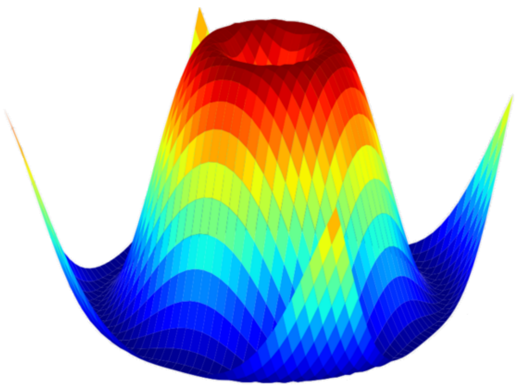
Skript zur Vorlesung

Lineare Algebra I

Private Mitschrift



Vektorraumtheorie

gelesen von

Prof. Dr. Alexander Prestel

Martin Gubisch

Konstanz, Wintersemester 2005/2006

Inhaltsverzeichnis

1. Vektorräume

In diesem Kapitel führen wir das Konzept des endlich-dimensionalen Vektorraumes ein. Dies sind Strukturen, in denen man Elemente addieren und skalieren, d.h. Vektoren aneinandersetzen und um gewisse Faktoren strecken kann. Die ersten vier Abschnitte konzentrieren sich auf die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$; hier lassen sich geometrische Größen wie Winkel, Längen und Abstände definieren, Geraden und Ebenen lassen sich als Untervektorräume auffassen. Die zentralen Begriffe Basis, Dimension und lineare Abhängigkeit, die in den Abschnitten fünf bis acht eingeführt werden, machen dagegen in beliebigen Vektorräumen Sinn. Sie dienen dazu, Vektorräume durch eine endliche Teilmenge zu repräsentieren, aus denen sich der gesamte Raum durch endliche Linearkombinationen zurückgewinnen lassen. In den Abschnitten neun und zehn werden Vektorraummethoden zum Lösen linearer Gleichungssysteme eingesetzt; die Menge aller Lösungen solcher Systeme kann als Untervektorraum aufgefasst werden.

1.1. Der n-dimensionale, reelle Raum

Wir versehen die Menge der reellen n-Tupel $\mathbb{R}^n$ mit einer Addition und einer sklalaren Multiplikation, die es erlauben, Vektoren aneinanderzufügen und um einen Faktor zu strecken. Geometrische Größen können dann analytisch bestimmt werden, was es erlaubt, zentrale Sätze der Euklidischen Geometrie wie die Parallelogrammregel, den Satz des Pythagoras oder den Satz des Thales analytisch zu beweisen.

Bemerkung 1.1.

Wir führen die folgenden Konventionen ein:

1. $\mathbb{R}$ bezeichnet die Menge der reellen Zahlen.
2. Operationen auf $\mathbb{R}$, d.h. Abbildungen $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, sind $+$ und $\cdot$.
3. $\mathbb{R}^n$ bezeichnet die Menge der reellen n-Tupel $(a_1, ..., a_n)$, also $\{(a_1, ..., a_n) \mid a_1, ..., a_n \in \mathbb{R}\}$.
4. Charakteristische Eigenschaft eines n-Tupels: $(a_1, ..., a_n) = (b_1, ..., b_n) :\Leftrightarrow a_1 = b_1, \; ..., \; a_n = b_n$.
5. Im $\mathbb{R}^n$ heißen $a_1, ..., a_n$ die Komponenten oder Koordinaten des Punktes $P = (a_1, ..., a_n)$.
6. Seien $x = (x_1, ..., x_n) \in \mathbb{R}^n, \; y = (y_1, ..., y_n) \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Wir definieren:
   a) $x + y = (x_1 + y_1, ..., x_n + y_n)$. Die Operation $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ heißt Vektoraddition auf $\mathbb{R}^n$.
   b) $0 = (0, ..., 0)$.
   c) $-x = (-x_1, ..., -x_n)$.
   d) $\alpha \cdot x = (\alpha x_1, ..., \alpha x_n)$ Die Abbildung $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ heißt skalare Multiplikation auf $\mathbb{R}^n$. $\Diamond$

Definition 1.2.

Seien G eine Menge und $+$ eine Operation auf G, dann heißt $(G, +)$ eine Gruppe, falls

1. $(x + y) + z = x + (y + z)$ für alle $x, y, z \in G$ (Assoziativität).
2. Es gibt ein $0 \in G$ mit $x + 0 = x = 0 + x$ für alle $x \in G$. 0 heißt neutrales Element.
3. Zu jedem $x \in G$ gibt es ein $-x \in G$ mit $x + (-x) = 0 = (-x) + x$. $-x$ heißt das zu x inverse Element.

Falls zusätzlich gilt

4. $x + y = y + x$ für alle $x, y \in G$ (Kommutativität),

dann heißt $(G, +)$ eine abelsche Gruppe.

Ist $\cdot$ eine weitere Operation auf G, so dass $G^\times = G \backslash \{0\}$ eine abelsche Gruppe ist (wir schreiben 1 für das Neutrale und x^{-1} für das Inverse zu x) und gilt

5. $x \cdot (y + z) = x \cdot y + x \cdot z$ und $(x + y) \cdot z = x \cdot z + y \cdot z$ für alle $x, y, z \in \mathbb{R}$ (Distributivität),

dann heißt $(G, +, \cdot)$ ein Körper.

Bemerkung 1.3.

1. Die reellen Zahlen $\mathbb{R}$ bilden einen Körper, ebenso die rationalen Zahlen $\mathbb{Q}$.
2. $(\mathbb{R}^n, +)$ ist eine abelsche Gruppe: Seien $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n)$ und $z = (z_1, ..., z_n)$ Elemente aus $\mathbb{R}^n$, dann gelten

$$\begin{aligned}(x+y)+z &= (x_1+y_1, ..., x_n+y_n) + (z_1, ..., z_n)\\ &= ((x_1+y_1)+z_1, ..., (x_n+y_n)+z_n)\\ &= (x_1+(y_1+z_1), ..., x_n+(y_n+z_n))\\ &= (x_1, ..., x_n) + (y_1+z_1, ..., y_n+z_n)\\ &= x+(y+z);\end{aligned}$$

$$\begin{aligned}x+y &= (x_1+y_1, ..., x_n+y_n)\\ &= (y_1+x_1, ..., y_n+x_n)\\ &= y+x;\end{aligned}$$

$$\begin{aligned}x+0 &= (x_1+0, ..., x_n+0)\\ &= (x_1, ..., x_n)\\ &= (0+x_1, ..., 0+x_n)\\ &= 0+x;\end{aligned}$$

$$\begin{aligned}x+-x &= (x_1+(-x_1), ..., x_n+(-x_n))\\ &= (0, ..., 0)\\ &= ((-x_1)+x_1, ..., (-x_n)+x_n)\\ &= -x+x.\end{aligned}$$

$\Diamond$

Definition 1.4.

Sind $(V, +)$ eine abelsche Gruppe und $\cdot : \mathbb{R} \times V \to V$ eine Abbildung mit

1. $(\alpha+\beta)x = \alpha x + \beta x$ für alle $\alpha, \beta \in \mathbb{R}$ und alle $x \in V$,
2. $\alpha(x+y) = \alpha x + \alpha y$ für alle $\alpha \in \mathbb{R}$ und alle $x, y \in V$,
3. $(\alpha\beta)x = \alpha(\beta x)$ für alle $\alpha, \beta \in \mathbb{R}$ und alle $x \in V$ und
4. $1x = x$ für alle $x \in V$,

dann heißt $(V, +, \cdot)$ ein $\mathbb{R}$-Vektorraum.

Bemerkung 1.5.

$(V, +, \cdot)$ ist ein $\mathbb{R}$-Vektorraum: Seien $\alpha, \beta \in \mathbb{R}$, dann gelten:

$$\begin{aligned}(\alpha+\beta)x &= ((\alpha+\beta)x_1, ..., (\alpha+\beta)x_n)\\ &= (\alpha x_1+\beta x_1, ..., \alpha x_n+\beta x_n)\\ &= (\alpha x_1, ..., \alpha x_n) + (\beta x_1, ..., \beta x_n)\\ &= \alpha(x_1, ..., x_n) + \beta(x_1, ..., x_n)\\ &= \alpha x+\beta x;\end{aligned}$$

$$\begin{aligned}\alpha(x+y) &= \alpha(x_1+y_1, ..., x_n+y_n)\\ &= (\alpha(x_1+y_1), ..., \alpha(x_n+y_n))\\ &= (\alpha x_1+\alpha y_1, ..., \alpha x_n+\alpha y_n)\\ &= (\alpha x_1, ..., \alpha x_n) + (\alpha y_1, ..., \alpha y_n)\\ &= \alpha x+\alpha y;\end{aligned}$$

$$\begin{aligned}(\alpha\beta)x &= ((\alpha\beta)x_1, ..., (\alpha\beta)x_n)\\ &= (\alpha(\beta x_1), ..., \alpha(\beta x_n))\\ &= \alpha(\beta x_1, ..., \beta x_n)\\ &= \alpha(\beta x);\end{aligned}$$

$$\begin{aligned}1x &= (1x_1, ..., 1x_n)\\ &= (x_1, ..., x_n)\\ &= x.\end{aligned}$$

◊

Bemerkung 1.6.

Addition und Multiplikation im $\mathbb{R}^n$ entsprechen geometrisch dem Aneinandersetzen und Strecken (Skalieren) von Vektoren. Wir werden im Folgenden weitere geometrische Objekte analytisch einführen: Längen, Abstände und Winkel. Allgemein bekannte geometrische Konzepte lassen sich so durch direktes Nachrechnen beweisen; dieses Vorgehen bezeichnet man als die analytische Geometrie. Beispiele sind:

1. Satz des Pythagoras: In einem rechtwinklingen Dreieck ist das Quadrat der längsten Seite gleich der Summe der Quadrate der anderen beiden Seiten.
2. Satz des Thales: Spannt man über einer Sehne eines Kreises zu einem beliebigen Punkt auf dem Kreis ein Dreieck auf, so beträgt der in diesem Punkt eingeschlossene Winkel stets 90°.
3. Parallelogrammgleichung: Die Summe der Längen der beiden Diagonalen in einem Parallelogramm ist stets identisch mit dem Doppelten der Summe der beiden Seitenlängen, d.h. der das Parallelogramm aufspannenden Vektoren. ◊

Definition 1.7.

Seien $(V, +, \cdot)$ ein $\mathbb{R}$-Vektorraum und $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ eine positiv definite, symmetrische Bilinearform, d.h. eine Abbildung mit

1. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$ und $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$, $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$ für alle $x, y \in V$ und alle $\lambda \in \mathbb{R}$ (Bilinearität);
2. $\langle x, y\rangle = \langle y, x\rangle$ für alle $x, y \in V$ (Symmetrie);
3. $\langle x, x\rangle \geq 0$ (Positivität) und
4. $\langle x, x\rangle = 0 \Rightarrow x = 0$ (Definitheit).

Dann heißt $\langle\cdot,\cdot\rangle$ ein euklidisches Skalarprodukt auf V und $(V, \langle\cdot,\cdot\rangle)$ heißt ein euklidischer Raum.

Bemerkung 1.8.

Zu $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$ definieren wir

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + ... + x_n y_n.$$

Dann definiert $\langle\cdot,\cdot\rangle$ ein euklidisches Skalarprodukt auf $\mathbb{R}^n$:

$$\begin{aligned}\langle x+y, z\rangle &= (x_1+y_1)z_1 + \cdots + (x_n+y_n)z_n\\ &= x_1z_1 + y_1z_1 + \cdots + x_nz_n + y_nz_n\\ &= (x_1z_1 + \cdots + x_nz_n) + (y_1z_1 + \cdots + y_nz_n)\\ &= \langle x, z\rangle + \langle y, z\rangle;\end{aligned}$$

$$\begin{aligned}\langle \lambda x, y\rangle &= (\lambda x_1)y_1 + \cdots + (\lambda x_n)y_n\\ &= \lambda(x_1y_1) + \cdots + \lambda(x_ny_n)\\ &= \lambda(x_1y_1 + \cdots + x_ny_n)\\ &= \lambda\langle x, y\rangle;\end{aligned}$$

$$\begin{aligned}\langle x,y\rangle &= x_1y_1+\dots+x_ny_n\\ &= y_1x_1+\dots+y_nx_n\\ &= \langle y,x\rangle;\end{aligned}$$

$$\langle x,x\rangle = x_1^2+\dots+x_n^2 \geq 0;$$

$$\begin{aligned}\langle x,x\rangle = 0 &\Rightarrow x_1^2+\dots+x_n^2 = 0\\ &\Rightarrow x_1^2=0,\ \dots,\ x_n^2=0\\ &\Rightarrow x_1=0,\ \dots,\ x_n=0\\ &\Rightarrow x=0.\end{aligned}$$

Die Linearität in der zweiten Komponenten folgt mit der Symmetrie aus der Linearität in der ersten Komponenten. $\Diamond$

Definition 1.9.
Sei $(V,+,\cdot)$ ein $\mathbb{R}$-Vektorraum. Eine Abbildung $||\cdot||: V\to\mathbb{R}$ heißt eine Norm auf V, falls gelten:

1. $||x||\geq 0$ (Positivität) und $||x||=0\Rightarrow x=0$ für alle $x\in V$ (Definitheit),
2. $||\alpha x|| = |\alpha|\cdot||x||$ für alle $\alpha\in\mathbb{R},\ x\in V$ (Multiplikativität) und
3. $||x+y||\leq||x||+||y||$ für alle $x,y\in V$ (Dreiecksungleichung).

$||x||$ heißt die Länge von x.

$(V,||\cdot||)$ heißt ein normierter $\mathbb{R}$ Vektorraum.

Bemerkung 1.10.
Seien $(V,\langle\cdot,\cdot\rangle)$ ein euklidischer Raum und $x\in V$. Wir definieren

$$||x|| = \sqrt{\langle x,x\rangle}.$$

$||\cdot||$ heißt euklidische Norm auf V oder auch die zu $\langle\cdot,\cdot\rangle$ gehörige Norm.

$||\cdot||$ definiert eine Norm: $||\cdot||$ ist positiv definit, da $\langle\cdot,\cdot\rangle$ es ist. Seien $x,y\in V$ und $\alpha\in\mathbb{R}$, dann

$$\begin{aligned}||\alpha x|| &= \sqrt{\langle\alpha x,\alpha x\rangle}\\ &= \sqrt{\alpha^2\langle x,x\rangle}\\ &= |\alpha|\cdot||x||;\end{aligned}$$

$$\begin{aligned}||x+y||^2 &= ||x||^2+||y||^2+2\langle x,y\rangle\\ &\leq ||x||^2+2||x||\,||y||+||y||^2\\ &= (||x||+||y||)^2\end{aligned}$$

nach Satz 1.11, also $||x+y||\leq||x||+||y||$. $\Diamond$

Satz 1.11.
Seien $(V,\langle\cdot,\cdot\rangle)$ ein euklidischer $\mathbb{R}$-Vektorraum und $x,y\in V$. Dann gelten:

1. $||x+y||^2 = ||x||^2+2\langle x,y\rangle+||y||^2$ (Binomische Formel)
2. $||x-y||^2 = ||x||^2-2\langle x,y\rangle+||y||^2$ (Satz des Pythagoras)
3. $||x+y||^2+||x-y||^2 = 2||x||^2+2||y||^2$ (Parallelogrammregel)
4. $|\langle x,y\rangle|\leq||x||\,||y||$ (Cauchy-Schwarzsche Ungleichung).

Beweis.

1. Seien $x, y \in V$, dann gelten:

$$\begin{aligned} ||x+y||^2 &= \langle x+y, x+y\rangle \\ &= \langle x, x+y\rangle + \langle y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle \\ &= ||x||^2 + 2\langle x,y\rangle + ||y||^2; \end{aligned}$$

$$\begin{aligned} ||x-y||^2 &= \langle x-y, x-y\rangle \\ &= \langle x, x-y\rangle + \langle -y, x-y\rangle \\ &= \langle x,x\rangle + \langle x,-y\rangle + \langle -y,x\rangle + \langle -y,-y\rangle \\ &= ||x||^2 - 2\langle x,y\rangle + ||y||^2; \end{aligned}$$

$$\begin{aligned} ||x+y||^2 + ||x-y||^2 &= ||x||^2 + 2\langle x,y\rangle + ||y||^2 + ||x||^2 - 2\langle x,y\rangle + ||y||^2 \\ &= 2||x||^2 + 2||y||^2. \end{aligned}$$

2. Für alle $\lambda, \mu \in \mathbb{R}$ gilt:

$$0 \leq ||\lambda x + \mu y||^2 = \langle \lambda x + \mu y, \lambda x + \mu x\rangle = \lambda^2\langle x,x\rangle + 2\lambda\mu\langle x,y\rangle + \mu^2\langle y,y\rangle.$$

Setze $\lambda = \langle y,y\rangle$ und $\mu = -\langle x,y\rangle$; dann gilt:

$$0 \leq \langle y,y\rangle\langle y,y\rangle\langle x,x\rangle - 2\langle y,y\rangle\langle x,y\rangle\langle x,y\rangle + \langle x,y\rangle\langle x,y\rangle\langle y,y\rangle,$$

also

$$0 \leq \langle y,y\rangle(\langle x,x\rangle\langle y,y\rangle - \langle x,y\rangle\langle x,y\rangle).$$

Für $\langle y,y\rangle = 0$, d.h. $y = 0$, ist die Ungleichung erfüllt. Sei nun also $y \neq 0$. Dann $\langle y,y\rangle > 0$, also

$$0 \leq \langle x,x\rangle\langle y,y\rangle - \langle x,y\rangle^2 \qquad \Longrightarrow \qquad \langle x,y\rangle^2 \leq \langle x,x\rangle\langle y,y\rangle,$$

d.h. $|\langle x,y\rangle| \leq ||x||\,||y||$. □

Definition 1.12.

Seien $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Raum und $x, y \in V$. Wir sagen, x, y sind orthogonal zueinander oder stehen senkrecht aufeinander (in Zeichen: $x \perp y$), falls $\langle x,y\rangle = 0$.

Bemerkung 1.13.

1. Seien $x, y \in V$, dann gilt:

$$x \perp y \qquad \Longleftrightarrow \qquad ||x-y|| = ||x+y||,$$

denn

$$\begin{aligned} ||x+y|| = ||x-y|| &\Leftrightarrow ||x+y||^2 = ||x-y||^2 \\ &\Leftrightarrow ||x||^2 + 2\langle x,y\rangle + ||y||^2 = ||x||^2\langle x,y\rangle + ||y||^2 \\ &\Leftrightarrow 4\langle x,y\rangle = 0 \\ &\Leftrightarrow x \perp y. \end{aligned}$$

2. Der Nullvektor steht orthogonal auf jedem $x \in V$:

$$\langle x, 0\rangle = 0 \cdot \langle x, 0\rangle = 0.$$

◊

Definition 1.14.
Sei X eine nicht-leere Menge. Eine Abbildung $\mathrm{d}(\cdot,\cdot) : X \times X \to \mathbb{R}$ heißt eine Metrik, falls gelten:

1. $\mathrm{d}(x,y) = 0 \Leftrightarrow x = y$ für alle $x, y \in V$.
2. $\mathrm{d}(x,y) = \mathrm{d}(y,x)$ für alle $x, y \in V$. (Symmetrie)
3. $\mathrm{d}(x,y) \leq \mathrm{d}(x,z) + \mathrm{d}(z,y)$ für alle $x, y, z \in V$ (Dreiecksungleichung).

$\mathrm{d}(x,y)$ heißt der Abstand zwischen x und y.

d heißt eine Metrik und das Paar (X, d) heißt metrischer Raum.

Bemerkung 1.15.
Sei $(V, ||\cdot||)$ ein normierter $\mathbb{R}$-Vektorraum. Zu $x, y \in V$ definiere

$$\mathrm{d}(x,y) = ||x-y||.$$

d heißt euklidische Metrik oder zu $||\cdot||$ gehörige Metrik.

Dann ist (V, d) ein metrischer Raum: 1. folgt aus der Definitheit von $||\cdot||$. Seien $x, y, z \in V$, dann gelten:

$$\begin{aligned}
\mathrm{d}(x,y) &= ||x-y|| \\
&= |-1| \cdot ||y-x|| \\
&= \mathrm{d}(y,x);
\end{aligned}$$

$$\begin{aligned}
\mathrm{d}(x,y) &= ||x-y|| \\
&= ||(x-z)+(z-y)|| \\
&\leq ||x-z|| + ||z-y|| \\
&= \mathrm{d}(x,z) + \mathrm{d}(z,y).
\end{aligned}$$

$\Diamond$

Definition 1.16.
Seien $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Raum und $x, y \in V\backslash\{0\}$.

$$\sphericalangle(x,y) = \cos^{-1}\left(\frac{\langle x,y\rangle}{||x||\,||y||}\right).$$

$\sphericalangle(x,y)$ heißt der Winkel zwischen x und y.

Bemerkung 1.17.
Im $\mathbb{R}^2$ stimmt diese Winkeldefinition mit der „Üblichen“ überein:

1. Wegen $\sphericalangle(x,y) = \cos^{-1}\langle \frac{x}{||x||}, \frac{y}{||y||}\rangle$ hängt der Winkel zwischen x und y nicht von der Länge von x bzw. y ab. Wir können also Œ $||x|| = ||y|| = 1$ annehmen, d.h. $\sphericalangle(x,y) = \langle x,y\rangle$.
2. Seien $\alpha = \sphericalangle(x,(1,0))$ und $\beta = \sphericalangle(y,(1,0))$, dann $x = (\cos(\alpha), \sin(\alpha))$ und $y = (\cos(\beta), \sin(\beta))$. Nach Additionstheorem gilt

$$\langle x,y\rangle = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \cos(\alpha-\beta) = \sphericalangle(x,y).$$

$\Diamond$

Definition 1.18.
Zwei Vektoren x und y aus $\mathbb{R}^n$ heißen linear abhängig, wenn es ein $\alpha \in \mathbb{R}$ gibt mit $x = \alpha y$ oder $y = \alpha x$.

Existiert kein solches α, so heißen die Vektoren linear unabhängig.

Bemerkung 1.19.
Beschränkt man sich nur auf eine der beiden Gleichungen, z.B. auf $x = \alpha y$, so stößt man beim Nullvektor auf einen Widerspruch:

Für $x = 0$ und $y \neq 0$ findet man kein $\alpha \in \mathbb{R}$ mit $\alpha y = x$, d.h. 0 wäre linear unabhängig von allen anderen Vektoren. Gleichzeitig gilt mit $y = 0$ und $x \neq 0$ für $\alpha \in \mathbb{R}$ beliebig die Beziehung $\alpha x = y$, d.h. 0 wäre linear abhängig von allen anderen Vektoren, ein Widerspruch.

Nach der „richtigen“ Definition ist der Nullvektor tatsächlich linear abhängig von allen Vektoren. $\Diamond$

Korollar 1.20. (zur Cauchy-Schwarzschen Ungleichung)
Für alle $x, y \in \mathbb{R}^n$ gilt: $|\langle x, y\rangle| = ||x||\,||y|| \Leftrightarrow x, y$ sind linear abhängig.

Beweis. Wir dürfen Œ annehmen, dass $y \neq 0$ gilt (andernfalls ist die Behauptung trivial).

1. Gelte $\langle x, y\rangle^2 = ||x||^2||y||^2$. Setze $\lambda = \frac{||x||^2}{||y||^2}$, dann gilt

$$||x - \lambda y||^2 = ||x||^2 - 2\lambda\langle x, y\rangle + \lambda^2||y||^2 = ||x||^2 - 2\lambda||x||\,||y|| + \lambda^2||y||^2 = (||x|| - \lambda||y||)^2 = 0,$$

d.h. $x - \lambda y = 0$ und damit $x = \lambda y$, d.h. x, y sind linear abhängig.

2. Gelte umgekehrt $y = \lambda x$ für ein $\lambda \in \mathbb{R}$ (insbesondere $\lambda \neq 0$), dann ist

$$||x||^2||y||^2 = \lambda^2||x||^4 = \lambda^2\langle x, x\rangle^2 = \lambda^2 \left\langle x, \frac{1}{\lambda}y \right\rangle^2 = \frac{\lambda^2}{\lambda^2}\langle x, y\rangle^2,$$

also $||x||\,||y|| = |\langle x, y\rangle|$. $\square$

1.2. Die komplexen Zahlen

Wir führen die komplexen Zahlen ein, um einen Zahlbereich zur Verfügung zu haben, in dem quadratische Gleichungen und auch polynomiale Gleichungen höherer Ordnung stets auflösbar sind – die Gleichung $x^2 + 1 = 0$ besitzt bekanntlich keine Lösung x in $\mathbb{R}$.

Bemerkung 1.21.
In $\mathbb{R}$ besitzt jede polynomiale Gleichung

$$x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0 = 0$$

mit $n \in \mathbb{N}$ ungerade und Koeffizienten $\alpha_1, ..., \alpha_{n-1} \in \mathbb{R}$ eine Lösung.

Die Gleichung

$$x^2 + 1 = 0$$

hat dagegen keine Lösung in $\mathbb{R}$. Wir suchen also einen Oberkörper von $\mathbb{R}$, in dem $x^2 + 1 = 0$ eine Lösung hat. $\Diamond$

Bemerkung 1.22.
Falls dieser Oberkörper existiert, müssen gelten:

1. $i = \sqrt{-1}$ existiert, die imaginäre Einheit.
2. Seien $a, b \in \mathbb{R}$. Ist $a + bi = 0$, dann sind bereits $a = 0$ und $b = 0$:
   $a + bi = 0 \Leftrightarrow -a = bi$. Wäre $b \neq 0$, dann $i = -\frac{a}{b} \in \mathbb{R}$, was nicht sein kann. Also $b = 0$ und damit auch $a = 0$, denn $0i = (0 + 0)i = 0i + 0i \Rightarrow 0 = 0i$.
3. $(a + bi) + (c + di) = (a + c) + (b + d)i$ für alle $a, b, c, d \in \mathbb{R}$: In einem Körper gelten Assoziativ-, Kommutativ- und Distributivgesetz.
4. $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$ für alle $a, b, c, d \in \mathbb{R}$:

$$(a + bi)(c + di) = ac + adi + bic + bidi = ac + (ad + bc)i - bd = (ac - bd) + (ad + bc)i. \quad \Diamond$$

Satz 1.23.
Definieren wir als Multiplikation auf $\mathbb{R}^2$
$$(a,b)(c,d) = (ac - bd, ad + bc),$$
so bildet die abelsche Gruppe $(\mathbb{R}^2, +)$ zusammen mit dieser Multiplikation einen Körper, den Körper der komplexen Zahlen $\mathbb{C}$, d.h. es gelten:

1. $\forall z, z', z'' \in \mathbb{R}^2 : z(z'z'') = (zz')z''$. (Assoziativität)
2. $\forall z, z' \in \mathbb{R}^2 : zz' = z'z$. (Kommutativität)
3. $\exists 1 \in \mathbb{R}^2 : \forall z \in \mathbb{R}^2 : 1z = z$. $1 = (1,0)$ ist das multiplikativ Neutrale.
4. $\forall z \in \mathbb{R}^2 : \exists z^{-1} \in \mathbb{R}^2 : zz^{-1} = 1$. $(a,b)^{-1} = \frac{1}{a^2+b^2}(a,-b)$ ist das multiplikativ Inverse zu $z = (a,b)$.
5. $\forall z, z', z'' \in \mathbb{R}^2 : z(z' + z'') = zz' + zz''$. (Distributivität)

Beweis.
Seien $z = (a,b),\ z' = (a',b')$ und $z'' = (a'',b'')$ Elemente aus $\mathbb{R}^2$. Dann gelten:
$$\begin{aligned}
(a,b)((a',b')(a'',b'')) &= (a,b)(a'a'' - b'b'', a'b'' + b'a'') \\
&= (a(a'a'' - b'b'') - b(a'b'' + b'a''), a(a'b'' + b'a'') - b(a'a'' - b'b'')) \\
&= (aa'a'' - ab'b'' - ba'b'' - bb'a'', aa'b'' + ab'a'' - ba'a'' + bb'b'') \\
&= (aa'a'' - bb'a'' - ab'b'' - ba'b'', aa'b'' + bb'b'' + ab'a'' - ba'a'') \\
&= ((aa' - bb')a'' - (ab' + ba')b'', (aa' + bb')b'' + (ab' - ba')a'') \\
&= (aa' - bb', ab' + ba')(a'', b'') \\
&= ((a,b)(a',b'))(a'',b'');
\end{aligned}$$

$$\begin{aligned}
(a,b)(a',b') &= (aa' - bb', ab' + ba') \\
&= (a'a - b'b, a'b + b'a) \\
&= (a',b')(a,b);
\end{aligned}$$

$$\begin{aligned}
(a,b)(1,0) &= (a1 - b0, a0 + b1) \\
&= (a,b);
\end{aligned}$$

$$\begin{aligned}
(a,b)\frac{1}{a^2+b^2}(a,-b) &= \left(\frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}, \frac{-ab}{a^2+b^2} + \frac{ba}{a^2+b^2}\right) \\
&= \left(\frac{a^2+b^2}{a^2+b^2}, \frac{ab-ab}{a^2+b^2}\right) \\
&= (1,0);
\end{aligned}$$

$$\begin{aligned}
(a,b)((a',b') + (a'',b'')) &= (a,b)(a' + a'', b' + b'') \\
&= (a(a' + a'') - b(b' + b''), a(b' + b'') + b(a' + a'')) \\
&= (aa' + aa'' - bb' + bb'', ab' + ab'' + ba' + ba'') \\
&= (aa' - bb', ab' + ba') + (aa'' - bb'', ab'' + ba'') \\
&= (a,b)(a',b') + (a,b)(a'',b'').
\end{aligned}$$
□

Satz 1.24. (Fundamentalsatz der Algebra)
Jedes Polynom
$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$
mit $a_0, ..., a_{n-1} \in \mathbb{C}$ und $n \geq 1$ hat eine Nullstelle in $\mathbb{C}$, d.h. es gibt ein $x \in \mathbb{C}$ mit $p(x) = 0$.

