Algebra

Arne Vorweg

December 17, 2025

Contents

1 Groups

1.1 Semigroups, monoids and groups

A **semigroup** $(G; \circ)$ is a pair of a **set** G and a **map** $\circ : G \times G \to G$ with

1. the **associative law** $(a \circ b) \circ c = a \circ (b \circ c)$ for every $a, b, c \in G$

$(G; \circ)$ is **abelian** resp. **regular** iff it satisfies

2. the **commutative law** $a \circ b = b \circ a$ for every $a, b \in G$.
3. the **division rule**: $a \circ b = a \circ c \Leftrightarrow a = c$.

A **semigroup** $(G; \circ)$ is a **monoid** iff it has

4. a **left neutral element** $e \in G$ such that $e \circ a = a$ for every $a \in G$

A **monoid** $(G; \circ)$ is a **group** iff it has

5. a **left inverse element** $a' \in G$ such that $a' \circ a = e$ for every $a \in G$

1.2 Translations

A **semigroup** $(G; \circ)$ is a **group** iff for every $a \in G$ the **left translation** $l_a : G \to G$ with $l_a(x) = a \circ x$ and the **right translation** $r_a : G \to G$ with $r_a(x) = x \circ a$ are both **surjective**. In that case they are both **injective** hence **bijective**.

1.3 Properties of a group

For elements $a, b, c \in G$ of a **group** G and **left neutral elements** $e; e_0$ resp. **left inverse elements** $a^{-1}; a_0^{-1}$ we have

1. **right inverse property**: $a \circ a^{-1} = e \circ a \circ a^{-1} = \left(a^{-1}\right)^{-1} \circ a^{-1} \circ a \circ a^{-1} = \left(a^{-1}\right)^{-1} \circ e \circ a^{-1} = \left(a^{-1}\right)^{-1} \circ a^{-1} = e$
2. **right neutral property**: $a \circ e = a \circ a^{-1} \circ a = e \circ a = e$
3. **uniqueness of the inverse**: $a_0^{-1} = e \circ a_0^{-1} = a^{-1} \circ a \circ a_0^{-1} = a^{-1} \circ e = a'$
4. **uniqueness of the** neutral element: $e_0 = e \circ e_0 = e$
5. **Division rule**: $a \circ b = a \circ c \Leftrightarrow b = a^{-1} \circ a \circ b = a^{-1} \circ a \circ c = c$ resp. $b \circ a = c \circ a \Leftrightarrow b = b \circ a \circ a^{-1} = b \circ a \circ a^{-1} = c$.

1.4 Direct products and subgroups

The **direct product** $\left(\prod_{i \in I} G_i; \circ\right)$ of groups $(G_i)_{i \in I}$ for any index set I refers to **componentwise composition** $(x_i)_{i \in I} \circ (y_i)_{i \in I} = (x_i \circ y_i)_{i \in I}$ on the **set theoretic product** $\prod_{i \in I} G_i = (x_i)_{i \in I} : I \to \bigcup_{i \in I} G_i : x_i \in G_i$. A **subgroup** $H \subset G$ is a group included in G. A set $H \subset G$ is a subgroup iff for every $a; b \in H$ we have $a \circ b^{-1} \in H$. Any set $S \subset G$ may be the **generator** of a subgroup $\langle S \rangle = \left\{ \prod_{i=1}^{n} x_i : x_i \vee x_i^{-1} \in S \forall 1 \le i \le n \in \mathbb{N} \right\}$. Obviously $\langle S \rangle$ is the **smallest** subgroup containing S and equal to the **intersection** of all such subgroups.

Examples: There are two non-abelian groups of order 8:

1. The **symmetry group of the square** is generated by the **rotation** $\sigma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and the **reflection** $\tau = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ such that $\sigma^4 = \tau^2 = e = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

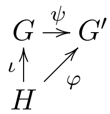
2. The **quaternion group** is generated by $\iota = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$ and $\kappa = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ with $\iota^4 = \kappa^4 = e$.

1.5 Homomorphisms

A mapping $\varphi : G \to G'$ between two groups $(G; \circ)$ and $(G'; \circ')$ is a **homomorphism** resp. **endomorphism** in the case of $G' = G$ iff $\varphi(a \circ b) = \varphi(a) \circ' \varphi(b)$ for every $a; b \in G$. The **left translation** l_a and the **right translation** r_a are homomorphisms iff $a = e$. The mapping $a \mapsto l_a$ is **always** a homomorphism but $a \mapsto r_a$ only iff G is **abelian**. The **composition** $\psi \circ \varphi : G \to G''$ of two homomorphisms $\varphi : G \to G'$ resp $\psi : G' \to G''$ is again a homomorphism. Since we have $\varphi(e) = e'$ and $\varphi(a^{-1}) = \varphi(a)^{-1}$ the **image** $\mathrm{Im}\varphi = \varphi[H] \subset G'$ as well as the **inverse image** $\varphi^{-1}[H'] \subset G$ of subgroups $H \subset G$ resp. $H' \subset G'$ under a homomorphism φ are again subgroups. A special case is the **kernel** $\ker\varphi = \varphi^{-1}[\{e'\}]$, i.e. the inverse image of the **trivial subgroup** $\{e\}$. A homomorphism φ is **injective** iff $\ker\varphi = \{e\}$. In this case it is also called an **embedding**. A **bijective** homomorphism is an **isomorphism** resp. an **automorphism** in the case of $G' = G$. The **bijections** on an arbitrary set X constitute the **symmetric group** $(S(X); \circ)$ with reference to the composition of mappings. Any subgroup of $S(X)$ is called a **permutation group**. In the case of a **group** G the family $\mathrm{Aut}G \subset S(G)$ of **automorphisms** on G is a subgroup of $S(G)$. In particular the **left translations** $l_a \in \mathrm{Aut}G$ are a **permutation group** and since $l : G \to \mathrm{Aut}G$ with $l(a) = l_a$ is an **isomorphism** we have **Cayley's theorem**: **Every group is isomorphic to a permutation group**.

1.6 Extension of semigroups to groups

For every **abelian** and **regular semigroup** H exists an abelian group G and an embedding $\iota : H \to G$ such that for every homomorphism $\varphi : H \to G'$ into an abelian group G' there is a unique homomorphism $\psi : G \to G'$ with $\psi \circ \iota = \varphi$.



Proof: The set $\sim = \{(a;b);(a';b') \in H^4 : \exists x \in H : ab'x = a'bx\}$ is an **equivalence relation** with obvious **reflexivity** resp. **symmetry** and **transitivity** since $a'bx = ab'x$ and $a''b'y = a'b''y$ imply $ab''(a'b'xy) = (ab'x)(a'b''y) = (a'bx)(a''b'y) = a''b(a'b'xy)$.

For $[a;b];[c;d] \in G = H \times H$ the mapping $\circ : [a;b] \circ [c;d] \mapsto [ac;bd]$ is **independent of the representants** since $ab'x = a'bx$ and $cd'y = c'dy$ imply $abc'd'(xy) = a'b'cd(xy)$. It is obviously **associative, commutative** with the **neutral element** is $0 = [a;a]$ and the **inverse** $[a;b]^{-1} = [b;a]$ for each $[a;b]$.

The mapping $\iota : H \to H^2$ with $\iota(a) = [a^2;a]$ is a **homomorphism** since $\iota(ab) = [a^2b^2;ab] = [a^2;a] \circ [b^2;b] = \iota(a) \circ \iota(b)$. In the case of **regularity** and owing to the **commutativity** we have $[a^2;a] = [b^2;b] \Leftrightarrow a^2bx = ab^2x \Leftrightarrow a(ab) = (ab)b \Leftrightarrow a = b$, i.e. ι is **injective**. It is also **surjective** since for every $[a;b] \in G$ we have $[a;b] = [a(ab);b(ab)] = [a^2;a] \circ [b;b^2] = \iota(a) \circ \iota(b)^{-1}$.

For a homomorphism $\varphi : H \to G'$ and $[a;b] \in G$ the mapping $\psi([a;b]) = \varphi(a) \circ' \varphi(b)^{-1}$ is a **homomorphism** since $\psi([a;b] \circ [c;d]) = \psi([ac;bd]) = \varphi(ac) \circ' \varphi(bd)^{-1} = \varphi(a) \circ' \varphi(c) \circ' \varphi(b)^{-1} \circ' \varphi(d)^{-1} = \psi([a;b]) \circ' \psi([c;d])$. It is **uniquely determined** since for every $[a;b] \in G$ the condition $(\psi \circ \iota)(x) = \varphi(x)$ implies $\psi([a;b]) = \psi\left(\iota(a) \circ \iota(b)^{-1}\right) = \varphi(a) \circ' \varphi(b)^{-1}$.

1.7 Index and order of a subgroup

Any subgroup $H \subset G$ of a group G defines an **equivalence relation** $a = b \bmod H \Leftrightarrow ab^{-1} \in H$. The **equivalence classes** $aH = l_a[H]$ or **left coset**s have the same **cardinality** or **order** $\mathrm{ord}H = (H:1)$ as H since the **left translation** l_a is bijective. The order $\mathrm{ind}H = (G:H) = (G/H:1)$ of the **quotient set** is called the **index** of H and in the case of two of theses indices being finite we have **Lagrange's theorem** $(G:H)(H:1) = (G:1)$.

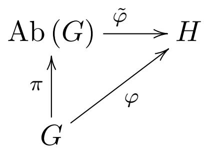
A second application of Lagrange's theorem to a further subgroup $K \subset H$ yields the **generalization** $(G : H)(H : K) = (G : K)$, cf. the **second isomorphism theorem** 1.12.

1.8 Normal subgroups

The composition $\circ$ extends to the **quotient set** G/H such that the **projection** $\pi : G \to G/H$ with $\pi(a) = a \circ H$ is a **homomorphism** iff $a \circ H \in G/H \Rightarrow a \circ H \circ a^{-1} \circ H = e \circ H \Leftrightarrow a \circ H \circ a^{-1} \in H$ $\Leftrightarrow a \circ H = H \circ a \Leftrightarrow \pi(a \circ b) = a \circ b \circ H = a \circ H \circ b \circ H = \pi(a) \circ \pi(b)\ \forall a; b \in G$. A subgroup satisfying this condition is **normal** and in this case the pair $(G/H; \circ)$ is the **factor group** with $\pi(a) \circ \pi(b) = a \circ H \circ b \circ H = a \circ b \circ H \circ H = a \circ b \circ H = \pi(a \circ b)$. In the following sections we abbreviate $ab = a \circ b$ if no ambiguity is caused.

1.9 Commutator subgroups

The **commutator subgroup** $[G, G] = \langle aba^{-1}b^{-1} : a; b \in G \rangle$ of a group G is a **normal** subgroup which assumes the trivial form iff G is abelian. The **abelianization** $\mathrm{Ab}(G) = G/[G, G]$ is the largest abelian quotient of G or equivalently the largest homomorphic image of G in the sense of its **universal property**: For any **abelian** group H and any homomorphism $\varphi : G \to H$ exists a unique homomorphism $\tilde{\varphi} : \mathrm{Ab}(G) \to H$ such that $\varphi = \tilde{\varphi} \circ \pi$ since in that case we have $\varphi(ab) = \varphi(ba) \Leftrightarrow \varphi\left((ba)^{-1}\right) = \varphi\left((ab)^{-1}\right) \Leftrightarrow \varphi(aba^{-1}b^{-1}) = \varphi(ab)\varphi\left((ba)^{-1}\right)$ $= \varphi(ab)\varphi\left((ab)^{-1}\right) = e$ which means that $G/R_\varphi = G/[G, g]$ whence $\tilde{\varphi} = \varphi \circ \pi^{-1}$ is well defined and satisfies the universal property. Isomorphic groups have isomorphic abelianizations because the isomorphism $F : G_1 \to G_2$ identifies $\mathrm{Ab}(G_2) = F[\mathrm{Ab}(G_1)]$ whence it descends to an isomorphism $\pi_2 \circ F \circ \pi_1^{-1} : \mathrm{Ab}(G_1) \to \mathrm{Ab}(G_2)$.



1.10 The fundamental theorem on homomorphisms

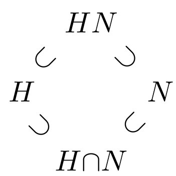
For every homomorphism $\varphi : G \to G'$

1. the **inverse image** $\varphi^{-1}[N'] \subset G$ of a **normal** subgroup $N' \subset G'$ is **normal** in G. In particular the **kernel** $\ker\varphi$ is **normal**.
2. If φ is **surjective** the **canonical injection** $\iota : G/\mathrm{Ker}\varphi \to G'$ with $\iota(a \circ \mathrm{Ker}\varphi) = \varphi(a)$ **is an isomorphism** and in that case the **image** $\varphi[N]$ of a **normal subgroup** N is **normal** in G'**.**

On account of $(xy)^{-1} = y^{-1}x^{-1}$ for any subset $S \subset G$ the **normalizer** $N_S = \{x \in G : xSx^{-1} = S\}$ and the **centralizer** $Z_S = \{x \in G : xsx^{-1} = s \forall s \in S\}$ are subgroups. The **center** Z_G is a **normal subgroup** and the normalizer N_H of a **subgroup** $H \subset G$ is the **largest subgroup** in which H is normal. Also in that case for any other subgroup $K \subset N_H$ the **product** KH is a group and H is normal in KH.

1.11 Noether's first isomorphism theorem

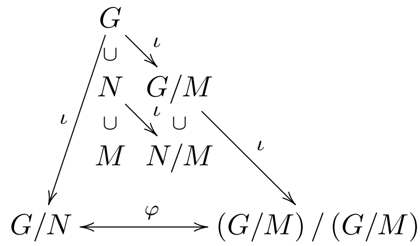
For every **subgroup** $H \subset G$ and every **normal subgroup** $N \subset G$.



1. the **product** HN is a subgroup of G.
2. N is a **normal subgroup** of HN.
3. $H \cap N$ is a **normal subgroup** of H.
4. the **injection** $\iota : H/(H \cap N) \to HN/N$ with $\iota(a(H \cap N)) = aN$ is an **isomorphism.**

1.12 Noether's second isomorphism theorem

For **normal subgroups** $M \subset N \subset G$



1. the **factor group** N/M is **normal** in G/M
2. the mapping

$$\varphi : (G/M) / (N/M) \to G/N$$

   with

$$\varphi ((aM) (N/M)) = aN$$

   is an **isomorphism**.

1.13 Cyclic groups

A single element $S = \{a\}$ generates a **cyclic group** $\langle a \rangle := \langle \{a\} \rangle = \{a^z : z \in \mathbb{Z}\}$ with the inductively defined **powers** $a^0 = e$, $a^{n+1} = a \circ a^n$ and $a^{-n} = \left(a^{-1}\right)^n$. A further induction yields $a^n \circ a^m = a^{n+m}$ such that for any **cyclic group** $\langle a \rangle$ of **order** $n = \text{ord} \langle a \rangle = \text{ord}a$ we have an **isomorphism** $\varphi : \mathbb{Z}/n\mathbb{Z} \to \langle a \rangle$ with $\varphi (m \bmod n) = a^m$. Hence every subgroup $H \subset \langle a \rangle$ contains a smallest $m \in \mathbb{N}$ with $a^m \in H$ and since there is no $m - k \in \mathbb{N}$with $a^{m-k} \in H$ we have $H = \langle a^m \rangle$. In particular from **Lagrange's theorem** we infer

1. $a^n = e \Leftrightarrow \text{ord}a|n$
2. **Every cyclic group is abelian.**
3. Every **subgroup** of a **cyclic** group is **cyclic.**
4. **Fermat's little theorem**: if $\text{ord}G < \infty$ for every $a \in G$ we have $a^{\text{ord}G} = e$.
5. If $\text{ord}G \in \mathbb{P}$ is a **prime number** for every $a \in G$ we also have $a^n \neq e$ for $n < \text{ord}G$. In that case G is **cyclic** and $G = \langle a \rangle$ for every $a \in G \setminus \{e\}$.
6. For every $a \in G$ with $\text{ord}G < \infty$ we have $\text{ord}a^m = \frac{\text{ord}a}{\text{GCD}(m;\text{ord}a)}$.
7. $\langle a \rangle = \langle b \rangle$ iff there is an $m \in \mathbb{N}$with $\text{GCD}\,(m; \text{ord}a) = 1$ and $b = a^m$.
8. For every $m \in \mathbb{N}$with $m|\text{ord}a$ resp. $\text{GCD}\,(m; \text{ord}a) = m$ there is a subgroup $\left\langle a^{\frac{\text{ord}a}{m}} \right\rangle \subset \langle b \rangle$.
9. Every group $G \neq \{e\}$ **without any subgroups** apart from $\{e\}$and G itself is of prime order $\text{ord}G \in \mathbb{P}$ and hence **cyclic**.

Proof of 1.13.6: Since there are coprime $m'; n'$ with $m = m' \cdot \text{GCD}\,(m; \text{ord}a)$ resp. $\text{ord}a = n' \cdot \text{GCD}\,(m; \text{ord}a)$ and $a^{m \cdot \text{ord}a^m} = e = a^{\text{ord}a}$ on the one hand we have an $n \geq 1$ with $m \cdot \text{ord}a^m = n \cdot \text{ord}a \Rightarrow m' \cdot \text{ord}a^m = n \cdot n' \Rightarrow n'|\text{ord}a^m$ and on the other hand $(a^m)^{n'} = a^{m \cdot n'} = a^{m' \cdot \text{GCD}(m;\text{ord}a) \cdot n'} = a^{m \cdot \text{ord}a} = e$ whence $\text{ord}a^m|n'$. This proves $\text{ord}a^m = n'$ and thus the assertion.

1.14 Operations and group actions

An **operation** is a homomorphism $\pi : G \to S\,(X)$ written $g \mapsto \pi_g : X \to X$ between a **group** G and the **symmetric group** of a set X. It can also be expressed both as a **right** rep. **left action** as a map $G \times X \to X$ written in the form $(g; x) \mapsto \pi_g\,(x)= x \cdot g = g^{-1} \cdot x$ with$\pi_{g_1 g_2} = \pi_{g_2} \circ \pi_{g_1}$ resp. $x \cdot (g_1 g_2) = (x \cdot g_1) \cdot g_2 = g_2^{-1} \cdot \left(g_1^{-1} \cdot x\right) = \left(g_1^{-1} g_2^{-2}\right) \cdot x = (g_2 g_1)^{-1} \cdot x$ and $\pi_e = \text{id}$ resp. $x \cdot e = e \cdot x = x$ for $g_1; g_2 \in G$ and $x \in X$. Usually the notation as a left action is preferred since it preserves the order of execution of the automorphisms $\pi_{g_2} \circ \pi_{g_1}$.

The **translation** $l : G \to S\,(G)$ with $l\,(a) = l_a : G \to G$ and $l_a\,(x) = ax$ due to $\text{Ker}l = l^{-1}\,(id) = l^{-1}\,(l_e) = \{e\}$ is an **injective** operation and observing that in general $(ab)\,x = a\,(bx)$ holds but **not**

$a(xy) = (ax)(ay)$ we note that l **is a homomorphism but** l_a **is not**. The group G may also operate by translation $l : G \to S(P(G))$ with $l(a) = l_a : P(G) \to P(G)$ and $l_a(H) = aH$ on its family $P(G)$ of subsets. Note again that even for a **subgroup** H in general the image aH will only be a **left coset**. Due to 1.7 the group G operates by translation on the **quotient set** G/H with $l(a) = l_a : G/H \to G/H$ and $l_a(xH) = axH$.

The **conjugation** $c : G \to S(G)$ is defined by $c(a) = c_a : G \to G$ with $c_a(x) = axa^{-1}$. The **inner automorphisms** $c_a \in \mathrm{Aut}G$ are **injective** but in general this is not true for c itself whose **kernel** $\mathrm{Ker}c = c^{-1}(\mathrm{id}) = Z_G$ is the **center** of G due to 1.10. As above the group G by conjugation also operates on the families of **subsets** resp. **subgroups** $H \in P(G)$. The resulting conjugations $c_a : P(G) \to P(G)$ with $c_a(H) = aHa^{-1}$ are obviously **bijective** and the **inverse image** $c_a^{-1}(H) = a^{-1}Ha = c_{a^{-1}}(H)$ is the **conjugate** of $ac_a^{-1}(H)a^{-1} = H$. The set $\{aHa^{-1} : a \in G\}$ is called the **conjugacy class** of the subset $H \subset G$.

1.15 Orbit and class formulae

For an operation $\pi : G \to S(X)$ and $x \in X$ the **isotropy group** is defined by $G_x = \{a \in G : \pi_a(x) = x\}$ such that its **kernel** is $\ker\pi = \pi^{-1}(\mathrm{id}) = \bigcap_{x\in X} G_x$. An $x \in X$ is a **fixed point** iff $G_x = G$ and π is **free** iff $G_x = \{e\}$. The **orbit** $\pi_G(x) = G \cdot x = \{\pi_a(x) : a \in G\} \subset X$ of $x \in X$ is a **group** with regard to $\pi_a(x) \bullet \pi_a(x) := \pi_{ab}(x)$ and π is **transitive** iff $G \cdot x = X$. The orbits are the elements of the **orbit space** X/G given by the **equivalence relation** $xRy \Leftrightarrow y \in G \cdot x$ which reduces to the single class $X/G = \{\bar{x}\}$ for any $x \in X$ iff π is **transitive**. Since for every $x \in X$ the map $\varphi : G/G_x \to \pi_G(x) \subset X$ with $\varphi(aG_x) = \pi_a(x)$ is **bijective** we have $\mathrm{ind}G_x = \mathrm{ord}\pi_G(x)$. Because $\pi_a(x) = \pi_b(y) \Leftrightarrow \pi_{b^{-1}a}(x) = y \Leftrightarrow y \in \pi_G(x)$ the orbits **partition** X and in the case of a **finite number** n **of orbits** we can choose $x_i \in X$ for $1 \le i \le n$ such that $i \ne j \Leftrightarrow \pi_G(x_i) \ne \pi_G(x_j)$ and $X = \mathring{\bigcup}_{1\le i\le n}\pi_G(x_i)$ yielding the **orbit decomposition formula** $\mathrm{card}X = \sum_{i=1}^{n} \mathrm{ord}\pi_G(x_i)$.

In the case of the **conjugation** and due to 1.10 the isotropy group coincides with the **normalizer** $G_x = N_x$ for $x \in X$ resp. $G_H = N_H$ for $H \in P(G)$. The **isotropy groups** of every pair $x; y \in X$ with $\pi_a(x) = y$ **are conjugate** with $aG_x = G_y a$ since $b \in G_y \Leftrightarrow \pi_b(y) = y \Leftrightarrow \pi_{a^{-1}ba}(x) = \pi_{a^{-1}}(y)$ $= x \Leftrightarrow a^{-1}ba \in G_x \Leftrightarrow b \in aG_xa^{-1}$. Because of $\pi_a(x) \ne \pi_b(x) \Leftrightarrow axa^{-1} \ne bxb^{-1} \Leftrightarrow b^{-1}ax \ne xb^{-1}a$ $\Leftrightarrow b^{-1}a \notin N_x \Leftrightarrow aN_x \ne bN_x$ we have $\mathrm{ord}\pi_G(x) = \mathrm{ind}N_x = \mathrm{ind}G_x$ such that in the case of the conjugation on a group G of finite order the **orbit decomposition formula** becomes the **class formula** $\mathrm{ord}G = \sum_{i=1}^{n} \mathrm{ind}N_{x_i}$ with $x_i \in X$ for $1 \le i \le n$ chosen such that $i \ne j \Leftrightarrow N_{x_i} = G_{x_i} \ne G_{x_j} = N_{x_j}$.

Examples:

1. The **general linear group** $\mathrm{GL}(n;\mathbb{R})$ acts **freely** on the left on $\mathbb{R}^n$ by scalar multiplication with two orbits $\mathbb{R}^n \setminus \{\mathbf{0}\}$ and $\{\mathbf{0}\}$.
2. The **orthogonal group** $\mathrm{O}(n;\mathbb{R})$ acts **freely** on the left on $\mathbb{R}^n$ by scalar multiplication with the orbits $r\mathbb{S}^{n-1}$ for $0 \le r < \infty$. Its restriction to $\mathbb{S}^{n-1}$ yields a **transitive** action.
3. The multiplicative abelian group $\mathbb{R}_*$ acts **freely** on $\mathbb{R}^n \setminus \{\mathbf{0}\}$ by scalar multiplication with the orbits $\{r\boldsymbol{x} : r \in \mathbb{R}_*\}$ for $\boldsymbol{x} \in \mathbb{R}^n \setminus \{0\}$.
4. The multiplicative abelian group $\{\pm 1\}$ acts **freely** on $\mathbb{S}^{n-1}$ by scalar multiplication with the orbits $\{\pm\boldsymbol{x}\}$ for $\boldsymbol{x} \in \mathbb{S}^{n-1}$.

1.16 Isotropy groups of transitive G-sets

For a **transitive right** G**-set** S endowed with a right transitive action by a group G holds $G_{s\cdot g} = g^{-1}G_s g$ for all $s \in S$ and $g \in G$ such that $\{G_s : s \in S\} = \{g^{-1}G_s g = G_{s\cdot g} : g \in G\}$ is a **uniquely determined conjugacy class** which is called the **isotropy type** of S.

Proof: For all $s \in S$ and $g \in G$ holds $g' \in G_{s \cdot g} \Leftrightarrow (s \cdot g) \cdot g' = s \cdot g \Leftrightarrow s \cdot (gg'g^{-1}) = s \Leftrightarrow gg'g^{-1} \in G_s$ $\Leftrightarrow g' \in g^{-1}G_s g$.

1.17 G-equivariant maps

A map $\varphi : S \to T$ between two right-transitive G-sets S and T is **G-equivariant** iff $\varphi(s) \cdot g = \varphi(s \cdot g)$ for every $s \in S$ and $g \in G$.

1. For any two G-equivariant maps $\varphi; \psi : S \to T$ holds $\varphi(s_0) = \psi(s_0)$ for some $s_0 \in S \Leftrightarrow \varphi \equiv \psi$.
2. Every G-equivariant map $\varphi : \emptyset \neq S \to T$ is **surjective**.
3. For every pair $(s_0; t_0) \in S \times T$ exists a uniquely defined G-equivariant $\varphi : S \to T$ with $\varphi(s_0)=$ t_0 iff $G_S \subset G_T$.

Proof:

1. By transitivity for every $s \in S$ exists a $g \in G$ with $s = s_0 \cdot g$ whence follows $\varphi(s) = \varphi(s_0 \cdot g) =$ $\varphi(s_0) \cdot g = \psi(s_0) \cdot g = \psi(s_0 \cdot g) = \psi(s)$.
2. Since S is nonempty and T is G-transitive for every $t \in T$ exist a $g \in G$ and an $s \in S$ with $t =$ $\varphi(s) \cdot g = \varphi(s \cdot g)$.
3. The necessity of the condition is evident by $s \cdot g = s \Rightarrow t \cdot g = \varphi(s) \cdot g = \varphi(s \cdot g) = \varphi(s) = t$. Conversely we define $\varphi : S \to T$ by $\varphi(s) = t_0 \cdot g_s$ with $g_s \in G$ such that $s = s_0 \cdot g_s$. This satisfies $\varphi(s_0) = t_0 \cdot e = t_o$ and due to $G_S \subset G_T$ the map is well defined since for any other g'_s with $s =$ $s_0 \cdot g_s$' follows $g_s (g'_s)^{-1} \in G_S \subset G_T$ which means $t_0 \cdot g_s = t_0 \cdot g'_s$. Concerning the G-invariance for any $s \in S$ and $h \in G$ holds $\varphi(s \cdot h) = \varphi(s_0 \cdot g_s h) = t_0 \cdot g_s h = (t_0 \cdot g) \cdot h = \varphi(s) \cdot h$.

1.18 The G-set isomorphism criterion

Two G-sets S and T are **G-isomorphic** iff there is a G-equivariant bijection $\varphi : S \to T$ which is called a **G-isomorphism**. Any pair of transitive right G-sets S and T

1. for a given pair $(s_0; t_0) \in S \times T$ by a uniquely determined **G-isomorphism** $\varphi : S \to T$ with $\varphi(s_0) = t_0$ is **G-isomorphic** iff $G_{s_0} = G_{t_0}$
2. are **G-isomorphic** iff they have the same **isotropy type**.

Proof:

1. For every $g \in G_{t_0}$we have $\varphi(s_0 \cdot g) = \varphi(s_0) \cdot g = t_0 \cdot g = t_0$ whence by the **injectivity** follows $s_0 \cdot g = s_0$ and thereby $g \in G_s$ or generally $G_t \subset G_s$. Equality then follows from the preceding result 1.17.3. Conversely the **injectivity** of the map φ defined in 1.17.3 follows since transitivity implies the existence of a $h; g_1; g_2 \in G$ with $\varphi(s_1) = \varphi(s_2) \Rightarrow t_0 = \varphi(s_0 \cdot g_1) \cdot h = \varphi(s_0 \cdot g_2) \cdot h$ $\Rightarrow t_0 \cdot (g_1 h) = t_0 \cdot (g_2 h) \Rightarrow t_0 = (t_0 \cdot (g_2 h)) \cdot (g_2 h)^{-1} = t_0 \cdot \left((g_2 h)(g_2 h)^{-1}\right) = t_0 \cdot (g_2 g^{-1}) \Rightarrow g_2 g_1^{-1}$ $\in G_{t_0} \subset G_{s_0} \Rightarrow s_2 = s_0 \cdot g_2 = s_0 \cdot g_1 = s_1$.
2. Assuming the existence of the unique g-isomorphism $\varphi : S \to T$ the first part 1.18.2 shows that for **every** s holds $G_s = G_{\varphi(s)}$, so the conjugacy classes must coincide. Conversely assuming that $G_s g = g G_t$ for every $s \in S$, $t \in T$ and $g \in G$ theorem 1.16 shows that we can choose a $t' \in T$ with $G_s = G_{t'}$ whence the first part 1.18.2 provides a G-isomorphism $\varphi : S \to T$ with $\varphi(s) = t'$.

1.19 The orbit criterion for G-automorphisms

The group of all **G-automorphisms** $\varphi : S \to S$ is called the **G-automorphism group** $\mathrm{Aut}_G(S)$. For any pair $(s; t) \in S^2$ in a G-transitive set S exists a unique G-automorphism $\varphi \in \mathrm{Aut}_G(S)$ with $\varphi(s) = t$ iff $G_s = G_t$.

Proof: Immediately follows from the preceding theorem 1.18.

1.20 The algebraic characterization of G-automorphism groups

For every point $s \in S$ in a G-**transitive** set S and every element $n \in N_G(G_s) = \{nG_sn^{-1} : n \in G\}$ of the **normalizer** in G of its **isotropy** group G_s exists a unique G-**automorphism** $\varphi : S \to S$ such that $\varphi_n(s) = s{\cdot}n$. The map $\varphi : N_G(G_s) \to \mathrm{Aut}_G(S)$ given by $\varphi(n) = \varphi_n$ is a **surjective homomorphism** with the **kernel** $\ker\varphi = G_s$ such that it descends to an **isomorphism** $\bar{\varphi} : N_G(G_s)/G_s \to \mathrm{Aut}_G(S)$.

Proof: The existence of the unique φ follows from the preceding result 1.19 with 1.10 and 1.16 since $G_s = nG_sn^{-1} = G_{s\cdot n}$. For $m; n \in N_G(G_s)$ and $s \in G_s$ we have $\varphi(mn)(s) = s \cdot mn = (s \cdot m) \cdot n = (\varphi(m) \circ \varphi(n))(s)$ which shows that φ is a **homomorphism**. Concerning **surjectivity** for every $\psi \in \mathrm{Aut}_G(S)$ and the given point $s \in S$ the **transitivity** of the action provides an element $g_s \in G$ such that $\psi(s) = s \cdot g_s$. Due to 1.19 we have $G_s = G_{s\cdot g_s} = g_sG_sg_s^{-1}$ which means $g_s \in N_G(G_s)$ whence the unique $\varphi_{g_s} \in \mathrm{Aut}_G(S)$ with $\varphi_{g_s}(s) = s \cdot g_s$ provided by 1.19 must coincide with ψ. Finally $\varphi_g = \mathrm{id}_S \Leftrightarrow \varphi_g(s) = s \Leftrightarrow s \cdot g = s \Leftrightarrow g \in G_s$ shows that $\ker\varphi = G_s$.

1.21 Permutations

A bijection $\sigma \in S(X)$ is a **cycle** of **length** $n \in \mathbb{N}$ iff there are $x_1; ...; x_n \in X$ with $x_{i+1} = \sigma(x_i) = ... = \sigma^{i-1}(x_1)$ for $1 \le i \le n-1$ resp. $x_1 = \sigma(x_n) = ... = \sigma^n(x_1)$ and $\sigma(x) = x$ for every other $x \in X \setminus \{x_1; ...; x_n\}$. The set of all cycles is $C(X) \subset S(X)$. In a simplified notation we only mention the elements $x_i \in X$ affected by σ and write $\sigma[X] = \{...; \sigma(x_1); ...; \sigma(x_n); ...\} = \langle x_1; ...; x_n\rangle = \langle x_1; \sigma^1(x_1); ...; \sigma^{n-1}(x_1)\rangle$, e.g. $\{1; 4; 3; 6; 5; 2; 7; 8; 9; 10\} = \langle 2; 4; 6\rangle$. Also if there is no ambiguity between the **mapping** σ and its **image** $\sigma[S]$ the argument may be suppressed and we write $\sigma = \sigma[S]$ as in e.g. $\sigma = \sigma[\{1; 2; 3; 4\}] = \{4; 3; 2; 1\}$. For every $1 \le j \le n$ the **image** $\langle x_1; ...; x_n\rangle = \langle \sigma^1(x_j); ...; \sigma^n(x_j)\rangle = \langle\sigma\rangle_{x_j}$ is the **orbit** of the **cyclic subgroup** generated by σ on the single element x_j. A cycle $\tau = \langle x_i; x_j\rangle$ of length $n = 2$ is a **transposition** τ with $\tau(x_i) = x_j$ and the set of all transpositions is $T(X) \subset C(X) \subset S(X)$. Note that neither set is closed under composition, e.g. $\langle 1; 2\rangle \circ \langle 3; 4\rangle = \{2; 1; 4; 3\}$ is not a cycle any more. (cf. 1.22.2) By $\mathbf{x} : \{1; ...; n\} \to \{x_1; ...; x_n\}$ with $\mathbf{x}(n) = x_n$ and $\sigma(x_n) = (\sigma \circ \mathbf{x})(n)$ every **symmetric group** $S(\{x_1; ...; x_n\})$ of order n is **isomorphic** to $S_n = S(\{1; ...; n\}) = S(\mathbf{x}[\{1; ...; n\}])$ such that any **permutation** $\sigma \in S_n$ may be expressed simply in the form $\sigma[1; ...; n] = \{\sigma(1); ...; \sigma(n)\}$. For example the **Klein Vierergruppe** $H = \langle \pi; \rho\rangle = \{\mathrm{id}; \pi; \rho; \pi \circ \rho\} \subset S_4$ with $\pi = \{2; 1; 4; 3\}$, $\rho = \{3; 4; 1; 2\}$ and $\pi \circ \rho = \rho \circ \pi = \{4; 3; 2; 1\}$ is an **abelian** subgroup of S_4.

1.22 The symmetric group

For $n \ge 2$ the symmetric group S_n has the following properties:

1. $S_n = \langle T_n\rangle$ is **generated** by the transpositions $T_n = \{\tau_{i;j} = \langle i; j\rangle : 1 \le i < j \le n\}$ and is of **order** $\mathrm{ord}S_n = n!$.
2. Every $\rho \in S_n$ is a finite product of disjoint cycles.
3. Disjoint cycles commutate: $\sigma \circ \rho = \rho \circ \sigma$ for every $\sigma, \rho \in C(X)$with $\sigma \cap \rho = \emptyset$.
4. Every $\pi \in S_n$ is a finite product of disjoint transpositions.

Proof:

1. Follows by **induction** from the observation that for $n \ge 2$ there are n transpositions $\tau_{i;n}$ and $S_n = \{\tau_{i;n} \circ \sigma_{n-1} : \sigma_{n-1} \in S_{n-1}\}$ whence $\mathrm{ord}S_n = n \cdot \mathrm{ord}S_{n-1}$.
2. Since with $x_i \in X$ for $1 \le i \le m \le n$ such that $i \ne j \Leftrightarrow \pi_{\langle\rho\rangle}(x_i) \ne \pi_{\langle\rho\rangle}(x_j)$ the orbits $\pi_{\langle\rho\rangle}(x_i) = \{\rho(x_i); ...; \rho^{m_i+1}(x_i) = x_i\} = \langle x_i; \rho(x_i); ...; \rho^{m_i}(x_i)\rangle$ resp. images of cycles $\sigma_i \in C(X)$ with $\sigma_i(x_i) = \rho(x_i)$ partition $X = \mathring{\bigcup}_{1\le i\le m}\pi_{\langle\rho\rangle}(x_i)$ such that for every $x \in X$ there is an $1 \le i \le m$ and a $1 \le j \le m_i$ such that $x = \sigma_i^j(x_i) \in \pi_{\langle\rho\rangle}(x_i)$ whence $\pi(x) = \pi \circ \sigma_i^j(x_i) = \sigma_i^{j+1}(x_i)$.
3. obvious.
4. Follows from 2. since for every cycle we have $\langle x_1; ...; x_n\rangle = \langle x_1; x_n\rangle \circ \langle x_1; x_{n-1}\rangle \circ ... \circ \langle x_1; x_2\rangle$.

1.23 The signum of a permutation

For every $n \geq 2$ the signum $\operatorname{sgn} : S_n \to \{\pm 1\}$ with $\operatorname{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(i)-\sigma(j)}{i-j}$ is a **homomorphism** with $\operatorname{sgn}(\sigma \circ \rho) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\rho)$ and $\operatorname{sgn}(\tau) = -1$ for every **transposition** τ. Hence $\operatorname{sgn}(\tau_1 \circ ... \circ \tau_n) = (-1)^n$ and for $A_n = \sigma^{-1}[\{1\}]$ the map $\sigma \mapsto \sigma \circ \tau$ is a **bijection** $A_n \to A_n \circ \tau$ such that from $S_n = A_n \mathring{\cup} A_n \circ \tau$ follows $|A_n| = |A_n \circ \tau| = \frac{1}{2} n!$.

Proof:

$$
\begin{aligned}
\operatorname{sgn}(\sigma \circ \rho) &= \prod_{i<j} \frac{(\sigma \circ \rho)(i) - (\sigma \circ \rho)(j)}{i-j} \\
&= \prod_{i<j} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \prod_{i<j} \frac{\rho(i) - \rho(j)}{i-j} \\
&= \prod_{\substack{i<j \\ \rho(i)<\rho(j)}} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \prod_{\substack{i<j \\ \rho(i)>\rho(j)}} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \epsilon(\rho) \\
&= \prod_{\substack{i<j \\ \rho(i)<\rho(j)}} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \prod_{\substack{j>i \\ \rho(j)<\rho(i)}} \frac{\sigma(\rho(j)) - \sigma(\rho(i))}{\rho(j) - \rho(i)} \cdot \epsilon(\rho) \\
&= \prod_{\substack{i<j \\ \rho(i)<\rho(j)}} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \prod_{\substack{i>j \\ \rho(i)<\rho(j)}} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \epsilon(\rho) \\
&= \prod_{\rho(i)<\rho(j)} \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \epsilon(\rho) \\
&= \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\rho).
\end{aligned}
$$

1.24 Free products

Similar to the polygonal presentation of compact surfaces in [6, p. 23.2] we define a **word** in a family of **groups** $(G_\alpha)_{\alpha \in A}$ as an ordered m-tuple $(g_1; ...; g_m)$ of group elements $g_i \in \bigsqcup_{\alpha \in A} G_\alpha$. By concatenation $(g_1; ...; g_m)(h_1; ...; h_n) = (g_1; ...; g_m; h_1; ...; h_n)$ the set $\mathcal{W}$ of all words becomes a **monoid** with the **neutral element** $()$. An **elementary reduction** is an operation of one of the following forms: $(g_1; ...; g_i; g_{i+1}; ...; g_m) \mapsto (g_1; ...; g_i g_{i+1}; ...; g_m)$ for $g_i; g_{i+1} \in G_\alpha$ or $(g_1; ...; g_{i-1}; e_\alpha; g_{i+1}; ...; g_m) \mapsto (g_1; ...; g_{i-1}; g_{i+1}; ...; g_m)$. Two words $W \sim W'$ are **equivalent** if they differ by a **finite sequence of elementary reductions** and the **free product** $\ast_{\alpha \in A} G_\alpha = \mathcal{W}/\sim$ is a **group**: Multiplication is **well defined** since $V \sim V'$ implies $VW \sim V'W$ as well as $WV \sim WV'$ whence by induction over the number of elementary reduction follows $V \sim V' \wedge W \sim W' \Rightarrow VW \sim V'W'$. **Associativity** and the **neutral element** carry over from $\mathcal{W}$ and the **inverse** of $(g_1; ...; g_m)$ is $\left(g_m^{-1}; ...; g_1^{-1}\right)$. In particular we have $(g_1; ...; g_m) = (g_1) ... (g_m)$ such that because the ordered character of the words generally implies $(g_i; g_{i+1}) \neq (g_{i+1}; g_i)$ **every free product of nontrivial groups is nonabelian and infinite**. E.g. In the free product $(\mathbb{Z}/2\mathbb{Z}) \ast (\mathbb{Z}/2\mathbb{Z}) = \langle g_1 \rangle \ast \langle g_2 \rangle$ we have $(g_1)(g_1 g_2) = (g_2)$ and $(g_1 g_2)(g_1) = (g_1 g_2 g_2)$.

1.25 Reduced words in free products

Every equivalence class of $\ast_{\alpha \in A} G_\alpha$ has a **unique** shortest representative which is called a **reduced word**.

Proof: Firstly by simple iteration over the concatenation over the length of a word $W \in \mathcal{W}$ we construct a **reduction operator** $\bullet : \mathcal{R} \times \mathcal{W} \to \mathcal{R}$ into the set $\mathcal{R}$ of reduced words: For length 0 we start with $R \bullet () = R$ for every reduced word $R \in \mathcal{R}$. Next, for length 1 and $R = (h_1; ...; h_k)$ with $g \in G_\alpha$ we define $(h_1; ...; h_k) \bullet (g) = \begin{cases} () & \text{for } k = 0 \text{ and } g = e_\alpha \\ (g) & \text{for } k = 0 \text{ and } g \neq e_\alpha \\ (h_1; ...; h_{k-1}) & \text{for } h_k \in G_\alpha \text{ and } h_k g = e_\alpha \\ (h_1; ...; h_k g) & \text{for } h_k \in G_\alpha \text{ and } h_k g \neq e_\alpha \\ (h_1; ...; h_k) & \text{for } h_k \notin G_\alpha \text{ and } g = e_\alpha \\ (h_1; ...; h_k; g) & \text{for } h_k \notin G_\alpha \text{ and } g \neq e_\alpha \end{cases}$.

Now we can perform the induction step for length $m > 1$ by letting $(h_1; ...; h_k) \bullet (g_1; ...; g_m) = (h_1; ...; h_k) \bullet (g_1; ...; g_{m-1}) \bullet (g_m) = (h_1; ...; h_k) \bullet (g_1) \bullet ... \bullet (g_m)$. Obviously for an already reduced $W \in \mathcal{R}$ nothing happens with $R \bullet W = RW$. For equivalent words of the form $W = (g_1; ...; g_i; g_{i+1}; ...; g_m)$ and $W' = (g_1; ...; g_i g_{i+1}; ...; g_m)$ with $g_i; g_{i+1} \in G_\alpha$ writing $R \bullet (g_1) \bullet ... \bullet (g_{i-1}) = (h_1; ...; h_k)$ we compute

$$(h_1; ...; h_k) \bullet (g_i) \bullet (g_{i+1}) = \begin{cases} () & \text{for } k = 0 \text{ and } g_i g_{i+1} = e_\alpha \\ (g_i g_{i+1}) & \text{for } k = 0 \text{ and } g_i g_{i+1} \neq e_\alpha \\ (h_1; ...; h_{k-1}) & \text{for } h_k \in G_\alpha \text{ and } h_k g_i g_{i+1} = e_\alpha \\ (h_1; ...; h_k g_i g_{i+1}) & \text{for } h_k \in G_\alpha \text{ and } h_k g_i g_{i+1} \neq e_\alpha \\ (h_1; ...; h_k) & \text{for } h_k \notin G_\alpha \text{ and } g_i g_{i+1} = e_\alpha \\ (h_1; ...; h_k; g_i g_{i+1}) & \text{for } h_k \notin G_\alpha \text{ and } g_i g_{i+1} \neq e_\alpha \end{cases} = (h_1; ...; h_k) \bullet (g_i g_{i+1})$$

whence follows $R \bullet W = R \bullet W'$. The same equation results in the case $W = (g_1; ...; g_{i-1}; e_\alpha; g_{i+1}; ...; g_m)$ and $W' = (g_1; ...; g_{i-1}; g_{i+1}; ...; g_m)$ from $R \bullet (e_\alpha) = R$. Hence the map $r : \mathcal{W} \to \mathcal{R}$ defined by $r(W) = () \bullet W$ satisfies $r(W) = W$ for every $W \in \mathcal{R}$and $r(W) = r(W')$ for every equivalent pair $W \sim W'$. For two equivalent reduced words $V \sim V' \in \mathcal{R}$ then follows equality by $V = r(V) = r(V')= V'$.

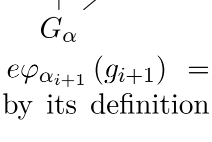
1.26 Components of free products

Each component G_α of a free product $\ast_{\alpha \in A} G_\alpha$ is **isomorphic** to a **subgroup** of the free product since the canonical map $\iota_\alpha : G_\alpha \to \ast_{\alpha \in A} G_\alpha$ defined by $\iota_\alpha(g) = (g)$ is a **homomorphism** due to $(gh) = (g)(h)$ and it is **injective** because the preceding result 1.25 implies $(g) = (h) \Rightarrow g = h$ for $g; h \neq e_\alpha$ resp. $\iota_\alpha(g) = () \Rightarrow g = e_\alpha$. Therefore we may identify words g with equivalence classes (g) and abbreviate $g_1 ... g_m \overset{3.6}{\cong} (g_1) ... (g_m) \overset{3.4}{=} (g_1; ...; g_m) \overset{3.6}{\cong} g_1; ...; g_m$.

1.27 The universal property of a free product

A group $\ast_{\alpha \in A} G_\alpha$ is a free product of groups $(G_\alpha)_{\alpha \in A}$ iff there is a family of homorphisms $\iota_\alpha : G_\alpha \to \ast_{\alpha \in A} G_\alpha$ such that for every family $(g_\alpha)_{\alpha \in A}$ of homomorphisms $\varphi_\alpha : G_\alpha \to H$ to a group H exists a unique homomorphism $\Phi : \ast_{\alpha \in A} G_\alpha \to H$ with $\varphi_\alpha = \Phi \circ \iota_\alpha$.

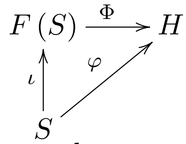
Proof:



$\Rightarrow$: The map $\Phi : \ast_{\alpha \in A} G_\alpha \to H$ defined by $\Phi(g_1; ...; g_m) = \varphi_{\alpha_1}(g_i) ... \varphi_{\alpha_m}(g_m)$ is **well defined** since for equivalent words differing by a **single elementary reduction** with $g_i; g_{i+1} \in G_\alpha$ we have $\Phi(g_i; g_{i+1}) = \varphi_\alpha(g_i) \varphi_\alpha(g_{i+1}) = \varphi_\alpha(g_i g_{i+1}) = \Phi(g_i g_{i+1})$ and with the neutral elements $e_\alpha \in G_\alpha$ resp. $e \in H$ we obtain $\Phi(g_{i-1}; e_\alpha; g_{i+1}) = \varphi_{\alpha_{i-1}}(g_{i-1}) \varphi_\alpha(e_\alpha) \varphi_{\alpha_{i+1}}(g_{i+1}) = \varphi_{\alpha_{i-1}}(g_{i-1}) e \varphi_{\alpha_{i+1}}(g_{i+1}) = \varphi_{\alpha_{i-1}}(g_{i-1}) \varphi_{\alpha_{i+1}}(g_{i+1}) = \Phi(g_{i-1}; g_{i+1})$. Obviously it is a **homomorphism** and by its definition satisfies $\varphi_\alpha = \Phi \circ \iota_\alpha$.

