# Basic linear algebra

Friedrich-Schiller-Universität Jena, SS 2016

Jürgen Müller

# Contents

## 1 Linear equations

**(1.1) Example: Linear optimization.** A peasant has got an area of 40ha = 400.000m$^2$ of farm land, and a cow shed 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 maximize 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**:

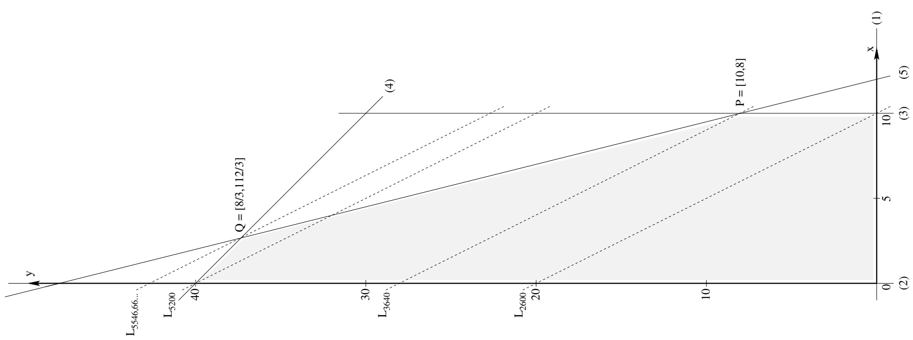
$$\left\{ \begin{array}{rcrcr} x & + & y & = & 40 \\ 200x & + & 50y & = & 2400 \end{array} \right.$$

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 maximize 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 $\varphi$ 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.

**(1.2) 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 are various ways to do so, where these descriptions are by no means unique,

Table 1: Geometric picture of constraints and cost function.




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 **parameterized 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 $[u,v]$ describes 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 adding two 'point vectors', respectively.

For example, the lines in (1.1) 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) | $4x+y=48$ | $[12,0]+t\cdot[1,-4]$ | $[12,0],[0,48]$ |
| $\varphi$ | $260x+130y=c$ | $[0,\frac{c}{130}]+t\cdot[1,-2]$ | $[\frac{c}{260},0],[0,\frac{c}{130}]$ |

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 parameterizations $\{[0,48]+t\cdot[1,-4] \in \mathbb{R}^2; t \in \mathbb{R}\}$, 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 $4x+y=48$.

**(1.3) 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** $v:=[x_1,\ldots,x_n]^{\mathrm{tr}}\in\mathbb{R}^{n\times 1}$ and $w:=[y_1,\ldots,y_m]^{\mathrm{tr}}\in\mathbb{R}^{m\times 1}$, 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** $a_{ij}\in\mathbb{R}$, 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 $[\sum_{j=1}^n a_{1j}x_j,\ldots,\sum_{j=1}^n a_{mj}x_j]^{\mathrm{tr}}\in\mathbb{R}^{m\times 1}$. The set of **solutions** is defined as $\mathcal{L}(A,w):=\{v\in\mathbb{R}^{n\times 1}; Av=w\}$; 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 equation $i$ with $0\neq a\in\mathbb{R}$, adding the $a$-fold of equation $j$ to equation $i\neq j$, for $a\in\mathbb{R}$, and interchanging equations $i$ and $j$, translates into **row operations** on $[A|w]$, namely multiplying row $i$ entrywise with $0\neq a\in\mathbb{R}$, adding entrywise the $a$-fold of row $j$ to row $i$, and interchanging rows $i$ and $j$, respectively. Since we are replacing equations by consequences of given equations, the set of solutions might become larger, but since all steps are invertible, 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 (1.1); 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 $A\cdot[x,y]^{\mathrm{tr}}=w$, 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 $\mathcal{L}(A,w)=\{[0,0,6,-1]^{\text{tr}}+t\cdot[2,1,-6,0]^{\text{tr}}\in\mathbb{R}^{4\times 1};t\in\mathbb{R}\}$, a line in $\mathbb{R}^{4\times 1}$.

**(1.4) Theorem: Gauß algorithm.** Let $A \in \mathbb{R}^{m\times n}$, where $m, n \in \mathbb{N}$. Then using row operations $A$ can be transformed into **Gaussian normal form**

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

having $r = r(A) \in \{0, \ldots, \min\{m, n\}\}$ non-zero rows, being called the **Gauß rank** of $A$. The '1's occur in the uniquely determined **pivot columns** $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ of $A$; hence the Gaussian normal form is uniquely determined.

**Proof.** Here, we only show existence, being sufficient for the problem of solving systems of linear equations, and postpone the question of uniqueness to (3.8):

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 $A$. We may assume that $\widetilde{A} \neq 0$, 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 $k \neq i \in \{1, \ldots, m\}$. ♯

**(1.5) Solving linear equations. a)** Let $A = [a_{ij}]_{ij} \in \mathbb{R}^{m\times n}$, where $m, n \in \mathbb{N}$, and $w = [y_1, \ldots, y_m]^{\mathrm{tr}} \in \mathbb{R}^{m\times 1}$. We consider the system of linear equations in the indeterminates $x_1, \ldots, x_n \in \mathbb{R}$ given by $A \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = w$; the latter is called **homogeneous** if $w = 0$, otherwise it is called **inhomogeneous**. Hence the set of solutions is $\mathcal{L}(A, w) := \{v \in \mathbb{R}^{n\times 1}; Av = w\}$, where the homogeneous system associated with $A$ we also write $\mathcal{L}(A) := \mathcal{L}(A, 0)$.

Since for $0 \in \mathbb{R}^{n\times 1}$ we have $A \cdot 0 = 0 \in \mathbb{R}^{m\times 1}$, we infer $0 \in \mathcal{L}(A)$, in particular any homogeneous system is solvable. Moreover, it is immediate from the explicit description $A \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = [\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j]^{\mathrm{tr}} \in \mathbb{R}^{m\times 1}$ of the matrix product, for all $[x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathbb{R}^{n\times 1}$, that we have $A(v+v') = Av + Av'$ and $A(av) = a \cdot Av$, for all $v, v' \in \mathbb{R}^{n\times 1}$ and $a \in \mathbb{R}$, where addition of tuples and taking scalar multiples is performed entrywise. Hence we infer that $\mathcal{L}(A)$ is closed with respect to these operations.

Moreover, if $v, v' \in \mathcal{L}(A, w)$ then we have $A(v - v') = w - w = 0$, that is $v - v' \in \mathcal{L}(A)$; conversely, for all $u \in \mathcal{L}(A)$ 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 $\mathcal{L}(A, w) = v_0 + \mathcal{L}(A) := \{v_0 + u \in \mathbb{R}^{n\times 1}; u \in \mathcal{L}(A)\}$, where $v_0 \in \mathcal{L}(A, w)$ is a fixed **particular solution**.

Hence solving the system amounts to prove solvability, and in this case to find a particular solution and the solutions $\mathcal{L}(A)$ of the associated homogeneous system. To decide solvability, we apply the Gauß algorithm to the extended matrix $[A|w] \in \mathbb{R}^{m\times(n+1)}$, which shows that $\mathcal{L}(A,w) \neq \emptyset$ if and only if $n+1$ is not a pivot column, which is equivalent to saying that for the associated Gauß ranks we have $r(A) = r([A|w])$. In this case the solutions $v = [x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathcal{L}(A,w)$ are described as follows:

**b)** Let $[A|w]$ have Gaussian normal form $[A'|w'] \in \mathbb{R}^{m\times(n+1)}$ with pivot columns $1 \leq j_1 < \cdots < j_r \leq n$, where $r := r(A) \in \mathbb{N}_0$. By the above analysis we have $\mathcal{L}(A,w) = \{v \in \mathbb{R}^{n\times 1}; Av = w\} = \{v \in \mathbb{R}^{n\times 1}; A'v = w'\} = \mathcal{L}(A', w')$.

The $n-r$ entries $x_j \in \mathbb{R}$, where $j \in \mathcal{J} := \{1,\ldots,n\} \setminus \{j_1,\ldots,j_r\}$ are the non-pivot columns, can be chosen arbitrarily, thus are considered as **parameters**. Then the $r$ entries $x_{j_k} \in \mathbb{R}$, for the pivot columns $\{j_1,\ldots,j_r\}$, are uniquely determined by $x_{j_k} := y'_k - \sum_{j\in\mathcal{J}} a'_{kj}x_j$, for $k \in \{1,\ldots,r\}$.

To facilitate the explicit description of solutions, for $j \in \{1,\ldots,n\}$ let $e_j := [0,\ldots,0,1,0,\ldots,0]^{\mathrm{tr}} \in \mathbb{R}^{n\times 1}$ be the **unit tuple** with entry $j$ non-zero; then we have $[x_1,\ldots,x_n]^{\mathrm{tr}} = \sum_{j=1}^n x_je_j$. Now, the particular solution given by letting $x_j := 0$, for all $j \in \mathcal{J}$, equals $v_0 := \sum_{k=1}^r y'_k e_{j_k} \in \mathbb{R}^{n\times 1}$.

Considering the homogeneous system, specifying the parameters to be unit tuples on $\mathcal{J}$, we get the **basic solutions** $v_j := e_j - \sum_{k=1}^r a'_{kj}e_{j_k} \in \mathbb{R}^{n\times 1}$, for $j \in \mathcal{J}$, and thus $\mathcal{L}(A) = \{\sum_{j\in\mathcal{J}} x_jv_j \in \mathbb{R}^{n\times 1}; x_j \in \mathbb{R} \text{ for } j \in \mathcal{J}\}$, hence $\mathcal{L}(A,w) = v_0 + \mathcal{L}(A)$. Note that any element of $\mathcal{L}(A)$ can be written uniquely as an **$\mathbb{R}$-linear combination** of the basic solutions $\{v_j \in \mathbb{R}^{n\times 1}; j \in \mathcal{J}\}$.

**c)** From a more general point of view we make the following observations: We have $\mathcal{L}(A,w) \neq \emptyset$ for all $w \in \mathbb{R}^{m\times 1}$ if and only if $r = m$; hence in this case we have $m = r \leq n$. Moreover, if $\mathcal{L}(A,w) \neq \emptyset$ for some $w \in \mathbb{R}^{m\times 1}$, then we have $|\mathcal{L}(A,w)| = 1$ if and only if $|\mathcal{L}(A)| = 1$, which holds if and only if $n - r = 0$; hence in this case we have $m \geq r = n$.

Thus in the case $m = n$ of quadratic systems the following are equivalent:
**i)** We have $r = n$, that is the Gaussian normal form is the **identity matrix**.
**ii)** We have $|\mathcal{L}(A)| = 1$, that is $\mathcal{L}(A) = \{0\}$.
**iii)** There is $w \in \mathbb{R}^{n\times 1}$ such that $|\mathcal{L}(A,w)| = 1$.
**iv)** For all $w \in \mathbb{R}^{n\times 1}$ we have $|\mathcal{L}(A,w)| = 1$.
**v)** For all $w \in \mathbb{R}^{n\times 1}$ we have $\mathcal{L}(A,w) \neq \emptyset$.

**d)** Note that in practice it is usually not necessary to compute the Gaussian normal form in order to solve a system of linear equations, but an intermediate step typically is sufficient, as long as it already exhibits an **echelon form**; such a form is also indicated in example (i) below. Here are a couple of examples; recall also those in (1.1) and (1.3), the former being of the type described in (c):

**i)** The Gauß algorithm yields:

$$[A|w|\widetilde{w}] := \left[\begin{array}{ccccc|c|c} 1 & 0 & 4 & 0 & 1 & 0 & 1 \\ 2 & 1 & 0 & 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 & 2 & 2 \\ 6 & 2 & 5 & 3 & 3 & 3 & 3 \\ 7 & 2 & 12 & 4 & 5 & 2 & 2 \end{array}\right]$$

$$\mapsto \left[\begin{array}{ccccc|c|c} 1 & . & 4 & . & 1 & . & 1 \\ . & 1 & -8 & 2 & -1 & 1 & -1 \\ . & . & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & . \\ . & . & . & . & . & . & 1 \\ . & . & . & . & . & . & . \end{array}\right] \mapsto \left[\begin{array}{ccccc|c|c} 1 & . & . & -\frac{4}{3} & -\frac{1}{3} & \frac{4}{3} & 1 \\ . & 1 & . & \frac{14}{3} & \frac{5}{3} & -\frac{5}{3} & -1 \\ . & . & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & . \\ . & . & . & . & . & . & 1 \\ . & . & . & . & . & . & . \end{array}\right]$$

Hence we have $\mathcal{L}(A,\widetilde{w}) = \emptyset$, while $\mathcal{L}(A,w) \neq \emptyset$. The latter is given as $\mathcal{L}(A,w) = v_0 + \mathcal{L}(A)$, where a particular solution can be chosen as $v_0 := [\frac{4}{3}, -\frac{5}{3}, -\frac{1}{3}, 0, 0]^{\mathrm{tr}} \in \mathbb{R}^{5\times 1}$, while the associated homogeneous system has the set of solutions $\mathcal{L}(A) = \{sv_4 + tv_5 \in \mathbb{R}^{5\times 1}; s,t \in \mathbb{R}\}$, where in turn the basic solutions are given as $v_4 := [\frac{4}{3}, -\frac{14}{3}, -\frac{1}{3}, 1, 0]^{\mathrm{tr}} \in \mathbb{R}^{5\times 1}$ and $v_5 := [\frac{1}{3}, -\frac{5}{3}, -\frac{1}{3}, 0, 1]^{\mathrm{tr}} \in \mathbb{R}^{5\times 1}$.