Definition 1.25.

Seien G und H Mengen und f eine Abbildung (Funktion) von G nach H, d.h. eine Zuordnungsvorschrift, die **jedem** Element der Menge G **genau ein** Element der Menge H zuordnet.

f heißt surjektiv, wenn für jedes $h \in H$ ein $g \in G$ existiert mit $f(g) = h$.

f heißt injektiv, wenn zwei verschiedenen Elementen $g_1, g_2 \in G$ stets zwei verschiedene Elemente h_1, h_2 zugeordnet werden, d.h. wenn für alle $g_1, g_2 \in G$ gilt: $g_1 \neq g_2 \Rightarrow f(g_1) \neq f(g_2)$.

Ist eine Abbildung sowohl surjektiv als auch injektiv, so heißt sie bijektiv.

Bemerkung 1.26.

1. $\iota : \mathbb{R} \to \mathbb{R}^2,\ a \mapsto (a, 0)$ ist eine Einbettung, d.h. eine injektive, mit den Körperoperationen verträgliche Abbildung: Seien $a, b \in \mathbb{R}$, dann gelten:
   a) Ist $\iota(a) = \iota(b)$, d.h. $(a, 0) = (b, 0)$, dann ist bereits $a = b$.
   b) $\iota(a + b) = (a + b, 0) = (a, 0) + (b, 0) = \iota(a) + \iota(b)$.
   c) $\iota(ab) = (ab, 0) = (a, 0)(b, 0) = \iota(a)\iota(b)$.

   Wir können $a \in \mathbb{R}$ also mit $(a, 0) \in \mathbb{R}^2$ identifizieren, d.h. wir setzen $a \in \mathbb{R}^2$ via $a = (a, 0)$.
2. Definieren wir noch $i = (0, 1)$, dann gelten:
   a) $i^2 = -1$, denn $i^2 = (0, 1)(0, 1) = (0 - 1, 0 + 0) = (-1, 0) = -1$.
   b) $ib = (0, b)$ für alle $b \in \mathbb{R}$, denn $ib = (0, 1)(b, 0) = (0 - 0, 0 + b) = (0, b)$.
   c) $a + ib = (a, b)$ für alle $a, b \in \mathbb{R}$, denn $a + ib = (a, 0) + (0, b) = (a + 0, 0 + b) = (a, b)$. $\diamond$

Definition 1.27.

Zu $z \in \mathbb{C},\ z = (a, b) = a + ib$, heißt a der Realteil ($\mathrm{Re}(z)$) von z und b heißt der Imaginärteil ($\mathrm{Im}(z)$) von b.

Ist $z = a + ib$, dann heißt $\overline{z} = a - ib$ die zu z komplex konjugierte Zahl.

Satz 1.28.

Die komplexe Konjugation $\varkappa : \mathbb{C} \to \mathbb{C},\ z \mapsto \kappa(z) = \overline{z}$ ist bijektiv.

Beweis.

1. $\varkappa$ ist idempotent, d.h. $\varkappa(\varkappa(x)) = x$ für alle $x \in \mathbb{C}$: Sei $x = a + ib$, dann ist
$$\varkappa(\varkappa(x)) = \varkappa(\varkappa(a + ib)) = \varkappa(a - ib) = a + ib.$$
2. $\varkappa$ ist injektiv: Gelte $\overline{z} = \overline{z'}$ für $z, z' \in \mathbb{C}$, dann ist
$$z = \overline{\overline{z}} = \overline{\overline{z'}} = z',$$
3. Ist $z \in \mathbb{C}$, dann ist $z = \varkappa(\overline{z})$, d.h. $\varkappa$ ist surjektiv. $\square$

Satz 1.29.

Für $z_1, z_2 \in \mathbb{C}$ gilt weiter:

1. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$.
2. $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.
3. Für $|z| = z\overline{z}$ gilt: $|z| = ||z||^2 = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$.

Beweis.

Seien $z_1 = (a_1, b_1),\ z_2 = (a_2, b_2)$. Dann gelten:

$$\begin{aligned}
\overline{z_1 + z_2} &= \overline{(a_1, b_1) + (a_2, b_2)} \\
&= \overline{(a_1 + a_2, b_1 + b_2)} \\
&= (a_1 + a_2, -(b_1 + b_2)) \\
&= (a_1, -b_1) + (a_2, -b_2) \\
&= \overline{(a_1, b_1)} + \overline{(a_2, b_2)} \\
&= \overline{z_1} + \overline{z_2}
\end{aligned}$$

$$\begin{aligned}
\overline{z_1 z_2} &= \overline{(a_1, b_1)(a_2, b_2)} \\
&= \overline{(a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2)} \\
&= (a_1 a_2 - b_1 b_2, -(a_1 b_2 + b_1 a_2)) \\
&= (a_1 a_2 - b_1 b_2, a_1(-b_2) + (-b_1) a_2) \\
&= (a_1, -b_1)(a_2, -b_2) \\
&= \overline{(a_1, b_1)}\,\overline{(a_2, b_2)} \\
&= \overline{z_1}\,\overline{z_2}
\end{aligned}$$

$$\begin{aligned}
z\overline{z} &= (a, b)(a, -b) \\
&= (aa + bb, -ab + ab) \\
&= (a^2 + b^2, 0) \\
&= a^2 + b^2 \\
&= ||(a, b)||^2 \\
&= ||z||^2.
\end{aligned}$$

□

Bemerkung 1.30.

1. Ist $z \neq 0$, dann besitzt das zu z multiplikativ Inverse $\overline{z}$ die Darstellung

$$\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{|z|^2}.$$

2. $|\cdot|$ ist multiplikativ: Seien $z, z' \in \mathbb{C}$, dann

$$|zz'| = (zz')\overline{zz'} = (z\overline{z})(z'\overline{z'}) = |z|\,|z'|.$$

◇

1.3. Der n-dimensionale, komplexe Raum

Wir versehen nun auch die Menge der komplexen n-Tupel $\mathbb{C}^n$ mit einer Vektorraumstruktur und übertragen die geometrischen Konzepte des $\mathbb{R}^n$ ins Komplexe.

Definition 1.31.

Seien $(V, +)$ eine abelsche Gruppe und $\cdot : \mathbb{C} \times V \to V$ eine Abbildung, so dass für alle $z, z' \in \mathbb{C}^n$ und alle $\alpha, \beta \in \mathbb{C}$ gelten:

$$\begin{aligned}
\alpha(z + z') &= \alpha z + \beta z'; \\
(\alpha + \beta)z &= \alpha z + \beta z; \\
\alpha(\beta z) &= (\alpha\beta)z; \\
1z &= z.
\end{aligned}$$

Dann heißt $(V, +, \cdot)$ ein $\mathbb{C}$-Vektorraum.

Bemerkung 1.32.
Versehen wir die Menge

$$\mathbb{C}^n = \{(z_1, ..., z_n) \mid z_1, ..., z_n \in \mathbb{C}\}$$

mit der Vektoraddition

$$(z_1, ..., z_n) + (z'_1, ..., z'_n) = (z_1 + z'_1, ..., z_n + z'_n)$$

und skalaren Multiplikation

$$\alpha(z_1, ..., z_n) = (\alpha z_1, ..., \alpha z_n),$$

$z = (z_1, ..., z_n),\ z' = (z'_1, ..., z'_n) \in \mathbb{C}^n$ und $\alpha \in \mathbb{C}$.
Dann wird $(\mathbb{C}^n, +, \cdot)$ zu einem $\mathbb{C}$-Vektorraum:
Seien $x, y, z \in \mathbb{C}^n$ und $\alpha, \beta \in \mathbb{C}$, dann

$$\begin{aligned}
(x+y)+z &= (x_1+y_1, ..., x_n+y_n) + (z_1, ..., z_n) \\
&= ((x_1+y_1)+z_1, ..., (x_n+y_n)+z_n) \\
&= (x_1+(y_1+z_1), ..., x_n+(y_n+z_n)) \\
&= (x_1, ..., x_n) + (y_1+z_1, ..., y_n+z_n) \\
&= x+(y+z);
\end{aligned}$$

$$\begin{aligned}
x+y &= (x_1+y_1, ..., x_n+y_n) \\
&= (y_1+x_1, ..., y_n+x_n) \\
&= y+x;
\end{aligned}$$

$$\begin{aligned}
x+0 &= (x_1+0, ..., x_n+0) \\
&= (x_1, ..., x_n) \\
&= (0+x_1, ..., 0+x_n) \\
&= 0+x;
\end{aligned}$$

$$\begin{aligned}
x+-x &= (x_1+(-x_1), ..., x_n+(-x_n)) \\
&= (0, ..., 0) \\
&= ((-x_1)+x_1, ..., (-x_n)+x_n) \\
&= -x+x;
\end{aligned}$$

$$\begin{aligned}
(\alpha+\beta)x &= ((\alpha+\beta)x_1, ..., (\alpha+\beta)x_n) \\
&= (\alpha x_1+\beta x_1, ..., \alpha x_n+\beta x_n) \\
&= (\alpha x_1, ..., \alpha x_n) + (\beta x_1, ..., \beta x_n) \\
&= \alpha(x_1, ..., x_n) + \beta(x_1, ..., x_n) \\
&= \alpha x+\beta x;
\end{aligned}$$

$$\begin{aligned}
\alpha(x+y) &= \alpha(x_1+y_1, ..., x_n+y_n) \\
&= (\alpha(x_1+y_1), ..., \alpha(x_n+y_n)) \\
&= (\alpha x_1+\alpha y_1, ..., \alpha x_n+\alpha y_n) \\
&= (\alpha x_1, ..., \alpha x_n) + (\alpha y_1, ..., \alpha y_n) \\
&= \alpha x+\alpha y;
\end{aligned}$$

$$\begin{aligned}
(\alpha\beta)x &= ((\alpha\beta)x_1, ..., (\alpha\beta)x_n) \\
&= (\alpha(\beta x_1), ..., \alpha(\beta x_n)) \\
&= \alpha(\beta x_1, ..., \beta x_n) \\
&= \alpha(\beta x);
\end{aligned}$$

$$\begin{aligned}1x &= (1x_1, ..., 1x_n)\\ &= (x_1, ..., x_n)\\ &= x.\end{aligned}$$

◊

Definition 1.33.

Seien $(V, +, \cdot)$ ein $\mathbb{C}$-Vektorraum und $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ eine positiv definite, antisymmetrische Sesquilinearform, d.h. eine Abbildung mit

1. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$ und $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$, $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$ für alle $x, y \in V$ und alle $\lambda \in \mathbb{R}$ (Sesquilinearität);
2. $\langle x, y\rangle = \overline{\langle y, x\rangle}$ für alle $x, y \in V$ (Antisymmetrie);
3. $\langle x, x\rangle \geq 0$ (Positivität) und
4. $\langle x, x\rangle = 0 \Rightarrow x = 0$ (Definitheit).

Dann heißt $\langle\cdot,\cdot\rangle$ ein hermitesches Skalarprodukt auf V und $(V, \langle\cdot,\cdot\rangle)$ heißt ein unitärer Raum.

Bemerkung 1.34.

Definiere $\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ durch

$$\langle z, z'\rangle = \sum_{i=1}^{n} z_i\overline{z_i'} = z_1\overline{z_1'} + \cdots + z_n\overline{z_n'}.$$

Dann ist $\langle\cdot,\cdot\rangle$ ein hermitesches Skalarprodukt auf $\mathbb{C}^n$, d.h. $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ ist ein unitärer Raum:

$$\begin{aligned}\langle z, z'\rangle &= z_1\overline{z_1'} + \cdots + z_n\overline{z_n'}\\ &= \overline{z_1'}\,\overline{\overline{z_1}} + \cdots + \overline{z_n'}\,\overline{\overline{z_n}}\\ &= \overline{z_1'\overline{z_1} + \cdots + z_n'\overline{z_n}}\\ &= \overline{\langle z', \overline{z}\rangle};\end{aligned}$$

$$\begin{aligned}\langle z, z' + z''\rangle &= z_1(\overline{z_1' + z_1''}) + \cdots + z_n(\overline{z_n' + z_n''})\\ &= z_1\overline{z_1'} + z_1\overline{z_1''} + \cdots + z_n\overline{z_n'} + z_n\overline{z_n''}\\ &= (z_1\overline{z_1'} + \cdots + z_n\overline{z_n'}) + (z_1\overline{z_1''} + \cdots + z_n\overline{z_n''})\\ &= \langle z, z'\rangle + \langle z, z''\rangle;\end{aligned}$$

$$\begin{aligned}\langle \alpha z, z'\rangle &= \alpha z_1 + \cdots + \alpha z_n\overline{z_n'}\\ &= \alpha(z_1\overline{z_1'} + \cdots + z_n\overline{z_n'})\\ &= \alpha\langle z, z'\rangle;\end{aligned}$$

$$\langle z, z\rangle = |z_1|^2 + \cdots + |z_n|^2 \geq 0;$$

$$\begin{aligned}\langle z, z\rangle = 0 &\Rightarrow |z_1|^2 + \cdots + |z_n|^2 = 0\\ &\Rightarrow |z_1|^2 = 0,\ ...,\ |z_n|^2 = 0\\ &\Rightarrow z_1 = 0,\ ...,\ z_n = 0\\ &\Rightarrow z = 0.\end{aligned}$$

Die Sesquilinearität in der zweiten Komponenten folgt wiederum mit Hilfe der Antisymmetrie aus der Linearität in der ersten Komponenten. ◊

Definition 1.35.
Sei $(V, +, \cdot)$ ein unitärer Raum. Eine Abbildung $||\cdot|| : V \to \mathbb{R}$ heißt eine Norm auf V, falls gelten:

1. $||x|| \geq 0$ (Positivität) und $||x|| = 0 \Rightarrow x = 0$ für alle $x \in V$ (Definitheit),
2. $||\alpha x|| = |\alpha| \cdot ||x||$ für alle $\alpha \in \mathbb{C},\ x \in V$ (Multiplikativität) und
3. $||x + y|| \leq ||x|| + ||y||$ für alle $x, y \in V$ (Dreiecksungleichung).

$||x||$ heißt die Länge von x.
$(V, ||\cdot||)$ heißt ein normierter $\mathbb{C}$-Vektorraum.

Bemerkung 1.36.
Seien $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Raum und $x \in V$.

$$||x|| = \sqrt{\langle x, x\rangle}$$

definiert eine Norm auf V, die unitäre Norm oder auch die zu $\langle\cdot,\cdot\rangle$ gehörige Norm.
Für alle $x, y \in V$ gelten:

1. $||x + y||^2 = ||x||^2 + 2\mathrm{Re}\langle x, y\rangle + ||y||^2$ (Binomische Formel)
2. $||x - y||^2 = ||x||^2 - 2\mathrm{Re}\langle x, y\rangle + ||y||^2$ (Satz des Pythagoras)
3. $||x + y||^2 + ||x - y||^2 = 2||x||^2 + 2||y||^2$ (Parallelogrammregel)
4. $|\langle x, y\rangle| \leq ||x||\,||y||$ (Cauchy-Schwarz-Ungleichung)
5. $|\langle x, y\rangle| = ||x||\,||y|| \Leftrightarrow x = \lambda y$ für ein $\lambda \in \mathbb{C}$. $\Diamond$

Bemerkung 1.37.
Sei $(V, ||\cdot||)$ ein normierter $\mathbb{C}$-Vektorraum. Dann definiert

$$\mathrm{d}(x, y) = ||x - y||$$

eine Metrid $\mathrm{d} : V \times V \to \mathbb{R}$ auf V, die unitäre Metrik oder zu $||\cdot||$ gehörige Metrik. $\Diamond$

1.4. Geraden und Ebenen im $\mathbb{R}^n$

Wir übertragen in diesem Abschnitt die geometrischen Objekte "Gerade" und "Ebene" der Anschauungsräume $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ auf den n-dimensionalen $\mathbb{R}$-Vektorraum und beweisen den Satz vom Lotpunkt: Die kürzeste Verbindung eines Punktes zu einer Geraden steht stets senkrecht auf dieser.

Definition 1.38.
Seien $x, w \in \mathbb{R}^n$ mit $w \neq 0$. Dann heißt die Menge

$$G = \{z \in \mathbb{R}^n \mid \exists \lambda \in \mathbb{R} : z = x + \lambda w\}$$

eine Gerade im $\mathbb{R}^n$.
Zu $z = x + \lambda w \in G$ heißt λ Parameter des Punktes z.

Satz 1.39.
Durch zwei verschiedene Punkte $x, y \in \mathbb{R}^n$ gibt es genau eine Gerade $G(x, y)$.
Diese besitzt die Darstellung

$$G(x, y) = \{x + \lambda(y - x) \mid \lambda \in \mathbb{R}\}.$$

Beweis.

1. Existenz: Für $\lambda = 0$ gilt $x = x + \lambda(y-x)$, also $x \in G$, und für $\lambda = 1$ ist $y = x + \lambda(y-x)$, also $y \in G$.
2. Eindeutigkeit: Seien $H = \{z \in \mathbb{R}^n \mid \exists \sigma \in \mathbb{R} : z = u + \sigma(v-u)\}$ und $u, v \in \mathbb{R}^n$ mit $u \neq v$ sowie $x, y \in H$. Zu zeigen: $G = H$.

   Seien $x, y \in H$, dann gibt es $\sigma_x, \sigma_y \in \mathbb{R}$ mit $x = u + \sigma_x(v-u)$ und $y = u + \sigma_y(v-u)$.

   a) Zu $\supseteq$: Sei $z \in H$, d.h. $z = u + \sigma(v-u)$ für ein $\sigma \in \mathbb{R}$. Dann
   $$z = x - \sigma_x(v-u) + \sigma(v-u) = x + (\sigma - \sigma_x)(v-u) = x + \frac{\sigma - \sigma_x}{\sigma_y - \sigma_x}(y-x).$$
   Setze
   $$\lambda = \frac{\sigma - \sigma_x}{\sigma_y - \sigma_x},$$
   dann folgt $z \in G$.

   b) Zu $\subseteq$: Sei $z \in G$. Dann $z = x + \lambda(y-x)$ für ein $\lambda \in \mathbb{R}$, also
   $$z = u + \sigma_x(v-u) + \lambda(\sigma_y - \sigma_x)(v-u) = u + (\sigma_x + \lambda(\sigma_y - \sigma_x))(v-u).$$
   Setze
   $$\sigma = (\sigma_x + \lambda(\sigma_y - \sigma_x)),$$
   dann folgt $z \in H$. $\diamond$

Satz 1.40. (Satz vom Lotpunkt)

Seien $x \neq y \in \mathbb{R}^n$ und $u \in \mathbb{R}^n$.

Dann gibt es ein eindeutig bestimmtes $v_0 \in G(x,y)$ mit $(v_0 - u) \perp (y-x)$ und es gilt
$$\mathrm{d}(u, v_0) = \min\{\mathrm{d}(u,v) \mid v \in G(x,y)\}.$$

Beweis.

1. Existenz: Setze
   $$v_0 = x + \lambda_0(y-x), \qquad \lambda_0 = \frac{\langle u-x, y-x\rangle}{||y-x||^2}.$$
   Dann gilt:
   $$\begin{aligned}\langle v_0 - u, y-x\rangle &= \langle x + \lambda_0(y-x) - u, y-x\rangle \\ &= \langle x-u, y-x\rangle + \lambda_0\langle y-x, y-x\rangle \\ &= \langle x-u, y-x\rangle + \lambda_0||y-x||^2 \\ &= \langle x-u, y-x\rangle - \langle x-u, y-x\rangle \\ &= 0.\end{aligned}$$
2. Eindeutigkeit: Sei $v_1 = x + \lambda_1(y-x)$ mit $\langle v_1 - u, y-x\rangle = 0$. Dann
   $$\langle x - u + \lambda_1(y-x), y-x\rangle = 0 \qquad \Longrightarrow \qquad \langle x-u, y-x\rangle + \lambda_1||y-x||^2 = 0,$$
   d.h. λ_1 besitzt die Darstellung
   $$\lambda_1 = \frac{\langle u-x, y-x\rangle}{||y-x||^2},$$
   also $\lambda_1 = \lambda_0$ und damit $v_1 = v_0$.
3. Sei $v \in G(x,y)$, dann $v = x + \lambda(y-x)$, d.h.
   $$v - v_0 = x + \lambda(y-x) - x - \lambda_0(y-x) = (\lambda - \lambda_0)(y-x).$$
   Da $(v_0 - u)\perp(y-x)$, folgt $(v_0-u)\perp(\lambda-\lambda_0)(y-x)$, also $(v_0-u)\perp(v-v_0)$ und damit
   $$||u-v||^2 = ||u-v_0||^2 + ||v-v_0||^2 - 2\langle u-v_0, v-v_0\rangle = ||u-v_0||^2 + ||v-v_0||^2 \geq ||u-v_0^2||^2. \quad \square$$

Definition 1.41.
Seien $u, v, w \in \mathbb{R}^n$, so dass v, w linear unabhängig sind. Dann heißt die Menge

$$\{z \in \mathbb{R}^n \mid \exists \alpha, \beta \in \mathbb{R} : z = u + \alpha v + \beta w\}$$

eine Ebene im $\mathbb{R}^n$.

Satz 1.42.
Drei Punkte $u, x, y \in \mathbb{R}^n$, die nicht auf einer Geraden liegen bzw. identisch sind, bestimmen eindeutig eine Ebene $E(u, x, y)$ im $\mathbb{R}^n$; für diese gilt:

$$E(u, x, y) = \{u + \alpha(x - u) + \beta(y - u) \mid \alpha, \beta \in \mathbb{R}\}.$$

Beweis.

1. Existenz.
   a) $u = u + \alpha(x - u) + \beta(y - u) \Leftrightarrow \alpha = 0, \beta = 0 \Rightarrow u \in E$.
   b) $x = u + \alpha(x - u) + \beta(y - u) \Leftrightarrow \alpha = 1, \beta = 0 \Rightarrow x \in E$.
   c) $y = u + \alpha(x - u) + \beta(y - u) \Leftrightarrow \alpha = 0, \beta = 1 \Rightarrow u \in E$.
2. Lineare Unabhängigkeit.
   a) Sei $(x - u) = \lambda(y - u)$ für ein $\lambda \in \mathbb{R}$, d.h. $x = u + \lambda(y - u)$, dann $x \in G(u, y)$, Widerspruch.
   b) Sei $(y - u) = \lambda(x - u)$ für ein $\lambda \in \mathbb{R}$, d.h. $x = u + \lambda(x - u)$, dann $x \in G(x, u)$, Widerspruch.

   Also sind $(x - u), (y - u)$ linear unabhängig.
3. Eindeutigkeit: vgl. Übungen. $\square$

Satz 1.43.
Die Geraden im $\mathbb{R}^2$ liegen genau in der Menge

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid \exists a, b, c \in \mathbb{R},\ a \neq 0 \text{ oder } b \neq 0 : ax_1 + bx_2 = c\}.$$

Beweis.
Seien $u \neq v \in \mathbb{R}^2$ mit $u = (u_1, u_2)$, $v = (v_1, v_2)$. Sei $x \in G$ beliebig, dann existiert ein $\lambda \in \mathbb{R}$ mit $x_1 = u_1 + \lambda(v_1 - u_1)$, d.h. $\lambda = \frac{x_1 - u_1}{v_1 - u_1}$. Dann gilt für x_2:

$$\begin{array}{rrcl} & x_2 & = & u_2 + \lambda(v_2 - u_2) \\ & & = & u_2 + \frac{x_1 - u_1}{v_1 - u_1}(v_2 - u_2) \\ \Leftrightarrow & (x_2 - u_2)(v_1 - u_1) & = & (x_1 - u_1)(v_2 - u_2) \\ \Leftrightarrow & x_2(v_1 - u_1) - u_2(v_1 - u_1) & = & x_1(v_2 - u_2) - u_1(v_2 - u_2) \\ \Leftrightarrow & x_1 \underbrace{(v_2 - u_2)}_{a} - x_2 \underbrace{(v_1 - u_1)}_{b} & = & \underbrace{u_1(v_2 - u_2) - u_2(v_1 - u_1)}_{c} \end{array}$$

Noch zu zeigen: $ax_1 + bx_2 = c \Rightarrow x = (x_1, x_2) \in G(u, v)$. Es gilt

$$ax_1 + bx_2 = c \Leftrightarrow (x_1 - u_1)(v_2 - u_2) = (x_2 - u_2)(v_1 - u_1) \Leftrightarrow x_1 - u_1 = \frac{x_2 - u_2}{v_2 - u_2}(v_1 - u_1).$$

Setze $\lambda = \frac{x_2 - u_2}{v_2 - u_2}$. Dann $x_1 - u_1 = \lambda(v_1 - u_1)$ und $x_2 - u_2 = \lambda(v_2 - u_2)$, d.h. $x = u + \lambda(v - u)$ und damit $x \in G(u, v)$.
Sei umgekehrt $G = \{(x_1, x_2) \mid \exists a, b, c \in \mathbb{R},\ a \neq 0 : ax_1 + bx_2 = c\}$, dann $G = G((\frac{c}{a}, 0), (b + \frac{c}{a}, -a))$. $\square$

1.5. Der endlich-dimensionale K-Vektorraum

Die Definition eines $\mathbb{R}$- bzw. $\mathbb{C}$-Vektorraums lässt sich mühelos auf Vektorräume über beliebigen Körpern verallgemeinern – man muss in diesen Strukturen lediglich addieren und skalieren können. Iteratives Anwenden dieser Operationen liefert die Möglichkeit, neue Vektoren als Linearkombination zu konstruieren. Wir führen in diesem Abschnitt abstrakte Vektorräume und ihre kanonischen Unterstrukturen, die Untervektorräume, ein.