$\Leftarrow$: The hypothesis coincides with the **universal property** of the **categorical coproduct** from whence the group $\ast_{\alpha \in A} G_\alpha$ is the up to isomorphism uniquely defined coproduct of the objects $(G_\alpha)_{\alpha \in A} \subset \text{Grp}$.

1.28 Free groups



According to 1.4 every subset S of a group G generates the smallest group $\langle S\rangle$ including $S \subset G$ by taking all finite products of integer powers of elements of S. Analogously for an arbitrary object σ we can define a **cyclic free group** $F(\sigma) = \{\sigma\} \times \mathbb{Z}$ with $(\sigma; m)(\sigma; n) = (\sigma; m+n)$ and for an arbitrary set S the **free group on** S given by $F(S) = \ast_{\sigma\in S} F(\sigma)$ with the canonical injection $\iota : S \to F(S)$ given by $\iota(\sigma) = (\sigma; 1)$. Thus we may identify S with the subset $\iota[S]$ of $F(S)$ such that every element of $F(S)$ has a **unique expression as a reduced word** $\sigma_1^{n_1}...\sigma_k^{n_k}$ with $\sigma_i \in S$, $\sigma_i \neq \sigma_i^0$, $\sigma_i \neq \sigma_{i+1}$ and $n_i \in \mathbb{Z}\setminus\{0\}$ for $1 \leq i < k$ whence follows that $F(S)$ is a **normal group** and that the **center** of a free group on a set with more than one element **consists of the identity alone**. For convenience we also define the **free group of the empty set** as $F() = \{1\}$. The **cyclic character** of the components $F(\sigma)$ of a free group $F(S)$ allows the hypothesis of the the **universal property** 1.27 to be weakened as follows: For every map $\varphi : S \to H$ between an arbitrary set S and a group H exists a unique homomorphism $\Phi : F(S) \to H$ with $\varphi = \Phi \circ \iota$ defined by $\Phi((\sigma; m)) = (\varphi(\sigma))^m$ for $m \in \mathbb{Z}$ since $F(S)$ $= \ast_{\sigma\in S} F(\sigma)$ and $\varphi_\sigma : F(\sigma) \to H$ given by $\varphi_\sigma((\sigma; m)) = (\varphi(\sigma))^m$ is a homomorphism coinciding with $\varphi(\sigma) = \varphi_\sigma(\sigma)$ for every $\sigma \in S$. Generally a group G is a **free group** iff there is a subset $S \subset G$ satisfying one of the following two equivalent conditions:

1. The homomorphism $\Phi : F(S) \to G$ induced by the inclusion $\iota : S \to G$ via $\iota = \Phi \circ \iota$ is an isomorphism
2. Every $g \in G$ has a **unique** expression as a reduced product $g = \sigma_1^{n_1}...\sigma_k^{n_k}$ with $\sigma_i \in S$, $\sigma_i \neq \sigma_i^0$, $\sigma_i \neq \sigma_{i+1}$ and $n_i \in \mathbb{Z}\setminus\{0\}$ for $1 \leq i < k$.

Proof:

1. $\Rightarrow$ 2. : If ι extends to a homomorphism its image $\iota[S]$ has the form described in 2. and if the extension Φ is an isomorphism we have $\iota[S] = G$.

2. $\Rightarrow$ 1. : obvious

1.29 Group presentations

A **group presentation** is the **quotient group** $\langle S|R\rangle = F(S)/\bar{R}$ of a set of **generators** S with respect to the **normal closure** $\bar{R} \subset F(S)$, i.e. the intersection of all normal subgroups including R,of a subset of **relators** $R \subset F(S)$. For any generator $S \subset G$ with $\langle S\rangle = G$ the **unique homomorphism** $\Phi : F(S) \to G$ supplied by the **universal property** 1.27 must be **surjective** since every element of G can be expressed as a (not necessarily unique) product of integer powers of **generators**. If we demand every **relator** to be equivalent to **unity** in G the **kernel** of Φ must include R and since due to 1.10.1 the kernel is **normal** we conclude $\bar{R} \subset \ker\Phi$. Hence every group G generated by S such that all elements of R are equivalent to unity in G must be isomorphic to $F(S)/\ker\Phi$, i.e. to the free group on S divided by a normal subgroup including $\bar{R}$. Thus division by $\bar{R}$ obtains the **maximal** group generated by S with R equivalent to unity. A **presentation** of the group G is a **group presentation** $\langle S|R\rangle$ together with the **unique isomorphism** $\bar{\Phi} : \langle S|R\rangle \to G$. In the **finite** case it is explicitly written in the form $\langle s_1; ...; s_n|r_1; ...; r_m\rangle$. For convenience the **relations** to unity are spelled out in the form $\langle s_1; ...; s_n|r_1 = 1; ...; r_m = 1\rangle$ or sometimes even $\langle s_1; ...; s_n|r_1 = q_1; ...; r_m = q_m\rangle$ $= \left\langle s_1; ...; s_n|r_1q_1^{-1}; ...; r_mq_m^{-1}\right\rangle$.

Examples:

1. $F(\alpha_1; ...; \alpha_n) \cong \langle\alpha_1; ...; \alpha_n|\emptyset\rangle$. In particular $F(\alpha) \cong \langle\alpha|\emptyset\rangle \cong \langle\alpha\rangle \cong \mathbb{Z}$ is the **only free group which is abelian** since the isomorphism $\Phi : \mathbb{Z} \to F(\alpha)$ given by $\Phi(m) = (\alpha^m)$ satisfies $\alpha^m\alpha^k$ $= \Phi(m+k) = \Phi(k+m) = \alpha^k\alpha^m$.
2. $\mathbb{Z}^n \cong \mathrm{Ab}(F(\alpha_1; ...; \alpha_n)) = \langle\alpha_1; ...; \alpha_n|\alpha_i\alpha_j = \alpha_j\alpha_i : 1 \leq i; j \leq n\rangle$ is isomorphic to the **abelianization** of the free group $F(\alpha_1; ...; \alpha_n)$ since the homomorphism $\Phi : F(\alpha_1; ...; \alpha_n) \to \mathbb{Z}^n$ given by $\Phi(\alpha_i) = \boldsymbol{e}_j$ implies $\Phi\left(\alpha_{i_1}^{k_1(i_1)}...\alpha_{i_m}^{k_m(i_m)}\right) = \left(\sum_{i_j=1} k_j(i_j); ...; \sum_{i_j=n} k_j(i_j)\right)$ whence $\Phi(\alpha\beta) =$

$\Phi(\beta\alpha) \Leftrightarrow \Phi\left(\alpha\beta\alpha^{-1}\beta^{-1}\right) = \Phi(\alpha\beta) + \Phi\left(\alpha^{-1}\beta^{-1}\right) = \Phi(\alpha\beta) - \Phi(\beta\alpha) = 0 \Leftrightarrow \alpha\beta\alpha^{-1}\beta^{-1} \in \ker\Phi$ for every $\alpha, \beta \in F(\alpha_1; ...; \alpha_n)$ such that we obtain the **injective** canonical homomorphism $\bar{\Phi} = \Phi \circ \pi^{-1} : F(\alpha_1; ...; \alpha_n) / \ker\Phi = \langle \alpha_1; ...; \alpha_n | \alpha\beta = \beta\alpha \, \forall \alpha, \beta \in F(\alpha_1; ...; \alpha_n) \rangle = \langle \alpha_1; ...; \alpha_n | \alpha_i\alpha_j = \alpha_j\alpha_i : 1 \leq i; j \leq n \rangle \to \mathbb{Z}^n$. To prove its **surjectivity** y $\bar{\Psi}(m_1; ...; m_n) = \alpha_1^{m_1} ... \alpha_n^{m_n}$ we define a map $\bar{\Psi} : \mathbb{Z}^n \to \langle \alpha_1; ...; \alpha_n | \alpha_i\alpha_j = \alpha_j\alpha_i : 1 \leq i; j \leq n \rangle$ which by the **abelian** character of its domain is a homomorphism and also an inverse to $\bar{\Phi}$ since obviously $\bar{\Phi} \circ \bar{\Psi} = \mathrm{id}$ and $\bar{\Psi} \circ \bar{\Phi} = \mathrm{id}$.

3. $\mathbb{Z}/n \cong \langle \alpha | \alpha^n = 1 \rangle$ with the **quotient group** $\mathbb{Z}/n = \mathbb{Z} \bmod n = \mathbb{Z}/n\mathbb{Z}$ because the identity $\Phi[n\mathbb{Z}] = \alpha^{n\mathbb{Z}}$ of he isomorphism $\Phi : \mathbb{Z} \to F(\alpha)$ from 1.29.1 implies that $\Phi \circ \pi^{-1} : \mathbb{Z}/n \to F(\alpha)/\alpha^{n\mathbb{Z}} = \langle \alpha | \alpha^n = 1 \rangle$ is also an isomorphism.

4. $\mathbb{Z}/m \times \mathbb{Z}/n \cong \langle \alpha\beta | \alpha^m = 1; \beta^n = 1; \alpha\beta = \beta\alpha \rangle$ because the identity $\Phi[m\mathbb{Z} \times n\mathbb{Z}] = \alpha^{m\mathbb{Z}} \times \beta^{n\mathbb{Z}}$ of he isomorphism $\Phi : \mathbb{Z} \times \mathbb{Z} \to F(\alpha; \beta)$ from 1.29.2 implies that $\Phi \circ \pi^{-1} : (\mathbb{Z} \times \mathbb{Z}) / (m\mathbb{Z} \times n\mathbb{Z}) \to F(\alpha; \beta) / \left(\alpha^{m\mathbb{Z}} \times \beta^{n\mathbb{Z}}\right)$ is also an isomorphism and $(\mathbb{Z} \times \mathbb{Z}) / (m\mathbb{Z} \times n\mathbb{Z}) \cong \mathbb{Z}/m \times \mathbb{Z}/n$ by $\pi(z_\alpha; z_\beta) = (z_\alpha; z_\beta) + (m\mathbb{Z}; n\mathbb{Z}) \mapsto (z_\alpha + m\mathbb{Z}; z_\beta + n\mathbb{Z}) = (\pi(z_\alpha); \pi(z_\beta))$.

5. For homomorphisms $f_i : G_i \to H_i$ between groups G_i and H_i for $i \in \{1; 2\}$ exists a unique homomorphism $f_1 * f_2 : G_1 * G_2 \to H_1 * H_2$ defined by $(f_1 * f_2)(g_i) = f_i(g_i)$ for $g_i \in G_i$ such that the diagram on the right hand side with the canonical injections $\iota_i : G_i \to G_1 * G_2$ and $\iota_i' : H_i \to H_1 * H_2$ commutes.

$$\begin{array}{ccc} G_1 * G_2 & \xrightarrow{f_1 * f_2} & H_1 * H_2 \\ \uparrow \iota_i & & \uparrow \iota_i' \\ G_i & \xrightarrow[f_i]{} & H_i \end{array}$$

5. For **generator** sets S_i and **relator** sets $R_i \subset F(S_i)$ for $i \in \{1; 2\}$ the **free product** $\langle S_1 | R_1 \rangle * \langle S_2 | R_2 \rangle$ has the presentation $\langle S_1 \cup S_2 | R_1 \cup R_2 \rangle$ because the union $\bar{R}_1 \cup \bar{R}_2$ of the **normal** subgroups $\bar{R}_1$ and $\bar{R}_2$ is still a **normal** set whence $\bar{R}_1\bar{R}_2 = \left\{\prod_{k=1}^m r_{1,k}^{\alpha_k} \prod_{\ell=1}^n r_{2,\ell}^{\alpha_\ell} : r_{1,k} \in \bar{R}_1; r_{2,\ell} \in \bar{R}_2; \alpha_k, \alpha_\ell \in \mathbb{Z}_*; m, n \in \mathbb{N}_*\right\} = \left\{\prod_{k=1}^m r_k^{\alpha_k} : r_k \in \bar{R}_1 \cup \bar{R}_2; \alpha_k \in \mathbb{Z}_*; m \in \mathbb{N}_*\right\} = \left\langle \bar{R}_1 \cup \bar{R}_2 \right\rangle = \overline{R_1 \cup R_2}$.

6. For any two subsets $R; R' \subset F(S)$ of the free group generated by a set S the quotient group $\langle S | R \rangle / \overline{\pi[R']}$ with the canonical projection $\pi : F(S) \to \langle S | R \rangle$ has the presentation $\langle S | R \cup R' \rangle$, because for any pair $\sigma; \rho \in F(S)$ with $\sigma\rho^{-1} \in \overline{R \cup R'} = \left\langle \bar{R} \cup \bar{R}' \right\rangle$ the **normal** character of $\bar{R}$ and $\bar{R}'$ implies $\sigma\rho^{-1} = \prod_{k=1}^m r_k^{\alpha_k} \prod_{\ell=1}^n r_\ell'^{\alpha_\ell}$ for some $r_k \in \bar{R}; r_\ell' \in \bar{R}'; \alpha_k, \alpha_\ell \in \mathbb{Z}_*; m, n \in \mathbb{N}_*$ or equivalently $\sigma\rho^{-1} \prod_{\ell=1}^n r_\ell'^{-\alpha_\ell} = \prod_{k=1}^m r_k^{\alpha_k}$. For $\pi(\sigma); \pi(\rho^{-1}) \in \langle S | R \rangle$ this means $\pi(\sigma)\pi(\rho^{-1}) = \pi(\sigma\rho^{-1}) \in \pi\left[\bar{R}'\right] = \overline{\pi[R']}$ since due to 1.8 the projection $\pi\left[\bar{R}'\right] \subset F(S)/\bar{R}$ is already a normal subgroup.

The uniqueness of the group representation is far from solved: In the 1950s it was shown that there is no algorithm for solving either the **isomorphism problem** or the **word problem** formulated around 1910 by Max Dehn and Heinrich Tietze for every group representation in a finite amount of time. The former stands for the question whether the groups generated by two finite group presentations are isomorphic and the latter asks to decide whether two words formed from elements in S are equal in $\langle S | R \rangle$.

1.30 Free abelian groups

The analogue of free groups in the category of abelian groups are described in the additive style so that the **ordered finite sequences** or **words** $\sigma_1^{n_1} ... \sigma_k^{n_k} = \prod_{1 \leq i \leq k} \sigma_i^{n_i} \in F(S) \subset \mathbb{Z}^S$ with $\sigma_i \in S$, $\sigma_i \neq \sigma_{i+1}$ and $n_i \in \mathbb{Z} \setminus \{0\}$ for $1 \leq i < k$ resp. finite products of integer powers of generators $\sigma_i \in S$ in the free group $F(S)$ are replaced by **unordered finite collections** or **formal linear combinations** of the form $\sum_{1 \leq i \leq k} n_i\sigma_i \in \mathbb{Z}S$ with with $\sigma_i \in S$, $\sigma_i \neq \sigma_{i+1}$ and $n_i \in \mathbb{Z} \setminus \{0\}$ for $1 \leq i < k$. The resulting **free abelian group** is denoted by $\mathbb{Z}S$ with $\mathbb{Z}\emptyset = F(\emptyset) = \{0\}$. The **universal property** is reduced to the statement that any map $\varphi : S \to G$ from a set of generators S into an **abelian** group G extends to a **unique homomorphism** $\Phi : \mathbb{Z}S \to G$ given by $\Phi\left(\sum_{1 \leq i \leq k} n_i\sigma_i\right) = \sum_{1 \leq i \leq k} n_i\varphi(\sigma_i)$. Note that the **abelian** character expressed by $\sigma_i + \sigma_{i+1} = \sigma_{i+1} + \sigma_i$ in the formal linear combinations of $\mathbb{Z}S$ can only be preserved in an **abelian** target group G with $\varphi(\sigma_i) + \varphi(\sigma_{i+1}) = \varphi(\sigma_{i+1}) + \varphi(\sigma_i)$.

In the case of a finite set of generators $S = \{\sigma_1; ...; \sigma_k\}$ and $G = \mathbb{Z}^k$ we even obtain an **isomorphism** $\Phi : \mathbb{Z}\{\sigma_1; ...; \sigma_k\} \to \mathbb{Z}^k$ defined by $\Phi\left(\sum_{1\leq i\leq k} n_i\sigma_i\right) = (n_1; ...; n_k)$. A nonempty subset $S \subset G$ of an abelian group G is **linearly independent** iff $\sum_{1\leq i\leq k} n_i\sigma_i = 0 \Leftrightarrow n_1 = ... = n_k = 0$. If S also generates G with $G = \langle S | \sigma_i + \sigma_{i+1} = \sigma_{i+1} + \sigma_i : 1 \leq i < k\rangle = \mathbb{Z}S$ it is a **basis** for G and every element $g \in G$ has a **unique linear combination of elements** of S in analogy to the **reduced words** in a free group whence an **abelian group is free abelian iff it has a basis.** E.g. the familiar basis $(\boldsymbol{e}_i)_{1\leq i\leq k}$ of the **vector space** $\mathbb{R}^k$ is also the **standard basis** for the free abelian group $\mathbb{Z}^k$. As mentioned in 1.29.1 apart from the trivial group $\mathbb{Z}\emptyset = F(\emptyset) = \{0\}$ the **integers** $\mathbb{Z} \cong \mathbb{Z}\{\alpha\} = F(\{\alpha\})$ are the **only abelian group which is also a free group**. Every abelian group G with a **finite basis** has a minimal basis whose cardinality is the **rank of** G since every element g_i of a longer basis $\{g_1; ...; g_{k+1}\}$ can be expressed as a unique linear combination $g_i = n^{ij}e_j$ of elements of a minimal basis $\{e_j; ...e_k\}$ so that it can also be expressed as the corresponding linear combination of $\{g_1; ...; g_{k+1}\}$, e.g. $g_{k+1} = n^{k+1;j}e_j = m^i g_i$ with $(m_i)_{1\leq k\leq k} = \boldsymbol{m} = N^{-1} * \boldsymbol{n}_{k+1}$ for the transformation matrix $N = (n^{i;j}) \in GL(n;\mathbb{Z})$ provided that $g_{k+1} \neq e_j$ for $1 \leq j \leq k$.

Theorem:

1. Every subgroup of a free abelian group of finite rank is free abelian of rank less than or equal to that of G.
2. Every subgroup of a finitely generated abelian group is itself finitely generated.

Proof:

1. W.l.o.g. we assume $G = \mathbb{Z}^n$and proceed by induction starting with the cyclic group $\mathbb{Z}$ of rank 1 and referring to 1.13.3. Assuming the hypothesis for $\mathbb{Z}^{n-1} \cong \mathbb{Z} \times ... \times \mathbb{Z} \times \{0\} \subset \mathbb{Z}^n$ the intersection $H \cap \mathbb{Z}^{n-1}$ with an arbitrary subgroup $H \subset G$ must be abelian with rank $m - 1 \leq n - 1$ such that it has a basis $\{h_1; ...; h_{m-1}\}$. Then the **cyclic** subgroup $H \cap (\{0\} \times ... \times \{0\} \times \mathbb{Z}) \cong H \cap \mathbb{Z} = \langle c\rangle$ for some $c \in \mathbb{Z}$ must have a generator $(0; ...; 0; c)$ which is the n-th projection of an element $h_m \in H$, i.e. $\pi_n(h_m) = c \in \mathbb{Z}$. Applying π_n to the equation $a_1h_1 + ... + a_mh_m = 0$ yields $a_mc = 0$, hence $a_m = 0$. Since $\{h_1; ...; h_{m-1}\}$ is a basis, we conclude $a_1 = ... = a_{m-1} = 0$ whence $\{h_1; ...; h_m\}$ is linearly independent. For every $h \in H$ exists an integer $z \in \mathbb{Z}$ with $\pi_n(h) = zc$, so that $h - zh_m \in H \cap \mathbb{Z}^{n-1}$ is a lnear combination of $\{h_1; ...; h_{m-1}\}$ which shows that $\{h_1; ...; h_m\}$ is a basis of H and proves the proposition.
2. By the hypothesis and the universal property of free abelian groups the inclusion $\iota : S \to G$ extends to a surjective homomorphism $I : \mathbb{Z}S \to G$. The preimage $I^{-1}[H] \subset \mathbb{Z}S$ of any subgroup $H \subset G$ is a subgroup og $\mathbb{Z}S$ so that the preceding result 1.30.1 implies the existence of a basis for $I^{-1}[H]$ which is taken to a (not necessarily linearly independent) finite set of generators for H.

1.31 Torsion subgroups

An element g of an abelian group G is a **torsion element** if $zg = 0$ for some $z \in \mathbb{Z} \setminus \{0\}$. If $zg = z'g' = 0$ then $zz'(g + g') = 0$ so the set of a torsion elements G_{tor} is a group, the **torsion subgroup** of G. G is **torsion-free** iff the only torsion element of G is 0. Obviously G/G_{tor} is always torsion-free.

Theorem: A finitely generated torsion-free abelian group is free abelian of finite rank.

Examples: The rank of $\mathbb{Z}^n$ is n and the rank of every finite abelian group is obviously 0. The rank of a product of the form $G = \mathbb{Z}^n \times \mathbb{Z}/k_1 \times ... \times \mathbb{Z}/k_m$ is n, because $G_{tor} = \{(0; ...; 0)\} \times \mathbb{Z}/k_1 \times ... \times \mathbb{Z}/k_m$.

Proof: By induction over the cardinality of the generator T of G we prove the existence of a finite linearly independent set $S \subset G$ and a $z \in \mathbb{Z}_*$ such that $zG \subset \langle S\rangle$: The start for $T = \{g_1\}$ follows from $G = \langle g_1\rangle$, which means $z = 1$. If $T = \{g_1; ...; g_m\}$ is not already linearly independent w.l.o.g. we can assume $a_1; ...; a_m \in \mathbb{Z}$ with $a_1g_1 + ... + a_{m-1}g_{m-1} = a_mg_m$. The hypothesis provides for a linearly independent set $S \subset \langle g_1; ...; g_{m-1}\rangle$ and a $z' \in \mathbb{Z}_*$ such that $z'(a_1g_1 + ... + a_{m-1}g_{m-1}) = z'a_mg_m \in \langle S\rangle$. Then for $z = z'a_m$ and every $g = b_1g_1 + ... + b_mg_m \in G$ we obtain $zg = z'a_m(b_1g_1 + ... + b_{m-1}g_{m-1})$

$+ z'a_m b_m g_m \in \langle S\rangle$. The homomorphism $t_z : G \to G$ given by $t_z(g) = zg$ is **injective** because G is **torsion-free** and the result of the induction above implies $t_z[G] \subset \langle S\rangle$ such that the assertion follows from 1.30.1.

1.32 The rank-nullity law

An abelian group G is finitely generated iff both $\mathrm{Im}f$ and $\mathrm{Ker}f$ of a homomorphism $f : G \to H$ into another abelian group H are finitely generated and in that case holds an analogue to the **rank-nullity law** for vector spaces from 3.8: $\mathrm{rank}G = \mathrm{rank}(\mathrm{Im}f) + \mathrm{rank}(\mathrm{Ker}f)$.

Proof: W.l.o.g. we can assume that $f : G \to \mathrm{Im}f \subset H$ is **surjective**. By 1.30.2 the subgroup $\mathrm{Ker}f \subset G$ is finitely generated and so is $\mathrm{Im}f = \langle f[T]\rangle$ for $G = \langle T\rangle$. Conversely for finitely generated $\mathrm{Ker}f = \langle g_1; ...; g_p\rangle$ resp. $\mathrm{Im}f = \langle f(h_1); ...; f(h_q)\rangle$ and an arbitrary $g \in G$ we have $f(g) = z_1 f(h_1) + ... + z_q f(h_q) = f(z_1 h_1 + ... + z_q h_q)$ whence follows $g = z_1 h_1 + ... + z_q h_q + g_0$ with $g_0 = b_1 g_1 + ... + b_p g_p \in \mathrm{Ker}f$ such that $G = \langle g_1; ...; g_p; h_1; ...; h_q\rangle$. This result already implies the **rank-nullity law** in the case of **free abelian** $\mathrm{Im}f$ and G whence $\mathrm{Ker}f$ is also free abelian by 1.30.1. In the general case we note that f preserves the **torsion group** whence it descends to a surjective homomorphism $\bar{f} : G/G_{tor} \to \mathrm{Im}f/(\mathrm{Im}f)_{tor}$ with $\mathrm{Ker}f/(\mathrm{Ker}f)_{tor} = \mathrm{Ker}f/(\mathrm{Ker}f \cap G_{tor}) \subset \mathrm{Ker}\bar{f} \subset G/G_{tor}$. Because the latter is free abelian it follows from 1.30.1 that both $\mathrm{Ker}f/(\mathrm{Ker}f)_{tor}$ and $\mathrm{Ker}\bar{f}$ are free abelian with $\mathrm{rank}(\mathrm{Ker}f/(\mathrm{Ker}f)_{tor}) \leq \mathrm{rank}\left(\mathrm{Ker}\bar{f}\right)$. Because $(\mathrm{Im}f)_{tor}$ is a finitely generated torsion group there is a common integer $z \in \mathbb{Z}_*$ such that $zt = 0$ for **every** $t \in (\mathrm{Im}f)_{tor}$. Thus for any $\bar{g} \in \mathrm{Ker}\bar{f} \subset G/G_{tor}$, it follows that $zf(g) = f(zg) = 0$, i.e. $zg \in \mathrm{Ker}f$,and therefore $z\bar{g} = \overline{zg} \in \mathrm{Ker}f/(\mathrm{Ker}f)_{tor} \subset \mathrm{Ker}\bar{f}$. Thus the **injective** homomorphism $t_z : G/G_{tor} \to G/G_{tor}$ given by $t_z(\bar{g}) = z\bar{g}$ maps $\mathrm{Ker}\bar{f}$ into $\mathrm{Ker}f/(\mathrm{Ker}f)_{tor}$ whence by 1.30.1 follows $\mathrm{rank}\left(\mathrm{Ker}\bar{f}\right) \leq \mathrm{rank}(\mathrm{Ker}f/(\mathrm{Ker}f)_{tor})$ and by the preceding result **equality**. By 1.31 we can apply the same argument as above to obtain $\mathrm{rank}G = \mathrm{rank}(G/G_{tor}) = \mathrm{rank}(\mathrm{Im}f/(\mathrm{Im}f)_{tor}) + \mathrm{rank}\left(\mathrm{Ker}\bar{f}\right) = \mathrm{rank}(\mathrm{Im}f/(\mathrm{Im}f)_{tor}) + \mathrm{rank}(\mathrm{Ker}f/(\mathrm{Ker}f)_{tor}) = \mathrm{rank}(\mathrm{Im}f) + \mathrm{rank}(\mathrm{Ker}f)$.

2 Rings

2.1 Rings

A **ring** $(R; +\cdot)$ is a triple of a **set** R and two **maps** $+; \cdot : R \times R \to R$ iff for every $a; b; c \in R$ we have:

1. $(R; +)$ is an **abelian group**
2. **associativity** of the **multiplication :** $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
3. **distributiv**ity $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$
4. a **unit element** $1 \in R$ with $1 \cdot a = a \cdot 1 = a$
5. A ring is **commutative** iff this property holds for the multiplication.
6. The unit elemt is **uniquely determined** and in the case of $0 = 1$we have $R = \{0\}$
7. The ring is an **integral domain** iff $a \cdot b = 0 \Leftrightarrow a = 0 \vee b = 0$ for every $a; b \in R$, i.e. it is free (*nullteilerfrei*) of **left** resp. **right zero divisors** $a \in R : \exists 0 \neq b \in R : a \cdot b = 0$ resp $b \cdot a = 0$.
8. $0 \cdot a = a \cdot 0 = 0$ since $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$ and vice versa.
9. $a \cdot (-b) = -(a \cdot b)$ and $(-a) \cdot (-b) = a \cdot b$ since $a \cdot (-b) + a \cdot b = a(-b + b) = a \cdot 0 = 0$.

2.2 Examples

1. The set of maps $G \to G$ on an **additive group** $(G; +)$ forms a **ring** with respect to the **composition**. Special cases are the **endomorphisms** $(\mathrm{End}G; +; \circ)$ on G , **linear maps** $(\mathrm{L}(X); +; \circ)$ on

a complex **vector space** X and in the finite dimensional case the isomorphic set $(\mathrm{M}(n;\mathbb{C});+;*)$ of complex **quadratic matrices** with respect to the matrix product.

2. The set of maps $R \to R$ on a **ring** $(R;+;\cdot)$ forms a **ring** with respect to the **multiplication** with the **unit** $1 : r \to 1$ for $r \in R$ as well as with respect to the **composition** with the **unit** $\mathrm{id} : R \to R$.

3. The set $\left(L^1(\mathbb{C});+;*\right)$ of **Lebesgue integrable** complex functions with respect to the **convolution** is **not a ring** since the **convolution lacks a neutral element**.

2.3 Ideals

A **subring** $I \subset R$ of a ring $(R;+;\cdot)$ is a **left** resp. **right ideal** iff $RI \subset I$ resp. $IR \subset R$ whence $RI = I$ resp. $IR = I$ since $1 \in R$. Only for a **two-sided** or simply **ideal** I the **factor group** $(R/I;+)$ with the multiplication $(r+I)\cdot(s+I) = \pi(r)\cdot\pi(s) = \pi(rs) = rs + I$ becomes a **factor ring** $(R/I;+;\cdot)$ since only in that case $r' = r \bmod I$ and $s' = s \bmod I$ satisfy $r's' - rs = r'(s'-s) - (r'-r)s \in I$ whence $\pi(r's') = \pi(rs) \bmod I$ while the **associativity** resp. the **distributivity** obviously extend from $(R;+;\cdot)$ to $(R/I;+;\cdot)$.

The simplest ideals are the **left principal ideals** Ri for any **generator** $i \in R$. This can be extended to finitely many generators $(i_k)_{1\le k\le n} \subset R$ such that $\sum\limits_{k=1}^{n} Ri_k = \left\{\sum\limits_{k=1}^{n} r_k i_k : r_k \in R\,\forall 1 \le k \le n\right\}$ and likewise for **right principal** resp. **two-sided principal ideals**. In the latter case we use the notation $\langle i_k\rangle_{1\le k\le n} = \sum\limits_{k=1}^{n} Ri_k = \sum\limits_{k=1}^{n} i_k R$. A **commutative nontrivial** ring is **principal** iff every ideal is principal.

The **sum** $I + J = \{i + j : i \in I \wedge j \in J\} = I \cup J$ of two ideals is again an **ideal**. In general this is not true for the **product** $IJ = \{ij : i \in I \wedge j \in J\} \subset I \cap J$. The **product of two principal ideals** can be represented by $\langle i\rangle\langle j\rangle = \left\{\sum\limits_{k=1}^{n} i_k j_k : i_k \in I; j_k \in J : 1 \le k \le n \in \mathbb{N}\right\}$.

Examples:

1. The ring $\mathbb{Z}$ of **integers** is **principal** since for the **smallest positive** integer $d \in \mathbb{N} \cap I$ of a given ideal $I \subset \mathbb{Z}$and any other $n \in I$ according to the **Euclidean division** there exist integers q and $0 \le r < d$ such that $n = dq + r \Leftrightarrow r = n - dq \in I$ whence $r = 0$ such that we obtain $I = d\mathbb{Z}$.

2. The ring $K[x]$ of **polynomials in one variable** x **over a field** K is **principal** since for any polynomial $d \in K[x] \cap I$ with **minimal degree** $\deg d$ in a given ideal $I \subset K[x]$and any other $n \in I$ according to the **Euclidean division** there exist polynomials $q, r \in K[x]$ with $\deg r < \deg d$ such that $n = dq + r \Leftrightarrow r = n - dq \in I$ whence $r = 0$ such that we obtain $I = dK[x]$.

3. The ring $H(\mathbb{C})$of **entire functions** on the complex plane is **principal** since according to the **finite multiplicity of zeros of holomorphic functions** [2, p. 2.11] the generators f_k of the ideal $I = \langle f_k\rangle_{1\le k\le n}$ have at most finitely many common zeros a_k with at most finite common multiplicity n_k such that due to the **Weierstrass factorization theorem** [1, th. 5.14] there is an $f \in H(\mathbb{C})$with exactly these zeros of matching multiplicity. Hence we have $I = \langle f\rangle$

2.4 Commutative rings

An ideal $I \subset R$ in a **commutative** ring R is

1. **prime** iff the factor ring R/I is an **integral domain**

2. **maximal** iff there is no ideal $I \subsetneq M \subsetneq R$.

In a **commutative** ring R the following statements hold:

1. Every **maximal** ideal is **prime**.

2. Every ideal is contained in some maximal ideal.
3. $\{0\}$ is **prime** ideal iff R is an **integral domain**.
4. For every **maximal** ideal $M$the factor ring R/M is a **field**.
5. If the factor ring R/M of an ideal M is a **field**, then M is **maximal**.

Proof:

1. For a maximal ideal M and $r; s \in R$ with $rs \in M$. In the case of $r \notin M$due to the maximal character of M we have $M \subset M + Rr = R$ whence there are $n \in M$ and $t \in R$ with $1 = n + tr$. Multiplication by s yields $s = ns + trs \in M$ whence M is prime.
2. For any given ideal $I \subset R$ the family $\mathcal{I}$ of all ideals $I \subset J \subsetneq R$ in R is **inductively ordered** by inclusion since every linearly ordered chain $(I_k)_{k\in K} \subset \mathcal{I}$ has an upper bound $1 \notin \bigcup_{k\in K} I_k \subsetneq R$ which is an ideal due to the increasing charachter of the chain such that **Zorn's lemma** [5, p. 14.2.4] provides the desired maximal ideal $M \subset R$.
3. obvious.
4. From $1 \notin M$ follows $1 \neq 0 \mathrm{mod} M$ whence for every $r \neq 0 \mathrm{mod} M$ due to the maximal character of M the ideal $M \subset M + Rr = R$ such that there are $m \in M$ and $t \in R$ with $m + tr = 1$, i.e. $tr = 1 \mathrm{mod} M$ resp. $\pi(r)^{-1} = \pi(t)$.
5. For every $i \in I$ of an ideal $M \subset I \subsetneq R$ there is a $j \in R$ with $ij = 1 \mathrm{mod} M$ resp. an $m \in M$ with $1 = ij + m$ such that we obtain $i \in M$. Hence $I \subset M$ and since obviously $M \subsetneq R$ the assertion follows.

2.5 The Chinese remainder theorem

For **ideals** $(I_k)_{1\leq k\leq n}$ with $I_k + I_l = R$ for $k \neq l$ in a **commutative** ring R and any set $(r_m)_{1\leq m\leq n} \subset R$ there is an $r \in R$ with $r = r_k \mathrm{mod} I_k$ for every $1 \leq k \leq n$.

Proof: According to the hypothesis for $n = 2$ there are $i_1 \in I_1$ and $i_2 \in I_2$ with $i_1 + i_2 = 1$ such that $r = r_2 i_1 + r_1 i_2$ due to $r - r_1 = (r_2 - r_1) i_1$ and vice versa satisfies the given congruences. For $k \geq 2$ there are $a_k \in I_1$ and $b_k \in I_k$ with $a_k + b_k = 1$. Due to 2.3 this implies $1 = \prod_{k=2}^{n} (a_k + b_k) \in I_1 + \prod_{k=2}^{n} I_k$, i.e. $1 = i_1 + \prod_{k=2}^{n} i_k$ for some $i_k \in I_k$. But then for every $r \in R$ follows $r \cdot 1 = r \cdot i_1 + r \cdot \prod_{k=2}^{n} i_k \in I_1 + \prod_{k=2}^{n} I_k$ whence $I_1 + \prod_{k=2}^{n} I_k = R$. By the proven case for $n = 2$ we can find an $s_1 \in R$ with $s_1 = 1 \mathrm{mod} I_1$ resp. $s_1 = 0 \mathrm{mod} \left(\prod_{k=2}^{n} i_k\right)$ whence in particular $s_1 = 0 \mathrm{mod} I_k$ for $k \neq 1$. Similarly we obtain $(s_m)_{2\leq m\leq n}$ with $s_m = 1 \mathrm{mod} I_m$ resp. $s_m = 0 \mathrm{mod} I_k$ for $k \neq m$. Then $r = \sum_{m=1}^{n} r_m s_m$ is a solution for the given system of congruences.

2.6 Fields

A triple of a **set** K and two **maps** $+; \cdot : K \times K \to K$ is a **field** (*Körper*) $(K; + \cdot)$ iff

1. $(K; +; \cdot)$ is a **ring**
2. $(K \setminus \{0\}; \cdot)$ is an **abelian group**

Examples:

1. The ring $\mathbb{Z} \mathrm{mod}\ p = \mathbb{Z}/p\mathbb{Z}$ with the **equivalence relation** $r = s \bmod p \Leftrightarrow \exists z \in \mathbb{Z} : r - s = z \cdot p$ resp. the **equivalence classes** $r \bmod p$ with $0 \leq r < p$ (cf. [5, p. 8.9]) is a **domain** iff $p \in \mathbb{P}$ is a **prime number**, since for $k < p$ and $b, l, m \in \mathbb{Z}$ we have $(mp + k) \cdot b = lp \Leftrightarrow kb = (l - mb) \cdot p \Leftrightarrow$

$p|b$. For every $p \in \mathbb{P} \setminus \{2\}$ the pair $(\mathbb{Z}/p\mathbb{Z};\cdot)$ forms a **cyclic** and hence **abelian** group since for $l;m < p$ and $n \in \mathbb{N}$ we have $l \cdot m = n \cdot p + l \Leftrightarrow l \cdot (m-1) = n \cdot p$. In these cases $(\mathbb{Z}/p\mathbb{Z};+;\cdot)$ is a **field**.

2. More generally **every domain R of finite order with multiplicative unit is a field** since in that case for every $a \in R$ the mapping $x \mapsto ax$ is **injective** and hence **surjective**. Conversely the **order of a finite field is a prime number** since assuming $n \cdot m \cdot 1 = 0$ implies $m = 0$ or $n = 0$.

3. The ring $\mathbb{R}(x) \mod (x^2+1) = \mathbb{R}(x) / (x^2+1) \cdot \mathbb{R}(x)$ with the **equivalence relation** $p(x) = q(x) \mod x^2+1 \Leftrightarrow \exists r(x) \in \mathbb{R}(x) : p(x) - q(x) = r(x) \cdot (x^2+1)$ resp. the **equivalence classes** $(ax+b) \mod (x^2+1)$ with $a;b \in \mathbb{R}$ is a **field** since with $x^2 = -1 \mod (x^2+1)$ we have $(ax+b)\cdot(cx+d) = 1 \mod (x^2+1) \Leftrightarrow bd - ac = 1 \wedge ad + bc = 0 \Leftrightarrow b^2c + a^2c = -a \Leftrightarrow c = \frac{-a}{a^2+b^2} \wedge d = \frac{b}{a^2+b^2}$, i.e. $(ax+b)^{-1} = \frac{-ax+b}{a^2+b^2}$. With the isomorphism $x \mapsto i$ we obtain the **complex numbers**: $\mathbb{R}(x) \mod (x^2+1) \simeq \mathbb{C}$.

2.7 Polynomials

According to the **fundamental theorem of algebra** [2, p. 2.10] every **non constant polynomial** $p(z) = \sum\limits_{k=0}^{n} a_k z^k \in \mathbb{C}(z)$ with $a_k \in \mathbb{C}$, $a_n \neq 0$, **degree** $\deg p = n \geq 1$ and $0 \leq k \leq n$ has a complex **root** $\lambda \in \mathbb{C}$ with $p(\lambda) = 0$. Due to the **Euclidean polynomial division** for every root $\lambda \in \mathbb{C}$ we have $p(z) = q(z) \cdot (z - \lambda)$ with $q(z) \in \mathbb{C}(z)$ and $\deg q = \deg p - 1$. According to the rules for **complex conjugation** for every **non real root** $\lambda \in \mathbb{C}$ of a **real polynom** $p(z) \in \mathbb{R}(z)$we have $0 = p(\lambda) = \overline{p(\lambda)} = \overline{p}\left(\overline{\lambda}\right) = p\left(\overline{\lambda}\right)$ whence $p(z) = q(z) \cdot (z-\lambda)^{\mu(\lambda)} \left(z - \overline{\lambda}\right)^{\mu(\overline{\lambda})}$ with multiplicities $\mu(\lambda) = \mu\left(\overline{\lambda}\right)$ and the **real polynom** $(z-\lambda)^{\mu(\lambda)}\left(z-\overline{\lambda}\right)^{\mu(\overline{\lambda})}$ of **even degree** $2\mu(\lambda)$. Hence every **real polynom** can be factorized in the form $p(z) = \prod\limits_{i=1}^{k} (z - \lambda_i) \cdot \prod\limits_{i=k}^{(n-k)/2} (z - \lambda_i)\left(z - \overline{\lambda}_i\right)$ with $\mathrm{Im}\lambda_i = 0$ for $i < k$ and $\mathrm{Im}\lambda_i \neq 0$ for $i \geq k$ such that every **real polynom of odd degree must have at least one real root** $\lambda_1 \in \mathbb{R}$.

2.8 Descartes' rule of signs

The **number Z_f of strictly positive real roots** (counting multiplicity) of a **real polynomial** $p(x) = \sum\limits_{k=0}^{n} a_k z^{b_k} \in \mathbb{R}(x)$ with integer powers $0 \leq b_0 < b_1 < ... < b_n$ and real coefficients $a_i \in \mathbb{R} \setminus \{0\}$ is equal to the number $V_f = \sum_{\substack{0 \leq k < n \\ a_k \cdot a_{k+1} < 0}} 1$ of **sign changes** in the coefficients of f **minus a nonnegative even number**.

Proof:

1. W.l.o.g. we assume $b_0 = 0$ since otherwise a division by z^{b_0}would not change the number of strictly positive roots.

2. Z_f is **even** iff $a_n a_0 > 0$ since in the case of $f(0) = a_0 > 0$ and $a_n > 0$ we have $f(x) \to +\infty$ for $x \to +\infty$ and due to the **intermediate value theorem** [6, p. 5.1] it must cross the positive x-axis an even number of times (each of which contributes an odd number of roots) and glance without crossing an arbitrary number of times (each of which contributes an odd number of roots) such that Z_f must be even. The other cases are dealt with analogously.

3. Since every coefficient a_k with a $a_k a_0 < 0$ produces a pair of sign changes it follows from 2. that Z_f and V_f **have the same parity**.

4. It remains to show that $Z_f \leq V_f$: For $n = 0$ and $n = 1$ the proposition is obvious. Assuming $n \geq 2$ by the induction hypothesis we have $Z_{df/dx} = V_{df/dx} - 2m$ for some integer $m \geq 0$. By

the **mean value theorem** [2, th. 1.9] there is at least one positive root of $\frac{df}{dx}$ between any two different roots of f. Due to the **product rule** [2, th. 4.4] any k-multiple positive root of f is a $k-1$-multiple root of $\frac{df}{dx}$, i.e. $Z_{df/dx} \geq Z_f - 1$. Since $V_{df/dx} = V_f$ in the case of $a_1 a_0 > 0$ and $V_{df/dx} = V_f - 1$ otherwise we have $V_{df/dx} \leq V_f$. Hence $Z_f \leq Z_{df/dx} + 1 = V_{df/dx} - 2m + 1 \leq V_f - 2m + 1 \leq V_f + 1$ whence the assertion follows from 3.

3 Vector spaces

3.1 Vector spaces

The Quadruple $(X; K; +; \cdot)$ of a set V, a **field** $K \in \{\mathbb{R}; \mathbb{C}\}$, an **internal addition** $+ : X \times X \to X$ and an **external multiplication** $\cdot : K \times X \to X$ is a **vector space** over K iff

1. $(X; +)$is an **abelian group**
2. For $\lambda; \mu \in K$ and $\boldsymbol{x}; \boldsymbol{y} \in X$ we have
   a) **distribution laws** $(\lambda + \mu) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \mu \cdot \boldsymbol{x}$ and $\lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y}$
   b) **associative law** $\lambda \cdot (\mu \cdot \boldsymbol{x}) = (\lambda\mu) \cdot \boldsymbol{x}$
   c) compatibility of the **neutral element** $1 \cdot \boldsymbol{x} = \boldsymbol{x}$

These axioms imply the following properties:

3. $0 \cdot \boldsymbol{x} = 0$ and $\lambda \cdot \mathbf{0} = \mathbf{0}$
4. $\lambda \cdot \boldsymbol{x} = \mathbf{0} \Rightarrow \lambda = 0 \vee \boldsymbol{x} = \mathbf{0}$
5. $(-1) \cdot \boldsymbol{x} = -\boldsymbol{x}$

3.2 Vector subspaces

For a **vector space** X over a **field** $K \in \{\mathbb{R}; \mathbb{C}\}$ with subsets $A, B \subset X$ and **vectors** $\boldsymbol{x} \in X$ as well as **scalars** $\alpha \in K$ we define $\alpha A := \{\lambda \boldsymbol{a} : \boldsymbol{a} \in A\}$, $\boldsymbol{x} + A := \{\boldsymbol{x} + \boldsymbol{a} : \boldsymbol{a} \in A\}$ and $A + B := \{A + B : \boldsymbol{a} \in A, \boldsymbol{b} \in B\}$ with $-A = (-1)\, A$. For vector subspaces A and B the sets αA, $\boldsymbol{x} + A$ and $A + B$ are still **algebraically closed**. We have $2A \subset A + A$ with equality if A is a vector subspace. An arbitrary subset A generates its **linear** span $(A) = \langle A \rangle = \left\{ \sum_{k=1}^{n} \alpha_k \boldsymbol{x}_k : \alpha_k \in K, \boldsymbol{x}_k \in A, n \in \mathbb{N} \right\}$. A family $(\boldsymbol{x}_i)_{i \in I} \subset X$ is **linearly independent** iff $\sum_{i \in H} \alpha_i \boldsymbol{x}_i = 0 \Leftrightarrow \alpha_i = 0 \, \forall i \in H$ for every **finite** $H \subset I$. It is a **basis** of the subspace $E \subset X$ iff it **generates** $E = \langle \boldsymbol{x}_i \rangle_{i \in I} = \{\sum_{i \in H} \alpha_i \boldsymbol{x}_i : H$ **finite** in $I\}$. The rank (A) of a **matrix** $A = (a_{ij})_{1 \leq i \leq n; 1 \leq j \leq m} \in M\,(n \times m; \mathbb{C})$ is the maximal number of linearly independent **column vectors** $\boldsymbol{x}_j = (x_{ij})_{1 \leq i \leq n}$.

3.3 The basis of a vector space

Every **linearly independent** family $(\boldsymbol{x}_i)_{i \in I} \subset X$ can be extended to a **basis** $\langle \boldsymbol{x}_i \rangle_{i \in J} = X$ with $I \subset J$.

Proof: The set L of all linearly independent families $\mathcal{N} \subset X$ containing the given set $(\boldsymbol{x}_i)_{i \in I} \subset \mathcal{N}$ is **inductively ordered** by **inclusion** since for every **linearly ordered** chain $(\mathcal{N}_j)_{j \in J}$ with $\mathcal{N}_j = (\boldsymbol{x}_i)_{i \in I_j}$the index sets $(I_j)_{j \in J}$ are also linearly ordered such that $\mathcal{N} = \bigcup_{j \in J} \mathcal{N}_j = (\boldsymbol{x}_i)_{i \in \bigcup_{j \in J} I_j} \in L$ is a **supremum** of $(\mathcal{N}_j)_{j \in J}$. According to **Zorn's lemma** [5, th. 14.1.4] there is a **maximal family** $\mathcal{M} \in L$. Since for every $\boldsymbol{x} \in X$ we have $\langle \mathcal{M} \rangle \subset \langle \mathcal{M} \cup \{\boldsymbol{x}\} \rangle \in L$ such that from the maximal character of $\mathcal{M}$ follows $\boldsymbol{x} \in \langle \mathcal{M} \rangle$ whence we conclude that $X = \langle \mathcal{M} \rangle$.

3.4 The dimension of a vector space

All bases of a vector space X have the same cardinal number which is called the **dimension** $\dim X$ of X.

