\{v + U \in V/U; v \in B\}$ und $\{v + W \in V/W; v \in C\}$ jeweils K-Basen von V/U und V/W sind.

(11.52) Aufgabe: Linearformen.
Es seien K ein Körper und V ein K-Vektorraum. Man zeige:
a) Es gilt $\{U \leq V; \dim_K(V/U) = 1\} = \{\mathrm{Kern}(\tau) \leq V; 0 \neq \tau \in \mathrm{Hom}_K(V, K)\}$.
b) Ist $U \leq V$ mit $\dim_K(V/U) = n \in \mathbb{N}$, so gibt es eine K-linear unabhängige Folge $[\tau_i \in \mathrm{Hom}_K(V, K); i \in \{1, \ldots, n\}]$ mit $U = \bigcap_{i=1}^n \mathrm{Kern}(\tau_i)$.

(11.53) Aufgabe: Natürliche Abbildungen.
Es seien K ein Körper, V ein K-Vektorraum und $U, W \leq V$. Man zeige:
a) Es ist $V/U \times V/W$ ein K-Vektorraum mit komponentenweiser Addition und Skalarmultiplikation.
b) Die Abbildung $\nu\colon V \to V/U \times V/W : v \mapsto [\nu_U(v), \nu_W(v)]$ ist K-linear. Wann ist ν ist injektiv, wann ist ν surjektiv?

(11.54) Aufgabe: Teilräume und Quotienten.
Es seien K ein Körper, V ein K-Vektorraum, $U \leq V$ und $\varphi \in \mathrm{End}_K(V)$ mit $\varphi(U) \leq U$. Man zeige:
a) Es ist $\varphi|_U \in \mathrm{End}_K(V)$, und es gibt eine eindeutig bestimmte Abbildung $\widetilde{\varphi} \in \mathrm{End}_K(V/U)$ mit $\nu_U\varphi = \widetilde{\varphi}\nu_U \in \mathrm{Hom}(V, V/U)$.

b) Sind $\varphi|_U$ und $\widetilde{\varphi}$ injektiv, so ist auch φ injektiv; und sind $\varphi|_U$ und $\widetilde{\varphi}$ surjektiv, so ist auch φ surjektiv. Gilt jeweils auch die Umkehrung?
c) Ist V endlich erzeugt, so ist φ genau dann bijektiv, wenn $\varphi|_U$ und $\widetilde{\varphi}$ es sind.
d) Sind V endlich erzeugt, $\varphi|_U = 0$ und $\widetilde{\varphi} = 0$, so ist $\mathrm{Rang}(\varphi) \leq \frac{\dim_K(V)}{2}$.

(11.55) Aufgabe: Matrixkommutatoren.
Es seien K ein Körper, $V := K^{2\times 2}$ und für $A := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in V$ sei $\psi_A \colon V \to V \colon X \mapsto AX - XA$ die zugehörige **Kommutatorabbildung**. Man zeige: Die Abbildung ψ_A ist K-linear. Man gebe die Abbildungsmatrix ${}_B(\psi_A)_B$ bezüglich der K-Standardbasis B von V, sowie K-Basen für $\mathrm{Kern}(\psi_A)$ und $\mathrm{Bild}(\psi_A)$ an. Wann gilt $V = \mathrm{Kern}(\psi_A) \oplus \mathrm{Bild}(\psi_A)$?

(11.56) Aufgabe: Matrixspur.
Es seien K ein Körper und $n \in \mathbb{N}$. Man zeige:
a) Die **Spurabbildung** $\tau \colon K^{n\times n} \to K \colon [a_{ij}]_{ij} \mapsto \sum_{i=1}^n a_{ii}$ ist K-linear. Man bestimme $\dim_K(\mathrm{Kern}(\tau))$. Man zeige: Für alle $A, B \in K^{n\times n}$ gilt $\tau(AB) = \tau(BA)$. Gilt auch $\tau(AB) = \tau(A) \cdot \tau(B)$?
b) Ist $n \neq 0 \in K$, so gibt es keine Matrizen $A, B \in K^{n\times n}$ mit $AB - BA = E_n$.

(11.57) Aufgabe: Matrixgruppen.
Es sei K ein Körper. Man zeige: Die folgenden Mengen sind Untergruppen von $\mathrm{GL}_n(K)$, wobei $n \in \mathbb{N}$, und jedes Element $b \in B$ kann in eindeutiger Weise als Produkt $b = tu$ für geeignete $t \in T$ und $u \in U$ geschrieben werden; die Elemente von B heißen **(untere) Dreiecksmatrizen**:

$$\begin{array}{rcl} B & := & \{[a_{ij}]_{ij} \in K^{n\times n}; a_{ij} = 0 \text{ für alle } i < j\} \\ U & := & \{[a_{ij}]_{ij} \in K^{n\times n}; a_{ij} = 0, a_{ii} = 1 \text{ für alle } i < j\} \\ T & := & \{[a_{ij}]_{ij} \in K^{n\times n}; a_{ij} = 0, a_{ii} \neq 0 \text{ für alle } i \neq j\} \end{array}$$

(11.58) Aufgabe: Invertierbare Matrizen.
Es seien K ein Körper und $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in K^{2\times 2}$. Man gebe ein notwendiges und hinreichendes Kriterium an die Matrixeinträge $a, b, c, d \in K$ für die Invertierbarkeit von A an, und bestimme in diesem Fall die inverse Matrix.

(11.59) Aufgabe: Matrixprodukt.
Es seien die folgenden Matrizen über dem Körper K gegeben:

$$A_1 := \begin{bmatrix} -3 & -2 \\ 6 & 4 \end{bmatrix}, \quad A_2 := \begin{bmatrix} 4 & -2 & 1 \\ 2 & -1 & 1 \end{bmatrix},$$

$$A_3 := \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad A_4 := \begin{bmatrix} -6 & 2 & -1 \\ 2 & -3 & 3 \\ 20 & -10 & 8 \end{bmatrix}.$$

Man berechne die Produkte $A_i \cdot A_j$ für $i, j \in \{1, \ldots, 4\}$, soweit definiert.

(11.60) Aufgabe: Matrixmultiplikation.

a) Man zeige, daß die Menge $\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in \mathbb{R}^{2\times 2}; a, b \in \mathbb{R}\right\}$ zusammen mit Matrixaddition und -multiplikation ein Körper ist. Was hat dieser mit $\mathbb{C}$ zu tun?
b) Man zeige, daß es keine Matrizen $A, B \in \mathbb{R}^{2\times 2}$ mit $A^2 = B^2 = -E_2$ und $AB = -BA$ gibt, daß man solche Matrizen aber in $\mathbb{C}^{2\times 2}$ finden kann.
c) Man zeige, daß die Menge $\left\{\begin{bmatrix} a & \overline{b} \\ -b & \overline{a} \end{bmatrix} \in \mathbb{C}^{2\times 2}; a, b \in \mathbb{C}\right\}$ zusammen mit Matrixaddition und -multiplikation ein nichtkommutativer Schiefkörper ist. Was hat dieser mit den Hamilton-Quaternionen $\mathbb{H}$ zu tun?

(11.61) Aufgabe: Äquivalenz von Matrizen.
Es seien K ein Körper sowie $m, n \in \mathbb{N}_0$. Matrizen $A, B \in K^{m\times n}$ heißen **äquivalent**, falls es $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ gibt mit $B = PAQ$. Man zeige, daß dies eine Äquivalenzrelation auf $K^{m\times n}$ definiert, und bestimme die zugehörigen Äquivalenzklassen.

(11.62) Aufgabe: Matrixrang.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p\}$, für eine Primzahl $p \in \mathbb{N}$. Man bestimme den Rang der folgenden Matrix, in Abhängigkeit von K:

$$\begin{bmatrix} 4 & 5 & 2 & 2 \\ 2 & 1 & -2 & -2 \\ -3 & 0 & 4 & 3 \\ 5 & 5 & 0 & 1 \end{bmatrix} \in K^{4\times 4}$$

(11.63) Aufgabe: Matrixrang mit Parameter.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man bestimme den Rang der folgenden Matrizen, wobei $a \in K$. Wie hängen die Rechnungen von K ab?

a) $\begin{bmatrix} 1 & 2 \\ 3 & a \end{bmatrix} \in K^{2\times 2}$ **b)** $\begin{bmatrix} 1 & a & 1 \\ 0 & 1 & a \\ a & 1 & 0 \end{bmatrix} \in K^{3\times 3}$ **c)** $\begin{bmatrix} 1 & 2 & 3 & a \\ 2 & 3 & a & 1 \\ 3 & a & 1 & 2 \\ a & 1 & 2 & 3 \end{bmatrix} \in K^{4\times 4}$

(11.64) Aufgabe: Matrixrang mit zwei Parametern.
Es sei K ein Körper. Man bestimme den Rang der folgenden Matrizen, wobei $a, b \in K$. Wie hängen die Rechnungen von K ab?
a) $[a_{ij}]_{ij} \in K^{n\times n}$, wobei $n \in \mathbb{N}$, mit $a_{ij} := a$ für $j \geq i$ und $a_{ij} := b$ für $j < i$.

b) $\begin{bmatrix} 0 & a & a \\ b & 0 & a \\ b & b & 0 \end{bmatrix} \in K^{3\times 3}$ **c)** $\begin{bmatrix} 0 & a & a & a \\ b & 0 & a & a \\ b & b & 0 & a \\ b & b & b & 0 \end{bmatrix} \in K^{4\times 4}$

(11.65) Aufgabe: Lineare Gleichungssysteme.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und

$$A := \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 2 & 3 & 4 & 5 & 6 \\ 0 & 2 & 2 & 4 & 4 \end{bmatrix} \in K^{4\times 5}.$$

Man bestimme die Lösungen des Systems $A \cdot [x_1, \ldots, x_5]^{\mathrm{tr}} = w \in K^{4\times 1}$ für
a) $w := [0,0,0,0]^{\mathrm{tr}}$, **b)** $w := [1,1,1,1]^{\mathrm{tr}}$, **c)** $w := [1,1,1,-1]^{\mathrm{tr}}$.

(11.66) Aufgabe: Generische lineare Gleichungssysteme.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und

$$A := \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix} \in K^{3\times 3}.$$

Für welche $[a,b,c]^{\mathrm{tr}} \in K^{3\times 1}$ ist das lineare Gleichungssystem $A \cdot [x_1, x_2, x_3]^{\mathrm{tr}} = [a,b,c]^{\mathrm{tr}} \in K^{3\times 1}$ lösbar? Gegebenenfalls bestimme man die Lösungsmenge. Wie hängen die Rechnungen von K ab?