Definition 1.44.

Seien $K = (K, \vdash, \circ)$ ein Körper und $V = (V, \dashv, 0)$ eine abelsche Gruppe.

$(V, \dashv, 0)$ heißt ein K-Vektorraum, falls für alle $x, y \in V$ und alle $\alpha, \beta \in K$ gelten:

1. $(\alpha \vdash \beta)x = \alpha x \dashv \beta x$;
2. $\alpha(x \dashv y) = \alpha x \dashv \alpha y$;
3. $(\alpha\beta)x = \alpha(\beta x)$;
4. $1x = x$.

Bemerkung 1.45.

1. Im Kontext mit Vektorräumen haben wir also vier Operationen zu betrachten:

$$\begin{aligned}(x, y) &\mapsto x \dashv y \in V;\\ (\alpha, \beta) &\mapsto \alpha \vdash \beta \in K;\\ (\alpha, \beta) &\mapsto \alpha\beta \in K;\\ (\alpha, x) &\mapsto \alpha x \in V.\end{aligned}$$

2. Als Körper K können wir etwa $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ und $\mathbb{Z}_2 = \{0, 1\}$ wählen, den zweielementigen Körper, bestehend nur aus den Neutralen von Addition und Multiplikation:

$\vdash$	0	1
0	0	1
1	1	0

$\circ$	0	1
0	0	0
1	0	1

Es gilt $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$, aber $\mathbb{Z}_2$ liegt in keinem dieser Körper.

3. Sei $(K, +, \cdot)$ ein Körper. Wir versehen den Raum K^n der n-Tupel über K mit folgender Addition und skalaren Multiplikation:

$$\begin{aligned}(a_1, ..., a_n) + (b_1, ..., b_n) &= (a_1 + b_1, ..., a_n + b_n);\\ \alpha(a_1, ..., a_n) &= (\alpha a_1, ..., \alpha a_n).\end{aligned}$$

Dann ist K^n ein K-Vektorraum. $\diamondsuit$

Satz 1.46.

Sei $(V, \dashv, 0)$ ein K-Vektorraum. Dann gelten:

1. $0x = 0$;
2. $\alpha 0 = 0$;
3. $\alpha x = 0 \Rightarrow \alpha = 0$ oder $x = 0$;
4. $(-1)x = -x$;
5. $(-\alpha)x = -(\alpha x)$.

Beweis.

1. $0x = (0 \vdash 0)x = 0x \dashv 0x \Rightarrow 0 = 0x$.
2. $0\alpha = (0 \vdash 0)\alpha = 0\alpha \vdash 0\alpha \Rightarrow 0 = 0\alpha$.
3. Seien $\alpha \neq 0$ und $\alpha x = 0$. Dann $0 = \alpha^{-1}0 = \alpha^{-1}(\alpha x) = (\alpha^{-1}\alpha)x = 1x = x$.
4. $0x = (1 - 1)x = (1 \vdash (-1))x = 1x \dashv (-1)x = x \dashv (-1)x \Rightarrow (-1)x = -x$.
5. $-\alpha x = (-1)\alpha x = -(\alpha x)$. ◊

Beispiel 1.47.

1. Sei U ein Unterkörper des Körpers K. Dann ist K ein U-Vektorraum, wenn man setzt $V = (K, +, 0)$ und $\alpha x = \alpha \cdot x$ für $\alpha \in U$ und $x \in K$. Insbesondere sind $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum und $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum. Aber: $\mathbb{Q}$ ist kein $\mathbb{Z}_2$-Vektorraum.
2. Seien X eine nicht-leere Menge und K ein Körper. Dann ist $V = \mathrm{Abb}(X, K) = \{f : X \to K,\ x \mapsto f(x)\}$ ein K-Vektorraum via

$$
\begin{aligned}
(f+g)(x) &= f(x) + g(x) \\
(\alpha f)(x) &= \alpha f(x).
\end{aligned}
$$

für $f, g \in V$ und beliebiges $x \in K$, denn mit $f, g \in V$ und $\alpha \in K$ liegen dann auch $f + g$ und αf in V und es gelten die Vektorraumaxiome. ◊

Definition 1.48.
Seien V ein K-Vektorraum und $x_1, ..., x_n \in V$. Dann heißt

$$
x = \alpha_1 x_1 + \cdots + \alpha_n x_n = \sum_{k=1}^{n} \alpha_k x_k
$$

eine Linearkombination der Vektoren $x_1, ..., x_n$.
$\alpha_1, ..., \alpha_n$ heißen die Koeffizienten von $x \in V$ und $(\alpha_1, ..., \alpha_n) \in K^n$ der Koeffizientenvektor von x.

Bemerkung 1.49.
Seien $x_1, ..., x_n$ Vektoren aus V.
Es gelten:

$$
\begin{aligned}
0x_1 + \cdots + 0x_n &= 0, \\
(\alpha_1 x_1 + \cdots + \alpha_n x_n) + (\beta_1 x_1 + \cdots + \beta_n x_n) &= (\alpha_1 + \beta_1)x_1 + \cdots + (\alpha_n + \beta_n)x_n, \\
\lambda(\alpha_1 x_1 + \cdots + \alpha_n x_x) &= (\lambda\alpha_1)x_1 + \cdots + (\lambda\alpha_n)x_n.
\end{aligned}
$$

Damit ist die Menge

$$
U = Kx_1 + \cdots + Kx_n = \{\alpha_1 x_1 + \cdots + \alpha_n x_n \mid \alpha_1, ..., \alpha_n \in K\}
$$

ein Vektorraum, ein "Untervektorraum": ◊

Definition 1.50.
$U \subseteq V$ heißt ein Untervektorraum von V, falls gelten:

1. $U \neq \emptyset$ (bzw. dazu äquivalent: $0 \in U$).
2. $U + U \subseteq U$ (d.h. $u_1, u_2 \in U \Rightarrow u_1 + u_2 \in U$).
3. $KU \subseteq U$ (d.h. $\alpha \in K,\ u \in U \Rightarrow \alpha u \in U$).

Beispiel 1.51.

1. Der Nullraum $U = \{0\} \subseteq V$ ist Untervektorraum jedes Vektorraums.
2. Geraden und Ebenen durch 0 im $\mathbb{R}^n$ sind Unterräume des $\mathbb{R}^n$:
   a) $G = \{0 + \lambda v \mid \lambda \in \mathbb{R}\} = \mathbb{R}v$ für $v \in \mathbb{R}^n,\ v \neq 0$.
   b) $E = \{0 + \alpha v + \beta w \mid \alpha, \beta \in \mathbb{R}\} = \mathbb{R}v + \mathbb{R}w$ für linear unabhängige $v, w \in \mathbb{R}^n$.
3. $M = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \leq x_2\}$ ist *kein* Untervektorraum des $\mathbb{R}^2$, denn es ist zwar $(-1, 1) \in M$, aber $(-1)(-1, 1) = (1, -1) \notin M$. ♢

1.6. Erzeugendensysteme und lineare Abhängigkeit

Aus Teilmengen der einem Vektorraum zugrunde liegenden Menge lassen sich Unterräume erzeugen. Besonders wichtige solche Teilmengen sind diejenigen, die "linear unabhängig" sind: Durch geeignete Linearkombinationen der enthaltenen Elemente lassen sich alle Vektoren des Raumes eindeutig darstellen.

Definition 1.52.

Sei $X \subseteq V$ nicht-leer. Die Menge

$$U = \{\alpha_1 x_1 + \cdots + \alpha_m x_m \mid m \in \mathbb{N},\ \alpha_1, ..., \alpha_m \in K\}$$

heißt der von X erzeugte oder aufgespannte Untervektorraum und wird mit $\operatorname{span}(X)$ bezeichnet.

X heißt dann ein Erzeugendensystem von $\operatorname{span}(X)$ und $\operatorname{span}(X)$ die lineare Hülle von X.

Beispiel 1.53.

1. Sei V ein Vektorraum, dann ist $\operatorname{span}(V) = V$ und V ist Unterraum von sich selbst.
2. $\operatorname{span}(\emptyset) = \{0\}$. Beachte: Die "leere Linearkombination" ist die 0.
3. $\{e^{(1)}, ..., e^{(n)}\}$ ist ein Erzeugendensystem für den $\mathbb{R}^n$ bzw. K^n, wobei wir setzen:

$$\begin{array}{rcl} e^{(1)} & = & (1, 0, 0, ..., 0), \\ e^{(2)} & = & (0, 1, 0, ..., 0), \\ & \vdots & \\ e^{(n)} & = & (0, 0, ..., 0, 1). \end{array}$$

   Denn für beliebiges $a \in \mathbb{R}^n$ ist $a = (a_1, ..., a_n) = a_1 e^{(1)} + \cdots + a_n e^{(n)} \in \operatorname{span}(\{e^{(1)}, ..., e^{(n)}\})$. ♢

Definition 1.54.

X heißt minimales Erzeugendensystem für V, falls $V = \operatorname{span}(X)$ und $V \neq \operatorname{span}(X \backslash \{v\})$ für alle $v \in X$ ist.

Beispiel 1.55.

$\{e^{(1)}, ..., e^{(n)}\}$ ist ein minimales Erzeugendensystem von $\mathbb{R}^n$, denn (etwa)

$$(\{e^{(2)}, ..., e^{(n)}\}) = \{\alpha_2 e^{(2)} + ... + \alpha_n e^{(n)} \mid \alpha_i \in \mathbb{R}\} = \{(0, \alpha_2, ..., \alpha_n) \mid \alpha_i \in \mathbb{R}\} = \{0\} \times \mathbb{R}^{n-1} \neq \mathbb{R}^n. \quad ♢$$

Definition 1.56.

$\{v_1, ..., v_n\}$ heißt linear abhängig, falls es ein $i \in \{1, ..., n\}$ und $a_j \in K$ $(j \neq i)$ gibt mit

$$v_i = \sum_{j \neq i} \alpha_j v_j = \alpha_1 v_1 + ... + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + ... + \alpha_n v_n.$$

Andernfalls heißt $\{v_1, ..., v_n\}$ linear unabhängig.

Beispiel 1.57.

1. $\{(1,0),(0,1),(1,1)\} \subseteq \mathbb{R}^2$ ist linear abhängig: $(1,1) = (1,0) + (0,1)$.
2. $\{v_1, v_2\} \subseteq V$ ist linear abhängig $\Leftrightarrow \exists \lambda \in K : v_1 = \lambda v_2$ oder $v_2 = \lambda v_1$. $\Diamond$

Satz 1.58.
$\{v_1, ..., v_n\}$ linear unabhängig $\Leftrightarrow \alpha_1 v_1 + ... + \alpha_n v_n = 0$ gilt nur für $\alpha_1 = ... = \alpha_n = 0$.

Beweis.

1. Sei $\{v_1, ..., v_n\}$ linear unabhängig. Es gilt:

$$v_i = \sum_{j \neq i} \alpha_j v_j \qquad \Longrightarrow \qquad 0 = \alpha_1 v_1 + ... + (-1) v_i + ... + \alpha_n v_n,$$

   d.h. sind die Vektoren $v_1, ..., v_n$ linear abhängig, so gibt es eine nicht triviale Darstellung der 0.

2. Gelte $0 = \alpha_1 v_1 + ... + \alpha_n v_n$, etwa $\alpha_1 \neq 0$. Dann ist

$$-\alpha_1 v_1 = \sum_{j=2}^{n} \alpha_j v_j \qquad \Longrightarrow \qquad v_1 = \sum_{j=2}^{n} (\frac{\alpha_j}{-\alpha_1}) v_j,$$

   d.h. die Vektoren sind nicht linear unabhängig. $\square$

Bemerkung 1.59.
$\{v\}$ ist linear unabhängig $\Leftrightarrow v \neq 0$.
Beachte: Der Nullvektor darf in einem System linear unabhängiger Vektoren nicht vorkommen! $\Diamond$

Satz 1.60. (Austauschlemma)
Sei $V = \text{span}(v_1, ..., v_n)$ mit $v_1, ..., v_n$ linear unabhängig. Weiter sei $w = \sum_{j=1}^n \alpha_j v_j$ mit $\alpha_i \neq 0$.
Dann sind $v_1, ..., v_{i-1}, w, v_{i+1}, ..., v_n$ linear unabhängig und $V = \text{span}(v_1, ..., v_{i-1}, w, v_{i+1}, ..., v_n)$.

Beweis.
Gelte Œ $\alpha_1 \neq 0$, dann ist $v_1 = \frac{1}{\alpha_1}(w - \sum_{i=2}^n \alpha_i v_i)$. Sei $v \in V$ beliebig, dann gilt:

$$\begin{aligned}
v &= \sum_{i=1}^{n} \beta_i v_i \in \text{span}(v_1, ..., v_n) \\
&= \beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n \\
&= \beta_1 \frac{1}{\alpha_1} \left(w - \sum_{i=2}^{n} \alpha_i v_i \right) + \beta_2 v_2 + ... + \beta_n v_n \\
&= \beta_1 \alpha_1^{-1} w + (-\alpha_2 \beta_1 \alpha_1^{-1} + \beta_2) v_2 + ... + (-\alpha_n \beta_1 \alpha_1^{-1}) v_n \in \text{span}(w, v_2, ..., v_n)
\end{aligned}$$

Noch zu zeigen: $w, v_2, ..., v_n$ sind linear unabhängig.
Gelte $\lambda_1 w + \lambda_2 v_2 + ... + \lambda_n v_n = 0$. Zu zeigen: $\lambda_1 = ... = \lambda_n = 0$.

$$\begin{aligned}
0 &= \lambda_1 w + \lambda_2 v_2 + ... + \lambda_n v_n \\
&= \lambda_1 (\alpha_1 v_1 + ... + \alpha_n v_n) + \lambda_2 v_2 + ... + \lambda_n v_n \\
&= \lambda_1 \alpha_1 v_1 + (\lambda_1 \alpha_2 + \lambda_2) v_2 + ... + (\lambda_1 \alpha_n + \lambda_n) v_n.
\end{aligned}$$

Da $v_1, ..., v_n$ linear unabhängig, folgt $\lambda_1 \alpha_1 = 0$, $\lambda_1 \alpha_2 + \lambda_2 = 0$, ..., $\lambda_1 \alpha_n + \lambda_n = 0$; da nach Voraussetzung $\alpha_1 \neq 0$, folgt daraus $\lambda_1 = 0 \Rightarrow \lambda_2, ..., \lambda_n = 0$. $\square$

Korollar 1.61.

Sei V endlich erzeugt mit $V = \text{span}(M_1)$ und $V = \text{span}(M_2)$, wobei M_1, M_2 endlich und minimal.

Dann folgt: $\#M_1 = \#M_2$.

Beweis.

Seien $M_1 = \{v_1, ..., v_n\}$ und $M_2 = \{w_1, ..., w_m\}$ mit $m > n$. Wir zeigen: $w_1, ..., w_{n+1}$ sind linear abhängig, d.h. M_2 ist nicht minimal.

Da M_1 Erzeugendensystem von V und $M_2 \subseteq V$, gilt für alle $1 \leq i \leq n$ und alle $1 \leq j \leq n+1$, dass $w_j = \alpha_1^{(j)} v_1 + ... + \alpha_n^{(j)} v_n$. Für alle $i \in \{1, ..., n\}$ gibt es ein $\alpha_i^{(i)} \in K \backslash \{0\}$ mit $w_i = \alpha_i^{(i)} v_i + ...$ (nach geeigneter Umnummerierung der v_i und w_i), denn sonst wäre $V = \text{span}(w_1, ..., w_{n+1}) \subseteq \text{span}(v_1, ..., v_{-1})$ (etwa), d.h. M_1 wäre nicht minimal im Widerspruch zu den Voraussetzungen. Tausche nun mit dem Austauschlemma $v_1, ..., v_n$ gegen $w_1, ..., w_n$ aus. Da $w_{n+1} \in \text{span}(v_1, ..., v_n)$, ist $w_{n+1} \in \text{span}(w_1, ..., w_n)$, d.h. $w_1, ..., w_{n+1}$ sind linear anhängig. □

Bemerkung 1.62.

Die Anzahl der Elemente einer Menge M wird mit $\#M$ bezeichnet und Kardinalität von M genannt. ◊

Satz 1.63. (Austauschsatz von Steinitz)

Es sei $V = \text{span}(v_1, ..., v_n)$ und $v_1, ..., v_n$ seien linear unabhängig. Weiter seien $w_1, ..., w_m \in V$ ebenfalls linear unabhängig.

Dann ist $m \leq n$ und es gibt Indizes $i_1, ..., i_m$, so dass $M = (\{v_1, ..., v_n\} \backslash \{v_{i_1}, ..., v_{i_m}\}) \cup \{w_1, ..., w_m\}$ linear unabhängig sind und $V = \text{span}(M)$ ist.

Beweis. (per Induktion über m)

Der Fall $m = 1$ wird gerade durch das Austauschlemma abgedeckt.

Nach Induktionsvoraussetzung gilt die Behauptung also für m. Also $m \leq n$ und (nach Umnummerierung) $\{w_1, ..., w_m, v_{m+1}, ..., v_n\} =: N$ linear unabhängig und $\text{span}(N) = V$. Seien nun $w_1, ..., w_{m+1}$ linear unabhängig. Wegen $V = \text{span}(N)$ ist $w_{m+1} = \sum_{i=1}^m \alpha_i w_i + \sum_{i=m+1}^n \alpha_i v_i$. Da $w_1, ..., w_{m+1}$ linear unabhängig, gilt $n \geq m+1$ und es gibt $i \geq m+1$ mit $\alpha_i \neq 0$. Mit dem Lemma tauschen wir v_i gegen w_i aus. □

Korollar 1.64.

1. Ist $V = \text{span}(v_1, ..., v_n)$, so gibt es in V nicht mehr als n linear unabhängige Vektoren.
2. Sei V ein endlich erzeugter K-Vektorraum. Dann gibt es ein endliches, minimales Erzeugendensystem von V und alle solche haben die gleiche Kardinalität.

1.7. Basis und Dimension

Minimale Erzeugendensysteme, sog. "Basen", erlauben es, einen Vektorraum, der aus unendlich vielen Elementen besteht, durch eine endliche Teilmenge zu repräsentieren; aus dieser lassen sich dann alle Vektoren des Raumes durch Linearkombinationen rekonstruieren. Die Anzahl der Basiselemnte ist ein Maßfür die Größe des Vektorraums.

Definition 1.65.

Die Kardinalität eines endlichen, minimalen Erzeugendensystems eines K-Vektorraums V heißt die K-Dimension von V und wird mit $\dim_K(V)$ bezeichnet. $n = \dim_K(V)$ heißt die Länge der Basis.

Das n-Tupel $(v_1, ..., v_n)$ heißt eine K-Basis von V, falls $\{v_1, ..., v_n\}$ ein minimales Erzeugendensystem aus n Elementen von V ist.

Bemerkung 1.66.

Gemäß Korollar 1.64 besitzt jeder endlich-dimensionale K-Vektorraum eine Basis und alle diese Basen besitzen die gleiche Länge; die Dimension eines endlich-dimensionalen Raumes ist somit "wohldefiniert", d.h. eindeutig bestimmt.

Wir setzen ab jetzt stets voraus, dass K-Vektorräume endlich-dimensional sind. $\Diamond$

Beispiel 1.67.

1. $(e^{(1)}, ..., e^{(n)})$ ist eine Basis des K^n.
2. $B = ((1,2,3), (2,3,1), (3,1,2))$ ist eine Basis des $\mathbb{R}^3$. Dazu ist nur zu überprüfen, ob $\operatorname{span}(B) = \mathbb{R}^3$ gilt, denn dann ist B automatisch auch linear unabhängig: Wir wissen bereits, dass jede Basis des $\mathbb{R}^3$ genau drei Elemente enthält.
3. $\dim G = 1$ und $\dim E = 2$ für jede Gerade G und jede Ebene E durch 0 im $\mathbb{R}^n$. $\Diamond$

Bemerkung 1.68.

Eine Gerade im $\mathbb{R}^n$ ist ein Untervektorraum des $\mathbb{R}^n$, wenn sie durch den Ursprung geht, und damit selbst wieder ein K-Vektorraum. $\Diamond$

Satz 1.69. (Basisergänzungssatz)

Seien V ein K-Vektorraum und $w_1, ..., w_m$ linear unabhängige Vektoren aus V.

Dann lässt sich $w_1, ..., w_m$ zu einer Basis ergänzen, so dass $(w_1, ..., w_m, v_{m+1}, ..., v_n)$ eine Basis von V ist für gewisse $v_{m+1}, ..., v_n \in V$.

Beweis.

Sei $\{v_1, ..., v_n\}$ ein minimales Erzeugendensystem von V. Tausche mit dem Austauschsatz von Steinitz (Œ) $v_1, ..., v_m$ gegen $w_1, ...w_m$ aus. $\square$

Korollar 1.70.

1. Seien $\dim(V) = n$ und $w_1, ..., w_n$ linear unabhängig. Dann ist $(w_1, ..., w_n)$ eine Basis von V.
2. Sei W ein Untervektorraum von V, dann ist $\dim(W) \leq \dim(V)$.
3. Ist $\dim(W) = \dim(V)$, dann gilt $V = W$.

Bemerkung 1.71.

Der $\mathbb{R}^3$ ist ein $\mathbb{R}$-Vektorraum und enthält vier Untervektorraumtypen: den $\mathbb{R}^3$ selbst, den Nullraum $\{0\}$ sowie Geraden durch 0 und Ebenen durch 0. $\Diamond$

Lemma 1.72.

Seien V ein K-Vektorraum und $v_1, ..., v_m \in V$. Dann gelten:

1. $\operatorname{span}(v_1, ..., v_m) = \operatorname{span}(v_1, ..., v_{i-1}, \lambda v_i, v_{i+1}, ..., v_m)$ für alle $\lambda \in K \backslash \{0\}$.
2. $\operatorname{span}(v_1, ..., v_m) = \operatorname{span}(v_1, ..., \underbrace{v_i + \lambda v_j}_{i\text{–te Stelle}}, ..., v_j, ..., v_m)$ $(i \neq j)$.

Beweis.

1. Sei $v \in \operatorname{span}(v_1, ..., v_m)$, d.h. es gibt $\alpha_1, ..., \alpha_m \in K$ mit $v = \sum_{i=1}^n \alpha_i v_i$. Dann gilt

$$v = \alpha_1 v_1 + ... + \frac{\alpha_i}{\lambda}(\lambda v_i) + ... + \alpha_m v_m,$$

d.h. $v \in \text{span}(v_1, ..., \lambda v_i, ..., v_m)$.

2. Sei wieder $v \in \text{span}(v_1, ..., v_m)$, $v = \sum_{i=1}^m \alpha_i v_i$. Dann ist
$$v = \alpha_1 v_1 + ... + \alpha_i(v_i + \lambda v_j) + ... + (\alpha_j - \alpha_i\lambda)v_j + ... + \alpha_m v_m,$$
d.h. $v \in \text{span}(v_1, ..., v_i + \lambda v_j, ..., v_m)$. □

Bemerkung 1.73.

Genau dann bildet $(v_1, ..., v_n)$ eine Basis des K-Vektorraumes K^n, wenn zu jedem $w \in V$ ein eindeutig bestimmter Koeffizientenvektor $\alpha = (\alpha_1, ..., \alpha_n) \in K^n$ existiert mit $w = \sum_{i=1}^n \alpha_i v_i$, d.h. wenn das "lineare Gleichungssystem"
$$\begin{array}{ccccccccc} X_1v_{11} & + & X_2v_{12} & + & \cdots & + & X_nv_{1n} & = & w_1 \\ X_1v_{21} & + & X_2v_{22} & + & \cdots & + & X_nv_{2n} & = & w_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ X_nv_{n1} & + & X_2v_{n2} & + & \cdots & + & X_nv_{nn} & = & w_1 \end{array}$$
eindeutig lösbar ist (durch $X_1 = \alpha_1$, ..., $X_n = \alpha_n$). ♢

1.8. Summen und direkte Summen von Unterräumen

In diesem Abschnitt wird aus gegebenen Unterräumen eines K-Vektorraums ein neuer Unterraum gebildet und eine Dimensionsformel für Summenräume wird bewiesen. Anschließend werden spezielle Summenräume untersucht: Direkte Summen, bei denen die Summanden im Vektorraumkontext disjunkt sind, und orthogonale Summen von Unterräumen des K^n, bei denen die zu summierenden Vektoren jeweils orthogonal zueinander sind.

Satz 1.74.

Sei V ein K-Vektorraum und seien U, W zwei K-Untervektorräume von V. Dann gelten:

1. $U \cap W$ ist ein K-Untervektorraum von V.
2. $U + W$ ist ein K-Untervektorraum von V.

Beweis.

1. a) $0 \in U$ und $0 \in V \Rightarrow 0 \in U \cap V$.

   b) Seien $x, y \in U \cap W$, dann $x, y \in U$ und $x, y \in V$, also auch $x + y \in U$ und $x + y \in V$, d.h. $x + y \in U \cap V$.

   c) Seien $\alpha \in K$, $x \in u \cap V$. Dann $\alpha x \in U$ und $\alpha x \in V$, folglich $\alpha x \in U \cap V$.

2. Zur Erinnerung: $U + W = \{u + w \mid u \in U \ \& \ w \in W\}$.

   a) $0 \in U$ und $0 \in W \Rightarrow 0 = 0 + 0 \in U + W$.

   b) Seien $x, y \in U + W$, dann gibt es $u_1, u_2 \in U$ und $w_1, w_2 \in W$ mit $u_1 + w_1 = x$ und $u_2 + w_2 = y$, d.h.
   $$x + y = u_1 + w_1 + u_2 + w_2 = (u_1 + u_2) + (w_1 + w_2) \in U + W.$$

   c) Seien $\alpha \in K$ und $x \in U + W$. Dann gibt es $u \in U, w \in W$ mit $u + w = x$ und es gelten $\alpha u \in U$ und $\alpha w \in W$. Damit
   $$\alpha x = \alpha(u + w) = \alpha u + \alpha w \in U + W.$$ □

Satz 1.75. (Erste Dimensionsformel)

Seien V ein K-Vektorraum und U, W zwei K-Untervektorräume von V. Dann gilt:
$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Beweis.

Sei $(v_1, ..., v_n)$ eine Basis von $U \cap W$, $(v_1, ..., v_n, u_1, ..., u_k)$ eine Basis von U und $(v_1, ..., v_n, w_1, ..., w_l)$ eine Basis von W. Es genügt zu zeigen, dass $(v_1, ..., v_n, u_1, ..., u_k, w_1, ..., w_l)$ eine Basis von $U + W$ ist, denn dann folgt

$$\dim(U + W) = n + l + k = (n + k) + (n + l) - n = \dim(U) + \dim(W) - \dim(U \cap W).$$

1. Erzeugendensystem: Sei $v \in U + W$, d.h. $v = u + w$ für gewisse $u \in U, w \in W$ mit

$$u = \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \beta_i u_i, \qquad w = \sum_{i=1}^{n} \gamma_i v_i + \sum_{i=1}^{l} \delta_i w_i.$$

Also ist

$$v = \sum_{i=1}^{n} (\alpha_i + \gamma_i) v_i + \sum_{i=1}^{k} \beta_i u_i + \sum_{i=1}^{l} \delta_i w_i,$$

d.h. $v \in \operatorname{span}(v_1, ..., v_n, u_1, ..., u_k, w_1, ..., w_l)$.