Proof: For two bases B and C of X the family Φ of injective maps $\varphi : B \supset \mathrm{dom}\varphi \to \mathrm{im}\varphi \subset C$ with linearly independent sets $\mathrm{im}\varphi \cup B \setminus \mathrm{dom}\varphi$ is inductively odered by inclusion since for every linearly ordered chain Φ_0 the map $\varphi_0 = \bigcup_{\varphi\in\Phi_0} \varphi \in \Phi$ is an upper bound of Φ_0; note that $\bigcup_{\varphi\in\Phi_0} \mathrm{im}\varphi \cup B \setminus \bigcup_{\varphi\in\Phi_0} \mathrm{dom}\varphi = \bigcup_{\varphi\in\Phi_0} \mathrm{im}\varphi \cup \bigcap_{\varphi\in\Phi_0} B \setminus \mathrm{dom}\varphi$ is still linearly independent. By **Zorn's lemma** [5, th. 14.2.4] the family Φ has a maximal element φ. Since any $b \in B \setminus \mathrm{dom}\varphi$ is linearly independent of $\mathrm{im}\varphi$ we infer that $\mathrm{im}\varphi \subset C$ is not a basis whence there exists a $c_0 \in C \setminus \mathrm{im}\varphi$.

On the one hand if this c_0 is linearly independent of $\mathrm{im}\varphi \cup B \setminus \mathrm{dom}\varphi$ there is an extension $\varphi' \supset \varphi$ defined by $\varphi'(b_0) = c_0$ for any $b_0 \in B \setminus \mathrm{dom}\varphi$ and $\varphi'(b) = \varphi(b)$ for every $b \in \mathrm{dom}\varphi$ contrary to the maximal character of φ.

On the other hand if c_0 is linearly dependent of $\mathrm{im}\varphi \cup B \setminus \mathrm{dom}\varphi$ it follows that $c_0 = \sum_{c\in\mathrm{im}\varphi} \lambda_c c + \sum_{b\notin\mathrm{dom}\varphi} \mu_b b$ with at least one $\mu_{b_0} \neq 0$ for some $b_0 \notin \mathrm{dom}\varphi$ Again we define an extension $\varphi' \supset \varphi$ by $\varphi'(b_0) = c_0$ and since c_0 is linearly independent of $\mathrm{im}\varphi \cup B \setminus \mathrm{dom}\varphi'$ the set $\mathrm{im}\varphi' \cup B \setminus \mathrm{dom}\varphi'$ is linearly independent whence $\varphi' \in \Phi$ contrary to the maximal character of φ.

Thus we have shown that $|X| \subset |Y|$ whence by the symmetry of the argument and the **Schroeder-Bernstein theorem** [5, th. 15.4] follows the assertion.

3.5 The Steinitz basis exchange lemma

For every basis $\mathcal{A} = (\boldsymbol{a}_i)_{i\in I}$ of a vector space $X = \langle \boldsymbol{a}_i \rangle_{i\in I}$ and $\boldsymbol{x} = \sum_{i\in J} \alpha_i \boldsymbol{a}_i \in X$ with finite $J \subset I$ and $\alpha_k \neq 0$ for $k \in J$ the set $\mathcal{A}' = (\boldsymbol{a}_i)_{i\in I'} \cup \{\boldsymbol{x}\}$ with $I' = I/\{k\}$ is again a **basis** since for every $\mathbf{y} = \sum_{i\in H} \beta_i \boldsymbol{a}_i$ w.l.o.g. we can assume $J \subset H$ whence $\boldsymbol{y} = \frac{\beta_k}{\alpha_k}\boldsymbol{x} + \sum_{i\in H}\left(\beta_i - \frac{\beta_k\alpha_i}{\alpha_k}\right)\boldsymbol{a}_i$. Also $\mathcal{A}'$ is **linearly independent** since any nontrivial solution $(\gamma_i)_{i\in I} \neq (0)_{i\in H}$ for $\gamma_k \boldsymbol{y} + \sum_{i\in H} \gamma_i \boldsymbol{a}_i = \mathbf{0}$ and finite $H \subset I$ would either entail a nontrivial solution $(\gamma_i)_{i\in H} \neq (0)_{i\in H}$ for $\sum_{i\in H} \gamma_i \boldsymbol{a}_i = \mathbf{0}$ in the case of $\gamma_k = 0$ or $-\sum_{i\in H} \frac{\gamma_i}{\gamma_k} \boldsymbol{a}_i = \boldsymbol{x}$ whence $\alpha_k = 0$ in the case of $\gamma_k \neq 0$ both in contradiction to the hypotheses. Hence the **dimension of a finite dimensional vector space is uniquely determined**.

3.6 Direct sums

Due to the exchange lemma for any vector subspace $E = \langle \mathbf{v}_l \rangle_{l\in L}$ w.l.o.g. we can assume $E = \langle \boldsymbol{a}_i \rangle_{i\in J}$ with $J \subset I$ such that X can be decomposed into a **direct sum** $X = E \oplus F$ with the **complementary space** $F = \langle \boldsymbol{a}_i \rangle_{i\in I\setminus J}$ and in the case of **finite** I follows

$$\dim X = \dim E + \dim F.$$

Obviously the vector subspaces $E; F \subset X$ are complementary to each other iff $E + F = X$ and $E \cap F = \{\mathbf{0}\}$. In the case of **vector spaces**, **rings** and **abelian groups**, the **direct sum** by $\varphi : E \times F \to E \oplus F$ with $\varphi((\boldsymbol{v}; \mathbf{w})) = (\boldsymbol{v}; \mathbf{0}) + (\mathbf{0}; \mathbf{w})$ is **isomorphic** to the **direct product**. Hence we have

$$\mathrm{card}\ X = \mathrm{card}\ E \cdot \mathrm{card}\ F.$$

The isomorphism fails for **nonabelian groups** since

$$\begin{aligned}
\varphi((\boldsymbol{v}_1 + \boldsymbol{v}_2; \boldsymbol{w}_1 + \boldsymbol{w}_2)) &= (\boldsymbol{v}_1 + \boldsymbol{v}_2; \mathbf{0}) + (\mathbf{0}; \boldsymbol{w}_1 + \boldsymbol{w}_2) \\
&= (\boldsymbol{v}_1; \mathbf{0}) + (\boldsymbol{v}_2; \mathbf{0}) + (\mathbf{0}; \boldsymbol{w}_1) + (\mathbf{0}; \boldsymbol{w}_2) \\
&\neq (\boldsymbol{v}_1; \mathbf{0}) + (\mathbf{0}; \boldsymbol{w}_1) + (\boldsymbol{v}_2; \mathbf{0}) + (\mathbf{0}; \boldsymbol{w}_2). \\
&= \varphi((\boldsymbol{v}_1; \boldsymbol{w}_1)) + \varphi((\boldsymbol{v}_2; \boldsymbol{w}_2)).
\end{aligned}$$

3.7 Linear maps

To avoid excessive cluttering of the notation by indices we follow the **Einstein summation convention** i.e. the summation sign is omitted for any index occurring twice. For the same reason we introduce a special case of the **index notation** from 3.13 with uppercase indices for coordinate vectors. A **linear map** $\boldsymbol{f} : X \to Y$ between vector spaces X and Y satisfies $\boldsymbol{f}(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha \boldsymbol{f}(\boldsymbol{x}) + \beta \boldsymbol{f}(\boldsymbol{y})$ for every $\alpha; \beta \in \mathbb{C}$ and $\boldsymbol{x}; \boldsymbol{y} \in X$. In particular it is a **homomorphism** on the **additive group** $(X; +)$ such that the corresponding terms resp. properties from 1.5 apply. Especially the **image** $\mathrm{Im}\boldsymbol{f} = \boldsymbol{f}[E] \subset Y$ as well as the **inverse image** $\boldsymbol{f}^{-1}[F] \subset X$ of vector subspaces $E \subset X$ resp. $F \subset Y$ under a linear map $\boldsymbol{f}$ are again vector subspaces and $\boldsymbol{f}$ is **injective** iff $\ker \boldsymbol{f} = \{\boldsymbol{0}\}$. The **ring** $L(X;Y)$ of **linear maps** $\boldsymbol{f} : X \to Y$ between **finite dimensional vector spaces** $X = \langle \boldsymbol{a}_i \rangle_{1 \le i \le m}$ resp. $Y = \langle \boldsymbol{b}_i \rangle_{1 \le j \le n}$ generated by **bases** $\mathcal{A} = (\boldsymbol{a}_i)_{1 \le i \le m}$ resp. $\mathcal{B} = (\boldsymbol{b}_j)_{1 \le j \le n}$ with $m, n \in \mathbb{N}$ by $M_{\mathcal{B}}^{\mathcal{A}} : L(X;Y) \to M(n \times m; \mathbb{C})$ defined by $M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f}) = \left(f^j(\boldsymbol{a}_i)\right)_{1 \le i \le m; 1 \le j \le n}$ for $\boldsymbol{f}(\boldsymbol{a}_i) = f^j(\boldsymbol{a}_i) \cdot \boldsymbol{b}_j \in L(X;Y)$ with **components** $f^j \in L(X; \mathbb{C})$ is **isomorphic** to the **ring** $M(n \times m; \mathbb{C})$ of **complex matrices**. Since $M(n \times m; \mathbb{C})$ is also a complex vector space of dimension $n \cdot m$ we obtain

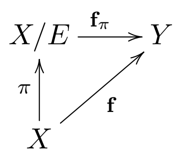
$$\dim L(X;Y) = \dim X \cdot \dim Y.$$

For $\boldsymbol{x} = x_{\mathcal{A}}^i \boldsymbol{a}_i \in X$ resp. $\boldsymbol{y} = y_{\mathcal{B}}^j \boldsymbol{b}_j \in Y$ with **coordinate vectors** $\boldsymbol{x}_{\mathcal{A}} = x_{\mathcal{A}}^i \boldsymbol{e}_i \in \mathbb{C}^m$ resp. $\boldsymbol{y}_{\mathcal{B}} = y_{\mathcal{B}}^j \boldsymbol{e}_j \in \mathbb{C}^n$ with regard to **orthonormal bases** of $\mathbb{C}^m = \langle \boldsymbol{e}_i \rangle_{1 \le i \le m}$ resp. $\mathbb{C}^n = \langle \boldsymbol{e}_j \rangle_{1 \le j \le n}$ defined in 6.4 we compute $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x_{\mathcal{A}}^i \boldsymbol{a}_i) = x_{\mathcal{A}}^i \boldsymbol{f}(\boldsymbol{a}_i) = x_{\mathcal{A}}^i f_j(\boldsymbol{a}_i) \cdot \boldsymbol{b}_j$. With the **canonical inner product** $A * \boldsymbol{x} = a_j^i \cdot x^j \cdot \boldsymbol{e}_i \in \mathbb{C}^n$ between a vector $\boldsymbol{x} = x^i \boldsymbol{e}_i \in \mathbb{C}^m$ and a matrix $A = \left(a_i^j\right)_{1 \le i \le m; 1 \le j \le n} \in M(n \times m; \mathbb{C})$ this assumes the form

$${}^T\boldsymbol{y}_{\mathcal{B}} = M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f}) * \boldsymbol{x}_{\mathcal{A}}.$$

3.8 Quotient spaces and rank

For any vector subspace $E \subset X$ the **quotient space** X/E is again a **vector space**. Its elements $\pi(\boldsymbol{x}) = \boldsymbol{x} + E$ for $\boldsymbol{x} \in X$ with $\pi(\boldsymbol{x}) = \pi(\boldsymbol{x}') \Leftrightarrow \boldsymbol{x} - \boldsymbol{x}' \in E$ for the **canonical projection** $\pi : X \to X/E$ are **affine space**s as defined in 8.1. In the case of a **finite dimensional** vector space $X = \langle \boldsymbol{a}_i \rangle_{i \in I}$ with $I = \{1; ...; n\}$ and a vector subspace $E = \langle \boldsymbol{a}_i \rangle_{i \in J}$ with $J \subset I$ according to the **Steinitz lemma** 3.5 w.l.o.g. we can assume $E = \langle \boldsymbol{a}_i \rangle_{i \in J}$ whence $X/E = \langle \boldsymbol{a}_i + E \rangle_{i \in I \setminus J}$ and



$$\dim X = \dim X/E + \dim E$$

in analogy to **Lagrange's theorem** 1.7 for finite groups. For every **linear** map $\boldsymbol{f} : X \to Y$ defined in the obvious way in following section 3.7 into another vector space Y with $E \subset \ker \varphi$ exists a **uniquely determined** and **linear** $\boldsymbol{f}_\pi : X/E \to Y$ with $\boldsymbol{f} = \boldsymbol{f}_\pi \circ \pi$ and

$$\ker \boldsymbol{f}_\pi = (\ker \boldsymbol{f})/E.$$

These properties are obvious if we define $\boldsymbol{f}_\pi(\pi(\boldsymbol{x})) = \boldsymbol{f}(\boldsymbol{x})$ for every $\boldsymbol{x} \in X$. The following **dimension formula** is a useful application: For every matrix $A \in M(n \times m; \mathbb{C})$ we define its **rank** $\mathrm{rank}\, A = \dim \mathrm{im}\boldsymbol{f} \le \min\{m; n\}$ with regard to the corresponding linear map $\boldsymbol{f} : \mathbb{C}^m \to \mathbb{C}^n$ with $A = M(\boldsymbol{f})$. Then by $\boldsymbol{f}_\pi : X/\ker\boldsymbol{f} \to \mathrm{im}\boldsymbol{f}$ with $\boldsymbol{f}_\pi \circ \pi = \boldsymbol{f}$ we obtain the **rank-nullity law**

$$\dim X = \dim \mathrm{im}\boldsymbol{f} + \dim \ker\boldsymbol{f}$$

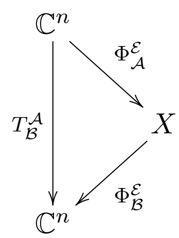
3.9 Endomorphisms

The **group** End (X) of **endomorphisms** $\boldsymbol{f} : X \to X$ on a **finite dimensional vector spaces** $X = \langle \boldsymbol{a}_i \rangle_{1 \leq i \leq n} = \langle \boldsymbol{b}_i \rangle_{1 \leq j \leq n}$ generated by the **bases** $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq n}$ and $\mathcal{B} = (\boldsymbol{b}_j)_{1 \leq j \leq n}$ by the map $M_{\mathcal{B}}^{\mathcal{A}} : \mathrm{End}\,(X) \to GL\,(n; \mathbb{C})$ defined as above is **isomorphic** to the **ring** (*groupe lineáire*) $GL\,(n;\mathbb{C})$ of **invertible complex matrices** of **rank** n. For a matrix $M = (f_{ij})_{1 \leq i;j \leq n} \in M\,(n;\mathbb{C})$ with the corresponding endomorphism $\boldsymbol{f} \in \mathrm{End}\,(X)$ defined by $\boldsymbol{f}\,(\boldsymbol{a}_i) = \sum_{j=1}^{n} f_{ij} \boldsymbol{b}_j$ for given bases $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq n}$ and $\mathcal{B} = (\boldsymbol{b}_j)_{1 \leq j \leq n}$ the following conditions are equivalent:

1. The column vectors $\boldsymbol{f}_j = \sum_{i=1}^{n} f_{ij} \boldsymbol{e}_i$ are linearly independent.
2. $\mathrm{Ker}\boldsymbol{f} = \{\boldsymbol{0}\}$
3. $\boldsymbol{f}$ is injective
4. $M \in GL\,(n;\mathbb{C})$
5. $\boldsymbol{f}$ is surjective.

Proof: In the chain 1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 5. $\Rightarrow$ 1. only the third step may require a comment: For **injective** $\boldsymbol{f} : X \to \mathrm{im}\boldsymbol{f} \subset X$ we have an inverse $\boldsymbol{f}^{-1} : \mathrm{im}\boldsymbol{f} \to X$ whence for every $\boldsymbol{x} \in X$ we infer $\boldsymbol{x} = \boldsymbol{f}^{-1}\,(\boldsymbol{f}\,(\boldsymbol{x})) \in \mathrm{im}\boldsymbol{f}$.

3.10 Coordinate transformations



As before every element $\boldsymbol{v} = \sum_{i=1}^{n} x_{\mathcal{A}i} \boldsymbol{a}_i \in X$ of a finite dimensional vector space $X = \langle \boldsymbol{a}_i \rangle_{1 \leq i \leq n}$ is determined by the **basis** $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq n}$ and a **coordinate vector** $\boldsymbol{x}_{\mathcal{A}} = \sum_{i=1}^{n} x_{\mathcal{A}i} \boldsymbol{e}_i \in \mathbb{C}^n$ with reference to the **orthonormal basis** $\mathcal{E} = (\boldsymbol{e}_i)_{1 \leq i \leq n}$ of $\mathbb{C}^n$. The corresponding **coordinate system** $\Phi_{\mathcal{A}}^{\mathcal{E}} : \mathbb{C}^n \to X$ with $\Phi_{\mathcal{A}}^{\mathcal{E}}\,(\boldsymbol{x}_{\mathcal{A}}) = \sum_{i=1}^{n} x_{\mathcal{A}i} \boldsymbol{a}_i$ is an isomorphism with the representing matrix $M\left(\Phi_{\mathcal{A}}^{\mathcal{E}}\right) = E_n$ (cf 3.7). For brevity in the **canonical case** $X = \mathbb{C}^n$ we will omit the symbol $\mathcal{E}$ for the **canonical basis** and also use the same notation for a **matrix** $A \in M\,(n \times m;\mathbb{C})$ and its corresponding **linear map** $A : \mathbb{C}^m \to \mathbb{C}^n$ with $A\,(\boldsymbol{x}) = A * \boldsymbol{x}$. The transition from the basis $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq n}$ to another basis $\mathcal{B} = (\boldsymbol{b}_j)_{1 \leq j \leq n}$ of $X = \langle \boldsymbol{a}_i \rangle_{1 \leq i \leq n} = \langle \boldsymbol{b}_j \rangle_{1 \leq j \leq n}$ is given by the **coordinate transformation**

$$T_{\mathcal{B}}^{\mathcal{A}} = \left(\Phi_{\mathcal{B}}^{\mathcal{E}}\right)^{-1} \circ \Phi_{\mathcal{A}}^{\mathcal{E}} : \mathbb{C}^n \to \mathbb{C}^n$$

with

$$T_{\mathcal{B}}^{\mathcal{A}}\,(\boldsymbol{e}_i) = \sum_{j=1}^{n} t_{ji} \boldsymbol{e}_j$$

such that the **column vectors** $(t_{ji})_{1 \leq j \leq n}$ of the **transformation matrix**

$$T_{\mathcal{B}}^{\mathcal{A}} = (t_{ji})_{1 \leq ij \leq n}$$

coincide with the **coordinate vectors** of the **original** basis $\mathcal{A} \subset X$ expressed by the **new** basis $\mathcal{B}$. For an arbitrary vector $\boldsymbol{v} = \sum_{i=1}^{n} x_{\mathcal{A}i} \boldsymbol{a}_i = \sum_{j=1}^{n} x_{\mathcal{B}j} \boldsymbol{b}_j$ we have $T_{\mathcal{B}}^{\mathcal{A}}\,(\boldsymbol{v}) = \sum_{i=1}^{n} x_{\mathcal{A}i} \sum_{j=1}^{n} t_{ji} \boldsymbol{b}_j$ whence $x_{\mathcal{B}j} = \sum_{i=1}^{n} t_{ji} x_{\mathcal{A}i}$, i.e.

$$\boldsymbol{x}_{\mathcal{B}} = T_{\mathcal{B}}^{\mathcal{A}} * \boldsymbol{x}_{\mathcal{A}}.$$

Vice versa the **column** vectors of the **inverse** $\left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} = T_{\mathcal{A}}^{\mathcal{B}}$ coincide with the **coordinate vectors** of the **new** basis $\mathcal{B}$ expressed by the **original** basis $\mathcal{A}$. In the **orthogonal** case according to 6.6 these coincide with the **row vectors** of $T_{\mathcal{B}}^{\mathcal{A}}$, i.e. $\left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} = {}^{T}T_{\mathcal{B}}^{\mathcal{A}}$ and vice versa.

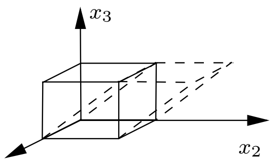
3.11 Change of bases

According to the **Gauss algorithm** every **automorphism** resp. every **invertible matrix** is the product of **elementary transformations** resp. **elementary matrices** of the two following types:

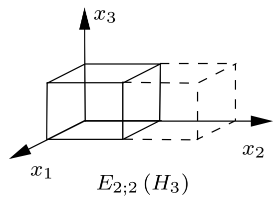
$$E_{kl} = \begin{array}{c} \begin{matrix} 1 & \cdots & k & \cdots & l & \cdots & n \end{matrix} \\ \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & 1 & & \\ & & & \ddots & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \end{array} \begin{matrix} 1 \\ \vdots \\ k \\ \vdots \\ l \\ \vdots \\ n \end{matrix} \qquad E_{k\alpha} = \begin{array}{c} \begin{matrix} 1 & \cdots & k & \cdots & n \end{matrix} \\ \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{array} \begin{matrix} 1 \\ \vdots \\ k \\ \vdots \\ n \end{matrix}$$

$$E_{22} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

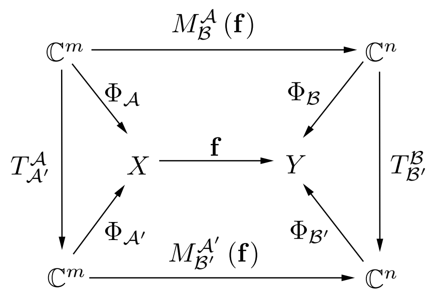


$E_{2;3}(H_3)$



$E_{2;2}(H_3)$

Multiplication of a matrix $A \in M(n \times m; \mathbb{C})$ with $E_{kl} \in GL(n; \mathbb{C})$ from the **left** results in an addition of the l-th **row** to the k-th row resp. a **shear** of the **hyperplane** span $\{\boldsymbol{e}_i : 1 \le i \le n;\, i \ne l\}$ in the direction of e_k whereas multiplication with $E_{kl} \in GL(m; \mathbb{C})$ from the **right** results in an addition of the k-th **column** to the l-th column and the corresponding shear of span $\{\boldsymbol{e}_i : 1 \le i \le n;\, i \ne k\}$ in the direction of e_l.



Multiplication with $E_{k\alpha}$ results in a multiplication of the k th row with the factor $\alpha \in \mathbb{C}$ resp. a **dilation** in the direction of $\boldsymbol{e}_k$ with factor α.

Hence for every **homomorphism** $\boldsymbol{f} : X \to Y$ between **finite dimensional vector spaces** X and Y with bases $\mathcal{A} \subset X$, $\mathcal{B} \subset Y$ there are bases $\mathcal{A}' \subset X$, $\mathcal{B}' \subset Y$ resp. coordinate transformations $T = T_{\mathcal{A}'}^{\mathcal{A}} \in GL(m; \mathbb{C})$ resp. $S = T_{\mathcal{B}'}^{\mathcal{B}} \in GL(n; \mathbb{C})$ such that

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\mathbf{f}) = \begin{array}{c} \begin{matrix} \overbrace{}^{k} & \overbrace{}^{m-k} \end{matrix} \\ \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \end{array} \begin{matrix} \}\, k \\ \}\, n-k \end{matrix}$$

$$S * F * T^{-1} = E_k$$

with $k = \mathrm{rank} A$ for $F = M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f})$. The corresponding map is

$$\boldsymbol{f}_{\mathcal{B}'}^{\mathcal{A}'} = T_{\mathcal{B}'}^{\mathcal{B}} \circ \boldsymbol{f}_{\mathcal{B}}^{\mathcal{A}} \circ \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} = \mathrm{id}|_{X'} + \boldsymbol{0}|_{\mathrm{Ker}\boldsymbol{f}} : X = X' \oplus \mathrm{Ker}\boldsymbol{f} \to Y.$$

3.12 Dual spaces

The **dual space** X^* of a vector space X is the vector space of all linear functionals $\boldsymbol{x}^* : X \to \mathbb{C}$. If the equation $\boldsymbol{x}^* \alpha \boldsymbol{x} = \alpha \boldsymbol{x}^* \boldsymbol{x}$ only holds for **real** $\alpha \in \mathbb{R}$ we have **real linearity**. In the case of **complex**

linearity we have $\mathrm{Re}\boldsymbol{x}^*i\boldsymbol{x} + i\mathrm{Im}\boldsymbol{x}^*i\boldsymbol{x} = \boldsymbol{x}^*i\boldsymbol{x} = i\boldsymbol{x}^*\boldsymbol{x} = -\mathrm{Im}\boldsymbol{x}^*\boldsymbol{x} + i\mathrm{Re}\boldsymbol{x}^*\boldsymbol{x} \Leftrightarrow \mathrm{Re}\boldsymbol{x}^*i\boldsymbol{x} = -\mathrm{Im}\boldsymbol{x}^*\boldsymbol{x}$ whence the functional $\boldsymbol{x}^*$ is **uniquely determined** by its **real part** $\mathrm{Re}\boldsymbol{x}^*$. Hence every **complex linear** $\boldsymbol{x}^*\boldsymbol{x} = \mathrm{Re}\boldsymbol{x}^*\boldsymbol{x} + i\mathrm{Im}\boldsymbol{x}^*\boldsymbol{x} = \mathrm{Re}\boldsymbol{x}^*\boldsymbol{x} - i\mathrm{Re}\boldsymbol{x}^*i\boldsymbol{x}$ is **real linear** and conversely for every **real linear** $\boldsymbol{u}^* : X \to \mathbb{R}$ the functional $\boldsymbol{x}^* : X \to \mathbb{C}$ with $\boldsymbol{x}^*\boldsymbol{x} = \boldsymbol{u}^*\boldsymbol{x} - i\boldsymbol{u}^*i\boldsymbol{x}$ is **complex linear**, since for $\alpha = \beta + i\gamma$ we have $\boldsymbol{x}^*\alpha\boldsymbol{x} = \beta\boldsymbol{u}^*\boldsymbol{x} + \gamma\boldsymbol{u}^*i\boldsymbol{x} - i\left(\beta\boldsymbol{u}^*i\boldsymbol{x} - \gamma\boldsymbol{u}^*\boldsymbol{x}\right) = (\beta + i\gamma)\left(\boldsymbol{u}^*\boldsymbol{x} - i\boldsymbol{u}^*i\boldsymbol{x}\right) = \alpha\boldsymbol{x}^*\boldsymbol{x}$.

In a **topological vector space** $\boldsymbol{x}^* \in X^*$ is **continuous** iff its real part is **continuous** and the dual space X^* usually is defined as the vector space of all **continuous** resp. due to [3, th. 5.1] **bounded** linear functionals on X.

For a **basis** $(\boldsymbol{e}_i)_{i\in I}$ of X the dual space $X^* = \langle \boldsymbol{e}_i^* \rangle_{i\in I}$ is generated by the **dual basis** $(\boldsymbol{e}_i^*)_{i\in I}$ defined by $\boldsymbol{e}_i^*\boldsymbol{e}_j = \delta_{ij}$. Hence by the **transposition** $\tau_X : X \to X^*$ with $\tau_X(\boldsymbol{e}_i) = \boldsymbol{e}_i^*$ every vector space X is **isomorphic** to its dual space X^*.

The transformation $\Phi_{\mathcal{B}^*}^{\mathcal{A}^*} : X^* \to X^*$ of the dual basis $\mathcal{A}^* = (\boldsymbol{a}_i^*)_{1\le i\le n}$ into another dual basis $\mathcal{B}^* = \left(\boldsymbol{b}_j^*\right)_{1\le j\le n}$ of $X^* = \langle \boldsymbol{a}_i^* \rangle_{1\le i\le n} = \left\langle \boldsymbol{b}_j^* \right\rangle_{1\le j\le n}$ is determined by the invariance with regard to **coordinate transformation** of the linear functional $\boldsymbol{x}^*\boldsymbol{x} = {}^T\boldsymbol{x}_{\mathcal{B}^*} * \boldsymbol{x}_{\mathcal{B}} = {}^T\left(T_{\mathcal{B}^*}^{\mathcal{A}^*} * \boldsymbol{x}_{\mathcal{A}^*}\right) * T_{\mathcal{B}}^{\mathcal{A}} * \boldsymbol{x}_{\mathcal{A}}$ $= {}^T\boldsymbol{x}_{\mathcal{A}^*} * {}^T T_{\mathcal{B}^*}^{\mathcal{A}^*} * T_{\mathcal{B}}^{\mathcal{A}} * \boldsymbol{x}_{\mathcal{A}} = {}^T\boldsymbol{x}_{\mathcal{A}^*} * \boldsymbol{x}_{\mathcal{A}}$ whence

$$ {}^T T_{\mathcal{B}^*}^{\mathcal{A}^*} = \left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1}. $$

3.13 The index notation

The coordinate vectors of $\boldsymbol{x} = \sum_{i=1}^{n} x_{\mathcal{A}i}\boldsymbol{a}_i = \sum_{i=1}^{n} x_{\mathcal{B}i}\boldsymbol{b}_i$ resp. its dual $\boldsymbol{x}^* = \sum_{i=1}^{n} x_{\mathcal{A}i}^*\boldsymbol{a}_i^* = \sum_{i=1}^{n} x_{\mathcal{B}i}^*\boldsymbol{b}_i^*$ are transformed from the **original** basis $\mathcal{A}$ to the **new** basis $\mathcal{B}$ by $\boldsymbol{x}_{\mathcal{B}} = T_{\mathcal{B}}^{\mathcal{A}} * \boldsymbol{x}_{\mathcal{A}}$ resp. $\boldsymbol{x}_{\mathcal{B}}^* = {}^T\left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} * \boldsymbol{x}_{\mathcal{A}}^*$ resp. ${}^T\boldsymbol{x}_{\mathcal{B}}^* * T_{\mathcal{B}}^{\mathcal{A}} = {}^T\boldsymbol{x}_{\mathcal{A}}^*$. The coordinate vectors $\boldsymbol{x}_{\mathcal{A}}$ and $\boldsymbol{x}_{\mathcal{B}}$ are called **contravariant** since the **column vectors** of the transformation matrix $T_{\mathcal{B}}^{\mathcal{A}} = (t_{jk})_{1\le j;k\le n}$ coincide with the coordinate vectors of the **original** basis $\mathcal{A}$ expressed by the **new** basis $\mathcal{B}$, i.e. $\boldsymbol{a}_k = \sum_{j=1}^{n} t_{jk}\boldsymbol{b}_j$ "contrary" to the **new** basis vectors. The **dual** coordinate vectors ${}^T\boldsymbol{x}_{\mathcal{A}}^*$ and ${}^T\boldsymbol{x}_{\mathcal{B}}^*$ are **covariant vectors** or **covectors** since the **row vectors** of the transformation matrix $T_{\mathcal{B}}^{\mathcal{A}}$ coincide with the coordinate vectors of the **new** basis $\mathcal{B}$ expressed by the **original** basis $\mathcal{A}$.

The **basis** vectors $\boldsymbol{a}_i$ are transformed to $\boldsymbol{b}_i = \sum_{i=1}^{n} s'_{ik}\boldsymbol{a}_k = \sum_{i=1}^{n}\sum_{j=1}^{n} s'_{ik}t_{jk}\boldsymbol{b}_j$ whence $\sum_{i=1}^{n}\sum_{j=1}^{n} s'_{ik}t_{jk} = \delta_{ij}$ resp. $(s'_{ik})_{1\le i;k\le n} = {}^T\left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} = T_{\mathcal{B}^*}^{\mathcal{A}^*}$. Hence the **basis vectors** of X are of **covariant** type and correspondingly the **basis covectors** of X^* are of **contravariant** type**.**

The transformation behaviour of a vector is indicated by the **index notation** denoting **contravariant** vectors with **uppercase** indices and **covariant** ones with **lowercase** indices.

Also we follow the **Einstein summation convention** introduced in 3.7 so that we have **vectors** $\boldsymbol{x} = \sum_{i=1}^{n} x_{\mathcal{A}}^i\boldsymbol{a}_i = x_{\mathcal{A}}^i\boldsymbol{a}_i$ with **contravariant coordinate vectors** resp. **covariant basis vectors** or **covectors** $\boldsymbol{x}^* = \sum_{i=1}^{n} x_{\mathcal{A}i}\boldsymbol{a}^i = x_{\mathcal{A}i}\boldsymbol{a}^i$. We will use **both notations** $\boldsymbol{a}^i = \boldsymbol{a}_i^*$ depending on the context of the behaviour of $\boldsymbol{a}^i$ under **coordinate transformation** resp. the role of $\boldsymbol{a}_i^*$ as a **functional**.

The representing matrix $m_k^j = M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f}) \in M(n\times m;\mathbb{C})$ of a **homomorphism** $\boldsymbol{f} : X \to Y$ between **finite dimensional complex vector spaces** X resp. Y with $\dim X = m$ resp. $\dim Y = n$ for bases $\mathcal{A} \subset X$ resp. $\mathcal{B} \subset Y$ has **contravariant column vectors** $\boldsymbol{m}_k = m_k^j\boldsymbol{e}_j \in \mathbb{C}^n$; $1 \le k \le m$ and **covariant row vectors** $\boldsymbol{m}^j = m_k^j\boldsymbol{e}^k \in (\mathbb{C}^m)^*$; $1 \le j \le n$ since the transformation into $(m')_l^i = M_{\mathcal{B}'}^{\mathcal{A}'}\boldsymbol{f} \in M(n\times m;\mathbb{C})$ for bases $\mathcal{A}' \subset X$, $\mathcal{B}' \subset Y$ with coordinate transformations $(t^{-1})_j^i = \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} \in$

$GL(m;\mathbb{C})$ resp. $s_l^k = T_{\mathcal{B}'}^{\mathcal{B}} \in GL(n;\mathbb{C})$ is given by

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\boldsymbol{f}) = S_{\mathcal{B}'}^{\mathcal{B}} * M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f}) * \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1}$$

resp.

$$(m')_l^i = s_l^k \cdot m_k^j \cdot \left(t^{-1}\right)_j^i .$$

Accordingly the **basis transformation** of **contravariant** $\boldsymbol{x} = x_{\mathcal{A}}^i \boldsymbol{a}_i = x_{\mathcal{B}}^j \boldsymbol{b}_j$ resp. **covariant** $\boldsymbol{y}^* = y_{\mathcal{A}i}\boldsymbol{a}^i = y_{\mathcal{B}j}\boldsymbol{a}^j$ by $t_j^i = T_{\mathcal{B}}^{\mathcal{A}} \in GL(n;\mathbb{C})$ is given by

$$\boldsymbol{x}_{\mathcal{B}} = T_{\mathcal{B}}^{\mathcal{A}} * \boldsymbol{x}_{\mathcal{A}} \text{ and } {}^T\boldsymbol{y}_{\mathcal{B}}^* = {}^T\boldsymbol{y}_{\mathcal{A}}^* * \left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1}$$

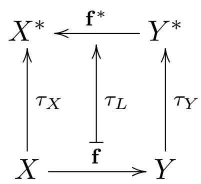
resp.

$$x_{\mathcal{B}}^j = t_i^j \cdot x_{\mathcal{A}}^i \text{ and } y_{\mathcal{B}j} = y_{\mathcal{A}i} \cdot \left(t^{-1}\right)_k^i \text{ with } t_i^j \cdot \left(t^{-1}\right)_k^i = \delta_k^j$$

Note that the distinction between **column** and **row** vectors as well as the **transposition** of matrices becomes **obsolete** since the information about the assignment of the corresponding summands is completely determined by the indices. In 7.4 we will encounter representing matrices $m_{j;k} = M_{\mathcal{A}}(s) \in M(m;\mathbb{C})$ of **sesquilinear forms** $s : X \times X \to \mathbb{C}$ with **covariant column vectors** as well as **covariant row vectors** $(\boldsymbol{m}^j);\left(\boldsymbol{m}^k\right) \in (\mathbb{C}^m)^*$ leading to the definition of the **tensor** concept generalizing vectors and matrices.

3.14 Dual linear maps

The vector space $L(X;Y)$ of linear maps $\boldsymbol{f} : X \to Y$ between the vector spaces $X = \langle \boldsymbol{a}_i \rangle_{1 \le i \le m}$ and $Y = \langle \boldsymbol{b}_j \rangle_{1 \le j \le n}$ is **isomorphic** to the dual space $L(Y^*;X^*)$ of linear maps between $X^* = \langle \boldsymbol{a}^i \rangle_{1 \le i \le m}$ and $Y^* = \left\langle \boldsymbol{b}^j \right\rangle_{1 \le j \le n}$ with the dual bases $\boldsymbol{a}^i = \tau_X(\boldsymbol{a}_i)$ resp. $\boldsymbol{b}^j = \tau_Y(\boldsymbol{b}_j)$ provided by the **transpositions** τ_X resp. τ_Y according to 3.12. The isomorphism is given by another **transposition** $\tau_L : L(X;Y) \to L(Y^*;X^*)$ with $\tau_L(\boldsymbol{f}) = \boldsymbol{f}^*$ defined by $\boldsymbol{f}(\boldsymbol{y}^*) = \boldsymbol{y}^* \circ \boldsymbol{f}$ for every **linear** $\boldsymbol{f} = f_i^j \boldsymbol{b}_j \boldsymbol{a}^i \in L(X;Y)$ with $\boldsymbol{a}^i \boldsymbol{a}_k = \delta_k^i$ whence $\boldsymbol{f}(\boldsymbol{a}_i) = f_i^j \boldsymbol{b}_j$ and every linear form $\boldsymbol{y}^* = y_j \boldsymbol{b}^j \in Y^*$.



For $\boldsymbol{x} = x^k \boldsymbol{a}_k \in X$ on the one hand we have

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = f_i^j \boldsymbol{b}_j \boldsymbol{a}^i x^k \boldsymbol{a}_k = f_i^j x^i \boldsymbol{b}_j$$

resp. in coordinate vectors

$$\begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = \begin{pmatrix} f_1^1 & \cdots & f_m^1 \\ \vdots & & \vdots \\ f_1^n & \cdots & f_m^n \end{pmatrix} * \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$$

and on the other hand due to $\boldsymbol{b}^l \boldsymbol{b}_j = \delta_j^l$ holds

$$\boldsymbol{x}^* = \boldsymbol{f}^*(\boldsymbol{y}^*) = y_l \boldsymbol{b}^l f_i^j \boldsymbol{b}_j \boldsymbol{a}^i = f_i^j y_j \boldsymbol{a}^i$$

resp. in coordinate vectors

$$(x_1;...;x_m) = (y_1;...;y_n) * \begin{pmatrix} f_1^1 & \cdots & f_m^1 \\ \vdots & & \vdots \\ f_1^n & \cdots & f_m^n \end{pmatrix}$$

resp.

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = {}^T\begin{pmatrix} f_1^1 & \cdots & f_m^1 \\ \vdots & & \vdots \\ f_1^n & \cdots & f_m^n \end{pmatrix} * \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

whence

$$M_{\mathcal{A}^*}^{\mathcal{B}^*}(\boldsymbol{f}^*) = {}^T M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f})$$

.

3.15 Annihilator and rank

For the **annihilator** $E^0 = \{\boldsymbol{x}^* \in X^* : \boldsymbol{x}^*\boldsymbol{x} = 0\,\forall \boldsymbol{x} \in E\}$ of a vector subspace $E \subset X$ and every $\boldsymbol{f} \in L(X;Y)$ holds

$$\ker \boldsymbol{f}^* = (\mathrm{im}\boldsymbol{f})^0 \text{ and } \mathrm{im}\boldsymbol{f}^* = (\ker\boldsymbol{f})^0$$

since $\boldsymbol{y}^* \in \mathrm{Ker}\boldsymbol{f}^* \Leftrightarrow \boldsymbol{y}^* \circ \boldsymbol{f} = \boldsymbol{0} \Leftrightarrow \boldsymbol{y}^*(\boldsymbol{f}(\boldsymbol{x})) = 0\,\forall \boldsymbol{x} \in X \Leftrightarrow \boldsymbol{y}^* \in (\mathrm{Im}\boldsymbol{f})^0$ and vice versa. Due to 3.12 in the **finite** case we have $\boldsymbol{a}_i^*\boldsymbol{a}_j = \delta_{ij}$ for every basis $(\boldsymbol{a}_j)_{1\le j\le n}$ with $X = \langle \boldsymbol{a}_j \rangle_{1\le j\le n}$ resp. $X^* = \langle \boldsymbol{a}_i^* \rangle_{1\le i\le n}$ and hence

$$\mathrm{dim}X = \mathrm{dim}E + \mathrm{dim}E^0.$$

According to 3.7 it follows for every matrix $F = M(\boldsymbol{f}) \in \mathrm{M}(n \times m;\mathbb{C})$ associated to the uniquely determined linear map $\boldsymbol{f} = M^{-1}(F) : \mathbb{C}^m \to \mathbb{C}^n$ on **canonical bases** $\mathbb{C}^m = \langle \boldsymbol{e}_j \rangle_{1\le j\le m}$ resp. $\mathbb{C}^n = \langle \boldsymbol{e}_i \rangle_{1\le i\le n}$ that

$$\mathrm{rank}^T F = \dim\, \mathrm{im}\boldsymbol{f}^* = \dim(\ker\boldsymbol{f})^0 = \dim\, \mathrm{im}\boldsymbol{f} = \mathrm{rank}F$$

3.16 Dual bases

$\boldsymbol{x}^* \in X^*$ is a linear combination of the **linearly independent** family $(\boldsymbol{x}_i^*)_{1\le i\le n} \subset X^*$ iff

$$\ker \boldsymbol{x}^* \supset \bigcap_{i=1}^{n} \ker \boldsymbol{x}_i^*$$

.

Proof: $\Rightarrow$ is trivial and concerning $\Leftarrow$ we consider the linear map $\boldsymbol{f} = \sum_{i=1}^{n} \boldsymbol{x}_i^*\boldsymbol{e}_i : X \to Y = \mathbb{C}^n = \langle \boldsymbol{e}_i \rangle_{1\le i\le n}$ with $\ker \boldsymbol{f} = \bigcap_{i=1}^{n} \ker \boldsymbol{x}_i^* \subset \ker \boldsymbol{x}^*$ whence $\mathrm{im}\boldsymbol{f}^* = (\ker\boldsymbol{f})^0 \supset (\ker \boldsymbol{x}^*)^0 \ni \boldsymbol{x}^*$, i.e. there is an $\boldsymbol{a}^* = \sum_{i=1}^{n} a_i\boldsymbol{e}_i^* \in Y^*$ such that $\boldsymbol{x}^* = \boldsymbol{f}^*(\boldsymbol{a}^*) = \boldsymbol{a}^*(\boldsymbol{f}) = \sum_{i=1}^{n} a_i\boldsymbol{x}_i^*$ on account of $\boldsymbol{e}_i^*(\boldsymbol{e}_j) = \delta_{ij}$.

4 Determinants

4.1 The Weierstrass axioms

In this section we write quadratic matrices as representing matrices of endomorphisms, i.e. as tensors of type $(1;1)$ with **contravariant column vectors** and **covariant row vectors**. The general determinant as defined below is a **function of a matrix** resp. tensor of dgree 2 of arbitrary type $(2;0)$, $(1;1)$ or $(0;2)$ (cf. section 7) without regard to its transformation properties. In the followng section 5 the matrix will be defined as a **function of an endomorphism** and only in that context resp. only on tensors of type $(1;1)$ it is invariant under coordinate transformations. Also in this section we will **not use the Einstein summation convention**.

The map $\det : M(n;\mathbb{C}) \to \mathbb{C}$ is a **determinant**, iff it is

1. **linear** in every row, i.e. $\det\begin{pmatrix} \vdots \\ \lambda\boldsymbol{a}+\mu\boldsymbol{b} \\ \vdots \end{pmatrix} = \lambda\det\begin{pmatrix} \vdots \\ \boldsymbol{a} \\ \vdots \end{pmatrix} + \mu\det\begin{pmatrix} \vdots \\ \boldsymbol{b} \\ \vdots \end{pmatrix}$ for $\lambda;\mu\in\mathbb{C}$, $A\in M((i-1)\times n;\mathbb{C})\,; B\in M((n-i)\times n;\mathbb{C})$, $0\le i\le n$ and row vectors $\boldsymbol{a};\boldsymbol{b}\in\mathbb{C}^n$.
2. **alternating**, i.e. $\det\begin{pmatrix} \boldsymbol{a}^1 \\ \vdots \\ \boldsymbol{a}^n \end{pmatrix} = 0$ iff $\boldsymbol{a}^i = \boldsymbol{a}^j$ for some $1\le i<j\le n$
3. **normed**, i.e. $\det E_n = 1$.

The following properties are direct consequences of the definitions:

4. $\det(\lambda\cdot A) = \lambda^n\cdot\det A$
5. $\det\begin{pmatrix} \vdots \\ \boldsymbol{0} \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ \boldsymbol{a} \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ -\boldsymbol{a} \\ \vdots \end{pmatrix} = 0$.
6. $\det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{b} \end{pmatrix} + \det\begin{pmatrix} \boldsymbol{b} \\ \vdots \\ \boldsymbol{a} \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{b} \end{pmatrix} + \det\begin{pmatrix} \boldsymbol{b} \\ \vdots \\ \boldsymbol{b} \end{pmatrix} + \det\begin{pmatrix} \boldsymbol{b} \\ \vdots \\ \boldsymbol{a} \end{pmatrix} + \det\begin{pmatrix} \boldsymbol{b} \\ \vdots \\ \boldsymbol{a} \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{a}+\boldsymbol{b} \\ \vdots \\ \boldsymbol{a}+\boldsymbol{b} \end{pmatrix}$ $= 0$.
7. $\det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{b}+\lambda\boldsymbol{a} \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{b} \end{pmatrix} + \lambda\det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{a} \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{a} \\ \vdots \\ \boldsymbol{b} \end{pmatrix}$.
8. $\det\begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1\cdot\ldots\cdot\lambda_n$ due to the **Gauss algorithm** and 4.1.7.
9. $\det\begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix} = \det A_1\cdot\det A_2$ for **quadratic** matrices A_1 and A_2 due to 4.1.8.
10. $\det A = 0 \Leftrightarrow \operatorname{rang} A < n$ due to the **Gauss algorithm** and 4.1.7.
11. $\det(A*B) = \det A\cdot\det B$ which in the case of $\operatorname{rang}A = \operatorname{rang}B = n$ due to the **Gauss algorithm** and 4.1.7 can be deduced from the diagonal case

$$\det\left(\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\cdot\begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}\right) = \lambda_1\cdot\mu_1\cdot\ldots\cdot\lambda_n\cdot\mu_n$$
$$= \det\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\cdot\det\begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$$

and in the case of $\operatorname{rang}A < n$ or $\operatorname{rang}B < n$ is trivial due to 4.1.10.
12. **Antisymmetry**: According to 1.22.1, 1.23 and 4.1.6 for every **permutation** $\sigma = \tau_1\circ\ldots\circ\tau_n\in S_n$ we have

$$\det\begin{pmatrix}\boldsymbol{a}^{\sigma(1)}\\ \vdots\\ \boldsymbol{a}^{\sigma(n)}\end{pmatrix}=\operatorname{sgn}(\sigma)\cdot\det\begin{pmatrix}\boldsymbol{a}^{1}\\ \vdots\\ \boldsymbol{a}^{n}\end{pmatrix}\text{ and in particular }\det\begin{pmatrix}\boldsymbol{e}^{\sigma(1)}\\ \vdots\\ \boldsymbol{e}^{\sigma(n)}\end{pmatrix}=\operatorname{sgn}(\sigma)\,.$$

In the case of a transposition $\tau_{i;j}=\langle i;j\rangle$ exchanging to identical row vectors $\boldsymbol{a}_i=\boldsymbol{a}_j$ we have

$$\det\begin{pmatrix}\boldsymbol{a}^{\tau(1)}\\ \vdots\\ \boldsymbol{a}^{\tau(n)}\end{pmatrix}=-\det\begin{pmatrix}\boldsymbol{a}^{1}\\ \vdots\\ \boldsymbol{a}^{n}\end{pmatrix}=0$$

whence the **antisymmetry is equivalent to the alternating character** 4.1.2 of the determinant.