**ii)** The Gauß algorithm yields for generic right hand side $w := [a,b,c]^{\mathrm{tr}} \in \mathbb{R}^{3\times 1}$; we will come back to this point of view in (1.7):

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

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

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

**(1.6) Example: Geometric interpretation.** For $n \in \{2,3\}$ 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 $r := r(A) \in \{0,\ldots,n\}$. 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 $\mathcal{L} = \mathbb{R}^{n\times 1}$. 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 $\mathcal{L} = \{[a - sy, y]^{\rm tr} = [a, 0]^{\rm tr} + y \cdot [-s, 1]^{\rm tr} \in \mathbb{R}^{2\times 1}; y \in \mathbb{R}\}$, a line in $\mathbb{R}^{2\times 1}$; it intersects the first coordinate axis in $[a, 0]^{\rm tr}$, and the second coordinate axis in $[0, \frac{a}{s}]^{\rm tr}$ if $s \neq 0$, while it is parallel to it if $s = 0$, coinciding with it if and only if $a = 0$. In the second case we have $\mathcal{L} = \{[x, a]^{\rm tr} = [0, a]^{\rm tr} + x \cdot [1, 0]^{\rm tr} \in \mathbb{R}^{2\times 1}; x \in \mathbb{R}\}$, a line in $\mathbb{R}^{2\times 1}$; it intersects the second coordinate axis in $[0, a]^{\rm tr}$, and is parallel to the first coordinate axis, coinciding 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 $\mathcal{L} = \{[a - sz, b - tz, z]^{\rm tr} = [a, b, 0]^{\rm tr} + z \cdot [-s, -t, 1]^{\rm tr} \in \mathbb{R}^{3\times 1}; z \in \mathbb{R}\}$ and $\mathcal{L} = \{[a - sy, y, b]^{\rm tr} = [a, 0, b]^{\rm tr} + y \cdot [-s, 1, 0]^{\rm tr} \in \mathbb{R}^{3\times 1}; y \in \mathbb{R}\}$ and $\mathcal{L} = \{[x, a, b]^{\rm tr} = [0, a, b]^{\rm tr} + x \cdot [1, 0, 0]^{\rm tr} \in \mathbb{R}^{3\times 1}; x \in \mathbb{R}\}$, 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 $\mathcal{L} = \{[a - sy - tz, y, z]^{\rm tr} = [a, 0, 0]^{\rm tr} + y \cdot [-s, 1, 0]^{\rm tr} + z \cdot [-t, 0, 1]^{\rm tr} \in \mathbb{R}^{3\times 1}; y, z \in \mathbb{R}\}$ and $\mathcal{L} = \{[x, a - zs, z]^{\rm tr} = [0, a, 0]^{\rm tr} + x \cdot [1, 0, 0]^{\rm tr} + z \cdot [0, -s, 1]^{\rm tr} \in \mathbb{R}^{3\times 1}; x, z \in \mathbb{R}\}$ and $\mathcal{L} = \{[x, y, a]^{\rm tr} = [0, 0, a]^{\rm tr} + x \cdot [1, 0, 0]^{\rm tr} + y \cdot [0, 1, 0]^{\rm tr} \in \mathbb{R}^{3\times 1}; x, y \in \mathbb{R}\}$, respectively, hyperplanes in $\mathbb{R}^{3\times 1}$.

**(1.7) Inversion.** We again come back to the system of linear equations in (1.1), given by $A := \begin{bmatrix} 1 & 1 \\ 200 & 50 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, which is of the type described in (1.5)(c). We have seen how to solve the specific system $A \cdot [x, y]^{\rm tr} = [40, 2400]^{\rm tr}$. Now any system with the same matrix $A$ and some right hand side $w = [c, d]^{\rm tr} \in \mathbb{R}^{2\times 1}$ has a unique solution, and it is immediate that we can similarly solve it, but at the cost of doing the explicit computation for any $w$ separately. Hence we are tempted to do this only once, for a generic right hand side, that is for the system $A \cdot [x, y]^{\rm tr} = [c, d]^{\rm tr}$, where now $c, d \in \mathbb{R}$ are indeterminates. Note that example (ii) in (1.5) already points in the same direction.

To see how $\mathcal{L}(A, [c, d]^{\rm tr})$ 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 $\mathcal{L}(A,[c,d]^{\mathrm{tr}})$ is the 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 recover $B\cdot[40,2400]^{\mathrm{tr}} = \frac{1}{150}\cdot[400,5600]^{\mathrm{tr}} = \frac{1}{3}\cdot[8,112]^{\mathrm{tr}}$. From a formal point of view, for the maps $\alpha\colon \mathbb{R}^{2\times 1}\to\mathbb{R}^{2\times 1}\colon [x,y]^{\mathrm{tr}} \mapsto A\cdot[x,y]^{\mathrm{tr}}$ and $\beta\colon \mathbb{R}^{2\times 1}\to\mathbb{R}^{2\times 1}\colon [c,d]^{\mathrm{tr}} \mapsto B\cdot[c,d]^{\mathrm{tr}}$ we have shown that $\alpha\beta = \mathrm{id}$; indeed we can check that $\beta\alpha = \mathrm{id}$ as well, hence $\alpha$ is actually bijective with inverse $\beta$.

Since the answer is given in terms of a matrix product, we reformulate this again: Let $v_i \in \mathbb{R}^{2\times 1}$ be the unique tuples such that $Av_i = e_i$, for $i \in \{1,2\}$, for the unit tuple $e_i \in \mathbb{R}^{2\times 1}$. Then we get $A(cv_1+dv_2) = ce_1+de_2 = [c,d]^{\mathrm{tr}}$, saying that the solution of the system with right hand side $[c,d]^{\mathrm{tr}}$ is given as an $\mathbb{R}$-linear combination of the solutions of the systems with right hand side $e_i$. Hence, 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 the identity matrix $E_2 \in \mathbb{R}^{2\times 2}$:

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

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

Thus we have $v_1 = [-\frac{1}{3},\frac{4}{3}]^{\mathrm{tr}}$ and $v_2 = [\frac{1}{150},-\frac{1}{150}]^{\mathrm{tr}}$, and the inverse matrix $B \in \mathbb{R}^{2\times 2}$ is obtained as the concatenation of the columns $v_i \in \mathbb{R}^{2\times 1}$.

**(1.8) 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

$$\mathbb{R}^{2\times 1}\to\mathbb{R}^{2\times 1}\colon \begin{bmatrix} a \\ b \end{bmatrix} \mapsto A\cdot\left(\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right), \quad \text{where} \quad A := \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2};$$

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

Since encryption essentially is the matrix product with $A$, decryption amounts to solving the system of linear equations with coefficient matrix $A$, for various right hand sides. Hence decryption is given as

$$\mathbb{R}^{2\times 1}\to\mathbb{R}^{2\times 1}\colon \begin{bmatrix} c \\ d \end{bmatrix} \mapsto \left(B\cdot\begin{bmatrix} c \\ d \end{bmatrix}\right) - \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{where} \quad B := \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

The latter 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

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

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'.

## 2 Vector spaces

**(2.1) Groups. a)** A set $A$ together with an **addition** $+\colon A \times A \to A$ fulfilling the following conditions is called a **commutative group**: We have **commutativity** $a+b = b+a$ for all $a, b \in A$; we have **associativity** $(a+b)+c = a+(b+c)$ for all $a, b, c \in A$; there is a **neutral element** $0 \in A$ such that $a + 0 = a$ for all $a \in A$; and for any $a \in A$ there is an **inverse** $-a \in A$ such that $a + (-a) = 0$.

In particular, we have $A \neq \emptyset$. For all $a_1, \ldots, a_n \in A$, where $n \in \mathbb{N}$, the sum $a_1 + a_2 + \cdots + a_n \in A$ is well-defined, independently from the bracketing and the order of the summands. Moreover, the neutral element and inverses are uniquely defined: If $0' \in A$ is a neutral element, then we have $0' = 0' + 0 = 0$; and if $a' \in A$ is an inverse of $a \in A$, then we have $a' = a' + 0 = a' + a + (-a) = 0 + (-a) = -a$.

**b)** For example, letting $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, the set $R$ becomes an **additive** commutative group, with neutral element $0 \in R$ and inverses given by $-a \in R$, for $a \in R$. Moreover, letting $K \in \{\mathbb{Q}, \mathbb{R}\}$, the set $K^* := K \setminus \{0\}$ becomes a **multiplicative** commutative group with neutral element $1 \in K^*$, and inverses given by $a^{-1} = \frac{1}{a} \in K^*$, for $a \in K^*$. Actually, $K$ becomes a **field**, inasmuch we also have **distributivity** $a(b + c) = ab + ac$, for all $a, b, c \in K$.

Note that just from these rules we get $0 \cdot a = 0$ and $(-1) \cdot a = -a$, for all $a \in K$: From $0 + 0 = 0$ we get $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$, and hence $0 = 0 \cdot a - (0 \cdot a) = (0 \cdot a + 0 \cdot a) - (0 \cdot a) = 0 \cdot a$; and we have $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1 + 1) \cdot a = 0 \cdot a = 0$.

**(2.2) Vector spaces. a)** An additive commutative group $V$, together with a **scalar multiplication** $\cdot\colon \mathbb{R} \times V \to V$ fulfilling the following conditions is called an $\mathbb{R}$**-vector space**, its elements being called **vectors**: We have **unitarity** $1 \cdot v = v$ and $\mathbb{R}$**-linearity** $a(v+w) = av+aw$, as well as distributivity $(a+b)v = av + bv$ and $(ab)v = a(bv)$, for all $v, w \in V$ and $a, b \in \mathbb{R}$.

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 $a \in \mathbb{R}$; note that we write both $0 \in \mathbb{R}$ and $0 \in V$: We have $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, hence $a \cdot 0 = 0$; as well as $0 \cdot v + 0 \cdot v = (0+0) \cdot v = 0 \cdot v$, 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, $av = 0 \in V$ implies $v = 0 \in V$ or $a = 0 \in \mathbb{R}$: If $a \neq 0$, then we have $0 = a^{-1} \cdot 0 = a^{-1}(av) = (a^{-1}a)v = 1 \cdot v = v$.

**b)** Let $\mathcal{I}$ be a set. Then $\mathrm{Maps}(\mathcal{I}, \mathbb{R})$ is an $\mathbb{R}$-vector space with **pointwise** addition $f + g \colon \mathcal{I} \to \mathbb{R} \colon x \mapsto f(x) + g(x)$ and scalar multiplication $af \colon \mathcal{I} \to \mathbb{R} \colon x \mapsto a \cdot f(x)$, for $f, g \in \mathrm{Maps}(\mathcal{I}, \mathbb{R})$ and $a \in \mathbb{R}$. For example, for $\mathcal{I} = \emptyset$ we get the **zero space** $\{0\}$; for $\mathcal{I} = \mathbb{R}$ and $\mathcal{I} = \mathbb{N}$ we get the $\mathbb{R}$-vector spaces $\mathrm{Maps}(\mathbb{R}, \mathbb{R})$ of all real-valued functions on the real numbers, and $\mathrm{Maps}(\mathbb{N}, \mathbb{R})$ of all sequences of real numbers, respectively. Moreover, here are the most important examples:

If $\mathcal{I} = \{1, \ldots, n\}$, where $n \in \mathbb{N}$, the bijection $\mathrm{Maps}(\{1, \ldots, n\}, \mathbb{R}) \to \mathbb{R}^n \colon f \mapsto [f(1), \ldots, f(n)]$ shows that the **row space** $\mathbb{R}^n$ of all of $n$-tuples $[a_1, \ldots, a_n]$ with entries $a_i \in \mathbb{R}$, becomes an $\mathbb{R}$-vector space with respect to **componentwise** addition $[a_1, \ldots, a_n] + [b_1, \ldots, b_n] := [a_1 + b_1, \ldots, a_n + b_n]$ and scalar multiplication $a \cdot [a_1, \ldots, a_n] := [aa_1, \ldots, aa_n]$, for all $[a_1, \ldots, a_n], [b_1, \ldots, b_n] \in \mathbb{R}^n$ and $a \in \mathbb{R}$. In particular, for $n = 1$ we recover $\mathbb{R} = \mathbb{R}^1$, where scalar multiplication is just given by left multiplication; and for $n = 0$ we let $\mathbb{R}^0 := \{0\}$.