2. Lineare Unabhängigkeit: Gelte

$$0 = \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \beta_k u_k + \sum_{i=1}^{l} \gamma_i w_i \qquad \Longrightarrow \qquad \underbrace{\sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \beta_i u_i}_{\in U} = \underbrace{\sum_{i=1}^{l} (-\gamma_i) w_i}_{\in W}.$$

Dann liegt $x = \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \beta_i u_i = \sum_{i=1}^{l} (-\gamma_i) w_i$ in $U \cap W$. Also

$$\begin{aligned} x = \sum_{i=1}^{n} \lambda_i v_i \qquad &\Longrightarrow \qquad \sum_{i=1}^{n} \lambda_i v_i + \sum_{i=1}^{l} \gamma_i w_i = 0 \\ &\Longrightarrow \qquad \text{alle } \lambda_i = 0 \text{ und alle } \gamma_i = 0, \text{ da } v_i, w_i \text{ linear unabhängig} \\ &\Longrightarrow \qquad \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \beta_i u_i = 0 \\ &\Longrightarrow \qquad \text{alle } \alpha_i = 0 \text{ und alle } \beta_i = 0, \text{ da } v_i, u_i \text{ linear unabhängig.} \end{aligned}$$

Also folgt

$$0 = \sum \alpha_i v_i + \sum \beta_i u_i + \sum \gamma_i w_i = 0 \qquad \Longleftrightarrow \qquad \text{alle } \alpha_i = 0 \ \& \text{ alle } \beta_i = 0 \ \& \text{ alle } \gamma_i = 0.$$

Damit sind u_i, v_i, w_i linear unabhängig. □

Beispiel 1.76.

Betrachte die Unterräume $U = \mathbb{R}(1, 0)$ und $W = \mathbb{R}(0, 1)$ des $\mathbb{R}^2$. Es gelten $U + W = \mathbb{R}^2$ und $U \cap W = \{0\}$, damit $2 = \dim(U + W) = \dim(U) + \dim(W) - \dim(\{0\}) = 1 + 1 - 0$. ◊

Definition 1.77.

Seien U, W Unterräume von V. Dann setzen wir $V = U \oplus W :\Leftrightarrow V = U + W$ und $U \cap W = \{0\}$.

$U + W$ heißt die Summe von U, W; $U \oplus W$ die direkte Summme von U, W.

Bemerkung 1.78.

$V = U \oplus W \Rightarrow \dim(V) = \dim(U) + \dim(W)$. ◊

Lemma 1.79.

$V = U \oplus W \Leftrightarrow$ jedes $v \in V$ lässt sich eindeutig als Summe $v = u + w$ mit $u \in U, w \in W$ schreiben.

Beweis.

1. Sei $V = U \oplus W$. Gelte $v = u_1 + w_1 = u_2 + w_2$ mit $u_1, u_2 \in U,\ w_1, w_2 \in W$. Dann $0 = (u_1 - u_2) + (w_1 - w_2)$, d.h. $u_1 - u_2 = w_2 - w_1 \in U \cap W = \{0\}$. Also sind $u_1 = u_2$ und $w_1 = w_2$.
2. Sei $v = u + w$ eindeutig mit $v \in U \cap W$. Dann $v = v + 0 = 0 + v$, also $v = 0$, d.h. $U \cap W = \{0\}$. □

Definition 1.80.
Seien K ein Körper und $n \in \mathbb{N}$. Dann setzen wir

$$(a_1, ..., a_n) \circ (b_1, ..., b_n) = a_1 b_1 + ... + a_n b_n.$$

Wir schreiben $a \perp b$ für $a \circ b = 0$; in dem Fall nennen wir a, b orthogonal zueinander.

Für $M \subseteq K^n$ setzen wir $M^\perp = \{x \in K^n \mid x \perp a \text{ für alle } a \in M\}$.

Bemerkung 1.81.
Für $K = \mathbb{R}$ ist $\circ$ das euklidische Skalarprodukt. Für $K = \mathbb{C}$ ist $\circ$ dagegen nicht das Hermitesche Skalarprodukt, denn z.B. ist

$$(1, i) \cdot (1, i) = 1 \cdot 1 + i \cdot (-i) = 2 \neq 0 = 1 \cdot 1 + i \cdot i = (1, i) \circ (1, i).$$ ◊

Lemma 1.82.
Für $\circ$ gelten folgende Rechenregeln:

1. $a \circ b = b \circ a$ für alle $a, b \in K^n$.
2. $(\alpha a + \alpha' a') \circ b = \alpha(a \circ b) + \alpha'(a' \circ b)$ für alle $a, a', b \in K^n$ und alle $\alpha, \alpha' \in K$.
3. $a \perp b \Rightarrow \alpha a \perp \beta b$ für alle $a, b \in K^n$ und alle $\alpha, \beta \in K$.

Beweis.
Seien $a = (a_1, ..., a_n)$, $a' = (a'_1, ..., a'_n)$ und $b = (b_1, ..., b_n)$. Dann gelten:

$$\begin{aligned} a \circ b &= (a_1, ..., a_n) \circ (b_1, ..., b_n) \\ &= a_1 b_1 + ... + a_n b_n \\ &= b_1 a_1 + ... + b_n a_n \\ &= (b_1, ..., b_n) \circ (a_1, ..., a_n) \\ &= b \circ a; \end{aligned}$$

$$\begin{aligned} (\alpha a + \alpha' a') \circ b &= (\alpha a_1 + \alpha' a'_1 + ... + \alpha a_n + \alpha' a'_n) \circ (b_1, ..., b_n) \\ &= \alpha a_1 b_1 + \alpha' a'_1 b_1 + ... + \alpha a_n b_n + \alpha' a'_n b_n \\ &= \alpha(a_1 b_1 + ... + a_n b_n) + \alpha'(a'_1 b_1 + ... + a'_n b_n) \\ &= \alpha(a_1, ..., a_n) \circ b + \alpha'(a'_1, ..., a'_n) \circ b \\ &= \alpha(a \circ b) + \alpha'(a' \circ b); \end{aligned}$$

$$\begin{aligned} (\alpha a) \circ (\beta b) &= \alpha(a \circ (\beta b)) \\ &= \alpha((\beta b) \circ a) \\ &= \alpha\beta(b \circ a) \\ &= \alpha\beta(a \circ b) \\ &= \alpha\beta 0 \\ &= 0, \end{aligned}$$

d.h. $a \circ b = 0$ impliziert $(\alpha a) \perp (\beta b)$. □

Lemma 1.83.
Für $M \subseteq K^n$ ist $M^\perp$ ein Untervektorraum des K^n und es gilt $M^\perp = (\text{span}(M))^\perp$.

Beweis.
Seien $x, y \in M^\perp$ und $\alpha \in K$. Dann gelten:

1. $0 \circ a = 0 \Rightarrow 0 \perp a$ für alle $a \in M$, also $0 \in M^\perp$.
2. $(x \circ a = 0$ und $y \circ a = 0) \Rightarrow x \circ a + y \circ a = (x+y) \circ a = 0 \Rightarrow x + y \in M^\perp$.
3. $x \circ a = 0 \Rightarrow (\alpha x) \circ a = 0 \Rightarrow \alpha x \in M^\perp$.

Also ist $M^\perp$ ein Untervektorraum des K^n.
Weiter gilt: $M \subseteq \text{span}(M) \Rightarrow M^\perp \supseteq (\text{span}(M))^\perp$; umgekehrt impliziert $x \circ a = 0$ für alle $a \in M$, dass $x \circ (\alpha_1 a_1 + ... + \alpha_m a_m) = 0$ für alle $a_1, ..., a_m \in M, \alpha_1, ..., \alpha_m \in K$ und damit $M^\perp \subseteq (\text{span}(M))^\perp$. □

1.9. Homogene lineare Gleichungssysteme

In diesem Abschnitt führen wir homogene lineare Gleichungssysteme ein und eine kompakte Form, die darzustellen – die Matrizen. Der Gauß-Algorithmus erlaubt es, die Einträge einer solchen Matrix so zu manipulieren, dass alle Lösungen des zugehörigen Gleichungssystems abgelesen werden können.

Definition 1.84.
Seien K ein Körper und $a_{ij} \in K$ für $1 \leq i \leq m$, $1 \leq j \leq m$. Dann heißt

$$A = (a_{ij})_{1 \leq i \leq m}^{1 \leq j \leq n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

eine $m \times n$-Matrix und die a_{ij} die Komponenten oder Einträge von A.

i bezeichnet den Zeilenindex und j den Spaltenindex von A.

Id bezeichnet die Einheitsmatrix (Identitätsmatrix), die auf der Diagonalen Einsen und sonst überall Nullen stehen hat.

Betrachtet man die m Zeilen als n-Tupel $v_1, ..., v_m$, dann nennt man $\text{span}(v_1, ..., v_m)$ den Zeilenraum von A. Entsprechend bilden die n Spalten $w_1, ..., w_n$ als m-Tupel den Spaltenraum von A.

$\dim(\text{span}(v_1, ..., v_n))$ heißt der Zeilenrang von A und $\dim(\text{span}(w_1, ..., w_m))$ der Spaltenrang.

Bemerkung 1.85.
Der Zeilenraum von A ändert sich nicht bei ...

1. Multiplikation der i-ten Zeile mit $\lambda \neq 0$.
2. Addieren eines λ-fachen der j-ten Zeile zur i-ten Zeile (wobei $i \neq j$).
3. Vertauschen der i-ten mit der j-ten Zeile. ◊

Durch iterierte Anwendung von 1., 2., 3. kann man jede Matrix in eine Stufenmatrix der folgenden Gestalt überführen:

$$\begin{array}{cc} \left(\begin{array}{cc|c|c|cc|c|c|c|c} 0 & 0 & 1 & 0 & * & * & 0 & \cdots & 0 & * \\ 0 & 0 & 0 & 1 & * & * & 0 & \cdots & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array}\right) & \begin{array}{l} \leftarrow 1 \\ \leftarrow 2 \\ \leftarrow 3 \\ \vdots \\ \leftarrow r \\ \end{array} \end{array}$$

Algorithmus 1.86. (Gauß-Algorithmus)

1. Suche die erste Spalte j_1 mit einem $a_{ij_1} \neq 0$.
2. Dividiere die i-te Zeile durch a_{ij_1}, vertausche sie mit der ersten Zeile und mache mit 2. alle anderen Komponenten der j_1-ten Spalte zu 0.
3. Suche eine Spalte j_2 rechts von j_1 mit einem $a_{ij_2} \neq 0$ für $i \neq 1$.
4. Dividiere die i-te Zeile durch a_{ij_2}, vertausche sie mit der zweiten Zeile und mache mit 2. alle anderen Komponenten der j_2-ten Spalte zu 0.
5. u.s.w.

Lemma 1.87.
Die ersten r Zeilen der neuen Matrix sind linear unabhängig und bilden damit eine Basis des Zeilenraums von A.

Beweis.
Seien $w_i = (b_{i1}, ..., b_{in})$ die ersten r Zeilen der neuen Matrix und $\alpha_i \in K$, $i = 1, ..., r$, so dass gilt $0 = \alpha_1 w_1 + ... + \alpha_r w_r$. Dann ist

$$0 = (0, ..., 0, \underbrace{\alpha_1}_{j_1}, 0, ..., 0, \underbrace{\alpha_2}_{j_2}, 0, ..., \underbrace{\alpha_3}_{j_3}, ..., \underbrace{\alpha_r}_{j_r}, ...),$$

also $\alpha_1 = ... = \alpha_r = 0$. □

Definition 1.88.
Seien K ein Körper, $a_{ij} \in K$. Das System

$$\begin{array}{ccccccc} a_{11}X_1 & + & ... & + & a_{1n}X_n & = & 0 \\ & \vdots & & \vdots & & \vdots & \\ a_{m1}X_1 & + & ... & + & a_{mn}X_n & = & 0 \end{array} \qquad (*)$$

heißt ein homogenes lineares Gleichungssystem über K in den Unbestimmten oder Variablen $X_1, ..., X_n$.

Die zugehörige Koeffizientenmatrix ist

$$A = (a_{ij})_{1 \leq i \leq m}^{1 \leq j \leq n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Eine Lösung des Systems ist ein n-Tupel $x = (x_1, ..., x_n) \in K$, für das gilt

$$\begin{array}{ccccccc} a_{11}x_1 & + & ... & + & a_{1n}x_n & = & 0 \\ & \vdots & & \vdots & & \vdots & \\ a_{m1}x_1 & + & ... & + & a_{mn}x_n & = & 0. \end{array}$$

$\mathbb{L} = \{x \in K \mid x \text{ ist eine Lösung von } (*)\}$ heißt die Lösungsmenge des Systems.

Beispiel 1.89.

1. Betrachte das Gleichungssystem

$$\begin{array}{ccccc} x_1 & + & 2x_2 & = & 0 \\ x_1 & - & x_2 & = & 0. \end{array}$$

Die zugehörige Koeffizientenmatrix ist

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

die (eindeutige) Lösung ist $x_1 = 0$ & $x_2 = 0$.

2. $x_1 + 2x_2 = 0$ hat als Lösungsmenge alle Vektoren $(x_1, x_2) \in \mathbb{R}^2$ mit $(x_1, x_2) \perp (1, 2)$, d.h. $\mathbb{L} = \mathbb{R}(2, -1)$.
3. Das homogene lineare Gleichungssystem

$$\begin{array}{ccccccc} x_1 & + & 2x_2 & & & = & 0 \\ x_1 & + & 2x_2 & + & x_3 & = & 0 \end{array}$$

hat als Lösungsmenge $\mathbb{L} = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x \perp (1, 2, 0) \ \& \ x \perp (1, 2, 1)\}$. $\diamondsuit$

Satz 1.90.
Die Lösungsmenge von $(*)$ ist gerade $\mathbb{L} = \{(x_1, ..., x_n) \in K^n \mid (x_1, ..., x_n) \perp (a_{i1}, ..., a_{in}) \text{ für } i = 1, ..., m\}$.

Beweis.
Es gilt:

$$\begin{array}{rcl} x \text{ löst } (*) & \iff & Ax = 0 \\ & \iff & A^{(i)}x = 0 \text{ für jede Zeile } A^{(i)} \\ & \iff & A^{(i)} \perp x \text{ für jede Zeile } A^{(i)}. \end{array}$$

$\square$

Korollar 1.91.
Bezeichne M den Zeilenraum von A. Dann ist der Vektorraum $M^\perp$ gerade die Lösungsmenge von $(*)$.
$M^\perp$ wird der Lösungsraum von $(*)$ genannt.

Bemerkung 1.92. (Bestimmung einer Basis des Lösungsraums von $(*)$)
Gegeben sei das lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{11}X_1 & + & ... & + & a_{1n}X_n & = & 0 \\ & \vdots & & \vdots & & \vdots & \\ a_{m1}X_1 & + & ... & + & a_{mn}X_n & = & 0 \end{array} \qquad (*)$$

mit zugehöriger Koeffizientenmatrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Durch elementare Zeilenumformungen 1., 2., 3. wird A auf Stufenform transformiert:

$$\left(\begin{array}{cc|c|c|cc|c|c|c|c} 0 & 0 & 1 & 0 & * & * & 0 & \cdots & 0 & * \\ 0 & 0 & 0 & 1 & * & * & 0 & \cdots & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * \\ \underbrace{0}_{k_1} & \underbrace{0}_{k_2} & \underbrace{0}_{j_1} & \underbrace{0}_{j_2} & \underbrace{0}_{k_3} & \underbrace{0}_{k_4} & \underbrace{0}_{j_3} & \cdots & \underbrace{0}_{j_r} & \underbrace{0}_{k_{n-r}} \end{array}\right) \begin{array}{l} \leftarrow 1 \\ \leftarrow 2 \\ \leftarrow 3 \\ \vdots \\ \leftarrow r \\ \\ \end{array} = (a'_{ij}) = A',$$

d.h. $\{1, ..., n\} = \{j_1, ..., j_r\} \ \cup \ \{k_1, ..., k_{n-r}\}$.
Dann hat das zu A' gehörige Gleichungssystem

$$\begin{array}{cccccccccccc} X_{j_1} & & & & + & a'_{1k_1}X_{k_1} & + & \cdots & + & a'_{1k_{n-r}}X_{k_{n-r}} & = & 0 \\ & X_{j_1} & & & + & a'_{2k_1}X_{k_1} & + & \cdots & + & a'_{2k_{n-r}}X_{k_{n-r}} & = & 0 \\ & & \ddots & & \vdots & & \vdots & & \vdots & & & \vdots \\ & & & X_{j_r} & + & a'_{rk_1}X_{k_1} & + & \cdots & + & a'_{rk_{n-r}}X_{k_{n-r}} & = & 0 \end{array} \qquad (**)$$

den gleichen Lösungsraum wie $(*)$.

Spezielle Lösungen $(x_1, ..., x_n) \in K^n$ von $(**)$ sind dann:

$$\left.\begin{array}{cccccc} (1) & x_{k_1} & = & 1, & x_{j_l} & = & -a'_{lk_1} \\ (2) & x_{k_2} & = & 1, & x_{j_l} & = & -a'_{lk_2} \\ \vdots & & \vdots & & & \vdots & \\ (n-r) & x_{k_{n-r}} & = & 1, & x_{j_l} & = & -a'_{lk_{n-r}} \end{array}\right\} \text{für } 1 \le l \le r,\ x_i = 0 \text{ sonst.}$$

Wir bezeichnen diese Lösungen mit $c^{(1)}, ..., c^{(n-r)}$, d.h. es gilt:

$$c^{(i)} = (c_1^{(i)}, ..., c_n^{(i)}) \text{ mit } c_{k_i}^{(i)} = 1,\ c_{j_l}^{(i)} = -a'_{lk_i} \text{ für } 1 \le l \le r \text{ und } c_{k_s}^{(i)} = 0 \text{ für } s \ne i.$$ ◊

Lemma 1.93.
$(c^{(1)}, ..., c^{(n-r)})$ ist eine Basis des Lösungsraums von $(*)$ bzw. $(**)$.

Beweis.

1. Lineare Unabhängigkeit: Es gilt

$$(0, ..., 0) = (\alpha_1 c^{(1)} + ... + \alpha_{n-r} c^{(n-r)}) = (..., \underbrace{\alpha_1}_{k_1}, ..., \underbrace{\alpha_{n-r}}_{k_{n-r}}, ...) \implies \alpha_1 = ... = \alpha_{n-r} = 0.$$

2. Erzeugendensystem: Sei $b = (b_1, ..., b_n)$ eine beliebige Lösung von $(*)$. Dann ist auch das m-Tupel $x = (x_1, ..., x_n) = b - b_{k1} c^{(1)} - ... - b_{k,n-r} c^{n-r}$ eine Lösung von $(*)$, also $x_{k_i} = 0$ für alle k_i. Da x auch $(**)$ löst, sind auch alle $x_{j_l} = 0$, d.h. $x = 0$. Also $b = b_{k_1} c^{(1)} + ... + b_{k,n-r} c^{n-r}$. □

Korollar 1.94.
Ein homogenes lineares Gleichungssystem $(*)$ in n Variablen besitzt genau dann nur die triviale Lösung $x_1 = 0, ..., x_n = 0$, wenn der Spaltenrang der zugehörigen Koeffizientenmatrix n beträgt.

Beweis.
$(*)$ ist genau dann nur trivial lösbar, wenn der Lösungsraum von $(*)$ der Nullraum ist. Da die einzige Basis des Nullraums die leere Menge ist, kann es keine Lösungen der Gestalt $c^{(i)}$ geben, d.h. $n = r$. □

Korollar 1.95.
Für jeden Untervektorraum U des K^n gilt:

1. $\dim U + \dim U^\perp = n$.
2. $(U^\perp)^\perp = U$.

Beweis.

1. Sei $U = \operatorname{span}(v_1, ..., v_m)$ mit $v_i = (a_{i1}, ..., a_{in})$. Dann gilt:

$$\dim U^\perp = n - \text{Zeilenrang}(a_{ij}) = n - \dim \operatorname{span}(v_1, ..., v_m) = n - \dim U.$$

2. Es gilt $u \in U \Rightarrow u \perp v$ für alle $v \in U^\perp \Rightarrow u \in (U^\perp)^\perp \Rightarrow U \subseteq (U^\perp)^\perp$.Umgekehrt folgt wegen $\dim U + \dim U^\perp = n$ und $\dim U^\perp + \dim (U^\perp)^\perp = n$, dass $\dim U = \dim (U^\perp)^\perp$. Also ist insgesamt $(U^\perp)^\perp = U$. □

Bemerkung 1.96.
Im Allgemeinen gilt nicht $U \oplus U^\perp = V$: Betrachte etwa $V = \mathbb{C}^2$ mit $U = \operatorname{span}((1, i))$, dann ist $U = U^\perp$, da $(1, i) \perp (1, i)$, also $U + U^\perp = U \ne V$ und $U \cap U^\perp = U \ne \{0\}$. ◊

Beispiel 1.97.
Wir betrachten das Gleichungssystem

$$\begin{array}{rrrrrrrrrrrrl}
 & & + & 2x_3 & - & 2x_4 & + & 6x_5 & - & x_6 & = & 0 \\
 & 2x_2 & + & 6x_3 & + & 4x_4 & + & 2x_5 & - & 4x_6 & = & 0 \\
- & x_2 & - & x_3 & - & 4x_4 & + & 7x_5 & + & x_6 & = & 0 \\
- & 2x_2 & - & x_3 & - & 9x_4 & + & 17x_5 & + & 2x_6 & = & 0 \\
 & & - & x_3 & + & x_4 & - & 3x_5 & - & \frac{1}{2}x_6 & = & 0
\end{array}.$$

Mit dem Gauß-Algorithmus erhalten wir

$$\begin{pmatrix} 0 & 0 & 2 & -2 & 6 & -1 \\ 0 & 2 & 6 & 4 & 2 & -4 \\ 0 & -1 & -1 & -4 & 7 & 1 \\ 0 & -2 & -1 & -9 & 17 & 2 \\ 0 & 0 & -1 & 1 & -3 & -0,5 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 & 0 & 5 & -10 & 0 \\ 0 & 0 & 1 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \underbrace{0}_{k_1} & \underbrace{0}_{j_1} & \underbrace{0}_{j_2} & \underbrace{0}_{k_2} & \underbrace{0}_{k_3} & \underbrace{0}_{j_3} \end{pmatrix}$$

$$\Longrightarrow \begin{array}{rcllllll}
 & & k_1 & & & k_2 & k_3 & \\
c^{(1)} & = & (1, & 0, & 0, & 0, & 0, & 0) \\
c^{(2)} & = & (0, & -5, & 1, & 1, & 0, & 0) \\
c^{(3)} & = & (0, & 10, & -3, & 0, & 1, & 0)
\end{array}$$

sind die speziellen Lösungen zu $(*)$. $\Diamond$

Beispiel 1.98.
Wir betrachten das System

$$\begin{array}{rrrrrrrrrrrrl}
x_1 & + & 2x_2 & - & x_3 & - & 2x_4 & & & - & x_6 & = & 0 \\
2x_1 & + & 2x_2 & - & 2x_3 & - & x_4 & + & x_5 & + & 2x_6 & = & 0 \\
2x_1 & + & 4x_2 & + & (5/2)x_3 & + & 4x_4 & - & 2x_5 & + & 2x_6 & = & 0 \\
x_1 & & & - & x_3 & - & x_4 & & & + & x_6 & = & 0
\end{array}.$$

Der Gauß-Algorithmus liefert

$$\begin{aligned}
& \begin{pmatrix} 1 & 2 & -1 & -2 & 0 & -1 \\ 2 & 2 & -2 & -1 & 1 & 2 \\ 2 & 4 & (5/2) & 4 & -2 & 2 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{pmatrix} \\
\longrightarrow & \begin{pmatrix} 1 & 2 & -1 & -2 & 0 & -1 \\ 0 & 2 & 0 & -3 & -1 & -4 \\ 0 & 0 & -(9/2) & -8 & 2 & -4 \\ 0 & 2 & 0 & -1 & 0 & -2 \end{pmatrix} \\
\longrightarrow & \begin{pmatrix} 1 & 0 & -1 & 1 & 1 & 3 \\ 0 & 1 & 0 & -(3/2) & -(1/2) & -2 \\ 0 & 0 & -(9/2) & -8 & 2 & -4 \\ 0 & 0 & 0 & -2 & -1 & -2 \end{pmatrix} \\
\longrightarrow & \begin{pmatrix} 1 & 0 & 0 & (25/9) & (5/9) & (35/9) \\ 0 & 1 & 0 & -(3/2) & -(1/2) & -2 \\ 0 & 0 & 1 & (16/9) & -(4/9) & (8/9) \\ 0 & 0 & 0 & -2 & -1 & -2 \end{pmatrix} \\
\longrightarrow & \begin{pmatrix} 1 & 0 & 0 & 0 & -(5/6) & (10/9) \\ 0 & 1 & 0 & 0 & (1/4) & -(1/2) \\ 0 & 0 & 1 & 0 & -(4/3) & -(8/9) \\ 0 & 0 & 0 & 1 & (1/2) & 1 \end{pmatrix};
\end{aligned}$$

Spezielle Lösungen sind

$$c^{(1)} = \left(\frac{5}{6}, -\frac{1}{4}, \frac{4}{3}, -\frac{1}{2}, 1, 0\right) \quad \text{und} \quad c^{(2)} = \left(-\frac{10}{9}, \frac{1}{2}, \frac{8}{9}, -1, 0, 1\right).$$ $\Diamond$

1.10. Inhomogene lineare Gleichungssysteme

Inhomogene Gleichungssysteme sind solche, bei denen die rechte Seite von Null verschieden ist. Die Lösungsmenge eines solchen Systems erhält man durch Lösen eines homogenen Systems, sofern eine spezielle Lösung des inhomogenen Systems bekannt ist; sie stellt dann einen um die spezielle Lösung verschobenen Untervektorraum dar. Im Gegensatz zu homogenen Systemen, die stets zumindest die Nulllösung besitzen, muss ein inhomogenes System keine Lösung haben.

Definition 1.99.
Seien K ein Körper, $a_{ij}, b_i \in K$ und $X_1, ..., X_n$ Unbestimmte. Das System

$$\begin{array}{ccccccc} a_{11}X_1 & + & ... & + & a_{1n}X_n & = & b_1 \\ & \vdots & & \vdots & & \vdots & \\ a_{m1}X_1 & + & ... & + & a_{mn}X_n & = & b_m \end{array} \tag{+}$$

heißt ein inhomogenes lineares Gleichungssystem über K in den Unbestimmten $X_1, ..., X_n$ mit einfacher Koeffizientenmatrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

und erweiterter Koeffizientenmatrix

$$(A|b) = (a_{ij}|b_i) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

Das System

$$\begin{array}{ccccccc} a_{11}X_1 & + & ... & + & a_{1n}X_n & = & 0 \\ & \vdots & & \vdots & & \vdots & \\ a_{m1}X_1 & + & ... & + & a_{mn}X_n & = & 0 \end{array} \tag{$*$}$$

heißt das zugehörige homogene System.

Bemerkung 1.100.
Sei $A = (a_{ij})$ die einfache Matrix von (+). $A^{(i)}$ bezeichne die i-te Zeile von A und $X = (X_1, ..., X_n)$ den Variablenvektor. Dann besagt (+) gerade:

$$A_1 \circ X = b_1, \qquad ..., \qquad A_m \circ X = b_m.$$ ◊

Definition 1.101.
Seien V ein K-Vektorraum und U ein Untervektorraum von V. Für $v \in V$ heißt

$$v + U = \{v + u \mid u \in U\}$$

der um v verschobene oder affine Unterraum.

Bemerkung 1.102.

1. $v + U$ ist genau dann ein Untervektorraum von V, wenn $v \in U$ gilt.
2. Sei $v \notin U$. Dann ist $(v + U, \vdash, \circ)$ ein K-Vektorraum, wenn man setzt

$$(v + u_1) \vdash (v + u_2) = v + (u_1 + u_2),$$
$$\alpha \circ (v + u) = v + \alpha u.$$

   Nullelement dieses Raums ist $v = v + 0$. ◊

Lemma 1.103.

Bezeichne $\mathbb{L}$ den Lösungsraum von $(*)$. Sei x' eine spezielle Lösung von $(+)$.