4.2 Leibniz' formula

There exists a **uniquely determined** and **continuous** map $\det : M(n;\mathbb{C})\to\mathbb{C}$ with $\det A=\sum_{\sigma\in S_n}\prod_{i=1}^{n}a^i_{\sigma(i)}\operatorname{sgn}(\sigma)$ satisfying the three conditions 4.1.1 - 4.1.3. In particular we have

$$\det{}^T A=\sum_{\sigma\in S_n}\prod_{k=1}^{n}a_k^{\sigma(k)}=\sum_{\sigma^{-1}\in S_n}\prod_{\sigma(k)=1}^{n}a^{\sigma(k)}_{\sigma^{-1}(\sigma(k))}=\sum_{\sigma^{-1}\in S_n}\prod_{m=1}^{n}a^m_{\sigma^{-1}(m)}=\det A.$$

Proof: Applying 4.1.1 to the row vectors ${}^T\boldsymbol{a}^i=\sum_{j=1}^{n}a^i_j\,{}^T\boldsymbol{a}^j$ we obtain

$$\begin{aligned}\det\begin{pmatrix}\boldsymbol{a}^{1}\\ \boldsymbol{a}^{2}\\ \vdots\\ \boldsymbol{a}^{n}\end{pmatrix}&\overset{4.1.1}{=}\sum_{i_1=1}^{n}a^1_{i_1}\det\begin{pmatrix}\mathbf{e}^{i_1}\\ \boldsymbol{a}^{2}\\ \vdots\\ \boldsymbol{a}^{n}\end{pmatrix}\\
&\overset{4.1.1}{=}\sum_{i_1=1}^{n}a^1_{i_1}\sum_{i_2=1}^{n}a^2_{i_2}\det\begin{pmatrix}\mathbf{e}^{i_1}\\ \mathbf{e}^{i_2}\\ \vdots\\ \boldsymbol{a}^{n}\end{pmatrix}\\
&\vdots\\
&\overset{4.1.1}{=}\sum_{i_1=1}^{n}a^1_{i_1}\sum_{i_2=1}^{n}a^2_{i_2}\dots\sum_{i_n=1}^{n}a^n_{i_n}\det\begin{pmatrix}\boldsymbol{e}^{i_1}\\ \boldsymbol{e}^{i_2}\\ \vdots\\ \boldsymbol{e}^{i_n}\end{pmatrix}\\
&\overset{4.1.2}{=}\sum_{\sigma\in S_n}\prod_{i=1}^{n}a^i_{\sigma(i)}\det\begin{pmatrix}\boldsymbol{e}^{\sigma(1)}\\ \boldsymbol{e}^{\sigma(2)}\\ \vdots\\ \boldsymbol{e}^{\sigma(n)}\end{pmatrix}\\
&\overset{4.1.3}{=}\sum_{\sigma\in S_n}\prod_{i=1}^{n}a^i_{\sigma(i)}\operatorname{sgn}(\sigma)\,.\end{aligned}$$

The above defined function satisfies

1. since

$$\det\begin{pmatrix} \vdots \\ \lambda \boldsymbol{a}^i + \mu \boldsymbol{b}^i \\ \vdots \end{pmatrix} = \sum_{\sigma\in S_n} a^1_{\sigma(1)}\cdot \ldots \cdot \left(\lambda a^i_{\sigma(i)} + \mu b^i_{\sigma(i)}\right)\cdot \ldots \cdot a^n_{\sigma(n)}\operatorname{sgn}(\sigma)$$
$$= \lambda \sum_{\sigma\in S_n} a^1_{\sigma(1)}\cdot \ldots \cdot a^i_{\sigma(i)}\cdot \ldots \cdot a^n_{\sigma(n)}\operatorname{sgn}(\sigma)$$
$$+ \mu \sum_{\sigma\in S_n} a^1_{\sigma(1)}\cdot \ldots \cdot b^i_{\sigma(i)}\cdot \ldots \cdot a^n_{\sigma(n)}\operatorname{sgn}(\sigma)$$
$$= \lambda \det\begin{pmatrix} \vdots \\ \boldsymbol{a}^i \\ \vdots \end{pmatrix} + \mu \det\begin{pmatrix} \vdots \\ \boldsymbol{b}^i \\ \vdots \end{pmatrix}$$

2. since in the case of $\mathbf{a}^k = \mathbf{a}^l$ due to 1.23 we have a bijection $A_n \to A_n \circ \tau$ with $\tau(k) = l$, $\tau(i) = i$ for $i \neq k; l$ and $\operatorname{sgn}[A_n] = 1 = -\operatorname{sgn}[A_n \circ \tau]$ such that

$$\det\begin{pmatrix} \boldsymbol{a}^k \\ \vdots \\ \boldsymbol{a}^l \end{pmatrix} = \sum_{\sigma\in A_n} a^1_{\sigma(1)}\cdot \ldots \cdot a^k_{\sigma(k)}\cdot \ldots \cdot a^l_{\sigma(l)}\cdot \ldots \cdot a^n_{\sigma(n)}$$
$$- \sum_{\sigma\in A_n} a^1_{\sigma(\tau(1))}\cdot \ldots \cdot a^k_{\sigma(\tau(k))}\cdot \ldots \cdot a^l_{\sigma(\tau(l))}\cdot \ldots \cdot a^n_{\sigma(\tau(n))}$$
$$= \sum_{\sigma\in A_n} a^1_{\sigma(1)}\cdot \ldots \cdot a^k_{\sigma(k)}\cdot \ldots \cdot a^l_{\sigma(l)}\cdot \ldots \cdot a^n_{\sigma(n)}$$
$$- \sum_{\sigma\in A_n} a^1_{\sigma(1)}\cdot \ldots \cdot a^k_{\sigma(l)}\cdot \ldots \cdot a^l_{\sigma(k)}\cdot \ldots \cdot a^n_{\sigma(n)}$$
$$= 0$$

3. since $\det E_n = 1^n = 1$.

4.3 Cramer's rule

For $A = \left(a^i_j\right)_{1\leq i;j\leq n} \in M(n;\mathbb{C})$ and

$$A_{ij} = \det\begin{pmatrix} a^1_1 & \cdots & a^1_{j-1} & 0 & a^1_{j+1} & \cdots & a^1_n \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a^{i-1}_1 & \cdots & a^{i-1}_{j-1} & 0 & a^{i-1}_{j-1} & & a^{i-1}_n \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a^{i+1}_1 & \cdots & a^{i+1}_{j-1} & 0 & a^{i+1}_{j+1} & \cdots & a^{i+1}_n \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a^n_1 & \cdots & a^n_{j-1} & 0 & a^n_{j+1} & \cdots & a^n_n \end{pmatrix} \text{ and } A'_{ij} = \det\begin{pmatrix} a^1_1 & \cdots & a^1_{j-1} & a^1_{j+1} & \cdots & a^1_n \\ \vdots & & \vdots & \vdots & & \vdots \\ a^{i-1}_1 & \cdots & a^{i-1}_{j-1} & a^{i-1}_{j+1} & \cdots & a^{i-1}_n \\ a^{i+1}_1 & \cdots & a^{i+1}_{j-1} & a^{i+1}_{j+1} & \cdots & a^{i+1}_n \\ \vdots & & \vdots & \vdots & & \vdots \\ a^n_1 & \cdots & a^n_{j-1} & a^n_{j+1} & \cdots & a^n_n \end{pmatrix}$$

and the **complementary** $A^\natural = \left(a^i_j\right)_{1\leq i;j\leq n}$ with $a_{i,j} = \det A_{ji}$ we have

1. $A^{-1} = \frac{A^\natural}{\det A}$

such that for every $A \in GL(n;\mathbb{C})$ and $\boldsymbol{b} \in \mathbb{C}^n$ the solution of the linear equation $A * \boldsymbol{x} = \boldsymbol{b}$ is given by

2. $x_i = \frac{\det(\boldsymbol{a}_1;\ldots;\boldsymbol{a}_{i-1};\boldsymbol{b};\boldsymbol{a}_{i+1};\ldots;\boldsymbol{a}_n)}{\det A}$ for $1 \leq i \leq n$.

Proof:

First we note that

1. $\det A_{ij} = (-1)^{i+j} \det A'_{ij}$
2. $\det A_{ij} = \det(\boldsymbol{a}_1; ...; \boldsymbol{a}_{j-1}; \boldsymbol{e}_i; \boldsymbol{a}_{j+1}; ...; \boldsymbol{a}_n) = \det \begin{pmatrix} \boldsymbol{a}^1 \\ \\ \boldsymbol{a}^{i-1} \\ \boldsymbol{e}^j \\ \boldsymbol{a}^{i+1} \\ \\ \boldsymbol{a}^n \end{pmatrix}$

since

1. A_{ij} can be transformed into A'_{ij} by $(i-1)$ row **transpositions** and $(j-1)$ column **transpositions** due to 4.1.6
2. The **row** vector $\boldsymbol{e}^i$ can be transformed into $\boldsymbol{a}^i$ by adding $a^i_j \boldsymbol{e}_i$ to the j-th **column** vectors of A_{ij} and
   the **column** vector $\boldsymbol{e}_j$ can be transformed into $\boldsymbol{a}_j$ by adding $a^i_j \boldsymbol{e}^j$ to the i-th **row** vectors of A_{ij} due to 4.1.7.

In order to prove the formula for the **inverse matrix** we compute the components of $A * A^\natural$:

$$\begin{aligned}\sum_{j=1}^n a^{\natural i}_j a^j_k &= \sum_{j=1}^n a^j_k \cdot \det A_{ji} \\ &= \sum_{j=1}^n a^j_k \cdot \det(\boldsymbol{a}_1; ...; \boldsymbol{a}_{i-1}; \boldsymbol{e}_j; \boldsymbol{a}_{i+1}; ...; \boldsymbol{a}_n) \\ &= \det\left(\boldsymbol{a}_1; ...; \boldsymbol{a}_{i-1}; \sum_{j=1}^n a^j_k \boldsymbol{e}_j; \boldsymbol{a}_{i+1}; ...; \boldsymbol{a}_n\right) \\ &= \det(\boldsymbol{a}_1; ...; \boldsymbol{a}_{i-1}; \boldsymbol{a}_k; \boldsymbol{a}_{i+1}; ...; \boldsymbol{a}_n) \\ &= \delta_{ik} \cdot \det A\end{aligned}$$

Applying this formula to the components of $\mathbf{x} = A^{-1} * \boldsymbol{b} = \frac{A^\natural \boldsymbol{b}}{\det A}$ yields

$$\begin{aligned}x_i &= \frac{1}{\det A} \sum_{j=1}^n b_j \cdot \det A_{ji} \\ &= \frac{1}{\det A} \sum_{j=1}^n b_j \cdot \det(\boldsymbol{a}_1; ...; \boldsymbol{a}_{i-1}; \boldsymbol{e}_j; \boldsymbol{a}_{i+1}; ...; \boldsymbol{a}_n) \\ &= \frac{\det(\boldsymbol{a}_1; ...; \boldsymbol{a}_{i-1}; \boldsymbol{b}; \boldsymbol{a}_{i+1}; ...; \boldsymbol{a}_n)}{\det A}.\end{aligned}$$

4.4 Laplace's formula

For $A \in M(n; \mathbb{C})$, $n \geq 2$ and every $1 \leq i; j \leq n$ we have

$$\det A = \sum_{j=1}^n (-1)^{i+j} \cdot a^i_j \cdot \det A'_{ij} = \sum_{i=1}^n (-1)^{i+j} \cdot a^i_j \cdot \det A'_{ij}$$

.

Proof: According to 4.3.1 and the subsequent formula 1. in the proof we have

$$\det A = \sum_{j=1}^n a^i_j a^{\natural j}_i = \sum_{j=1}^n a^i_j \cdot \det A_{ij} = \sum_{j=1}^n a^i_j \cdot (-1)^{i+j} \cdot a^i_j \cdot \det A'_{ij}.$$

4.5 Orientation

Two **matrices** $A, B \in GL(n;\mathbb{R})$ have the same **orientation**, i.e.

$$\det A \cdot \det B > 0$$

iff they are **connected** (cf. [6, p. 5.8]), i.e. there is a **continuous path**

$$\varphi : [0;1] \to GL(n;\mathbb{R}) \text{ with } \varphi(0) = A \text{ and } \varphi(1) = B.$$

Hence the family of all **invertible matrice**s resp. **bases** $\mathcal{A} = \left(\sum_{i=1}^{n} a_j^i \boldsymbol{e}_i\right)_{1 \le j \le n}$ and $\mathcal{B} = \left(\sum_{i=1}^{n} b_j^i \boldsymbol{e}_i\right)_{1 \le j \le n}$ is decomposed into two equivalence classes resp. **connected components** with **right** resp. **left handed orientation**.

Proof:

$\Rightarrow$: We show that every $A \in GL(n;\mathbb{R})$ with $\det A > 0$ is **path connected** to E_n.

Step I: According to the **Gauss-algorithm** the **invertible** matrix A can be transformed into a **diagonal** matrix

$$L = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ with } (\lambda_i)_{1 \le i \le n} \subset \mathbb{C}$$

by adding multiples of rows to other rows. Due to 4.1.7 these operations leave the determinant unchanged such that $\det A = \det L = \prod_{i=1}^{n} \lambda_i$. Each row operation can be represented by a **path** as e.g. for addition of the μ-multiple of the i-th row $\boldsymbol{a}^i$ to the j-th row $\boldsymbol{a}^j$

$$\varphi(t) = \begin{pmatrix} \vdots \\ \boldsymbol{a}_{j-1} \\ \boldsymbol{a}^j + t \cdot \mu \boldsymbol{a}^i \\ \boldsymbol{a}_{j+1} \\ \vdots \end{pmatrix} \text{ with } \det \varphi(t) = \det A \text{ for } 0 \le t \le 1.$$

Step II: The values of the diagonal elements are reduced to ± 1 without leaving $GL(n;\mathbb{R})$: Due to 4.1.1 we have $\det L = \prod_{i=1}^{n} \lambda_i = \prod_{i=1}^{n} |\lambda_i| \det E$ with

$$E = \begin{pmatrix} \frac{\lambda_1}{|\lambda_1|} & & 0 \\ & \ddots & \\ 0 & & \frac{\lambda_n}{|\lambda_n|} \end{pmatrix} \text{ and a path } \varphi(t) = \begin{pmatrix} \lambda_1 + t\left(\frac{\lambda_1}{|\lambda_1|} - \lambda_1\right) & & 0 \\ & \ddots & \\ 0 & & \lambda_n + t\left(\frac{\lambda_n}{|\lambda_n|} - \lambda_n\right) \end{pmatrix}.$$

such that $\det \varphi(t) = \prod_{i=1}^{n} \left(\lambda_i + t\left(\frac{\lambda_i}{|\lambda_i|} - \lambda_i\right)\right) \neq 0$ for $0 \le t \le 1$, $\varphi(0) = L$ and $\varphi(1) = E$.

Step III: Since $\det A = \det L > 0$ the number $|I_n^-|$ with $I_n^- = \left\{1 \le i \le n : \frac{\lambda_i}{|\lambda_i|} = -1\right\}$ must be **even** such that for each pair $\{i;j\} \subset I_n^-$ there is a path

$$\varphi(t)=\begin{pmatrix}\ddots & & & & & & & 0\\ & \frac{\lambda_{i-1}}{|\lambda_{i-1}|} & & & & & & \\ & & -\cos\pi t & 0 & \cdots & 0 & -\sin\pi t & \\ & & 0 & \frac{\lambda_{i+1}}{|\lambda_{i+1}|} & & & & \\ & & \vdots & & \ddots & & \vdots & \\ & & 0 & & & \frac{\lambda_{j-1}}{|\lambda_{j-1}|} & 0 & \\ & & \sin\pi t & 0 & \cdots & 0 & -\cos\pi t & \\ & & & & & & & \frac{\lambda_{j+1}}{|\lambda_{j+1}|} \\ 0 & & & & & & & \ddots\end{pmatrix}$$

with $\varphi(0)=E$ and $\varphi(0)=E$' with the values $e_i^i=e_j^j=-1$ transformed to $(e')_i^i=(e')_j^j=-1$ without without leaving $\mathrm{GL}(n;\mathbb{R})$:

$$\det\varphi(t)=\frac{\lambda_1}{|\lambda_1|}\cdot\ldots\cdot\frac{\lambda_{i-1}}{|\lambda_{i-1}|}\cdot\left((\cos\pi t)^2+(\sin\pi t)^2\right)\left(\frac{\lambda_{i+1}}{|\lambda_{i+1}|}\cdot\ldots\cdot\frac{\lambda_{j-1}}{|\lambda_{j-1}|}\right)\cdot\frac{\lambda_{j+1}}{|\lambda_{j+1}|}\cdot\ldots\cdot\frac{\lambda_n}{|\lambda_n|}\neq 0.$$

$\Leftarrow$: According to the hypothesis we have $\varphi(t)\in GL(n;\mathbb{R})$ whence $\det\varphi(t)\neq 0$ for $0\leq t\leq 1$ such that $\det\varphi A\cdot\det\varphi B=\det\varphi(0)\cdot\det\varphi(1)>0$ follows from the **continuity** of $\det: M(n;\mathbb{R})\to\mathbb{R}$ shown in 4.2 and the **intermediate value theorem** [6, th. 5.1].

4.6 The Vandermonde determinant

For $n\geq 1$ and $x\in\mathbb{C}$ with $x_i^n=(x_i)^n$ meaning the n-th power of x_i the **Vandermonde determinant** is defined as

$$\Delta_n(x)=\det\begin{pmatrix}1 & x_1 & \cdots & x_1^{n-1}\\ 1 & x_2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & x_n & \cdots & x_n^{n-1}\end{pmatrix}=\prod_{1\leq i<j\leq n}(x_j-x_i).$$

Proof by induction: The case $n=2$ is obvious. In order to prove the induction step $n\Rightarrow n+1$ we write

$$\det\begin{pmatrix}1 & x_1 & \cdots & x_1^{n-1} & x_1^n\\ 1 & x_2 & \cdots & x_2^{n-1} & x_2^n\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & x_{n+1} & \cdots & x_{n+1}^{n-1} & x_{n+1}^n\end{pmatrix}=\det\begin{pmatrix}1 & x_1-x_{n+1} & \cdots & x_1^{n-1}-x_1^{n-2}\cdot x_{n+1} & x_1^n-x_1^{n-1}\cdot x_{n+1}\\ 1 & x_2-x_{n+1} & \cdots & x_2^{n-1}-x_2^{n-2}\cdot x_{n+1} & x_2^n-x_2^{n-1}\cdot x_{n+1}\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & x_{n+1}-x_{n+1} & \cdots & x_{n+1}^{n-1}-x_{n+1}^{n-1} & x_{n+1}^n-x_{n+1}^n\end{pmatrix}$$

$$=\det\begin{pmatrix}1 & (x_1-x_{n+1}) & \cdots & x_1^{n-2}\cdot(x_1-x_{n+1}) & x_1^{n-1}\cdot(x_1-x_{n+1})\\ 1 & (x_2-x_{n+1}) & \cdots & x_2^{n-2}\cdot(x_2-x_{n+1}) & x_2^{n-1}\cdot(x_2-x_{n+1})\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & 0 & \cdots & 0 & 0\end{pmatrix}$$

$$=1\cdot(x_1-x_{n+1})\cdot\ldots\cdot(x_n-x_{n+1})\cdot\det\begin{pmatrix}1 & x_1 & \cdots & x_1^{n-1}\\ 1 & x_2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & x_n & \cdots & x_n^{n-1}\end{pmatrix}$$

$$=\prod_{1\leq i<j\leq n+1}(x_j-x_i).$$

5 Eigendecomposition

5.1 Eigenvectors and Eigenvalues

A vector $\mathbf{0} \neq \boldsymbol{v} \in X$ is an **Eigenvector** for the **Eigenvalue** $\lambda \in \mathbb{C}$ of the **endomorphism** $\boldsymbol{f} : X \to X$ iff $\boldsymbol{f}(\boldsymbol{v}) = \lambda \boldsymbol{v}$. Obviously the **Eigenspace** $\mathrm{Eig}(\boldsymbol{f}, \lambda) := \mathrm{Ker}(\boldsymbol{f} - \lambda \mathrm{id}) \subset X$ of all Eigenvectors for the Eigenvalue λ is a **vector subspace**. Eigenvectors for different Eigenvalues are **linearly independent**. This is obvious for $n = 2$ and follows by induction for $n > 2$. In the case of a finite dimensional vector space $X = \langle \mathrm{u}_i \rangle_{1 \leq i \leq n}$ the Eigenvalues of $\boldsymbol{f}$ are exactly the zeros of the **characteristic polynom**

$$P_{\mathbf{f}}(t) = \det(F - t \cdot E_n) = \det \begin{pmatrix} f_1^1 - t & & f_n^1 \\ & & \\ f_1^n & & f_n^n - t \end{pmatrix} = (-t)^n + t^{n-1} \cdot \sum_{i=1}^{n} f_i^i + ... + \det F$$

with the **representing matrix** $F = M_{\mathcal{A}}^{\mathcal{A}}(\boldsymbol{f})$ for the given **basis** $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq n}$. The characteristic polynom and the eigenvalues are **independent of the basis** since for any transformation matrix $T \in \mathrm{GL}(n; \mathbb{C})$ we have $\det(T * F * T^{-1} - t \cdot E_n) = \det(T * F * T^{-1} - t \cdot T * E_n * T^{-1}) = \det T * (F - t \cdot E_n) * T^{-1} = \det T \cdot \det(F - t \cdot E_n) \cdot \frac{1}{\det T} = \det(F - t \cdot E_n)$. The basis $(\boldsymbol{v}_i)_{1 \leq i \leq m}$ of $\mathrm{Eig}(\boldsymbol{f}, \lambda) = \langle \boldsymbol{v}_i \rangle_{1 \leq i \leq m}$ can be complemented to a basis $\mathcal{A}' = \{\boldsymbol{v}_1; ...; \boldsymbol{v}_m; \boldsymbol{a}_{m+1}; ...; \boldsymbol{a}_n\}$ of X with $M_{\mathcal{A}'}(\boldsymbol{f}) = \begin{pmatrix} \lambda \cdot E_m & 0 \\ 0 & F' \end{pmatrix}$ and $F' = \left(f_j^i\right)_{m+1 \leq i; j \leq n}$ such that the **dimension** $m = \dim \mathrm{Eig}(\boldsymbol{f}; \lambda) \leq \mu(P_{\boldsymbol{f}}; \lambda)$ **of the Eigenspace cannot exceed the multiplicity of the Eigenvalue** λ in $P_{\boldsymbol{f}}$.

5.2 Trigonalization of complex endomorphisms

For every **endomorphism** $\boldsymbol{f} \in \mathrm{End}(X)$ on an n-dimensional **complex** vector space X there is a basis $\mathcal{B} = (\boldsymbol{v}_i)_{1 \leq i \leq n}$ of $\mathbb{C}^n$ such that $\boldsymbol{f}(\boldsymbol{v}_k) \subset \langle \boldsymbol{v}_i \rangle_{1 \leq i \leq k}$ and

$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$$

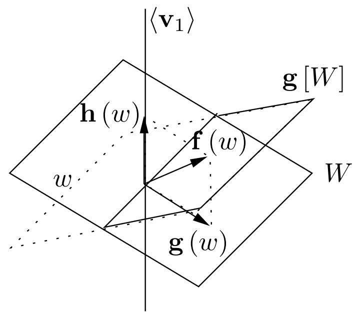
with $P_{\boldsymbol{f}}(\lambda_i) = 0$ for every $1 \leq i \leq n$ and the λ_i not necessarily distinct from each other.

Proof: According to the **fundamental theorem of algebra** [2, th. 5.11] there are (not necessarily distinct!) **eigenvalues** $(\lambda_i)_{1 \leq i \leq n}$ such that $P_{\boldsymbol{f}}(t) = \prod_{1=1}^{n} (\lambda_i - t)$. For $n = 1$ the case is obvious. Assuming the hypothesis for $n - 1$ we choose an **eigenvalue** $\lambda_1 \in \mathbb{C}$ and an **eigenvector** $\mathbf{v}_1 = \sum_{i=1}^{n} v^{i1} \boldsymbol{a}_i \in \mathbb{C}^n$ expressed as linear combination of the basis $\mathcal{A} = (\boldsymbol{a}_1; ...; \boldsymbol{a}_n)$ with $\boldsymbol{f}(\boldsymbol{v}_1) = \lambda_1 \boldsymbol{v}_1$. W.l.o.g. assuming $v_1 \neq 0$ and replacing $\boldsymbol{a}_1$ by $\boldsymbol{v}_1$ we obtain a basis $\mathcal{B}' = \{\boldsymbol{v}_1; \boldsymbol{a}_2; ...; \boldsymbol{a}_n\}$ with the transformation matrix

$$M_{\mathcal{A}}^{\mathcal{B}'}(\mathrm{id}) = \begin{pmatrix} v_1^1 & 0 & \cdots & 0 \\ v_1^2 & 1 & & 0 \\ \vdots & & \ddots & \\ v_1^n & 0 & & 1 \end{pmatrix}$$

comprised of the column vectors of the new basis $\mathcal{B}'$ expressed in linear combinations of the old basis. Hence we obtain $\boldsymbol{f} = T_{\mathcal{B}'}^{\mathcal{A}} \circ \boldsymbol{f}_{\mathcal{A}}^{\mathcal{A}} \circ T_{\mathcal{A}}^{\mathcal{B}'}$ resp. the transition from

$$M_{\mathcal{A}}^{\mathcal{A}}(\boldsymbol{f}) = \begin{pmatrix} f_1^1 & f_2^1 & \cdots & f_n^1 \\ f_1^2 & f_2^2 & \cdots & f_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ f_1^n & f_2^n & \cdots & f_n^n \end{pmatrix} \text{ to } M_{\mathcal{B}'}^{\mathcal{B}'}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 & (f')_2^1 & \cdots & (f')_n^1 \\ 0 & f_2^2 & \cdots & f_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f_2^n & \cdots & f_n^n \end{pmatrix}.$$



In the case of $\dim \operatorname{Eig}(\lambda) < \mu(\lambda)$ the subspace $W_1 = \langle \boldsymbol{a}_i \rangle_{1 \le i \le n-1}$ is not $\mathbf{f}$-invariant and the coefficients $(f')_2^1 ; ...; (f')_n^1$ do not vanish. We circumvene this complication by **splitting** the restriction $\boldsymbol{f}|_W = \boldsymbol{g} + \boldsymbol{h}$ into $\boldsymbol{g} : W_1 \to W_1$ with $\boldsymbol{g}(\boldsymbol{a}_j) = \sum_{i=2}^{n} f_j^i \boldsymbol{a}_i$ and $\boldsymbol{h} : W_1 \to \langle \boldsymbol{v}_1 \rangle$ with $\boldsymbol{h}(\boldsymbol{a}_j) = (f')_j^1 \boldsymbol{v}_1$. Now we can apply the hypothesis to $\boldsymbol{g}$ and find a basis $\mathcal{B}_{W_1} = (\boldsymbol{v}_i)_{2 \le i \le n}$ of $W_1 = \langle \boldsymbol{a}_i \rangle_{2 \le i \le n}$ such that $\boldsymbol{g}(\boldsymbol{v}_k) \subset \langle \boldsymbol{v}_i \rangle_{2 \le i \le k}$. Since the basis $\mathcal{A}_{W_1} = \mathcal{B}'_{W_1}$ did not change on the subspace W_1 the **coordinate transformation** $T_{\mathcal{A}_{W_1}}^{\mathcal{B}_{W_1}} = \left(\Phi_{\mathcal{A}_{W_1}}^{-1}(\boldsymbol{v}_2) ; ...; \Phi_{\mathcal{A}_{W_1}}^{-1}(\boldsymbol{v}_n) \right) : \mathbb{C}^{n-1} \to \mathbb{C}^{n-1}$ given by the **coordinate vectors** $\Phi_{\mathcal{A}_{W_1}}^{-1}(\boldsymbol{v}_i)$ yields $\boldsymbol{g}_{\mathcal{B}_{W_1}}^{\mathcal{B}_{W_1}} = T_{\mathcal{B}_{W_1}}^{\mathcal{A}_{W_1}} \circ \boldsymbol{g}_{\mathcal{A}_{W_1}}^{\mathcal{A}_{W_1}} \circ T_{\mathcal{A}_{W_1}}^{\mathcal{B}_{W_1}}$ with

$$M_{\mathcal{B}_{W_1}}^{\mathcal{B}_{W_1}}(\boldsymbol{g}) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}.$$

Then the basis $\mathcal{B} = \{\boldsymbol{v}_1 ; ...; \boldsymbol{v}_n\}$ with the transformation $T_{\mathcal{B}}^{\mathcal{A}} = T_{\mathcal{B}}^{\mathcal{B}'} \circ T_{\mathcal{B}'}^{\mathcal{A}}$ represented by

$$\left(T_{\mathcal{B}}^{\mathcal{B}'} \right)^{-1} = T_{\mathcal{B}'}^{\mathcal{B}} = (\boldsymbol{v}_1 ; ...; \boldsymbol{v}_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & v_2^2 & & 0 \\ \vdots & \vdots & \ddots & \\ 0 & v_2^n & \cdots & v_n^n \end{pmatrix}$$

expressed in $\mathcal{B}' = \{\boldsymbol{v}_1 ; \boldsymbol{a}_2 ; ...; \boldsymbol{a}_n\}$ by $\boldsymbol{v}_1 = \boldsymbol{v}_1$ resp. $\boldsymbol{v}_i = \sum_{j=2}^{n} v_{ji} \boldsymbol{a}_j$

resp. directly by

$$\left(T_{\mathcal{B}}^{\mathcal{A}} \right)^{-1} = T_{\mathcal{A}}^{\mathcal{B}} = (\boldsymbol{v}_1 ; ...; \boldsymbol{v}_n) = \begin{pmatrix} v_1^1 & & 0 \\ \vdots & \ddots & \\ v_1^n & \cdots & v_n^n \end{pmatrix}$$

expressed in $\mathcal{A} = \{\boldsymbol{a}_1 ; ...; \boldsymbol{a}_n\}$ by $\boldsymbol{v}_1 = \sum_{j=1}^{n} v_1^j \boldsymbol{a}_j$ resp. $\boldsymbol{v}_i = \sum_{j=2}^{n} v_i^j \boldsymbol{a}_j$

results in

$$M_{\mathcal{B}}^{\mathcal{B}}(\mathbf{f}) = \begin{pmatrix} \lambda_1 & (f'')_2^1 & \cdots & (f'')_n^1 \\ 0 & \lambda_2 & & * \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

satisfying the assertion.

Note:

1. The transformation matrix has an **inverse triangular structure** with zeroes obove the main diagonal since all subsequent basis changes only affect the corresponding subspaces W_i in the chain $X \supset W_1 \supset ... \supset W_n$.
2. Every transformation changes every element of $M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f})$ above the main diagonal as indicated by the double dashes in the first row.

Example: For brevity we identify the representing matrices with the corresponding canonical maps on $\mathbb{R}^3$. For $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$ we have $P_A(t) = -(t-2)^3$ with **eigenvalue** $\lambda = 2$, $A - 2E_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$ and $\operatorname{rank}(A - 2E_3) = 2$ such that $\dim \operatorname{Eig}(A;2) = 1 < 3 = \mu(P_A;2)$ whence A cannot be diagonalized. With the **eigenvector** $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ and the completed **basis** $\mathcal{B}' = (\boldsymbol{v}_1; \boldsymbol{e}_2; \boldsymbol{e}_3)$ we obtain $S_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ resp. $S_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ and $A_2 = S_1 * A * S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix}$. On the vector subspace $W = \langle \boldsymbol{e}_2; \boldsymbol{e}_3 \rangle$ we choose an eigenvector $\boldsymbol{v}_2 = \boldsymbol{e}_2 - \boldsymbol{e}_3$ of the restriction with $M(\boldsymbol{f}|_W - 2\mathrm{id}) = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}$. The transformation is carried out by $S_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ resp. $S_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ such that $A_3 = S_2 * A * S_2^{-1} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$.

5.3 The Cayley-Hamilton theorem

For every endomorphism $\boldsymbol{f} \in \operatorname{End}(X)$ on a finite-dimensional vector space X we have $P_{\mathbf{f}}(\boldsymbol{f}) = 0$ with the **characteristic polynom** $P_{\mathbf{f}} \in \mathbb{C}[\boldsymbol{g}]$ on the **commutative subring** $\mathbb{C}[\boldsymbol{g}] \subset \operatorname{End}(X)$ of **polynoms** $\sum_{i=0}^{n} a_i \boldsymbol{g}^i$ with **complex coefficients** $a_i \in \mathbb{C}$ resp. **variables** $\boldsymbol{g} \in (\operatorname{End}(X); +; \circ)$ with $\boldsymbol{g}^0 := \mathrm{id}$, $\boldsymbol{g}^1 := \boldsymbol{g}$, and $\boldsymbol{g}^i := \boldsymbol{g}^{i-1} \circ \boldsymbol{g}$ for $1 \leq i \leq n \in \mathbb{N}$.

Proof: With the notations from the preceding theorem 5.2 we have $P_{\boldsymbol{f}}(\boldsymbol{g}) = \prod_{i=1}^{n} (\lambda_i \mathrm{id} - \boldsymbol{g})$ for every $\boldsymbol{g} \in \operatorname{End}(X)$ and the product referring to the composition. We prove that $\prod_{i=1}^{k} (\lambda_i \mathrm{id} - \boldsymbol{f}) \left[\langle \mathbf{v}_j \rangle_{1 \leq j \leq k} \right] = \{0\}$ for every $1 \leq j \leq k$ and the basis $\mathcal{B} = (\boldsymbol{v}_i)_{1 \leq i \leq n}$ for the **trigonalized form**. For $n = 1$ the case is obvious. Assuming the hypothesis for $k-1$ and choosing an arbitrary $\boldsymbol{w} + \mu \boldsymbol{v}_k$ with $\boldsymbol{w} \in \langle \boldsymbol{v}_j \rangle_{1 \leq j \leq k-1}$ and $\mu \in \mathbb{C}$ the matrix

$$M_{\mathcal{B}}^{\mathcal{B}} \boldsymbol{f} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$$

shows that $\boldsymbol{f}(\boldsymbol{v}_k) - \lambda_k \boldsymbol{v}_k \in \langle \boldsymbol{v}_i \rangle_{1 \leq i \leq k-1}$ and $\boldsymbol{f}(\boldsymbol{w}) \in \langle \boldsymbol{v}_i \rangle_{1 \leq i \leq k-1}$. This implies $(\lambda_k \mathrm{id} - \boldsymbol{f})(\boldsymbol{w} + \mu \boldsymbol{v}_k) \in \langle \boldsymbol{v}_i \rangle_{1 \leq i \leq k-1}$ whence $\prod_{i=1}^{k} (\lambda_i \mathrm{id} - \boldsymbol{f})(\boldsymbol{w} + \mu \boldsymbol{v}_k) = \prod_{i=1}^{k-1} (\lambda_i \mathrm{id} - \boldsymbol{f}) \circ ((\lambda_k \mathrm{id} - \boldsymbol{f})(\boldsymbol{w} + \mu \boldsymbol{v}_k)) = \mathbf{0}$.

5.4 Decomposition of real endomorphisms

For every **endomorphism** $\boldsymbol{f} \in \mathrm{End}(X)$ on an n-dimensional **real** vector space X there is a basis $\mathcal{B} = (\boldsymbol{v}_i)_{1\leq i\leq k} \cup (\boldsymbol{w}_i; \boldsymbol{f}(\boldsymbol{w}_i))_{1\leq i\leq l}$ with $k+2l=n$ of $\mathbb{R}^n$ such that $\boldsymbol{f}(\boldsymbol{v}_j) \subset \langle \boldsymbol{v}_i\rangle_{1\leq i\leq j}$ for $j\leq k$, $\boldsymbol{f}[\langle \boldsymbol{w}_m; \boldsymbol{f}(\boldsymbol{w}_m)\rangle] \subset \langle \boldsymbol{v}_i\rangle_{1\leq i\leq k} \oplus \langle \boldsymbol{w}_i; \boldsymbol{f}(\boldsymbol{w}_i)\rangle_{1\leq i\leq m}$ for $1\leq m\leq l$ and

$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 & & & & & & & * \\ & \ddots & & & & & & \\ & & \lambda_k & & & & & \\ & & & 0 & -q_1 & & & \\ & & & 1 & -p_1 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & -q_m \\ 0 & & & & & & 1 & -p_m \end{pmatrix}.$$

Proof: According to the **fundamental theorem of algebra** [2, th. 2.10] resp. the **Euclidean division algorithm** for polynomials there are **real eigenvalues** $(\lambda_j)_{1\leq i\leq k} \subset \mathbb{R}$ and **real coefficients** $(p_i; q_i)_{1\leq i\leq l} \subset \mathbb{R}$ such that $P_{\boldsymbol{f}}(t) = \prod_{i=1}^{k} (\lambda_i - t) \prod_{i=1}^{l} (t^2 + p_i t + q_i)$. Note that the λ_i are not necessarily different from each other and that the trigonalisation of the restriction $\boldsymbol{f}|_V$ for $V = \langle \boldsymbol{v}_i\rangle_{1\leq i\leq k}$ is guaranteed by 5.2. Similarly to the trigonalization we now split the the restriction $\boldsymbol{f}|_W = \boldsymbol{h}_1 + \boldsymbol{g}_1$ with $\boldsymbol{h}_1 : W \to W$ and $\boldsymbol{g}_1 : W \to V$ on the complementing vector subspace W with $X = V \oplus W$. For an arbitrary $\boldsymbol{w} \in W$ according to the **Cayley-Hamilton theorem** 5.3 we can arrange the order of the factors in the characteristic polynomial such that the iterated composition holds

$$\boldsymbol{w}_1 := \left(\prod_{i=2}^{m} \left(\boldsymbol{h}_1^2 + p_i \cdot \boldsymbol{h}_1 + q_i \cdot \mathrm{id}|_W\right)\right)(\boldsymbol{w}) \neq \boldsymbol{0}$$

and

$$\boldsymbol{h}_1(\boldsymbol{h}_1(\boldsymbol{w}_1)) + p_1 \cdot \boldsymbol{h}_1(\boldsymbol{w}_1) + q_1 \cdot \boldsymbol{w}_1 = \left(\boldsymbol{h}_1^2 + p_1 \cdot \boldsymbol{h}_1 + q_1 \cdot \mathrm{id}|_W\right)(\boldsymbol{w}_1) = \boldsymbol{0},$$

i.e. $\boldsymbol{h}_1(\boldsymbol{h}_1(\boldsymbol{w}_1)) = -p_1 \cdot \boldsymbol{h}_1(\boldsymbol{w}_1) - q_1 \cdot \boldsymbol{w}_1$. Since $\boldsymbol{h}_1$ has no eigenvectors $\boldsymbol{w}_1$ and $\boldsymbol{h}_1(\boldsymbol{w}_1)$ are linearly independent such that we have obtained an $\boldsymbol{h}_1$-invariant subspace $W_1 = \langle \boldsymbol{w}_1; \boldsymbol{h}_1(\boldsymbol{w}_1)\rangle \subset W$. W.l.o.g. assuming nonzero coefficients in the linear combinations for $\boldsymbol{w}_1$ and $\boldsymbol{h}_1(\boldsymbol{w}_1)$ in terms of the original basis we replace the first two of the previous basis vectors of W by $\boldsymbol{w}_1$ and $\boldsymbol{h}_1(\boldsymbol{w}_1)$ the representing matrix with reference to the new basis $\mathcal{B}_1$ has the form

$$M_{\mathcal{B}_1}^{\mathcal{B}_1}(\boldsymbol{h}_1) = \begin{pmatrix} 0 & -q_1 & * & \cdots & * \\ 1 & -p_1 & \vdots & & \vdots \\ 0 & 0 & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \cdots & * \end{pmatrix}.$$

By induction we proceed with the restriction $\boldsymbol{f}|_{Y_1}$ on the complementing vector subspace Y_1 with $W = W_1 \oplus Y_1$ until we have a decomposition $X = V \oplus W = V \oplus W_1 \oplus Y_1 = V \oplus W_1 \oplus W_2 \oplus Y_2 = ... = V \oplus W_1 \oplus ... \oplus W_l$ such that $\boldsymbol{h}_i[W_i] \subset W_i$. Similarly to the trigonalization of complex matrices 5.2 the transformation matrices have the form

$$M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = \begin{pmatrix} v_1^1 & & & & & & & & & 0 \\ \vdots & \ddots & & & & & & & & \\ & & v_k^k & & & & & & & \\ & & \vdots & w_1^{k+1} & (h_1)^{k+1}(\boldsymbol{w}_1) & & & & & \\ & & & \vdots & \vdots & & & & & \\ & & & & & & \ddots & & & \\ \vdots & & \vdots & \vdots & \vdots & & & w_l^{n-1} & (h_l)^{n-1}(\boldsymbol{w}_l) \\ v_1^n & \cdots & v_k^n & w_1^n & (h_1)^n(\boldsymbol{w}_1) & & \cdots & w_l^n & (h_1)^n(\boldsymbol{w}_l) \end{pmatrix}.$$

Hence the transition from $\mathcal{B}_i$ to $\mathcal{B}_{i+1}$ will change the elements **above** the 2×2 units in an undetermined manner but leave the zeroes **beneath** unaffected such that we finally arrive at the asserted structure of $M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f})$.

5.5 Minimal polynoms

According to 2 for every $\boldsymbol{f} \in \operatorname{End}(X)$ with $\dim X = n \in \mathbb{N}$ the **ideal** $I_{\boldsymbol{f}} = \{Q \in \mathbb{C}[z] : Q(\boldsymbol{f}) = 0\}$ is **principal**, i.e. $I_{\boldsymbol{f}} = \langle M_{\boldsymbol{f}} \rangle$ with the **minimal polynom** $M_{\boldsymbol{f}}$. Since $M_{\boldsymbol{f}} \in I_{\boldsymbol{f}}$ the minimal polynom must have at least the same zeroes λ_j as $P_{\boldsymbol{f}}$. Since it is a divisor of $P_{\boldsymbol{f}}$ their multiplicities cannot exceed those of $M_{\boldsymbol{f}}$. Hence for $P_{\boldsymbol{f}}(t) = \prod_{j=1}^{k} (\lambda_j - t)^{r_j}$ with $\sum_{j=1}^{k} r_j = n$ we have $M_{\boldsymbol{f}}(t) = \prod_{j=1}^{k} (\lambda_j - t)^{d_j}$ with $1 \le d_j \le r_j$.

For an endomorphism $\boldsymbol{g} \in \operatorname{End}(X)$ with $\dim X = n \in \mathbb{N}$ the following conditions are equivalent:

1. $\boldsymbol{g}$ is **nilpotent**, i.e. $\boldsymbol{g}^k = 0$ for some $k \in \mathbb{N}$
2. $\boldsymbol{g}^k = 0$ for some $1 \le k \le n$
3. $P_{\boldsymbol{g}}(t) = \pm t^n$
4. There is a **basis** $\mathcal{B}$ of X such that $M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{g}) = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}$

Proof: 1. ⇒ 2. ⇒ 3. ⇒ 4. ⇒ 1. directly follows from the definitions resp. the preceding paragraph.

5.6 Fitting's lemma

For an endomorphism $\boldsymbol{g} \in \operatorname{End}(X)$ with $\dim X = n \in \mathbb{N}$ and $d = \min\left\{l : \ker \boldsymbol{g}^l = \ker \boldsymbol{g}^{l+1}\right\}$ we have

1. $d = \min\left\{l : \operatorname{im}\boldsymbol{g}^l = \operatorname{im}\boldsymbol{g}^{l+1}\right\}$.
2. $\ker \boldsymbol{g}^{d+i} = \ker \boldsymbol{g}^d$ and $\operatorname{im}\boldsymbol{g}^{d+i} = \operatorname{im}\boldsymbol{g}^d$ for every $i \in \mathbb{N}$.
3. $\boldsymbol{g}[U] \subset U$ for $U = \ker \boldsymbol{g}^d$ and $\boldsymbol{g}[V] \subset V$ for $V = \operatorname{im}\boldsymbol{g}^d$.
4. $(\boldsymbol{g}|_U)^d = 0$ and $\boldsymbol{g}|_V$ is an isomorphism.
5. $M_{\boldsymbol{g}|U}(t) = t^d$.
6. $X = U \oplus V$ with $\dim U = r \ge d$ with $r = \mu(P_{\boldsymbol{g}}; 0)$.
7. There is a basis $\mathcal{B}$ of X such that $M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{g}) = \begin{pmatrix} N & 0 \\ 0 & C \end{pmatrix}$ with $N^d = 0$ and $C \in \operatorname{GL}(n - r; \mathbb{C})$.

Proof:

1. – 4.: By 3.7 we conclude that

$$\begin{array}{ccccc} \ker \mathbf{g}^l & \subset & V & \xrightarrow{\mathbf{g}^l} & \text{im } \mathbf{g}^l \\ \cap & & \| & & \cup \\ \ker \mathbf{g}^{l+1} & \subset & V & \xrightarrow{\mathbf{g}^{l+1}} & \text{im } \mathbf{g}^{l+1} \end{array}$$

$$\begin{aligned} \text{im}\boldsymbol{g}^{l+1} = \text{im}\boldsymbol{g}^{l} &\Leftrightarrow \dim \text{im}\boldsymbol{g}^{l+1} = \dim \text{im}\boldsymbol{g}^{l} \\ &\Leftrightarrow \dim\ker \boldsymbol{g}^{l+1} = \dim\ker \boldsymbol{g}^{l} \\ &\Leftrightarrow \ker \boldsymbol{g}^{l+1} = \ker \boldsymbol{g}^{l} \end{aligned}$$

whence $g|_{\text{im}\boldsymbol{g}^l} : \text{im}\boldsymbol{g}^l \to \text{im}\boldsymbol{g}^{l+1}$ is an isomorphism.

5.: Assuming $M_{\boldsymbol{g}|U}(t) = t^{d-1}$resp. $(\boldsymbol{g}|_U)^{d-1} = \mathbf{0}$ would imply $\ker \boldsymbol{g}^d \subset \ker \boldsymbol{g}^{d-1}$ contrary to the minimal character of d.
6.: For $v \in U \cap W$ we have $\boldsymbol{g}^d(\boldsymbol{v}) = \mathbf{0}$ and a $\boldsymbol{w} \in V$ with $\boldsymbol{g}^d(\boldsymbol{w}) = \boldsymbol{v}$ whence $\boldsymbol{g}^{2d}(\boldsymbol{w}) = \mathbf{0}$, i.e. $w \in \ker \boldsymbol{g}^{2d} = \ker \boldsymbol{g}^d$ such that $\mathbf{0} = \boldsymbol{g}^d(\boldsymbol{w}) = \boldsymbol{v}$. Hence we conclude that $X = U \oplus V$. Owing to $\ker \boldsymbol{g}^{i-1} \subsetneq \ker \boldsymbol{g}^i$ whence $\dim\ker \boldsymbol{g}^{i-1} < \dim\ker \boldsymbol{g}^i$ for $1 \le i \le d$ we have $\dim U \ge d$. With $r = \mu(P_{\boldsymbol{g}}; 0)$ we have $t^r \cdot Q(t) = P_{\boldsymbol{g}}(t) = P_{\boldsymbol{g}|_U}(t) \cdot P_{\boldsymbol{g}|_V}(t)$ for some polynomial Q with $Q(0) \neq 0$. By 5.5.3 we have $P_{\boldsymbol{g}|_U}(t) = \pm t^m$ with $m = \dim U$ whence $\mu(P_{\boldsymbol{g}}; 0) = r = m = \dim U$ since for the characteristic polynomial of the isomorphism $\boldsymbol{g}|_V$ holds $P_{\boldsymbol{g}|_V}(0) \neq 0$.

5.7 Generalized eigenspaces

For every $\boldsymbol{f} \in \text{End}(X)$ with $\dim X = n \in \mathbb{N}$ and characteristic polynomial $P_{\boldsymbol{f}}(t) = \prod_{j=1}^{k} (\lambda_j - t)^{r_j}$ with $\sum_{j=1}^{k} r_j = n$ there exists a decomposition into **generalized eigenspaces** (*Haupträume*) $U_j = \text{Hau}(\boldsymbol{f}; \lambda_i) = \langle \mathcal{B}_j \rangle$ with bases $\mathcal{B}_j$ for $1 \le j \le k$ such that

1. $\boldsymbol{f}[U_j] \subset U_j$ and $\dim U_j = r_j$
2. $X = U_1 \oplus ... \oplus U_k$
3. $M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 E_{r_1} + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k E_{r_k} + N_k \end{pmatrix}$ with nilpotent matrices $N_j \in \text{M}(r_j; \mathbb{C})$.

Proof by induction over the number k of eigenvalues: For $\boldsymbol{g} = \boldsymbol{f} - \lambda_1 \text{id}$ we have $P_{\boldsymbol{g}}(t - \lambda_1) = P_{\boldsymbol{f}}(t)$ whence $r_1 = \mu(P_{\boldsymbol{g}}; 0) = \mu(P_{\boldsymbol{f}}; \lambda_1)$ such that **Fitting's lemma** 5.6 yields $X = U_1 \oplus W$ with $\boldsymbol{g}[U_1] \subset U_1$ resp. $\boldsymbol{f}[U_1] \subset U_1$ and $\boldsymbol{g}[W] \subset W$ resp. $\boldsymbol{f}[W] \subset W$. The representing matrices have the form

$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} N_1 & 0 \\ 0 & C \end{pmatrix} \text{ with } N_1 = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix} \in M(r_1; \mathbb{C}) \text{ and } M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{g}|_W) = C \in GL(n - r_1; \mathbb{C})$$

resp.