(11.67) Aufgabe: Lineare Gleichungssysteme mit Parameter.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Für welche Parameter λ ist das lineare Gleichungssystem mit Koeffizientenmatrix $A \in K^{n\times n}$ wie folgt für jede rechte Seite $w = [y_1, \ldots, y_n]^{\mathrm{tr}} \in K^{n\times 1}$ lösbar? In diesen Fällen gebe man die (eindeutige) Lösung in Abhängigkeit von $y_1, \ldots, y_n \in K$ an.

In den übrigen Fällen gebe man eine K-Basis des Lösungsraums $\mathcal{L}(A,0)$ des zugehörigen homogenen Systems an, sowie eine K-Basis des K-Teilraums der rechten Seiten $w \in K^{n\times 1}$ mit $\mathcal{L}(A,w) \neq \emptyset$. Hängen die Rechnungen von K ab?

a) $A := \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & \lambda \\ 1 & \lambda & 3 \end{bmatrix}$ **b)** $A := \begin{bmatrix} 2 & -\lambda & 1 \\ 3 & 4 & 2 \\ \lambda & 4-\lambda & 3 \end{bmatrix}$ **c)** $A := \begin{bmatrix} 1 & 1 & -\lambda & 1 \\ 1 & -\lambda & 1 & 6 \\ 2 & 0 & -1 & \lambda \\ 0 & 1 & \lambda & -1 \end{bmatrix}$

(11.68) Aufgabe: Gleichungssysteme über $\mathbb{C}$.
Man bestimme die Lösungsmenge $\mathcal{L}(A,w)$ der folgenden linearen Gleichungssysteme über $\mathbb{C}$, und gebe zusätzlich eine $\mathbb{C}$-Basis des Lösungsraums $\mathcal{L}(A,0)$ des zugehörigen homogenen Systems an:

a) $A := \begin{bmatrix} 1+i & 2+i \\ 3+i & 4+i \end{bmatrix}$, $\quad w := \begin{bmatrix} 5+2i \\ 11-6i \end{bmatrix}$

b) $A := \begin{bmatrix} 1+i & i & -2 \\ -i & 2+i & 1 \\ 2-3i & 1-i & 3+i \end{bmatrix}$, $\quad w := \begin{bmatrix} 3-3i \\ -4+13i \\ -5-7i \end{bmatrix}$

c) $A := \begin{bmatrix} i-1 & 0 & 2-2i & 4i-3 \\ 1-3i & -i & 6i-2 & 2-12i \\ 2 & i & -4 & 7 \\ i-1 & i & 3-3i & 6i-5 \end{bmatrix}$, $\quad w := \begin{bmatrix} 4-5i \\ 16i-3 \\ -9-i \\ 7-9i \end{bmatrix}$

(11.69) Aufgabe: Gleichungssysteme über endlichen Körpern.
Man bestimme die Lösungsmenge der folgenden linearen Gleichungssysteme; dabei gebe man eine Parameterdarstellung als auch die Lösungen explizit an:

a) $A := \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 1 & 2 \end{bmatrix} \in \mathbb{Z}_3^{3\times 4}$, $\quad w := \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \in \mathbb{Z}_3^{3\times 1}$

b) $A := \begin{bmatrix} 2 & 1 & 4 & 1 & 0 & 1 \\ 4 & 4 & 2 & 3 & 4 & 4 \\ 1 & 2 & 0 & 3 & 0 & 1 \\ 4 & 2 & 3 & 3 & 4 & 2 \\ 4 & 2 & 3 & 0 & 2 & 4 \end{bmatrix} \in \mathbb{Z}_5^{5\times 6}$, $\quad w := \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 2 \end{bmatrix} \in \mathbb{Z}_5^{5\times 1}$

(11.70) Aufgabe: Geraden in der reellen Ebene.
Man stelle die folgenden Geraden $g_1, \ldots, g_4 \subseteq \mathbb{R}^{2\times 1}$ durch eine Parametrisierung bzw. Gleichung, sowie in einem kartesischen Koordinatensystem dar:
i) $g_1 := \{[4,-3]^{\text{tr}} + t \cdot [0,2]^{\text{tr}} \in \mathbb{R}^{2\times 1}; t \in \mathbb{R}\}$
ii) $g_2 := \{[0,-3]^{\text{tr}} + t \cdot [-3,4]^{\text{tr}} \in \mathbb{R}^{2\times 1}; t \in \mathbb{R}\}$
iii) $g_3 := \{[x_1,x_2]^{\text{tr}} \in \mathbb{R}^{2\times 1}; 4x_1 - 10x_2 = 0\}$
iv) $g_4 := \{[x_1,x_2]^{\text{tr}} \in \mathbb{R}^{2\times 1}; 2x_1 - 11x_2 = -43\}$

(11.71) Aufgabe: Geraden und ihre Durchschnitte.
Man stelle die folgenden Geraden $g_1, g_2, g_3 \subseteq \mathbb{R}^{2\times 1}$ in einem kartesischen Koordinatensystem dar, und bestimme ihre paarweisen Durchschnitte:
i) $g_1 := \{[x_1,x_2] \in \mathbb{R}^{2\times 1}; 2x_1 + \frac{5}{3}x_2 = -12\}$
ii) $g_2 := \{[-6,0]^{\text{tr}} + t \cdot [4,5]^{\text{tr}} \in \mathbb{R}^{2\times 1}; t \in \mathbb{R}\}$
iii) g_3 ist die Gerade durch die Punkte $[15,0]^{\text{tr}}$ und $[0,-\frac{45}{8}]^{\text{tr}}$.

(11.72) Aufgabe: Geraden über endlichen Körpern.
Man bestimme Parameterdarstellungen der folgenden Geraden $g, h \subseteq \mathbb{Z}_5^{2\times 1}$, gebe jeweils die darauf liegenden Punkte an, und berechne ihren Durchschnitt:

$$\begin{array}{rcl} g & := & \{[x,y] \in \mathbb{Z}_5^{2\times 1}; 3x + 2y = 0\} \\ h & := & \{[x,y] \in \mathbb{Z}_5^{2\times 1}; 4x + 3y = 1\} \end{array}$$

(11.73) Aufgabe: Zentralisatoren.
Es sei K ein Körper, und für $S \subseteq K^{n \times n}$, wobei $n \in \mathbb{N}$, sei $C(S) := \{A \in K^{n \times n}; AM = MA \text{ für alle } M \in S\}$ der **Zentralisator** von S.

a) Man zeige: C ist sowohl ein K-Teilraum als auch ein Teilring von $K^{n \times n}$. Man bestimme $C(K^{n \times n})$ und zeige $\dim_K(C(K^{n \times n})) = 1$.

b) Es sei $A := \operatorname{diag}[a_1, \ldots, a_n] \in K^{n \times n}$ mit paarweise verschiedenen Einträgen a_i. Man bestimme $C(A)$ und zeige $\dim_K(C(A)) = n$.

c) Es sei $B := [b_{ij}]_{ij} \in K^{n \times n}$ mit $b_{ij} := 1$ für alle i, j. Man bestimme $C(B)$ und zeige $\dim_K(C(B)) = n^2 - 2n + 2$.

d) Es sei $C := [c_{ij}]_{ij} \in K^{n \times n}$ mit $c_{ii} := 1$ und $c_{i,i+1} := 1$ sowie $c_{ij} := 0$ für alle $j \neq i, i+1$. Man bestimme $C(C)$ und zeige $\dim_K(C(C)) = n$.

e) Man bestimme $C(\mathcal{B})$ und $C(\mathcal{U})$ für die Mengen $\mathcal{B}$ und $\mathcal{U}$ aus Aufgabe (11.57).

(11.74) Aufgabe: Zassenhaus-Algorithmus.

a) Es seien K ein Körper und $U, V \leq K^n$, wobei $n \in \mathbb{N}$. Wie kann man mittels des Gauß-Algorithmus aus K-Basen von U und V gleichzeitig K-Basen von $U + V \leq K^n$ und $U \cap V \leq K^n$ berechnen?

b) Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ sowie $U := \langle [3, -7, 2, -4], [2, -1, 5, 1] \rangle_K \leq K^4$ und $V := \langle [5, 7, 1, 6], [1, 7, -4, 4] \rangle_K \leq K^4$. Man berechne K-Basen von $U + V \leq K^4$ und $U \cap V \leq K^4$. Wie hängen die Rechnungen von K ab?

Hinweis zu a). Sind $B \in K^{m \times n}$ und $C \in K^{l \times n}$ Matrizen mit Zeilenraum U bzw. V, so betrachte man $\left[\begin{array}{c|c} B & 0 \\ \hline C & C \end{array}\right] \in K^{(m+l) \times 2n}$.

(11.75) Aufgabe: Taylor-Entwicklung.
Für $n \in \mathbb{N}_0$ sei $\mathcal{P}_n(\mathbb{R}) := \langle P_n \rangle_{\mathbb{R}} \leq C^{\mathrm{pol}}(\mathbb{R})$ der $\mathbb{R}$-Vektorraum der **Polynomfunktionen vom Grad** $\leq n$, wobei $P_n := [p_0, \ldots, p_n]$ und $p_n \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto x^n$.

a) Man zeige: Durch $\tau \colon f \mapsto (x \mapsto f(x+1))$ wird ein $\mathbb{R}$-Automorphismus von $\mathcal{P}_n(\mathbb{R})$ definiert. Man gebe ${}_{P_n}(\tau)_{P_n}$ und ${}_{P_n}(\tau^{-1})_{P_n}$ an, und zeige

$$\sum_{j=0}^{i} (-1)^{j-i} \binom{i}{j} \binom{j}{k} = \begin{cases} 1, & \text{für } i = k, \\ 0, & \text{für } i \neq k. \end{cases}$$

b) Man zeige: In $\operatorname{End}_{\mathbb{R}}(\mathcal{P}_n(\mathbb{R}))$ gelten $\frac{\partial}{\partial x} \cdot \tau = \tau \cdot \frac{\partial}{\partial x}$ und die **Taylor-Formel** $\tau = \sum_{i=0}^{n} \frac{1}{i!} (\frac{\partial}{\partial x})^i$ sowie $(\tau - \mathrm{id})^{n+1} = 0$ und $\frac{\partial}{\partial x} = \sum_{i=1}^{n} \frac{1}{i!} (\tau - \mathrm{id})^i$.

(11.76) Aufgabe: Adjazente Transpositionen.
Man zeige: Für $n \geq 2$ kann $(1, n) \in \mathcal{S}_n$ als Produkt von $2n - 3$ adjazenten Transpositionen geschrieben werden. Gibt es eine kürzere solche Darstellung?