Dann ist der Lösungsraum von $(+)$ gerade der affine Raum $\mathbb{L}_b = x' + \mathbb{L}$.

Beweis.

Es gilt:

$$\begin{aligned} x \in K^n \text{ löst } (+) \quad &\Longleftrightarrow \quad A_1 \circ x = b_1, ..., A_m \circ x = b_m \\ &\Longleftrightarrow \quad A_1 \circ x = A_1 \circ x', ..., A_m \circ x = A_m \circ x' \\ &\Longleftrightarrow \quad A_1 \circ (x - x') = 0, ..., A_m \circ (x - x') = 0 \\ &\Longleftrightarrow \quad x - x' \in \mathbb{L} \\ &\Longleftrightarrow \quad x \in x' + M. \end{aligned}$$

Also ist $\mathbb{L}_b = x' + \mathbb{L}$. $\square$

Lemma 1.104.

Bezeichne $(A^{(i)}, b_i)$ die i-te Zeile der erweiterten Matrix $(a_{ij}|b_i)$ und $(X, X_{n+1}) = (X_1, ..., X_n, X_{n+1})$.

Betrachte das homogene System

$$(A_1, b_1) \circ (X, X_{n+1}) = 0, \qquad ..., \qquad (A_m, b_m) \circ (X, X_{n+1}) = 0. \tag{++}$$

Dann gilt:

$$x = (x_1, ..., x_n) \text{ löst } (+) \quad \Longleftrightarrow \quad (x, -1) \text{ löst } (++).$$

Beweis.

Es gilt:

$$a_{i1}x_1 + ... + a_{in}x_n = b_i \quad \Longleftrightarrow \quad a_{i1}x_1 + ... + a_{in}x_n + b_i(-1) = 0.$$

$\square$

Bemerkung 1.105.

Wir betrachten die erweiterte Matrix $(A|b) = (a_{ij}|b_i)$. Durch elementare Zeilenumformungen erhalten wir

$$\left(\begin{array}{ccc|c|ccc|c|ccc||c}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & b'_1 \\
 & & & 0 & & & & 0 & * & \cdots & * & \\
\vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
 & & & 0 & & & & 0 & * & \cdots & * & \\
0 & \cdots & 0 & 0 & * & \cdots & * & 1 & * & \cdots & * & b'_r \\
 & & & 0 & & & & 0 & & & & b'_{r+1} \\
 & 0 & & \vdots & & 0 & & \vdots & & 0 & & \vdots \\
 & & & \underbrace{0}_{j_1} & & & & \underbrace{0}_{j_r} & & & & b'_m
\end{array}\right) = (a'_{ij}|b'_i).$$

Aus dieser Darstellungsform $(A'|b')$ lassen sich Aussagen über die Lösbarkeit von $(+)$ treffen und – sofern vorhanden – eine spezielle Lösung ablesen: $\Diamond$

Satz 1.106.

1. $(+)$ ist genau dann lösbar, wenn $b'_{r+1} = 0, ..., b'_m = 0$ erfüllt ist.
2. Eine spezielle Lösung ist dann $x' = (0, ..., 0, b'_1, 0, ..., 0, b'_r, 0, ..., 0)$.

Beweis.

1. Seien $b'_{r+1}, ..., b'_m = 0$. Dann löst $x'' = (x', -1)$ das homogene System $(++)'$, denn

$$\begin{aligned}
0? + ... + 0? + b'_1 1 + 0? + ... + 0? + b'_2 0 + 0? + ... + ... + 0? + b'_r 0 + 0? + ... + 0? - b'_1 &= 0\\
0? + ... + 0? + b'_1 0 + 0? + ... + 0? + b'_2 1 + 0? + ... + ... + 0? + b'_r 0 + 0? + ... + 0? - b'_2 &= 0\\
&\vdots\\
0? + ... + 0? + b'_1 0 + 0? + ... + 0? + b'_2 0 + 0? + ... + ... + 0? + b'_r 1 + 0? + ... + 0? - b'_3 &= 0.
\end{aligned}$$

   Also löst x' das System $(+)'$ und damit auch $(+)$.

2. Sei etwa $b'_{r+1} \neq 0$. Wir multiplizieren die $(r+1)$-te Zeile mit $(b'_{r+1})^{-1}$ und erreichen:

$$\left(\begin{array}{ccc|c|ccc|c|ccc||c}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & b'_1\\
 & & & 0 & & & & 0 & * & \cdots & * & \\
\vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
 & & & 0 & & & & 0 & * & \cdots & * & \\
0 & \cdots & 0 & 0 & * & \cdots & * & 1 & * & \cdots & * & b'_r\\
 & & & 0 & & & & 0 & & & & 1\\
 & 0 & & \vdots & & 0 & & \vdots & & 0 & & \vdots\\
 & & & 0 & & & & 0 & & & & b'_m
\end{array}\right).$$

   Also fordert die $(r+1)$-te Zeile des zugehörigen inhomogenen Gleichungssystems:

$$0 = 0X_1 + 0X_2 + ... + 0X_n = 1,$$

   ein Widerspruch. Also ist das System in dem Fall nicht lösbar. □

Satz 1.107.
Das inhomogene lineare Gleichungssystem $(+)$ ist genau dann lösbar, wenn der Zeilenrang von A mit dem Zeilenrang von $(A|b)$ übereinstimmt.

Beweis.
Wir zeigen:

$$b'_{r+1} = 0, ..., b'_m = 0 \qquad \Longleftrightarrow \qquad \text{Zeilenrang}(A) = \text{Zeilenrang}(A|b).$$

"$\Rightarrow$" ist klar. Zu "$\Leftarrow$": Sei (etwa) $b'_{r+1} \neq 0$, dann ist $\text{Zeilenrang}(A|b) = r + 1 > r = \text{Zeilenrang}(A)$. □

Definition 1.108.
$(+)$ heißt universell lösbar, falls $(+)$ bei *jeder* Wahl der b_i lösbar ist.
$(+)$ heißt eindeutig lösbar, falls $(+)$ zu jeder Wahl der b_i *höchstens eine* Lösung hat.

Bemerkung 1.109.
Beachte: Ein System, das für *gar keine* Wahl der b_i lösbar ist, wird auch eindeutig lösbar genannt. ◊

Satz 1.110.
1. $(+)$ ist genau dann eindeutig lösbar, wenn der Spaltenrang von A gleich der Spaltenzahl n ist.
2. $(+)$ ist genau dann universell lösbar, wenn der Zeilenrang von A gleich der Zeilenzahl m ist.

Beweis.

1. Eindeutiger Fall: Für $b_1, ..., b_m \in K$ beliebig gilt: $\text{Spaltenrang}(A) = n \Leftrightarrow (*)$ ist nur trivial lösbar. Also ist die Lösungsmenge von $(+)$ entweder $x' + \{0\} = x'$ oder $\emptyset$.

2. Universeller Fall:
   a) Gelte Zeilenrang$(A) = r =$ Zeilenzahl $= m$, dann Zeilenrang$(A|b) = r =$ Zeilenrang(A).
   b) Sei der Zeilenrang r von A kleiner als die Zeilenzahl m. Wir bezeichnen die r linear unabhängigen Zeilen von A mit $A^{(i_1)}, ..., A^{(i_r)}$.

      Wähle ein $i_0 \in \{1, ..., m\} \setminus \{i_1, ..., i_r\}$ und setze $b_i = 1$ für $i = i_0$ und andernfalls $b_i = 0$. Dann ist $(A^{(i_0)}, 1) \notin \operatorname{span}((A^{(i_1)}, 0), ..., (A^{i_r}, 0))$, d.h. Zeilenrang$(A|b) > r =$ Zeilenrang(A), also ist $(+)$ nicht lösbar. □

2. Lineare Abbildungen

In diesem Kapitel untersuchen wir Abbildungen, die mit der Vektorraum-Struktur verträglich sind, sogenannte Homomorphismen. Diese haben die Eigenschaft, dass Bilder und Urbilder von Vektorräumen stets Unterräume sind. Im ersten Abschnitt betrachten wir die grundlegenden Eigenschaften linearer Abbildungen. In Abschnitt zwei erhalten wir eine wichtige Repräsentation linearer Abbildungen: Die Darstellungsmatrizen. In Abschnitt drei wird die Invertierbarkeit von Matrizen behandelt, die zum Auflösen linearer Gleichungssysteme sehr nützlich ist. Der vierte Abschnitt beschäftigt sich mit Basiswechseln und Transformationsmatrizen. Im letzen Abschnitt werden Determinanten eingeführt, die zentrale Eigenschaften von Matrizen in einer Zahl repräsentieren.

2.1. Vektorraum-Homomorphismen

In diesem Abschnitt führen wir die Begriffe Linearität, Kern und Bild ein und beweisen eine zweite Dimensionsformel. Als zentrale Aussage erhalten wir, dass jeder n-dimensionale K-Vektorraum die gleiche Vektorraum-Struktur besitzt wie der kanonische K-Vektorraum K^n.

Definition 2.1.

Seien V, W zwei K-Vektorräume. $f : V \to W$ heißt eine K-lineare Abbildung bzw. ein K-Vektorraum-Homomorphismus, falls für alle $v_1, v_2 \in V$ und alle $\lambda \in K$ gelten:

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$;
2. $f(\lambda v_1) = \lambda f(v_1)$.

Bemerkung 2.2.

1. a) $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)$:
$$f(\alpha v_1 + \beta v_2) = f(\alpha v_1) + f(\beta v_2) = \alpha f(v_1) + \beta f(v_2).$$
   b) $f(0) = 0$:
$$0 = f(0) - f(0) = f(0 + 0) - f(0) = f(0) + f(0) - f(0) = f(0).$$
   c) $f(v_1 - v_2) = f(v_1) - f(v_2)$:
$$f(v_1 - v_2) = f(v_1 + (-v_2)) = f(v_1) + f(-v_2) = f(v_1) - f(v_2).$$
2. $v_1, ..., v_m$ linear anhängig $\Rightarrow f(v_1), ..., f(v_m)$ linear abhängig:

   Gelte $0 = \alpha_1 v_1 + ... + \alpha_m v_m$ mit $\alpha_i \neq 0$, dann auch
$$0 = f(0) = f(\alpha_1 v_1 + ... + \alpha_m v_m) = \alpha_1 f(v_1) + ... + \alpha_m f(v_m)$$
   mit $\alpha_i \neq 0$, d.h. $f(v_1), ..., f(v_m)$ sind linear abhängig. ◊

Satz 2.3.

1. Ist $V' \subseteq V$ ein Untervektorraum, dann auch das Bild $f(V') = \{f(v') \mid v' \in V'\} \subseteq W$.
2. Der Kern eines Homomorphismus $\operatorname{Kern}(f) = \{v \in V \mid f(v) = 0\}$ ist ein Untervektorraum von V.

Beweis.

1. a) Es gilt $0 \in V' \Rightarrow 0 = f(0) \in f(V')$.

   b) Seien $w_1, w_2 \in f(V')$, dann $w_1 = f(v_1)$ und $w_2 = f(v_2)$ für gewisse $v_1, v_2 \in V'$. Aus $v_1 + v_2 \in V'$ folgt $f(v_1 + v_2) = f(v_1) + f(v_2) \in f(V')$, also $w_1 + w_2 \in f(V')$.

   c) Seien $w \in f(V')$ und $\lambda \in K$, dann gibt es $v \in V'$ mit $f(v) = w$, also $\lambda w = \lambda f(v) = f(\lambda v) \in f(V')$.

2. a) Es gilt $f(0) = 0$, also $0 \in \mathrm{Kern}(f)$.

   b) Seien $v_1, v_2 \in \mathrm{Kern}(f)$, dann gilt $0 = f(v_1) = f(v_2)$, also $0 = f(v_1) + f(v_2) = f(v_1 + v_2)$, d.h. $v_1 + v_2 \in \mathrm{Kern}(f)$.

   c) Seien $\lambda \in K$ und $v \in \mathrm{Kern}(f)$, dann ist $f(v) = 0$, also auch $f(\lambda v) = \lambda f(v) = \lambda 0 = 0$, d.h. $\lambda v \in \mathrm{Kern}(f)$. □

Beispiel 2.4.

1. Die Nullabbildung $f : V \to W$, $f(v) = 0$ ist ein Vektorraum-Homomorphismus mit $\mathrm{Kern}(f) = V$ und $f(V) = \{0\}$.

2. Sei $\lambda \in K$, dann ist die Gerade $f : V \to V$ mit $f(v) = \lambda v$ linear, denn $f(0) = \lambda 0 = 0$ und
$$f(\alpha v + \beta w) = \lambda(\alpha v + \beta w) = \lambda\alpha v + \lambda\beta w = \alpha f(v) + \beta f(w).$$

3. Seien allgemeiner $a_1, ..., a_n \in K$. Die Linearkombination $f : K^n \to K$, $f(x_1, ..., x_n) = \sum_{i=1}^n a_i X_i$ ist linear: Mit $= (a_1, ..., a_n)$ und $X = (x_1, ..., x_n)$ ist $f(X) = A \circ X$ und es gelten:

   a) $f(0) = 0$,

   b) $f(X + Y) = A \circ (X + Y) = A \circ X + A \circ Y = f(X) + f(Y)$ und

   c) $f(\alpha X) = A \circ (\alpha X) = \alpha(A \circ X) = \alpha f(X)$.

   Speziell für $n = 2$ ist f eine Ebene im K^n.

4. Die Abbildung $f : K^n \to K^m$, gegeben durch $f_j(x_1, ..., x_n) = \sum_{i=1}^n a_{ji} x_i$ mit $a_{ji} \in K$ ist linear: Setze $A_j = (a_{j1}, ..., a_{jn})$, dann ist $f(X) = (A_1 \circ X, ..., A_m \circ X) = (A_j \circ X)_j$ und es folgen:

   a) $f(0) = (0, ..., 0) = 0$,

   b) $f(X + Y) = (A_j \circ (X + Y))_j = (A_j \circ X + A_j \circ Y)_j = (A_j \circ X)_j + (A_j \circ Y)_j = f(X) + f(Y)$ und

   c) $f(\alpha X) = f(A_j \circ (\alpha X))_j = (\alpha(A_j \circ X))_j = \alpha(A_j \circ X)_j = \alpha f(X)$. ◊

Satz 2.5.
Ist $f : V \to W$ eine K-lineare Abbildung, dann gilt $\dim f(V) \leq \dim V$.

Beweis.
Sei $(v_1, ..., v_n)$ eine Basis von v, d.h. $\dim V = n$. Dann ist
$$\begin{aligned} f(V) &= \{f(v) \mid v \in V\} \\ &= \{f(\alpha_1 v_1 + ... + \alpha_n v_n \mid \alpha_1, ..., \alpha_n \in K)\} \\ &= \{\alpha_1 f(v_1), ..., \alpha_n f(v_n) \mid \alpha_1, ..., \alpha_n \in K\} \\ &= \mathrm{span}(f(v_1), ..., f(v_n)), \end{aligned}$$
d.h. $\dim f(V) \leq n = \dim V$. □

Satz 2.6.
Ist $f : V \to W$ eine K-lineare Abbildung, dann gilt: f ist injektiv $\Leftrightarrow \mathrm{Kern}(f) = \{0\}$.

Beweis.

1. Sei $x \in \text{Kern}(f) \backslash \{0\}$. Dann $f(x) = f(0) = 0$, d.h. f ist nicht injektiv.
2. Gelte $f(x) = f(y) \Rightarrow 0 = f(x) - f(y) = f(x-y) \Rightarrow x - y \in \text{Kern}(f) = \{0\} \Rightarrow x - y = 0 \Rightarrow x = y$. □

Satz 2.7.

Seien V und W zwei K-Vektorräume, $(v_1, ..., v_n)$ eine Basis von V und $v \in V,\ w_1, ..., w_n \in W$ beliebig. Dann gelten:

1. Die Darstellung $v = \alpha_1 v_1 + ... + \alpha_n v_n$ ist eindeutig.
2. Es gibt genau eine K-lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i,\ 1 \leq i \leq n$.
3. $\text{span}(w_1, ..., w_n) = f(W)$.
4. f ist injektiv $\Leftrightarrow w_1, ..., w_n$ sind K-linear unabhängig.

Beweis.

1. Gelte $v = \alpha_1 v_1 + ... + \alpha_n v_n = \beta_1 v_1 + ... + \beta_n v_n$, dann $0 = v - v = (\alpha_1 - \beta_1)v_1 + ... + (\alpha_n - \beta_n)v_n$. Da $v_1, ..., v_n$ linear unabhängig, folgt $\alpha_1 - \beta_1 = 0, ..., \alpha_n - \beta_n = 0$, also $\alpha_1 = \beta_1, ..., \alpha_n = \beta_n$.
2. a) Existenz: Sei $v \in V$ mit der Darstellung $v = \alpha_1 v_1 + ... + \alpha_n v_n$. Setze $f(v) = \alpha_1 w_1 + ... + \alpha_n w_n$, dann ist $f(v_i) = w_i$ und es gelten:

$$\begin{aligned} f(u+w) &= f((\lambda_1 v_1 + ... + \lambda_n v_n) + (\mu_1 v_1 + ... + \mu_n v_n)) \\ &= f((\lambda_1 + \mu_1) + ... + (\lambda_n + \mu_n)v_n) \\ &= (\lambda_1 + \mu_1)w_1 + ... + (\lambda_n + \mu_n)w_n \\ &= (\lambda_1 w_1 + ... + \lambda_n w_n) + (\mu_1 w_1 + ... + \mu_n w_n) \\ &= f(u) + f(w); \end{aligned}$$

$$\begin{aligned} f(\lambda v) &= f(\lambda(\alpha_1 v_1 + ... + \alpha_n v_n)) \\ &= f((\lambda\alpha_1)v_1 + ... + (\lambda\alpha_n)v_n) \\ &= (\lambda\alpha_1)w_1 + ... + (\lambda\alpha_n)w_n \\ &= \lambda(\alpha_1 w_1 + ... + \alpha_n w_n) \\ &= \lambda f(v). \end{aligned}$$

   b) Eindeutigkeit: Seien f, g zwei K-lineare Abbildungen mit $w_i = f(v_i) = g(v_i)$ und $v \in V$ mit der Darstellung $v = \alpha_1 v_1 + ... + \alpha_n v_n \in V$. Dann gilt:

$$\begin{aligned} f(v) &= f(\alpha_1 v_1 + ... + \alpha_n v_n) \\ &= \alpha_1 f(v_1) + ... + \alpha_n f(v_n) \\ &= \alpha_1 w_1 + ... + \alpha_n w_n \\ &= \alpha_1 g(v_1) + ... + \alpha_n g(v_n) \\ &= g(\alpha_1 v_1 + ... + \alpha_n v_n) \\ &= g(v). \end{aligned}$$

3. Sei $w \in \text{span}(w_1, ..., w_n)$, d.h. $w = \sum \alpha_i w_i$. Setze $v = \sum \alpha_i v_i$, dann $f(v) = \sum \alpha_i f(v_i) = w$, d.h. $w \in f(V)$ und damit $\text{span}(w_1, ..., w_n) \subseteq f(V)$.

   Sei umgekehrt $w \in f(V)$, d.h. es gibt ein $v \in V$ mit $f(v) = w$. Habe v die Darstellung $v = \sum \alpha_i w_i$, dann ist $w = f(v) = \sum \alpha_i w_i$, d.h. $w \in \text{span}(w_1, ..., w_n)$, also auch $f(V) \subseteq \text{span}(w_1, ..., w_n)$.
4. Seien f injektiv und $\sum \alpha_i w_i = 0$, dann $\sum \alpha_i f(f_i) = f(\sum \alpha_i v_i) = 0$, d.h. $\sum \alpha_i v_i = 0$ und damit alle $\alpha_i = 0$, d.h. $w_1, ..., w_n$ linear unabhängig.

   Seien umgekehrt $w_1, ..., w_n$ linear unabhängig und $v \in \text{Kern}(f)$. Habe v die Darstellung $v = \sum \alpha_i v_i$, dann ist $0 = f(v) = \sum \alpha_i w_i$, also alle $\alpha_i = 0$ und damit $v = 0$, d.h. $\text{Kern}(f) = \{0\}$ und damit f injektiv. □

Satz 2.8. (Zweite Dimensionsformel)

Seien V, W zwei K-Vektorräume und $f : V \to W$ ein K-Vektorraum-Homomorphismus. Dann gilt:

$$\dim V = \dim \operatorname{Kern}(f) + \dim \operatorname{Bild}(f)$$

Beweis.

Seien $(w_1, ..., w_r)$ eine Basis von $\operatorname{Bild}(f)$, $(u_1, ..., u_k)$ eine Basis von $\operatorname{Kern}(f)$ und $(v_1, ..., v_r \in V$ mit $w_1 = f(v_1), ..., w_r = f(v_r)$. Wir zeigen: $(v_1, ..., v_r, u_1, ..., u_k)$ ist eine Basis von V, dann gilt nämlich: $\dim V = r + k = \dim \operatorname{Bild}(f) + \dim \operatorname{Kern}(f)$.

1. Erzeugendensystem: Sei $v \in V$. Dann existieren $\lambda_1, ..., \lambda_r$, so dass gilt $v = \lambda_1 v_1 + ... + \lambda_r v_r$ und damit $f(v) = f(\lambda_1 v_1 + ... + \lambda_r v_r) = \lambda_1 w_1 + ... + \lambda_r w_r \Rightarrow f(v) - \lambda_1 f(v_1) - ... - \lambda_r f(v_r) = f(v - \lambda_1 v_1 - ... - \lambda_r v_r) = 0$, d.h. $v - \lambda_1 v_1 - ... - \lambda_r v_r \in \operatorname{Kern}(f)$. Also gibt es $\delta_1, ..., \delta_k$ mit $v - \lambda_1 v_1 - ... - \lambda_r v_r = \delta_1 u_1 + ... + \delta_k u_k$, d.h. $v = \lambda_1 v_1 + ... + \lambda_r v_r + \delta_1 u_1 + ... + \delta_k u_k$, d.h. $v \in \operatorname{Kern}(f) + \operatorname{Bild}(f)$.

2. Lineare Unabhängigkeit: Sei $\lambda_1 v_1 + ... + \lambda_r v_r + \delta_1 u_1 + ... + \delta_k u_k = 0$. Dann gilt:

$$0 = f(0) = f\left(\sum_{i=1}^r \lambda_i v_i + \sum_{i=1}^k \delta_i u_i\right) = f\left(\sum_{i=1}^r \lambda_i v_i\right) + \underbrace{f\left(\sum_{i=1}^k \delta_i u_i\right)}_{\in \operatorname{Kern}(f)} = \sum_{i=1}^r \lambda_i f(v_i) = \sum_{i=1}^r \lambda_i w_i.$$

   Da $w_1, ..., w_r$ linear unabhängig, sind alle $\lambda_1, ..., \lambda_r = 0$, d.h. $\sum_{i=1}^k \delta_i u_i = 0$, also alle $\delta_1, ..., \delta_k = 0$, da $u_1, ..., u_k$ linear unabhängig. Also sind $v_1, ..., v_r, u_1, ..., u_k$ linear unabhängig. □

Definition 2.9.

Seien V, W zwei K-Vektorräume und $f : V \to W$ ein K-Vektorraum-Homomorphismus.

Man bezeichnet f als einen ...

1. Monomorphismus, falls f injektiv ist.
2. Epimorphismus, falls f surjektiv ist.
3. Isomorphismus, falls f bijektiv ist.
4. Endomorphismus, falls $V = W$.
5. Automorphismus, falls f bijektiv und $V = W$.

Bemerkung 2.10.

Ist f ein Isomorphismus, dann heißen die Räume V, W isomorph. ◊

Satz 2.11.

Jeder n-dimensionale K-Vektorraum V ist isomorph zum K^n.

Beweis.

Sei $(v_1, ..., v_n)$ eine Basis von V und bezeichne $(e^{(1)}, ..., e^{(n)})$ die kanonische Basis des K^n. Wir setzen $f(v_i) = e^{(i)}$. Es wurde bereits gezeigt, dass dies einen injektiven Homomorphismus definiert. Da außerdem $\operatorname{Bild}(f) = \operatorname{span}(f(v_1), ..., f(v_n)) = \operatorname{span}(e^{(1)}, ..., e^{(n)}) = K^n$, ist f auch surjektiv. □

Bemerkung 2.12.

Ein Homomorphismus ist durch seine Werte auf einer Basis also bereits eindeutig festgelegt. ◊

Satz 2.13.
Seien V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann gilt:

$$f \text{ ist injektiv} \iff f \text{ ist surjektiv} \iff f \text{ ist bijektiv.}$$

Beweis.

1. f injektiv $\Rightarrow$ f surjektiv: Aus $\operatorname{Kern}(f) = \{0\}$ folgt mit der Dimensionsformel: $\dim V = \dim \operatorname{Bild}(f)$. Da $\operatorname{Bild}(f) = f(V) \subseteq V$, folgt $V = \operatorname{Bild}(f)$, d.h. f ist surjektiv.
2. f surjektiv $\Rightarrow$ f injektiv: Aus $\operatorname{Bild}(f) = V$ folgt mit der Dimensionsformel, dass $\dim \operatorname{Bild}(f) = \dim V$, d.h. $\dim \operatorname{Kern}(f) = 0$, also $\operatorname{Kern}(f) = 0$, d.h. f ist injektiv. □

Bemerkung 2.14.

1. Seien U, V, W Vektorräume, $f : V \to W$ sowie $g : W \to U$ lineare Abbildungen. Dann ist auch die Verkettung $g \circ f : V \to U$ linear:
$$(g \circ f)(\alpha u + \beta w) = g(f(\alpha u + \beta w)) = g(\alpha f(u) + \beta f(w)) = \alpha(g \circ f)(u) + \beta(g \circ f)(w).$$
2. Für $f, g : V \to W$ und $\alpha \in K$ setzen wir $(f + g) : V \to W$, $(f + g)(v) = f(v) + g(v)$ und $(\alpha f) : V \to W$, $(\alpha f)(v) = \alpha f(v)$. Die Abbildungen von V nach W bilden mit dieser Addition und Skalarmultiplikation einen K-Vektorraum, bezeichnet mit $\operatorname{Abb}(V, W)$. Die Menge der K-linearen Abbildungen $\operatorname{Hom}(V, W) \subseteq \operatorname{Abb}(V, W)$ bildet einen Untervektorraum. Insbesondere sind für $f, g \in \operatorname{Hom}(V, W)$ und $\alpha \in K$ auch $f + g, \alpha f \in \operatorname{Hom}(V, W)$:
$$(f + g)(\alpha u + \beta w) = \alpha f(u) + \beta f(w) + \alpha g(u) + \beta g(w) = \alpha((f + g)(u)) + \beta((f + g)(w))$$
und
$$(\lambda f)(\alpha u + \beta w) = \lambda f(\alpha u + \beta w) = \lambda\alpha f(u) + \lambda\alpha f(w) = \lambda(\alpha f)(u) + \lambda(\alpha f)(w).$$
3. Sei $f : V \to W$ ein Isomorphismus. Dann existiert zu jedem $w \in W$ genau ein $v \in V$ mit $f(v) = w$, d.h. es existiert die Umkehrabbildung $f^{-1} : W \to V$ mit $f(v) \mapsto v$ bzw. $f^{-1}(w) =$ dasjenige $v \in V$ mit $f(v) = w$. Insbesondere gilt $f(f^{-1}(w)) = w$ und $f^{-1}(f(v)) = v$, d.h. $f \circ f^{-1}$ ist die Identität id auf W und $f^{-1} \circ f$ ist die Identität auf V.

   Mit f ist auch f^{-1} linear, denn es gelten:

   a) $f^{-1}(0) = 0$, da $f(f^{-1}(0)) = 0$,

   b) $f^{-1}(u + w) = f^{-1}(u) + f^{-1}(w)$, da $f(f^{-1}(u) + f^{-1}(w)) = f(f^{-1}(u)) + f(f^{-1}(w)) = u + w$ und

   c) $f^{-1}(\alpha v) = \alpha f^{-1}(v)$, da $f(\alpha f^{-1}(v)) = \alpha f(f^{-1}(v)) = \alpha v$.