$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 E_{r_1} + N_1 & 0 \\ 0 & D \end{pmatrix} \text{ with } M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}|_W) = D \in GL(n - r_1; \mathbb{C}).$$

The induction hypothesis then applies to the isomorphism $\boldsymbol{f}|_W$ with the characteristic polynom $P_{\boldsymbol{f}|_W}(t) = \prod_{j=2}^{k} (\lambda_j - t)^{r_j}$ which proves the theorem.

5.8 The Jordan decomposition

For every **nilpotent** endomorphism $\boldsymbol{g} \in \text{End}(X)$ with $\dim X = r \in \mathbb{N}$ and $d = \min\left\{l \in \mathbb{N} : \boldsymbol{g}^l = \mathbf{0}\right\}$ there exist uniquely determined numbers $s_j \in \mathbb{N}$ such that $\sum_{j=1}^{d} j \cdot s_j = r$ and a basis $\mathcal{B}$ of X such that

$$M_{\mathcal{B}}^{\mathcal{B}}\left(\boldsymbol{g}\right)=\begin{pmatrix} J_d & & & & & & 0 \\ & \ddots & & & & & \\ & & J_d & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix}$$

with the **Jordan matrices** $J_j=\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}\in M\left(j;\mathbb{R}\right)$ occurring s_j times for $1\leq j\leq d$.

Proof: Consider the chain $\{0\}=U_0\subsetneq U_1\subsetneq\ldots\subsetneq U_d=X$ for $U_j=\ker\boldsymbol{g}^j$ with $\boldsymbol{g}^{-1}\left[U_{j-1}\right]=U_j$ and in particular $\boldsymbol{g}\left[U_j\right]\subset U_{j-1}$. Since for every vector subspace W with $W\cap U_j=\{\mathbf{0}\}$ the restriction $\boldsymbol{g}|_W$ is injective for every $1\leq j\leq d$ there is a vector subspace W_j such that $U_j=U_{j-1}\oplus W_j$ with $\boldsymbol{g}\left[W_j\right]\subset U_{j-1}$ and $\boldsymbol{g}\left[W_j\right]\subset U_{j-2}=\emptyset$. Hence we obtain a decomposition according to the following diagram:

$$\begin{array}{cccccccccccc}
 & & & & & & & & X & = & U_d \\
 & & & & & & & & & \supset\mathbf{g}(U_d) & \cup \\
 & & & & & & = & & U_{d-1} & \oplus & W_d \\
 & & & & & & \supset\mathbf{g}(U_{d-1}) & & \cup & \supset\mathbf{g}(W_d) & \| \\
 & & & & = & & U_{d-2} & \oplus & W_{d-1} & \oplus & W_d \\
 & & & & \supset\mathbf{g}(U_1)\supset\ldots\supset\mathbf{g}(U_{d-2}) & & \supset\mathbf{g}(W_2)\supset\ldots\supset\mathbf{g}(W_{d-1}) & & \| & & \| \\
 & & & & & & & & \vdots & & \vdots \\
 & & & & & & & & \| & \supset\mathbf{g}(W_d) & \| \\
= & U_0 & \oplus & W_1 & \oplus & \cdots & & \oplus & W_{d-1} & \oplus & W_d \\
 & \| & & & & & & & & & \\
 & \{0\} & & & & & & & & &
\end{array}$$

In order to provide the corresponding bases we complete each U_{j-1} with some W_j such that $U_j=U_{j-1}\oplus W_j$ and making use of the basis of the previous completion by $\boldsymbol{g}\left[W_{j+1}\right]\subset W_j$:

$$\begin{array}{lcl}
W_d & = & \left\langle \boldsymbol{w}_1^{(d)}\ ,\ldots,\ \boldsymbol{w}_{s_d}^{(d)}\right\rangle \\
W_{d-1} & = & \left\langle \boldsymbol{g}\left(\boldsymbol{w}_1^{(d)}\right)\ ,\ldots,\ \boldsymbol{g}\left(\boldsymbol{w}_{s_d}^{(d)}\right)\ ,\ \boldsymbol{w}_1^{(d-1)}\ ,\ldots,\ \boldsymbol{w}_{s_{d-1}}^{(d-1)}\right\rangle \\
\vdots & & \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
W_1 & = & \left\langle \boldsymbol{g}^{d-1}\left(\boldsymbol{w}_1^{(d)}\right)\ ,\ldots,\ \boldsymbol{g}^{d-1}\left(\boldsymbol{w}_{s_{d-1}}^{(d)}\right)\ ,\ \boldsymbol{g}^{d-2}\left(\boldsymbol{w}_1^{(d-1)}\right)\ ,\ldots,\ \boldsymbol{g}^{d-2}\left(\boldsymbol{w}_{s_{d-1}}^{(d-1)}\right)\ ,\ldots,\ \boldsymbol{w}_1^{(1)}\ ,\ldots,\ \boldsymbol{w}_{s_1}^{(1)}\right\rangle
\end{array}$$

with

$$W_1=U_1=\ker\boldsymbol{g}$$

The matrix $M_{\mathcal{B}}^{\mathcal{B}}\left(\boldsymbol{g}\right)$ obtains the asserted form if the basis vectors in the above pattern are taken from each column upwards starting on the left column with $\boldsymbol{g}^{d-1}\left(\boldsymbol{w}_1^{(d)}\right)$, moving upwards to $\boldsymbol{w}_1^{(d)}$ then working up the the next column starting with $\boldsymbol{g}^{d-1}\left(\boldsymbol{w}_2^{(d)}\right)$ up to $\boldsymbol{w}_2^{(d)}$ and so on until we close with s_1 null vectors for $\boldsymbol{w}_1^{(1)}, ..., \boldsymbol{w}_{s_1}^{(1)}$.

Example:

As before we restrict the exposition to the matrix level. We want to separate as far as possible the components of $\boldsymbol{f} : \mathbb{R}^5 \to \mathbb{R}^5$ defined by

$$\mathcal{M}(\boldsymbol{f}) = A = \begin{pmatrix} 2 & 1 & 1 & 0 & -2 \\ 1 & 1 & 1 & 0 & -1 \\ 1 & 0 & 2 & 0 & -1 \\ 1 & 0 & 1 & 2 & -2 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

with the **characteristic polynomial**

$$P_A(t) = -(t-2)^2 \cdot (t-1)^3$$

Its **eigenvalues** are $\lambda_1 = 2$ with **mulitplicity** $r_1 = 2$ and $\lambda_2 = 1$ wit $r_2 = 3$. For the sake of simplification in the following calculations we identify matrices with maps. The map defined by

$$A - \lambda_1 E_5 = \begin{pmatrix} 0 & 1 & 1 & 0 & -2 \\ 1 & -1 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 & -2 \\ 1 & 0 & 1 & 0 & -2 \end{pmatrix} \cong \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

provides the **eigenspace**

$$V_1 = \ker(A - \lambda_1 E_5) = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}\right).$$

Since $\dim V_1 = 2 = r_2$ the restriction $\boldsymbol{f}|_{V_1}$can be **diagonalized**:

$$\text{With } S^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ we obtain } B = S * A * S^{-1} = \begin{pmatrix} 2 & 0 & 1 & 1 & -2 \\ 0 & 2 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

The second **eigenspace** $V_2 = \operatorname{Ker}(B - \lambda_2 E_5)$ is determined by

$$B - \lambda_2 E_5 = \begin{pmatrix} 1 & 0 & 1 & 1 & -2 \\ 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \cong \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

with $\dim\ker(B - \lambda_2 E_5) = 1 < 3 = r_2$ such that it is **not diagonizable**. By

$$\ker(B - \lambda_2 E_5) = \operatorname{span}\left(\begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}\right) \text{ and } T^{-1} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

we obtain

$$C = T * B * T^{-1} = \begin{pmatrix} 2 & 0 & 0 & 0 & -1 \\ 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \text{ with } C - \lambda_2 E_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

whence

$$\ker(C - \lambda_2 E_5) = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}\right) \text{ and } U^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

and the **trigonalized form**

$$D = U * C * U^{-1} = \begin{pmatrix} 2 & 0 & 0 & -1 & -1 \\ 0 & 2 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

For the **Jordan decomposition** we consider the **kernels** $U_i^l := \ker \boldsymbol{g}_i^l$ of the **powers** of $\boldsymbol{g}_i = \mathrm{f} - \lambda_i \mathrm{id}$. Concerning V_1 we have

$$(A - \lambda_1 E_5)^2 = \begin{pmatrix} 0 & -1 & -1 & 0 & 2 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 & 1 \end{pmatrix} \cong \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

with

$$\ker\left((A - \lambda_1 E_5)^2\right) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

$$\text{which by } S^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ leads to } B = S * A * S^{-1} = \begin{pmatrix} 2 & 0 & 1 & 1 & -2 \\ 0 & 2 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

as before. Concerning $V_2 = \operatorname{Ker}(B - \lambda_2 E_5)$ we observe

$$B - \lambda_2 E_5 = \begin{pmatrix} 1 & 0 & 1 & 1 & -2 \\ 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \cong \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

with $\dim\ker(B-\lambda_2 E_5) = 1 < 3 = r_2$ such that B is **not diagonizable**. But its power

$$(B-\lambda_2 E_5)^2 \cong \begin{pmatrix} 1 & 0 & 1 & 1 & -2 \\ 0 & 1 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

has $\dim\ker(B-\lambda_2 E_5)^2 = 3 = r_2$ such that

$$\ker(B-\lambda_2 E_5)^2 = \left\langle \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \text{ and } T^{-1} = \begin{pmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

yields the **eigen decomposition** resp. **separation of eigenspaces** by

$$C = T * B * T^{-1} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

The **Jordan decomposition** is attained by a further transformation of the mixed component

$$\ker\left(C|_{\mathrm{Eig}(C,\lambda_2)} - \lambda_2 E_3\right) = \ker \begin{pmatrix} 0 & -1 & -1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle = \mathrm{Eig}(C,\lambda_2)$$

such that

$$U^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \text{ yields } D = U * C * U^{-1} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

In a final step we reduce the **nilpotent endomorphism** represented by $M(\boldsymbol{g}) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ with $\boldsymbol{g}^2 = 0$ into a **Jordan matrix**. With $d = 2$ we have

$$\mathbb{C}^3 = U_2 = U_1 \oplus W_2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

with

$$\boldsymbol{w}_1^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{g}\left(\boldsymbol{w}_1^{(2)}\right) = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \text{ and } \boldsymbol{w}_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

such that

$$V^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \text{ whence finally } E = V * D * V^{-1} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

6 Unitary and euclidean vector spaces

6.1 Sesquilinear forms

A map $\langle\rangle : X \times X \to \mathbb{C}$ on a **complex vector space** X is

1. **sesquilinear** iff for $\boldsymbol{x};\boldsymbol{y};\boldsymbol{z} \in X$ and $\alpha;\beta \in \mathbb{C}$ holds
   $\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}; \boldsymbol{z}\rangle = \alpha\langle\boldsymbol{x};\boldsymbol{z}\rangle + \beta\langle \mathrm{y};\boldsymbol{z}\rangle$ (**linearity** in the first component)
   $\langle \boldsymbol{x}; \alpha\mathrm{y} + \beta\boldsymbol{z}\rangle = \overline{\alpha}\langle\boldsymbol{x};\boldsymbol{y}\rangle + \overline{\beta}\langle\boldsymbol{x};\boldsymbol{z}\rangle$ (**conjugate linearity** in the second component)
2. **hermitian** iff for $\boldsymbol{x};\boldsymbol{y} \in X$ holds $\langle\boldsymbol{x};\boldsymbol{y}\rangle = \overline{\langle\boldsymbol{y};\boldsymbol{x}\rangle}$ (**conjugate symmetry**)
3. **positive definite** iff for $\boldsymbol{x} \in X \setminus \{\boldsymbol{0}\}$ holds $\langle\boldsymbol{x};\boldsymbol{x}\rangle > 0$.

With 1. and 2. the map s is a **scalar product** and with all three properties it is an **inner product**. Note that 2. implies $\langle\boldsymbol{x};\boldsymbol{x}\rangle \in \mathbb{R}$. According to [6, th. 1.3] every inner product by $\|\boldsymbol{x}\| = \sqrt{\langle\boldsymbol{x};\boldsymbol{x}\rangle}$ generates a **norm** $\|\| : X \to \mathbb{R}_0^+$ which by $d(\boldsymbol{x};\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$ produces a **metric** $d : X \times X \to \mathbb{R}_0^+$.

A **unitary space** resp**. euclidean** is a pair $(X;\langle\rangle)$ of a **complex** resp. **real** vector space X and a **scalar product**. In the case of a **real vector space** the properties 6.1.1 resp. 6.1.2 become **bilinearity** resp. **symmetry**. According to [6, th. 14.8] a space $(X;\langle\rangle)$ with an **inner product** can be **embedded** into a **complete Hilbert space**.

6.2 Bases

A scalar product $\langle\rangle : X^2 \to \mathbb{C}$ is determined by its values $\langle\boldsymbol{a}_i;\boldsymbol{a}_j\rangle_{i;j\in I}$ on a **basis** $\mathcal{A} = (\boldsymbol{a}_i)_{i\in I}$ of X. In the finite dimensional case with $\dim X = n \in \mathbb{N}$ it is represented by a **hermitian covariant matrix** $S_\mathcal{A} = s_{\mathcal{A}ij} = \langle\boldsymbol{a}_i;\boldsymbol{a}_j\rangle$ with $\langle\boldsymbol{x};\boldsymbol{y}\rangle = x^i_\mathcal{A} s_{\mathcal{A}ij}\overline{y}^j_\mathcal{A} = {}^T\boldsymbol{x}_\mathcal{A} * S_\mathcal{A} * \overline{\boldsymbol{y}}_\mathcal{A}$ for $\boldsymbol{x} = x^i_\mathcal{A}\boldsymbol{a}_i$ resp. $\boldsymbol{y} = y^j_\mathcal{A}\boldsymbol{a}_j$. Owing to 6.1.2 a quadratic matrix $S \in \mathrm{M}(n;\mathbb{C})$ is **hermitian** iff ${}^T S = \overline{S}$ and it is **positive definite** iff ${}^T\boldsymbol{x} * S * \overline{\boldsymbol{x}} \in \mathbb{R}_0^+$ for every $\boldsymbol{x} \in \mathbb{C}^n$.

6.3 Coordinate transformation

According to 3.10 the transformation from the basis $\mathcal{A} = (\boldsymbol{a}_i)_{1\le i\le n}$ to another basis $\mathcal{B} = (\boldsymbol{b}_j)_{1\le j\le n}$ with $\boldsymbol{b}_j = t^i_j\boldsymbol{a}_i$ is determined by the **transformation matrix** $T^\mathcal{B}_\mathcal{A} = t^i_j$ such that the **coordinate vectors** $\boldsymbol{x}_\mathcal{A} = x^i_\mathcal{A}\boldsymbol{e}_i$ resp. $\boldsymbol{x}_\mathcal{B} = x^j_\mathcal{B}\boldsymbol{e}_j$ of every $\boldsymbol{x} = x^k_\mathcal{B}\boldsymbol{b}_k = x^i_\mathcal{A}\boldsymbol{a}_i = x^k_\mathcal{B}t^i_k\boldsymbol{a}_i \in X$ are transformed by $\boldsymbol{x}_\mathcal{A} = T^\mathcal{B}_\mathcal{A} * \boldsymbol{x}_\mathcal{B}$. Consequently we have

$$\begin{aligned}
\langle\boldsymbol{x};\boldsymbol{y}\rangle = x^i_\mathcal{A}s_{\mathcal{A}ij}\overline{y}^j_\mathcal{A} \quad &= {}^T\boldsymbol{x}_\mathcal{A} * S_\mathcal{A} * \overline{\boldsymbol{y}_\mathcal{A}} \\
= x^k_\mathcal{B}t^i_k s_{\mathcal{A}ij}\overline{y}^l_\mathcal{B}\overline{t}^j_l \quad &= {}^T\left(T^\mathcal{B}_\mathcal{A} * \boldsymbol{x}_\mathcal{B}\right) * S_\mathcal{A} * \overline{T^\mathcal{B}_\mathcal{A} * \boldsymbol{y}_\mathcal{B}} \\
= x^k_\mathcal{B}t^i_k s_{\mathcal{A}ij}\overline{t}^j_l\overline{y}^l_\mathcal{B} \quad &= {}^T\boldsymbol{x}_\mathcal{B} * {}^T T^\mathcal{B}_\mathcal{A} * S_\mathcal{A} * \overline{T^\mathcal{B}_\mathcal{A}} * \overline{\boldsymbol{y}_\mathcal{B}} \\
= x^k_\mathcal{B}t^i_k s_{\mathcal{B}kl}\overline{t}^j_l\overline{y}^l_\mathcal{B} \quad &= {}^T\boldsymbol{x}_\mathcal{B} * S_\mathcal{B} * \overline{\boldsymbol{y}_\mathcal{B}}
\end{aligned}$$

with

$$s_{\mathcal{B}kl} = t^i_k s_{\mathcal{A}ij} t^j_l$$

resp.

$$S_\mathcal{B} = {}^T T^\mathcal{B}_\mathcal{A} * S_\mathcal{A} * \overline{T^\mathcal{B}_\mathcal{A}}$$

which proves the **covariant** character resp. the type $(0;2)$ of the representing tensor S.

6.4 The Gram-Schmidt-Orthonormalization

Two vectors $\boldsymbol{u};\boldsymbol{v}\in X$ on a **unitary vector space** $(X,\langle\rangle)$ are **orthogonal** iff $\langle\boldsymbol{u},\boldsymbol{v}\rangle=0$ and they are **normal** iff $\|\boldsymbol{u}\|=\|\boldsymbol{v}\|=1$. Every finite dimensional vector space X has an **orthonormal** basis since according to 3.3 for any given basis $(\boldsymbol{a}_i)_{1\le i\le n}$ the basis $(\boldsymbol{b}_i)_{1\le i\le n}$ **inductively** defined by

$$\boldsymbol{b}_1=\boldsymbol{a}_1 \text{ and } \boldsymbol{b}_i=\boldsymbol{a}_i-\sum_{k=1}^{i-1}\frac{\langle\boldsymbol{a}_i,\boldsymbol{b}_k\rangle}{\langle\boldsymbol{a}_k,\boldsymbol{b}_k\rangle}\boldsymbol{b}_k \text{ for } 2\le i\le n$$

is **orthogonal** and the basis $(\boldsymbol{q}_i)_{1\le i\le n}$ with $\boldsymbol{q}_i=\frac{\boldsymbol{b}_i}{\|\boldsymbol{b}_i\|}$ is **orthonormal** with $\left\langle\boldsymbol{q}_i;\boldsymbol{q}_j\right\rangle=\delta_{ij}$ and $\langle\boldsymbol{q}_i;\boldsymbol{a}_i\rangle=\langle\boldsymbol{q}_i;\boldsymbol{b}_i\rangle=\|\boldsymbol{b}_i\|$.

Obviously for every vector subspace $V\subset X$ its **orthogonal complement** $V^\perp=\{\boldsymbol{u}\in X:\langle\boldsymbol{u},\boldsymbol{v}\rangle=0\,\forall\boldsymbol{v}\in V\}$ is a vector subspace. Due to the above desribed **Gram-Schmidt orthonormalisation** for every vector subspace $V\subset X$ we have an **orthogonal decomposition** $X=V\oplus V^\perp$ with

$$\dim X=\dim V+\dim V^\perp.$$

Also every **invertible matrix** $A\in GL(n;\mathbb{C})$ has a QR**-decomposition**

$$A=Q*R$$

into a **unitary matrix** $Q\in U(n;\mathbb{C})$ with $Q^{-1}={}^T\overline{Q}$ (cf. 6.6.1) and an **upper triangular matrix** $R\in GL(n;\mathbb{C})$ since the **Gram-Schmidt-orthonormalization** of the column vectors of the given matrix $A=(\boldsymbol{a}_1;...;\boldsymbol{a}_n)$ produces **orthonormal** column vectors of a **unitary** $Q=(\boldsymbol{q}_1;...;\boldsymbol{q}_n)\in GL(n;\mathbb{C})$ with $\boldsymbol{q}_j=\sum_{i=1}^{j}s_j^i\boldsymbol{a}_i$ such that $S=\left(s_j^i\right)_{1\le i;j\le n}\in GL(n;\mathbb{C})$ with $s_j^i=0\Leftrightarrow i>j$ is an upper triangular matrix with $q_j^k=a_i^k s_j^i$. Solving the Gram-Schmidt equations for the original basis $(\boldsymbol{a}_i)_{1\le i\le n}$ yields $\boldsymbol{a}_i=\sum_{j=1}^{i}\left\langle\boldsymbol{q}_j;\boldsymbol{a}_i\right\rangle\boldsymbol{q}_j$ whence $A=Q*R$ with the **inverse** $R=S^{-1}=r_i^j=\left\langle\boldsymbol{q}_j;\boldsymbol{a}_i\right\rangle_{1\le j\le i\le n}\in GL(n;\mathbb{C})$ and $r_i^j=0\Leftrightarrow j>i$, i.e. R is again an **upper triangular matrix**.

6.5 Geometric formulae

For any $\boldsymbol{u};\boldsymbol{v};\boldsymbol{u}_i\in X$ on a **unitary vector space** $(X,\langle\rangle)$ resp. **real vectors** $\boldsymbol{x}_i\in\mathbb{R}^n$ and $1\le i\le n$ we have

1. The **Cauchy-Schwarz inequality**: $|\langle\boldsymbol{u},\boldsymbol{v}\rangle|\le\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|$
2. The **Triangle inequality I**: $\|\boldsymbol{u}+\boldsymbol{v}\|\le\|\boldsymbol{u}\|+\|\boldsymbol{v}\|$
   with **equality** if **x** and **y** are **orthogonal**. (**Pythagoras equality**)
3. The **Triangle inequality II**: $|\|\boldsymbol{u}\|-\|\boldsymbol{v}\||\le\|\boldsymbol{u}-\boldsymbol{v}\|$
4. The **Parallelogram equality**: $\|\boldsymbol{u}+\boldsymbol{v}\|+\|\boldsymbol{u}-\boldsymbol{v}\|=2\|\boldsymbol{u}\|+2\|\boldsymbol{v}\|$
5. The **Polarisation equality**: $\langle\boldsymbol{u},\mathbf{v}\rangle=\frac{1}{4}\left(\|\boldsymbol{u}+\boldsymbol{v}\|^2-\|\boldsymbol{u}-\boldsymbol{v}\|^2+i\|\boldsymbol{u}+i\boldsymbol{v}\|^2-i\|\boldsymbol{u}-i\boldsymbol{v}\|^2\right)$
6. **Gram's determinant**: $\lambda^n\left(\left\{\sum_{i=1}^{n}t_i\boldsymbol{x}_i:0\le t_i\le1\right\}\right)=\det(\boldsymbol{x}_1;...;\boldsymbol{x}_n)=\sqrt{\det\left(\langle\boldsymbol{x}_i,\boldsymbol{x}_j\rangle_{1\le i,j\le n}\right)}$
7. **Hadamard's inequality**: $\det(\boldsymbol{x}_1;...;\boldsymbol{x}_n)\le\prod_{i=1}^{n}\|\boldsymbol{x}_i\|$ with equality iff the $\boldsymbol{x}_i$ are **orthogonal** with $\langle\boldsymbol{x}_i,\boldsymbol{x}_j\rangle=0$ for $1\le i\ne j\le n$.

Proof:

1. For all $\boldsymbol{u};\boldsymbol{v}\in X$ holds
$$0\leq\langle\langle\boldsymbol{v},\boldsymbol{v}\rangle\,\boldsymbol{u}-\langle\boldsymbol{u},\boldsymbol{v}\rangle\,\boldsymbol{v},\langle\boldsymbol{v},\boldsymbol{v}\rangle\,\boldsymbol{u}-\langle\boldsymbol{u},\boldsymbol{v}\rangle\,\boldsymbol{v}\rangle$$
$$=\langle\boldsymbol{v},\boldsymbol{v}\rangle^2\langle\boldsymbol{u},\boldsymbol{u}\rangle-\langle\boldsymbol{u},\boldsymbol{v}\rangle\langle\boldsymbol{v},\boldsymbol{v}\rangle\langle\boldsymbol{v},\boldsymbol{u}\rangle-\langle\boldsymbol{v},\boldsymbol{v}\rangle\langle\boldsymbol{u},\boldsymbol{v}\rangle\langle\boldsymbol{u},\boldsymbol{v}\rangle+\langle\boldsymbol{u},\boldsymbol{v}\rangle^2\langle\boldsymbol{v},\boldsymbol{v}\rangle$$
$$=\langle\boldsymbol{v},\boldsymbol{v}\rangle\left(\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|-\langle\boldsymbol{u},\boldsymbol{v}\rangle\overline{\langle\boldsymbol{u},\boldsymbol{v}\rangle}\right)$$
$$=\langle\boldsymbol{v},\boldsymbol{v}\rangle\left(\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|-|\langle\boldsymbol{u},\boldsymbol{v}\rangle|\right).$$

2. we have
$$\|\boldsymbol{u}+\mathbf{v}\|^2=\langle\boldsymbol{u}+\boldsymbol{v};\boldsymbol{u}+\boldsymbol{v}\rangle$$
$$=\langle\boldsymbol{u},\boldsymbol{u}\rangle+\langle\boldsymbol{u},\boldsymbol{v}\rangle+\overline{\langle\boldsymbol{u},\boldsymbol{v}\rangle}+\langle\boldsymbol{v},\boldsymbol{v}\rangle$$
$$=\|\boldsymbol{u}\|^2+2\mathrm{Re}\,\langle\boldsymbol{u},\boldsymbol{v}\rangle+\|\boldsymbol{v}\|^2$$
$$\leq\|\boldsymbol{u}\|^2+2\,|\langle\boldsymbol{u},\boldsymbol{v}\rangle|+\|\boldsymbol{v}\|^2$$
$$\leq\|\boldsymbol{u}\|^2+2\,\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|+\|\boldsymbol{v}\|^2$$
$$\leq(\|\boldsymbol{u}\|+\|\boldsymbol{v}\|)^2.$$

3. Follows from 2. by
$$\|\boldsymbol{u}\|-\|\boldsymbol{v}\|=\|\boldsymbol{u}-\boldsymbol{v}+\boldsymbol{v}\|-\|\boldsymbol{v}\|$$
$$\leq\|\boldsymbol{u}-\boldsymbol{v}\|+\boldsymbol{v}-\|\boldsymbol{v}\|$$
$$=\|\boldsymbol{u}-\boldsymbol{v}\|$$ and vice versa.

4. obvious

5. obvious

6. according to [4, p. 8.9.3] with the matrix $(\boldsymbol{x}_1;...;\boldsymbol{x}_n)=x_i^k$ formed by the **coordinate vectors** in $\boldsymbol{x}_i=x_i^k\boldsymbol{e}_k$ we have
$$\det\left(\langle\boldsymbol{x}_i,\boldsymbol{x}_j\rangle_{1\leq i,j\leq n}\right)=\det\left(\left(x_i^kx_j^k\right)_{1\leq i,j\leq n}\right)$$
$$=\det\left({}^T(\boldsymbol{x}_1;...;\boldsymbol{x}_n)\cdot(\boldsymbol{x}_1;...;\boldsymbol{x}_n)\right)$$
$$=\det{}^T(\boldsymbol{x}_1;...;\boldsymbol{x}_n)\cdot\det(\boldsymbol{x}_1;...;\boldsymbol{x}_n)$$
$$=(\det(\boldsymbol{x}_1;...;\boldsymbol{x}_n))^2$$
$$=\left(\lambda^n\left(\left\{\sum_{i=1}^n t_i\boldsymbol{x}_i:0\leq t_i\leq 1;1\leq i\leq n\right\}\right)\right)^2$$

7. On account of the previous result it suffices to consider linearly independent $(\boldsymbol{x}_i)_{1\leq i\leq n}$ such that the QR-decomposition 6.4 applies with $A=(\boldsymbol{x}_1;...;\boldsymbol{x}_n)$, $Q=(\boldsymbol{q}_1;...;\boldsymbol{q}_n)\in\bar{O}(n)$ and the upper triangular matrix $R=\left\langle\boldsymbol{q}_j;\boldsymbol{x}_i\right\rangle_{1\leq j\leq i\leq n}\in GL(n;\mathbb{C})$ with $r_i^j=0\Leftrightarrow j>i$ whence $\det A=\det(Q*R)=\det Q\cdot\det R$, i.e. $\det(\boldsymbol{x}_1;...;\boldsymbol{x}_n)=1\cdot\prod\limits_{i=1}^n\langle\boldsymbol{q}_i;\boldsymbol{x}_i\rangle=\prod\limits_{i=1}^n\|\boldsymbol{b}_i\|$. The equality $A=Q*R$ also yields $\boldsymbol{a}_i=\sum\limits_{j=1}^i r_i^j\boldsymbol{q}_j$, i.e. $\boldsymbol{x}_i=\sum\limits_{j=1}^i\left\langle\boldsymbol{q}_j;\boldsymbol{x}_i\right\rangle\boldsymbol{q}_j$ whence from $\left\langle\boldsymbol{q}_j;\boldsymbol{q}_i\right\rangle=\delta_{ji}$ follows $\|\boldsymbol{x}_i\|^2=\left\|\sum\limits_{j=1}^i\left\langle\boldsymbol{q}_j;\boldsymbol{x}_i\right\rangle\boldsymbol{q}_j\right\|^2=\sum\limits_{j=1}^i\left\langle\boldsymbol{q}_j;\boldsymbol{x}_i\right\rangle^2\left\|\boldsymbol{q}_j\right\|^2=\sum\limits_{j=1}^{i-1}\left\langle\boldsymbol{q}_j;\boldsymbol{x}_i\right\rangle^2\cdot 1+\|\boldsymbol{b}_i\|^2\cdot 1$. Thus we conclude that $\|\boldsymbol{b}_i\|\leq\|\boldsymbol{x}_i\|$ and the assertion is proved.

6.6 Unitary and orthogonal endomorphisms

1. An **endomorphism** $\boldsymbol{f} \in \operatorname{End}(X)$ on a **unitary vector space** $(X; \langle\,\rangle)$ is **unitary** iff $\langle \boldsymbol{f}(\boldsymbol{u}), \boldsymbol{f}(\mathbf{v})\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle$ for all $\boldsymbol{u}, \boldsymbol{v} \in X$. Correspondingly an invertible matrix $A \in GL(n; \mathbb{C})$ is **unitary** iff $A^{-1} = {}^T\overline{A}$ and these matrices form the **normal subgroup** $U(n) \subset GL(n; \mathbb{C})$. An invertible matrix $A = a_j^i \in GL(n; \mathbb{C})$ is unitary iff $A * {}^T\overline{A} = E_n$, i.e. iff its **column vectors** $(\boldsymbol{a}_j)_{1 \le j \le n}$ resp. its **row vectors** $(\boldsymbol{a}^i)_{1 \le i \le n}$ are an **orthonormal basis** of $(\mathbb{C}^n; \langle\,\rangle)$. In the case of the **canonical scalar product** $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = x_i \overline{y}_i = {}^T\boldsymbol{x} * \overline{\boldsymbol{y}}$ for coordinate vectors $\boldsymbol{x} = x^i \boldsymbol{e}_i$ and $\boldsymbol{y} = y^i \boldsymbol{e}_i$ on the **canonical basis** $\mathcal{B} = (\boldsymbol{e}_i)_{1 \le i \le n}$ the two definitions **coincide**, i.e. $\boldsymbol{f} \in \operatorname{End}(X)$ is **unitary** iff $F = M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f})$ is **unitary** since

$$\langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{f}(\boldsymbol{y})\rangle = {}^T(F * \boldsymbol{x}) * \overline{F\boldsymbol{y}} = {}^T\boldsymbol{x} * {}^TF * \overline{F} * \boldsymbol{y} = {}^T\boldsymbol{x} * \overline{\boldsymbol{y}} = \langle \boldsymbol{x}, \boldsymbol{y}\rangle \Leftrightarrow {}^TF * \overline{F} = E_n.$$

2. For unitary matrices $A \in \mathrm{U}(n)$ we have $1 = \det A \cdot \det \overline{A} = \det A \cdot \overline{\det A} = |\det A|^2$ whence $|\det A| = 1$. An invertible matrix $A \in GL(n; \mathbb{R})$ is **orthogonal** iff $A^{-1} = {}^TA$ and these matrices form the **normal subgroup** $O(n) \subset GL(n; \mathbb{C})$ with $\det A \in \{\pm 1\}$ for $A \in \mathrm{O}(n)$. The normal subgroup $SL(n) = \{A \in \mathrm{O}(n) : \det A = 1\} \subset O(n)$ (*special linear group*) according to 4.5 preserves the **orientation** of the basis.
3. Every **eigenvalue** $\lambda \in \mathbb{C}$ of a **unitary** endomorphism $\mathbf{f} : X \to X$ has an **absolute value** of $|\lambda| = 1$ and in particular $\overline{\lambda} = \frac{1}{\lambda}$ since for the corresponding eigenvector $\boldsymbol{v} \in X$ we have $\langle \boldsymbol{v}; \boldsymbol{v}\rangle = \langle \boldsymbol{f}(\boldsymbol{v}); \boldsymbol{f}(\boldsymbol{v})\rangle = \langle \lambda\boldsymbol{v}; \lambda\boldsymbol{v}\rangle = \lambda\overline{\lambda}\langle \boldsymbol{v}; \boldsymbol{v}\rangle$ whence form $\langle \boldsymbol{v}; \boldsymbol{v}\rangle \ne 0$ follows $|\lambda|^2 = \lambda\overline{\lambda} = 1$.
4. **Eigenvectors** $\boldsymbol{v}$ resp. $\boldsymbol{w}$ for different eigenvalues λ resp. μ of a **unitary** endomorphism $\boldsymbol{f} : X \to X$ are **orthogonal** to each other since $\langle \boldsymbol{v}; \boldsymbol{w}\rangle = \langle \boldsymbol{f}(\boldsymbol{v}); \boldsymbol{f}(\boldsymbol{w})\rangle = \langle \lambda\boldsymbol{v}; \mu\boldsymbol{w}\rangle = \lambda\overline{\mu}\langle \boldsymbol{v}; \boldsymbol{w}\rangle$ whence form $\langle \boldsymbol{v}; \boldsymbol{w}\rangle \ne 0$ follows $\lambda \cdot \frac{1}{\mu} = \lambda\overline{\mu} = 1$, i.e. $\lambda = \mu$.

6.7 Decomposition of orthogonal endomorphisms

For every **orthogonal endomorphism** $\boldsymbol{f} \in SL(n)$ on an n-dimensional **euclidean** vector space X there is an orthonormal basis $\mathcal{B} = (\boldsymbol{u}_i)_{1 \le i \le k} \cup (\boldsymbol{v}_i; \boldsymbol{v}_{i+1})_{1 \le i \le l}$ with $k + 2l = n$ of $\mathbb{R}^n$ such that $\boldsymbol{f}[\operatorname{span}\{\boldsymbol{u}_i\}] = \operatorname{span}\{\boldsymbol{u}_i\}$ for $i \le k$, $\boldsymbol{f}[\operatorname{span}\{\boldsymbol{v}_j; \boldsymbol{v}_{j+1}\}] = \operatorname{span}\{\boldsymbol{v}_j; \boldsymbol{v}_{j+1}\}$ for $1 \le j \le l$ and

$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \lambda_k & & & & & & \\ & & & \cos\alpha_1 & -\sin\alpha_1 & & & & \\ & & & \sin\alpha_1 & \cos\alpha_1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & \cos\alpha_l & -\sin\alpha_l \\ 0 & & & & & & \sin\alpha_l & \cos\alpha_l \end{pmatrix}$$

with $\lambda_i \in \{\pm 1\}$ for $1 \le i \le k$ resp. $\alpha_j \in \,]0; 2\pi[\, \setminus \{\pi\}$ for $1 \le j \le l$.

Proof: This almost diagonal separation is an improvement on the general decomposition of real endomorphisms presented in 5.4 and can be proved in a similar way by utilizing the **Cayley-Hamilton theorem**. In this case the invariance resp. separation of the two-dimensional subspaces extends to the **whole function $\boldsymbol{f}$** instead of only the **restriction $\boldsymbol{f}|_W$** to a subspace such that the representing matrix has zeros **both below and above the diagonal.**

As usual we proceed by induction over the dimension n and assume the hypothesis for $\dim W = n - 1$. Similarly to the first part of the proof of 5.4 we show the existence of a vector subspace $V \subset X$ with $1 \le \dim V \le 2$ and $f[V] = V$. The othogonality implies important additional properties of the subspace V.

First we extend the **euclidean** vector space $(X;\mathbb{R};+;\cdot;\langle\rangle)$ to a **unitary** space $\left(\overline{X};\mathbb{C};+;\cdot;\langle\rangle\right)$ by admitting complex scalars and generalizing the **inner product** $\langle\boldsymbol{u},\boldsymbol{v}\rangle = {}^T\boldsymbol{x}_{\mathcal{A}}*\overline{\boldsymbol{y}_{\mathcal{A}}}$ for complex coordinate vectors $\boldsymbol{x}_{\mathcal{A}} = x^i_{\mathcal{A}}\boldsymbol{e}_i \in \mathbb{C}^n$ resp. $\boldsymbol{y}_{\mathcal{A}} = y^i_{\mathcal{A}}\boldsymbol{e}_i \in \mathbb{C}^n$ of $\boldsymbol{u} = x^i_{\mathcal{A}}\boldsymbol{a}_i$ and $\boldsymbol{v} = y^i_{\mathcal{A}}\boldsymbol{a}_i$ referring to the original basis $\mathcal{A} = (\boldsymbol{a}_i)_{1\le i\le n} \subset X$. The **characteristic polynomial** $P_{\boldsymbol{f}}(t) = \prod_{i=1}^{k}(\lambda_i - t)\prod_{j=1}^{l}(t^2+p_jt+q_j)$ with $k+2l=n$ in the case of $k\ge 1$ provides an **eigenvector** $\boldsymbol{u}_1$with $\boldsymbol{f}(\boldsymbol{u}_1) = \lambda_1\boldsymbol{u}_1$ and **eigenvalue** $\lambda_1 \in \{\pm 1\}$ such that we have $\boldsymbol{f}[V] = V$ for $V = \operatorname{span}\{\boldsymbol{u}_1\}$.

In the case of $k=0$ we have pair of **conjugated** complex **eigenvalues** λ resp. $\overline{\lambda} = -\frac{p}{2}\pm i\sqrt{\left(\frac{p}{2}\right)^2 - q} \in \mathbb{C}$ as zeros of the corresponding factor in with $p = p_1$ resp. $q = q_1$.

The corresponding **orthogonal** eigenvectors are also **conjugated** to each other since with $F = M^{\mathcal{A}}_{\mathcal{A}}(\boldsymbol{f}) \in M(n;\mathbb{R})$ and eigenvector $\boldsymbol{v} = x^i_{\mathcal{A}}\boldsymbol{a}_i$ the identity $\lambda\boldsymbol{v} = \boldsymbol{f}(\boldsymbol{v})$ resp. $\lambda\boldsymbol{x}_{\mathcal{A}} = \Phi^{-1}_{\mathcal{A}}(\lambda\boldsymbol{v}) = \Phi^{-1}_{\mathcal{A}}(\boldsymbol{f}(\boldsymbol{v})) = F*\boldsymbol{x}_{\mathcal{A}}$ implies $\overline{\lambda}\overline{\boldsymbol{x}_{\mathcal{A}}} = \overline{\lambda\boldsymbol{x}_{\mathcal{A}}} = \overline{F*\boldsymbol{x}_{\mathcal{A}}} = A*\overline{\boldsymbol{x}_{\mathcal{A}}} \Leftrightarrow \overline{\lambda}\overline{\boldsymbol{v}} = \Phi_{\mathcal{A}}\left(\overline{\lambda}\overline{\boldsymbol{x}_{\mathcal{A}}}\right) = \Phi_{\mathcal{A}}\left(F*\overline{\boldsymbol{x}_{\mathcal{A}}}\right) = \boldsymbol{f}(\overline{\boldsymbol{v}})$. Hence we have $\langle\boldsymbol{v};\boldsymbol{v}\rangle = \|\boldsymbol{v}\|^2 = 1 = \|\overline{\boldsymbol{v}}\|^2 = \langle\overline{\boldsymbol{v}};\overline{\boldsymbol{v}}\rangle$ and $\langle\boldsymbol{v};\overline{\boldsymbol{v}}\rangle = \langle\overline{\boldsymbol{v}};\boldsymbol{v}\rangle = 0$. Note that due to the definition of the canonical inner product on $\mathbb{C}^n$ these equations imply ${}^T\mathbf{x}_{\mathcal{A}}*\overline{\mathbf{x}_{\mathcal{A}}} = 1$ but ${}^T\mathbf{x}_{\mathcal{A}}*\mathbf{x}_{\mathcal{A}} = 0$.

By **polarisation** we obtain **real valued orthonormal** vectors $\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}(\boldsymbol{v}+\overline{\boldsymbol{v}})$ and $\boldsymbol{v}_2 = \frac{1}{i\sqrt{2}}(\boldsymbol{v}-\overline{\boldsymbol{v}})$ resp. $\boldsymbol{v} = \sqrt{2}(\boldsymbol{v}_1 + i\boldsymbol{v}_2)$ and $\overline{\boldsymbol{v}} = \sqrt{2}(\boldsymbol{v}_1 - i\boldsymbol{v}_2)$ such that

$\boldsymbol{f}(\boldsymbol{v}_1) = \frac{1}{\sqrt{2}}\left(\lambda\boldsymbol{v}+\overline{\lambda}\overline{\boldsymbol{v}}\right) = \frac{1}{\sqrt{2}}\left(\lambda\boldsymbol{v}+\overline{\lambda\boldsymbol{v}}\right) = \sqrt{2}\mathrm{Re}(\lambda\boldsymbol{v}) = 2\mathrm{Re}(\lambda(\boldsymbol{v}_1+i\boldsymbol{v}_2)) = 2(\mathrm{Re}\lambda)\boldsymbol{v}_1 - 2(\mathrm{Im}\lambda)\boldsymbol{v}_2$
and

$\boldsymbol{f}(\boldsymbol{v}_2) = \frac{1}{i\sqrt{2}}\left(\lambda\boldsymbol{v}-\overline{\lambda}\overline{\boldsymbol{v}}\right) = \frac{1}{i\sqrt{2}}\left(\lambda\boldsymbol{v}-\overline{\lambda\boldsymbol{v}}\right) = \sqrt{2}\mathrm{Im}(\lambda\boldsymbol{v}) = 2\mathrm{Im}(\lambda(\boldsymbol{v}_1+i\boldsymbol{v}_2)) = 2(\mathrm{Im}\lambda)\boldsymbol{v}_1 + 2(\mathrm{Re}\lambda)\boldsymbol{v}_2$
, i.e.

$\boldsymbol{f}[V] \subset V$ for $V = \operatorname{span}\{\boldsymbol{v}_1;\boldsymbol{v}_2\}$.

According to 3.7 and since orthogonal maps are **injective** we infer $\mathbf{f}[V] = V$.

Since $\boldsymbol{f}^{-1}$is orthogonal as well for any $\boldsymbol{v}\in V$ and $\boldsymbol{w}\in V^{\perp} = (\boldsymbol{f}[V])^{\perp}$ follows $\langle\boldsymbol{f}(\boldsymbol{w}),\boldsymbol{v}\rangle = \langle\boldsymbol{w},\boldsymbol{f}(\boldsymbol{v})\rangle = 0$ whence $\boldsymbol{f}\left[V^{\perp}\right] = V^{\perp}$.

By the **Gram-Schmidt-orthonormalisation** 6.4 we find an orthonormal basis $\mathcal{B} = (\boldsymbol{v}_i)_{1\le i\le n}$ such that the representing matrix for $\boldsymbol{h} = \boldsymbol{f}|_V : V\to V$ has the form

$$M^{\mathcal{B}}_{\mathcal{B}}(\boldsymbol{h}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $1 = \|\boldsymbol{h}(\boldsymbol{v}_1)\| = \sqrt{a^2+c^2} = \sqrt{b^2+d^2} = \|\boldsymbol{h}(\boldsymbol{v}_2)\|$ and $0 = \langle\boldsymbol{v}_1;\boldsymbol{v}_2\rangle = \langle\boldsymbol{h}(\boldsymbol{v}_1);\boldsymbol{h}(\boldsymbol{v}_2)\rangle = ab+cd$ whence

$$M^{\mathcal{B}}_{\mathcal{B}}(\boldsymbol{h}) = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \end{pmatrix} \text{ or } \begin{pmatrix} \cos\alpha_1 & \sin\alpha_1 \\ -\sin\alpha_1 & \cos\alpha_1 \end{pmatrix} \text{ or } \begin{pmatrix} \sin\alpha_1 & \cos\alpha_1 \\ -\cos\alpha_1 & \sin\alpha_1 \end{pmatrix} \text{ or } \begin{pmatrix} \sin\alpha_1 & -\cos\alpha_1 \\ \cos\alpha_1 & \sin\alpha_1 \end{pmatrix}$$

for $0 < \alpha_1 < \frac{\pi}{2}$. By extending the range of the argument to $\alpha_1 \in]0;2\pi[\setminus\{\pi\}$ all four possibilities can be expressed by the first formula alone.

Thus $\boldsymbol{f}$ can be decomposed into $\boldsymbol{f}|_V : V\to V$ and $\boldsymbol{f}|_{V^\perp} : V^\perp \to V^\perp$ with $V\oplus V^\perp = X$ and an orthonormal basis $\mathcal{B}$ with

$$\text{either } M^{\mathcal{B}}_{\mathcal{B}}(\boldsymbol{h}) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{pmatrix} \text{ or } M^{\mathcal{B}}_{\mathcal{B}}(\boldsymbol{h}) = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 & 0 & \cdots & 0 \\ \sin\alpha_1 & \cos\alpha_1 & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{pmatrix}$$

such that we can apply the induction hypothesis to obtain the assertion.

6.8 Self-adjoint endomorphisms

An endomorphism $\boldsymbol{f}^{\mathrm{ad}} : X \to X$ is the **adjoint** to the endomorphism $\boldsymbol{f} : X \to X$ on a **unitary** vector space $(X; \langle\,\rangle)$ iff $\langle \boldsymbol{f}(\boldsymbol{v}), \boldsymbol{w}\rangle = \left\langle \boldsymbol{v}, \boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{w})\right\rangle$ for every $\boldsymbol{v}; \boldsymbol{w} \in X$. In the case of an **orthonormal basis** $\mathcal{A} = (\boldsymbol{a}_i)_{1\le i\le n}$ of a **finite dimensional** $X = \langle \boldsymbol{a}_i\rangle_{1\le i\le n}$ with $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{x}_{\mathcal{A}} * \boldsymbol{y}_{\mathcal{A}}$ for $\boldsymbol{u} = \sum\limits_{i=1}^{n} x_{\mathcal{A}i}\boldsymbol{a}_i$ resp. $\boldsymbol{v} = \sum\limits_{i=1}^{n} y_{\mathcal{A}i}\boldsymbol{a}_i$ we have $M_{\mathcal{A}}^{\mathcal{A}}\left(\boldsymbol{f}^{\mathrm{ad}}\right) = {}^{T}\overline{M_{\mathcal{A}}^{\mathcal{A}}(\boldsymbol{f})}$. The endomorphism $\boldsymbol{f}$ is **self-adjoint** iff $\boldsymbol{f}^{\mathrm{ad}} = \boldsymbol{f}$ resp. in the case of an orthonormal basis and finite dimension iff the representing matrix $F = M_{\mathcal{A}}^{\mathcal{A}}(\boldsymbol{f})$ is **hermitian** with $F = {}^{T}\overline{F}$. In the **real case** we have $F = {}^{T}F$ and the matrix is **symmetric**. The **vector spaces** (cf. 6.6!) of hermitian resp. symmetric matrices are denoted as $S(n)$ resp. $H(n)$.

Every **eigenvalue** λ with $\lambda\boldsymbol{v} = \boldsymbol{f}(\boldsymbol{v})$ of a **self-adjoint** endomorphism $\boldsymbol{f}$ is **real** since $\lambda\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{f}(\boldsymbol{v}), \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{f}(\boldsymbol{v})\rangle = \overline{\lambda}\langle \boldsymbol{v}, \boldsymbol{v}\rangle$.