(11.77) Aufgabe: Permutationen.
Man schreibe die folgenden Permutationen $\pi \in \mathcal{S}_9$ und ihre Inversen $\pi^{-1} \in \mathcal{S}_9$ als Produkte disjunkter Zykel sowie als Produkte von Transpositionen und bestimme jeweils $\operatorname{sgn}(\pi)$:

a) $\pi = (1,2,3)(2,3,4)(4,5)(5,6,7,8,9)$ **b)** $\pi = (1,4,7)^{-1}(3,9)(4,5,6,2)$
c) $\pi = (1,9,2,8)(3,4,5)(5,4)(6,1)(7,8,9)$ **d)** $\pi = (3,7,4,6,5)^{-1}(1,2)(1,3)$

(11.78) Aufgabe: Kästchen-Matrizen.
Es seien $R \neq \{0\}$ ein kommutativer Ring und $m, n \in \mathbb{N}_0$. Weiter seien $B \in R^{m\times m}$ und $C \in R^{m\times n}$ sowie $D \in R^{n\times n}$ und $A := \left[\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right] \in R^{(m+n)\times(m+n)}$. Man zeige: Es gilt $\det(A) = \det(B) \cdot \det(D) \in R$.

(11.79) Aufgabe: Determinanten.
a) Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man berechne die Determinante der folgenden Matrizen $A, B \in K^{4\times 4}$ jeweils auf zwei verschiedene Weisen: indem sie auf Dreiecksgestalt gebracht werden, und indem mindestens einmal der Entwicklungssatz benutzt wird. Wie hängen die Rechnungen von K ab?

$$A := \begin{bmatrix} 1 & -5 & 9 & -13 \\ -2 & 6 & -10 & 12 \\ 4 & -8 & 14 & -16 \\ -3 & 7 & -11 & 15 \end{bmatrix} \qquad B := \begin{bmatrix} 14 & -2 & 4 & 15 \\ 5 & 6 & 10 & -9 \\ -1 & 0 & 1 & -3 \\ 12 & 0 & 8 & 6 \end{bmatrix}$$

b) Man berechne die Determinanten der Produkte $A \cdot B$ und $B \cdot A$.

(11.80) Aufgabe: Determinanten über verschiedenen Körpern.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p\}$, für eine Primzahl $p \in \mathbb{N}$, und $a_1, \ldots, a_4 \in \mathbb{R}$. Man berechne die Determinante der folgenden Matrizen, bestimme ihren Rang und berechne gegebenenfalls die jeweilige inverse Matrix. Wie hängen die Rechnungen in a) von K ab? Was hat b) mit den Hamilton-Quaternionen $\mathbb{H}$ zu tun?

$$\textbf{a)} \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 \\ 4 & 5 & 2 & 3 \\ 5 & 2 & 3 & 4 \end{bmatrix} \in K^{4\times 4} \qquad \textbf{b)} \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

(11.81) Aufgabe: Determinanten von Matrixserien.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}$. Man berechne die Determinante der folgenden Matrizen $[a_{ij}]_{ij} \in K^{n\times n}$:
a) $a_{ii} := 2$, und $a_{ij} := 1$ für $i \neq j$.
b) $a_{ii} := 2$, und $a_{ij} := -1$ für $|i-j| = 1$, sowie $a_{ij} := 0$ für $|i-j| \geq 2$.
c) $a_{ii} := 0$, und $a_{ij} = -a_{ji} := 1$ für $i < j$.
d) $a_{ii} := c_i$, und $a_{ij} := a$ für $i < j$, und $a_{ij} := b$ für $i > j$, für $a, b, c_1, \ldots, c_n \in K$.
e) $a_{ii} := a$, und $a_{i,n+1-i} := b$ für $i \neq j$, und $a_{ij} := 0$ sonst, für $a, b \in K$.
f) $a_{ij} := \max\{i, j\}$.

Hinweis zu d). Man betrachte die Funktion $K \to K \colon z \mapsto \det([a_{ij} + z])$ und das Polynom $\prod_{i=1}^{n}(X - c_i) \in K[X]$, und behandle den Fall $a = b$ gesondert.

(11.82) Aufgabe: Cauchy-Determinante.
Es sei K ein Körper, und für $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$, wobei $n \in \mathbb{N}$, sei $\Delta(a_1, \ldots, a_n; b_1, \ldots, b_n) := \prod_{1 \leq i < j \leq n} (a_j - a_i)(b_j - b_i) \in K$.
a) Es sei $A := [\frac{1}{1-a_i b_j}]_{ij} \in K^{n \times n}$, wobei $\prod_{i=1}^n \prod_{j=1}^n (1 - a_i b_j) \neq 0$. Man zeige: Es gilt $\det(A) \cdot \prod_{i=1}^n \prod_{j=1}^n (1 - a_i b_j) = \Delta(a_1, \ldots, a_n; b_1, \ldots, b_n)$.
b) Es sei $A := [\frac{1}{a_i + b_j}]_{ij} \in K^{n \times n}$, wobei $\prod_{i=1}^n \prod_{j=1}^n (a_i + b_j) \neq 0$. Man zeige: Es gilt $\det(A) \cdot \prod_{i=1}^n \prod_{j=1}^n (a_i + b_j) = \Delta(a_1, \ldots, a_n; b_1, \ldots, b_n)$.

Hinweis. Man subtrahiere die erste Spalte von den anderen Spalten, die erste Zeile von den anderen Zeilen, und wende Induktion nach n an.

(11.83) Aufgabe: Determinanten von Endomorphismen.
Es seien K ein Körper und $V := K^{n \times n}$, wobei $n \in \mathbb{N}_0$. Für $A \in V$ zeige man:
a) Die Abbildungen $\rho_A \colon V \to V \colon B \mapsto BA$ und $\lambda_A \colon V \to V \colon B \mapsto AB$ sind K-linear, und es gilt $\det(\rho_A) = \det(\lambda_A) = \det(A)^n \in K$.
b) Man berechne $\det(\rho_A - \lambda_A) \in K$.

(11.84) Aufgabe: Matrixinversion über $\mathbb{Z}$.
Es sei $A \in \mathbb{Z}^{n \times n}$, wobei $n \in \mathbb{N}$. Man zeige:
a) Es ist genau dann $A \in \mathrm{GL}_n(\mathbb{Q})$, wenn $A \in \mathrm{GL}_n(\mathbb{C})$ ist.
b) Es ist genau dann $A \in \mathrm{GL}_n(\mathbb{Z})$, wenn $A \in \mathrm{GL}_n(\mathbb{C})$ und $\det(A) \in \{\pm 1\}$ ist.

(11.85) Aufgabe: Adjungierte Matrizen.
Es seien $R \neq \{0\}$ ein kommutativer Ring und $A, B \in R^{n \times n}$, wobei $n \in \mathbb{N}$.
a) Man zeige: Es gelten $\mathrm{adj}(AB) = \mathrm{adj}(B)\mathrm{adj}(A)$ und $\det(\mathrm{adj}(A)) = \det(A)^{n-1}$.
b) Man zeige: Für $A \in \mathrm{GL}_n(R)$ ist $\mathrm{adj}(A) \in \mathrm{GL}_n(R)$ mit $\mathrm{adj}(A^{-1}) = \mathrm{adj}(A)^{-1}$.
c) Für $n \geq 2$ zeige man: Es gilt $\mathrm{adj}(\mathrm{adj}(A)) = \det(A)^{n-2} \cdot A$.

(11.86) Aufgabe: Cramersche Regel.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, und für $\lambda \in K$ sei

$$A := \begin{bmatrix} 7 & \lambda & -3 \\ \lambda & 5 & 2 \\ -3 & 2 & -1 \end{bmatrix} \in K^{3 \times 3}.$$

a) Man berechne die adjungierte Matrix $\mathrm{adj}(A)$ und die Determinante $\det(A)$. Für welche Werte von $\lambda \in K$ ist A invertierbar?
b) In diesem Fall gebe man die inverse Matrix A^{-1} an, und bestimme die (eindeutige) Lösung des linearen Gleichungssystems mit Koeffizientenmatrix A und rechter Seite $w := [3, 3, 3]^{\mathrm{tr}} \in K^{3 \times 1}$. Hängen die Rechnungen von K ab?

(11.87) Aufgabe: Basisergänzung und Determinanten.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, sowie $v := [7, 2, 21, 4] \in K^{4 \times 1}$ und $w := [-3, 5, -9, 10] \in K^{4 \times 1}$. Man bestimme alle Paare $\{e_i, e_j\}$, für $1 \leq i < j \leq 4$ von Einheitsvektoren, die $\{v, w\}$ zu einer K-Basis von $K^{4 \times 1}$ ergnzen, durch Berechnung der

Determinante der Matrix $[v, w, e_i, e_j] \in K^{4\times 4}$ mittels zweimaliger Entwicklung nach der letzten Spalte. Wie hängen die Rechnungen von K ab?

12 Exercises for Part II (in German)

(12.1) Aufgabe: Matrixringe.
Es seien K ein Körper und $R := K^{n\times n}$, wobei $n \in \mathbb{N}$.
a) Man zeige: R hat nur die trivialen Ideale, ist aber für $n \geq 2$ kein Schiefkörper.
b) Ein $0 \neq a \in R$ heißt **Nullteiler**, falls es $0 \neq b, c \in R$ gibt mit $ba = ac = 0$. Man zeige: Jedes Element von $R \setminus \{0\}$ ist entweder Einheit oder Nullteiler.

(12.2) Aufgabe: Homomorphiesatz.
Es seien R ein Ring und $I \trianglelefteq R$ ein Ideal. Man zeige:
a) Ist $S \subseteq R$ ein Teilring, so ist $S \cap I \trianglelefteq S$ ein Ideal, $S + I \subseteq R/I$ ist ein Teilring, und es gilt $S/(S \cap I) \cong (S + I)/I$.
b) Ist $J \trianglelefteq R$ ein Ideal mit $I \subseteq J$, so ist $J/I \trianglelefteq R/I$, und es gilt $(R/I)/(J/I) \cong R/J$.
c) Die Menge der Ideale von R/I steht in natürlicher Bijektion zu $\{J \trianglelefteq R; I \subseteq J\}$.