If $\mathcal{I} = \{1, \ldots, m\} \times \{1, \ldots, n\}$, where $m, n \in \mathbb{N}$, we get the $\mathbb{R}$-vector space $\mathbb{R}^{m \times n}$ of all $(m \times n)$**-matrices** $[a_{ij}]_{ij}$ with entries $a_{ij} \in \mathbb{R}$; if $m = 0$ or $n = 0$ we let $\mathbb{R}^{m \times n} := \{0\}$. In particular, for $m = 1$ we recover the row space $\mathbb{R}^{1 \times n} = \mathbb{R}^n$, and for $n = 1$ we get the **column space** $\mathbb{R}^{n \times 1}$ consisting of all column $n$-tuples $[a_1, \ldots, a_n]^{\mathrm{tr}}$ with entries $a_i \in \mathbb{R}$.

**(2.3) Subspaces.** Let $V$ be an $\mathbb{R}$-vector space. A subset $\emptyset \neq U \subseteq V$ is called a $\mathbb{R}$**-subspace**, if addition and scalar multiplication restrict to maps $+ \colon U \times U \to U$ and $\cdot \colon \mathbb{R} \times U \to U$, respectively.

From $0 \cdot v = 0 \in V$ and $-v = -1 \cdot v \in V$, for all $v \in V$, we conclude that $0 \in U$, and that $U$ is closed with respect to taking inverses, hence $U$ again is a commutative group, and thus again is an $\mathbb{R}$-vector space; we write $U \leq V$. For example, we have $\{0\} \leq V$ and $V \leq V$.

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.

But the prototypical example is as follows: Given $A = [a_{ij}]_{ij} \in \mathbb{R}^{m\times n}$, where $m, n \in \mathbb{N}_0$, we consider the homogeneous system of linear equations in the indeterminates $x_1, \ldots, x_n \in \mathbb{R}$ given by $A \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = 0$. Then the solutions $\mathcal{L}(A) := \{v \in \mathbb{R}^{n\times 1}; Av = 0 \in \mathbb{R}^{m\times 1}\}$ form an $\mathbb{R}$-subspace of $\mathbb{R}^{n\times 1}$: We have $0 \in \mathcal{L}(A)$, and for $v, v' \in \mathcal{L}(A) \subseteq \mathbb{R}^{n\times 1}$ and $a \in \mathbb{R}$ we have $A(v+v') = Av + Av' = 0$ and $A(av) = a \cdot Av = 0$, hence $v + v' \in \mathcal{L}(A)$ and $av \in \mathcal{L}(A)$; see (1.5).

**(2.4) Linear combinations. a)** Let $V$ be an $\mathbb{R}$-vector space. For any subset $S \subseteq V$ let $\langle S\rangle_{\mathbb{R}} := \{\sum_{i=1}^k a_i v_i \in V; k \in \mathbb{N}_0, a_i \in \mathbb{R}, v_i \in S \text{ for all } i \in \{1, \ldots, k\}\}$, where the finite sums $\sum_{i=1}^k a_i v_i \in V$ are called **$\mathbb{R}$-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, where $n \in \mathbb{N}_0$, we also write $\langle S\rangle_{\mathbb{R}} = \langle v_1, \ldots, v_n\rangle_{\mathbb{R}} = \{\sum_{i=1}^n a_i v_i \in V; a_i \in \mathbb{R} \text{ for all } i \in \{1, \ldots, n\}\}$; in particular, for $S = \emptyset$ and $S = \{v\}$ we have $\langle\emptyset\rangle_{\mathbb{R}} = \{0\}$ and $\langle v\rangle_{\mathbb{R}} = \{av \in V; a \in \mathbb{R}\}$, respectively.

Then we have $S \subseteq \langle S\rangle_{\mathbb{R}} \leq V$, and since any $\mathbb{R}$-subspace of $V$ containing $S$ also contains $\langle S\rangle_{\mathbb{R}}$, we conclude that $\langle S\rangle_{\mathbb{R}}$ is the smallest $\mathbb{R}$-subspace of $V$ containing $S$. Hence $\langle S\rangle_{\mathbb{R}}$ is called the $\mathbb{R}$-subspace of $V$ **generated** by $S$. If $\langle S\rangle_{\mathbb{R}} = V$, then $S \subseteq V$ is called an **$\mathbb{R}$-generating set** of $V$; if $V$ has a finite $\mathbb{R}$-generating set then $V$ is called **finitely generated**.

**b)** Let $U, U' \leq V$. Then $U \cap U' \leq V$ is an $\mathbb{R}$-subspace of $V$ again. But we have $U \cup U' \leq V$ if and only if $U \leq U'$ or $U' \leq U$ holds:

For $U \leq U'$ we have $U \cup U' = U' \leq V$, while for $U' \leq U$ we have $U \cup U' = U \leq V$. Assume that $U \cup U' \leq V$, where $U \not\subseteq U'$ and $U' \not\subseteq U$. Then 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. $\sharp$

The problem, that $U \cup U'$ in general is not an $\mathbb{R}$-subspace again, is remedied by going over to the $\mathbb{R}$-subspace generated by $U \cup U'$: The set $U + U' := \{u + u' \in V; u \in U, u' \in U'\} = \langle U \cup U'\rangle_{\mathbb{R}} \leq V$ is called the **sum** of $U$ and $U'$. Note that hence for subsets $S, S' \subseteq V$ we have $\langle S\rangle_{\mathbb{R}} + \langle S'\rangle_{\mathbb{R}} = \langle S \cup S'\rangle_{\mathbb{R}}$; in particular, the sum of finitely generated $\mathbb{R}$-subspaces is finitely generated again.

**(2.5) Linear independence.** Let $V$ be an $\mathbb{R}$-vector space. Then a sequence $\mathcal{S} := [v_i \in V; i \in \mathcal{I}]$, where $\mathcal{I}$ is a set, is called **$\mathbb{R}$-linearly independent**, if for all finite subsets $\mathcal{J} \subseteq \mathcal{I}$ and for all sequences $[a_j \in \mathbb{R}; j \in \mathcal{J}]$ 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 **$\mathbb{R}$-linearly dependent**. A subset $S \subseteq V$ is called **$\mathbb{R}$-linearly independent**, if $[f(i) \in V; i \in \mathcal{I}]$ is $\mathbb{R}$-linearly independent for some bijection $f : \mathcal{I} \to S$.

If $\mathcal{S}$ is $\mathbb{R}$-linearly independent, then $[v_j \in V; j \in \mathcal{J}]$ is as well, for any $\mathcal{J} \subseteq \mathcal{I}$. If $v_i = 0$ for some $i \in \mathcal{I}$, or if $v_i = v_j$ for some $i \neq j \in \mathcal{I}$, then $\mathcal{S}$ is $\mathbb{R}$-linearly dependent. The empty set $S = \emptyset$ is $\mathbb{R}$-linearly independent. Since $av = 0$, where $v \in V$ and $a \in \mathbb{R}$, implies $a = 0$ or $v = 0$, we infer that the singleton set $S = \{v\}$, where $v \neq 0$, is $\mathbb{R}$-linearly independent.

For example, $[f_k \in \mathrm{Maps}(\mathbb{R},\mathbb{R}); k \in \mathbb{N}]$ where $f_k \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \exp(kx)$ is $\mathbb{R}$-linearly independent:

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 $a_n = 0$. $\sharp$

**(2.6) Bases. a)** Let $V$ be an $\mathbb{R}$-vector space. Then an $\mathbb{R}$-linearly independent $\mathbb{R}$-generating set $B \subseteq V$ is called a **$\mathbb{R}$-basis** of $V$.

Given an $\mathbb{R}$-basis $B = \{v_1, \ldots, v_n\} \subseteq V$, where $n \in \mathbb{N}_0$, we have the principle of **comparison of coefficients**: Given $v \in V$, if $a_1, \ldots, a_n \in \mathbb{R}$ and $a_1', \ldots, a_n' \in \mathbb{R}$ fulfill $v = \sum_{i=1}^n a_i v_i = \sum_{i=1}^n a_i' v_i \in V$, then $0 = v - v = \sum_{i=1}^n (a_i - a_i') \cdot v_i$, which by $\mathbb{R}$-linear independence implies $a_i = a_i'$ for all $i \in \{1, \ldots, n\}$.

Since $B$ is an $\mathbb{R}$-generating set of $V$, any $v \in V$ is an $\mathbb{R}$-linear combination of $B$, that is there are $a_1, \ldots, a_n \in \mathbb{R}$ such that $v = \sum_{i=1}^n a_i v_i \in V$. Thus there is a unique representation $v = \sum_{i=1}^n a_i v_i$, leading to the **coordinate vectors** $v_B := [a_1, \ldots, a_n] \in \mathbb{R}^n$ and ${}_B v := [a_1, \ldots, a_n]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}$.

**b)** Here are the prototypical examples: Firstly, we consider $V = \mathbb{R}^{n \times 1}$, where $n \in \mathbb{N}_0$: For $i \in \{1, \ldots, n\}$ let $e_i := [0, \ldots, 0, 1, 0, \ldots, 0]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}$ be the **unit vector** with entry $i$ non-zero. Then we have $[a_1, \ldots, a_n]^{\mathrm{tr}} = \sum_{i=1}^n a_i e_i \in \mathbb{R}^{n \times 1}$, and thus $\langle e_1, \ldots, e_n \rangle_{\mathbb{R}} = \mathbb{R}^{n \times 1}$. From $\sum_{i=1}^n a_i e_i = [a_1, \ldots, a_n]^{\mathrm{tr}} = 0 \in \mathbb{R}^{n \times 1}$ we get $a_i = 0 \in \mathbb{R}$ for all $i \in \{1, \ldots, n\}$, hence $S := \{e_1, \ldots, e_n\} \subseteq \mathbb{R}^{n \times 1}$ is $\mathbb{R}$-linearly independent. Thus $S \subseteq \mathbb{R}^{n \times 1}$ is an $\mathbb{R}$-basis, being called its **standard** $\mathbb{R}$-basis; the empty set $\emptyset$ is the only $\mathbb{R}$-basis of $\mathbb{R}^{0 \times 1} = \{0\}$. Note that it is a particular feature of the standard basis that any vector coincides with its coordinate vector, that is ${}_S([a_1, \ldots, a_n]^{\mathrm{tr}}) = [a_1, \ldots, a_n]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}$.

Similarly, for $m, n \in \mathbb{N}_0$, let the **matrix unit** $E_{ij} = [a_{kl}]_{kl} \in \mathbb{R}^{m \times n}$ 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 $\{E_{ij} \in \mathbb{R}^{m \times n}; i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \subseteq \mathbb{R}^{m \times n}$ is an $\mathbb{R}$-basis, being called its **standard** $\mathbb{R}$-basis. Note that in particular we get the **diagonal matrix** $\mathrm{diag}[a_1, \ldots, a_n] := \sum_{i=1}^n a_i E_{ii} \in \mathbb{R}^{n \times n}$, where $a_i \in \mathbb{R}$ for $i \in \{1, \ldots, n\}$, and the **identity matrix** $E_n = \mathrm{diag}[1, \ldots, 1] = \sum_{i=1}^n E_{ii} \in \mathbb{R}^{n \times n}$.

Secondly, we consider the homogeneous system of linear equations associated with a matrix $A = [a_{ij}]_{ij} \in \mathbb{R}^{m \times n}$, where $m, n \in \mathbb{N}_0$: Letting $\mathcal{J} \subseteq \{1, \ldots, n\}$ be the set of non-pivot columns, any solution $v \in \mathcal{L}(A)$ can be written uniquely as an $\mathbb{R}$-linear combination of the basic solutions $\{v_j \in \mathbb{R}^{n \times 1}; j \in \mathcal{J}\}$, see (1.5), implying that the latter is an $\mathbb{R}$-basis of the $\mathbb{R}$-subspace $\mathcal{L}(A) \leq \mathbb{R}^{n \times 1}$.