   Die Menge der invertierbaren Endomorphismen $\operatorname{Iso}(V, W)$ der Isomorphismen von V nach W bildet im Allgemeinen allerdings keinen Untervektorraum von $\operatorname{Hom}(V, W)$, da die Nullabbildung nur dann in $\operatorname{Iso}(V, W)$ liegt, wenn sowohl V als auch W der Nullraum sind.
4. $\operatorname{Iso}(V, V)$, versehen mit der Verkettung, bildet eine Gruppe mit neutralem Element id. Insbesondere ist mit f, g auch $f \circ g$ invertierbar; genauer gilt $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. ◊

2.2. Darstellungsmatrizen und lineare Abbildungen

Wir haben gesehen, dass eine lineare Abbildung $f : V \to W$ genau durch ihre Werte auf einer Basis festgelegt ist. Sind speziell $V = K^n$ und $W = K^m$, dann ist f durch die n Vektoren $f(v_1), ..., f(v_n)$ im K^m festgelegt, d.h. durch die Matrix im $K^{m \times n}$, deren n Spalten gerade die m-Tupel $f(v_1), ..., f(v_n)$ sind. Allgemeiner kann man auch eine lineare Abbildung $f : V \to W$ mit n- bzw. m-dimensionalen K-Vektorräumen als Matrix repräsentieren, da V zum K^n und W zum K^m isomorph sind: Seien $\mathfrak{V}$ eine Basis von V und $\mathfrak{W}$ eine Basis von W, dann vermitteln die Isomorphismen $\Psi_{\mathfrak{V}} : V \to K^n$ und $\Psi_{\mathfrak{W}} : W \to K^m$ zusammen mit f eine lineare Abbildung $\Psi_{\mathfrak{W}} \circ f \circ \Psi_{\mathfrak{V}}^{-1}$, die von K^n nach K^m abbildet und damit als $(m \times n)$-Matrix dargestellt werden kann.

Definition 2.15.

Seien K ein Körper und $m, n \in \mathbb{N}$. Weiter bezeichne $K^{m\times n} = \mathfrak{M}_K(m,n)$ die Menge der $(m \times n)$-Matrizen über K: $A = (a_{ij})_{1\le j\le n}^{1\le i\le m}$ mit $a_{ij} \in K$.

Wir definieren auf $\mathfrak{M}(m,n)$ eine Addition, eine skalare Multiplikation und ein Nullelement:

$$\begin{aligned} A + B &= (a_{ij} + b_{ij}) \in \mathfrak{M}(m,n), \\ \lambda A &= (\lambda a_{ij}) \in \mathfrak{M}(m,n) \\ 0 &= (0) \in \mathfrak{M}(m,n). \end{aligned}$$

Zu $A = (a_{ij})$ setzen wir $A^{\mathrm{T}} = (a_{ji})$. Dann heißt A^{T} die zu A transponierte Matrix.

Bemerkung 2.16.

1. Es gilt $A^{\mathrm{T}} \in \mathfrak{M}(n,m)$; diese Matrix entsteht aus A, indem man die Spalten von A zu den Zeilen von A^{T} macht.

   Weiter gelten $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$ und $(\lambda A)^{\mathrm{T}} = \lambda A^{\mathrm{T}}$ sowie $0^{\mathrm{T}} = 0$. Die Transpositionsabbildung $f : \mathfrak{M}(m,n) \to \mathfrak{M}(n,m)$, $A \mapsto f(A) = A^{\mathrm{T}}$ ist somit ein Vektorraum-Homomorphismus. f ist sogar ein Isomorphismus mit $f^{-1}(A) = A^{\mathrm{T}}$, denn $(A^{\mathrm{T}})^{\mathrm{T}} = A$.

2. Die $(m \times n)$-Matrizen bilden einen K-Vektorraum. Dieser ist isomorph zum $K^{n\cdot m}$ via

   $$f : \mathfrak{M}(m,n) \to K^{m\cdot n}, \qquad (a_{ij}) \mapsto (a_{11}, ..., a_{1n}, ..., a_{m1}, ..., a_{mn}).$$

   Damit gilt insbesondere: $\dim \mathfrak{M}(m,n) = m \cdot n$. ◊

Definition 2.17.

Zu $A \in \mathfrak{M}(m,n)$ und $B \in \mathfrak{M}(n,r)$ definieren wir ein Matrixprodukt:

$$(a_{ij})_{1\le i\le m}^{1\le j\le n} \cdot (b_{ij})_{1\le i\le n}^{1\le j\le r} = (c_{ij})_{1\le i\le m}^{1\le j\le r} \text{ mit } c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

Bemerkung 2.18.

1. Wir bezeichnen ab jetzt zu $A = (a_{ij}) \in \mathfrak{M}(m,n)$ mit $A_i = (a_{i1}, ..., a_{in}) \in K^n$ (genauer: $\mathfrak{M}(1,n)$) die i-te Zeile der Matrix A und mit $A^{(j)} = (a_{1j}; ...; a_{mj}) \in K^m$ (genauer: $\mathfrak{M}(m,1)$) die j-te Spalte. Dann können wir auch schreiben:

   $$c_{ij} = A_i \circ (B^{(j)})^{\mathrm{T}} \in K.$$

2. Zu $A, B \in \mathfrak{M}(n,n)$ ist $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$. Im Allgemeinen gilt nicht $(AB)^{\mathrm{T}} = A^{\mathrm{T}} B^{\mathrm{T}}$.

3. Wir fassen von nun an die Elemente des K^n als Spaltenvektoren auf, identifizieren sie also mit Elementen des $\mathfrak{M}(n,1)$. ◊

Definition 2.19.

Sei $A \in \mathfrak{M}(m,n)$. Dann heißt

$$\mathrm{Lin}_m^n(A) : K^n \to K^m, \qquad x \mapsto \mathrm{Lin}_m^n(A)x = Ax$$

mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto Ax = \begin{pmatrix} a_{11} & + \cdots + & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & + \cdots + & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 & + \cdots + & a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n \end{pmatrix} = \begin{pmatrix} A_1 \circ x \\ \vdots \\ A_m \circ x \end{pmatrix}$$

eine lineare Abbildung.

Bemerkung 2.20.

1. Anwenden von $\mathrm{Lin}(A)$ auf einen Vektor $x \in K^n$ entspricht also der Multiplikation mit $A \in \mathfrak{M}(m,n)$ von rechts.
2. Speziell: $\mathrm{Lin}(A)(e^{(i)}) = Ae^{(i)}$ ist gerade die i-te Spalte $A^{(i)}$ von A.
3. $\dim \mathrm{Bild}(\mathrm{Lin}(A))$ ist gerade der Spaltenrang von A. $\diamondsuit$

Satz 2.21.

Für $A \in \mathfrak{M}(m,n)$ gilt stets:

$$\mathrm{Zeilenrang}(A) = \mathrm{Spaltenrang}(A) = \mathrm{rg}(A).$$

$\mathrm{rg}(A)$ heißt der Rang von A.

Beweis.

Bezeichne $Z = \mathrm{span}(A_1, ..., A_m)$ den Zeilenraum, n die Anzahl der Spalten und $\mathbb{L}$ den Lösungsraum von A. Wir haben bereits gezeigt, dass dann $\mathbb{L} = Z^\perp$ und $\dim(\mathbb{L}) = n - \dim(Z)$ gelten. Außerdem ist $\mathbb{L} = \{x \in K^m \mid Ax = 0\} = \mathrm{Kern}(A)$. Mit der Dimensionsformel erhalten wir:

$$\begin{aligned}\mathrm{Spaltenrang}(A) &= \dim \mathrm{Bild}(\mathrm{Lin}(A)) \\ &= \dim(K^n) - \dim \mathrm{Kern}(\mathrm{Lin}(A)) \\ &= n - (n - \mathrm{Zeilenrang}(A)) \\ &= \mathrm{Zeilenrang}(A).\end{aligned}$$

$\square$

Satz 2.22.

$\mathrm{Lin} : \mathfrak{M}(m,n) \to \mathrm{Hom}(K^n, K^m)$, gegeben durch $A \mapsto \mathrm{Lin}(A)$, ist ein K-Vektorraum-Isomorphismus.

Beweis.

1. Lin ist surjektiv: Sei $f \in \mathrm{Hom}(K^n, K^m)$. Dann ist f eindeutig bestimmt durch $f(e^{(i)}) = (a_{1i}, ..., a_{mi})$. Setze $A = (a_{ij}) \in \mathfrak{M}(m,n)$, dann ist $\mathrm{Lin}(A)(e^{(i)}) = Ae^{(i)} = (a_{1i}; ...; a_{mi}) = f(e^{(i)})$, d.h. $\mathrm{Lin}(A)$ und f stimmen auf einer Basis des K^n überein, also $\mathrm{Lin}(A) = f$.
2. Lin ist K-linear: Seien $A, B \in \mathfrak{M}(m,n)$ und $\alpha \in K$. Dann gelten

   $$\mathrm{Lin}(A+B)(e^{(i)}) = (A+B)e^{(i)} = A^{(i)} + B^{(i)} = Ae^{(i)} + Be^{(i)} = \mathrm{Lin}(A)(e^{(i)}) + \mathrm{Lin}(A)(e^{(i)}),$$

   also $\mathrm{Lin}(A+B) = \mathrm{Lin}(A) + \mathrm{Lin}(B)$, und

   $$\mathrm{Lin}(\alpha A)(e^{(i)}) = (\alpha A)e^{(i)} = \alpha A^{(i)} = \alpha A e^{(i)} = \alpha \mathrm{Lin}(A(e^{(i)}),$$

   also $\mathrm{Lin}(\alpha A) = \alpha \mathrm{Lin}(A)$.

   Beachte: Es genügt, die Linearität auf einer Basis nachzurechnen.
3. Lin ist injektiv: Sei $A \in \mathrm{Kern}(\mathrm{Lin})$. Angenommen, es gilt $A \neq 0$, dann gibt es ein $a_{ij} \neq 0$. Damit ist die j-te Komponente von $\mathrm{Lin}(A)e^{(i)}$ gerade $a_{ij} \neq 0$, d.h. $\mathrm{Lin}(A) \neq 0$, ein Widerspruch. $\square$

Bemerkung 2.23.

Damit ist auch $\mathrm{Mat} : \mathrm{Hom}(K^n, K^m) \to \mathfrak{M}(m,n)$, gegeben durch $f \mapsto \mathrm{Lin}^{-1}(f) = (f(e^1), ..., f(e^{(n)}))$, ein Isomorphismus.

Zu $f \in \mathrm{Hom}(K^n, K^m)$ heißt $A = \mathrm{Mat}(f) \in \mathfrak{M}(m,n)$ die Darstellungsmatrix von f. $\diamondsuit$

Satz 2.24.

Seien $A \in \mathfrak{M}(m, n)$ und $B \in \mathfrak{M}(n, r)$. Dann gilt:

$$\mathrm{Lin}_m^n(A) \circ \mathrm{Lin}_n^r(B) = \mathrm{Lin}_m^r(AB).$$

Beweis.

Nach der Definition des Matrixproduktes ist $AB \in \mathfrak{M}(m, r)$. Weiter gilt:

$$(\mathrm{Lin}(A) \circ \mathrm{Lin}(B))e^{(i)} = \mathrm{Lin}(A)B^{(i)} = AB^{(i)} = (AB)^{(i)} = (AB)e^{(i)} = \mathrm{Lin}(AB)e^{(i)}. \qquad \square$$

Definition 2.25.

Seien V ein n- und W ein m-dimensionaler K-Vektorraum und seien $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V sowie $\mathfrak{W} = (w_1, ..., w_m)$ eine Basis von W

$\Psi_\mathfrak{V} : V \to K^n$ mit $\Psi_\mathfrak{V}(v_i) = e^{(i)}$ und $\Psi_\mathfrak{W} : W \to K^m$ mit $\Psi_\mathfrak{W}(w_i) = e^{(i)}$ heißen die kanonischen Isomorphismen oder Koeffizientenabbildungen.

Sei $f \in \mathrm{Hom}(V, W)$. Dann heißt $A = \mathrm{Mat}(\Psi_\mathfrak{W} \circ f \circ \Psi_\mathfrak{V}^{-1}) \in \mathfrak{M}(n, m)$ die Darstellungsmatrix von f bzgl. der Basen $\mathfrak{V}, \mathfrak{W}$. Wir schreiben $A = \mathrm{Mat}_\mathfrak{W}^\mathfrak{V}(f)$.

Bemerkung 2.26.

Die i-te Spalte von $\mathrm{Mat}_\mathfrak{W}^\mathfrak{V}(f)$ ist

$$A^{(i)} = \mathrm{Mat}(\Psi_\mathfrak{W} \circ f \circ \Psi_\mathfrak{V}^{-1})e^{(i)} = \Psi_\mathfrak{W}(f(v_i)).$$

Sei $A^{(i)} = (a_{1i}; ...; a_{mi})$, dann gilt

$$f(v_i) = a_{1i}w_1 + ... + a_{mi}w_m. \qquad \lozenge$$

Bemerkung 2.27.

1. Zu $v \in V$ ist $\Psi_\mathfrak{V}(v) = x \in \mathbb{K}^n$ der Koeffizientenvektor von v: Es gilt $v = \sum x_i v_i$.
2. Bezeichne $\mathfrak{E} = (e^{(1)}, ..., e^{(n)})$ die kanonische Basis des K^n.

   Wegen $\Psi_\mathfrak{E} = \mathrm{id}$ gilt für jeden Endomorphismus $f \in \mathrm{Hom}(K^n, K^n)$, dass $\mathrm{Mat}_\mathfrak{E}^\mathfrak{E}(f) = \mathrm{Mat}(f)$.
3. Im Fall $W = V$, $\mathfrak{W} = \mathfrak{V}$ und $f = \mathrm{id}$ ist
   $$\mathrm{Mat}_\mathfrak{W}^\mathfrak{V}(f) = \mathrm{Mat}(\Psi_\mathfrak{V} \circ \mathrm{id} \circ \Psi_\mathfrak{V}^{-1}) = \mathrm{Mat}(\mathrm{id}) = \mathrm{Id},$$
   d.h. die Darstellungsmatrix der Identitätsfunktion bzgl. identischer Basen ist die Identitätsmatrix.
4. Die Verkettung von Funktionen ist assoziativ, d.h. es gilt $f \circ (g \circ h) = (f \circ g) \circ h$. Ebenso ist die Multiplikation von Matrizen assoziativ, d.h. $A(BC) = (AB)C$. Verkettung und Matrizenmultiplikation sind allerdings im Allgemeinen nicht kommutativ, d.h. $f \circ g \neq g \circ f$ und $AB \neq BA$.
5. Darstellungsmatrix einer einfachen Verkettung: Seien V_1, V_2, V_3 K-Vektorräume mit Basen $\mathfrak{V}_1$, $\mathfrak{V}_2$, $\mathfrak{V}_3$ und $f \in \mathrm{Hom}(V_1, V_2)$, $g \in \mathrm{Hom}(V_2, V_3)$. Dann erhalten wir die Darstellungsmatrix der Verkettung $g \circ f : V_1 \to V_3$ via
   $$\begin{aligned}\mathrm{Mat}_{\mathfrak{V}_3}^{\mathfrak{V}_1}(g \circ f) &= \mathrm{Mat}(\Psi_{\mathfrak{V}_3} \circ (g \circ f) \circ \Psi_{\mathfrak{V}_1}^{-1}) \\ &= \mathrm{Mat}((\Psi_{\mathfrak{V}_3} \circ g \circ \Psi_{\mathfrak{V}_2}^{-1}) \circ (\Psi_{\mathfrak{V}_2} \circ f \circ \Psi_{\mathfrak{V}_1}^{-1})) \\ &= \mathrm{Mat}(\Psi_{\mathfrak{V}_3} \circ g \circ \Psi_{\mathfrak{V}_2}^{-1})\mathrm{Lin}^{-1}(\Psi_{\mathfrak{V}_2} \circ f \circ \Psi_{\mathfrak{V}_1}^{-1}) \\ &= \mathrm{Mat}_{\mathfrak{V}_3}^{\mathfrak{V}_2}(g)\mathrm{Mat}_{\mathfrak{V}_2}^{\mathfrak{V}_1}(f),\end{aligned}$$
   denn für $A = \mathrm{Mat}_{\mathfrak{V}_2}^{\mathfrak{V}_1}(f)$ und $B = \mathrm{Mat}_{\mathfrak{V}_3}^{\mathfrak{V}_2}(g)$ gilt ja $\mathrm{Lin}(BA) = \mathrm{Lin}(B)\mathrm{Lin}(A)$.
6. Darstellungsmatrix einer zweifachen Verkettung: Sei V_4 ein weiterer K-Vektorraum mit Basis $\mathfrak{V}_4$ und sei $h \in \mathrm{Hom}(V_3, V_4)$, dann erhalten wir die Darstellungsmatrix der Verkettung $h \circ g \circ f : V_1 \to V_4$ via
   $$\mathrm{Mat}_{\mathfrak{V}_4}^{\mathfrak{V}_1}(h \circ g \circ f) = \mathrm{Mat}_{\mathfrak{V}_4}^{\mathfrak{V}_3}(h)\mathrm{Mat}_{\mathfrak{V}_3}^{\mathfrak{V}_2}(g)\mathrm{Mat}_{\mathfrak{V}_2}^{\mathfrak{V}_1}(f). \qquad \lozenge$$

2.3. Invertierbarkeit von Matrizen und Homomorphismen

In diesem Abschnitt stellen wir das Konzept der Matrixinvertierung vor, mit dem sich lineare Gleichungssysteme kompakt auflösen lassen. Außerdem wird ein Zusammenhang zwischen inversen Matrizen und den zugehörigen linearen Abbildungen bzw. zwischen inversen Homomorphismen und den zugehörigen Darstellungsmatrizen hergestellt.

Definition 2.28.

Eine Matrix $A \in \mathfrak{M}(n,n)$ heißt invertierbar, falls es ein $B \in \mathfrak{M}(n,n)$ gibt mit $AB = BA = \mathrm{Id}$.

B heißt die zu A inverse Matrix und wird mit A^{-1} bezeichnet.

Bemerkung 2.29.

1. Ist $A \in \mathfrak{M}(n,n)$ invertierbar, dann ist A^{-1} eindeutig bestimmt: Gelten $AB = BA\mathrm{Id}$ und $AC = CA\mathrm{Id}$, dann ist
$$B = B\mathrm{Id} = B(AC) = (BA)C = \mathrm{Id}C = C.$$
2. Die Menge $\mathfrak{I}(n) \subseteq \mathfrak{M}(n,n)$ der invertierbaren $(n \times n)$-Matrizen, versehen mit der Multiplikation, bildet eine Gruppe, da gelten:
   a) $A, B \in \mathfrak{I}(n) \Rightarrow AB \in \mathfrak{I}(n)$ mit $(AB)^{-1} = B^{-1}A^{-1}$.
   b) $(AB)C = A(BC)$, $\mathrm{Id}A = A = A\mathrm{Id}$ und zu jedem $A \in \mathfrak{I}(n)$ gibt es $B = A^{-1} \in \mathfrak{I}(n)$ mit $AB = \mathrm{Id} = BA$.

   $\mathfrak{I}(n)$ ist aber im Allgemeinen nicht kommutativ.
3. Die Gruppen $\mathrm{Iso}(K^n, K^n)$ und $\mathfrak{I}(n)$ sind homomorph, d.h. es gibt ein $G : \mathrm{Iso}(K^n, K^n) \to \mathfrak{I}(n)$ mit $G(f \circ g) = G(f)G(g)$. Wir können nämlich $G(f) = \mathrm{Mat}(f)$ wählen.

   G ist umkehrbar mit $G^{-1}(A) = \mathrm{Lin}(A)$; es gilt $G^{-1}(AB) = \mathrm{Lin}(A)\mathrm{Lin}(B)$, d.h. auch G^{-1} ist ein Homomorphismus von Gruppen.
4. Zu $A \in \mathfrak{I}(n)$ ist $\mathrm{Lin}(A) \in \mathrm{Iso}(K^n, K^n)$ mit $\mathrm{Lin}(A^{-1}) = (\mathrm{Lin}(A))^{-1}$; umgekehrt ist für $f \in \mathrm{Iso}(K^n, K^n)$ die Matrix $\mathrm{Mat}(f)$ invertierbar mit $(\mathrm{Mat}(f))^{-1} = \mathrm{Mat}(f^{-1})$. ♢

Satz 2.30.

Für $A \in \mathfrak{M}(n,n)$ sind äquivalent:

1. A ist invertierbar.
2. Es gibt ein $B \in \mathfrak{M}(n,n)$ mit $AB = \mathrm{Id}$.
3. Es gibt ein $B \in \mathfrak{M}(n,n)$ mit $BA = \mathrm{Id}$.
4. $\mathrm{rg}(A) = n$.
5. $\mathrm{Lin}(A)$ ist bijektiv.

Beweis.

(1) $\Rightarrow$ (2) und (2) $\Rightarrow$ (3) sind klar nach Definition.

(3) $\Rightarrow$ (4): Es gilt $\mathrm{rg}(B) = \dim \mathrm{Bild}\mathrm{Lin}(B) \leq \dim K^n = n$, also
$$n = \mathrm{rg}(\mathrm{Id}) = \mathrm{rg}(BA) = \dim \mathrm{Bild}\mathrm{Lin}(BA) = \dim \mathrm{Bild}(\mathrm{Lin}(B)\mathrm{Lin}(A)) \leq \dim \mathrm{Bild}\mathrm{Lin}(A) = \mathrm{rg}(A),$$
d.h. $\mathrm{rg}(A) \geq n$. Außerdem gilt trivialerweise $\mathrm{rg}(A) \leq n$, insgesamt also $\mathrm{rg}(A) = n$

(4) $\Rightarrow$ (5): Wegen $\mathrm{rg}(A) = \dim \mathrm{Bild}\mathrm{Lin}(A)$ ist $\mathrm{Lin}(A)$ surjektiv, also bijektiv.

(5) $\Rightarrow$ (1): Setze $f = \mathrm{Lin}(A) \Rightarrow f \circ f^{-1} = \mathrm{id}$, d.h. $\mathrm{Mat}(f \circ f^{-1}) = \mathrm{Mat}(f)\mathrm{Mat}(f^{-1}) = \mathrm{Mat}(\mathrm{id}) = \mathrm{Id}$. Wegen $\mathrm{Mat}(\mathrm{Lin}(A)) = A$ folgt $A^{-1} = \mathrm{Mat}(f^{-1})$. □

Bemerkung 2.31. (Berechnung der inversen Matrix)
Sei $A \in \mathfrak{M}(n,n)$ invertierbar mit Inverser $A^{-1} = B$. Dann löst die i-te Spalte $B^{(i)}$ von B das lineare Gleichungssystem

$$A_1 \circ x = 0, \qquad ..., \qquad A_i \circ x = 1, \qquad ..., \qquad A_n \circ x = 0,$$

d.h. das Lösen der Matrixgleichung $AX = \mathrm{Id}$ ist äquivalent zum simultanen Lösen der n inhomogenen, linearen Gleichungssysteme $Ax = e^{(i)}$.

Die inverse Matrix kann also mit dem Gauß-Algorithmus bestimmt werden, indem man die $(n \times 2n)$-Matrix $(A|\mathrm{Id})$ mit elementaren Zeilenumformungen auf die Gestalt $(\mathrm{Id}|B)$ bringt; B ist dann gerade die inverse Matrix zu A. ◊

Beispiel 2.32.
Wir suchen die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & 1 & 1 \\ 1 & -2 & -2 & -1 \end{pmatrix}.$$

Diese berechnen wir nach dem Schema

$$\begin{array}{rrrr|rrrr}
1 & 3 & 2 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 3 & 0 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & -2 & -2 & -1 & 0 & 0 & 0 & 1 \\
\hline
1 & 3 & 2 & 1 & 1 & 0 & 0 & 0 \\
0 & 2 & 1 & -2 & 1 & -1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & -1 & 0 \\
0 & 5 & 4 & 2 & 1 & 0 & 0 & -1 \\
\hline
1 & 0 & \frac{1}{2} & 4 & -\frac{1}{2} & \frac{3}{2} & 0 & 0 \\
0 & 1 & \frac{1}{2} & -1 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} & -1 & 0 \\
0 & 0 & \frac{3}{2} & 7 & \frac{3}{2} & \frac{5}{2} & 0 & -1 \\
\hline
1 & 0 & 0 & 3 & -1 & 1 & 1 & 0 \\
0 & 1 & 0 & -2 & 0 & -1 & 1 & 0 \\
0 & 0 & 1 & 2 & 1 & 1 & -2 & 0 \\
0 & 0 & 0 & 4 & -3 & 1 & 3 & -1 \\
\hline
1 & 0 & 0 & 0 & \frac{5}{4} & \frac{1}{4} & -\frac{5}{4} & \frac{3}{4} \\
0 & 1 & 0 & 0 & -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\
0 & 0 & 1 & 0 & \frac{5}{2} & \frac{1}{2} & -\frac{7}{2} & \frac{1}{2} \\
0 & 0 & 0 & 1 & -\frac{3}{4} & \frac{1}{4} & \frac{3}{4} & -\frac{1}{4}
\end{array}$$

Also ist A^{-1} gegeben als

$$A^{-1} = \begin{pmatrix} \frac{5}{4} & \frac{1}{4} & -\frac{5}{4} & \frac{3}{4} \\ -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ \frac{5}{2} & \frac{1}{2} & -\frac{7}{2} & \frac{1}{2} \\ -\frac{3}{4} & \frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \end{pmatrix}.$$ ◊

Definition 2.33.
Unter einer Teilmatrix einer gegebenen Matrix versteht man das Ergebnis der Streichung gewisser Zeilen und Spalten.

Beispiel 2.34.
Die (4×5)-Matrix

$$\left(\begin{array}{cccccc} & & \times & & \times & \times \\ & 1 & 2 & 3 & 1 & 1 \\ & 0 & 7 & 1 & 8 & 5 \\ \times & 2 & 3 & 0 & 1 & 9 \\ & 7 & 8 & 0 & 6 & 2 \end{array} \right)$$

ergibt nach Streichung der Spalten 3, 5, 6 und der Zeile 3 die (2×3)-Teilmatrix

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \\ 7 & 0 \end{pmatrix}. \qquad \Diamond$$

Lemma 2.35.
Der Rang einer Matrix A ist das größte n, so dass es eine invertierbare $n \times n$-Teilmatrix von A gibt.

Beweis.
Seien $r = \mathrm{rg}(A)$ und n maximal mit inveriertbarer $(n \times n)$-Teilmatrix B. Zu zeigen: $r = n$.

1. $\mathrm{rg}(B) = n$. Verlängern der Zeilen von B liefert n linear unabhängige Zeilen von A, also $r \geq n$.
2. Wegen $\mathrm{rg}(A) = r$ gibt es r linear unabhängige Zeilen von A. Streichung der restlichen Zeilen ergibt eine Teilmatrix A'. Wegen $\mathrm{rg}(A') = r$ hat A' r linear unabhängige Spalten. Streichung der restlichen Spalten von A' ergibt eine quadratische $(r \times r)$-Matrix A''; diese ist wegen $\mathrm{rg}(A'') = r$ invertierbar. Also $r \leq n$. $\square$

2.4. Transformationsmatrizen und Basiswechsel

Dieser Abschnitt behandelt einen Spezialfall der Darstellungsmatrizen: Die Basistransformationen. Hier wird die Identitätsabbildung bezüglich zwei verschiedener Basen eines Vektorraums dargestellt. Wir werden sehen, dass die Rücktransformation durch die inverse Matrix der Transformation beschrieben werden kann. Dies erlaubt es insbesondere, die Darstellungsmatrix eines Homomorphismus bezüglich zweier neuen Basen mittels der Darstellungsmatrix der alten Basen und der beidenTransformationsmatrizen auszudrücken.