(12.3) Aufgabe: Maximale Ideale.
Es sei R ein Ring. Der Ring $R \neq \{0\}$ heißt **einfach**, wenn R nur die Ideale $\{0\}$ und R besitzt. Ein echtes Ideal $I \triangleleft R$ heißt **maximal**, wenn es kein echtes Ideal $J \triangleleft R$ mit $I \subset J$ gibt. Man zeige:
a) Ein Ideal $I \trianglelefteq R$ ist genau dann maximal, wenn der Ring R/I einfach ist.
b) Ist R kommutativ, so ist $I \trianglelefteq R$ genau dann maximal, wenn R/I Körper ist.

(12.4) Aufgabe: Funktionenringe.
Es seien R ein Ring, M eine Menge, und $U, V \subseteq M$. Man zeige:
a) Die Menge $\mathrm{Abb}(M, R)$ wird durch punktweise Addition $f + g \colon M \to R \colon x \mapsto f(x) + g(x)$ und Multiplikation $f \cdot g \colon M \to R \colon x \mapsto f(x)g(x)$ zu einem Ring. Wann ist $\mathrm{Abb}(M, R)$ kommutativ? Wann ist $\mathrm{Abb}(M, R)$ der Nullring?
b) Die Menge $I_U := \{f \in \mathrm{Abb}(M, R); f(U) = \{0\}\}$ ist ein Ideal von $\mathrm{Abb}(M, R)$, und es gilt genau dann $I_V \subseteq I_U$, wenn $U \subseteq V$ ist. Für $x \in M$ gebe man einen natürlichen Ringisomorphismus $\mathrm{Abb}(M, R)/I_{\{x\}} \cong R$ an.

(12.5) Aufgabe: Teilbarkeit.
Es sei R ein Integritätsbereich.
a) Man zeige: Sind $p \in R$ irreduzibel und $a \in R$, so sind entweder a und p teilerfremd oder es gilt $p \mid a$. Sind $p, q \in R$ irreduzibel, so sind entweder p und q teilerfremd oder es gilt $p \sim q$.
b) Man gebe eine formale Definition von **kleinsten gemeinsamen Vielfachen** von Teilmengen von R an, und formuliere eine Eindeutigkeitsaussage.
c) Man zeige: In R gibt es genau dann immer kleinste gemeinsame Vielfache, wenn es immer größte gemeinsame Teiler gibt. In welcher Beziehung stehen kleinste gemeinsame Vielfache und größte gemeinsame Teiler zueinander?

(12.6) Aufgabe: Faktorielle Ringe.
Es sei R ein faktorieller Ring.
a) Es seien $0 \neq a, b \in R$ teilerfremd und $c \in R$. Man zeige: Es gilt $a \mid bc$ genau dann, wenn $a \mid c$. Gilt $a \mid c$ und $b \mid c$, so folgt $ab \mid c$.
b) Man zeige: In R gibt es immer kleinste gemeinsame Vielfache. Wie kann man sie berechnen?
c) Es sei $K := \mathrm{Q}(R)$ der Quotientenkörper von R. Man zeige: Ist $c \in K$, so gibt es bis auf Assoziiertheit eindeutige teilerfremde $a, b \in R$ mit $c = \frac{a}{b} \in K$.

(12.7) Aufgabe: Hauptidealbereiche.
Es seien R ein Integritätsbereich und $0 \neq p \in R$. Man zeige:
a) Es ist p genau dann prim, wenn R/pR ein Integritätsbereich ist.
b) Ist R ein Hauptidealbereich, so ist p genau dann prim, wenn R/pR ein Körper ist. Welche bekannte Aussage folgt daraus für $R := \mathbb{Z}$?

(12.8) Aufgabe: Euklidische Ringe.
Es sei R ein Euklidischer Ring. Man zeige: Dann gibt es eine Gradabbildung $\epsilon\colon R \setminus \{0\} \to \mathbb{N}_0$ mit $\epsilon(a) \leq \epsilon(ab)$ für alle $0 \neq a, b \in R$.

(12.9) Aufgabe: Euklidischer Algorithmus in $\mathbb{Z}$.
Für $a, b \in \mathbb{N}_0$ sei $\mathrm{ggT}_+(a, b) \in \mathbb{N}_0$ der **nichtnegative** größte gemeinsame Teiler.
a) Für $a := 223\,092\,870$ und $b := 143\,197\,215$ bestimme man $\mathrm{ggT}_+(a, b)$ und zugehörige Bézout-Koeffizienten.
b) Für $k, m, n \in \mathbb{N}$ zeige man: Es gilt $\mathrm{ggT}_+(k^m - 1, k^n - 1) = k^{\mathrm{ggT}_+(m,n)} - 1$.
c) Es sei $[x_n \in \mathbb{N}_0; n \in \mathbb{N}_0]$ die durch $x_0 := 0$, $x_1 := 1$ und $x_{n+2} := x_n + x_{n+1}$ rekursiv definierte Folge der **Fibonacci-Zahlen**. Man zeige: Für $m, n \in \mathbb{N}$ sind x_n und x_{n-1} teilerfremd, und es gilt $\mathrm{ggT}_+(x_m, x_n) = x_{\mathrm{ggT}_+(m,n)}$.

(12.10) Aufgabe: Quadratische Zahlringe.
Für $n \in \mathbb{Z} \setminus \{0, 1\}$ quadratfrei seien $\sqrt{n} \in \mathbb{R}_{\geq 0} \subseteq \mathbb{C}$ für $n > 0$, und $\sqrt{n} := i \cdot \sqrt{|n|} \in \mathbb{C}$ für $n < 0$. Weiter seien $R := \mathbb{Z}[\sqrt{n}] := \{x + y\sqrt{n} \in \mathbb{C}; x, y \in \mathbb{Z}\}$ und $K := \mathbb{Q}[\sqrt{n}] := \{x + y\sqrt{n} \in \mathbb{C}; x, y \in \mathbb{Q}\}$. Man zeige:
a) In der Darstellung von $x + y\sqrt{n} \in K$ sind die Koeffizienten $x, y \in \mathbb{Q}$ eindeutig bestimmt. Es sind $R \subseteq K \subseteq \mathbb{C}$ Teilringe, und R und K sind Integritätsbereiche. Außerdem ist K ein $\mathbb{Q}$-Vektorraum, man bestimme $\dim_{\mathbb{Q}}(K)$.
b) Es ist R ist kein Körper, aber K ist ein Körper. Außerdem ist K der Quotientenkörper $\mathrm{Q}(R)$ von R.

(12.11) Aufgabe: Struktur quadratischer Zahlringe.
Es seien $n \in \mathbb{Z} \setminus \{0, 1\}$ quadratfrei, $K := \mathbb{Q}[\sqrt{n}]$ und $R := \mathbb{Z}[\sqrt{n}]$, sowie $N\colon K \to \mathbb{Q}\colon x + y\sqrt{n} \mapsto x^2 - ny^2$, wobei $x, y \in \mathbb{Q}$, die **Normabbildung**. Man zeige:
a) Es gilt $N(ab) = N(a)N(b)$, für alle $a, b \in K$, und es ist $N(R) \subseteq \mathbb{Z}$. Für $a \in R$ ist genau dann $a \in R^*$, wenn $N(a) \in \mathbb{Z}^* = \{\pm 1\}$ ist. Ist $a \in R$ mit $N(a) \in \mathbb{Z}$ irreduzibel, so ist $a \in R$ irreduzibel.
b) Man bestimme R^* für $n < 0$, und zeige, daß $\mathbb{Z}[\sqrt{3}]^*$ unendlich ist.

c) Die Elemente $2, 3 \in \mathbb{Z}[\sqrt{-5}]$ sind irreduzibel, aber nicht prim.

(12.12) Aufgabe: Quadratische Zahlringe als Euklidische Ringe.
a) Man zeige: Der Ring $\mathbb{Z}[i]$ der **Gaußschen Zahlen** ist Euklidisch bezüglich der Normabbildung $N\colon \mathbb{Z}[i] \to \mathbb{Z}\colon x + iy \mapsto |x + iy|^2 = x^2 + y^2$, wobei $x, y \in \mathbb{Z}$.
b) Man bestimme $\mathrm{ggT}(23 + 11i, 1 - 21i) \subseteq \mathbb{Z}[i]$ und zugehörige Bézout-Koeffizienten. Man gebe die Faktorisierung der Elemente 2, 3 und 5 in $\mathbb{Z}[i]$ an.
c) Man zeige, daß auch die Ringe $\mathbb{Z}[\sqrt{n}]$ für $n \in \{-2, 2, 3\}$ Euklidisch bezüglich der Normabbildung $N\colon \mathbb{Z}[\sqrt{n}] \to \mathbb{Z}\colon x + y\sqrt{n} \mapsto x^2 - ny^2$, wobei $x, y \in \mathbb{Z}$, sind.

Hinweis zu a) und c). Man dividiere in $\mathbb{C}$ und approximiere den Quotienten in $\mathbb{Z}[\sqrt{n}]$.

(12.13) Aufgabe: Größte gemeinsame Teiler in $\mathbb{Z}[X]$.
a) Man zeige: Die Elemente 2 und X besitzen einen größten gemeinsamen Teiler in $\mathbb{Z}[X]$, aber das Ideal $\langle 2, X\rangle \trianglelefteq \mathbb{Z}[X]$ ist kein Hauptideal. Gibt es zugehörige zugehörige Bézout-Koeffizienten?
b) Man zeige: Für einen Integritätsbereich R sind äquivalent:
i) $R[X]$ ist Euklidisch. **ii)** $R[X]$ ist Hauptidealbereich. **iii)** R ist Körper.

(12.14) Aufgabe: Polynomdivision.
Es seien $1 \neq n \in \mathbb{N}$ und $R \in \{\mathbb{Q}, \mathbb{Z}, \mathbb{Z}_n\}$, sowie $f := X^4 + 3X^3 + 2X^2 + X - 1 \in R[X]$ und $g := X^2 - 3X + 1 \in R[X]$. Man bestimme Quotient und Rest der Division von f durch g in $R[X]$. Für welche R ist g ein Teiler von f in $R[X]$?

(12.15) Aufgabe: Euklidischer Algorithmus in $\mathbb{Q}[X]$.
Man bestimme jeweils einen größten gemeinsamen Teiler von f und g in $\mathbb{Q}[X]$ und zugehörige Bézout-Koeffizienten:
a) $f := 3X^3 - 7X^2 + 5X - 1$ und $g := -6X^2 + 5X - 1$.
b) $f := X^8 + X^6 - 3X^4 - 3X^3 + X^2 + 2X - 5$ und $g := 3X^6 + 5X^4 - 4X^2 - 9X + 21$.
c) $f := nX^{n+1} - (n + 1)X^n + 1$ und $g := X^n - nX + n - 1$, wobei $n \in \mathbb{N}$.