Eigenvectors $\boldsymbol{v}; \boldsymbol{w}$ with different eigenvalues $\lambda; \mu \in \mathbb{R}$ are **orthogonal** since $\lambda\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{f}(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, \boldsymbol{f}(\boldsymbol{w})\rangle = \overline{\mu}\langle \boldsymbol{v}, \boldsymbol{w}\rangle$.

6.9 Trigonalization of self-adjoint endomorphisms

For every **self-adjoint endomorphism** $\boldsymbol{f} : X \to X$ on an n-dimensional **euclidean or unitary** vector space X there is an orthonormal basis $\mathcal{B} = (\boldsymbol{u}_i)_{1\le i\le n}$ of **eigenvectors** $\boldsymbol{u}_i$ with **real eigenvalues** $\lambda_i \in \mathbb{R}$ such that $\boldsymbol{f}[\mathrm{span}\{\boldsymbol{u}_i\}] = \mathrm{span}\{\boldsymbol{u}_i\}$ for $1 \le i \le n$ and

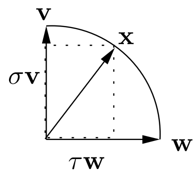
$$M_{\mathcal{B}}^{\mathcal{B}}(\boldsymbol{f}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Proof: Owing to the preceding paragraph it only remains to prove that $\dim \mathrm{Eig}(\boldsymbol{f}; \lambda) = \mu(P_{\boldsymbol{f}}; \lambda)$ for every eigenvalue λ resp. zero of the characteristic polynomial $P_{\boldsymbol{f}}$(cf. 5.1 and 5.7). As in 6.7 we proceed by induction over the dimension n. Assuming the hypothesis for $n-1$ we choose a real **eigenvalue** λ_1 with **eigenvector** $\boldsymbol{u}_1$ and according to the **Gram-Schmidt orthonormalisation** determine an orthonormal basis $\mathcal{B}' = \{\boldsymbol{u}_1; \boldsymbol{w}_2; ...; \boldsymbol{w}_n\}$ such that $X = V \oplus V^{\perp}$ with $V = \mathrm{span}\{\boldsymbol{u}_1\}$ and $V^{\perp} = \mathrm{span}\{\boldsymbol{w}_2; ...; \boldsymbol{w}_n\}$. We have $\boldsymbol{f}[V] = V$ but also $\boldsymbol{f}\left[V^{\perp}\right] = V^{\perp}$ since $\langle \boldsymbol{f}(\boldsymbol{w}), \boldsymbol{v}_1\rangle = \langle \boldsymbol{w}, \boldsymbol{f}(\boldsymbol{v}_1)\rangle$ $= \lambda\langle \boldsymbol{w}, \boldsymbol{v}_1\rangle\ 0$ for every $\boldsymbol{w} \in V^{\perp}$. The latter condition provides the existence of a further linearly independent eigenvector in the case of $\mu(P_{\boldsymbol{f}}; \lambda) \ge 2$. Hence both components are $\boldsymbol{f}$-invariant and by applying the induction hypothesis to $\boldsymbol{f}|_{V^{\perp}}$ we obtain the assertion.

6.10 Simultaneous determination of eigenvectors and eigenvalues

In the **real** case there is an effective optimizing procedure for the simultaneous determination of an eigenvalue λ and its eigenvector $\boldsymbol{v}$: For every **symmetric** matrix $A \in M(n; \mathbb{R})$ the **quadratic form** $q : \mathbb{R}^n \to \mathbb{R}$ with $q(\boldsymbol{x}) = {}^{T}\boldsymbol{x} * A * \boldsymbol{x}$ is **continuous** such that according to [6, p. 9.8] it attains its **supremum** $\lambda = \sup\{q(\boldsymbol{x}) : \boldsymbol{x} \in S\}$ on the **sphere** $S = \{\|\boldsymbol{x}\| = 1\}$, i.e. there is a $\boldsymbol{v} \in S$ with $\lambda = {}^{T}\boldsymbol{v} * A * \boldsymbol{v} \ge {}^{T}\boldsymbol{x} * A * \boldsymbol{x}$ for every $\boldsymbol{x} \in S$.

For every $\boldsymbol{w} \in S$ with $\langle \boldsymbol{w}, \boldsymbol{v}\rangle = 0$ we have $\mathbf{x} = \sigma\boldsymbol{v} + \tau\boldsymbol{w} \in S$ for $0 < \tau < 1$ and $\sigma = \sqrt{1-\tau^2}$. Hence with ${}^{T}\boldsymbol{w} * A * \boldsymbol{v} = {}^{T}\boldsymbol{v} * A * \boldsymbol{w}$ and $1 = \sigma^2 + \tau^2$ follows



$${}^{T}\boldsymbol{v} * A * \boldsymbol{v} \ge {}^{T}\boldsymbol{x} * A * \mathbf{x} = \sigma^{2T}\boldsymbol{v} * A * \boldsymbol{v} + 2\sigma\tau^{T}\boldsymbol{w} * A * \boldsymbol{v} + \tau^{2T}\boldsymbol{w} * A * \boldsymbol{w}$$

whence

$${}^{T}\boldsymbol{w} * A * \boldsymbol{v} \le \frac{1-\sigma^2}{2\sigma\tau}{}^{T}\boldsymbol{v} * A * \boldsymbol{v} - \frac{\tau}{2\sigma}{}^{T}\boldsymbol{w} * A * \boldsymbol{w} = \frac{\tau}{2\sigma}\left({}^{T}\boldsymbol{v} * A * \boldsymbol{v} - {}^{T}\boldsymbol{w} * A * \boldsymbol{w}\right).$$

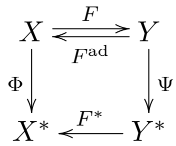
By exchanging $\boldsymbol{w}$ with $-\boldsymbol{w}$ we can assume ${}^T\boldsymbol{w} * A * \boldsymbol{v} \geq 0$ such that with ${}^T\boldsymbol{v} * A * \boldsymbol{v} - {}^T\boldsymbol{w} * A * \boldsymbol{w} \geq 0$ and τ arbitrary follows ${}^T\boldsymbol{w} * A * \boldsymbol{v} = 0$. Since this is true for every $\boldsymbol{w} \in S$ with $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = 0$ we have shown that $A * \boldsymbol{v} = \mu \boldsymbol{v}$ for some $\mu \in \mathbb{R}$ whence $\lambda = {}^T\boldsymbol{v} * A * \boldsymbol{v} = \mu\lambda = {}^T\boldsymbol{v} * \boldsymbol{v} = \mu$.

6.11 Adjoint maps

In a **unitary** vector space $(X, \langle\rangle)$ with an **inner product** the isomorphism $\Phi : X \to X^*$ with $\Phi(\boldsymbol{a}_i) = \boldsymbol{a}_i^*$ for a given **basis** $\mathcal{A} = (\boldsymbol{a}_i)_{i \in I}$ from 3.12 can be replaced by the canonical **semi-isomorphism** $\Phi(\boldsymbol{x}) = \langle\, , \boldsymbol{x}\rangle$ with $\Phi(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \overline{\alpha}\Phi(\boldsymbol{x}) + \overline{\beta}\Phi(\boldsymbol{y})$ being **independent of the basis**. Note that Φ is **injective** since $\langle\rangle$ is **positive definite**.

For every vector subspace $E \subset X$ and the **annihilator** $E^0 = \{\boldsymbol{x}^* \in X^* : \boldsymbol{x}^*\boldsymbol{x} = 0\,\forall \boldsymbol{x} \in E\}$ defined in 3.15 we obviously have $\Phi\left[E^\perp\right] = E^0$ and for every **orthonormal basis** $\mathcal{A} = (\boldsymbol{a}_i)_{i \in I}$ resp. the **dual othonormal basis** $\mathcal{A}^* = (\boldsymbol{a}_i^*)_{i \in I}$ determined by $\boldsymbol{a}_i^*\boldsymbol{a}_j = \delta_j^i$ holds $\Phi(\boldsymbol{a}_i) = \boldsymbol{a}_i^*$.

For every $\boldsymbol{f} \in L(X; Y)$ between **unitary finite dimensional** vector spaces $X = \langle \boldsymbol{a}_i \rangle_{1 \leq i \leq m}$ resp. $Y = \langle \boldsymbol{b}_j \rangle_{1 \leq j \leq n}$ generated by bases $\mathcal{A} = (\boldsymbol{a}_i)_{1 \leq i \leq m}$ resp. $\mathcal{B} = (\boldsymbol{b}_j)_{1 \leq j \leq n}$ the **adjoint map** $\boldsymbol{f}^{\text{ad}} : Y \to X$ is defined by $\langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{y} \rangle = \left\langle \boldsymbol{x}, \boldsymbol{f}^{\text{ad}}(\boldsymbol{y}) \right\rangle$. According to the definitions of the canonical semi-isomorphisms $\Phi : X \to X^*$ with $\Phi(\boldsymbol{x}) = \langle\, , \boldsymbol{x}\rangle$ resp. $\Psi : Y \to Y^*$ with $\Psi(\boldsymbol{y}) = \langle\, , \boldsymbol{y}\rangle$ and the **dual linear map** $\boldsymbol{f}^* : Y^* \to X^*$ with $\boldsymbol{f}(\boldsymbol{y}^*) = \boldsymbol{y}^* \circ \boldsymbol{f}$ we have



$$(\boldsymbol{f}^* \circ \Psi)(\boldsymbol{y}) = \langle \boldsymbol{f}(\,), \boldsymbol{y} \rangle = \left\langle\, , \boldsymbol{f}^{\text{ad}}(\boldsymbol{y}) \right\rangle = \left(\Phi \circ \boldsymbol{f}^{\text{ad}}\right)(\boldsymbol{y})$$

whence $\boldsymbol{f}^{\text{ad}} = \Phi^{-1} \circ \boldsymbol{f}^* \circ \Psi$. In particular for $\boldsymbol{x} = \sum_{i=1}^{m} x_{\mathcal{A}i}\boldsymbol{a}_i$, $\boldsymbol{y} = \sum_{j=1}^{n} y_{\mathcal{B}j}\boldsymbol{b}_j$ and $M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f}) = F \in M(n \times m; \mathbb{C})$ we have

$$\begin{aligned}
\left(\Phi \circ \boldsymbol{f}^{\text{ad}}\right)(\boldsymbol{y})(\boldsymbol{x}) &= (\boldsymbol{f}^* \circ \Psi)(\boldsymbol{y})(\boldsymbol{x}) \\
&= \langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{y} \rangle \\
&= {}^T(F * \boldsymbol{x}_{\mathcal{A}}) * \overline{\boldsymbol{y}_{\mathcal{B}}} \\
&= {}^T\boldsymbol{x}_{\mathcal{A}} * {}^TF * \overline{\boldsymbol{y}_{\mathcal{B}}} \\
&= {}^T\boldsymbol{x}_{\mathcal{A}} * \overline{{}^T\overline{F} * \boldsymbol{y}_{\mathcal{B}}}
\end{aligned}$$

whence $\boldsymbol{f}^{\text{ad}}(\boldsymbol{y}) = \sum_{i=1}^{m} z_{\mathcal{A}i}\boldsymbol{a}_i$ with the coordinate vector $\boldsymbol{z}_A = {}^T\overline{F} * \boldsymbol{y}_{\mathcal{B}}$ of $\boldsymbol{f}^{\text{ad}}(\boldsymbol{y})$ and the representing matrix

$$M_{\mathcal{A}}^{\mathcal{B}}\left(\boldsymbol{f}^{\text{ad}}\right) = {}^T\overline{F} = {}^T\overline{M_{\mathcal{B}}^{\mathcal{A}}(\boldsymbol{f})}$$

of $\boldsymbol{f}^{\text{ad}}$. According to 3.15 we also have

$$\ker \boldsymbol{f}^{\text{ad}} = (\text{im}\boldsymbol{f})^\perp \text{ and } \text{im}\boldsymbol{f}^{\text{ad}} = (\ker \boldsymbol{f})^\perp$$

6.12 Normal endomorphisms

An endomorphism $\boldsymbol{f}$ on a **unitary** vector space $(X;\langle\rangle)$ is **normal** iff $\boldsymbol{f}\circ\boldsymbol{f}^{\mathrm{ad}}=\boldsymbol{f}^{\mathrm{ad}}\circ\boldsymbol{f}$. Correspondingly a matrix $A\in M(n;\mathbb{C})$ is **normal** iff $A*{}^{T}\overline{A}={}^{T}\overline{A}*A$.

Since $\boldsymbol{x}\in\ker\boldsymbol{f}\Leftrightarrow 0=\langle\boldsymbol{f}(\boldsymbol{x}),\boldsymbol{f}(\boldsymbol{x})\rangle=\left\langle\boldsymbol{x},\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{f}(\boldsymbol{x}))\right\rangle=\left\langle\boldsymbol{x},\boldsymbol{f}\left(\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x})\right)\right\rangle=\overline{\left\langle\boldsymbol{f}\left(\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x})\right),\boldsymbol{x}\right\rangle}=\left\langle\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x}),\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x})\right\rangle=\overline{0}\Leftrightarrow\boldsymbol{x}\in\ker\boldsymbol{f}^{\mathrm{ad}}$ we have

$$\ker\boldsymbol{f}^{\mathrm{ad}}=\ker\boldsymbol{f}\text{ and }\mathrm{im}\boldsymbol{f}^{\mathrm{ad}}=\mathrm{im}\boldsymbol{f}$$

according to the preceding paragraph 6.11.

Also for any **eigenvalue** λ and $\boldsymbol{g}=\boldsymbol{f}-\lambda\mathrm{id}$ we have $\boldsymbol{g}^{\mathrm{ad}}=\boldsymbol{f}^{\mathrm{ad}}-\overline{\lambda}\mathrm{id}$ and for every $\boldsymbol{x}\in X$ holds $\boldsymbol{g}^{\mathrm{ad}}(\boldsymbol{g}(\boldsymbol{x}))=\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{f}(\boldsymbol{x}))-\overline{\lambda}\boldsymbol{f}(\boldsymbol{x})+\boldsymbol{f}^{\mathrm{ad}}(-\lambda\boldsymbol{x})+\overline{\lambda}\lambda\boldsymbol{x}=\boldsymbol{f}\left(\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x})\right)-\lambda\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{x})+\boldsymbol{f}\left(-\overline{\lambda}\boldsymbol{x}\right)+\lambda\overline{\lambda}\boldsymbol{x}=\boldsymbol{g}\left(\boldsymbol{g}^{\mathrm{ad}}(\boldsymbol{x})\right)$, i.e. $\boldsymbol{g}$ is **normal** such that from the preceding paragraph follows

$$\mathrm{Eig}(\boldsymbol{f},\lambda)=\ker\boldsymbol{g}=\ker\boldsymbol{g}^{\mathrm{ad}}=\mathrm{Eig}\left(\boldsymbol{f}^{\mathrm{ad}},\overline{\lambda}\right)$$

6.13 Diagonalization of normal endomorphisms

An endomorphism $\boldsymbol{f}$ on a **unitary** vector space $(X;\langle\rangle)$ is **normal** iff its **eigenvectors** form an **orthonormal basis** of X.

Proof:

$\Rightarrow$: For an orthonormal basis $\mathcal{B}=(\boldsymbol{v}_i)_{1\leq i\leq n}$ with $\boldsymbol{f}(\boldsymbol{v}_i)=\lambda_i\boldsymbol{v}_i$ for $1\leq i\leq n$ we have $\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{v}_i)=\overline{\lambda_i}\boldsymbol{v}_i$ and thus $\boldsymbol{f}\left(\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{v}_i)\right)=\boldsymbol{f}\left(\overline{\lambda_i}\boldsymbol{v}_i\right)=\lambda_i\overline{\lambda_i}\boldsymbol{v}_i=\overline{\lambda_i}\lambda_i\boldsymbol{v}_i=\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{f}(\boldsymbol{v}_i))$ for every basis vector $\boldsymbol{v}_i$ and hence for every $\mathbf{x}\in X$.

$\Leftarrow$: As usual we proceed by induction over the dimension n of X and assume the hypothesis for $n-1$. According to 5.2 we have the **characteristic polynomial** $P_f(t)=\pm\prod_{i=1}^{n}(t-\lambda_i)$ with **eigenvalues** $\lambda_i\in\mathbb{C}$ and an **eigenvector** $\boldsymbol{v}_1\in X$ with $\boldsymbol{f}(\boldsymbol{v}_1)=\lambda_1\boldsymbol{v}_1$. For $\boldsymbol{w}\in W=\langle\boldsymbol{v}_1\rangle^{\perp}$ holds $\langle\boldsymbol{f}(\boldsymbol{w}),\boldsymbol{v}_1\rangle=\left\langle\boldsymbol{w},\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{v}_1)\right\rangle=\left\langle\boldsymbol{w},\overline{\lambda_1}\boldsymbol{v}_1\right\rangle=\lambda_1\langle\boldsymbol{w},\boldsymbol{v}_1\rangle=0$ whence follows $\boldsymbol{f}[W]\subset W$. Also we have $\left\langle\boldsymbol{v}_1,\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{w})\right\rangle=\left\langle\boldsymbol{f}^{\mathrm{ad}}(\boldsymbol{v}_1),\boldsymbol{w}\right\rangle=\lambda_1\langle\boldsymbol{v}_1,\boldsymbol{w}\rangle=0$, i.e. $\boldsymbol{f}|_W$ is normal such that we can apply the induction hypothesis whence the assertion follows.

7 Multilinear algebra

7.1 Multilinear maps

1. A map $\varphi:\prod_{i\in I_p}X_i\to Y$ from a product $\prod_{i\in I_p}X_i=\{(\boldsymbol{x}_1;...;\boldsymbol{x}_p):\boldsymbol{x}_i\in X_i;i\in I_p\}$ of **complex** vector spaces X_i for $i\in I_p=\{1;...;p\}$ into a **real** vector space Y is p**-linear** iff every **projection** $\varphi_{\boldsymbol{x}_1;...;\boldsymbol{x}_{k-1};\boldsymbol{x}_{k+1};...;\boldsymbol{x}_p}:\boldsymbol{x}_k\to\varphi(\boldsymbol{x}_1;...;\boldsymbol{x}_p)$ is **linear** in $\boldsymbol{x}_k\in X_k$ for fixed $\boldsymbol{x}_i\in X_i$ and $k\in I_p$. For vector spaces X_i with **bases** $(\boldsymbol{e}_{i\mu})_{\mu\in J_i}$ and every function $\boldsymbol{y}:I^p\to Y$ there is a uniquely determined p-linear $\varphi:\prod_{i\in I_p}X_i\to Y$ with $\varphi\left(\boldsymbol{e}_{1\mu_1};...;\boldsymbol{e}_{p\mu_p}\right)=\boldsymbol{y}_{\mu_1;...;\mu_p}$ for every $(\mu_1;...;\mu_p)\in\prod_{1\leq i\leq p}J_i$. According to 3.2 every $\boldsymbol{x}_i\in X_i$ can be expressed as a **finite** sum $\boldsymbol{x}_i=x^{i\mu}\boldsymbol{e}_{i\mu}$ with **complex** coefficients $x^{i\mu}\neq 0$ for **finitely** many $\mu\in J_i$. Observing the **Einstein summation convention** 3.13 the desired p**-linearity** implies $\varphi(\boldsymbol{x}_1;...;\boldsymbol{x}_p)=\varphi\left(x^{1\mu_1}\boldsymbol{e}_{1\mu_1};...;x^{p\mu_p}\boldsymbol{e}_{p\mu_p}\right)=x^{1\mu_1}\cdot...\cdot x^{p\mu_p}\cdot\varphi\left(\boldsymbol{e}_{1\mu_1};...;\boldsymbol{e}_{p\mu_p}\right)=x^{1\mu_1}\cdot...\cdot x^{p\mu_p}\cdot\boldsymbol{y}_{1\mu_1;...;p\mu_p}$ and this is already a uniquely determined definition.

2. For every $k \in I_p$ by $\omega(\vartheta)(\boldsymbol{x}_1;...;\boldsymbol{x}_p) = \vartheta(\boldsymbol{x}_k)(\boldsymbol{x}_1;...;\boldsymbol{x}_{k-1};\boldsymbol{x}_{k+1};...;\boldsymbol{x}_p)$ defined for every **linear** $\vartheta : X_k \to L_p\left(\prod_{i\in I_p\setminus\{k\}} X_i;Y\right)$ with the inverse $\omega^{-1}(\eta)(\boldsymbol{x}_k)(\boldsymbol{x}_1;...;\boldsymbol{x}_{k-1};\boldsymbol{x}_{k+1};...;\boldsymbol{x}_p) = \eta(\boldsymbol{x}_1;...;\boldsymbol{x}_p)$ defined for every p-**linear** $\eta : \prod_{i\in I_p} X_i \to Y$ the spaces $L\left(X_k; L_{p-1}\left(\prod_{i\in I_p\setminus\{k\}} X_i;Y\right)\right)$ are **isomorphic** to $L_p\left(\prod_{i\in I_p} X_i;Y\right)$. For **finite dimensional** $X_i;Y$ due to 3.7 we have $\dim L(X_1;Y) = \dim X_1 \cdot \dim Y$ and by **induction** we conclude that

$$\dim L_p\left(\prod_{i=1}^{p} X_i;Y\right) = \dim X_k \cdot \dim L_{p-1}\left(\prod_{i\in I_p\setminus\{k\}} X_i;Y\right) = \prod_{i=1}^{p} \dim X_i \cdot \dim Y.$$

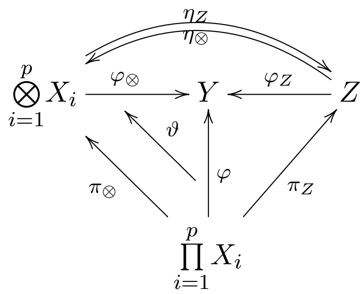
Example: The vector space $L_2(X^2;\mathbb{C})$ of **bilinear forms** $(\boldsymbol{x}_1;\boldsymbol{x}_2) \mapsto \omega(\boldsymbol{x}_1;\boldsymbol{x}_2)$ on $X = \mathbb{C}^n$ is **isomorphic** to

1. $M(n;\mathbb{C})$ by $\omega \mapsto a^{ij}\boldsymbol{e}_i\otimes\boldsymbol{e}_j$ with $a^{ij} = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$ and the basis $(\boldsymbol{e}_i\otimes\boldsymbol{e}_j)_{1\le i;j\le n}$ of $M(n;\mathbb{C})$ defined in 7.2
2. End $(M(n;\mathbb{C}))$ by $\omega \mapsto \left(m^{jk}\boldsymbol{e}_j\otimes\boldsymbol{e}_k \mapsto a_{ij}m^{jk}\boldsymbol{e}^i\otimes\boldsymbol{e}_k\right)$ with $a_{ij} = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$
3. End $(X) = L(X;X)$ by $\omega \mapsto \left(x^i\boldsymbol{e}_i \mapsto a_i^j x^i\boldsymbol{e}_j\right)$ with $a_i^j = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$
4. $L(X;X^*)$ by $\omega \mapsto (x^i\boldsymbol{e}_i \mapsto a_{ij}x^i\boldsymbol{e}^j)$ with $a_{ij} = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$
5. $L(X^*;X)$ by $\omega \mapsto (x_i\boldsymbol{e}^i \mapsto a^{ij}x_i\boldsymbol{e}_j)$ with $a^{ij} = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$
6. End $(X^*) = L(X^*;X^*)$ by $\omega \mapsto \left(x^i\boldsymbol{e}_i \mapsto a_j^i x_i\boldsymbol{e}_j\right)$ with $a_j^i = \omega(\boldsymbol{e}_i;\boldsymbol{e}_j)$

These cases are subsumed under the following generalization of the matrix:

7.2 Tensors

For every product $\prod_{i\in I_p} X_i = \{(\boldsymbol{x}_1;...;\boldsymbol{x}_p) : \boldsymbol{x}_i \in X_i; i\in I_p\}$ of **complex** vector spaces X_i for $i \in I_p = \{1;...;p\}$ exists a **complex** vector space $\bigotimes_{i\in I_p} X_i$ and a p-**linear** map $\pi_\otimes : \prod_{i\in I_p} X_i \to \bigotimes_{i\in I_p} X_i$ such that for every p-**linear** $\varphi : \prod_{i\in I_p} X_i \to Y$ into a **complex** vector space Y exists a linear $\varphi_\otimes : \bigotimes_{i\in I_p} X_i \to Y$ with $\varphi = \varphi_\otimes \circ \pi_\otimes$.



The vector space $\bigotimes_{i\in I_p} X_i$ is the **tensor product** of the **vector spaces** $(X_i)_{i\in I_p}$ and its elements are called **tensors**. The images $\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p = \pi_\otimes(\boldsymbol{x}_1;...;\boldsymbol{x}_p)$ are the **tensor products** of the **vectors** $(\boldsymbol{x}_i)_{i\in I_p}$. The tensor product is **uniquely determined** in the sense that every complex vector space Z is **isomorphic** to $\bigotimes_{i\in I_p} X_i$ iff there is a p-linear $\pi_Z : \prod_{i\in I_p} X_i \to Z$ such that for every p-**linear** $\varphi : \prod_{i\in I_p} X_i \to Y$ exists a **unique linear** $\varphi_Z : Z \to Y$ with $\varphi = \varphi_Z \circ \pi_Z$. Also the map $\vartheta : L_p\left(\prod_{i\in I_p} X_i;Y\right) \to L\left(\bigotimes_{i\in I_p} X_i;Y\right)$ with $\vartheta(\varphi) = \varphi_\otimes$ is an **isomorphism**.

Notes:

1. The **tensor product** of the vectors $\boldsymbol{x}_1 = x^{1\mu_1}\boldsymbol{e}_{1j_1};...;\boldsymbol{x}_p = x^{p\mu_p}\boldsymbol{e}_{pj_p}$ is $\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p = x^{1\mu_1}\cdot...\cdot x^{p\mu_p}\cdot\boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p}$ while a general **tensor** has the form $x^{1\mu_1;...;p\mu_p}\cdot\boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p}$ with arbitrary complex coefficients $x^{1\mu_1;...;p\mu_p}$.
2. In the **finite dimensional case** with **identical** factors $X_i = X$ and $\dim X = n$ the Einstein summation produces $\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p = \sum_{1\le\mu_1;...;\mu_p\le n} x^{\mu_1}\cdot...\cdot x^{\mu_p}\cdot\boldsymbol{e}_{\mu_1}\otimes...\otimes\boldsymbol{e}_{\mu_p}$ whence $\dim \prod_{i=1}^{p} X_i = p\cdot n$ and $\dim \bigotimes_{i=1}^{p} X_i = n^p$. In particular the family $E_p = (\boldsymbol{e}_{\mu_1};...;\boldsymbol{e}_{\mu_p})_{1\le\mu_1;...;\mu_p\le n}$ is not a **basis** and not even **linearly independent** in X^p since e.g. $(\boldsymbol{e}_1;\boldsymbol{0}) \notin \langle E_2\rangle$ and $(\boldsymbol{e}_1;\boldsymbol{e}_1) - (\boldsymbol{e}_1;\boldsymbol{0}) - (\boldsymbol{0};\boldsymbol{e}_1) = (\boldsymbol{0};\boldsymbol{0})$.

Proof:

Existence and p-**linearity** of $\pi_\otimes$: According to 7.1 the complex vector space $\bigotimes_{i\in I_p} X_i = L_p\left(\prod_{i\in I_p} X_i;\mathbb{C}\right)$ $= \left\langle \boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p}\right\rangle_{\mu_i\in J_i;i\in I_p}$ with **basis tensors** $\boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p} = \pi_\otimes^{1\mu_1;...;p\mu_p}\left(\boldsymbol{e}_{1\nu_1};...;\boldsymbol{e}_{p\nu_p}\right) = \delta^{\mu_1}_{\nu_1}\cdot...\cdot\delta^{\mu_p}_{\nu_p}$ is well defined and so is the map $\pi_\otimes:\prod_{i\in I_p} X_i\to\bigotimes_{i\in I_p} X_i$ given by $\pi_\otimes\left(\boldsymbol{e}_{1\nu_1};...;\boldsymbol{e}_{p\nu_p}\right) = \boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p}$.

Existence and **linearity** of $\varphi_\otimes$: For any given p-**linear** $\varphi:\prod_{i\in I_p} X_i\to Y$ into a **complex** vector space Y the map $\varphi_\otimes$ defined by $\varphi_\otimes\left(\boldsymbol{e}_{1\mu_1}\otimes...\otimes\boldsymbol{e}_{p\mu_p}\right) = \varphi\left(\boldsymbol{e}_{1\nu_1};...;\boldsymbol{e}_{p\nu_p}\right)$ is **linear** and satisfies $\varphi = \varphi_\otimes\circ\pi_\otimes$.

Uniqueness of $\varphi_\otimes$: For any given linear $\psi_\otimes:\bigotimes_{i=1}^p X_i\to Y$ with $\varphi=\psi_\otimes\circ\pi_\otimes$ and $\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p\in\bigotimes_{i=1}^p X_i$ we have $\psi_\otimes\left(\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p\right) = \left(\psi_\otimes\circ\pi_\otimes\right)\left(\boldsymbol{x}_1;...;\boldsymbol{x}_p\right) = \left(\varphi_\otimes\circ\pi_\otimes\right)\left(\boldsymbol{x}_1;...;\boldsymbol{x}_p\right) = \varphi_\otimes\left(\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_p\right)$ whence $\psi_\otimes=\varphi_\otimes$.

Uniqueness of $\bigotimes_{i=1}^p X_i$: Assuming there is an $\boldsymbol{a}\in Z\setminus V$ with $V=\operatorname{span}\left\{\pi_Z\left[\prod_{i=1}^p X_i\right]\right\}$ there exists a $\omega_Z\in\operatorname{End}(Z)$ with $\omega_Z(\boldsymbol{a})=\boldsymbol{a}$ and $V\subset\ker\omega_Z$ resp. $\prod_{i=1}^p X_i\subset\ker\left(\omega_Z\circ\pi_Z\right)$. Since $\prod_{i=1}^p X_i=\operatorname{dom}\pi_Z$ this means $\mathbf{0}=\omega_Z\circ\pi_Z=\mathbf{0}\circ\pi_Z$ hence 1. implies $\omega_Z=\mathbf{0}$ in contradiction to the assumption whence we conclude that $V=Z$. A second application of 1. yields a **linear** $\eta_\otimes:\bigotimes_{i=1}^p X_i\to Z$ with $\pi_Z=\eta_\otimes\circ\pi_\otimes$ while the hypothesis in 3. provides a **linear** $\eta_Z:Z\to\bigotimes_{i=1}^p X_i$ with $\pi_\otimes=\eta_Z\circ\pi_Z$. On the one hand we have $\left(\eta_Z\circ\eta_\otimes\right)\left(\boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_n\right) = \left(\eta_Z\circ\eta_\otimes\circ\pi_\otimes\right)\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right) = \left(\eta_Z\circ\pi_Z\right)\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right) = \pi_\otimes\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right) = \boldsymbol{x}_1\otimes...\otimes\boldsymbol{x}_n$ whence $\eta_Z\circ\eta_\otimes=\operatorname{id}:\bigotimes_{i=1}^p X_i\to\bigotimes_{i=1}^p X_i$ and on the other hand $\left(\eta_\otimes\circ\eta_Z\right)\left(\pi_Z\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right)\right) = \left(\eta_\otimes\circ\pi_\otimes\right)\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right) = \pi_Z\left(\boldsymbol{x}_1;...;\boldsymbol{x}_n\right)$ such that $\eta_\otimes\circ\eta_Z=\operatorname{id}:Z\to Z=\operatorname{span}\left\{\pi_Z\left[\prod_{i=1}^p X_i\right]\right\}$. Hence $\eta_\otimes$ and η_Z are **isomorphisms**.

Existence and **linearity** of ϑ: Due to the preceding arguments the map $\vartheta:L_p\left(\prod_{i=1}^p X_i;Y\right)\to L\left(\bigotimes_{i=1}^p X_i;Y\right)$ is well defined by $\varphi=\vartheta(\varphi)\circ\pi_\otimes$. Hence for p-**linear** maps $\varphi,\psi:\prod_{i=1}^p X_i\to Y$ we have $\varphi+\psi=\vartheta(\varphi)\circ\pi_\otimes+\vartheta(\psi)\circ\pi_\otimes=\left(\vartheta(\varphi)+\vartheta(\psi)\right)\circ\pi_\otimes$ whence $\vartheta(\varphi+\psi)=\vartheta(\varphi)+\vartheta(\psi)$ and $c\varphi=c\cdot\left(\vartheta(\varphi)\circ\pi_\otimes\right)=\left(c\cdot\vartheta(\psi)\right)\circ\pi_\otimes$ so that $\vartheta(c\cdot\varphi)=c\cdot\vartheta(\varphi)$.

Injectivity of ϑ: $\vartheta(\varphi)=\vartheta(\psi)\Rightarrow\varphi=\vartheta(\varphi)\circ\pi_\otimes=\vartheta(\psi)\circ\pi_\otimes=\psi$.

Surjectivity of ϑ: For any $\varphi_\otimes\in L\left(\bigotimes_{i=1}^p X_i;Y\right)$ follows $\varphi=\varphi_\otimes\circ\pi_\otimes\in L_p\left(\prod_{i=1}^p X_i;Y\right)$ and hence $\vartheta(\varphi)=\varphi_\otimes$.

Examples: The familiar notation can be used for $p\le 3$ e.g. for $\boldsymbol{x}\in X=\mathbb{R}^4$, $\boldsymbol{y}\in Y=\mathbb{R}^3$ and $\boldsymbol{z}\in Z=\mathbb{R}^2$:

$$\begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ x^4 \end{pmatrix}$$

$\boldsymbol{x} \in X$

vector

$$\begin{pmatrix} m^{11} & m^{12} & m^{13} \\ m^{21} & m^{22} & m^{23} \\ m^{31} & m^{32} & m^{33} \\ m^{41} & m^{42} & m^{43} \end{pmatrix}$$

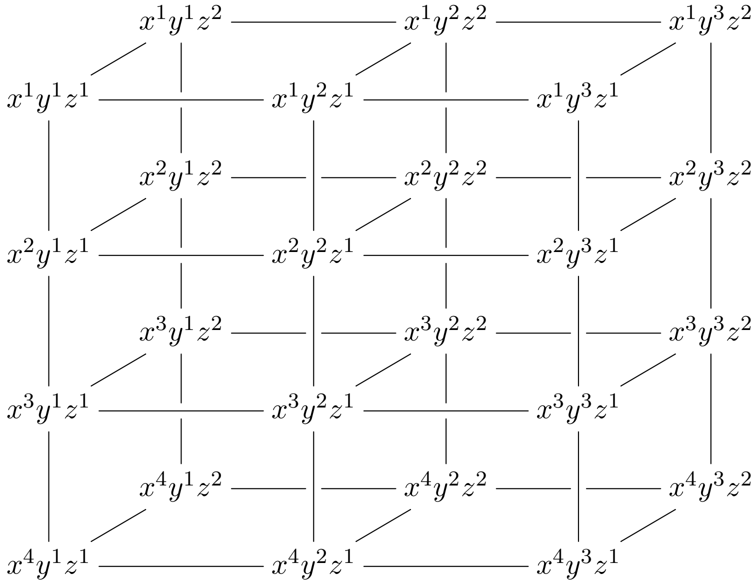
$m^{ij} \in X \otimes Y$

arbitrary tensor

$$\begin{pmatrix} x^1y^1 & x^1y^2 & x^1y^3 \\ x^2y^1 & x^2y^2 & x^2y^3 \\ x^3y^1 & x^3y^2 & x^3y^3 \\ x^4y^1 & x^4y^2 & x^4y^3 \end{pmatrix}$$

$\boldsymbol{x} \otimes \boldsymbol{y} \in X \otimes Y$

tensor product



$\boldsymbol{x} \otimes \boldsymbol{y} \otimes \boldsymbol{z} \in X \otimes Y \otimes Z$

tensor product

7.3 Tensors and multilinear forms

By $\vartheta : L_p\left(\prod_{i=1}^{p} X_i; \mathbb{R}\right) \to \bigotimes_{i=1}^{p} X_i$ with $\vartheta(\varphi) = \varphi\left(\boldsymbol{e}_{\mu_1}; ...; \boldsymbol{e}_{\mu_p}\right) \cdot \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_p}$ for $X_i = \operatorname{span}\{\boldsymbol{e}_1; ...; \boldsymbol{e}_n\}$ and $1 \leq \mu_i \leq n$ for $1 \leq i \leq p$ with **inverse** $\vartheta^{-1}\left(c^{\mu_1;...;\mu_p} \cdot \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_p}\right) = \varphi$ defined by $\varphi\left(\boldsymbol{e}_{\mu_1}; ...; \boldsymbol{e}_{\mu_p}\right)$ $= c^{\mu_1;...;\mu_p} \in \mathbb{C}$ the **p-multilinear forms** $L_p\left(\prod_{i=1}^{p} X_i; \mathbb{C}\right)$ are **isomorphic** to the **tensors** $\bigotimes_{i=1}^{p} X_i$. On account of the **linear character** of the X_i every p-linear $\varphi : \prod_{i=1}^{p} X_i \to \mathbb{C}$ is determined by $\varphi\left(\sum_{1 \leq \mu_1;...;\mu_p \leq n} c^{\mu_1;...;\mu_p} \cdot \left(\boldsymbol{e}_{\mu_1}; ...; \boldsymbol{e}_{\mu_p}\right)\right) = \sum_{1 \leq \mu_1;...;\mu_p \leq n} c^{\mu_1;...;\mu_p} \cdot \varphi\left(\boldsymbol{e}_{\mu_1}; ...; \boldsymbol{e}_{\mu_p}\right)$ whence follows again $\dim \bigotimes_{i=1}^{p} X_i = \dim \prod_{i=1}^{p} X_i = n^p$. (cf. 7.2)

7.3.1 Dyads and symmetric tensors

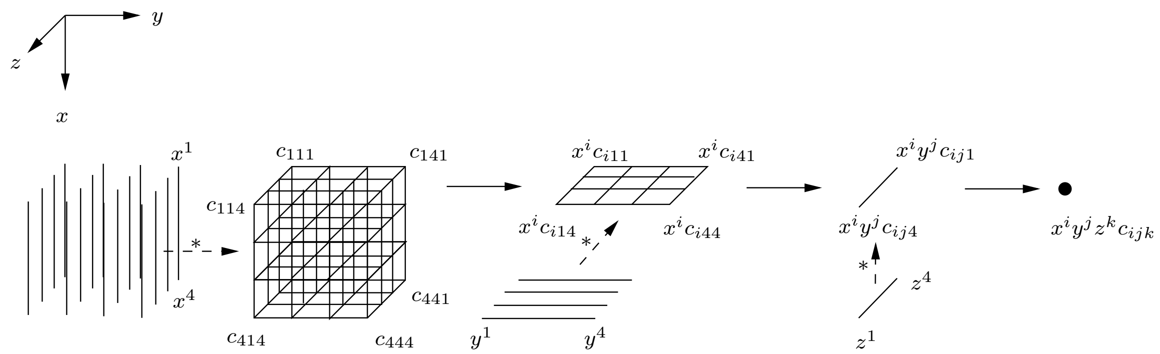
The $2n$**-dimensional** vector space $D_2 = \{\boldsymbol{x} \otimes \boldsymbol{y} : \boldsymbol{x}; \boldsymbol{y} \in X\} \simeq X^2$ of the **dyads** generates the n^2**-dimensional dyadic tensors** $\langle D_2 \rangle = \left\{\left(m_{ij}\mathbf{e}^i \otimes \mathbf{e}^j\right)_{1 \leq i;j \leq n} : m_{ij} \in \mathbb{R}\right\} = X_2$ since every $m_{ij}\mathbf{e}^i \otimes \mathbf{e}^j$ $\in X_2$ can be expressed as a linear combination

$$\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{pmatrix} = \begin{pmatrix} a_{n;1} \\ a_{n;2} \\ \vdots \\ a_{n;n} \end{pmatrix} \otimes \begin{pmatrix} b_{n;1} \\ b_{n;2} \\ \vdots \\ b_{n;n} \end{pmatrix} + \begin{pmatrix} 0 \\ a_{n-1;2} \\ \vdots \\ a_{n-1;n} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ b_{n-1;2} \\ \vdots \\ b_{n-1;n} \end{pmatrix} + ... + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{1;n} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{1;n} \end{pmatrix}$$

of dyads $\boldsymbol{a}_k \otimes \boldsymbol{b}_k \in D_2$. The n**-dimensional** subspace $SD_2 = \{\boldsymbol{x} \otimes \boldsymbol{x} : \boldsymbol{x} \in X\} \simeq X$ of the **symmetric dyads** forms a subspace of the $\frac{n(n+1)}{2}$**-dimensional** space of the **symmetric tensors** $S_2 = \{\boldsymbol{x} \otimes \boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{x} : \boldsymbol{x}; \boldsymbol{y} \in X\}$. Note that S_2 is closed under addtition since the distributive law of the tensor product resp. the bilinearity of $\pi_\otimes$ imply $(\boldsymbol{x} \otimes \boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{x}) + (\boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u}) = (\boldsymbol{x} + \boldsymbol{u}) \otimes (\boldsymbol{y} + \boldsymbol{v})$ $- (\boldsymbol{y} + \boldsymbol{v}) \otimes (\boldsymbol{x} + \boldsymbol{u}) - (\boldsymbol{u} \otimes \boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{u}) - (\boldsymbol{x} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{x})$.

7.3.2 Trilinear forms

For $\boldsymbol{x} = x^i \boldsymbol{e}_i, \boldsymbol{y} = y^j \boldsymbol{e}_j, \boldsymbol{z} = z^k \boldsymbol{e}_j \in X = \mathbb{R}^n$ every **cubic** tensor $C = (c_{ijk})_{1 \leq i;j;k \leq n}$ represents the **trilinear form** $\langle\,\rangle_C : X^3 \to \mathbb{C}$ with $\langle \boldsymbol{x}; \boldsymbol{y}; \boldsymbol{z} \rangle_C = x^i y^j z^k c_{ijk}$. The **graphical representation** of this computation shows the reduction of the cuboid tensor via matrix and vector to a number: Multiplying the first factor $(x^i)_{1 \leq i \leq n}$ with the n^2 corresponding vectors $(c_{ijk})_{1 \leq i \leq n}$ for $1 \leq j; k \leq n$ results in the matrix $(x^i c_{ijk})_{1 \leq j;k \leq n}$. Multiplying the second factor $(y^j)_{1 \leq j \leq n}$ with the n corresponding vectors $(x^i c_{ijk})_{1 \leq j \leq n}$ for $1 \leq k \leq n$ results in the vector $(x^i y^j c_{ijk})_{1 \leq k \leq n}$ which in turn combines with the third factor $\left(z^k\right)_{1 \leq k \leq n}$ to the final result $x^i y^j z^k c_{ijk}$:



7.4 Coordinate transformations

For a complex vector space $X = \operatorname{span} \mathcal{A}$ with **basis** $\mathcal{A} = (\boldsymbol{e}_i)_{1 \leq i \leq n}$ and its **dual space** $X^* = \operatorname{span} \mathcal{A}^*$ with the dual basis $\mathcal{A}^* = (\boldsymbol{e}^j)_{1 \leq j \leq n}$ defined by $\boldsymbol{e}^j \boldsymbol{e}_i = \delta_i^j$ the space $X_p^q = \bigotimes_{1 \leq i \leq p} X \bigotimes_{1 \leq j \leq q} X^*$ of p-**contravariant** and q-**covariant tensor** of **type** $(p; q)$ with $p; q \geq 1$ and $\dim X_p^q = n^{p+q}$ contains elements of the form

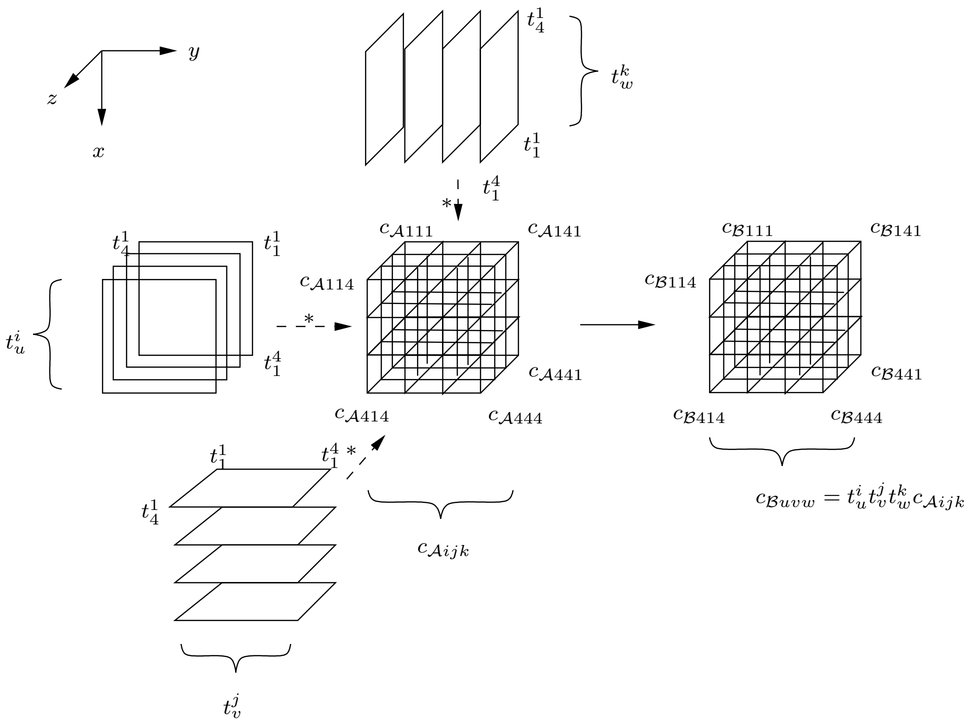
type	object
$(0;0)$	scalar
$(1;0)$	vector
$(0;1)$	linear form
$(1;1)$	endomorphism
$(2;0)$	quadratic matrix
$(0;2)$	bilinear form

$$\begin{aligned}\boldsymbol{x} &= \bigotimes_{1 \leq i \leq p} \boldsymbol{x}_i \bigotimes_{1 \leq j \leq q} \boldsymbol{x}^j \\ &= {}_{\mathcal{A}}x_1^{\mu_1} \cdot \ldots \cdot {}_{\mathcal{A}}x_p^{\mu_p} \cdot {}_{\mathcal{A}}x_{\nu_1}^1 \cdot \ldots \cdot {}_{\mathcal{A}}x_{\nu_q}^q \cdot \boldsymbol{e}_{\mu_1} \otimes \ldots \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes \ldots \otimes \boldsymbol{e}^{\nu_q} \\ &= {}_{\mathcal{A}}x_{\nu_1;\ldots;\nu_q}^{\mu_1;\ldots;\mu_p} \cdot \boldsymbol{e}_{\mu_1} \otimes \ldots \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes \ldots \otimes \boldsymbol{e}^{\nu_q}\end{aligned}$$

Also we define $X_0^0 = \mathbb{R}$; $X_1 = X$ and $X^1 = X^*$. According to 3.10 a tensor is transformed to a new basis $\mathcal{B}$ defined by the transformation matrix $T_{\mathcal{B}}^{\mathcal{A}} = t_i^j \in GL(n; \mathbb{R})$ by

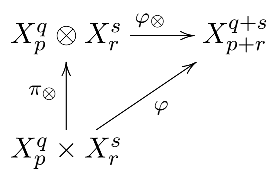
$$_{\mathcal{B}}x_{\beta_1;\ldots;\beta_q}^{\alpha_1;\ldots;\alpha_p} = \left(t^{-1}\right)_{\mu_1}^{\alpha_1} \cdot \ldots \cdot \left(t^{-1}\right)_{\mu_p}^{\alpha_p} \cdot {}_{\mathcal{A}}x_{\nu_1;\ldots;\nu_q}^{\mu_1;\ldots;\mu_p} \cdot t_{\beta_1}^{\nu_1} \cdot \ldots \cdot t_{\beta_q}^{\nu_q}.$$

Analogously to 7.3 the following drawing shows the **graphical representation** of these summations in the case of $p + q = 3$. Note the different transpositions of the transformation matrices according to the direction of the corresponding matrices resp. layers in the tensor:



7.5 The general tensor product

The definition of tensor product $\pi_\otimes X^p \to X_p$ from 7.2 can be extended to the general tensor product $\pi_\otimes : \left(X_p^q \times X_r^s\right) \to X_{p+r}^{q+s}$ by



$$\begin{aligned}\boldsymbol{x} \otimes \boldsymbol{y} &= \pi_\otimes(\boldsymbol{x};\boldsymbol{y}) \\ &= \left(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \otimes \boldsymbol{x}^1 \otimes ... \otimes \boldsymbol{x}^q\right) \otimes \left(\boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_r \otimes \boldsymbol{y}^1 \otimes ... \otimes \boldsymbol{y}^s\right) \\ &= \boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \otimes \boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_r \otimes \boldsymbol{x}^1 \otimes ... \otimes \boldsymbol{x}^q \otimes \boldsymbol{y}^1 \otimes ... \otimes \boldsymbol{y}^s\end{aligned}$$

with

1. **associativity**: $(\boldsymbol{x} \otimes \boldsymbol{y}) \otimes \boldsymbol{z} = \boldsymbol{x} \otimes (\boldsymbol{y} \otimes \boldsymbol{z})$
2. **associativity** with the **scalar multiplication**: $(c\boldsymbol{x}) \otimes \boldsymbol{y} = \boldsymbol{x} \otimes (c\boldsymbol{y}) = c \cdot (\boldsymbol{x} \otimes \boldsymbol{y})$
3. **associativity** with the **scalar product**: $(\boldsymbol{x} \otimes \boldsymbol{y}^*) * \boldsymbol{z} = x^i y_j z^j \boldsymbol{e}_i = \boldsymbol{x} \cdot (\boldsymbol{y}^* * \boldsymbol{z})$
4. **distributivity:** $\boldsymbol{x} \otimes (\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x} \otimes \boldsymbol{y} + \boldsymbol{x} \otimes \boldsymbol{z}$ resp. $(\boldsymbol{x} + \boldsymbol{y}) \otimes \boldsymbol{z} = \boldsymbol{x} \otimes \boldsymbol{z} + \boldsymbol{y} \otimes \boldsymbol{z}$

for $c \in \mathbb{C}$ and $\boldsymbol{x};\boldsymbol{y};\boldsymbol{z} \in X$ but in general **no commutativity**. Concerning the relationship with the **cross product** cf. 7.16.5. Hence every tensor may be written as the product of tensors of type $(1;0)$ resp. $(0;1)$ or even using the **basis** $\mathcal{B} = (\boldsymbol{e}_i)_{1 \le i \le n}$ of X in the form

$$\boldsymbol{x} = \boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \otimes \boldsymbol{x}^1 \otimes ... \otimes \boldsymbol{x}^q = x^{\mu_1;...;\mu_p}_{\nu_1;...;\nu_q} \cdot \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes ... \otimes \boldsymbol{e}^{\nu_q}$$

and its product with

$$\boldsymbol{y} = \boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_r \otimes \boldsymbol{y}^1 \otimes ... \otimes \boldsymbol{y}^s = y^{\lambda_1;...;\lambda_r}_{\chi_1;...;\chi_s} \boldsymbol{e}_{\lambda_1} \otimes ... \otimes \boldsymbol{e}_{\lambda_r} \otimes \boldsymbol{e}^{\chi_1} \otimes ... \otimes \boldsymbol{e}^{\chi_s}$$

can be written as

$$\boldsymbol{x} \otimes \boldsymbol{y} = x^{\mu_1;...;\mu_p}_{\nu_1;...;\nu_q} \cdot y^{\lambda_1;...;\lambda_r}_{\chi_1;...;\chi_s} \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}_{\lambda_1} \otimes ... \otimes \boldsymbol{e}_{\lambda_r} \otimes \boldsymbol{e}^{\nu_1} \otimes ... \otimes \boldsymbol{e}^{\nu_q} \otimes \boldsymbol{e}^{\chi_1} \otimes ... \otimes \boldsymbol{e}^{\chi_s}.$$

7.6 Contractions

The **contraction** $\gamma_i^k : X_p^q \to X_{p-1}^{q-1}$ is defined as

$$\gamma_i^k(\boldsymbol{x}) = x^{\mu_1;...;\mu_p}_{\nu_1;...;\nu_q} \cdot \left(\boldsymbol{e}^k \boldsymbol{e}_i\right) \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_{i-1}} \otimes \boldsymbol{e}_{\mu_{i+1}} \otimes ... \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes ... \otimes \boldsymbol{e}^{\nu_{k-1}} \otimes \boldsymbol{e}^{\nu_{k+1}} \otimes ... \otimes \boldsymbol{e}^{\nu_q}$$

$$= x^{\mu_1;...;\mu_{i-1};\lambda;\mu_{i+1};...;\mu_p}_{\nu_1;...;\nu_{k-1};\lambda;\nu_{k+1};...;\nu_q} \cdot \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_{i-1}} \otimes \boldsymbol{e}_{\mu_{i+1}} \otimes ... \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes ... \otimes \boldsymbol{e}^{\nu_{k-1}} \otimes \boldsymbol{e}^{\nu_{k+1}} \otimes ... \otimes \boldsymbol{e}^{\nu_q}$$

and due to $\boldsymbol{e}^k \boldsymbol{e}_i = \delta_i^k$ it can be computed by cancelling $\boldsymbol{e}_{\nu_k}$ resp. $\boldsymbol{e}^{\mu_i}$, replacing the respective indices by a common $\nu_k = \mu_i = \lambda$ and replacing the two independent summations over $1 \leq \nu_k; \mu_i \leq n$ by a single summation of

$$x^{\mu_1;...;\mu_{i-1};\lambda;\mu_{i+1};...;\mu_p}_{\nu_1;...;\nu_{k-1};\lambda;\nu_{k+1};...;\nu_q} = x_1^{\mu_1} \cdot ... \cdot x_{i-1}^{\mu_{i-1}} \cdot x_i^{\lambda} \cdot x_{i+1}^{\mu_{i+1}} \cdot ... \cdot x_p^{\mu_p} \cdot x_{\nu_1}^1 \cdot ... \cdot x_{\nu_{k-1}}^{k-1} \cdot x_\lambda^k \cdot x_{\nu_{k+1}}^{k+1} \cdot ... \cdot x_{\nu_q}^q$$

over $1 \leq \lambda \leq n$ each summand being multiplied with the identical tensor product of the basis vectors $\mathbf{a}_{\nu_k}$ resp. $\mathbf{a}^{\mu_i}$ with fixed values for ν_k resp. μ_i.