**c)** From an abstract point of view, standard bases cannot at all be distinguished from other bases; for example: We consider $V = \mathbb{R}^{2 \times 1}$ and let $B :=$

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

**(2.7) Theorem: Characterization of bases.** Let $V$ be an $\mathbb{R}$-vector space. Then for any subset $B \subseteq V$ the following are equivalent:
**i)** $B$ is an $\mathbb{R}$-basis of $V$.
**ii)** $B$ is a maximal $\mathbb{R}$-linearly independent subset, that is $B$ is $\mathbb{R}$-linearly independent and any subset $B \subset S \subseteq V$ is $\mathbb{R}$-linearly dependent.
**iii)** $B$ is a minimal $\mathbb{R}$-generating set, that is $B$ is an $\mathbb{R}$-generating set and for any subset $T \subset B$ we have $\langle T\rangle_{\mathbb{R}} < V$.

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

**ii)$\Rightarrow$iii)**: We show that $B$ is an $\mathbb{R}$-generating set: Let $v \in V$, and since $B \subseteq \langle B\rangle_{\mathbb{R}}$ we may assume that $v \notin B$. Then $B \,\dot\cup\, \{v\} \subseteq V$ is $\mathbb{R}$-linearly dependent, thus there are $v_1, \ldots, v_n \in B$ and $a, a_1, \ldots, a_n \in \mathbb{R}$, 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 $\mathbb{R}$-linearly dependent, a contradiction. Thus we have $a \neq 0$, and hence $v = a^{-1}(av) = -a^{-1} \cdot \sum_{i=1}^n a_i v_i \in \langle B\rangle_{\mathbb{R}}$.

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

**iii)$\Rightarrow$i)**: We show that $B$ is $\mathbb{R}$-linearly independent: Assume that there are $v_1, \ldots, v_n \in B$ and $a_1, \ldots, a_n \in \mathbb{R}$, 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 $m \in \mathbb{N}$ such that $m \geq n$, and $b_1, \ldots, b_m \in \mathbb{R}$ 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 $\langle B \setminus \{v_1\}\rangle_{\mathbb{R}} = V$, a contradiction. ♯

**(2.8) Theorem: Steinitz's base change theorem.** Let $V$ be an $\mathbb{R}$-vector space with finite $\mathbb{R}$-basis $B$, and let $S \subseteq V$ be $\mathbb{R}$-linearly independent. Then we have $|S| \leq |B|$, and there is $T \subseteq B \setminus S$ such that $S \,\dot\cup\, T$ is an $\mathbb{R}$-basis of $V$; note that necessarily $|S| + |T| = |B|$.

**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 $m$, where the assertion is trivial for $m = 0$, hence let $m \geq 1$.

The set $\{w_1, \ldots, w_{m-1}\}$ is $\mathbb{R}$-linearly independent as well, thus we may assume that $B' := \{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$ is an $\mathbb{R}$-basis of $V$. Hence there are $a_1, \ldots, a_n \in \mathbb{R}$ 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 $\mathbb{R}$-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 an $\mathbb{R}$-basis of $V$:

We have $v_m = a_m^{-1}(w_m - \sum_{i=1}^{m-1} a_i w_i - \sum_{i=m+1}^{n} a_i v_i) \in \langle B'' \rangle_{\mathbb{R}}$, and hence from $B' \setminus \{v_m\} \subseteq B''$ we conclude $\langle B'' \rangle_{\mathbb{R}} = V$, that is $B''$ is an $\mathbb{R}$-generating set. Let $b_1, \ldots, b_n \in \mathbb{R}$ 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 $\mathbb{R}$-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 $\mathbb{R}$-linearly independent. ♯

For example, $V := \mathbb{R}^{2\times 1}$ has standard $\mathbb{R}$-basis $\{e_1, e_2\}$, and $\{[1,1]^{\mathrm{tr}}\}$ is $\mathbb{R}$-linearly independent; 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 $\{[1,1]^{\mathrm{tr}}, [1,0]^{\mathrm{tr}}\} \subseteq V$ and $\{[1,1]^{\mathrm{tr}}, [0,1]^{\mathrm{tr}}\} \subseteq V$ are $\mathbb{R}$-linearly independent, and thus are $\mathbb{R}$-bases.

**(2.9) Dimension. a)** Let $V$ be a finitely generated $\mathbb{R}$-vector space. Then it follows from (2.7) and (2.8) that any finite $\mathbb{R}$-generating set of $V$ contains an $\mathbb{R}$-basis of $V$, any $\mathbb{R}$-linearly independent subset of $V$ can be extended to an $\mathbb{R}$-basis of $V$, and all $\mathbb{R}$-bases of $V$ are finite of the same cardinality.

The cardinality of any $\mathbb{R}$-basis of $V$ is called the **$\mathbb{R}$-dimension** $\dim_{\mathbb{R}}(V) \in \mathbb{N}_0$ of $V$; if $V$ is not finitely generated then we write $\dim_{\mathbb{R}}(V) = \infty$. For example, for $m, n \in \mathbb{N}_0$ we have $\dim_{\mathbb{R}}(\mathbb{R}^{n\times 1}) = n$ and $\dim_{\mathbb{R}}(\mathbb{R}^{m\times n}) = m \cdot n$.

**b)** We have the following further **numerical characterization of bases**: For any subset $B \subseteq V$ the following are equivalent:
**i)** $B$ is an $\mathbb{R}$-basis of $V$.
**ii)** $B$ is an $\mathbb{R}$-linearly independent subset of maximal cardinality.
**ii')** $B$ is an $\mathbb{R}$-linearly independent subset such that $|B| = \dim_{\mathbb{R}}(V)$.
**iii)** $B$ is an $\mathbb{R}$-generating set of minimal cardinality.
**iii')** $B$ is an $\mathbb{R}$-generating set such that $|B| = \dim_{\mathbb{R}}(V)$.

**c)** Let $U \leq V$. Then $U$ is finitely generated as well, and we have $\dim_{\mathbb{R}}(U) \leq \dim_{\mathbb{R}}(V)$, where equality holds if and only if $U = V$; note that this is the analogue for $\mathbb{R}$-vector spaces of a property of finite sets:

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

**(2.10) Theorem: Dimension formula for subspaces.** Let $V$ be an $\mathbb{R}$-vector space, and let $U, U' \leq V$ be finitely generated. Then we have $\dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(U') = \dim_{\mathbb{R}}(U + U') + \dim_{\mathbb{R}}(U \cap U')$.

**Proof.** Let $m := \dim_{\mathbb{R}}(U) \in \mathbb{N}_0$ and $l := \dim_{\mathbb{R}}(U') \in \mathbb{N}_0$. Then $U + U'$ is finitely generated, hence both $n := \dim_{\mathbb{R}}(U + U') \in \mathbb{N}_0$ and $k := \dim_{\mathbb{R}}(U \cap U') \in \mathbb{N}_0$ are well-defined. Let $C := \{v_1, \ldots, v_k\}$ be an $\mathbb{R}$-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 $\mathbb{R}$-bases of $U$ and $U'$, respectively. Hence we have $\langle C \cup B \cup B' \rangle_{\mathbb{R}} = U + U'$.

Let $a_1, \ldots, a_k, b_1, \ldots, b_{m-k}, b'_1, \ldots, b'_{l-k} \in \mathbb{R}$ 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 $\sum_{i=1}^{k} a_i v_i + \sum_{i=1}^{m-k} b_i w_i = -\sum_{i=1}^{l-k} b'_i w'_i \in U \cap U' = \langle C \rangle_{\mathbb{R}}$, which since $C \mathbin{\dot\cup} B$ is $\mathbb{R}$-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 $\mathbb{R}$-linearly independent. Thus we have $B \cap B' = \emptyset$, and hence $C \mathbin{\dot\cup} B \mathbin{\dot\cup} B'$ is an $\mathbb{R}$-basis of $U + U'$, where $|C \mathbin{\dot\cup} B \mathbin{\dot\cup} B'| = k + (m - k) + (l - k) = m + l - k$. ♯

## 3 Linear maps

**(3.1) Linear maps. a)** Let $V$ and $W$ be $\mathbb{R}$-vector spaces. Then a map $\varphi \colon V \to W$ fulfilling the following conditions is called **$\mathbb{R}$-linear** or a **$\mathbb{R}$-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 $a \in \mathbb{R}$.

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

We derive a few immediate consequences: For an $\mathbb{R}$-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 $\varphi(-v) = \varphi((-1) \cdot v) = (-1) \cdot \varphi(v) = -\varphi(v)$.

**b)** Hence we have $0 \in \mathrm{im}(\varphi)$, and for $w, w' \in \mathrm{im}(\varphi)$ and $a \in \mathbb{R}$, 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 $\mathrm{im}(\varphi) \leq W$. Hence $\mathrm{rk}(\varphi) := \dim_{\mathbb{R}}(\mathrm{im}(\varphi)) \in \mathbb{N}_0 \mathbin{\dot\cup} \{\infty\}$ is well-defined, being called the **rank** of $\varphi$.

Let $\ker(\varphi) := \{v \in V; \varphi(v) = 0\}$ be the **kernel** of $\varphi$. Hence we have $0 \in \ker(\varphi)$, and for $v, v' \in \ker(\varphi)$ and $a \in \mathbb{R}$ 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 $\ker(\varphi) \leq V$. 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 $\varphi$ is injective if and only if $\ker(\varphi) = \{0\}$.

If $\varphi\colon V \to W$ is an $\mathbb{R}$-isomorphism, that is bijective, then the inverse map $\varphi^{-1}\colon W \to V$ is $\mathbb{R}$-linear as well, showing that $V$ and $W$, together with addition and scalar multiplication, can be identified via $\varphi$ and $\varphi^{-1}$: For $w, w' \in W$ and $a \in \mathbb{R}$, 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 $\varphi^{-1}(aw) = av = a\varphi^{-1}(w)$.

**(3.2) Theorem: Linear maps and bases. a)** Let $V$ and $W$ be $\mathbb{R}$-vector spaces, let $B := \{v_1, \dots, v_n\} \subseteq V$ be an $\mathbb{R}$-basis, where $n = \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$, and let $C := [w_1, \dots, w_n] \subseteq W$. Then there is a unique $\mathbb{R}$-linear map $\varphi\colon V \to W$ such that $\varphi(v_i) = w_i$, for all $i \in \{1, \dots, n\}$.

This says that $\mathbb{R}$-linear maps can be defined, and then are uniquely determined, by prescribing arbitrarily the images of the elements of a chosen $\mathbb{R}$-basis.

**b)** We have $\mathrm{im}(\varphi) = \langle C \rangle_{\mathbb{R}} \leq W$; in particular, $\varphi$ is surjective if and only if $C \subseteq W$ is an $\mathbb{R}$-generating set. Moreover, $\varphi$ is injective if and only if $C \subseteq W$ is $\mathbb{R}$-linearly independent; thus $\varphi$ is bijective if and only if $C \subseteq W$ is an $\mathbb{R}$-basis.

**c)** The map $V \to \mathbb{R}^{n \times 1}\colon v \mapsto {}_B v$ is an $\mathbb{R}$-isomorphism.

This says that, having chosen an $\mathbb{R}$-basis, any $\mathbb{R}$-vector space of $\mathbb{R}$-dimension $n \in \mathbb{N}_0$ can be identified with the column space $\mathbb{R}^{n \times 1}$.