(12.16) Aufgabe: Irreduzibilität.
Es seien $R \in \{\mathbb{C}, \mathbb{R}, \mathbb{Q}, \mathbb{Z}, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_5, \mathbb{Z}_7\}$ und $f := X^2 + X + 1 \in R[X]$. Für welche R ist f in $R[X]$ irreduzibel?

(12.17) Aufgabe: Lagrange-Interpolation.
Es seien K ein unendlicher Körper, $n \in \mathbb{N}_0$, sowie $a_0, \ldots, a_n \in K$ paarweise verschiedene **Stützstellen** und $b_0, \ldots, b_n \in K$ **Stützwerte**. Weiter sei $L_i := \prod_{j \neq i} \frac{X - a_j}{a_i - a_j} \in K[X]$, für $i \in \{0, \ldots, n\}$. Man zeige:

a) Es ist $L := \sum_{i=0}^{n} b_i \cdot L_i \in K[X]$ das eindeutig bestimmte **Interpolationspolynom** mit $f(a_i) = b_i$, für alle $i \in \{0, \ldots, n\}$, so daß $L = 0$ oder $\mathrm{Grad}(L) \leq n$.

b) Es ist $\{L_0, \ldots, L_n\}$ eine K-Basis von $K[X]_{\leq n}$. Für $K = \mathbb{Q}$ und die Stützstellen $[a_0, \ldots, a_n] = [0, \ldots, n]$ gebe man die Polynome $L_i \in \mathbb{Q}[X]$ explizit an.

(12.18) Aufgabe: Eigenvektoren.
Es seien K ein Körper, V ein endlich erzeugter K-Vekorraum, $\varphi \in \mathrm{End}_K(V)$ mit Eigenvektoren $v, w \in V$, und $a, b \in K$. Wann ist $av + bw \in V$ ebenfalls ein Eigenvektor von φ?

(12.19) Aufgabe: Eigenräume.
Es sei K ein Körper. Für die folgende Matrix berechne man das charakteristische Polynom, die Eigenwerte, ihre geometrischen und algebraischen Vielfachheiten sowie die Eigenräume, und untersuche sie auf Diagonalisierbarkeit:

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ -2 & 2 & 0 & 2 \\ 1 & 0 & 2 & 0 \\ 2 & -1 & 0 & -1 \end{bmatrix} \in K^{4\times 4}$$

(12.20) Aufgabe: Matrixpotenzen.
Es sei $A := \begin{bmatrix} 5 & -4 \\ 3 & -2 \end{bmatrix} \in \mathbb{Z}^{2\times 2}$. Man berechne $A^{15} \in \mathbb{Z}^{2\times 2}$ unter Verwendung von höchstens zwei Matrixmultiplikationen.

(12.21) Aufgabe: Reelle Matrizen.
Man zeige: Es gibt genau dann eine Matrix in $\mathbb{R}^{n\times n}$ ohne Eigenwerte, wenn $n \in \mathbb{N}$ gerade ist.

(12.22) Aufgabe: Eigenwerte.
Es seien K ein Körper, $n \in \mathbb{N}$, sowie $a, b \in K$ und $A = [a_{ij}]_{ij} \in K^{n\times n}$ definiert durch $a_{ii} := a$ und $a_{ij} := b$ für $i \neq j$. Man zeige:
a) Es gilt $\chi_A = (X - a + b)^{n-1}(X - a - (n-1)b) \in K[X]$ und $\det(A) = (a-b)^{n-1} \cdot (a + (n-1)b)$. Man bestimme die geometrischen Vielfachheiten der Eigenwerte von A.
c) Es gilt $\mu_A = (X-a+b)(X-a-(n-1)b) \in K[X]$. Wann ist A diagonalisierbar?

Hinweis. Man betrachte zunächst den Fall $a = b$.

(12.23) Aufgabe: Charakteristisches Polynom.
Es seien K ein Körper, $m, n \in \mathbb{N}_0$, sowie $A \in K^{m\times n}$ und $B \in K^{n\times m}$. Man zeige: Es gilt $X^n \cdot \chi_{AB} = X^m \cdot \chi_{BA} \in K[X]$.

Hinweis. Man benutze $\left[\begin{array}{c|c} XE_m & -A \\ \hline 0 & E_n \end{array}\right], \left[\begin{array}{c|c} E_m & A \\ \hline B & XE_n \end{array}\right] \in K[X]^{(m+n)\times(m+n)}$.

(12.24) Aufgabe: Satz von Cayley-Hamilton.
Es seien K ein Körper und $A \in K^{n \times n}$, wobei $n \in \mathbb{N}$. Man zeige:
a) Es gibt ein Polynom $f \in K[X]$ mit $\operatorname{adj}(A) = f(A) \in K^{n \times n}$.
b) Ist $A \in \mathrm{GL}_n(K)$, so gibt es ein Polynom $g \in K[X]$ mit $A^{-1} = g(A) \in K^{n \times n}$.

(12.25) Aufgabe: Minimalpolynom.
Es seien K ein Körper und $A, B \in K^{n \times n}$, wobei $n \in \mathbb{N}_0$. Man zeige:
a) Für die zugehörigen Minimalpolynome gilt $\mu_{AB} \mid X \cdot \mu_{BA} \in K[X]$.
b) Ist $A \in \mathrm{GL}_n(K)$ oder $B \in \mathrm{GL}_n(K)$, so gilt sogar $\mu_{AB} = \mu_{BA} \in K[X]$. Kann man auf die Zusatzvoraussetzung verzichten?

(12.26) Aufgabe: Charakteristisches und Minimalpolynom.
Es seien K ein Körper und $f \in K[X]$ normiert mit $\mathrm{Grad}(f) = 2$.
a) Man zeige: Es gibt eine Matrix $A \in K^{2 \times 2}$ mit $\chi_A = \mu_A = f$.
b) Wann gibt es eine Matrix $A \in K^{2 \times 2}$ mit $\chi_A = f$ und $\mu_A \neq f$?

(12.27) Aufgabe: Hamilton-Quaternionen.
Es seien $a_1, \ldots, a_4 \in \mathbb{R}$ und

$$A := \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \in \mathbb{R}^{4 \times 4}.$$

Man zeige: Ist $[a_2, a_3, a_4] \neq [0, 0, 0]$, so gilt $\mu_A = X^2 - 2a_1X + (a_1^2 + a_2^2 + a_3^2 + a_4^2) \in \mathbb{R}[X]$ und $\chi_A = \mu_A^2$.

(12.28) Aufgabe: Fitting-Zerlegung.
Es seien K ein Körper, $V := K^{n \times 1}$ und $A \in K^{n \times n}$, wobei $n \in \mathbb{N}_0$.
a) Man zeige: $\mathrm{Kern}(A^i)$ und $\mathrm{Bild}(A^i)$ sind A-invariante Teilräume von V, und für alle $i \in \mathbb{N}_0$ gilt $\mathrm{Bild}(A^i) \geq \mathrm{Bild}(A^{i+1})$ und $\mathrm{Kern}(A^i) \leq \mathrm{Kern}(A^{i+1})$ sowie

$$\dim_K(\mathrm{Kern}(A^{i+1})/\mathrm{Kern}(A^i)) \geq \dim_K(\mathrm{Kern}(A^{i+2})/\mathrm{Kern}(A^{i+1})).$$

b) Es sei $m := \min\{i \in \mathbb{N}_0; \mathrm{Kern}(A^i) = \mathrm{Kern}(A^{i+1})\}$. Man zeige: Es gilt $\mathrm{Kern}(A^m) = \mathrm{Kern}(A^{m+i})$ und $\mathrm{Bild}(A^m) = \mathrm{Bild}(A^{m+i})$ für alle $i \in \mathbb{N}_0$, sowie $\dim_K(V) \leq m \cdot \dim_K(\mathrm{Kern}(A))$.
c) Man zeige: Es gilt $V = \mathrm{Kern}(A^m) \oplus \mathrm{Bild}(A^m)$. Man bestimme das Minimalpolynom von $\varphi_A|_{\mathrm{Kern}(A^m)}$, und zeige, daß $\varphi_A|_{\mathrm{Bild}(A^m)}$ invertierbar ist.

(12.29) Aufgabe: Simultane Diagonalisierbarkeit.
Es seien K ein Körper und V ein endlich erzeugter K-Vektorraum. Es heißen $\varphi, \psi \in \mathrm{End}_K(V)$ **simultan diagonalisierbar**, wenn es eine K-Basis von V aus Eigenvektoren von φ und ψ gibt. Man zeige: φ und ψ sind genau dann simultan diagonalisierbar, wenn sie diagonalisierbar sind und $\varphi\psi = \psi\varphi$ gilt.

(12.30) Aufgabe: Formale Ableitung.
Es seien K ein Körper und $K[X]_{\leq n} := \{f \in K[X]; f = 0 \text{ oder } \mathrm{Grad}(f) \leq n\}$, wobei $n \in \mathbb{N}_0$.
a) Man zeige: Durch $X^i \mapsto iX^{i-1}$, für $i \in \mathbb{N}$, und $1 \mapsto 0$ wird eine Abbildung $\frac{\partial}{\partial X} \in \mathrm{End}_K(K[X])$ definiert, für die zudem die **Produktregel** $\frac{\partial}{\partial X}(fg) = \frac{\partial}{\partial X}(f) \cdot g + f \cdot \frac{\partial}{\partial X}(g) \in K[X]$ gilt; $\frac{\partial}{\partial X}$ heißt die **formale Ableitung** nach X.
b) Für $f \in K[X]$ und $a \in K$ sei $\nu_a(f) \in \mathbb{N}_0$ die zugehörige **Nullstellenordnung**, das heißt die Vielfachheit von $X - a$ als Primfaktor von f. Man zeige: Es ist genau dann $\nu_a(f) = 1$, wenn $f(a) = 0$ und $\frac{\partial}{\partial X}(f)(a) \neq 0$ ist.
c) Man zeige: $K[X]_{\leq n}$ ist ein $\frac{\partial}{\partial X}$-invarianter K-Teilraum mit K-Basis $B_n := \{1, X, X^2, \ldots, X^n\}$. Man berechne die Abbildungsmatrix $A := {}_{B_n}(\frac{\partial}{\partial X})_{B_n}$, und das charakteristische Polynom und die Eigenwerte von A.
d) Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man bestimme die algebraischen und geometrischen Vielfachheiten der Eigenwerte, sowie das Minimalpolynom von A. Man berechne gegebenenfalls die Jordan-Normalform von A. Ist A diagonalisierbar?