Example: In the two dimensional real vector space $X = \text{span}\{\boldsymbol{e}_1; \boldsymbol{e}_2\}$ with its dual space $X^* = \text{span}\{\boldsymbol{e}^1; \boldsymbol{e}^2\}$ and

$$\boldsymbol{x} = 2\boldsymbol{e}_1 \otimes \boldsymbol{e}^1 + 3\boldsymbol{e}_1 \otimes \boldsymbol{e}^2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}^1 + 4\boldsymbol{e}_2 \otimes \boldsymbol{e}^2 \in X_1^1$$

$$\begin{aligned}\boldsymbol{y} &= 5\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 - 2\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \\ &+ 0\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}^1 - 0\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}^2 \\ &+ 0\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 - 0\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \\ &+ 0\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}^1 - 0\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}^2 \in X_2^1\end{aligned}$$

we have the **tensor product**

$$\begin{aligned}\boldsymbol{x} \otimes \boldsymbol{y} = {} & 10\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 \otimes \boldsymbol{e}^1 + 15\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \otimes \boldsymbol{e}^1 \\ & - 5\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 \otimes \boldsymbol{e}^1 + 20\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \otimes \boldsymbol{e}^1 \\ & - 4\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 \otimes \boldsymbol{e}^2 - 6\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \otimes \boldsymbol{e}^2 \\ & + 2\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 \otimes \boldsymbol{e}^2 - 8\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \otimes \boldsymbol{e}^2 \\ & \in X_3^2\end{aligned}$$

For the **contraction** γ_1^2 we replace the **first** contravariant vector $\boldsymbol{e}_i$ and the **second** covariant vector $\boldsymbol{e}^j$ by δ_i^jsuch that

$$\begin{aligned}\gamma_1^2(\boldsymbol{x} \otimes \boldsymbol{y}) &= 10\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 + 15\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \\ &+ 2\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 - 8\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2 \\ &= 12\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^1 + 7\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}^2\end{aligned}$$

A second **contraction** results e.g. in

$$\gamma_1^1\left(\gamma_1^2(\boldsymbol{x} \otimes \boldsymbol{y})\right) = 12\boldsymbol{e}_1.$$

7.7 Raising and lowering of indices

Symmetric and positive definite tensors $g_{ij}\boldsymbol{e}^i \otimes \boldsymbol{e}^j \in X_2$ resp. $g^{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \in X^2$ are called **metric tensors** since the corresponding **bilinear forms** $\langle\rangle_g : X^2 \to \mathbb{R}$ with $\langle \boldsymbol{x}; \boldsymbol{y}\rangle_g = x^i g_{ij} y^j$ for $\boldsymbol{x} = x^i \boldsymbol{e}_i$ resp. $\boldsymbol{y} = x^j \mathbf{e}_j$ define a **norm** $\|\,\| : X \to \mathbb{R}^+$ with $\|\boldsymbol{x}\| = \langle \boldsymbol{x}; \boldsymbol{x}\rangle_g$ and hence a **metric** on X resp. X^*. They also provide the coordinates $x_i = x^j g_{ij}$ of the associated **duals** $\boldsymbol{x}^* = \langle \boldsymbol{x};\ \rangle_g \in X^*$ for any $\boldsymbol{x} \in X$ with reference to the dual basis defined by $\mathbf{e}^j = \langle \boldsymbol{e}_j;\ \rangle_g$. In physics the **transposition** from $\boldsymbol{x} = x^j \boldsymbol{e}_j$ to $\tau_X(\boldsymbol{x}) = \boldsymbol{x}^* = x_i \boldsymbol{e}^i = g_{ij} x^j \boldsymbol{e}^i$ is called the **lowering of the index** and the reverse step from $\boldsymbol{x}^* = x_i \boldsymbol{e}^i$ to $\boldsymbol{x} = \tau_X^{-1}(\boldsymbol{x}^*) = x^j \boldsymbol{e}_j = g^{ji} x_i \boldsymbol{e}_j$ is the **raising of the index**.

7.8 Symmetric maps

For complex vector spaces X and Y for $1 \leq i \leq p \geq 2$ and the **symmetric tensors**

$$S_p = \text{span } \left\{ \left(\boldsymbol{x}_{\omega(1)} \otimes ... \otimes \boldsymbol{x}_{\omega(p)} \right) - \left(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \right) : \boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X; \, \omega \in S_p \right\}$$

from 7.3.1 the p**-linear** map $\varphi : X^p \to Y$ is **symmetric** iff it satisfies one of the following obviously equivalent conditions:

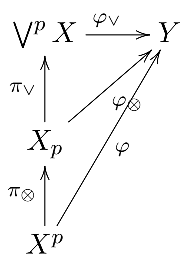
1. $\varphi \left(\boldsymbol{x}_{\omega(1)}; ...; \boldsymbol{x}_{\omega(p)} \right) = \varphi \left(\boldsymbol{x}_1; ...; \boldsymbol{x}_p \right)$ for every **permutation** $\omega \in S_p$
2. $S_p \subset \ker \varphi_{\otimes}$.

7.9 The symmetric product

For every **power** X^p of a **complex** vector space X to an exponent $p \geq 0$ exists a **complex** vector space $\bigvee^p X$ and a **symmetric** map $\vee : X^p \to \bigvee^p X$ such that for every **symmetric** $\varphi : X^p \to Y$ into a **complex** vector space Y exists a **uniquely determined linear** $\varphi_{\vee} : \bigvee^p X \to Y$ with $\varphi = \varphi_{\vee} \circ \vee$.

The **symmetric product** $\bigvee^p X = X_p / S_p$ is the **quotient space** of the **tensor product** $X_p = \bigotimes^p X$ defined in 7.2 and 7.4 over the subspace S_p from 7.8 and its elements $\boldsymbol{x}_1 \vee ... \vee \boldsymbol{x}_p = \pi_{\vee} \left(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \right)$ $= \vee \left(\boldsymbol{x}_1; ...; \boldsymbol{x}_p \right)$ for $\vee = \pi_{\vee} \circ \pi_{\otimes}$ with the p**-linear** map $\pi_{\otimes} : X^p \to X_p$ from 7.2 and the **linear projection** $\pi_{\vee} : X_p \to \bigvee^p X$ are the **symmetric products** of the vectors $\boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X$. According to 7.4 we have $\bigvee^0 X = X_0 = \mathbb{C}$ and $\bigvee^1 X = X_1 = X$. In the finite dimensional case with $\dim X = n$ resp. $\dim Y = r$ the symmetric product has the dimension

$$\dim S_p \left(X^p, Y \right) = \left(\begin{array}{c} n+p-1 \\ p \end{array} \right) \cdot r.$$



Proof: The **linearity** of $\pi_{\vee}$ and 7.2 and the p**-linearity** of $\pi_{\otimes}$ imply the p**-linearity** of $\vee = \pi_{\vee} \circ \pi_{\otimes}$, i.e. $... \vee (c\boldsymbol{y}_k + d\boldsymbol{z}_k) \vee ... = c \left(... \vee \boldsymbol{y}_k \vee ... \right) + d \left(... \vee \boldsymbol{z}_k \vee ... \right)$. Also we have $S_p \subset \ker \left(\pi_{\vee} \circ \pi_{\otimes} \right)$ whence according to 7.8.2 the map $\vee = \pi_{\vee} \circ \pi_{\otimes}$ is **symmetric**, i.e. for every **permutation** $\sigma \in S_p$ holds $\boldsymbol{x}_{\sigma(1)} \vee ... \vee \boldsymbol{x}_{\sigma(p)} = \boldsymbol{x}_1 \vee ... \vee \boldsymbol{x}_p$ and in particular $\boldsymbol{x}_2 \vee \boldsymbol{x}_1 = \boldsymbol{x}_1 \vee \boldsymbol{x}_2$. According to 7.2 for every **symmtric** $\varphi : X^p \to Y$ there is a **uniquely determined** and **linear** $\varphi_{\otimes} : X_p \to Y$ with $\varphi = \varphi_{\otimes} \circ \pi_{\otimes}$. Then due to 3.8 exists a **uniquely determined** and **linear** $\varphi_{\vee} : \bigvee^p X \to Y$ with $\varphi_{\vee} \circ \pi_{\vee}$ $= \varphi_{\otimes}$ whence follows $\varphi_{\vee} \circ \vee = \varphi_{\vee} \circ \pi_{\vee} \circ \pi_{\otimes} = \varphi_{\otimes} \circ \pi_{\otimes} = \varphi$. In the **finite dimensional** case with $\dim X = n$ and **basis** $\{\boldsymbol{e}_1; ...; \boldsymbol{e}_n\}$ we draw p vectors from an urn containing n marked basis vectors e_i plus $p-1$ unmarked dummy vectors for possible **repeats without replacing**. Due to the **symmetry** we do not consider the **order** of the **combinations** whence

$$\dim \bigvee^p = \text{card } \left\{ \boldsymbol{e}_{\mu_1} \vee ... \vee \boldsymbol{e}_{\mu_p} : 1 \leq \mu_1 \leq ... \leq \mu_p \leq n \right\} = \left(\begin{array}{c} n+p-1 \\ p \end{array} \right).$$

Combining these basis vectors with the r basis vectors of Y yields the desired dimension formula.

7.10 Antisymmetric maps

For complex vector spaces X and Y for $1 \leq i \leq p \geq 2$ with

$$A_p = \text{span } \left\{ \left(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \right) : \boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X; \, \exists 1 \leq i < j \leq p : \boldsymbol{x}_i = \boldsymbol{x}_j \right\}$$

the p**-linear** map $\varphi : X^p \to Y$ is **alternating** or **antisymmetric** iff it satisfies one of the following equivalent conditions:

1. $\varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = 0 \Leftrightarrow \exists 1 \leq i < j \leq p : \boldsymbol{x}_i = \boldsymbol{x}_j$.
2. $\varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = 0 \Leftrightarrow \boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X$ are **linearly dependent**.
3. $\varphi\left(\boldsymbol{x}_{\omega(1)}; ...; \boldsymbol{x}_{\omega(p)}\right) = \operatorname{sgn}(\omega) \cdot \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p)$ for every $\boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X$ and every **permutation** $\omega \in S_p$.
4. The **uniquely determined linear** map $\varphi_\otimes : X_p \to Y$ is **alternating**, i.e. $A_p \subset \ker \varphi_\otimes$.

The **vector subspaces** of the alternating p**-linear** resp. **linear** maps are denoted as $A_p(X^p; Y) \subset L_p(X^p; Y)$ resp. $A(X_p; Y) \subset L(X_p; Y)$. In the case of $Y = \mathbb{C}$ from $L_p(X^p; \mathbb{C}) \cong L(X_p; \mathbb{C}) = (X_p; \mathbb{C})^*$ follows $A_p(X^p; \mathbb{C}) \cong A(X_p; \mathbb{C})$. Note that for $p > \dim X$ every alternating map on X^p resp. X_p is the **null map**.

Proof:

1. $\Rightarrow$ 2. : As in 4.1.10 this follows from $\varphi(...; \boldsymbol{x}_i + \boldsymbol{x}_j; ...; \boldsymbol{x}_j; ...) \overset{2.}{=} \varphi(...; \boldsymbol{x}_i; ...; \boldsymbol{x}_j; ...)$ for every $1 \leq i < j \leq p$.

2. $\Rightarrow$ 1. : trivial.

1. $\Rightarrow$ 3. : As in 4.1.12 with $\tau = \langle i; j \rangle$ this follows from $\varphi\left(\boldsymbol{x}_{\tau(1)}; ...; \boldsymbol{x}_{\tau(p)}\right) = \varphi(...; \boldsymbol{x}_j; ...; \boldsymbol{x}_i; ...) \overset{2.}{=} \varphi(...; \boldsymbol{x}_i + \boldsymbol{x}_j; ...; \boldsymbol{x}_i; ...) \overset{2.}{=} \varphi(...; \boldsymbol{x}_i + \boldsymbol{x}_j; ...; -\boldsymbol{x}_j; ...) \overset{2.}{=} \varphi(...; \boldsymbol{x}_i; ...; -\boldsymbol{x}_j; ...) = -\varphi(...; \boldsymbol{x}_i; ...; \boldsymbol{x}_j; ...) = \operatorname{sgn}(\tau) \cdot \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p)$.

3. $\Rightarrow$ 1. : Obvious with the **transposition** $\omega = \tau_{i;j}$.

1. $\Leftrightarrow$ 4. : trivial.

7.11 Antisymmetrization

For every p**-linear** $\varphi : X^p \to Y$ its **antisymmetrical map** $\varphi_a : X^p \to Y$ defined by $\varphi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \varphi\left(\boldsymbol{x}_{\sigma(1)}; ...; \boldsymbol{x}_{\sigma(p)}\right)$ is **antisymmetric** since due to $\operatorname{sgn}(\sigma) = (\operatorname{sgn}(\omega))^2 \cdot \operatorname{sgn}(\sigma) = \operatorname{sgn}(\omega) \cdot \operatorname{sgn}(\sigma \circ \omega)$ for every $\omega \in S_n$we have

$$\begin{aligned}
\varphi_a\left(\boldsymbol{x}_{\omega(1)}; ...; \boldsymbol{x}_{\omega(p)}\right) &= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\omega) \cdot \operatorname{sgn}(\sigma \circ \omega) \cdot \varphi\left(\boldsymbol{x}_{(\sigma \circ \omega)(1)}; ...; \boldsymbol{x}_{(\sigma \circ \omega)(p)}\right) \\
&= \frac{1}{p!} \operatorname{sgn}(\omega) \cdot \sum_{\sigma \circ \omega \in S_p} \operatorname{sgn}(\sigma \circ \omega) \cdot \varphi\left(\boldsymbol{x}_{(\sigma \circ \omega)(1)}; ...; \boldsymbol{x}_{(\sigma \circ \omega)(p)}\right) \\
&= \frac{1}{p!} \operatorname{sgn}(\omega) \cdot \varphi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p).
\end{aligned}$$

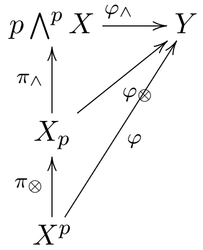
Equivalently it is **alternating** since in the case of $\boldsymbol{x}_i = \boldsymbol{x}_j$ due to 1.22.1 we have $S_p = \tau \circ S_p$ with $\tau = \langle i; j \rangle$ whence $\varphi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \varphi\left(\boldsymbol{x}_{\sigma(1)}; ...; \boldsymbol{x}_{\sigma(p)}\right) = \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\tau_{i;j} \circ \sigma) \cdot \varphi\left(\boldsymbol{x}_{\tau \circ \sigma(1)}; ...; \boldsymbol{x}_{\tau \circ \sigma(p)}\right) = -\frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \varphi\left(\boldsymbol{x}_{\sigma(1)}; ...; \boldsymbol{x}_{\sigma(p)}\right) = -\varphi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = 0$. The asymmetrical map of an already asymmetric map $\varphi : X^p \to Y$ is $\varphi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \varphi\left(\boldsymbol{x}_{\sigma(1)}; ...; \boldsymbol{x}_{\sigma(p)}\right) = \frac{1}{p!} \sum_{\sigma \in S_p} (\operatorname{sgn}(\sigma))^2 \cdot \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \frac{|S_p|}{p!} \cdot \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p)$.

7.12 The exterior product

For every **power** X^p of a **complex** vector space X to an exponent $p \geq 0$ exists a **complex** vector space $\bigwedge^p X$ and an **alternating** map $\wedge : X^p \to \bigwedge^p X$ such that for every **alternating** $\varphi : X^p \to Y$ into a **complex** vector space Y exists a **uniquely determined linear** $\varphi_\wedge : \bigwedge^p X \to Y$ with $\varphi = \varphi_\wedge \circ \wedge$.

The **exterior product** $\bigwedge^p X = X_p / A_p$ is the **quotient space** of the **tensor product** $X_p = \bigotimes^p X$ defined in 7.2 and 7.4 over the subspace A_p from 7.10 and its elements $\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p = \pi_\wedge(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p) = \wedge(\boldsymbol{x}_1; ...; \boldsymbol{x}_p)$ for $\wedge = \pi_\wedge \circ \pi_\otimes$ with the p**-linear** map $\pi_\otimes : X^p \to X_p$ from 7.2 and the **linear projection** $\pi_\wedge : X_p \to \bigwedge^p X$ are the **exterior products** of the vectors $\boldsymbol{x}_1; ...; \boldsymbol{x}_p \in X$. According to 7.4 we have $\bigwedge^0 X = X_0 = \mathbb{C}$ and $\bigwedge^1 X = X_1 = X$.

Proof: The **linearity** of $\pi_\wedge$ and 7.2 and the p-**linearity** of $\pi_\otimes$ imply the p-**linearity** of $\wedge = \pi_\wedge \circ \pi_\otimes$, i.e. $\boldsymbol{x}_1 \wedge ... \wedge (c\boldsymbol{y}_k + d\boldsymbol{z}_k) \wedge ... \wedge \boldsymbol{x}_p = c\,(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{y}_k \wedge ... \wedge \boldsymbol{x}_p)$ $+\, d\,(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{z}_k \wedge ... \wedge \boldsymbol{x}_p)$. Also we have $A_p \subset \ker\ (\pi_\wedge \circ \pi_\otimes)$ whence according to 7.10.4 the map $\wedge = \pi_\wedge \circ \pi_\otimes$ is **alternating**, i.e. for every **permutation** $\sigma \in S_p$ holds $\boldsymbol{x}_{\sigma(1)} \wedge ... \wedge \boldsymbol{x}_{\sigma(p)} = \operatorname{sgn}(\sigma) \cdot (\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p)$ and in particular $\boldsymbol{x}_2 \wedge \boldsymbol{x}_1 = -\boldsymbol{x}_1 \wedge \boldsymbol{x}_2$. Also we have $... \wedge (\boldsymbol{x}_k + c\boldsymbol{x}_l) \wedge ... \wedge \boldsymbol{x}_l \wedge ... = .. \wedge \boldsymbol{x}_k \wedge ... \wedge \boldsymbol{x}_l \wedge ...$ whence follows $\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p = 0$ iff $\boldsymbol{x}_1; ...; \boldsymbol{x}_p$ are **linearly independent**. In particular $X_p = \{\boldsymbol{0}\}$ if $p > n = \dim X$. According to 7.2 for every **alternating** $\varphi : X^p \to Y$ there is a **uniquely determined** and **linear** $\varphi_\otimes : X_p \to Y$ with $\varphi = \varphi_\otimes \circ \pi_\otimes$. Then due to 3.8 exists a **uniquely determined** and **linear** $\varphi_\wedge : \bigwedge^p X \to Y$ with $\varphi_\wedge \circ \pi_\wedge = \varphi_\otimes$ whence follows $\varphi_\wedge \circ \wedge = \varphi_\wedge \circ \pi_\wedge \circ \pi_\otimes = \varphi_\otimes \circ \pi_\otimes = \varphi$.



7.13 The finite dimensional case

For **finite dimensional** vector spaces X with $\dim X = n$ and **basis** $\{\boldsymbol{e}_1; ...; \boldsymbol{e}_n\}$ resp. Y with $\dim Y = r$ and **basis** $\{\boldsymbol{b}_1; ...; \boldsymbol{b}_r\}$ the vector space $A_p\,(X^p, Y)$ of all p-**linear alternating** maps $\varphi : X^p \to Y$ has the **basis** $\mathcal{B} = \left\{\psi^\rho_{\mu_1;...;\mu_p} \cdot \boldsymbol{b}_\rho : 1 \le \mu_1 < ... < \mu_p \le n; 1 \le \rho \le r\right\}$ with $\psi^\rho_{\mu_1;...;\mu_p}\,(\boldsymbol{e}_{\nu_1}; ...; \boldsymbol{e}_{\nu_p}) = \delta^{\nu_1}_{\mu_1} \cdot ... \cdot \delta^{\nu_p}_{\mu_p}$ and

$$\dim A_p\,(X^p, Y) = \begin{pmatrix} n \\ p \end{pmatrix} \cdot r.$$

Proof: Any p-**linear** map $\varphi : X^p \to Y$ has the form $\varphi\,(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \sum\limits_{\iota_k \in I} x_1^{\iota_1} \cdot ... \cdot x_p^{\iota_p} \cdot \varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p})$ for $(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) \in X^p$ with $\boldsymbol{x}_k = \sum\limits_{\iota_k \in I} x_k^{\iota_k} \boldsymbol{e}_{\iota_k}$. Since φ is **alternating** every summand with index vector $(\iota_1; ...; \iota_p)$ having two identical indices must vanish and every remaining summand is a permutation of exactly one ordered **combination** $(\iota_1 < ... < \iota_p)$ such that according to 7.10.3 follows

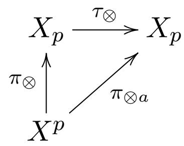
$$\begin{aligned} \varphi\,(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) &= \sum_{\iota_1 < ... < \iota_p \in I} \sum_{\sigma \in S_p} x_1^{\iota_{\sigma(1)}} \cdot ... \cdot x_p^{\iota_{\sigma(p)}} \cdot \varphi\left(\boldsymbol{e}_{\iota_{\sigma(1)}}; ...; \boldsymbol{e}_{\iota_{\sigma(p)}}\right) \\ &= \sum_{\iota_1 < ... < \iota_p \in I} \left(\sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot x_1^{\iota_{\sigma(1)}} \cdot ... \cdot x_p^{\iota_{\sigma(p)}}\right) .\varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p}) \\ &= \sum_{\iota_1 < ... < \iota_p \in I} \det \begin{pmatrix} x_1^{\iota_1} & \cdots & x_1^{\iota_p} \\ \vdots & \ddots & \vdots \\ x_p^{\iota_1} & \cdots & x_p^{\iota_p} \end{pmatrix} \cdot \varphi_\times\,(\boldsymbol{e}_{\iota_1} \wedge ... \wedge \boldsymbol{e}_{\iota_p}) \\ &= \varphi_\times\,(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p) \end{aligned}$$

due to 4.2 with the **linear** $\varphi_\times : \bigwedge^p X \to Y$ from 7.12 **uniquely determined** by $\varphi_\times\,(\boldsymbol{e}_{\iota_1} \wedge ... \wedge \boldsymbol{e}_{\iota_p})$ $= \varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p})$ for every $\iota_1 < ... < \iota_p \in I$. Note that in the case of coinciding indices follows $\boldsymbol{e}_{\iota_1} \wedge ... \wedge \boldsymbol{e}_{\iota_p} = 0$ whence $\varphi_\times\,(\boldsymbol{e}_{\iota_1} \wedge ... \wedge \boldsymbol{e}_{\iota_p}) = \varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p}) = 0$. Also for every permutation $\sigma \in S_p$ holds $\boldsymbol{e}_{\sigma(1)} \wedge ... \wedge \boldsymbol{e}_{\sigma(p)} = \operatorname{sgn}(\sigma) \cdot (\boldsymbol{e}_1 \wedge ... \wedge \boldsymbol{e}_p)$ whence $\varphi_\times\left(\boldsymbol{e}_{\sigma(1)} \wedge ... \wedge \boldsymbol{e}_{\sigma(p)}\right) = \operatorname{sgn}(\sigma) \cdot \varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p})$ $= \operatorname{sgn}(\sigma) \cdot \varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p})$. Hence follows $A_p\,(X^p, Y) \cong L\,(\bigwedge^p X; Y)$ and since $L\,(\bigwedge^p X; \mathbb{R}) \cong (\bigwedge^p X)^*$ $\cong \bigwedge^p X$ with $\varphi\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p}) = \sum\limits_{1 \le \rho \le r} \sum\limits_{\iota_1 < ... < \iota_p \in I} c^\rho_{\iota_1;...;\iota_p} \cdot \psi^\rho_{\iota_1;...;\iota_p}\,(\boldsymbol{e}_{\iota_1}; ...; \boldsymbol{e}_{\iota_p}) \cdot \boldsymbol{b}_\rho \in Y$ with coefficients $c^\rho_{\iota_1;...;\iota_p} \in \mathbb{C}$ for each of the $\begin{pmatrix} n \\ p \end{pmatrix}$ **combinations** $\iota_1 < ... < \iota_p \in I$ of **indices** and each of the r **basis vectors** $\boldsymbol{b}_\rho \in Y$ this implies the assertion.

The extension of the **exterior product** from **vectors** to **antisymmetric tensors** analogously to the extension of the **tensor product** from **vectors** in 7.2 to **tensors** in 7.4 will be introduced in 7.18.

7.14 Antisymmetric tensors

According to 7.11 for every vector space X and $p \geq 0$ the **antisymmetrical** map $\pi_{\otimes a} : X^p \to X_p$ of $\pi_\otimes : X^p \to X_p$ is **alternating** such that according to 7.12 there is a uniquely determined **alternating endomorphism** $\tau_\otimes \in \operatorname{end}(X_p)$ with $\pi_{\otimes a} = \tau_\otimes \circ \pi_\otimes$. The **antisymmetrical tensors** $\tau_\otimes(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p) = \pi_{\otimes a}(\boldsymbol{x}_1; ...; \boldsymbol{x}_p)$ for $\boldsymbol{x}_i \in X$ with $1 \leq i \leq p$ form the **vector subspace** $U_p = \tau_\otimes[X_p] \subset X_p$. Note that the cases $\bigwedge^0 X = X_0 = \mathbb{R}$ and $\bigwedge^1 X = X_1 = X$ are covered by $\pi_\otimes = \pi_{\otimes a} = \tau_\otimes = \mathrm{id}$ and that **every** 0**- resp.** 1**-dimensional vector is trivially antisymmetrical**.



Theorem: $A_p = \ker \tau_\otimes$. Hence **the antisymmetrical tensors** are **isomorphic** to the **exterior products** of **vectors** and to **alternating linear forms:** $U_p \cong \bigwedge^p X = X_p / A_p \cong A_p(X^p, \mathbb{R})$, i.e. $\tau_\otimes(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p) = \pi_{\otimes a}(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) \cong \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \cong (\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p)^*$. In particular the antisymmetrical tensor $\tau_\otimes(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p)$ **vanishes** iff the $(\boldsymbol{x}_i)_{i \in I_p}$ are **linearly dependent**.

Proof: According to 7.10.4 it suffices to show that for every $\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \in \ker \tau_\otimes$ and every **alternating** $\varphi : X^p \to Y$ holds $\varphi_\otimes(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p) = 0$ which implies $\boldsymbol{x} \in A_p$: According to 7.11 for $\varphi = \varphi_\otimes \circ \pi_\otimes$ there exists a p**-linear** $\psi : X^p \to Y$ with **antisymmetrical** $\psi_a = \varphi$. Then for every $(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) \in X^p$ and the corresponding **linear** $\psi_\otimes : X_p \to Y$ holds

$$\begin{aligned}
\varphi_\otimes(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p) &= \varphi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) \\
&= \psi_a(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) \\
&= \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \psi\left(\boldsymbol{x}_{\sigma(1)}; ...; \boldsymbol{x}_{\sigma(p)}\right) \\
&= \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \psi_\otimes\left(\boldsymbol{x}_{\sigma(1)} \otimes ... \otimes \boldsymbol{x}_{\sigma(p)}\right) \\
&= \psi_\otimes\left(\sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) \cdot \boldsymbol{x}_{\sigma(1)} \otimes ... \otimes \boldsymbol{x}_{\sigma(p)}\right) \\
&= (\psi_\otimes \circ \tau_\otimes)(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p)
\end{aligned}$$

whence follows $\varphi_\otimes = \psi_\otimes \circ \tau_\otimes$ which implies the assertion.

7.15 Exterior products in three dimensions

In the case of $p = 2$ we have a simple expression of antisymmetry by

$$\begin{aligned}
\pi_{\otimes a}(\boldsymbol{x}; \boldsymbol{y}) &= \pi_\otimes(\boldsymbol{x}; \boldsymbol{y}) - \pi_\otimes(\boldsymbol{y}; \boldsymbol{x}) \text{ resp.} \\
\boldsymbol{x} \wedge \boldsymbol{y} &= \boldsymbol{x} \otimes \boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{x} \text{ and} \\
\pi_{\otimes a}(\boldsymbol{y}; \boldsymbol{x}) &= \pi_\otimes(\boldsymbol{y}; \boldsymbol{x}) - \pi_\otimes(\boldsymbol{x}; \boldsymbol{y}) \text{ resp.} \\
\boldsymbol{y} \wedge \boldsymbol{x} &= \boldsymbol{y} \otimes \boldsymbol{x} - \boldsymbol{x} \otimes \boldsymbol{y} \\
&= -\boldsymbol{x} \wedge \boldsymbol{y}.
\end{aligned}$$

Hence the **exterior product** $\boldsymbol{x} \wedge \boldsymbol{y} \in \mathbb{R}^3 \wedge \mathbb{R}^3$ of two vectors $\boldsymbol{x} = x^i \boldsymbol{e}_i \in \mathbb{R}^3$ and $\boldsymbol{y} = y^j \boldsymbol{e}_j \in \mathbb{R}^3$ is computed by

$$\boldsymbol{x} \wedge \boldsymbol{y} = \begin{pmatrix} 0 & x^1y^2 - x^2y^1 & x^1y^3 - x^3y^1 \\ y^1x^2 - y^2x^1 & 0 & x^2y^3 - x^3y^2 \\ y^1x^3 - y^3x^1 & y^3x^2 - y^2x^3 & 0 \end{pmatrix}$$

The comparison with the **tensor product** 7.4 may be illustrated by some numerical computations with reference to the canonical basis, e.g.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^3 \otimes \mathbb{R}^3 \text{ but } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^3 \wedge \mathbb{R}^3$$

or

$$\begin{pmatrix}1\\2\\3\end{pmatrix}\otimes\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}0&1&0\\0&2&0\\0&3&0\end{pmatrix}\in\mathbb{R}^3\otimes\mathbb{R}^3 \text{ but } \begin{pmatrix}1\\2\\3\end{pmatrix}\wedge\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}0&1&0\\-1&0&-3\\0&3&0\end{pmatrix}\in\mathbb{R}^3\wedge\mathbb{R}^3$$

The representing **asymmetric** tensor $T = t_{ij}\boldsymbol{e}^i\wedge\boldsymbol{e}^j \in \left(\mathbb{R}^3\wedge\mathbb{R}^3\right)^*$ of a general **alternating** bilinear form $\varphi\in A_2\left(\mathbb{R}^3\times\mathbb{R}^3;\mathbb{R}\right)$ has the form

$$T=\begin{pmatrix}0&t_{12}&t_{13}\\-t_{12}&0&t_{23}\\-t_{13}&-t_{23}&0\end{pmatrix}$$

with $t_{ij}=\varphi\left(\boldsymbol{e}_i;\boldsymbol{e}_j\right)=-\varphi\left(\boldsymbol{e}_j;\boldsymbol{e}_i\right)=-t_{ji}$ whence $\dim\left(\mathbb{R}^3\wedge\mathbb{R}^3\right)=\dim A_2\left(\mathbb{R}^3\times\mathbb{R}^3;\mathbb{R}\right)=\frac{n(n-1)}{2}=3$.
The representing **symmetric** tensor s_{ij} of a **symmetric** bilinear form $\varphi\in S_2\left(\mathbb{R}^3\times\mathbb{R}^3;\mathbb{R}\right)$ has the form

$$S=\begin{pmatrix}t_{11}&t_{12}&t_{13}\\t_{12}&t_{22}&t_{21}\\t_{13}&t_{21}&t_{33}\end{pmatrix}$$

with $t_{ij}=\varphi\left(\boldsymbol{e}_i;\boldsymbol{e}_j\right)=\varphi\left(\boldsymbol{e}_j;\boldsymbol{e}_i\right)=t_{ji}$ whence $\dim S_2\left(\mathbb{R}^3\times\mathbb{R}^3;\mathbb{R}\right)=\frac{n(n+1)}{2}=6$.

7.16 The cross product

An **alternating** map $\times\in A_2\left(\left(\mathbb{R}^3\times\mathbb{R}^3\right);\mathbb{R}^3\right)$ is determined by $\begin{pmatrix}3\\2\end{pmatrix}=3$ conditions in the 3-dimensional image space $\mathbb{R}^3$. If we choose $\times\left(\boldsymbol{e}_1;\boldsymbol{e}_2\right)=\boldsymbol{e}_3$, $\times\left(\boldsymbol{e}_1;\boldsymbol{e}_3\right)=-\boldsymbol{e}_2$ and $\times\left(\boldsymbol{e}_2;\boldsymbol{e}_3\right)=\boldsymbol{e}_1$ we obtain the **cross product** :

$$\boldsymbol{x}\times\boldsymbol{y}=\times\left(\boldsymbol{x};\boldsymbol{y}\right)=\epsilon_{ijk}x^iy^j\boldsymbol{e}_k$$

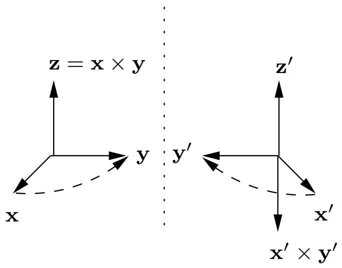
with the **Levi-Civita-symbol** $\epsilon_{ijk}=\begin{cases}\operatorname{sgn}(\sigma) & \text{for } (i;j;k)=\sigma\,(1;2;3) \text{ and } \sigma\in S_3\\ 0 & \text{for } i=j\vee j=k\vee i=k\end{cases}$. Explicitly we have

$$\begin{pmatrix}x^1\\x^2\\x^3\end{pmatrix}\times\begin{pmatrix}y^1\\y^2\\y^3\end{pmatrix}=\begin{pmatrix}x^2y^3-x^3y^2\\x^3y^1-x^1y^3\\x^1y^2-x^2y^1\end{pmatrix}=\begin{pmatrix}0&-x^3&x_2\\x^3&0&-x^1\\-x^2&x^1&0\end{pmatrix}*\begin{pmatrix}y^1\\y^2\\y^3\end{pmatrix}$$

with the corresponding **linear map**

$$\times_\otimes:\mathbb{R}^3\otimes\mathbb{R}^3\to\mathbb{R}^3 \text{ given by } \times_\otimes\begin{pmatrix}m_{11}&m_{12}&m_{13}\\m_{21}&m_{22}&m_{23}\\m_{31}&m_{32}&m_{33}\end{pmatrix}=\begin{pmatrix}m_{23}-m_{32}\\m_{31}-m_{13}\\m_{12}-m_{21}\end{pmatrix}.$$

It is **not injective** and its **kernel** are the **symmetric tensors** $S_2=\ker\times_\otimes$ introduced in 7.3.1.



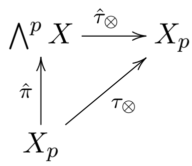
Pseudovectors or **axial vectors** like the the **angular velocity** ω and **every result from a cross product** as opposed to contravariant **polar vectors** are neither contravariant nor covariant. They transform in the usual contravariant way for coordinate transformations $T^{\mathcal{A}}_{\mathcal{B}}$ preserving **orientation** (cf. 4.5) with $\det T^{\mathcal{A}}_{\mathcal{B}} > 0$ e.g. for **translations**, **shears** and **rotations**. However in the case of **reflections** the **right-hand-orientation** of $\boldsymbol{x} \times \boldsymbol{y}$ towards $\boldsymbol{x}; \boldsymbol{y}$ is changed. Since the index notation cannot be applied the transformations are computed separately for each component.

Identities involving the cross product include:

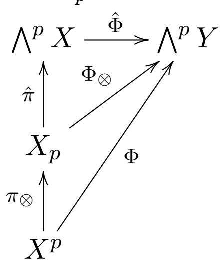
1. **Antisymmetry**: $\boldsymbol{x} \times \boldsymbol{y} = -\boldsymbol{y} \times \boldsymbol{x}$
2. **Distributivity**: $\boldsymbol{x} \times (\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x} \times \boldsymbol{y} + \boldsymbol{x} \times \boldsymbol{z}$
3. **Associativity with the scalar multiplication** $(c \cdot \boldsymbol{x}) \times \boldsymbol{y} = c \cdot (\boldsymbol{x} \times \boldsymbol{y})$
4. **The determinant formula with the canonical inner product** $\boldsymbol{x} * (\boldsymbol{y} \times \boldsymbol{z}) = \epsilon_{ijk} x_i y_j z_k = \det(\boldsymbol{x}; \boldsymbol{y}; \boldsymbol{z})$
5. **The** $BAC - CAB$**-formula**: (cf. 7.5.3)

$$\boldsymbol{x} \times (\boldsymbol{y} \times \boldsymbol{z}) = \begin{pmatrix} x_2 y_1 z_2 - x_2 y_2 z_1 + x_3 y_1 z_3 - x_3 y_3 z_1 \pm x_1 y_1 z_1 \\ -x_1 y_1 z_2 + x_1 y_2 z_1 - x_3 y_2 z_3 + x_3 y_3 z_2 \pm x_2 y_2 z_2 \\ x_1 y_1 z_2 - x_1 y_2 z_1 + x_2 y_2 z_3 - x_1 y_3 z_2 \pm x_3 y_3 z_3 \end{pmatrix} = \boldsymbol{y} \cdot (\boldsymbol{x} * \boldsymbol{z}) - \boldsymbol{z} \cdot (\boldsymbol{x} * \boldsymbol{y}) .$$

7.17 The exterior product of linear maps

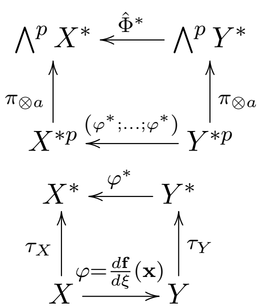


7.17.1 According to 7.12 for every real vector space X and $p \geq 0$ there is a uniquely determined **linear** $\hat{\tau}_\otimes : \bigwedge^p X \to X_p$ with $\tau_\otimes = \hat{\tau}_\otimes \circ \hat{\pi}$. Due to the preceding theorem 7.14 the map $\hat{\tau}_\otimes$ is **injective** whence $\bigwedge^p X \cong U_p = \tau_\otimes [X_p] \subset X_p$: **The exterior products can be identified with the antisymmetric tensors**.



7.17.2 For any vector spaces X and Y, every $p \geq 0$ and every **linear** $\varphi : X \to Y$ the map $\Phi : X^p \to \bigwedge^p Y$ defined by $\Phi(\boldsymbol{x}_1; ...; \boldsymbol{x}_p) = \varphi(\boldsymbol{x}_1) \wedge ... \wedge \varphi(\boldsymbol{x}_p)$ according to 7.12 is p**-linear** and **alternating.** In the case of $p = 0$ the definitions yield $\Phi = \mathrm{id} : X^0 = \mathbb{R} \to \bigwedge^0 Y = \mathbb{R}$ and for $p = 1$ we have $\Phi = \varphi : X^1 = X \to \bigwedge^1 Y = Y$. Due to 7.10 the corresponding **linear** map $\Phi_\otimes : X_p \to \bigwedge^p Y$ is again **alternating** with $\Phi = \Phi_\otimes \circ \pi_\otimes$. Finally and due to 7.12 we have a uniquely determined **linear alternating product** $\hat{\Phi} : \bigwedge^p X \to \bigwedge^p Y$ with $\Phi_\otimes = \hat{\Phi} \circ \hat{\pi}$ resp.

$$\hat{\Phi}(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p) = \varphi(\boldsymbol{x}_1) \wedge ... \wedge \varphi(\boldsymbol{x}_p) .$$



7.17.3 The **inverse** or **pullback image of a differential form** (cf. [2, th 5.3]) is provided by the corresponding **alternating product** $\hat{\Phi}^* : \bigwedge^p Y^* \to \bigwedge^p X^*$ **of the dual** $\varphi^* : Y^* \to X^*$ of the linear map given by the **derivative** $\varphi = D\boldsymbol{f}(\boldsymbol{x}) : X \to Y$ **at the point** $\boldsymbol{x} \in X$ such that

$$\hat{\Phi}^* \left(\boldsymbol{y}^1 \wedge ... \wedge \boldsymbol{y}^p\right) = \varphi^* \left(\boldsymbol{y}^1\right) \wedge ... \wedge \varphi^* (\boldsymbol{y}^p) = \boldsymbol{y}^1 \circ \varphi \wedge ... \wedge \boldsymbol{y}^p \circ \varphi$$

and

$$\hat{\Phi}^* (\mathbf{y}^p \wedge \mathbf{y}^q) = \hat{\Phi}^* (\mathbf{y}^p) \wedge \hat{\Phi}^* (\mathbf{y}^q)$$

for every $\boldsymbol{y}^p \in \bigwedge^p Y^*$ resp. $\boldsymbol{y}^q \in \bigwedge^q Y^*$. Consequently we have $\hat{\Phi}^* = \mathrm{id} : \mathbb{R}^* \to \mathbb{R}^*$ for $p = 0$ and $\hat{\Phi}^* = \varphi^* : Y^* \to X^*$ for $p = 1$.

7.18 The general exterior product

The general **exterior product** $\wedge : (\bigwedge^p X \times \bigwedge^q X) \to \bigwedge^{p+q} X$ defined by

$$\hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}} = (\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p) \wedge \left(\boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q\right) = \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \wedge \boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q$$

is a **bilinear** map between the **asymmetric tensors** $\hat{\boldsymbol{x}} = \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \in \bigwedge_p X$ and $\hat{\boldsymbol{y}} = \boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q \in \bigwedge^q X$. According to 7.12 it coincides with the **exterior product** for 1-vectors $\boldsymbol{x} = \hat{\boldsymbol{x}}; \boldsymbol{y} = \hat{\boldsymbol{y}} \in X = \bigwedge^1 X$ with $\boldsymbol{x} \wedge \boldsymbol{y} = \hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}}$. The following properties hold:

1. **associativity**: $(c\hat{\boldsymbol{x}}) \wedge \hat{\boldsymbol{y}} = \hat{\boldsymbol{x}} \wedge (c\hat{\boldsymbol{y}}) = c \cdot (\hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}})$ resp. $(\hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}}) \wedge \hat{\boldsymbol{z}} = \hat{\boldsymbol{x}} \wedge (\hat{\boldsymbol{y}} \wedge \hat{\boldsymbol{z}})$
2. **distributivity**: $\hat{\boldsymbol{x}} \wedge (\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}) = \hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}} + \hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{z}}$ resp. $(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}}) \wedge \hat{\boldsymbol{z}} = \hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{z}} + \hat{\boldsymbol{y}} \wedge \hat{\boldsymbol{z}}$
3. **anticommutativity**: $\hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}} = (-1)^{pq} (\hat{\boldsymbol{y}} \wedge \hat{\boldsymbol{x}})$.