**Proof. a)** We first show uniqueness: Since $B$ is an $\mathbb{R}$-generating set, for all $v \in V$ there are $a_1, \dots, a_n \in \mathbb{R}$ such that $v = \sum_{i=1}^n a_i v_i \in V$. Hence if $\varphi$ is as asserted, then we have $\varphi(v) = \sum_{i=1}^n a_i \varphi(v_i) = \sum_{i=1}^n a_i w_i \in W$; this also shows that $\mathrm{im}(\varphi) = \langle C \rangle_{\mathbb{R}}$. We now show that $\varphi$ as asserted exists:

Since $B$ is $\mathbb{R}$-linearly independent, the above representation $v = \sum_{i=1}^n a_i v_i \in V$ is unique. Hence there is a well-defined map $\varphi\colon V \to W$ given by $\varphi(v) := \sum_{i=1}^n a_i w_i \in W$. We show that $\varphi$ is $\mathbb{R}$-linear: Let $v' = \sum_{i=1}^n a'_i v_i \in V$ where $a'_1, \dots, a'_n \in \mathbb{R}$, and $a \in \mathbb{R}$, then we have $v + v' = \sum_{i=1}^n (a_i + a'_i) v_i$ and $av = \sum_{i=1}^n a a_i v_i$, showing that $\varphi(v + v') = \sum_{i=1}^n (a_i + a'_i) w_i = \sum_{i=1}^n a_i w_i + \sum_{i=1}^n a'_i w_i = \varphi(v) + \varphi(v')$, and $\varphi(av) = \sum_{i=1}^n a a_i w_i = a\varphi(v)$.

**b)** If $C$ is $\mathbb{R}$-linearly independent, then for $v = \sum_{i=1}^n a_i v_i \in \ker(\varphi)$ we have $0 = \varphi(v) = \sum_{i=1}^n a_i w_i$, implying $a_i = 0$ for all $i \in \{1, \dots, n\}$, showing that $\ker(\varphi) = \{0\}$. Conversely, if $\ker(\varphi) = \{0\}$, then for $\sum_{i=1}^n a_i w_i = 0$, where $a_1, \dots, a_n \in \mathbb{R}$, we have $\sum_{i=1}^n a_i v_i \in \ker(\varphi) = \{0\}$, implying $a_i = 0$ for all $i \in \{1, \dots, n\}$, showing that $C$ is $\mathbb{R}$-linearly independent.

**c)** Let $[e_1, \dots, e_n]$ be the standard $\mathbb{R}$-basis of $\mathbb{R}^{n \times 1}$, and let $\beta\colon V \to \mathbb{R}^{n \times 1}$ be defined by $v_i \mapsto e_i$, for $i \in \{1, \dots, n\}$. Then for $v = \sum_{i=1}^n a_i v_i \in V$, where $a_1, \dots, a_n \in \mathbb{R}$, we get $\beta(v) = \sum_{i=1}^n a_i e_i = [a_1, \dots, a_n]^{\mathrm{tr}} = {}_B v \in \mathbb{R}^{n \times 1}$. ♯

**(3.3) Theorem: Dimension formula for linear maps.** Let $V$ be a finitely generated $\mathbb{R}$-vector space.

**a)** Let $W$ be an $\mathbb{R}$-vector space, and let $\varphi\colon V \to W$ be an $\mathbb{R}$-linear map. Then we have $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(\ker(\varphi)) + \mathrm{rk}(\varphi) \in \mathbb{N}_0$.

**b)** For any $\mathbb{R}$-linear map $\varphi\colon V \to V$ the following are equivalent, analogously to the equivalence of injectivity and surjectivity of maps from a finite set to itself:
**i)** The map $\varphi$ is an $\mathbb{R}$-automorphism, that is $\varphi$ is bijective.
**ii)** The map $\varphi$ is an $\mathbb{R}$-monomorphism, that is $\varphi$ is injective.
**iii)** The map $\varphi$ is an $\mathbb{R}$-epimorphism, that is $\varphi$ is surjective.

**Proof. a)** Since $\mathrm{im}(\varphi) \leq W$ is finitely generated, we conclude that $r := \mathrm{rk}(\varphi) = \dim_{\mathbb{R}}(\mathrm{im}(\varphi)) \in \mathbb{N}_0$. Let $C := \{w_1, \ldots, w_r\} \subseteq \mathrm{im}(\varphi)$ be an $\mathbb{R}$-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 an $\mathbb{R}$-basis, where $k := \dim_{\mathbb{R}}(\ker(\varphi)) \leq \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$. We show that the sequence $[B', B] := [v_1', \ldots, v_k', v_1, \ldots, v_r] \subseteq V$ is an $\mathbb{R}$-basis, implying $\dim_{\mathbb{R}}(V) = k + r$:

Since $C \subseteq \mathrm{im}(\varphi)$ is an $\mathbb{R}$-generating set, for $v \in V$ we have $\varphi(v) = \sum_{j=1}^{r} a_j w_j \in \mathrm{im}(\varphi)$, for some $a_1, \ldots, a_r \in \mathbb{R}$. Hence we have $v - \sum_{j=1}^{r} a_j v_j \in \ker(\varphi)$, thus since $B' \subseteq \ker(\varphi)$ is an $\mathbb{R}$-generating set there are $a_1', \ldots, a_k' \in \mathbb{R}$ 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 an $\mathbb{R}$-generating set.

Let $a_1', \ldots, a_k', a_1, \ldots, a_r \in \mathbb{R}$ 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 $\mathbb{R}$-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 $\mathbb{R}$-linearly independent we get $a_1' = \cdots = a_k' = 0$, showing that $[B', B] \subseteq V$ is $\mathbb{R}$-linearly independent.

**b)** From $\dim_{\mathbb{R}}(\ker(\varphi)) = \dim_{\mathbb{R}}(V) - \mathrm{rk}(\varphi)$ we get $\dim_{\mathbb{R}}(\ker(\varphi)) = 0$ if and only if $\mathrm{rk}(\varphi) = \dim_{\mathbb{R}}(V)$, that is we have $\ker(\varphi) = \{0\}$ if and only if $\mathrm{im}(\varphi) = V$, which says that $\varphi$ is injective if and only if $\varphi$ is surjective. $\sharp$

**(3.4) Linear maps and matrices. a)** We first consider the prototypical example of $\mathbb{R}$-linear maps: Given a matrix $A \in \mathbb{R}^{m \times n}$, where $m, n \in \mathbb{N}_0$. we have $A(v + v') = Av + Av'$ and $A(av) = a \cdot Av$, for all $v, v' \in \mathbb{R}^{n \times 1}$ and $a \in \mathbb{R}$, hence the matrix product gives rise to the $\mathbb{R}$-linear map $\varphi_A\colon \mathbb{R}^{n \times 1} \to \mathbb{R}^{m \times 1}\colon v \mapsto Av$.

Let $\ker(A) := \ker(\varphi_A) = \{v \in \mathbb{R}^{n \times 1}; Av = 0\} \leq \mathbb{R}^{n \times 1}$ be the **(column) kernel** of $A$. Moreover, letting $w_j := [a_{1j}, \ldots, a_{mj}]^{\mathrm{tr}} = \varphi_A(e_j) \in \mathbb{R}^{m \times 1}$, for $j \in \{1, \ldots, n\}$, be the **columns** of $A$, that is the image of the standard $\mathbb{R}$-basis of $\mathbb{R}^{n \times 1}$, we have $\mathrm{im}(A) := \mathrm{im}(\varphi_A) = \langle w_1, \ldots, w_n \rangle_{\mathbb{R}} \leq \mathbb{R}^{m \times 1}$, being called the **image** or **column space** of $A$. Hence $\mathrm{rk}(A) := \mathrm{rk}(\varphi_A) = \dim_{\mathbb{R}}(\mathrm{im}(\varphi_A)) = \dim_{\mathbb{R}}(\mathrm{im}(A)) \in \{0, \ldots, \min\{m, n\}\}$ is called the **column rank** of $A$.

**b)** Let now $V$ and $W$ be $\mathbb{R}$-vector spaces, with $\mathbb{R}$-bases $B := [v_1, \ldots, v_n] \subseteq V$ and $C := [w_1, \ldots, w_m] \subseteq W$, where $n = \dim_{\mathbb{R}}(V)$ and $m = \dim_{\mathbb{R}}(W)$. Given an $\mathbb{R}$-linear map $\varphi\colon V \to W$, let $a_{ij} \in \mathbb{R}$, 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 $v = \sum_{j=1}^{n} b_j v_j \in V$, where $b_1, \ldots, b_n \in \mathbb{R}$, we get $\varphi(v) = \sum_{j=1}^{n} b_j \varphi(v_j) = \sum_{j=1}^{n} b_j (\sum_{i=1}^{m} a_{ij} w_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} b_j) w_i \in W$.

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

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

**c)** Let $\mathrm{Hom}_{\mathbb{R}}(V,W) := \{\varphi\colon V \to W; \varphi\ \mathbb{R}\text{-linear}\}$. Then $0 \in \mathrm{Hom}_{\mathbb{R}}(V,W)$, and for $\varphi,\psi \in \mathrm{Hom}_{\mathbb{R}}(V,W)$ and $a \in \mathbb{R}$ we have $\varphi+\psi \in \mathrm{Hom}_{\mathbb{R}}(V,W)$ and $a\varphi \in \mathrm{Hom}_{\mathbb{R}}(V,W)$, hence $\mathrm{Hom}_{\mathbb{R}}(V,W) \leq \mathrm{Maps}(V,W)$ is an $\mathbb{R}$-subspace.

Indeed, the map ${}_C\Phi_B\colon \mathrm{Hom}_{\mathbb{R}}(V,W) \to \mathbb{R}^{m\times n}\colon \varphi \mapsto {}_C\varphi_B$ is an $\mathbb{R}$-isomorphism; in particular, this says that, upon choice of $\mathbb{R}$-bases, any $\mathbb{R}$-linear map between $V$ and $W$ is described by a matrix product, for a uniquely determined matrix, and conversely any such matrix comes from an $\mathbb{R}$-linear map, and we have $\dim_{\mathbb{R}}(\mathrm{Hom}_{\mathbb{R}}(V,W)) = \dim_{\mathbb{R}}(\mathbb{R}^{m\times n}) = \dim_{\mathbb{R}}(V)\cdot\dim_{\mathbb{R}}(W)$:

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

**(3.5) Matrix products. a)** We consider the composition of $\mathbb{R}$-linear maps: Let $U,V,W$ be $\mathbb{R}$-vector spaces. If $\varphi\colon V \to W$ and $\psi\colon U \to V$ are $\mathbb{R}$-linear, then for $u,u' \in U$ and $a \in \mathbb{R}$ 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 $\varphi\psi\colon U \to W$ is $\mathbb{R}$-linear as well.

Let $R := [u_1,\ldots,u_l] \subseteq U$ and $S := [v_1,\ldots,v_n] \subseteq V$ and $T := [w_1,\ldots,w_m] \subseteq W$ be $\mathbb{R}$-bases, respectively, where $l := \dim_{\mathbb{R}}(U) \in \mathbb{N}_0$ and $m := \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$ and $n := \dim_{\mathbb{R}}(E) \in \mathbb{N}_0$. Moreover, let ${}_T\varphi_S = [a_{ij}]_{ij} \in \mathbb{R}^{m\times n}$ and ${}_S\psi_R = [b_{ij}]_{ij} \in \mathbb{R}^{n\times l}$ be the matrices associated with $\varphi$ and $\psi$.

Then for $k \in \{1,\ldots,l\}$ we have $\varphi\psi(u_k) = \varphi(\sum_{j=1}^n b_{jk}v_j) = \sum_{j=1}^n b_{jk}\varphi(v_j) = \sum_{j=1}^n \sum_{i=1}^m b_{jk}a_{ij}w_i = \sum_{i=1}^m (\sum_{j=1}^n a_{ij}b_{jk})w_i$, thus the composition $\varphi\psi$ is described by the matrix ${}_T(\varphi\psi)_R = [\sum_{j=1}^n a_{ij}b_{jk}]_{ik} \in \mathbb{R}^{m\times l}$. Then we have ${}_T(\varphi\psi)_R = {}_T\varphi_S \cdot {}_S\psi_R$, as soon as we adopt the following definition:

For matrices $A = [a_{ij}]_{ij} \in \mathbb{R}^{m\times n}$ and $B = [b_{ij}]_{ij} \in \mathbb{R}^{n\times l}$ we define the **ma-**

**trix product** $A \cdot B := [\sum_{j=1}^n a_{ij}b_{jk}]_{ik} \in \mathbb{R}^{m\times l}$. In particular, for $l = 1$, the elements of $\mathbb{R}^{n\times 1}$ are just columns, and for $[x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathbb{R}^{n\times 1}$ we recover $A \cdot [x_1, \ldots, x_n]^{\mathrm{tr}} = [\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j]^{\mathrm{tr}} \in \mathbb{R}^{m\times 1}$.

Identifying the matrix $A \in \mathbb{R}^{m\times n}$ with the $\mathbb{R}$-linear map $\varphi_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, associativity of maps implies $(AB)C = A(BC) \in \mathbb{R}^{m\times k}$, whenever $C \in \mathbb{R}^{l\times k}$ for some $k \in \mathbb{N}_0$, hence associativity holds for the matrix product as well.