(12.31) Aufgabe: Haupträume.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man bestimme jeweils das charakteristische und das Minimalpolynom, sowie die Haupträume der folgenden Matrizen in $K^{6 \times 6}$, zusammen mit geeigneten Transformationsmatrizen. Welche der Matrizen sind diagonalisierbar? Wie hängen die Rechnungen von K ab?

$$\textbf{a)} \begin{bmatrix} -1 & 1 & 2 & 2 & -1 & 5 \\ -2 & 2 & 2 & 2 & -1 & 5 \\ 2 & -1 & 1 & -2 & -1 & 5 \\ -1 & 1 & 0 & 1 & 0 & 0 \\ -2 & 0 & 2 & 4 & 1 & -4 \\ 1 & -2 & 1 & 2 & 1 & -3 \end{bmatrix} \qquad \textbf{b)} \begin{bmatrix} -3 & 4 & -2 & -1 & 4 & -1 \\ -3 & 3 & -2 & -1 & 4 & -2 \\ 1 & -4 & 1 & 2 & -2 & -1 \\ -2 & 2 & -2 & 0 & 4 & -1 \\ 2 & -3 & 1 & 1 & -1 & 0 \\ 2 & -2 & 2 & 1 & -4 & 2 \end{bmatrix}$$

(12.32) Aufgabe: Jordan-Normalform in Dimension 3.
Man bestimme jeweils das charakteristische und das Minimalpolynom, sowie die Jordan-Normalform der folgenden Matrizen in $\mathbb{C}^{3 \times 3}$, zusammen mit geeigneten Transformationsmatrizen. Welche der Matrizen sind diagonalisierbar?

$$\textbf{a)} \begin{bmatrix} -1 & 1 & -1 \\ 1 & 2 & -3 \\ 1 & 3 & -4 \end{bmatrix} \quad \textbf{b)} \begin{bmatrix} -1 & -2 & 2 \\ -1 & 0 & 1 \\ -3 & -3 & 4 \end{bmatrix} \quad \textbf{c)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \textbf{d)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(12.33) Aufgabe: Jordan-Normalform in Dimension 4.
Man bestimme jeweils das charakteristische und das Minimalpolynom, sowie die Jordan-Normalform der folgenden Matrizen in $\mathbb{C}^{4 \times 4}$, zusammen mit geeigneten Transformationsmatrizen. Welche der Matrizen sind diagonalisierbar?

a) $\begin{bmatrix} 2 & 0 & -3 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -4 & 1 \\ -5 & -2 & -1 & -1 \end{bmatrix}$ **b)** $\begin{bmatrix} 2 & -2 & 1 & -1 \\ 0 & 2 & 0 & 0 \\ -1 & -5 & 3 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix}$ **c)** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

d) $\begin{bmatrix} -2 & -1 & 0 & 3 \\ 3 & -2 & 3 & 0 \\ 0 & 1 & -2 & -3 \\ 1 & 0 & 1 & -2 \end{bmatrix}$ **e)** $\begin{bmatrix} -2 & -2 & 1 & 1 \\ 3 & 3 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ -3 & -2 & 1 & 2 \end{bmatrix}$ **f)** $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(12.34) Aufgabe: Jordan-Normalform.
Man bestimme jeweils das charakteristische und das Minimalpolynom, sowie die Jordan-Normalform der folgenden Matrizen über $\mathbb{C}$, zusammen mit geeigneten Transformationsmatrizen. Welche der Matrizen sind diagonalisierbar?

a) $\begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 0 & -1 & 1 \\ -2 & 3 & -1 & 1 & 0 & 1 \\ 1 & -6 & 6 & -1 & -6 & 0 \\ 0 & 3 & -3 & 1 & 2 & 0 \\ -6 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ **b)** $\begin{bmatrix} -4 & -4 & 4 & 0 & 4 & 0 \\ 0 & 2 & 0 & -2 & 0 & -1 \\ -2 & -4 & 2 & 0 & 4 & 0 \\ -1 & 2 & 2 & 4 & -2 & 1 \\ 1 & -2 & -2 & -2 & 4 & -1 \\ 2 & -4 & -4 & -4 & 4 & 0 \end{bmatrix}$

(12.35) Aufgabe: Jordan-Normalform von Matrixserien.
Es seien K ein Körper, $a \in K$ und $A \in K^{n \times n}$, wobei $n \in \mathbb{N}_0$. Man bestimme die Jordan-Normalform von $\left[\begin{array}{c|c} aE_n & 0 \\ \hline A & aE_n \end{array}\right] \in K^{2n \times 2n}$.

(12.36) Aufgabe: Jordan-Normalform in kleiner Dimension.
Für $n \in \{1, 2, 3\}$ gebe man Vertreter der Ähnlichkeitsklassen von Matrizen in $\mathbb{C}^{n \times n}$, und jeweils das charakteristische und das Minimalpolynom an. Welche davon sind diagonalisierbare Matrizen?

(12.37) Aufgabe: Matrixwurzeln.
Es seien K ein Körper und $n \geq 2$. Gibt es $A \in K^{n \times n}$ mit $A^2 = C_n(0)$?

(12.38) Aufgabe: Nilpotente Matrizen.
Es seien K ein Körper und $n \in \mathbb{N}_0$. Eine Matrix $[a_{ij}]_{ij} \in K^{n \times n}$ mit $a_{ij} = 0$ für alle $i \leq j$ heißt eine **echte (untere) Dreiecksmatrix**. Eine Matrix $A \in K^{n \times n}$ heißt **nilpotent**, falls es $m \in \mathbb{N}_0$ gibt mit $A^m = 0$. Man zeige die Äquivalenz der folgenden Aussagen:
i) A ist ähnlich zu einer echten Dreiecksmatrix. **ii)** A ist nilpotent.
iii) Es ist $\mu_A = X^l$ für ein $l \in \mathbb{N}_0$. **iv)** Es ist $\chi_A = X^n$.

(12.39) Aufgabe: Kreisteilung.
Man zeige: Jede quadratische Matrix über $\mathbb{C}$, die Nullstelle von $X^n - 1 \in \mathbb{C}[X]$ für ein $n \in \mathbb{N}$ ist, ist diagonalisierbar.

(12.40) Aufgabe: Smith-Normalform über Körpern.
Es seien K ein Körper sowie $m, n \in \mathbb{N}_0$. Matrizen $A, B \in K^{m \times n}$ heißen **äquivalent**, falls es $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ gibt mit $B = PAQ$. Man zeige: A und B sind genau dann äquivalent, wenn sie die gleiche Smith-Normalform besitzen. Man gebe die möglichen Smith-Normalformen an.

(12.41) Aufgabe: Allgemeine Smith-Normalform.
Man bestimme die Smith-Normalform der allgemeinen Matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{Z}^{2 \times 2}$.

(12.42) Aufgabe: Smith-Normalform über $\mathbb{Z}$.
Man berechne die Smith-Normalform der folgenden Matrizen, und bestimme ihre Invarianten-, Elementar- und Determinantenteiler; für a) und b) gebe man auch geeignete Transformationsmatrizen an:

a) $\begin{bmatrix} 20 & 15 & 16 & 8 \\ 14 & 7 & 10 & 8 \\ 10 & 7 & 8 & 4 \\ -18 & 16 & 12 & -30 \end{bmatrix} \in \mathbb{Z}^{4 \times 4}$ **b)** $\begin{bmatrix} 1 & 4 & 0 & 0 & -2 \\ 1 & 3 & 3 & -1 & 1 \\ -1 & 4 & 1 & -1 & 0 \\ -1 & 2 & -4 & 0 & 0 \\ 4 & -1 & 0 & 4 & 2 \end{bmatrix} \in \mathbb{Z}^{5 \times 5}$

c) $[a_{ij}]_{ij} \in \mathbb{Z}^{n \times n}$, wobei $n \in \mathbb{N}$, sowie $a_{ii} := 2$, und $a_{ij} := -1$ für $|i - j| = 1$, und $a_{ij} := 0$ für $|i - j| \geq 2$.

(12.43) Aufgabe: Lineare Gleichungssysteme über $\mathbb{Z}$.
In Abhängigkeit von $a, b, c, d \in \mathbb{Z}$ bestimme man die ganzzahligen Lösungen des linearen Gleichungssystems $AX^{\mathrm{tr}} = [a, b, c, d]^{\mathrm{tr}}$, wobei

$$A := \begin{bmatrix} 0 & 6 & 3 & -3 \\ 0 & 2 & 1 & -1 \\ 4 & -4 & 0 & 2 \\ 2 & 6 & 1 & -3 \end{bmatrix} \in \mathbb{Z}^{4 \times 4}.$$

(12.44) Aufgabe: Untermoduln freier Moduln.
Es seien R ein Integritätsbereich, V ein R-Modul mit endlicher R-Basis und $U < V$ ein echter R-Teilmodul.
a) Besitzt U dann notwendig eine R-Basis?
b) Ist R ein Hauptidealring, gilt dann notwendig $\mathrm{rk}_R(U) < \mathrm{rk}_R(V)$?
c) Man gebe bis auf Isomorphie alle R-Untermoduln des Hauptidealrings R an.

(12.45) Aufgabe: Abelsche Gruppen.
Es seien $p \in \mathbb{N}$ eine Primzahl und $n \in \{1, \ldots, 5\}$. Man gebe die Isomorphisklassen abelscher Gruppen der Kardinalität p^n an.

(12.46) Aufgabe: Ähnlichkeit.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man untersuche die folgenden Matrizen auf Ähnlichkeit:

a) $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 1 \\ 1 & -1 & 0 \end{bmatrix} \in K^{3\times 3}$ und $\begin{bmatrix} 1 & -2 & 0 \\ -3 & 3 & 2 \\ 3 & -5 & -1 \end{bmatrix} \in K^{3\times 3}$.

b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -2 & 0 \end{bmatrix} \in K^{4\times 4}$ und $\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \in K^{4\times 4}$.