Proof: Due to 7.3 for the $p+q$ **-linear** $\varphi : X^{p+q} \to \bigwedge^{p+q} X$ defined by

$$\varphi\left(\boldsymbol{x}_1; ...; \boldsymbol{x}_p; \boldsymbol{y}_1; ...; \boldsymbol{y}_q\right) = \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \wedge \boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q$$

there is a unique **linear** $\varphi_\otimes : X_{p+q} \to \bigwedge^{p+q} X$ with $\varphi = \varphi_\otimes \circ \pi_\otimes = \hat{\pi} \circ \pi_\otimes$, i.e.

$$\varphi_\otimes\left(\boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \otimes \boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_q\right) = \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \wedge \boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q$$

Hence owing to the **distributivity of the tensor product** 7.4 the map $\Psi : (X_p \times X_q) \to \bigwedge^{p+q} X$ with

$$\Psi(\boldsymbol{x}; \boldsymbol{y}) = \varphi_\otimes(\boldsymbol{x} \otimes \boldsymbol{y}) = \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_p \wedge \boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_q$$

is **bilinear** in $\boldsymbol{x} = \boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \in X_p$ and $\boldsymbol{y} = \boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_q \in X_q$. Due to the **linearity of the canonical projections** $\hat{\pi}_p : X_p \to \bigwedge^p X$ resp. $\hat{\pi}_q : X_q \to \bigwedge^q X$ the map $\wedge : (\bigwedge^p X \times \bigwedge^q X) \to \bigwedge^{p+q} X$ defined by $\Psi = \wedge \circ (\hat{\pi}_p; \hat{\pi}_q)$ is still **bilinear** and obviously coincides with the desired exterior product. The properties 1. - 3. directly follow from the definition.

7.19 The scalar product

For $p; q \geq 1$ and a **finite-dimensional unitary** vector space X with an **orthonormal basis** $\mathcal{B} = (\boldsymbol{e}_i)_{1 \leq i \leq n}$ resp. the dual space X^* with the canonical dual basis $\mathcal{B}^* = \left(\boldsymbol{e}^i\right)_{1 \leq i \leq n}$ determined by $\boldsymbol{e}^j \boldsymbol{e}_i = \langle \boldsymbol{e}_i; \boldsymbol{e}^j \rangle = \delta_i^j$ the **scalar product** $\langle\rangle : X_p^q \times X_q^p \to \mathbb{R}$ is a **bilinear** form defined by

$$\langle \boldsymbol{x}; \boldsymbol{y} \rangle = \prod_{i=1}^{p} \left\langle \boldsymbol{x}_i; \boldsymbol{y}^i \right\rangle \cdot \prod_{j=1}^{q} \left\langle \boldsymbol{y}_j; \boldsymbol{x}^j \right\rangle = x_{\mu_1; ...; \mu_q}^{\nu_1; ...; \nu_p} \cdot y_{\nu_1; ...; \nu_p}^{\mu_1; ...; \mu_q}$$

for

$$\boldsymbol{x} = \boldsymbol{x}_1 \otimes ... \otimes \boldsymbol{x}_p \otimes \boldsymbol{x}^1 \otimes ... \otimes \boldsymbol{x}^q = x_{\mu_1; ...; \mu_p}^{\nu_1; ...; \nu_q} \boldsymbol{e}_{\nu_1} \otimes ... \otimes \boldsymbol{e}_{\nu_p} \otimes \boldsymbol{e}^{\mu_1} \otimes ... \otimes \boldsymbol{e}^{\mu_q} \in X_p^q$$

resp.

$$\boldsymbol{y} = \boldsymbol{y}_1 \otimes ... \otimes \boldsymbol{y}_q \otimes \boldsymbol{y}^1 \otimes ... \otimes \boldsymbol{y}^p = y_{\nu_1; ...; \nu_p}^{\mu_1; ...; \mu_q} \boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_q} \otimes \boldsymbol{e}^{\nu_1} \otimes ... \otimes \boldsymbol{e}^{\nu_p} \in X_q^p$$

and the following properties for $c \in \mathbb{R}$ and $\boldsymbol{x}; \boldsymbol{y}; \mathbf{z} \in X$:

1. **Associativity**: $\langle (c\boldsymbol{x})\,; \boldsymbol{y}\rangle = c \cdot \langle \boldsymbol{x}; \boldsymbol{y}\rangle = \langle \boldsymbol{x}; (c\boldsymbol{y})\rangle$
2. **Distributivity**: $\langle \boldsymbol{x}; (\boldsymbol{y}+\boldsymbol{z})\rangle = \langle \boldsymbol{x}; \boldsymbol{y}\rangle + \langle \boldsymbol{x}; \boldsymbol{z}\rangle$ resp. $\langle (\boldsymbol{x}+\boldsymbol{y})\,; \mathbf{z}\rangle = \langle \boldsymbol{x}; \boldsymbol{z}\rangle + \langle \boldsymbol{y}; \boldsymbol{z}\rangle$
3. **Symmetry**: $\langle \boldsymbol{x}; \boldsymbol{y}\rangle = \langle \boldsymbol{y}; \boldsymbol{x}\rangle$

In the case of $(p;q) = (1;0)$ and a finite-dimensional X the **general scalar product** assumes the form $\langle\,\rangle : (X_1 \times X^1 = X \times X^* \to \mathbb{R})$ with $\left\langle \boldsymbol{e}_i; \boldsymbol{e}_j^*\right\rangle = \boldsymbol{e}_j^*\boldsymbol{e}_i = \langle \boldsymbol{e}_i; \boldsymbol{e}_j\rangle$ resp. $\langle \boldsymbol{x}; \boldsymbol{y}^*\rangle = \boldsymbol{y}^*\boldsymbol{x} = x^i y_i = \boldsymbol{x}_{\mathcal{A}} * \mathbf{y}_{\mathcal{A}} = \langle \boldsymbol{x}; \boldsymbol{y}\rangle$ for $\boldsymbol{x} = x^i\boldsymbol{e}_i$ resp. $\boldsymbol{y}^* = y_i\boldsymbol{e}^i$ hence **coinciding with the canonical bilinear form** $\langle\,\rangle : (X \times X \to \mathbb{R})$ defined by $\langle \boldsymbol{x}; \mathbf{y}\rangle = \mathbf{y}^*\boldsymbol{x}$. Thus the general scalar product provides a distinct interpretation of **row vectors** ${}^T\boldsymbol{x} \in X_1$ and **column vectors** $\boldsymbol{x} \in X^1$. However this coincidence is confined to **finite dimensional spaces** since **function space**s as e.g. $\mathcal{C}_c(\mathbb{R})$ (cf. [4, th. 10.12]) or $L^p(\lambda)$ (cf. [4, th. 9.13]) in general are **not isomorphic to their dual space any more**. Note also that the distinction between row and column vectors is meaningless for $p+q > 2$. (cf. 7.3)

7.20 The exterior algebra

For a real vector space X the vector space $\bigwedge X = \bigoplus_{p\geq 0} \bigwedge^p X = \left\{ \sum\limits_{p=0}^{m} \boldsymbol{x}_p : \boldsymbol{x}_p \in \bigwedge^p X; m \in \mathbb{N}\right\}$ with the **exterior product** $\wedge : \bigwedge X \to \bigwedge X$ with $\hat{\boldsymbol{x}} \wedge \hat{\boldsymbol{y}} = \sum\limits_{p=0}^{m}\sum\limits_{q=0}^{n} \left(\boldsymbol{x}_p \wedge \boldsymbol{y}_q\right)$ for $\hat{\boldsymbol{x}} = \sum\limits_{p=0}^{m} \boldsymbol{x}_p$ and $\hat{\boldsymbol{y}} = \sum\limits_{q=0}^{n} \boldsymbol{y}_q$ is the **exterior algebra** over X. In the case of a **finite dimensional** with $\dim X = n$ according to 7.12 we have $X_p = \{\mathbf{0}_n\}$ if $p > n = \dim X$ which implies $\bigwedge X = \bigoplus_{0\leq p\leq n} \bigwedge^p X$ with

$$\dim \bigwedge X = \sum_{p=0}^{n} \dim \bigwedge^{p} X = \sum_{p=0}^{n} \begin{pmatrix} n \\ p \end{pmatrix} = 2^n.$$

According to 7.19 the **scalar product** $\langle\,\rangle : X_p \times X^p \to \mathbb{R}$ defined by $\langle \boldsymbol{x}; \boldsymbol{y}^*\rangle = \prod\limits_{i=1}^{p} \langle \boldsymbol{x}_i; \boldsymbol{y}^i\rangle$ for $\boldsymbol{x} = \boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_p \in X_p$ resp. $\boldsymbol{y} = \boldsymbol{y}^1 \otimes \dots \otimes \boldsymbol{y}^p \in X^p$ provides an **isomorphism** $\eta : X^p \to X_p^*$ given by $\eta\mathbf{y}^* : X_p \to \mathbb{C}$ with $\eta\boldsymbol{y}^*\boldsymbol{x} = \langle \boldsymbol{x}; \boldsymbol{y}^*\rangle$. Hence ***p*-covariant tensors can be identified with linear maps on the tensor product** X_p. This isomorphism extends to the subspace $U^p = \tau_\otimes [X^p] \subset X^p$ of the **antisymmetric *p*-covariant tensors** $\tau_\otimes \boldsymbol{y}^*$ for $\boldsymbol{y}^* \in X^p$ defined in 7.14 and omitting the brackets for brevity resp. the subspace $A(X_p, \mathbb{R}) \subset X_p^*$ of all **linear alternating** forms $\varphi : X_p \to \mathbb{R}$ from 7.13: Due to 7.14 the **antisymmetric tensor** of $\boldsymbol{y}^*$ is $\tau_\otimes(\boldsymbol{y}^*) = \sum_{\sigma\in S_p} \operatorname{sgn}(\sigma) \cdot \boldsymbol{y}^{\sigma(1)} \otimes \dots \otimes \boldsymbol{y}^{\sigma(p)}$ such that

$$\begin{aligned}
(\eta \circ \tau_\otimes \circ \boldsymbol{y}^*)(\boldsymbol{x}) &= \langle \boldsymbol{x}; \tau_\otimes \boldsymbol{y}^*\rangle \\
&= \sum_{\sigma\in S_p} \operatorname{sgn}(\sigma) \cdot \left\langle \boldsymbol{x}_1; \boldsymbol{y}^{\sigma(1)}\right\rangle \cdot \dots \cdot \left\langle \boldsymbol{x}_1; \boldsymbol{y}^{\sigma(p)}\right\rangle \\
&= \sum_{\sigma^{-1}\in S_p} \operatorname{sgn}\left(\sigma^{-1}\right) \cdot \left\langle \boldsymbol{x}_{\sigma^{-1}(1)}; \boldsymbol{y}^1\right\rangle \cdot \dots \cdot \left\langle \boldsymbol{x}_{\sigma^{-1}(p)}; \boldsymbol{y}^p\right\rangle \\
&= \sum_{\sigma^{-1}\in S_p} \operatorname{sgn}\left(\sigma^{-1}\right) \cdot \eta\boldsymbol{y}^* \left(\boldsymbol{x}_{\sigma^{-1}(1)} \otimes \dots \otimes \boldsymbol{x}_{\sigma^{-1}(p)}\right) \\
&= (\eta\boldsymbol{y}^*)_a(\boldsymbol{x})
\end{aligned}$$

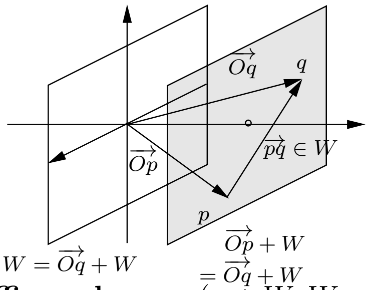
according to the definition of the **antisymmetric**al map in 7.11. Since the **linear alternating** forms $A(X_p, \mathbb{C}) \subset X_p^*$ defined in 7.10.4 are exactly the **antisymmetricals** of all linear forms in $L(X_p; \mathbb{C})$ we conclude that $U^p = \tau_\otimes [X^p] \cong A(X_p, \mathbb{C})$.

8 Affine spaces

8.1 Affine spaces

An **affine space** is a triple $(A; X_A; \rightarrow)$ of a set A, a **vector space** X_A and a map $\rightarrow: A \times A \rightarrow X_A$ such that for every $p; q; r \in A$ holds

1. $\forall \boldsymbol{a} \in X_A \, \exists p_{\boldsymbol{a}} \in A : \, \boldsymbol{a} = \overrightarrow{pp_{\boldsymbol{a}}}$
2. $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$



with immediate consequences

3. $\overrightarrow{pp} = \boldsymbol{0}$
4. $\overrightarrow{qp} = -\overrightarrow{pq}$

Its **dimension** is $\dim A = \dim X_A$. The most common example is the **affine subspace** $(\boldsymbol{v} + W; W; -)$ with $\boldsymbol{v} + V = \{\boldsymbol{x} \in X : \boldsymbol{x} - \boldsymbol{v} \in V\}$ generated by a vector subspace $V \subset X$ of a vector space X and any vector $\boldsymbol{v} \in X$. This example includes the vector space $(X; X; -)$ itself regarded as a point set. Geometrically speaking an affine space is a **vector space without a predetermined reference point** resp. origin. The reference point resp. support $\mathbf{v}$ can be chosen **arbitrarily** as a part of the **coordinate system**.

8.2 Affine subspaces

The set $U \subset A$ is an **affine subspace** iff $X_U = \{\overrightarrow{pq} : q \in U\}$ is a vector subspace for some $p \in U$ and this definition is **independent of the choice of** p. For any family $\mathcal{U}$ of affine subspaces its **intersection** $\bigcap \{U : U \in \mathcal{U}\}$ is again an affine subspace with $X_D = \bigcap \{X_U : U \in \mathcal{U}\}$. Hence any subset $M \subset A$ generates an affine subspace $[M]$ defined as the intersection of all affine subspaces containing M. The affine subspace $\bigvee \{U : U \in \mathcal{U}\}$=$[\bigcup \{U : U \in \mathcal{U}\}]$ generated by their **union** is their **affine hull**. The affine hull of a point $p \in A$ is the point itself with the corresponding vector subspace $X_p = \{\}$and $\dim p = 0$. The affine hull $p \vee q$ of two distinct points $p; q \in A$ is a **line** with $\dim p \vee q = 1$. The affine hull $p \vee q \vee r$ of three points $p; q; r \in A$ with linearly independent $\overrightarrow{pq}; \overrightarrow{qr}$ is a **plane** with $\dim p \vee q \vee r = 2$. An affine subspace $U \subsetneq A$ is a hyperplane iff there is a point p with $p \vee U = A$. The affine hull of finite dimensional affine subspaces $U; V \subset A$ has the dimension

$$\dim (U \vee V) = \dim U + \dim V - \dim (U \cap V) \, .$$

Two affine subspaces $U; V \subset A$ are **parallel**, in short $U \| V$, iff $X_U \subset X_V$ or $X_V \subset X_U$ and in that case they are either disjoint or one of them is contained in the other. A nonempty subspace U and a **hyperplane** H are either parallel or $\dim (U \cap H) = \dim U - 1$.

8.3 Affine coordinate systems

The points $\mathcal{P} = (p_i)_{0 \leq i \leq n} \subset A$ of an n-dimensional affine space A are a **coordinate system** of A iff $(\overrightarrow{p_0 p_i})_{1 \leq i \leq n} \subset X_A$ is a **basis** of X_A resp. iff $\bigvee (p_i)_{0 \leq i \leq n} = A$. Every point $q \in A$ has a uniquely determined **coordinate vector** $\boldsymbol{q}_{\mathcal{P}} \in \mathbb{C}^n$ with $\overrightarrow{p_0 q} = \sum_{i=1}^{n} q_{\mathcal{P}i} \overrightarrow{p_0 p_i}$ and for every other $r \in A$ we have $\overrightarrow{qr} = \sum_{i=1}^{n} (r_{\mathcal{P}i} - q_{\mathcal{P}i}) \overrightarrow{p_0 p_i}$, i.e. $\boldsymbol{qr}_{\mathcal{P}} = \boldsymbol{r}_{\mathcal{P}} - \boldsymbol{q}_{\mathcal{P}}$. The transformation from the affine coordinate system $\mathcal{P} = (p_i)_{0 \leq i \leq n}$ to the system $\mathcal{Q} = (q_i)_{0 \leq i \leq n}$ with $\overrightarrow{p_0 p_j} = \sum_{i=1}^{n} t_{i;j} \overrightarrow{q_0 q_i}$ and $\overrightarrow{p_0 q_0} = \sum_{i=1}^{n} s_{\mathcal{P}i} \overrightarrow{p_0 p_i}$ is determined by the **translation vector** $\mathbf{s} = \overrightarrow{p_0 q_0}$ with the coordinate vector $\boldsymbol{s}_P = (s_{\mathcal{P}i})_{0 \leq i \leq n}$ and the **transformation matrix** $T = (t_{i;j})_{0\leq i;j\leq n}$. The coordinate vector $\boldsymbol{r}_Q$ of a point $r \in A$ can be computed by

$$\begin{aligned}\overrightarrow{q_0r} &= \overrightarrow{q_0p_0} + \overrightarrow{p_0r}\\ &= -\sum_{j=1}^{n} s_{\mathcal{P}j}\overrightarrow{p_0p_j} + \sum_{j=1}^{n} r_{\mathcal{P}j}\overrightarrow{p_0p_j}\\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} t_{i;j}\left(r_{\mathcal{P}j} - s_{\mathcal{P}j}\right)\overrightarrow{q_0q_i}\end{aligned}$$

whence

$$\boldsymbol{r}_Q = T * (\boldsymbol{r}_P - \boldsymbol{s}_P) = T * \boldsymbol{r}_P - \boldsymbol{s}_Q \,\text{resp.}\, \boldsymbol{r}_P = T^{-1} * (\boldsymbol{r}_Q + \boldsymbol{s}_Q) = T^{-1} * \boldsymbol{r}_Q + \boldsymbol{s}_P.$$

8.4 Affine maps

A map $\Phi : A \to B$ between affine spaces A and B is **affine** iff there is a **linear** $\varphi : X_A \to X_B$ with $\varphi(\overrightarrow{pr}) = \overrightarrow{\Phi(p)\,\Phi(r)}$ for every $p; r \in A$. Hence the affine map is determined by a **linear map** φ and a **point** $r_0 = \Phi(p_0)$. Conversely for any points $(r_i)_{0\leq i\leq n} \subset B$ given by the coordinates there is a uniquely determined affine map $\Phi : A \to \bigvee_{0\leq i\leq n} r_i$ with $\Phi(p_i) = r_i$ for $0 \leq i \leq n$ and it is **bijective** iff $(r_i)_{0\leq i\leq n}$ is a **coordinate system**. For affine coordinate systems $\mathcal{P} = (p_i)_{0\leq i\leq n}$ of A resp. $\mathcal{Q} = (q_i)_{0\leq i\leq n}$ of B and image points $(r_i)_{0\leq i\leq n} \subset B$ with $r_i = \Phi(p_i)$ we have

$$\begin{aligned}\overrightarrow{q_0r_i} &= \overrightarrow{q_0r_0} + \overrightarrow{r_0r_i}\\ &= \overrightarrow{q_0r_0} + \overrightarrow{\Phi(p_0)\,\Phi(p_i)}\\ &= \overrightarrow{q_0r_0} + \varphi(\overrightarrow{p_op_i})\end{aligned}$$

resp.

$$\begin{aligned}\overrightarrow{q_0r_i} &= \sum_{j=1}^{n} r_{\mathcal{Q}j;i}\overrightarrow{q_0q_j}\\ &= \sum_{j=1}^{n}\left(r_{\mathcal{Q}j;0} + (r_{\mathcal{Q}j;i} - r_{\mathcal{Q}j;0})\right)\overrightarrow{q_0q_j}\\ &= \sum_{j=1}^{n} f_{\mathcal{Q}j;0}\overrightarrow{q_0q_j} + \sum_{j=1}^{n} f_{\mathcal{Q}j;i}\overrightarrow{q_0q_j}\end{aligned}$$

whence the coordinate vectors $\boldsymbol{r}_{\mathcal{Q}i} = \boldsymbol{f}_{Q0} + \boldsymbol{f}_{Qi}$ with $f_{\mathcal{Q}j;0} = r_{\mathcal{Q}j;0}$ and $f_{\mathcal{Q}j;i} = r_{\mathcal{Q}j;i} - r_{\mathcal{Q}j;0}$ of the points $r_i = \Phi(p_i)$ decompose into the **fixed** part $\boldsymbol{f}_{Q0}$ representing the translation $\overrightarrow{q_0r_0}$ and the **linear** part $\boldsymbol{f}_{Qi}$ of the map $\overrightarrow{r_0r_i} = \varphi(\overrightarrow{p_op_i})$. The image of an arbitrary point $s \in A$ with $\overrightarrow{p_0s} = \sum_{i=1}^{n} s_{\mathcal{P}i}\overrightarrow{p_0p_i}$ can be computed by

$$\begin{aligned}\overrightarrow{q_0\Phi(s)} &= \overrightarrow{q_0r_0} + \overrightarrow{r_0\Phi(s)}\\ &= \overrightarrow{q_0r_0} + \overrightarrow{\Phi(p_0)\,\Phi(s)}\\ &= \sum_{j=1}^{n} f_{\mathcal{Q}0;j}\overrightarrow{q_0q_j} + \varphi(\overrightarrow{p_0s})\\ &= \sum_{j=1}^{n} f_{\mathcal{Q}0;j}\overrightarrow{q_0q_j} + \sum_{i=1}^{n} s_{\mathcal{P}i}\cdot\varphi(\overrightarrow{p_op_i})\\ &= \sum_{j=1}^{n} f_{\mathcal{Q}0;j}\overrightarrow{q_0q_j} + \sum_{i=1}^{n}\sum_{j=1}^{n} s_{\mathcal{P}i}\cdot f_{\mathcal{Q}j;i}s_{\mathcal{P}i}\overrightarrow{q_0q_j}\end{aligned}$$

and its coordinate vector is

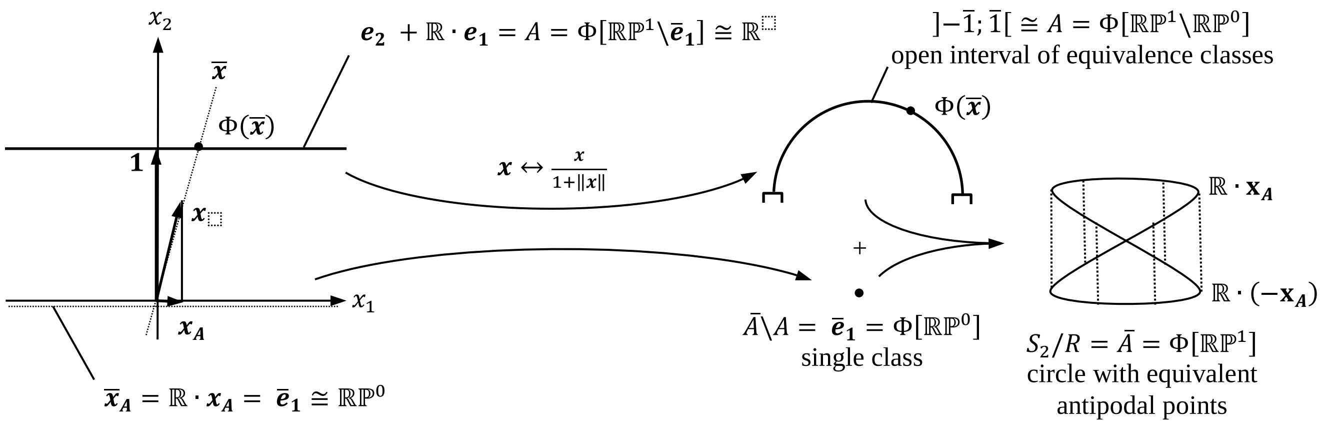
$$(\Phi(s))_{\mathcal{Q}} = \boldsymbol{f}_{Q0} + F * \boldsymbol{s}_{\mathcal{Q}}$$

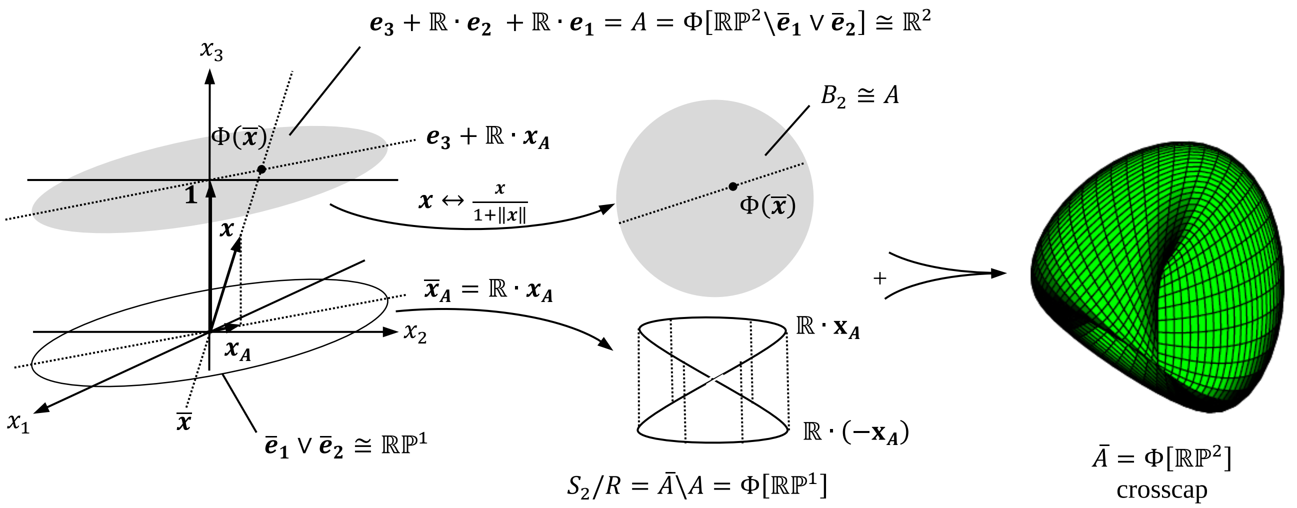
with the **representing matrix** $F = M_Q^P(\varphi) = (f_{\mathcal{Q}i;j})_{1 \leq i;j \leq n}$. The maps Φ resp. φ are **bijective** iff the representing matrix F is **invertible** resp. iff its column vectors $\boldsymbol{f}_{Qi}$ are **linearly independent** since these are the **coordinate vectors** of $\overrightarrow{r_0 r_i}$. In that case Φ is an **affinity** and A is **affine** to B. Every n-dimensional affine space $(A; X_A; \rightarrow)$ with $A = \bigvee_{i=0}^{n} p_i$ by $\Phi(p_i) = e_i$ is **affine** to the **canonical affine space** $(\mathbb{A}_n; \mathbb{C}^n; \rightarrow)$ with $\mathbb{A}_n = \bigvee_{i=0}^{n} e_i$ defined by an **arbitrary origin** e_0 and $\overrightarrow{e_0 e_i} = \boldsymbol{e}_i$ for the canonical basis $(\boldsymbol{e}_i)_{1 \leq i \leq n}$ of $\mathbb{C}^n$. The **image** $\Phi[U] \subset B$ of every affine subspace $U \subset A$ is again an affine subspace with $X_{\Phi[U]} = \varphi[X_U]$ and the **reverse image** $\Phi^{-1}[V] \subset A$ of every affine subspace $V \subset B$ is again an affine subspace with $X_{\Phi^{-1}[V]} = \varphi^{-1}[X_V]$. The **composition** $\Psi \circ \Phi : A \rightarrow C$ of affine maps $\Phi : A \rightarrow B$ and $\Psi : B \rightarrow C$ is again an affine map such that the set of affine **bijections** on an affine space A forms the **affine group**. In the case of $\varphi = \mathrm{id}_A$ we have a **translation** with $\overrightarrow{p\Phi(p)} = \overrightarrow{q\Phi(q)}$ for all points $p; q \in A$.

9 Projective spaces

9.1 Definitions

The **projective space** $\mathbb{P}X$ of the **finite dimensional vector space** X over a **field** K is the **orbit space** $X_* \setminus K_*$ generated as its quotient space with regard to the **equivalence relation** given by $\boldsymbol{x}K_*\boldsymbol{y} \Leftrightarrow \exists t \in K_* : t\boldsymbol{x} = \boldsymbol{y}$ resp. by the action of the multiplicative group K_* $K \setminus \{0\}$ on $X_* = X \setminus \{\mathbf{0}\}$ as defined in [6] Def. 22.1 resp. 1.15. Its equivalence classes are the **one-dimensional vector subspaces** in X **without the origin**. They are represented by **homogenous coordinates** $\overline{\boldsymbol{x}}_{\mathcal{B}} = [x_{\mathcal{B}1} : \ldots : x_{\mathcal{B}n+1}] = \left\{ t \cdot \sum_{n=1}^{n+1} x_{\mathcal{B}i}\boldsymbol{e}_i : t \in K \right\} = \pi(\boldsymbol{x}_{\mathcal{B}})$ with respect to the **basis** $\mathcal{B} = (\boldsymbol{e}_i)_{1 \le i \le n+1}$ of X. Although the **projective space is not a vector space** we define its **dimension** as $\dim \mathbb{P}X = \dim X - 1$. The projective space $K\mathbb{P}^n = \mathbb{P}K^{n+1}$ of the **vector space** K^{n+1}over a **field** $K \in \{\mathbb{R}; \mathbb{C}\}$ is a **compact K^n-manifold** as defined in [6] 20.8. It needs $n+1$ charts $(\pi\{x_i \neq 0\}; \varphi_i)$ with **coordinates** $\varphi_i : \pi\{x_i \neq 0\} \to K^n$ given by $\varphi_i[x_1 : \ldots : x_{n+1}] = \left(\frac{x_1}{x_i}; \ldots; \frac{x_{i-1}}{x_i}; \frac{x_{i+1}}{x_i}; \ldots; \frac{x_{n+1}}{x_i}\right)$ and **parametrizations** $\varphi_i^{-1}(x_1; \ldots; x_n) = [x_1 : \ldots : 1 : \ldots : x_n]$. Note that the projection $\pi\{x_i \neq 0\}$ of the **saturated open** set $\{x_i \neq 0\} = K^n \setminus \{x_i = 0\}$ is **open** in the **quotient topology** of $K\mathbb{P}^n$ as described in [6] 4.7. A **projective subspace** is the projective space of the corresponding vector subspace and the **projective hull** $\mathbb{P}X \vee \mathbb{P}Y = \mathbb{P}(X \oplus Y)$ is the projective space of the sum of the corresponding vector spaces with $\dim(\mathbb{P}X \vee \mathbb{P}Y) = \dim \mathbb{P}X + \dim \mathbb{P}X - \dim(\mathbb{P}X \cap \mathbb{P}Y)$.





9.2 Projective maps

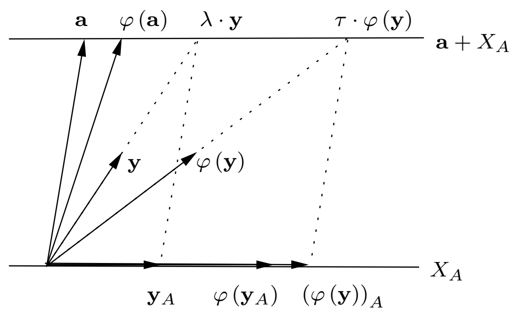
A map $\Phi : \mathbb{P}X \to \mathbb{P}Y$ is **projective** iff there is an **injective linear** $\varphi : X \to Y$ with $\Phi[\overline{\boldsymbol{x}}] = \overline{\varphi(\boldsymbol{x})}$ for every $\mathbf{0} \neq \boldsymbol{x} \in X$. A **bijective** projective map is a **projectivity**. For linear maps $\varphi, \varphi' : X \to Y$ with projective map $\Phi; \Phi' : \mathbb{P}X \to \mathbb{P}Y$ we have $\Phi = \Phi'$ iff there is a $\lambda \in K_*$ with $\lambda \cdot \varphi = \varphi'$ since for every pair of **linearly independent** $\boldsymbol{x}; \boldsymbol{y} \in X_*$ there are $\lambda; \mu; \nu \in K_*$ with $\Phi'(\boldsymbol{x}) = \lambda\Phi(\boldsymbol{x})$, $\Phi'(\boldsymbol{y}) = \mu\Phi(\boldsymbol{y})$ and $\Phi'(\boldsymbol{x}+\boldsymbol{y}) = \nu\Phi(\boldsymbol{x}+\boldsymbol{y})$ resp. $(\lambda-\mu)\Phi(\boldsymbol{x}) - (\lambda-\nu)\Phi(\boldsymbol{y}) = \mathbf{0}$ whence from the linear independence of $\Phi(\boldsymbol{x})$ and $\Phi(\boldsymbol{x})$ follows $\lambda = \mu = \nu$.

9.3 Projective completion

1. For every **complex vector space** X with **finite** $\dim X = n \geq 1$ and every vector subspace $X_A \subset X$ with $\dim X_A = n-1$ there is an **affine space** $(A; X_A; \to)$ and a **bijection** $\Phi : \mathbb{P}X \setminus \mathbb{P}X_A \to A$ such that for every **projectivity** $\Psi : \mathbb{P}X \to \mathbb{P}X$ with $\Psi[\mathbb{P}X_A] = \mathbb{P}X_A$ the **composition** $\mathbf{g} = \Phi \circ \Psi \circ \Phi^{-1} : A \to A$ is an **affinity**.
2. Conversely for every **affine** space $(A; X_A; \to)$ over a **complex vector space** Y with $X_A \subsetneq Y$ there is a **vector subspace** $X_A \subset X \subset Y$ and a **bijection** $\Phi : \mathbb{P}X \setminus \mathbb{P}X_A \to A$ such that for every **affinity** $g : A \to A$ the **composition** $\Psi = \Phi^{-1} \circ g \circ \Phi : \mathbb{P}X \setminus \mathbb{P}X_A \to \mathbb{P}X \setminus \mathbb{P}X_A$ is a **projectivity** .

The vector subspace X_A is the **infinitely distant hyperplane** and $\mathbb{P}X$ is the **projective completion** of the affine space A.

Proof:



$\Rightarrow$:We choose any $\boldsymbol{a} \in X \setminus X_A$ and consider the affine space $A = \boldsymbol{a} + X_A = \{\boldsymbol{a} + \boldsymbol{x}_A : \boldsymbol{x}_A \in X_A\}$. According to the **Steinitz basis exchange lemma** 3.5 there are bases $\mathcal{B}_A \subset \mathcal{B}$ of $X_A \subset X$ with $\mathcal{B} = \{\boldsymbol{a}\} \cup \mathcal{B}_A$ such that every $\boldsymbol{y} \in X \setminus X_A$ there is are uniquely determined $\mathbf{y}_A \in X_A$ resp. $\lambda \in K$ with $\boldsymbol{a} + \boldsymbol{y}_A = \lambda\boldsymbol{y}$. Hence the map $\Phi : \mathbb{P}X \setminus \mathbb{P}X_A \to A$ with $\Phi(K \cdot \boldsymbol{y}) = \boldsymbol{a} + \boldsymbol{y}_A$ is well defined. Furthermore for every **projectivity** $\Psi : \mathbb{P}X \to \mathbb{P}X$ with $\Psi[\mathbb{P}X_A] = \mathbb{P}X_A$ there is an **automorphism** $\psi : X \to X$ with $K \cdot \psi(\boldsymbol{y}) = \Psi(K \cdot \boldsymbol{y})$ and $\psi[X_A] = X_A$ and due to 9.2 by inserting a suitable factor $c \in K$ we can attain that $\psi(\boldsymbol{a}) \in X_A$. For every $\boldsymbol{y} \in X \setminus X_A$ follows that $(\psi(\boldsymbol{y}))_A - \psi(\boldsymbol{y}_A) = \tau \cdot \psi(\boldsymbol{y}) - \boldsymbol{a} - \psi(\lambda \cdot \boldsymbol{y}) + \psi(\boldsymbol{a}) = (\tau - \lambda)\psi(\boldsymbol{y}) + \psi(\boldsymbol{a}) - \boldsymbol{a} \in X_A$ which implies $(\tau - \lambda)\psi(\boldsymbol{y}) \in X_A$ whence $\tau = \lambda$ since $\psi(\boldsymbol{y}) \in X \setminus X_A$. Hence we have shown that $(\psi(\boldsymbol{y}))_A - \psi(\boldsymbol{y}_A) = \psi(\boldsymbol{a}) - \boldsymbol{a}$ independently of $\boldsymbol{y}$. The affine character of $\boldsymbol{g} = \Phi \circ \Psi \circ \Phi^{-1} : A \to A$ then follows by

$$\begin{aligned}
\psi\left(\overrightarrow{\boldsymbol{a}+\boldsymbol{y}_A; \boldsymbol{a}+\boldsymbol{z}_A}\right) &= \psi(\boldsymbol{z}_A + \boldsymbol{a} - \boldsymbol{y}_A - \boldsymbol{a}) \\
&= \psi(\boldsymbol{z}_A + \boldsymbol{a}) - \psi(\boldsymbol{y}_A + \boldsymbol{a}) \\
&= \psi(\boldsymbol{z}_A) - \psi(\boldsymbol{y}_A) \\
&= (\psi(\boldsymbol{z}))_A - (\psi(\boldsymbol{y}))_A \\
&= \overrightarrow{\boldsymbol{a} + (\psi(\boldsymbol{y}))_A; \boldsymbol{a} + (\psi(\boldsymbol{z}))_A} \\
&= \overrightarrow{\Phi(K \cdot \psi(\boldsymbol{y})); \Phi(K \cdot \psi(\boldsymbol{z}))} \\
&= \overrightarrow{(\Phi \circ \Psi)(K \cdot \boldsymbol{y}); (\Phi \circ \Psi)(K \cdot \boldsymbol{z})} \\
&= \overrightarrow{\mathbf{g}(\boldsymbol{a}+\boldsymbol{y}_A); \mathbf{g}(\boldsymbol{a}+\boldsymbol{z}_A)}.
\end{aligned}$$

$\Leftarrow$: As in the first part we choose an $\boldsymbol{a} \in Y \setminus X_A$ and consider the vector space $X = \operatorname{span}(\mathcal{B})$ with $\mathcal{B} = \{\boldsymbol{a}\} \cup \mathcal{B}_A$ and a basis $\mathcal{B}_A$ for $X_A = \operatorname{span}(\mathcal{B}_A)$. We define $\Phi : \mathbb{P}X \setminus \mathbb{P}X_A \to A$ with

$\Phi\left(K\cdot\mathbf{y}\right)=\boldsymbol{a}+\boldsymbol{y}_A$ and consider an **affinity** $\mathbf{g}:\boldsymbol{a}+X_A\to\boldsymbol{a}+X_A$ with a **linear injective** $\varphi:X_A\to X_A$ such that for any $\boldsymbol{y}_A;\boldsymbol{z}_A\in X_A$ holds $\varphi\left(\mathbf{z}_A-\boldsymbol{y}_A\right)=\varphi\left(\overrightarrow{\boldsymbol{a}+\boldsymbol{y}_A;\boldsymbol{a}+\boldsymbol{z}_A}\right)=\overrightarrow{\boldsymbol{g}\left(\boldsymbol{a}+\boldsymbol{y}_A\right)\boldsymbol{g}\left(\boldsymbol{a}+\boldsymbol{z}_A\right)}=\boldsymbol{g}\left(\boldsymbol{a}+\boldsymbol{z}_A\right)-\boldsymbol{g}\left(\boldsymbol{a}+\boldsymbol{y}_A\right)$. We define $\psi:X\to X$ by $\psi\left(\boldsymbol{y}\right)=\boldsymbol{g}\left(\boldsymbol{a}\right)+\varphi\left(\boldsymbol{y}_A\right)$ with the uniquely determined $\boldsymbol{y}_A=\tau\boldsymbol{y}-\boldsymbol{a}$ for $\boldsymbol{y}\in X\setminus X_A$ whence

$$\begin{aligned}\Psi\left(K\cdot\boldsymbol{y}\right)&=\left(\Phi^{-1}\circ\boldsymbol{g}\circ\Phi\right)\left(K\cdot\boldsymbol{y}\right)\\&=\left(\Phi^{-1}\circ\boldsymbol{g}\right)\left(\boldsymbol{a}+\boldsymbol{y}_A\right)\\&=\Phi^{-1}\left(\boldsymbol{g}\left(\boldsymbol{a}+\mathbf{0}\right)+\varphi\left(\boldsymbol{y}_A-\mathbf{0}\right)\right)\\&=K\cdot\left(\boldsymbol{g}\left(\boldsymbol{a}\right)+\varphi\left(\boldsymbol{y}_A\right)\right)\\&=K\cdot\psi\left(\boldsymbol{y}\right)\end{aligned}$$

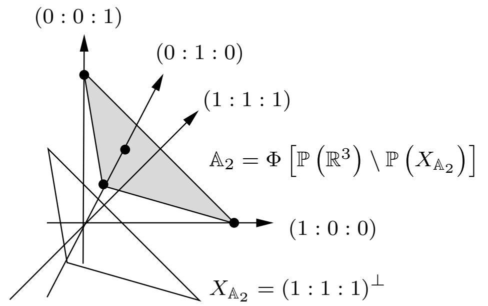
Hence $\Psi=\Phi^{-1}\circ\boldsymbol{g}\circ\Phi:\mathbb{P}X\setminus\mathbb{P}X_A\to\mathbb{P}X\setminus\mathbb{P}X_A$ is a **projectivity** and due to the first part of the proof we conclude that $\mathbb{P}X$ is the projective completion of A.

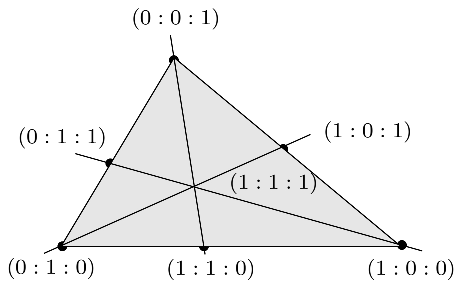
Example:

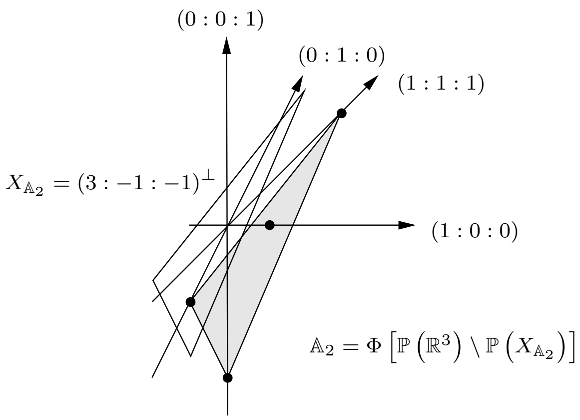
If we exclude the hyperplane $\mathbb{P}E_n=\{\overline{x}_{\mathcal{B}n}=0\}$ the remaining set $\mathbb{P}X\setminus\mathbb{P}E_n=\{\overline{x}_{\mathcal{B}n}\neq 0\}$ can be identified with the **affine space** $\boldsymbol{e}_n+E_n$ resp. the **vector space** E_n by the bijection $\Phi:\mathbb{P}X\setminus\mathbb{P}E_n\to\boldsymbol{e}_n+E_n$ with $\Phi\left(\overline{x}_{\mathcal{B}1}:...:\overline{x}_{\mathcal{B}n}\right)=\left(\frac{x_{\mathcal{B}1}}{x_{\mathcal{B}n}};...;\frac{x_{\mathcal{B}n-1}}{x_{\mathcal{B}n}};1\right)$. Geometrically speaking every line $\overline{\boldsymbol{x}}=\{t\boldsymbol{x}:t\in K\}$ in X except those parallel to the **infinitely distant plane** E_n will meet the affine plane $\boldsymbol{e}_n+E_n$ at a point $\Phi\left(\overline{\boldsymbol{x}}\right)$. Hence the theorem provides the mathematical basis for the projection of three-dimensional objects onto a twodimensional screen as explained by **Albrecht Dürer** in his **Underweysung mit dem Zirckel und Richtscheyt** from 1525 as shown below. As already mentioned in 9.1 **projective completion is not an affine space any more** but the **quotient space** obtained by **gluing** the two components together according to [6, th. 4.9] resp. [2, ex. 6.2.4] is homeomorph to a **closed manifold**.

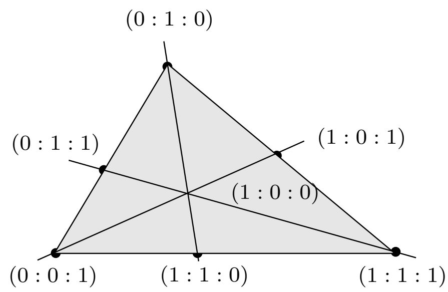
9.4 Projective coordinates

The elements $(\mathbb{C}\cdot\boldsymbol{x}_i)_{1\le i\le n}\subset\mathbb{P}X$ are **projectively independent** iff the $(\boldsymbol{x}_i)_{1\le i\le n}\subset X$ are **linearly independent**. The family $\mathcal{B}=(\mathbb{C}\cdot\boldsymbol{v}_i)_{1\le i\le n+2}\subset\mathbb{P}X$ with $\dim\mathbb{P}X=\dim X-1=n$ is a **projective basis** iff any subfamily of $n+1$ directions is projectively independent. A **projective coordinate** system is a **projectivity** $\kappa:\mathbb{P}\mathbb{C}^{n+1}\to\mathbb{P}X$ with **homogenous coordinates** $(x_1:...:x_{n+1}):=\mathbb{C}\cdot\sum\limits_{i=1}^{n+1}x_i\boldsymbol{e}_i$ for the **canonical basis** $(\boldsymbol{e}_i)_{1\le i\le n+1}\subset\mathbb{C}^{n+1}$. Usually we define $\kappa(x_1:...:x_{n+1}):=\mathbb{C}\cdot\sum\limits_{i=1}^{n+1}x_i\boldsymbol{v}_i$. The additional direction usually is defined by $\mathbb{C}\cdot\boldsymbol{v}_{n+2}=\mathbb{C}\cdot\sum\limits_{i=1}^{n+1}\boldsymbol{v}_i$ and describes the **orientation of the infinitely distant plane X_A with respect to the coordinate axes** in the representation of the projective space $\mathbb{P}X$ as an affine space $A=\Phi[\mathbb{P}X\setminus\mathbb{P}X_A]$ e.g. in the following two representations of $\mathbb{A}_2=\Phi\left[\mathbb{P}\mathbb{R}^3\setminus\mathbb{P}X_{\mathbb{A}_2}\right]$:









Example:

For a vector $\boldsymbol{a}\in\mathbb{C}^3\setminus\mathbb{C}\cdot(1;1;1)$ the projectivity $\Phi:\mathbb{P}\mathbb{C}^3\to\mathbb{P}\mathbb{C}^3$ defined by $\Phi(x_1:x_2:x_3)=(a_1x_1:a_2x_2:a_3x_3)$ with a corresponding linear $\varphi:\mathbb{C}^3\to\mathbb{C}^3$defined by $\varphi(x_1;x_2;x_3)=(a_1x_1;a_2x_2;a_3x_3)$ has three **fixed points** resp. directions along the **basis vectors**

$\Phi(1:0:0)=(a_1:0:0)=(1:0:0)$

$\Phi(0:1:0)=(0:a_2:0)=(0:1:0)$

$\Phi(0:0:1)=(0:0:a_3)=(0:0:1)$ but

$\Phi(1:1:0)=(a_1:a_2:0)\neq(1:1:0)$.

Index

References

[1] John B. Conway. *Functions of one complex variable*. 2nd ed. Springer, 1978.

[2] Arne Vorwerg. “Analysis”. In: (2023). URL: https://www.vorwerg-net.de/Mathematik/Analysis.pdf.

[3] Arne Vorwerg. “Functional Analysis”. In: (2023). URL: https://www.vorwerg-net.de/Mathematik/Functional%20Analysis.pdf.

[4] Arne Vorwerg. “Measure Theory”. In: (2023). URL: https://www.vorwerg-net.de/Mathematik/Measure%20Theory.pdf.

[5] Arne Vorwerg. “Mengenlehre”. In: (2022). URL: https://www.vorwerg-net.de/Mathematik/Mengenlehre.pdf.

[6] Arne Vorwerg. “Topology”. In: (2020). URL: https://www.vorwerg-net/Mathematik/Topology.pdf.