Definition 2.36.
Seien V ein K-Vektorraum und $\mathfrak{V}, \mathfrak{V}'$ zwei Basen von V. Dann heißt $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}'}(\mathrm{id})$ die Transformationsmatrix oder Übergangsmatrix des Basiswechsels von $\mathfrak{V}$ zu $\mathfrak{V}'$.

Bemerkung 2.37.

1. Wegen
$$\mathrm{Mat}^{\mathfrak{V}'}_{\mathfrak{V}}(\mathrm{id})\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}'}(\mathrm{id}) = \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(\mathrm{id}) = \mathrm{Id}$$
ist $\mathrm{Mat}^{\mathfrak{V}'}_{\mathfrak{V}}(\mathrm{id})$ die zu $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}'}(\mathrm{id})$ inverse Matrix: $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}'}(\mathrm{id})^{-1} = \mathrm{Mat}^{\mathfrak{V}'}_{\mathfrak{V}}(\mathrm{id})$.
2. Sei W ein weiterer K-Vektorraum und seien $\mathfrak{W}, \mathfrak{W}'$ zwei Basen von W. Dann gilt:
$$\mathrm{Mat}^{\mathfrak{V}'}_{\mathfrak{W}'}(f) = \mathrm{Mat}^{\mathfrak{W}}_{\mathfrak{W}'}(\mathrm{id})\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(f)\mathrm{Mat}^{\mathfrak{V}'}_{\mathfrak{V}}(\mathrm{id}) = \mathrm{Mat}^{\mathfrak{W}}_{\mathfrak{W}'}(\mathrm{id})\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(f)\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}'}(\mathrm{id})^{-1}.$$

3. Für $\Psi_{\mathfrak{V}} \in \mathrm{Hom}(K^n, K^n)$ gilt $\Psi_{\mathfrak{V}}^{-1}(e^{(i)}) = v_i$, d.h. $\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(\Psi_{\mathfrak{V}}^{-1}) = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(\Psi_{\mathfrak{V}})^{-1} = (v_1, ..., v_n)$.

   Zugleich ist $\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{id}) = \mathrm{Mat}(\Psi_{\mathfrak{E}} \circ \mathrm{id} \circ \Psi_{\mathfrak{V}}^{-1}) = (v_1, ..., v_n)$.

   Damit können wir $\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{V}}(f)$ im Falle $V = K^n$, $W = K^m$ konkret ausrechnen:

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{V}}(f) = \mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{E}}(\mathrm{id})\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f)\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{id}) = (w_1, ..., w_m)^{-1}(f(e^{(1)}), ..., f(e^{(n)}))(v_1, ..., v_n).$$

   Anstatt über die Basiswechsel $\mathfrak{V} \to \mathfrak{E}$ und $\mathfrak{E} \to \mathfrak{W}$ zu argumentieren, können wir auch folgendermaßen vorgehen:

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{V}}(f) = \mathrm{Mat}(\Psi_{\mathfrak{W}} \circ f \circ \Psi_{\mathfrak{V}}^{-1}) = \mathrm{Mat}(\Psi_{\mathfrak{W}})\mathrm{Mat}(f)\mathrm{Mat}(\Psi_{\mathfrak{V}}^{-1}) = \mathrm{Mat}(\Psi_{\mathfrak{W}})\mathrm{Mat}(f)\mathrm{Mat}(\Psi_{\mathfrak{V}})^{-1}. \quad \Diamond$$

Beispiel 2.38.

Betrachte den Homomorphismus $f : \mathbb{R}^3 \to \mathbb{R}^2$, gegeben durch $(x_1, x_2, x_3) \mapsto (x_1 + x_2, 2x_3 - x_1)$. Dann ist

$$\mathrm{Mat}(f) = (f(e^{(1)}), f(e^{(2)}), f(e^{(3)})) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

Wir stellen f nun bzgl. der Basen $\mathfrak{V} = ((1;0;-1),(1;1;1),(1;0;0))$ und $\mathfrak{W} = ((0;1),(1;0))$:

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{V}}(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix}. \quad \Diamond$$

Definition 2.39.

Zwei Matrizen $A, B \in \mathfrak{M}(n,n)$ heißen ähnlich (in Zeihen: $A \approx B$), wenn es eine invertierbare Matrix $P \in \mathfrak{M}(n,n)$ gibt mit $B = PAP^{-1}$.

Bemerkung 2.40.

$\approx$ definiert eine Äquivalenzrelation, d.h. es gilt:

1. $A \approx A$: es ist $A = \mathrm{Id}A\mathrm{Id}^{-1}$;
2. $A \approx B \Rightarrow B \approx A$: $B = PAP^{-1} \Rightarrow A = P^{-1}BP$;
3. $A \approx B$ und $B \approx C \Rightarrow A \approx C$: $B = PAP^{-1}$ und $C = QBQ^{-1} \Rightarrow C = (QP)A(QP)^{-1}$. $\Diamond$

Bemerkung 2.41.

Besonders nützlich ist es, zu einem gegebenen Endomorphismus $f \in \mathrm{Hom}(V, V)$ eine Basis $\mathfrak{V}$ von V zu finden, so dass $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ Diagonalgestalt hat, d.h.

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda$$

für gewisse $\lambda_1, ..., \lambda_n \in K$.

Sei nämlich $A \in \mathfrak{M}(n,n)$ mit $A = P\Lambda P^{-1}$ für eine gewisse Transformationsmatrix $P \in \mathfrak{I}(n)$, dann gelten:

1. $A^k = (P\Lambda P^{-1})^k = P\Lambda^k P^{-1} = P\mathrm{diag}(\lambda_1^k, ..., \lambda_n^k)P^{-1}$, wobei A^k das k-fache Produkt von A bezeichnet;
2. $\exp(A) = \exp(P\Lambda P^{-1}) = P\exp(\Lambda)P^{-1} = P\mathrm{diag}(\exp(\lambda_1), ..., \exp(\lambda_n))P^{-1}$, wobei $\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$.
3. Seien $\lambda_1, ..., \lambda_n \geq 0$, dann ist $\sqrt{(A)} = \sqrt{P\Lambda P^{-1}} = P\sqrt{\Lambda}P^{-1} = P\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})P^{-1}$, wobei $\sqrt{A}$ diejenige Matrix $B \in \mathfrak{M}(n,n)$ bezeichnet, für die gilt $B^2 = A$.

   Beachte: Im Allgemeinen ist weder klar, dass solch ein B existiert, noch, dass B eindeutig ist.

Im Kapitel über Eigenwerttheorie gehen wir näher auf dieses Thema ein. $\Diamond$

2.5. Die Determinantenabbildung

Die Determinante ist die Auswertung eines Polynoms n-ten Grades in den Einträgen einer quadratischen Matrix $A \in \mathfrak{M}(n,n)$. Sie enthält viele Informationen über A; beispielsweise ist $\det A \neq 0$ genau dann, wenn $A \in \mathfrak{I}(n)$. Leider ist das Auswerten der Determinantenabbildung recht aufwendig. Wir lernen in diesem Abschnitt verschiedene Darstellungsformen der Determinante kennen und Techniken, Determinanten zusammengesetzter Matrizen einfach zu bestimmen.

Bemerkung 2.42.

Wir definieren zu $A = (a, b; c, d) \in \mathfrak{M}(2,2)$: $\det A = ad - bc$. Dann gelten:

1. det ist linear in jeder Zeile, d.h.

$$\begin{aligned}
\det\begin{pmatrix} \alpha a + \beta a' & \alpha b + \beta b' \\ c & d \end{pmatrix} &= \alpha \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \beta \det\begin{pmatrix} a' & b' \\ c & d \end{pmatrix}, \\
\det\begin{pmatrix} a & b \\ \alpha c + \beta c' & \alpha d + \beta d' \end{pmatrix} &= \alpha \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \beta \det\begin{pmatrix} a & b \\ c' & d' \end{pmatrix},
\end{aligned}$$

   denn

$$\begin{aligned}
\det\begin{pmatrix} \alpha a + \beta a' & \alpha b + \beta b' \\ c & d \end{pmatrix} &= (\alpha a + \beta a')d - (\alpha b + \beta b')c \\
&= \alpha(ad - bc) + \beta(a'd - b'c) \\
&= \alpha \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \beta \det\begin{pmatrix} a' & b' \\ c & d \end{pmatrix};
\end{aligned}$$

   analog für die zweite Zeile.

2. det ist alternierend, d.h. sind zwei Zeilen von A identisch, dann ist $\det A = 0$:

$$\det\begin{pmatrix} a & b \\ a & b \end{pmatrix} = a \cdot b - b \cdot a = 0.$$

3. det ist normiert, d.h. $\det \mathrm{Id} = 1$:

$$\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1.$$

Beachte: Im Allgemeinen gilt $\det(\alpha A + \beta B) \neq \alpha \det A + \beta \det B$. ◊

Definition 2.43.

Eine Abbildung $\det : \mathfrak{M}(n,n) \to K$ heißt Determinantenabbildung, falls det linear in jeder Zeile, alternierend und normiert ist.

$\det A$ heißt die Determinante von A.

Bemerkung 2.44.

1. $\det A$ ändert sich nicht, wenn zu einer beliebigen Zeile A_i ein Vielfaches αA_j einer anderen Zeile addiert wird:

$$\det\begin{pmatrix} \vdots \\ A_i + \alpha A_j \\ \vdots \\ A_j \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \end{pmatrix} + \alpha \underbrace{\det\begin{pmatrix} \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \end{pmatrix}}_{=0} = \det A.$$

2. $\det A$ ändert sein Vorzeichen, wenn zwei Zeilen von A vertauscht werden:

$$0 = \det \begin{pmatrix} \vdots \\ A_i + A_j \\ \vdots \\ A_i + A_j \\ \vdots \end{pmatrix} = \underbrace{\det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \end{pmatrix}}_{=0} + \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \end{pmatrix} + \underbrace{\det \begin{pmatrix} \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \end{pmatrix}}_{=0},$$

also $\det(...; A_i; ...; A_j; ...) = -\det(...; A_j; ...; A_i; ...)$. ◊

Satz 2.45.
$\det : \mathfrak{M}(n,n) \to K$, iterativ gegeben durch

$$\det A = \sum_{k=1}^{n} (-1)^{k-1} a_{k1} \det A_{k1},$$

definiert eine Determinantenabbildung.

Hierbei bezeichnet $A_{k1} \in \mathfrak{M}(n-1, n-1)$ diejenige Matrix, die aus A nach Streichung der k-ten Zeile und der ersten Spalte entsteht.

Beweis.
Wir zeigen per Induktion über n, dass det multilinear, alternierend und normiert ist. Der Induktionsanfang $n = 1$ ist trivial; den Fall $n = 2$ haben wir eingangs schon gezeigt. Gelte die Behauptung also für $n - 1$.

1. Zur Multilinearität:

   Seien
   $$A = \begin{pmatrix} A_1 \\ \vdots \\ \alpha A_i' + \beta A_i'' \\ \vdots \\ A_n \end{pmatrix}, \qquad A' = \begin{pmatrix} A_1 \\ \vdots \\ A_i' \\ \vdots \\ A_n \end{pmatrix}, \qquad A'' = \begin{pmatrix} A_1 \\ \vdots \\ A_i'' \\ \vdots \\ A_n \end{pmatrix}.$$
   Dann gilt $\det A = \alpha \det A' + \beta \det A''$:
   $$\begin{aligned} \det A &= (-1)^{i-1}(\alpha a_{i1}' + \beta a_{i1}'') \det A_{i1} + \sum_{k \neq i} (-1)^{k-1} a_{k1} \det A_{k1} \\ &= \alpha (-1)^{i-1} a_{i1}' \det A_{i1} + \beta (-1)^{i-1} a_{i1}'' \det A_{i1} \\ &\quad + \alpha \sum_{k \neq i} (-1)^{k-1} a_{k1} \det A_{k1}' + \beta \sum_{k \neq i} (-1)^{k-1} a_{k1} \det A_{k1}'' \\ &= \alpha \sum_{k=1}^{n} a_{k1}' \det A_{k1}' + \beta \sum_{k=1}^{n} (-1)^{k-1} a_{k1}'' \det A_{k1} \\ &= \alpha \det A' + \beta \det A''. \end{aligned}$$

2. Zum Alternieren:

   Sei $A_i = A_j$ mit $i \neq j$, etwa $i < j$. Dann
   $$\det A = (-1)^{i-1} a_{i1} \det A_{i1} + (-1)^{j-1} a_{j1} \det A_{j1} + \sum_{k \neq i,j} (-1)^{k-1} a_{k1} \det A_{k1}.$$
   Wegen $(A_{k1})_i = (A_{k1})_j$ ist $\det A_{k1} = 0$. Also wird A_{i1} durch Vertauschen der Zeile $(A_{i1})_j$ mit den $(j-i)-1$ darüber liegenden Zeilen in die Matrix A_{j1} überführt. Also gilt $\det A_{i1} = (-1)^{j-i-1} \det A_{j1}$.

Damit folgt:

$$\begin{aligned}\det A &= (-1)^{i-1} a_{i1} (-1)^{j-i-1} \det A_{j1} + (-1)^{j-1} a_{j1} \det A_{j1} \\ &= (-1)^{j} a_{j1} \det A_{j1} + (-1)^{j-1} a_{j1} \det A_{j1} \\ &= 0.\end{aligned}$$

3. Zur Normiertheit:

   Sei $A = \mathrm{Id}(n)$. Dann gilt $a_{k1} = 0 \Leftrightarrow k = 1$. Also ist $\det \mathrm{Id}(n) = 1 \det \mathrm{Id}(n-1) = 1$. □

Satz 2.46. (Hauptsatz der Determinantentheorie)
Es gibt genau eine Abbildung $\det : \mathfrak{M}(n,n) \to K$, die multilinear in jeder Zeile, alternierend und normiert ist.

Beweis.
Sei $\mathrm{d} : \mathfrak{M}(n,n) \to K$ eine Determinantenabbildung. Zu zeigen ist, dass dann bereits $\mathrm{d} = \det$ gilt. Wir führen dazu wieder eine Induktion über n. Der Fall $n = 1$ ist klar, denn $\mathrm{d}(a) = a\mathrm{d}(1) = a = \det(a)$. Sei jetzt also $n > 1$; die Eindeutigkeit von $\det : \mathfrak{M}(n-1, n-1) \to K$ wird vorausgesetzt.

Sei $A \in \mathfrak{M}(n,n)$. Aus $A_1 = a_{11} e^{(1)} + \ldots + a_{1n} e^{(n)}$ folgt mit der Multilinearität: $\mathrm{d}(A) = a_{11} D_1 + \ldots + a_{1n} D_n$ mit

$$D_j = \mathrm{d}\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ a_{21} & \ldots & \ldots & a_{2j} & \ldots & \ldots & a_{2n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \ldots & \ldots & a_{nj} & \ldots & \ldots & a_{nn} \end{pmatrix} = \mathrm{d}\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ a_{21} & \ldots & \ldots & 0 & \ldots & \ldots & a_{2n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \ldots & \ldots & 0 & \ldots & \ldots & a_{nn} \end{pmatrix},$$

denn

$$\begin{aligned}\mathrm{d}(A) &= \mathrm{d}\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \\ &= \mathrm{d}\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} + \ldots + \mathrm{d}\begin{pmatrix} 0 & 0 & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \\ &= a_{11}\mathrm{d}\begin{pmatrix} 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} + \ldots + a_{n1}\mathrm{d}\begin{pmatrix} 0 & 0 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.\end{aligned}$$

Wir drücken nun D_j mit Hilfe von $\det$ aus.

Seien $B' \in \mathfrak{M}(n-1, j-1)$ und $B'' \in \mathfrak{M}(n-1, n-j)$. Dann ist $B = (B', B'') \in \mathfrak{M}(n-1, n-1)$. Setze $\Delta_j : \mathfrak{M}(n-1, n-1) \to K$ mit

$$\Delta(B) = \mathrm{d}\begin{pmatrix} 0 \quad \ldots \quad 0 & 1 & 0 \quad \ldots \quad 0 \\ & 0 & \\ B' & \vdots & B'' \\ & 0 & \end{pmatrix}.$$

Δ_j ist Zeilen-multilinear und alternierend; weiter gilt

$$\Delta_j(\mathrm{Id}(n-1)) = \mathrm{d}\left(\begin{array}{ccc|c|ccc} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ \hline 1 & & 0 & 0 & & & \\ & \ddots & & \vdots & & 0 & \\ 0 & & 1 & 0 & & & \\ \hline & & & 0 & 1 & & 0 \\ & 0 & & \vdots & & \ddots & \\ & & & 0 & 0 & & 1 \end{array}\right) = (-1)^{j-1}\Delta_j.$$

Also ist $(-1)^{j-1}\Delta_j$ eine Determinantenabbildung; nach Induktionsvoraussetzung ist $(-1)^{j-1}\Delta_j = \det$, d.h. $\Delta_j = (-1)^{j-1}\det$. Damit gilt:

$$D_j = \Delta_j(A_{1j}) = (-1)^{j-1}\det A_{1j} \qquad \Longrightarrow \qquad \mathrm{d}(A) = \sum_{l=1}^{n}(-1)^{l-1}a_{1l}\det A_{1l}.$$

Also gilt insbesondere

$$\det A = \sum_{l=1}^{n}(-1)^{l-1}a_{1l}\det A_{1l} = \mathrm{d}(A).$$

□

Korollar 2.47.
Für $A \in \mathfrak{M}(n,n)$ gilt:

$$\det A = \det A^{\mathrm{T}}.$$

Beweis.
Per Induktion über n. Der Fall $n = 1$ ist trivial. Gelte die Behauptung also für $n-1$. Dann folgt:

$$\det A = \sum_{l=1}^{n}(-1)^{l-1}a_{1l}\det A_{1l} = \sum_{l=1}^{n}(-1)^{l-1}a_{1l}\det A_{1l}^{\mathrm{T}} = \sum_{l=1}^{n}(-1)^{l-1}a_{1l}\det A_{1l}^{\mathrm{T}} = \det A^{\mathrm{T}}.$$

Beachte dabei: Mit $A^{\mathrm{T}} = (a'_{ij})$ gilt $a'_{l1} = a_{1l}$. □

Bemerkung 2.48. (Entwicklungsformeln)
Damit erhalten wir die folgenden Darstellungsformeln:

1. Entwicklung nach der i-ten Zeile:

$$\det A = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij};$$

2. Entwicklung nach der j-ten Spalte:

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}.$$

◊

Korollar 2.49. (Regel von Sarus)
Speziell für Matrizen im $\mathfrak{M}(3,3)$ gilt:

$$\det\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = (ab'c'' + a'b''c + a''bc') - (ab''c' + a'bc'' + a''b'c).$$

Beweis.
Durch Nachrechnen:

$$\begin{aligned}\det\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} &= a\det\begin{pmatrix} b' & c' \\ b'' & c'' \end{pmatrix} - a'\det\begin{pmatrix} b & c \\ b'' & c'' \end{pmatrix} + a''\det\begin{pmatrix} b & c \\ b'' & c'' \end{pmatrix} \\ &= a(b'c'' - b''c') - a'(bc'' - b''c) + a''(bc' - b'c) \\ &= (ab'c'' + a'b''c + a''bc') - (ab''c' + a'bc'' + a''b'c).\end{aligned}$$

□

Satz 2.50.
Für $A \in \mathfrak{M}(n,n)$ gilt:

$$\det A \neq 0 \qquad \Longleftrightarrow \qquad A \in \mathfrak{I}(n).$$

Beweis.

1. Sei A nicht invertierbar, d.h. die Zeilen von A sind nicht linear unabhängig: Es gibt ein i mit $A_i = \sum\limits_{l \neq i} \alpha_l A_l$ für ein $\alpha \in K$.
   Addiert man zur i-ten Zeile das $-\alpha_l$-fache der l-ten Zeile für alle $l \neq i$, so ist die i-te Zeile der entstehenden Matrix B die Nullzeile. Entwickelt man die Determinante von B nun nach der i-ten Zeile, so folgt $0 = \det B = \det A$.
2. Per Induktion über n. Der Fall $n = 1$ ist klar. Gelte „$\Leftarrow$" also für $n-1$. Angenommen, $\det A$ wäre 0. Dann
$$0 = \det \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} = \sum_{j=1}^{n} (-1)^{j-1} a_{1j} \det A_{1j}.$$
   Weiter gilt für $i \geq 2$:
$$0 = \det \begin{pmatrix} A_i \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \sum_{j=1}^{n} (-1)^{j-1} a_{ij} \det A_{ij}.$$
   Insgesamt ergibt sich also:
$$\sum_{j=1}^{n} (-1)^{j-1} a_{1j} \det A_{1j} + ... + \sum_{j=1}^{n} (-1)^{j-1} a_{nj} \det A_{1j} = \sum_{j=1}^{n} \underbrace{\left((-1)^{j-1} \det A_{1j}\right)}_{=:\alpha_j} A^{(j)} = 0.$$
   Da nach Voraussetzung A invertierbar ist, gilt $\operatorname{rg}(A) = n$, d.h. alle $\alpha_j = 0$ und damit alle $\det A_{1j} = 0$. Andererseits folgt aus $\operatorname{rg}(A) = n$ auch, dass $\operatorname{rg}(A_2, ..., A_n)^{\mathrm{T}} = n-1$, d.h. es gibt $n-1$ linear unabhängige Spalten von $(A_2, ..., A_n)^{\mathrm{T}}$ und damit auch eine invertierbare Teilmatrix A_{1j}. Nach Induktionsvoraussetzung folgt $\det A_{1j} \neq 0$, ein Widerspruch. □

Korollar 2.51.
Für $A \in \mathfrak{M}(n,n)$ gilt:

$$\det A \det B = \det AB.$$

Beweis.

Der Fall $\det B = 0$ ist klar, denn wäre $\det AB \neq 0$, d.h. AB invertierbar, dann gäbe es ein $C \in \mathfrak{M}(n,n)$ mit $\mathrm{Id} = C(AB) = (CA)B$, d.h. entgegen der Annahme wäre auch B invertierbar.

Gelte also jetzt $\det B \neq 0$. Wir definieren die Abbildung

$$\mathrm{d} : \mathfrak{M}(n,n) \to K, \qquad A \mapsto \mathrm{d}(A) = \frac{\det AB}{\det B}.$$

Dies ist eine Determinantenabbildung:

1. d ist Zeilen-multilinear, denn wegen
$$\begin{pmatrix} A_1 \\ \vdots \\ \alpha A_i' + \beta A_i'' \\ \vdots \\ A_n \end{pmatrix} B = \begin{pmatrix} A_1 \circ B \\ \vdots \\ \alpha A_i' \circ B + \beta A_i'' \circ B \\ \vdots \\ A_n \circ B \end{pmatrix}$$

ist $\det AB$ linear in der i-ten Zeile von A, also auch $\mathrm{d}(A) = \dfrac{\det AB}{\det B}$.

2. d ist alternierend, denn

$$\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} B = \begin{pmatrix} A_1 \circ B \\ \vdots \\ A_i \circ B \\ \vdots \\ A_j \circ B \\ \vdots \\ A_n \circ B \end{pmatrix},$$

also $\det AB = 0$, falls $i \neq j$.

3. d ist normiert:

$$\mathrm{d}(\mathrm{Id}) = \frac{\det \mathrm{Id}B}{\det B} = \frac{\det B}{\det B} = 1.$$

Mit der Eindeutigkeit der Determinantenabbildung folgt: $\mathrm{d} = \det$, also

$$\det A = \frac{\det AB}{\det B} \qquad \Longleftrightarrow \qquad \det A \det B = \det AB.$$

$\det : \mathfrak{I}(n) \to (K\backslash\{0\}, \cdot)$ ist also ein Gruppenhomomorphismus. □

Korollar 2.52.
Für $A \in \mathfrak{M}(n,n)$ gilt:

$$(\det A)^{-1} = \det A^{-1}.$$

Beweis.
Es ist

$$\det A \det A^{-1} = \det(AA^{-1}) = \det(\mathrm{Id}) = 1 \qquad \Longrightarrow \qquad \det A^{-1} = (\det A)^{-1}.$$

Determinantenberechnung und Invertierung vertauschen also miteinander. □

Korollar 2.53.
Für $A, B \in \mathfrak{M}(n,n)$ mit $A \approx B$ gilt:

$$\det A = \det B.$$

Beweis.
Es ist $A = PBP^{-1}$ für ein $P \in \mathfrak{I}(n)$. Dann gilt:

$$\det A = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \frac{\det P \det B}{\det P} = \det B.$$

Die Determinante ist also invariant unter Ähnlichkeitstransformationen. □

Bemerkung 2.54.
Zu $f \in \mathrm{Hom}(V,V)$ können wir daher definieren $\det f = \det \mathfrak{M}_{\mathfrak{V}}^{\mathfrak{V}}(f)$, wobei $\mathfrak{V}$ eine beliebige Basis von V bezeichnet: Nach obigem Korollar hängt $\det f$ nicht von der Wahl von $\mathfrak{V}$ ab. ◊

Korollar 2.55.
Seien $A \in \mathfrak{M}(n,n)$ und $B \in \mathfrak{M}(m,m)$. Dann gilt:

$$\det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} = \det A \det B.$$

Beweis.
Wir führen eine Induktion über n.

1. Zum Induktionsanfang: Entwicklung nach der ersten Spalte liefert

$$\det \begin{pmatrix} a & * \\ 0 & B \end{pmatrix} = (-1)^0 a \det B = \det(a) \det B.$$

2. Zum Induktionsschritt: Gelte die Behauptung für $n-1$. Dann

$$\begin{aligned} \det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} &= \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix}_{i1} \\ &= \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det A_{i1} \det B \\ &= \left(\sum_{i=1}^{n} (-1)^{i-1} \det A_{i1} \right) \det B \\ &= \det A \det B. \end{aligned}$$

□

Korollar 2.56.
Für $a_1, ..., a_n \in K$ gilt:

$$\det \begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix} = \prod_{i=1}^{n} a_i.$$

Beweis.
Durch induktives Anwenden des Korollars. □

Satz 2.57. (Regel von Cramer)
Sei $A \in \mathfrak{M}(n,n)$ mit $\det A \neq 0$.

Dann besitzt das Lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{11}X_1 & + & ... & + & a_{1n}X_n & = & b_1 \\ & \vdots & & \vdots & & \vdots & \\ a_{n1}X_1 & + & ... & + & a_{nn}X_n & = & b_n \end{array}$$

genau eine Lösung, nämlich

$$X_l = \frac{\det(A^{(1)} ... A^{(l-1)} B A^{(l+1)} ... A^{(n)})}{\det(A)},$$

wobei $b = (b_1; ...; b_n) \in K^n$.

Beweis.

Klar: Aus $Ax = b$ folgt $x = A^{-1}Ax = A^{-1}b$, d.h. das System ist eindeutig lösbar. Wir haben außerdem bereits gesehen, dass $Ax = b$ genau dann eindeutig lösbar ist, wenn $\operatorname{rg}(A) = n$ gilt, was wiederum äquivalent zu $A \in \mathfrak{I}(n)$ ist.