(12.47) Aufgabe: Charakteristische Matrix.
Es seien $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Für die folgenden Matrizen berechne man die Smith-Normalform ihrer charakteristischen Matrix, zusammen mit geeigneten Transformationsmatrizen, und bestimme ihre Invarianten-, Elementar- und Determinantenteiler. Daraus bestimme man ihre Frobenius-, Weierstraß- und Jordan-Normalform, und gebe geeignete Transformationsmatrizen an. Wie hängen die Rechnungen von K ab?

a) $\begin{bmatrix} -3 & -4 & 2 & -3 \\ -1 & -2 & 0 & -2 \\ -1 & 2 & 4 & 4 \\ 2 & 2 & -2 & 1 \end{bmatrix}$ **b)** $\begin{bmatrix} -2 & 4 & -2 & -2 \\ -2 & 4 & -1 & -1 \\ 2 & -2 & 3 & 1 \\ 2 & -2 & 1 & 3 \end{bmatrix}$ **c)** $\begin{bmatrix} 0 & -2 & -2 & 1 \\ 1 & -2 & -1 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{bmatrix}$

(12.48) Aufgabe: Normalformen.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man bestimme die Frobenius-, Weierstraß- und Jordan-Normalform einer quadratischen Matrix über K, deren charakteristische Matrix die nicht-konstanten Invariantenteiler $d_1 := X+1$ und $d_2 := (X+1)^2(X+2)$, sowie $d_3 := (X+1)^2(X+2)(X-2)^2$ und $d_4 := (X+1)^2(X+2)^2(X-2)^2$ hat.

(12.49) Aufgabe: Ähnlichkeitsklassen.
Es sei $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Man bestimme jeweils die Ähnlichkeitsklassen quadratischer Matrizen A über K mit folgender Eigenschaft; für a) gebe man auch die zugehörigen Frobenius-, Weierstraß- und Jordan-Normalformen an:
a) $\chi_A = (X+1)(X-1)^3(X^2+1)^2 \in K[X]$.
b) $\mu_A = (X+1)(X-1)(X^2+1) \in K[X]$.

(12.50) Aufgabe: Normalformen in kleiner Dimension.
Es seien K ein Körper und $n \in \{0, 1, 2, 3\}$. Man zeige: Zwei Matrizen in $K^{n\times n}$ sind genau dann ähnlich, wenn ihre charakteristischen und ihre Minimalpolynome übereinstimmen. Kann man auf die Voraussetzung $n \leq 3$ verzichten?

(12.51) Aufgabe: Begleitmatrizen.
Es seien K ein Körper und $f \in K[X]$ normiert. Man bestimme das charakteristische und das Minimalpolynom der Begleitmatrix $C(f)$, sowie die Invarianten-, Elementar- und Determinantenteiler der charakteristischen Matrix von $C(f)$.

(12.52) Aufgabe: Transponierte Matrizen.
Es seien K ein Körper und $A \in K^{n \times n}$, wobei $n \in \mathbb{N}_0$. Man zeige: A und A^{tr} sind ähnlich.

(12.53) Aufgabe: Permutationsmatrizen.
Eine Matrix $A := [a_{ij}]_{ij} \in \mathbb{R}^{n \times n}$, wobei $n \in \mathbb{N}_0$, heißt die zu $\pi \in \mathcal{S}_n$ gehörige **Permutationsmatrix**, falls $a_{\pi(j),j} = 1$ und $a_{i,j} = 0$ für alle $i, j \in \{1, \ldots, n\}$ mit $i \neq \pi(j)$ ist. Man zeige:
a) Ist $A \in \mathbb{R}^{n \times n}$ eine Permutationsmatrix, so ist orthogonal. Man bestimme $\det(A)$.
b) Ist $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ orthogonal mit $a_{ij} \geq 0$ für alle $i, j \in \{1, \ldots, n\}$, so ist A eine Permutationsmatrix.

(12.54) Aufgabe: Orthogonalräume.
Es seien K ein Körper, Φ eine α-Sesquilinearform auf dem endlich erzeugten K-Vektorraum V, sowie $\mathcal{I}$ eine Menge und $U_i \leq V$ für $i \in \mathcal{I}$. Man zeige:
a) Es gilt $(\sum_{i \in \mathcal{I}} U_i)^\perp = \bigcap_{i \in \mathcal{I}} U_i^\perp$.
b) Ist Φ nicht-ausgeartet, so gilt $(\bigcap_{i \in \mathcal{I}} U_i)^\perp = \sum_{i \in \mathcal{I}} U_i^\perp$.

(12.55) Aufgabe: Spurform.
Es seien $[K, \alpha] \in \{[\mathbb{R}, \mathrm{id}], [\mathbb{C}, \overline{\ }]\}$, und für $n \in \mathbb{N}_0$ seien $V := K^{n \times n}$ und $\tau \colon V \to K \colon [a_{ij}]_{ij} \mapsto \sum_{i=1}^n a_{ii}$.
a) Man zeige: Durch $\Phi(A, B) := \tau(A^\alpha B) \in K$ und $\Psi(A, B) := \tau(A^* B)$ werden hermitesche nicht-ausgeartete α-Sesquilinearformen auf V definiert. Welche davon sind Skalarprodukte?
b) Es sei $U_\epsilon := \{A \in V; A^{\mathrm{tr}} = \epsilon A\}$, wobei $\epsilon \in \{\pm 1\}$. Bezüglich Φ zeige man: Es gilt $U_\epsilon \leq V$ mit $(U_\epsilon)^\perp = U_{-\epsilon}$, sowie $V = U_1 \oplus U_{-1}$. Außerdem gilt $\epsilon\Phi(v, v) > 0$ für alle $0 \neq v \in U_\epsilon$.
c) Man untersuche die direkte Zerlegung von V aus b) bezüglich Ψ.

(12.56) Aufgabe: Signatur.
Man bestimme die Signatur der $\mathbb{R}$-Bilinearformen auf $V := \mathbb{R}^{n \times 1}$, wobei $n \geq 2$, deren Gram-Matrizen $[a_{ij}]_{ij} \in \mathbb{R}^{n \times n}$ bezüglich der $\mathbb{R}$-Standardbasis von V wie folgt gegeben sind, und gebe jeweils $V^\perp$ und eine $\mathbb{R}$-Orthogonalbasis an:
a) $a_{ii} := 2$, und $a_{ij} := -1$ für $|i - j| = 1$, und $a_{ij} := 0$ sonst.
b) $a_{ii} := 2$, und $a_{ij} := -1$ für $|i - j| \in \{1, n - 1\}$, und $a_{ij} := 0$ sonst.
c) $a_{ii} := 1$, und $a_{ij} := -1$ für $|i - j| = 1$, und $a_{ij} := 0$ sonst.

(12.57) Aufgabe: Normierte Vektorräume.
Es seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und V ein endlich erzeugter normierter K-Vektorraum mit Norm $\| \cdot \|$. Man zeige:
a) Ist V Euklidisch, so gilt der **Cosinus-Satz**: Für $0 \neq v, w \in V$ mit Zwischenwinkel $0 \leq \omega \leq \pi$ gilt $\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\cos(\omega) \cdot \|v\| \cdot \|w\|$. Welcher Spezialfall ergibt sich im Falle $v \perp w$?

b) Ist V unitär, so gilt die **Parallelogramm-Gleichung von von Neumann**: Für alle $v, w \in V$ gilt $\|v-w\|^2 + \|v+w\|^2 = 2(\|v\|^2 + \|w\|^2)$. Welche geometrische Interpretation hat diese Gleichung?
c) Gilt in V die Parallelogramm-Gleichung, so ist V ein Hilbert-Raum.
d) Es sei $V := K^n$, wobei $n \in \mathbb{N}$, und für $v = [a_1, \ldots, a_n] \in V$ seien $\|v\|_1 := \sum_{i=1}^n |a_i|$ die **1-Norm** und $\|v\|_\infty := \max\{|a_i|; i \in \{1, \ldots, n\}\}$ die **∞-Norm**. Man zeige: Mit $\| \cdot \|_1 \colon V \to \mathbb{R}$ und $\| \cdot \|_\infty \colon V \to \mathbb{R}$ wird V jeweils zu einem normierten K-Vektorraum, aber nicht zu einem Hilbert-Raum.

(12.58) Aufgabe: Quadratische Formen.
Es seien K ein Körper, Φ eine hermitesche α-Sesquilinearform auf dem K-Vektorraum V, wobei für $\alpha = \mathrm{id}_K$ zusätzlich $2 \neq 0 \in K$ vorausgesetzt werde. Man zeige: Φ ist bereits durch die zugehörige quadratische Form $q \colon V \to K \colon v \mapsto \Phi(v, v)$ eindeutig bestimmt.

(12.59) Aufgabe: Skalarprodukte.
Es seien $[K, \alpha] \in \{[\mathbb{R}, \mathrm{id}], [\mathbb{C}, \overline{\ }]\}$ und $A \in K^{n \times n}$, wobei $n \in \mathbb{N}_0$. Man zeige: Durch $\langle v, w \rangle := v^* \cdot A \cdot w$ wird genau dann ein Skalarprodukt auf $K^{n \times 1}$ definiert, wenn es $P \in \mathrm{GL}_n(K)$ gibt mit $A = P^* \cdot P$; in diesem Fall kann P sogar als Dreiecksmatrix mit positiven Diagonaleinträgen gewählt werden.

(12.60) Aufgabe: Gram-Schmidt-Orthonormalisierung.
Es sei $\mathbb{R}^{n \times 1}$, für $n \in \mathbb{N}_0$, versehen mit dem Standardskalarprodukt.
a) Man berechne die zur $\mathbb{R}$-Basis $\{[1,1,0]^{\mathrm{tr}}, [1,0,1]^{\mathrm{tr}}, [0,1,1]^{\mathrm{tr}}\} \subseteq \mathbb{R}^{3 \times 1}$ gehörige Gram-Schmidt-Basis.
b) Es sei $U := \{[a_1, \ldots, a_n]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}; \sum_{i=1}^n a_i = 0\} \leq \mathbb{R}^{n \times 1}$. Man berechne $\mathbb{R}$-Orthonormalbasen von U und von $U^\perp \leq \mathbb{R}^{n \times 1}$.

(12.61) Aufgabe: Legendre-Polynome.
Für $n \in \mathbb{N}_0$ sei $\mathcal{P}_n(\mathbb{R}) \leq C^{\mathrm{pol}}(\mathbb{R})$ der $\mathbb{R}$-Vektorraum der Polynomfunktionen vom Grad $\leq n$, mit $\mathbb{R}$-Basis $P_n := \{p_0, \ldots, p_n\}$, wobei $p_n \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto x^n$.
a) Man zeige: $\langle f, g \rangle := \int_{t \in [-1,1]} f(t)g(t)$ definiert ein Skalarprodukt auf $\mathcal{P}_n(\mathbb{R})$.
b) Es sei $L_n = \{l_0, \ldots, l_n\}$ die zu P_n gehörige Gram-Schmidt-Basis von $\mathcal{P}_n(\mathbb{R})$. Man zeige: Es ist $l_n \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \frac{1}{2^n \cdot n!} \cdot \sqrt{\frac{2n+1}{2}} \cdot (\frac{\partial}{\partial x})^n ((x^2 - 1)^n)$ das **Legendre-Polynom** vom Grad n.