**b)** In particular, $\mathrm{Hom}_{\mathbb{R}}(V, V)$ is closed under composition of maps. Since the composition of bijective maps is bijective again, and composition is associative, we infer that $\mathrm{GL}(V) := \{\varphi \in \mathrm{Hom}_{\mathbb{R}}(V, V); \varphi \text{ bijective}\}$ becomes an (in general non-commutative) group, with neutral element $\mathrm{id}_V \in \mathrm{GL}(V)$, and inverses given by $\varphi^{-1} \in \mathrm{GL}(V)$, for $\varphi \in \mathrm{GL}(V)$, called the **general linear group** on $V$.

Similarly, $\mathrm{GL}_n(\mathbb{R}) := \{A \in \mathbb{R}^{n\times n}; \varphi_A \text{ bijective}\}$ becomes an (in general non-commutative) group with respect to the matrix product, called the **general linear group** of degree $n$ over $\mathbb{R}$, with neutral element $E_n \in \mathrm{GL}_n(\mathbb{R})$, and inverses $A^{-1} \in \mathrm{GL}_n(\mathbb{R})$ given by the property $\varphi_{A^{-1}} = (\varphi_A)^{-1}$, for $A \in \mathrm{GL}_n(\mathbb{R})$, that is fulfilling $AA^{-1} = E_n = A^{-1}A$; its elements are called **invertible**.

For example, in $\mathrm{GL}_2(\mathbb{R})$, for the reflections mentioned above we have

$$\begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix} = \begin{bmatrix} . & 1 \\ -1 & . \end{bmatrix} \neq \begin{bmatrix} . & -1 \\ 1 & . \end{bmatrix} = \begin{bmatrix} -1 & . \\ . & 1 \end{bmatrix} \cdot \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix};$$

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

**(3.6) Base change. a)** Let $V$ be an $\mathbb{R}$-vector space, where $n := \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$, and let $B := [v_1, \ldots, v_n] \subseteq V$ and $B' := [v_1', \ldots, v_n'] \subseteq V$ be $\mathbb{R}$-bases. Then ${}_B\mathrm{id}_{B'} = [b_{ij}]_{ij} \in \mathbb{R}^{n\times n}$ 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 ${}_{B'}\mathrm{id}_B \cdot {}_B\mathrm{id}_{B'} = {}_{B'}\mathrm{id}_{B'} = E_n$, implying that ${}_B\mathrm{id}_{B'} \in \mathrm{GL}_n(\mathbb{R})$ with inverse $({}_B\mathrm{id}_{B'})^{-1} = {}_{B'}\mathrm{id}_B \in \mathrm{GL}_n(\mathbb{R})$.

Letting $W$ be a finitely generated $\mathbb{R}$-vector space, where $m := \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$, having $\mathbb{R}$-bases $C \subseteq W$ and $C' \subseteq W$, and $\varphi \colon V \to W$ be $\mathbb{R}$-linear, the matrices ${}_C\varphi_B \in \mathbb{R}^{m\times n}$ and ${}_{C'}\varphi_{B'} \in \mathbb{R}^{m\times n}$ are related by the **base change formula** ${}_{C'}\varphi_{B'} = {}_{C'}\mathrm{id}_C \cdot {}_C\varphi_B \cdot {}_B\mathrm{id}_{B'} = ({}_C\mathrm{id}_{C'})^{-1} \cdot {}_C\varphi_B \cdot {}_B\mathrm{id}_{B'}$, where ${}_C\mathrm{id}_{C'} \in \mathrm{GL}_m(\mathbb{R})$.

**b)** 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 ${}_B\mathrm{id}_C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$, and 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 $\sigma$ at the hyperplane perpendicular to $[-1, 1]^{\mathrm{tr}}$ with respect to the $\mathbb{R}$-basis $B$ we have ${}_B\sigma_B = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \in \mathbb{R}^{2\times 2}$. There are various (equivalent) ways to find the matrix of $\sigma$ with respect to the $\mathbb{R}$-basis $C$:

Geometrically, we have $\sigma(v_1) = v_1$ and $\sigma(v_2) = -v_2$. In terms of matrices with respect to the $\mathbb{R}$-basis $B$ this reads ${}_B(\sigma(v_1)) = {}_B\sigma_B \cdot {}_B(v_1) = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = {}_B(v_1)$ and ${}_B(\sigma(v_2)) = {}_B\sigma_B \cdot {}_B(v_2) = \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = {}_B(-v_2)$. Anyway, we thus get ${}_C\sigma_C = \begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$. Alternatively, using the base change formula we obtain

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

**(3.7) Matrix rank.** Having this machinery in place, we give another proof of the dimension formula for linear maps, saying that for an $\mathbb{R}$-linear map $\varphi \colon V \to W$, where $n := \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$, we have $\dim_{\mathbb{R}}(\ker(\varphi)) + \mathrm{rk}(\varphi) = n$; note that we may assume that $m := \dim_{\mathbb{R}}(V) \in \mathbb{N}_0$. To this end, we choose $\mathbb{R}$-bases of $V$ and $W$, respectively, in order to translate $\varphi$ into a map $\varphi_A$, for a matrix $A \in \mathbb{R}^{m\times n}$, for which we have to show that $\dim_{\mathbb{R}}(\ker(A)) + \mathrm{rk}(A) = n$.

**a)** Considering the system of linear equations associated with $A$, we observe that $\ker(A) = \{v \in \mathbb{R}^{n\times 1}; Av = 0\} = \mathcal{L}(A) \leq \mathbb{R}^{n\times 1}$ is the $\mathbb{R}$-subspace of solutions of the associated homogeneous system. Hence, using the Gaussian normal form $A' \in \mathbb{R}^{m\times n}$ of $A$, an $\mathbb{R}$-basis of $\ker(A)$ is given by the basic solutions $\{v_j \in \mathbb{R}^{n\times 1}; j \in \mathcal{J}\}$, where $\mathcal{J} := \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$ is the set of non-pivot columns of $A$; see (1.5). Thus we have $\dim_{\mathbb{R}}(\ker(A)) = |\mathcal{J}| = n - r$, where $r = r(A) \in \{0, \ldots, \min\{m, n\}\}$ is the number of non-zero rows of $A'$.

The number $r(A)$ can be interpreted as follows: Letting $v_i := [a_{i1}, \ldots, a_{in}] \in \mathbb{R}^n$, for $i \in \{1, \ldots, m\}$, be the **rows** of $A$, let $\langle v_1, \ldots, v_m \rangle_{\mathbb{R}} \leq \mathbb{R}^n$ be the **row space** of $A$, hence $\dim_{\mathbb{R}}(\langle v_1, \ldots, v_m \rangle_{\mathbb{R}})$ is called the **row rank** of $A$. Performing row operations on $A$ produces $\mathbb{R}$-linear combinations of the $v_i$, which belong to the row space of $A$, but since the admissible operations are reversible, we conclude that the row space of $A$ remains actually unchanged. In particular, letting $v_1', \ldots, v_m' \in \mathbb{R}^n$ be the rows of $A'$, we have $\langle v_1, \ldots, v_m \rangle_{\mathbb{R}} = \langle v_1', \ldots, v_r' \rangle_{\mathbb{R}} \leq \mathbb{R}^n$. Moreover, letting $a_1, \ldots, a_r \in \mathbb{R}$ such that $\sum_{i=1}^r a_i v_i' = 0 \in \mathbb{R}^n$, considering the pivot columns $[j_1, \ldots, j_r]$ shows that $a_i = 0 \in \mathbb{R}$, for all $i \in \{1, \ldots, r\}$, saying that $[v_1', \ldots, v_r']$ is $\mathbb{R}$-linearly independent. Thus $\{v_1', \ldots, v_r'\}$ is an $\mathbb{R}$-basis of the row space of $A$, hence the row rank of $A$ is $\dim_{\mathbb{R}}(\langle v_1, \ldots, v_m \rangle_{\mathbb{R}}) = r = r(A)$.

**b)** Thus we have to show that $\mathrm{rk}(A) = r(A)$, that is the row and column ranks of $A$ coincide, hence just being called the **rank** of $A$:

Row operations on the matrix $A$ can be described as follows: Multiplying row $i$ with $a \in \mathbb{R}^*$, where $i \in \{1, \ldots, m\}$, yields $A' = E_i(a) \cdot A$, where $E_i(a) := \mathrm{diag}[1, \ldots, 1, a, 1, \ldots, 1] \in \mathbb{R}^{m\times m}$; adding the $a$-fold of row $j$ to row $i$, where $i \neq$

$j \in \{1, \ldots, m\}$ and $a \in \mathbb{R}$, yields $A' = E_{i,j}(a) \cdot A$, where $E_{i,j}(a) := E_m + aE_{ij} \in \mathbb{R}^{m \times m}$; and interchanging rows $i$ and $j$ yields $A' = E(i,j) \cdot A$, where $E(i,j) := E_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} \in \mathbb{R}^{m \times m}$. Since $E_i(a)E_i(a^{-1}) = E_{i,j}(a)E_{i,j}(-a) = E(i,j)^2 = E_m$, for these **elementary matrices** we have $E_i(a), E_{i,j}(a), E(i,j) \in \mathrm{GL}_m(\mathbb{R})$, which of course is just expressing the invertibility of row operations.

Thus, for the Gaussian normal form of $A$ we have $A' = PA$, where $P \in \mathrm{GL}_m(\mathbb{R})$. Now, for $P \in \mathbb{R}^{m \times m}$ we have $\mathrm{rk}(PA) = \dim_{\mathbb{R}}(\mathrm{im}(\varphi_P \varphi_A)) \leq \dim_{\mathbb{R}}(\mathrm{im}(\varphi_A)) = \mathrm{rk}(A)$, hence for $P \in \mathrm{GL}_m(\mathbb{R})$ we have $\mathrm{rk}(A) = \mathrm{rk}(P^{-1}PA) \leq \mathrm{rk}(PA) \leq \mathrm{rk}(A)$, implying $\mathrm{rk}(PA) = \mathrm{rk}(A)$. In particular, we infer $\mathrm{rk}(A) = \mathrm{rk}(A')$, where considering the pivot columns $[j_1, \ldots, j_r]$ of $A'$ shows that the column space of $A'$ is given as $\langle e_1, \ldots, e_r \rangle_{\mathbb{R}} \leq \mathbb{R}^{m \times 1}$, implying that $\mathrm{rk}(A') = r = r(A)$. ♯

For example, let

$$A := \begin{bmatrix} 1 & 1 & . \\ 1 & . & 1 \\ . & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 3}.$$

For the columns $w_1, \ldots, w_3 \in \mathbb{R}^{4 \times 1}$ of $A$ we have $w_1 = w_2 + w_3$, hence $\{w_2, w_3\} \subseteq \mathbb{R}^{4 \times 1}$ being $\mathbb{R}$-linearly independent is an $\mathbb{R}$-basis of the column space of $A$; for the rows $v_1, \ldots, v_4 \in \mathbb{R}^3$ of $A$ we have $v_3 = v_1 - v_2$ and $v_4 = v_1 + v_2$, hence $\{v_1, v_2\} \subseteq \mathbb{R}^3$ being $\mathbb{R}$-linearly independent is an $\mathbb{R}$-basis of the row space of $A$.

**(3.8) Gaussian normal forms, revisited.** Let $A \in \mathbb{R}^{m \times n}$, where $m, n \in \mathbb{N}_0$, with Gaussian normal form $A'$ and pivot columns $1 \leq j_1 < \cdots < j_r \leq n$.

**a)** We prove that the Gaussian normal form of $A$ is uniquely determined; recall that it necessarily has $r = r(A) = \mathrm{rk}(A)$ of non-zero rows:

Let $\widetilde{A} \in \mathbb{R}^{m \times n}$ be a Gaussian normal form of $A$, with non-zero rows $\widetilde{v}_1, \ldots, \widetilde{v}_r \in \mathbb{R}^n$ 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 $\langle \widetilde{v}_k, \ldots, \widetilde{v}_r \rangle_{\mathbb{R}} \leq \langle v_{k+1}, \ldots, v_r \rangle_{\mathbb{R}}$, which since $\dim_{\mathbb{R}}(\langle \widetilde{v}_k, \ldots, \widetilde{v}_r \rangle_{\mathbb{R}}) = r - k + 1 > r - k = \dim_{\mathbb{R}}(\langle v_{k+1}, \ldots, v_r \rangle_{\mathbb{R}})$ 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 $i \in \{1, \ldots, r\}$. ♯

**b)** An $\mathbb{R}$-basis of the column space $\mathrm{im}(A) \leq \mathbb{R}^{m \times 1}$ is found as follows: Let $P \in \mathrm{GL}_m(\mathbb{R})$ such that $PA = A'$, and let $w_1, \ldots, w_n \in \mathbb{R}^{m \times 1}$ be the columns of $A$. Then $P \cdot [w_{j_1}, \ldots, w_{j_r}] \in \mathbb{R}^{m \times r}$ are just the pivot columns of $A'$, thus we infer $\mathrm{rk}([w_{j_1}, \ldots, w_{j_r}]) = \mathrm{rk}(P \cdot [w_{j_1}, \ldots, w_{j_r}]) = r$, hence $\{w_{j_1}, \ldots, w_{j_r}\}$ is an $\mathbb{R}$-basis of the column space of $A$.