(12.62) Aufgabe: Gram-Determinanten.
Es seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und Φ das Standardskalarprodukt auf $K^{n \times 1}$, wobei $n \in \mathbb{N}_0$. Für $r \leq n$ sei die $\Delta \colon K^{n \times r} \to K \colon [v_1, \ldots, v_r] \mapsto \det([\langle v_i, v_j \rangle]_{ij})$ die zugehörige **Gram-Determinante**. Man zeige:
a) Für alle $v_1, \ldots, v_r \in K^{n \times 1}$ gilt $\Delta(v_1, \ldots, v_r) \geq 0$, und es ist $\Delta(v_1, \ldots, v_r) = 0$ genau dann, wenn $\dim_K(\langle v_1, \ldots, v_r \rangle_K) < r$ ist.

b) Es ist $\mathbb{R}^{n\times r} \to \mathbb{R}\colon [v_1,\ldots,v_r] \mapsto \sqrt{\Delta(v_1,\ldots,v_r)}$ eine Determinantenform. Ist $\dim_K(\langle v_1,\ldots,v_r\rangle_K) = r$, so beschreibe man $\sqrt{\Delta(v_1,\ldots,v_r)}$ bezüglich einer $\mathbb{R}$-Orthonormalbasis von $\langle v_1,\ldots,v_r\rangle_K$.

(12.63) Aufgabe: Dualraum.
Es seien K ein Körper, Φ eine nicht-ausgeartete α-Sesquilinearform auf dem endlich erzeugten K-Vektorraum V mit K**-Dualraum** $V^* := \mathrm{Hom}_K(V,K)$.
a) Man zeige: Für $v \in V$ wird durch $V \to K\colon w \mapsto \Phi(v,w)$ eine K**-Linearform** $\Phi_v \in V^*$ definiert, und $\Phi^*\colon V \to V^*_\alpha\colon v \mapsto \Phi_v$ ist ein K-Isomorphismus.
b) Für $\varphi \in \mathrm{End}_K(V)$ sei $\varphi^\circ \in \mathrm{End}_K(V^*)$ gegeben durch $\varphi^\circ\colon \lambda \mapsto \lambda\varphi$. Man zeige: Es ist $\varphi^* := (\Phi^*)^{-1}\varphi^\circ\Phi^* \in \mathrm{End}_K(V)$.

(12.64) Aufgabe: Adjungierte Endomorphismen.
Es seien K ein Körper, Φ eine nicht-ausgeartete α-Sesquilinearform auf dem endlich erzeugten K-Vektorraum V und $\varphi \in \mathrm{End}_K(V)$; der von α induzierte Ringautomorphismus von $K[X]$ werde ebenfalls mit α bezeichnet.
a) Man zeige: Es gilt $\chi_{\varphi^*} = (\chi_\varphi)^\alpha \in K[X]$ und $\mu_{\varphi^*} = (\mu_\varphi)^\alpha \in K[X]$.
b) Es seien zudem Φ hermitesch und φ unitär. Man zeige: Es gilt $\chi_\varphi(X) = \det(\varphi)\cdot(-X)^n\cdot(\chi_\varphi)^\alpha(X^{-1}) \in K(X)$, wobei $n := \dim_K(V) \in \mathbb{N}_0$.
c) Es seien $2 \neq 0 \in K$, sowie zudem Φ symmetrisch und φ orthogonal. Man zeige: Ist $\det(\varphi) = -1$, so hat φ den Eigenwert -1; ist $(-1)^n\det(\varphi) = -1$, so hat φ den Eigenwert 1.

(12.65) Aufgabe: Normale Endomorphismen.
Es seien V ein Euklidischer oder unitärer K-Vektorraum und $\varphi \in \mathrm{End}_K(V)$.
a) Man zeige: φ ist genau dann normal, wenn es $f \in K[X]$ gibt mit $\varphi^* = f(\varphi)$.
b) Es sei φ normal. Man zeige: Es gilt $\mathrm{Kern}(\varphi) = \mathrm{Kern}(\varphi^*) = \mathrm{Bild}(\varphi)^\perp$ und $\mathrm{Bild}(\varphi) = \mathrm{Bild}(\varphi^*) = \mathrm{Kern}(\varphi)^\perp$ sowie $V = \mathrm{Kern}(\varphi) \oplus \mathrm{Bild}(\varphi)$; und Eigenvektoren von φ zu verschiedenen Eigenwerten sind orthogonal zueinander.
c) Es sei φ normal. Man zeige: Alle Eigenwerte von $\varphi\varphi^*$ sind nicht-negativ.
d) Es seien $\varphi, \psi \in \mathrm{End}_K(V)$ normal. Man zeige: Es gilt genau dann $\varphi\psi = 0$, wenn $\psi\varphi = 0$ ist.

(12.66) Aufgabe: Unitäre Endomorphismen.
Es seien V ein unitärer Vektorraum, und $v, w \in V$. Man zeige: Es gibt genau dann eine unitäre Abbildung $\varphi \in \mathrm{End}_\mathbb{C}(V)$ mit $\varphi(v) = w$, wenn $\|v\| = \|w\|$ gilt.

(12.67) Aufgabe: Orthogonale Endomorphismen.
Es seien V ein Euklidischer Vektorraum und $\varphi \in \mathrm{End}_\mathbb{R}(V)$ orthogonal. Man zeige: Es ist $\nu_{-1}(\varphi) = \gamma_{-1}(\varphi) \in \mathbb{N}_0$, und es gilt $\det(\varphi) = (-1)^{\nu_{-1}(\varphi)}$.

(12.68) Aufgabe: Orthogonale Normalform.
Es sei $A = [a_{ij}]_{ij} \in \mathbb{R}^{n\times n}$ gegeben durch $a_{i+1,i} := 1$ für alle $i \in \{1,\ldots,n-1\}$, sowie $a_{1,n} := 1$ und $a_{ij} := 0$ sonst. Man zeige: Die Matrix A ist orthogonal. Man bestimme die orthogonale Normalform von A.

(12.69) Aufgabe: Hermitesche Projektionen.
Es seien V ein Euklidischer oder unitärer K-Vektorraum und $\varphi \in \mathrm{End}_K(V)$ mit $\varphi^2 = \varphi$. Man zeige: Es gilt $V = \mathrm{Kern}(\varphi) \oplus \mathrm{Bild}(\varphi)$; und es gilt genau dann $\varphi = \varphi^*$, wenn $\mathrm{Kern}(\varphi) = \mathrm{Bild}(\varphi)^\perp$ ist.

(12.70) Aufgabe: Hermitesche Endomorphismen.
Es seien V ein Euklidischer oder unitärer Vektorraum und $\varphi \in \mathrm{End}_K(V)$.
a) Es sei $K = \mathbb{C}$. Man zeige: Es ist φ genau dann hermitesch, wenn $\langle \varphi(v), v \rangle \in \mathbb{R}$ für alle $v \in V$ gilt.
b) Es sei φ hermitesch. Man zeige: Es gilt genau dann $\langle \varphi(v), v \rangle \geq 0$ für alle $v \in V$, wenn alle Eigenwerte von φ nicht-negativ sind; es gilt genau dann $\langle \varphi(v), v \rangle > 0$ für alle $0 \neq v \in V$, wenn alle Eigenwerte von φ positiv sind.

(12.71) Aufgabe: Symmetrische Matrizen.
Es sei $A \in \mathbb{R}^{n \times n}$, wobei $n \in \mathbb{N}_0$. Man zeige: Es ist A genau dann symmetrisch, wenn es $P \in \mathbb{C}^{n \times n}$ gibt mit $A = P^{\mathrm{tr}} P$.

(12.72) Aufgabe: Hermitesche Sesquilinearformen.
Es sei V ein Euklidischer oder unitärer K-Vektorraum mit Skalarprodukt Γ. Dann heißen α-Sesquilinearformen Φ und Ψ auf V **ähnlich** bzw. **isometrisch**, falls $\varphi \in \mathrm{GL}(V)$ bzw. $\varphi \in \mathrm{GO}(V)$ existiert mit $\Psi(v, w) = \Phi(\varphi(v), \varphi(w))$, für alle $v, w \in V$.
a) Man zeige: Dies definiert Äquivalenzrelationen auf der Menge der α-Sesquilinearformen und auf der Menge der hermiteschen α-Sesquilinearformen auf V.
b) Man bestimme Vertreter der Ähnlichkeits- und Isometrieklassen hermitescher α-Sesquilinearformen auf V. Welche davon sind Skalarprodukte?

13 References

[1] A. Beutelspacher: Lineare Algebra, 7. Auflage, Teubner, 2010.

[2] S. Bosch: Lineare Algebra, 4. Auflage, Springer, 2008.

[3] G. Fischer: Lineare Algebra, eine Einführung für Studienanfänger, 15. Auflage, Vieweg, 2005.

[4] G. Fischer: Analytische Geometrie, eine Einführung für Studienanfänger, 7. Auflage, Vieweg, 2001.

[5] W. Greub: Linear algebra, 4. edition, Graduate Texts in Mathematics 23, Springer, 1981.

[6] J. Heinhold, B. Riedmüller: Lineare Algebra und analytische Geometrie, 2 Bände, 3. Auflage, Hanser, 1980.

[7] B. Huppert: Angewandte lineare Algebra, de Gruyter, 1990.

[8] B. Huppert, W. Willems: Lineare Algebra, 2. Auflage, Teubner, 2010.

[9] N. Jacobson: Lectures in abstract algebra, vol. II: linear algebra, 2. reprint of the 1951-1964 edition, Graduate Texts in Mathematics 31, Springer, 1984.

[10] K. Jänich: Lineare Algebra, mit 110 Testfragen, 11. Auflage, Springer, 2008.

[11] M. Koecher: Lineare Algebra und analytische Geometrie, 4. Auflage, Springer, 2003.

[12] H. Kowalsky, G. Michler: Lineare Algebra, 12. Auflage, de Gruyter, 2003.

[13] S. Lang: Linear algebra, 3. edition, Springer, 1996.

[14] F. Lorenz: Lineare Algebra, 2 Bände, 3. Auflage, BI Verlag, 1992.

[15] H. v. Mangoldt, K. Knopp: Höhere Mathematik: eine Einführung für Studierende und zum Selbststudium, Band 1, Hirzel-Verlag, 1990.

[16] G. Stroth: Lineare Algebra, 2. Auflage, Berliner Studienreihe zur Mathematik 7, Heldermann, 2008.

[17] H. Zieschang: Lineare Algebra und Geometrie, Teubner, 1997.