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

**d)** We consider the case $m = n$: Then $\dim_{\mathbb{R}}(\ker(A)) = n - \mathrm{rk}(A)$ implies that the following are equivalent:
**i)** The map $\varphi_A$ is bijective, that is $A \in \mathrm{GL}_n(\mathbb{R})$.
**ii)** The map $\varphi_A$ is injective, that is $\ker(A) = \{0\}$, that is $\dim_{\mathbb{R}}(\ker(A)) = 0$.
**iii)** The map $\varphi_A$ surjective, that is $\mathrm{im}(A) = \mathbb{R}^{n\times 1}$, that is $\mathrm{rk}(A) = n$.

In order to compare this with (1.5)(c), note that (i) is equivalent to $A' = E_n$, while (ii) is equivalent to $|\mathcal{L}(A)| = 1$, and (iii) is equivalent to $\mathcal{L}(A, w) \neq \emptyset$ for all $w \in \mathbb{R}^{n\times 1}$; for the remaining statements on $\mathcal{L}(A, w)$ recall that for $v, v' \in \mathbb{R}^{n\times 1}$ we have $\varphi_A(v) = \varphi_A(v')$ if and only if $v - v' \in \ker(A)$.

Hence to show that $A \in \mathbb{R}^{n\times n}$ is invertible, it suffices to exhibit $B \in \mathbb{R}^{n\times n}$ such that $AB = E_n$, since then $\varphi_A$ is surjective and thus $B = A^{-1} \in \mathrm{GL}_n(\mathbb{R})$; likewise it suffices to exhibit $C \in \mathbb{R}^{n\times n}$ such that $CA = E_n$, since then $\varphi_A$ is injective and thus $C = A^{-1} \in \mathrm{GL}_n(\mathbb{R})$. Moreover, we have $A \in \mathrm{GL}_n(\mathbb{R})$ if and only if $A' = E_n \in \mathbb{R}^{n\times n}$, in which case we have $PA = E_n$, that is $P = A^{-1} \in \mathrm{GL}_n(\mathbb{R})$, and thus running the Gauß algorithm on the extended matrix is an algorithm for **matrix inversion**; for an example, see (1.7). Note that it follows that any invertible matrix $P \in \mathrm{GL}_n(\mathbb{R})$ can be written as a product of elementary matrices.

## 4 Determinants

**(4.1) Oriented volumes. a)** Let $V$ be an $\mathbb{R}$-vector space and $n \in \mathbb{N}_0$. A map $\delta\colon V^n \to \mathbb{R}$ fulfilling the following conditions is called a **determinant form** of **degree** $n$: It is **$\mathbb{R}$-multilinear**, that is for $i \in \{1, \ldots, n\}$ and fixed $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$ the map $V \to \mathbb{R}\colon v \mapsto \delta(\ldots, v, \ldots)$ is $\mathbb{R}$-linear, and **alternating**, that is for $i < j$ and $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in V$ fixed and $v, v' \in V$ we have $\delta(\ldots, v, \ldots, v', \ldots) = -\delta(\ldots, v', \ldots, v, \ldots) \in \mathbb{R}$.

The set of determinant forms is an $\mathbb{R}$-subspace of $\mathrm{Maps}(V^n, \mathbb{R})$. Moreover, for $i < j$ and $v \in V$ we get $\delta(\ldots, v, \ldots, v, \ldots) = -\delta(\ldots, v, \ldots, v, \ldots)$, that is $\delta(\ldots, v, \ldots, v, \ldots) = 0$; actually, the latter property is equivalent to being alternating: For $i < j$ and $v_1, \ldots, v_n \in V$ we have $0 = \delta(\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) = \delta(\ldots, v_i, \ldots, v_i, \ldots) + \delta(\ldots, v_i, \ldots, v_j, \ldots) + \delta(\ldots, v_j, \ldots, v_i, \ldots) + \delta(\ldots, v_j, \ldots, v_j \ldots) = \delta(\ldots, v_i, \ldots, v_j, \ldots) + \delta(\ldots, v_j, \ldots, v_i, \ldots)$.

For $i \neq j$ and $v_1, \ldots, v_n \in V$ and $a \in \mathbb{R}$ we have $\delta(\ldots, v_i + av_j, \ldots, v_j, \ldots) = \delta(\ldots, v_i, \ldots, v_j, \ldots) + a\delta(\ldots, v_j, \ldots, v_j, \ldots) = \delta(v_1, \ldots, v_n)$. Hence we infer that $\delta(v_1, \ldots, v_n) = 0$ whenever $\dim_{\mathbb{R}}(\langle v_1, \ldots, v_n\rangle_{\mathbb{R}}) < n$: Let $\sum_{i=1}^n a_i v_i = 0$, where $0 \neq [a_1, \ldots, a_n] \in \mathbb{R}^n$, then we may assume that $a_n = 1$, hence we have $\delta(v_1, \ldots, v_n) = \delta(v_1, \ldots, v_{n-1}, v_n + \sum_{i=1}^{n-1} a_i v_i) = \delta(v_1, \ldots, v_{n-1}, 0) = 0$.

**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; and as normalization condition the **unit cube** $\{\sum_{i=1}^n a_i e_i \in \mathbb{R}^n; 0 \leq a_i \leq 1 \text{ for all } i \in \{1,\ldots,n\}\}$, being spanned by the standard $\mathbb{R}$-basis, has volume 1. Note that if the parallelotope is contained in a hyperplane, then it already follows from the above properties that the volume vanishes.

**(4.2) Permutations. a)** For $n \in \mathbb{N}_0$ let $\mathcal{S}_n$ be the set of all bijective maps $\pi\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$, where for $n = 0$ we just have $\{1,\ldots,0\} = \emptyset$. Actually, $\mathcal{S}_n$ is an (in general non-commutative) group with respect to the composition of maps, the neutral element is the identity map, and the inverse of a **permutation** $\pi \in \mathcal{S}_n$ is given by the inverse map $\pi^{-1} \in \mathcal{S}_n$.

Permutations $\pi \in \mathcal{S}_n$ are written as tuples $\pi = [\pi(1),\ldots,\pi(n)]$. For example, let $\tau_{ij} = [\ldots,j,\ldots,i,\ldots] := [1,\ldots,i-1,j,i+1,\ldots,j-1,i,j+1,\ldots,n] \in \mathcal{S}_n$ be a **transposition**, where $i < j \in \{1,\ldots,n\}$; moreover, we have $\mathcal{S}_1 = \{[1]\}$ and $\mathcal{S}_2 = \{[1,2],[2,1]\}$ and $\mathcal{S}_3 = \{[2,3,1],[3,2,1],[1,3,2],[3,1,2],[1,2,3],[2,1,3]\}$.

**b)** 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 $\pi$ 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 $\prod_{1\leq i<j\leq n} |\pi(j)-\pi(i)| = \prod_{1\leq i<j\leq n}(j-i) = \prod_{k\in\{1,\ldots,n-1\}}(n-k)!$, hence $\mathrm{sgn}\colon \mathcal{S}_n \to \{\pm1\}\colon \pi \mapsto (-1)^{l(\pi)}$, where $l(\pi) := |\{\{i,j\}; i<j, \pi(i)>\pi(j)\}| \in \mathbb{N}_0$ is called the **inversion number** of $\pi$.

For example, for $\pi \in \mathcal{S}_n$ from $l(\pi) = l(\pi^{-1})$ we infer $\mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})$; we have $l(\mathrm{id}) = 0$ implying $\mathrm{sgn}(\mathrm{id}) = 1$; and we have $l(\tau_{ij}) = 2(j-i)-1$, thus $\mathrm{sgn}(\tau_{ij}) = -1$, for $i < j \in \{1,\ldots,n\}$. Moreover, 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 $\mathrm{sgn}(\pi\rho) = (\prod_{1\leq i<j\leq n} \frac{\pi(j)-\pi(i)}{j-i}) \cdot (\prod_{1\leq i<j\leq n} \frac{\rho(j)-\rho(i)}{j-i}) = \mathrm{sgn}(\pi)\cdot\mathrm{sgn}(\rho)$. ♯

**(4.3) Determinants of matrices. a)** The **determinant** of $A := [a_{ij}]_{ij} \in \mathbb{R}^{n\times n}$, where $n \in \mathbb{N}_0$, is defined as $\det(A) := \sum_{\pi\in\mathcal{S}_n} \mathrm{sgn}(\pi)\cdot\prod_{j=1}^n a_{\pi(j),j} \in \mathbb{R}$.

For example, 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 $\mathcal{S}_2 = \{[1,2],[2,1]\}$ and hence we get $\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{22} - a_{12}a_{21}$; and for $n = 3$ we have $\mathcal{S}_3 = \{[1,2,3],[2,3,1],[3,1,2];[3,2,1],[2,1,3],[1,3,2]\}$ which thus yields $\det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{13}a_{22}a_{31} +$

$a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32})$, being called the **Sarrus rule**; note that this simple rule does not at all generalize to $n \geq 4$.

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

**b)** Then $\det\colon (\mathbb{R}^{n\times 1})^n \to \mathbb{R}\colon [v_1, \ldots, v_n] \mapsto \det(v_1, \ldots, v_n)$ is a determinant form, that is it is $\mathbb{R}$-multilinear and alternating:

Since each summand in the defining sum is $\mathbb{R}$-multilinear, det is as well. To show that det is alternating, we may assume that $n \geq 2$. Writing $v_j = [a_{1j}, \ldots, a_{nj}]^{\mathrm{tr}} \in \mathbb{R}^{n\times 1}$, for $i < j \in \{1, \ldots, n\}$ letting $\tau := \tau_{ij} \in \mathcal{S}_n$ we get $\det(\ldots, v_j, \ldots, v_i, \ldots) = \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\pi) \cdot \prod_{k=1}^n a_{\pi(k),\tau^{-1}(k)} = \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\pi) \cdot \prod_{k=1}^n a_{\tau\pi(k),k} = \mathrm{sgn}(\tau^{-1}) \cdot \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\tau\pi) \cdot \prod_{k=1}^n a_{\tau\pi(k),k} = -\det(v_1, \ldots, v_n)$. ♯

**c)** Moreover, we have $\det(A) = \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\pi) \cdot \prod_{j=1}^n a_{\pi(j),j} = \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\pi^{-1}) \cdot \prod_{j=1}^n a_{j,\pi^{-1}(j)} = \sum_{\pi \in \mathcal{S}_n} \mathrm{sgn}(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)}$. The latter expression says that $\det(A)$ is invariant under exchanging the roles of columns and rows of $A$. Hence in particular det is row $\mathbb{R}$-multilinear and row alternating, too.

Thus, performing row operations on $A \in \mathbb{R}^{n\times n}$ yielding $A' \in \mathbb{R}^{n \times n}$, changes $\det(A)$ as follows: Multiplying row $i$ with $a \in \mathbb{R}^*$, where $i \in \{1, \ldots, n\}$, yields $A' = E_i(a) \cdot A$ and $\det(A') = a \cdot \det(A) = \det(E_i(a))\det(A)$; adding a multiple of row $j$ to row $i$, where $j \neq i \in \{1, \ldots, n\}$, yields $A' = E_{i,j}(a) \cdot A$ and $\det(A') = \det(A) = \det(E_{i,j}(a))\det(A)$; interchanging row $i$ and row $j$ yields $A' = E(i,j) \cdot A$ and $\det(A') = -\det(A) = \det(E(i,j))\det(A)$.

Hence this allows to compute the determinant of $A$ by applying the Gauß algorithm and keeping track of the row operations made: If $\mathrm{rk}(A) < n$ then $\det(A) = 0$ anyway; and if $\mathrm{rk}(A) = n$ then as Gaussian normal form we get $A' = E_n$, hence $\det(A') = \det(E_n) = 1$.

**d)** For all $A, B \in \mathbb{R}^{n\times n}$ we have **multiplicativity** $\det(AB) = \det(A)\det(B)$:

If $\mathrm{rk}(A) < n$, then from $\mathrm{im}(AB) = \mathrm{im}(\varphi_A\varphi_B) \leq \mathrm{im}(\varphi_A) = \mathrm{im}(A)$ we get $\mathrm{rk}(AB) \leq \mathrm{rk}(A) < n$, thus $\det(AB) = 0 = \det(A) = \det(A)\det(B)$. If $\mathrm{rk}(A) = n$, then we have $A \in \mathrm{GL}_n(\mathbb{R})$, thus $A$ is a product of elementary matrices, hence it suffices to show that $\det(AB) = \det(A)\det(B)$ holds whenever $A$ is an elementary matrix, which indeed holds by the above observations. ♯

In particular, we conclude that for $A \in \mathrm{GL}_n(\mathbb{R})$ we have $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(E_n) = 1$, thus $\det(A^{-1}) = \det(A)^{-1} \in \mathbb{R}^*$. Thus for $A \in \mathbb{R}^{n\times n}$ we infer that $A \in \mathrm{GL}_n(\mathbb{R})$ if and only if $\det(A) \neq 0$.

**(4.4) Theorem: Laplace expansion.** Let $A = [a_{ij}]_{ij} \in \mathbb{R}^{n \times n}$, where $n \in \mathbb{N}$, and for $i, j \in \{1, \ldots, n\}$ let

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

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

Applying $\rho := [\ldots, n, i, i+1, \ldots, n-2, n-1] \in \mathcal{S}_n$ to the rows, and $\sigma := [\ldots, n, j, j+1, \ldots, n-2, n-1] \in \mathcal{S}_n$ to the columns of $A_j(e_i)$, we obtain $\left[\begin{array}{ccccc|c} \multicolumn{5}{c|}{A_{ij}} & . \\ \hline a_{i1} & \ldots & a_{i,j-1} \; a_{i,j+1} & \ldots & a_{in} & 1 \end{array}\right] \in \mathbb{R}^{n \times n}$. For the latter matrix, all summands in the defining sum of the determinant 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 $\mathrm{sgn}(\rho) = (-1)^{n-i}$ and $\mathrm{sgn}(\sigma) = (-1)^{n-j}$ yield $\det(A_j(e_i)) = (-1)^{i+j} \cdot \det(A_{ij})$. From this we get the column expansion $\det(A) = \det(A_j(w_j)) = \det(A_j(\sum_{i=1}^{n} a_{ij} e_i)) = \sum_{i=1}^{n} a_{ij} \cdot \det(A_j(e_i)) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$.

In order to obtain the row expansion formula, we exchange the roles of columns and rows of $A$: For $i \in \{1, \ldots, n\}$ and $v \in \mathbb{R}^n$ let $A_i(v) \in \mathbb{R}^{n \times n}$ be the matrix obtained from $A$ by replacing row $i$ by $v$. Then, similar to the argument above, we get $(-1)^{i+j} \cdot \det(A_{ij}) = \det(A_i(e'_j))$, where $e'_j \in \mathbb{R}^n$ is the $j$-th unit (row) vector, for $j \in \{1, \ldots, n\}$. Letting $v_1, \ldots, v_n \in \mathbb{R}^n$ be the rows of $A$, from this we get the row expansion $\det(A) = \det(A_i(v_i)) = \det(A_i(\sum_{j=1}^{n} a_{ij} e'_j)) = \sum_{j=1}^{n} a_{ij} \cdot \det(A_i(e'_j)) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$. ♯

**(4.5) Corollary: Cramer's rule. a)** Given a matrix $A \in \mathbb{R}^{n \times n}$, where $n \in \mathbb{N}$, let $\mathrm{adj}(A) := [(-1)^{i+j} \cdot \det(A_{ji})]_{ij} \in \mathbb{R}^{n \times n}$ be the associated **adjoint matrix**. Then we have $A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot E_n \in \mathbb{R}^{n \times n}$. In particular, if $A \in \mathrm{GL}_n(\mathbb{R})$ then we have $A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A) \in \mathrm{GL}_n(\mathbb{R})$.
**b)** If $A \in \mathrm{GL}_n(\mathbb{R})$, then the unique solution $[x_1, \ldots, x_n]^{\mathrm{tr}} \in \mathbb{R}^{n \times 1}$ of the system of linear equations with coefficient matrix $A$ and right hand side $w \in \mathbb{R}^{n \times 1}$, is by **Cramer's rule** given as $x_i := \det(A)^{-1} \cdot \det(A_i(w)) \in \mathbb{R}$, for $i \in \{1, \ldots, n\}$.

**Proof. a)** For $i, k \in \{1, \dots, n\}$ 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 $w_k \in \mathbb{R}^{n \times 1}$ 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 $\mathrm{adj}(A) \cdot A = \det(A) \cdot E_n$.

Similarly, we get $(A \cdot \mathrm{adj}(A))_{ik} = \sum_{j=1}^n a_{ij} \cdot (-1)^{j+k} \cdot \det(A_{kj}) = \sum_{j=1}^n a_{ij} \cdot \det(A_k(e'_j)) = \det(A_k(\sum_{j=1}^n a_{ij} e'_j)) = \det(A_k(v_i))$, where for $i \neq k$ we have $\det(A_k(v_i)) = 0$, while for $i = k$ we get $A_i(v_i) = A$, showing that $A \cdot \mathrm{adj}(A) = \det(A) \cdot E_n$ as well. Note that we can spare the second half of this argument in case $\det(A) \neq 0$, but a priorly not in case $\det(A) = 0$.

**b)** Letting $v := [x_1, \dots, x_n]^{\mathrm{tr}}$ and $w = [y_1, \dots, y_n]^{\mathrm{tr}}$, from $Av = w$ we get $v = A^{-1}w = \det(A)^{-1} \cdot \mathrm{adj}(A)w$, hence $\det(A) \cdot x_i = \sum_{j=1}^n (-1)^{i+j} \cdot \det(A_{ji}) \cdot y_j = \sum_{j=1}^n \det(A_i(e_j)) \cdot y_j = \det(A_i(\sum_{j=1}^n y_j e_j)) = \det(A_i(w))$. ♯

**(4.6) Example: Computation of determinants.** We come back to the matrix from (1.1), (1.3)(i) and (1.7):

**i)** For $A := \begin{bmatrix} 1 & 1 \\ 200 & 50 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ we get $\det(A) = 1 \cdot 50 - 1 \cdot 200 = -150$, expansion with respect to the first row yields $\det(A) = 1 \cdot \det([50]) - 1 \cdot \det([200]) = 1 \cdot 50 - 1 \cdot 200 = -150$, while expansion with respect to the first column yields $\det(A) = 1 \cdot \det([50]) - 200 \cdot \det([1]) = 1 \cdot 50 - 200 \cdot 1 = -150$; this implies that $A \in \mathrm{GL}_2(\mathbb{R})$. Moreover, we have $\mathrm{adj}(A) = \begin{bmatrix} \det([50]) & -\det([1]) \\ -\det([200]) & \det([1]) \end{bmatrix} = \begin{bmatrix} 50 & -1 \\ -200 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, thus we get $A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A) = \frac{1}{150} \cdot \begin{bmatrix} -50 & 1 \\ 200 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$.

**ii)** We recall the steps of the Gauß algorithm applied to $A$:

$$\begin{bmatrix} 1 & 1 \\ 200 & 50 \end{bmatrix} \overset{(1)}{\mapsto} \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \overset{(2)}{\mapsto} \begin{bmatrix} 1 & 1 \\ 3 & . \end{bmatrix} \overset{(3)}{\mapsto} \begin{bmatrix} 1 & 1 \\ 1 & . \end{bmatrix} \overset{(4)}{\mapsto} \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \overset{(5)}{\mapsto} \begin{bmatrix} 1 & . \\ . & 1 \end{bmatrix}$$

We keep track of the change of the determinant of the matrix under consideration: In step (1) we get a factor of 50, in step (3) we get a factor of 3, in step (5) we get a factor of $-1$, while in steps (2) and (4) we get a factor of 1 each. Hence we conclude that $\det(A) = 50 \cdot 3 \cdot (-1) = -150$.

**iii)** Letting $c, d \in \mathbb{R}$, the unique $v = [x, y]^{\mathrm{tr}} \in \mathbb{R}^{2 \times 1}$ fulfilling $Av = [c, d]^{\mathrm{tr}} \in \mathbb{R}^{2 \times 1}$ is given as $v = A^{-1} \cdot [c, d]^{\mathrm{tr}} = \frac{1}{150} \cdot \begin{bmatrix} -50 & 1 \\ 200 & -1 \end{bmatrix} \cdot [c, d]^{\mathrm{tr}} = \frac{1}{150} \cdot [-50c + d, 200c - d]^{\mathrm{tr}}$, which is also obtained from Cramer's rule as $x = \det(A)^{-1} \cdot \det\left(\begin{bmatrix} c & 1 \\ d & 50 \end{bmatrix}\right) = -\frac{1}{150} \cdot (50c - d)$ and $y = \det(A)^{-1} \cdot \det\left(\begin{bmatrix} 1 & c \\ 200 & d \end{bmatrix}\right) = -\frac{1}{150} \cdot (-200c + d)$.

**(4.7) Example: Vandermonde matrix.** Let $n \in \mathbb{N}_0$. For $a_1, \ldots, a_n \in \mathbb{R}$ let the associated **Vandermonde matrix** be defined as

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

Then we have $\det(A) = \Delta(a_1, \ldots, a_n) := \prod_{1 \le i < j \le n}(a_j - a_i) \in \mathbb{R}$, being called the associated **discriminant**; in particular, we have $A \in \mathrm{GL}_n(\mathbb{R})$ if and only if the $a_1, \ldots, a_n \in \mathbb{R}$ are pairwise distinct:

We proceed by induction; the cases $n \le 1$ being trivial, we let $n \ge 2$. 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

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

$\sharp$

**(4.8) Determinants of linear maps.** Let $V$ be an $\mathbb{R}$-vector space with finite $\mathbb{R}$-basis $B$, and let $\varphi\colon V \to V$ be $\mathbb{R}$-linear. Then $\det(\varphi) := \det({}_B\varphi_B) \in \mathbb{R}$ is independent from the $\mathbb{R}$-basis of $V$ chosen, and is called the **determinant** of $\varphi$: Indeed, if $C \subseteq V$ is an $\mathbb{R}$-basis, then we have $\det({}_C\varphi_C) = \det({}_C\mathrm{id}_B \cdot {}_B\varphi_B \cdot {}_B\mathrm{id}_C) = \det({}_B\mathrm{id}_C)^{-1} \cdot \det({}_B\varphi_B) \cdot \det({}_B\mathrm{id}_C) = \det({}_B\varphi_B)$.

In particular, in view of the geometric interpretation of determinants, if $V = \mathbb{R}^{n \times 1}$ and $B \subseteq \mathbb{R}^{n \times 1}$ is the standard $\mathbb{R}$-basis, then $\det(\varphi)$ is just the oriented volume of the image of the unit cube under $\varphi$, thus describes the change in volume application of the $\mathbb{R}$-linear map $\varphi$ entails.

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 ${}_C\mathrm{id}_B = ({}_B\mathrm{id}_C)^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. For the reflection $\sigma$ at the hyperplane perpendicular to $[-1,1]^{\mathrm{tr}}$ we get $\det(\sigma) = \det({}_B\sigma_B) = \det\left(\begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix}\right) = \det({}_C\sigma_C) = \det\left(\begin{bmatrix} 1 & . \\ . & -1 \end{bmatrix}\right) = -1$; for the rotation $\rho$ with angle $\alpha \in \mathbb{R}$ we get $\det(\rho) = \det({}_B\rho_B) = \det\left(\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\right) = \cos^2\alpha + \sin^2\alpha = 1$; and for the rotation-dilatation $\tau$ 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.

---

## 5 References

[1] S. Bosch: Lineare Algebra, 4. Auflage, Springer, 2008.

[2] B. Huppert, W. Willems: Lineare Algebra, 2. Auflage, Teubner, 2010.

[3] G. Stroth: Lineare Algebra, 2. Auflage, Berliner Studienreihe zur Mathematik 7, Heldermann, 2008.