Bezeichne $x = (x_1; ...; x_n) \in K^n$ die Lösung, dann ist $x_1 A^{(1)} + ... + x_n A^{(n)} = B$. Wegen der Linearität in der l-ten Spalte gilt dann:

$$\begin{aligned}\det(A^{(1)}...A^{(l-1)}|B|A^{(l+1)}...A^{(n)}) &= x_1 \det(A^{(1)}...|A^{(1)}|...A^{(n)}) + ... + x_l \det(A^{(1)}...|A^{(l)}|...A^{(n)}) \\ &\quad + ... + x_n \det(A^{(1)}...|A^{(n)}|...A^{(n)}) \\ &= x_l \det(A).\end{aligned}$$

Also ist $x_l = \dfrac{\det(A^{(1)}...|B|...A^{(n)})}{\det A}$. □

Bemerkung 2.58. (Berechnung der inversen Matrix)

Sei $A \in \mathfrak{I}(n)$. Wir suchen $X \in \mathfrak{I}(n)$ mit $AX = \mathrm{Id}$. Für die Spalten von X gilt:

$$A_k \circ X^{(j)} = \begin{cases} 1 & k = j \\ 0 & \text{sonst} \end{cases}$$

Mit der Regel von Cramer folgt:

$$x_{ij} = \frac{\det(A^{(1)}...A^{(i-1)}e^{(j)}A^{(i+1)}...A^{(n)})}{\det A}.$$

Es gilt also:

$$A^{-1} = \frac{1}{\det(A)}((-1)^{i+j}\det(A_{ij}))^{1\le i\le n}_{1\le j\le n},$$

denn mit Entwicklung nach der i-ten Spalte ist $\det(A^{(1)}...|e^{(j)}|...A^{(n)}) = (-1)^{i+j}\det A_{ij}$.

Beachte hierbei: $\det(A_{ij}))^{1\le i\le n}_{1\le j\le n}$ bezeichnet die Matrix, deren Einträge sich als die Determinanten der Streichmatrizen A_{ij} berechnen. Das Berechnen von A^{-1} auf diese Weise ist also sehr aufwändig. ◊

Bemerkung 2.59.

Für $A \in \mathfrak{M}(2,2)$ gilt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Insbesondere ist A genau dann invertierbar, wenn $ad - bc \neq 0$. ◊

3. Einführung in die Ringtheorie

In diesem Kapitel führen wir eine neue algebraische Struktur ein: den Ring. Ringe haben viel mit Körpern gemeinsam, unterscheiden sich aber dahingehend von diesen, dass nicht zu jedem Element die Existent eines multiplikativ inversen Elements gefordert wird. Außerdem fordern wir für die Multiplikation nicht das Kommutativgesetz. Dies hat zur Folge, dass die Gleichung $bx = a$ nicht unbedingt für alle Elemente a, b des Rings mit $b \neq 0$ lösbar ist.

3.1. Ringe

Der Prototyp eines Rings ist der der ganzen Zahlen $\mathbb{Z}$. Dieser besitzt einige für Ringe wünschenswerte zusätzliche Eigenschaften wie die Kommutativität und die Nullteilerfreiheit. Solche Ringe lassen sich zu Körpern vervollständigen, indem ihre Grundmenge um die multiplikativ Inversen erweitert werden – aus $\mathbb{Z}$ wird so beispielsweise der Körper der reellen Zahlen $\mathbb{R}$.

Definition 3.1.

Ein Ring $(R, +, \cdot)$ ist ein Tripel, bestehend aus einer Grundmenge R zusammen mit zwei Verknüpfungen $+$ und $\cdot$, genannt Addition und Multiplikation, geschrieben als Summe bzw. Produkt, so dass die folgenden Bedingungen erfüllt sind:

1. R bildet bezüglich der Addition eine abelsche Gruppe.
2. Die Multiplikation ist assoziativ, das heißt es gilt $(ab)c = a(bc)$ für alle $a, b, c \in R$.
3. Die Multiplikation besitzt ein neutrales Element, das heißt es gibt ein $e \in R$ mit $ae = a = ea$ für alle $a \in R$.
4. Die Distributivgesetze gelten, das heißt $(a+b)c = ac+bc$ und $c(a+b) = ca+cb$ für alle $a, b, c \in R$.

Bemerkung 3.2.

Sind e, e' neutrale Elemente bezüglich der Multiplikation, so gilt $e = ee' = e'$. Daher gibt es genau ein neutrales Element bezüglich der Multiplikation, welches wir mit 1 bezeichnen. Ebenso ist natürlich das neutrale Element 0 bezüglich der Addition eindeutig bestimmt. $\Diamond$

Definition 3.3.

Ein Ring R heißt Integritätsring, wenn folgende Bedingungen erfüllt sind:

1. R ist nicht trivial, d.h. es gilt $0 \neq 1$.
2. R ist kommutativ, das heißt es gilt $ab = ba$ für alle $a, b \in R$.
3. R ist nullteilerfrei, das heißt es gilt $ab = 0 \Rightarrow (a = 0 \text{ oder } b = 0)$ für alle $a, b \in R$.

Beispiel 3.4.

1. Jeder Körper ist ein Integritätsring.
2. Jede additive abelsche Gruppe R eines Körpers, die die Eins des Körpers enthält und multiplikativ abgeschlossen ist (das heißt $R \cdot R \subseteq R$) bildet einen Integritätsring.

   Damit sind z.B. Ring der ganzen Zahlen $\mathbb{Z} \subseteq \mathbb{R}$ und $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$ Integritätsringe.
3. Die $(n \times n)$-Matrizen über einem festen Körper K bilden einen Ring, den Matrizenring. Dieser ist i.A. nicht kommutativ.
4. Die Endomorphismen eines festen Vektorraums bilden einen Ring, den Endomorphismenring, wenn man zwei Endomorphismen f und g multipliziert durch Hintereinanderausführung: $fg = f \circ g$. $\Diamond$

Bemerkung 3.5.

In einem Integritätsring gilt die Kürzungsregel

$$ab = ac \qquad \Longrightarrow \qquad b = c$$

für beliebige $a, b, c \in R$ mit $a \neq 0$. $\Diamond$

Bemerkung 3.6.

Der „Prototyp“ aller Integritätsringe ist $\mathbb{Z}$. Was diesem Ring zum Körper fehlt, ist die Existenz multiplikativ Inverser, denn die Gleichung $bx = a$ besitzt nicht für alle $a \in \mathbb{Z}$ und $b \in \mathbb{Z}\backslash\{0\}$ eine Lösung $x \in \mathbb{Z}$.

Wir wollen die ganzen Zahlen durch Hinzufügung nicht allzu vieler Elemente zu einem größeren Zahlbereich mit schönen Eigenschaften erweitern, in dem solche Gleichungen stets eine Lösung besitzen. In der Tat kann man die ganzen Zahlen zum Körper $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z},\ b \neq 0\}$ der rationalen Zahlen $\mathbb{Q}$ erweitern.

Man kann sogar aus einem beliebigen Integritätsring R einen Körper K konstruieren, der grob gesprochen gerade aus Brüchen von Elementen von R besteht: $\Diamond$

Bemerkung 3.7.

1. Sei R ein Integritätsring. Dann gibt es genau einen Körper K mit $R \subseteq K$, so dass die Verknüpfungen von R sich durch Einschränkung der entsprechenden Verknüpfungen von K ergeben und alle Elemente x von K von der Form $x = \frac{a}{b} = ab^{-1}$ mit $a, b \in R$ und $b \neq 0$ sind.

   Wir nennen K den Quotientenkörper von R.

2. Die einzigen Integritätsringe sind also die additiven abelschen Gruppen eines Körpers, die multiplikativ abgeschlossen sind und das Einselement enthalten.

3. Es gibt nur eine Möglichkeit, im Quotientenkörper zu rechnen, denn man prüft anhand der Körperaxiome sofort nach, dass für jeden Körper K und alle $a, b, c, d \in K$ mit $c, d \neq 0$ die Bruchregeln gelten:
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \qquad \text{und} \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$
   Insbesondere gilt dies für alle $a, b, c, d \in R$ mit $c, d \neq 0$; da alle Elemente von K von der Form $\frac{a}{b}$ für $a \in R$ und $b \in R \backslash \{0\}$ sind, ist damit festgelegt, wie man in K rechnet: Die Ausdrücke $ad + bc$, bd und ac rechnet man ja in R aus. $\diamondsuit$

3.2. Polynomringe

Als wir zu den reellen Zahlen künstlich ein neues Element i hinzugefügt haben, welches die Beziehung $i^2 = -1$ erfüllt, haben wir die komplexen Zahlen erhalten, welche wieder einen Körper bilden. Formal haben wir dies bewerkstelligt, indem wir für $a, b \in \mathbb{R}$ die komplexe Zahl $a + bi$ durch das Paar $(a, b) \in \mathbb{R}^2$ modelliert haben.

Wir wollen nun zu den reellen Zahlen (oder allgemeiner zu einem Körper K) künstlich ein neues Element X hinzufügen, welches **keine** Beziehungen erfüllt. Auf diese Weise werden wir den *Polynomring* über den reellen Zahlen erhalten, der einen Integritätsring, aber keinen Körper bildet.

Formal modellieren wir für $a_0, ..., a_n \in \mathbb{R}$ das *Polynom* $a_n X^n + a_{n-1} X^{n-1} + ... + a_1 X + a_0$ durch die Folge $(a_0, a_1, ..., a_{n-1}, a_n, 0, 0, 0, ...)$. Dies ist nur ein bisschen aufwendiger als bei den komplexen Zahlen.

Satz 3.8.

Sei K ein Körper. Dann gibt es einen Integritätsring $K[X]$ mit folgenden Eigenschaften:

1. $K[X]$ umfasst den Körper K und setzt dessen Verknüpfungen fort.
2. X ist ein Element von $K[X]$.
3. Die Elemente von $K[X]$ sind alle von der Form
$$f = \sum_{i=0}^{n} a_i X^i = a_n X^n + a_{n-1} X^{n-1} + ... + a_1 X + a_0$$
   mit $a_0, ..., a_n \in K$.
4. Für alle $a_0, ..., a_n \in K$ gilt:
$$a_n X^n + a_{n-1} X^{n-1} + ... + a_1 X + a_0 = 0 \qquad \Longrightarrow \qquad a_0 = ... = a_n = 0.$$

Bemerkung 3.9.

1. Seien $f, g \in K[X]$, etwa $f = a_n X^n + ... + a_0$ und $g = b_m X^m + ... + b_0$. Dann gilt:
$$a_n X^n + ... + a_0 = b_n X^n + ... + b_0 \qquad \Longleftrightarrow \qquad n = m \text{ und } a_0 = b_0, ..., a_n = b_n.$$

2. Die Addition auf $K[X]$ ist komponentenweise gegeben: Sei $m \leq n$, dann gilt
$$\begin{aligned} f + g = (a_0 + b_0) + (a_1 + b_1)X + (a_2 + b_2)X^2 + \cdots \\ + (a_m + b_m)X^m + a_{m+1}X^{m+1} + \cdots + a_n X^n. \end{aligned}$$

Die Multiplikation ist komplizierter:

$$fg = a_0b_0 + (a_0b_1 + a_1b_0)X + (a_0b_2 + a_1b_1 + a_2b_0)X^2 + \cdots \\ + (a_{m-1}b_n + a_mb_{n-1})X^{m+n-1} + a_mb_nX^{m+n}.$$

In kompakter Schreibweise:

$$f + g = \sum_{i=0}^{n}(a_i + b_i)X^i, \qquad fg = \sum_{i=0}^{m+n}\Big(\sum_{j+k=i} a_jb_k\Big)X^i,$$

wobei wir setzen $b_{m+1} = \ldots = b_n = 0$. ◊

Definition 3.10.

Sei K ein Körper. $K[X]$ heißt der Polynomring über K, die Elemente von $K[X]$ nennt man Polynome in der Unbestimmten oder Variablen X mit Koeffizienten aus K.

Falls $a_n \neq 0$, so sagt man, $f = a_nX^n + \ldots + a_0$ habe den Grad n und schreibt $\deg(f) = n$. a_n bezeichnet man als den Leitkoeffizienten von f.

Bemerkung 3.11.

Seien $f, g \in K[X]$ mit $\deg(f) = n$ und $\deg(g) = m$. Dann hat fg den Grad $m + n$. ◊

Bemerkung 3.12.

Der Mangel eines Integritätsrings gegenüber einem Körper ist, dass man bei Division zweier Elemente zu einem größeren Rechenbereich, dem Quotientenkörper, übergehen muss. Die hier betrachteten Polynomringe verfügen allerdings (wie die ganzen Zahlen) auch über eine Art von Division, die sich innerhalb des Ringes abspielt. Dies ist eine Division mit Rest.

In den ganzen Zahlen besagt die Tatsache, dass sich $m \in \mathbb{Z}$ durch $n \in \mathbb{Z}\backslash\{0\}$ *mit Rest* dividieren lässt, gerade, dass es einen *Quotienten* $q \in \mathbb{Z}$ und einen $r \in \mathbb{Z}$ gibt mit $m = qn + r$ und $|r| < |n|$.

In $K[X]$ lautet die entsprechende Aussage wie folgt: ◊

Satz 3.13.

Sei K ein Körper und seien $f, g \in K[X]$, $g \neq 0$.

Dann gibt es $q, r \in K[X]$ mit $f = qg + r$, so dass $r = 0$ gilt (das heißt *die Division geht auf*) oder r einen kleineren Grad hat als g.

Beweis.

Wir führen eine Induktion über den Grad von f, wobei wir (nur für die Dauer dieses Beweises) dem Nullpolynom $0 \in K[X]$ den Grad -1 zuordnen.

Als Induktionsanfang betrachten wir den Fall, dass f einen kleineren Grad als g hat. Dann leisten $q = 0$ und $r = f$ das Gewünschte.

Im Induktionsschritt sei also nun $k = \deg(f) - \deg(g) > 0$. Bezeichne a den Leitkoeffizienten von f und b den von g. Dann ist $f - \frac{a}{b}X^kg \in K[X]$ ein Polynom von kleinerem Grad als f. Wir können also auf dieses Polynom die Induktionsvoraussetzung anwenden und erhalten $p, r \in K[X]$ mit $f - \frac{a}{b}X^kg = pg + r$, so dass r einen kleineren Grad als g hat.

Setzt man nun $q = \frac{a}{b}X^k + p$, so leisten q und r das Gewünschte. □

Beispiel 3.14.

Wie bei der Division in $\mathbb{Z}$ lässt sich die Polynomdivision mit Rest in Stufenform durchführen. Die Koeffizienten von q erhalten wir dabei folgendermaßen:

Seien $\deg(f) = n > \deg(g) = m$.

Algorithmus 3.15. (Polynomdivision)

1. Wähle $q_{n-m} = \frac{f_n}{g_m}$ und subtrahiere das Polynom $q_{n-m}g$ von f.

   Wegen $q_{n-m}g = f_n X^n + \cdots$ hat das Ergebnis f' der Subtraktion einen Grad $n' \leq n-1$.

2. Wähle $q_{n'-m} = \frac{f'_{n'}}{g_m}$ und subtrahiere das Polynom $q_{n'-m}g$ von f'.

   Wegen $q_{n'-m}g = f'_{n'} X^{n'} + \cdots$ hat das Ergebnis f'' der Subtraktion einen Grad $n'' \leq n'-1$.

3. u.s.w. bis $\deg(f'^{\cdots'}) < \deg g$.

   Wähle dann $r = f'^{\cdots'}$, dann gelten $f = qg + r$ und $\deg(r) < \deg(g)$.

Seien etwa $f = X^3 + 4X^2 + 2X + 1$ und $g = X - 1$, dann ist

$$\begin{array}{r} (1X^3 + 4X^2 + 2X + 1) : (X-1) = (1X^2 + 5X + 7), \\ -\underline{(1X^3 - 1X^2 + 0X + 0)} \\ (5X^2 + 2X + 1) \\ -\underline{(5X^2 - 5X + 1)} \\ (7X + 1) \\ -\underline{(7X - 7)} \\ (8) \end{array}$$

d.h. $q = X^2 + 5X + 7$ und $r = 8$. ◊

Definition 3.16.

Sei $f \in K[X]$ ein Polynom mit Koeffizienten $a_0, ..., a_n \in K$. Sei weiter $x \in K$. Da $a_0, ..., a_n$ durch f eindeutig bestimmt sind, können wir

$$f : K \to K, \qquad f(x) = a_n x^n + ... + a_0$$

definieren. Man spricht davon, dass x in f eingesetzt wird. f heißt dann Polynomfunktion.

Ein $x \in K$ heißt Nullstelle von f, falls $f(x) = 0$ gilt.

Bemerkung 3.17.

1. Seien $f, g \in K[X]$. Dann gelten
$$(f+g)(x) = f(x) + g(x) \qquad \text{und} \qquad (fg)(x) = (f(x))(g(x)).$$

2. Sei $x \in K$ eine Nullstelle von $f \in K[X]$. Dann gibt es ein $q \in K[X]$ mit $f = q(X - x)$. Man spricht davon, den Linearfaktor $(X - x)$ von f *abzuspalten.*

3. Ein Polynom aus $K[X] \backslash \{0\}$ vom Grad n hat somit höchstens n verschiedenen Nullstellen in K.

4. Sei x eine Nullstelle von $f \neq 0$, so dass f durch $(X - x)^k$ ohne Rest geteilt werden kann, aber nicht durch $(X - x)^{k+1}$. Dann heißt k die Vielfachheit der Nullstelle x.

5. Kässt sich f schreiben als $f = (X - x_1) \cdots (X - x_n)$ ($x_1, ..., x_n$ müssen nicht paarweise verschieden sein), dann sagt man, f faktorisiert über K. Beispielsweise faktorisiert $f = X^2 + 2X + 1$ über $\mathbb{R}$, denn $f = (X+1)^2$ – dabei ist $x = -1$ zweifache Nullstelle von f – während $f = X^2 + 1$ zwar über $\mathbb{C}$ faktorisiert: $f = (X + i)(X - i)$ mit den beiden einfachen Nullstellen $x = i$ und $x = -i$, aber nicht über $\mathbb{R}$.

6. Es besteht ein Unterschied zwischen einer Polynomfunktion und dem zugehörigen Polynom. Nimmt man für K etwa den zweielementigen Körper $\mathbb{Z}_2$, so gibt es sicher unendlich viele Polynome über K. Es gibt allerdings nur vier Polynomfunktionen $K \to K$, da es überhaupt nur vier Funktionen $K \to K$ gibt.

7. Ist K ein *unendlicher* Körper und sind f, g Polynome in $K[X]$, die die selbe Polynomfunktion darstellen (das heißt $f(x) = g(x)$ für alle $x \in K$), dann gilt bereits $f = g$. ◊

Definition 3.18.

Sei $f \in K[X]$ ein Polynom mit Koeffizienten $a_0, ..., a_n \in K$. Weiter sei V ein K-Vektorraum. Wir fassen $\mathrm{Hom}(V,V)$ als Endomorphismenring von V mit der Verkettung als Multiplikation.

Da $a_0, ..., a_n$ durch f eindeutig bestimmt sind, können wir zu $\varphi \in \mathrm{Hom}(V,V)$ definieren

$$f(\varphi) = a_n\varphi^n + ... + a_0 \mathrm{id}_V \in \mathrm{Hom}(V,V).$$

Man spricht davon, dass der Endomorphismus φ in das Polynom f eingesetzt wird.

Bemerkung 3.19.

Seien $f, g \in K[X]$, V ein K-Vektorraum und $\varphi \in \mathrm{Hom}(V,V)$. Dann gelten:

$$(f+g)(\varphi) = f(\varphi) + g(\varphi) \qquad \text{und} \qquad (fg)(\varphi) = (f(\varphi))(g(\varphi)),$$

d.h. die Einsetzung vertauscht mit den Ringoperationen. $\Diamond$

4. Eigenwerttheorie

In diesem Kapitel befassen wir uns damit, zu Endomorphismen Basen zu finden, bezüglich denen diese besonders einfach dargestellt werden können. Optimal ist es, eine Basis aus Vektoren zu finden, deren Bilder Streckungen um gewisse Faktoren entsprechen. Man spricht in diesem Kontext von Eigenvektoren und Eigenwerten. Hat man eine solche Basis gefunden, dann hat die zugehörige Darstellungsmatrix Diagonalgestalt, d.h. alle Einträge, die nicht auf der Diagonalen liegen, sind Null.

Im ersten Abschnitt führen wir Eigenwerte und Eigenvektoren ein. Der zweite Abschnitt beschäftigt sich damit, die Eigenwerte konkret als Nullstellen geeigneter Polynome auszurechnen – den sogenannten charakteristischen Polynomen. Im dritten Abschnitt wollen wir Endomorphismen bzw. quadratische Matrizen auf Diagonalgestalt oder wenigstens Dreiecksgestalt bringen. Wir werden sehen, dass das über $\mathbb{R}$ nicht immer möglich ist, über $\mathbb{C}$ hingegen schon. Im vierten Abschnitt beweisen wir den Satz von Hamilton-Cayley, der besagt, dass ein Endomorphismus, eingesetzt in sein charakteristisches Polynom, stets Null ist.

4.1. Eigenwerte und Eigenvektoren

Eigenvektoren eines Endomorphismus f sind von Null verschiedene Vektorraumelemente, die von f gestreckt werden, d.h. $f(v) = \lambda v$ für ein gewisses $\lambda \in K$, den Eigenwert zu v. Wir erhalten in diesem Abschnitt als einfache Folgerung unserer Überlegungen zur Basistransformation den Hauptsatz der Eigenwerttheorie: Endomorphismen haben bzgl. einer Basis aus Eigenvektoren Diagonalgestalt. Anschließend stellen wir einen Zusammenhang zwischen einem Eigenwert λ von f und dem um λ verschobenen Endomorphismus $f - \lambda$ her.

Definition 4.1.

Seien V ein K-Vektorraum und $f \in \mathrm{Hom}(V,V)$.

$\lambda \in K$ heißt ein Eigenwert von f, wenn es ein $v \in V\backslash\{0\}$ gibt mit $f(v) = \lambda v$. v heißt dann der zu λ gehörige Eigenvektor.

Bemerkung 4.2.

Seien V ein K-Vektorraum und $f \in \mathrm{Hom}(V,V)$. Offenbar sind äquivalent:

1. $v \in V$ ist ein Eigenvektor zu f mit zugehörigem Eigenwert $\lambda \in K$.
2. Es gibt ein $v \in V\backslash\{0\}$ mit $f(v) = \lambda v$.
3. Es gibt ein $v \in V\backslash\{0\}$ mit $(f - \lambda)v = 0$.
4. $\mathrm{Kern}(f - \lambda) \neq \{0\}$.
5. $\det(f - \lambda) = 0$. $\Diamond$

Beispiel 4.3. (Streckungen)

1. Betrachte eine Streckung um einen festen Faktor: $f(v) = \lambda v$ im K^n. Die zugehörige Darstellungsmatrix ist
$$\mathrm{Mat}(f) = \mathrm{Mat}^{\mathfrak{E}}_{\mathfrak{E}}(f) = \lambda \mathrm{Id}.$$
Offenbar ist λ der einzige Eigenwert von f und jedes $v \neq 0$ ist ein Eigenvektor.

2. Wir können auch längs mehrerer Achsen strecken: Seien $v_1, ..., v_n$ linear unabhängig mit $f(v_i) = \lambda_i v_i$. Dann ist
$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda.$$
Ist $\lambda_i \neq 0$, dann ist die Gerade $Kv_i = \{\alpha v_i \mid \alpha \in K\}$ invariant unter f, d.h. $f(Kv_i) = Kv_i$. Ist dagegen $\lambda_i = 0$, dann wird Kv_i auf den Nullpunkt abgebildet. $\diamondsuit$

Satz 4.4. (Hauptsatz der Eigenwerttheorie)

Seien V ein K-Vektorraum, $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V und $f \in \mathrm{Hom}(V, V)$. Dann gilt:

$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) \text{ hat Diagonalgestalt} \quad \Longleftrightarrow \quad \mathfrak{V} \text{ besteht aus Eigenvektoren.}$$

Beweis.

1. Habe $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ Diagonalgestalt, d.h. $A = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda$. Dann gilt:
$$f(v_i) = a_{1i}v_1 + \cdots + a_{ni}v_n = a_{ii}v_i = \lambda_i v_i,$$
d.h. v_i ist ein Eigenvektor zu f mit zugehörigem Eigenwert λ_i.

2. Sei umgekehrt $\mathfrak{V}$ eine Basis aus Eigenvektoren zu f, dann besitzt die i-te Spalte von $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ die Darstellung
$$\mathrm{Mat}(\Psi_{\mathfrak{V}} \circ f \circ \Psi_{\mathfrak{V}}^{-1})e^{(i)} = (\Psi_{\mathfrak{V}} \circ f \circ \Psi_{\mathfrak{V}}^{-1})(e^{(i)}) = (\Psi_{\mathfrak{V}} \circ f)(v_i) = \Psi_{\mathfrak{V}}(\lambda_i v_i) = \lambda_i \Psi_{\mathfrak{V}}(v_i) = \lambda_i e^{(i)},$$
d.h. $A = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda$ hat Diagonalgestalt. $\square$

Bemerkung 4.5.

1. Der Nullvektor $v = 0$ erfüllt zu jedem $f \in \mathrm{Hom}(V, V)$ und $\lambda \in K$ stets $f(v) = \lambda v$, ist aber niemals ein Eigenvektor.

2. $\lambda = 0$ kann aber ein Eigenwert sein. Beispielsweise ist 0 einziger Eigenwert der Nullabbildung $f = 0$ und jedes $v \neq 0$ ist ein Eigenvektor zu f

3. Sei $A \in \mathfrak{M}(n, n)$. Wir nennen $\lambda \in K$ einen Eigenwert zu A mit zugehörigem Eigenvektor $v \in K^n \backslash \{0\}$, wenn $Av = \lambda v$ gilt. Offenbar ist v genau dann ein Eigenvektor zu A, wenn v ein Eigenvektor zu $\mathrm{Lin}(A)$ ist und die zugehörigen Eigenwerte sind identisch. $\diamondsuit$

Lemma 4.6.

Seien $v_1, ..., v_m$ Eigenvektoren von $f \in \mathrm{Hom}(V, V)$ mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$. Dann sind $v_1, ..., v_m$ linear unabhängig. Insbesondere besitzt f höchstens n Eigenwerte.

Beweis.

Per Induktion über m. Der Fall $m = 1$ ist klar, da $v_1 \neq 0$ nach Definition eines Eigenvektors. Gelte die Behauptung also für $m - 1$ und sei $\alpha_1 v_1 + ... + \alpha_m v_m = 0$ mit gewissen $\alpha_1, ..., \alpha_m \in K$. Wir müssen zeigen, dass alle $\alpha_i = 0$. Wegen $\lambda_m \alpha_1 v_1 + ... + \lambda_m \alpha_m v_m = 0$ und $0 = f(0) = \alpha_1 \lambda_1 v_1 + ... + \alpha_m \lambda_m v_m$ sowie der Induktionsvoraussetzung ist

$$0 = \alpha_1 \underbrace{(\lambda_m - \lambda_1)}_{\neq 0} v_1 + ... + \alpha_{m-1} \underbrace{(\lambda_m - \lambda_{m-1})}_{\neq 0} v_{m-1} \quad \Longrightarrow \quad \alpha_1 = ... = \alpha_{m-1} = 0.$$

Also $0 = \alpha_m v_m$ mit $v_m \neq 0$, d.h. auch $\alpha_m = 0$. Damit sind $v_1, ..., v_m$ linear unabhängig. $\square$

Beispiel 4.7. (Drehungen)
Sei $\theta \in [0, 2\pi)$. Wir drehen den Einheitsvektor $e^{(1)} = (1; 0) \in \mathbb{R}^2$ um den Winkel θ gegen den Uhrzeigersinn. Dann sind die Koeffizienten des neuen Vektors $f(e^{(1)})$ gegeben durch

$$\begin{cases} \cos(\theta) = \dfrac{\text{Ankathete}}{\text{Hypothenuse}} = \dfrac{x_1}{1} \\[2ex] \sin(\theta) = \dfrac{\text{Gegenkathete}}{\text{Hypothenuse}} = \dfrac{x_2}{1} \end{cases} \qquad \Longrightarrow \qquad f(e^{(1)}) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

Entsprechend erhalten wir für die Drehung von $e^{(2)} = (0; 1) \in \mathbb{R}^2$ um θ:

$$f(e^{(2)}) = \begin{pmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}.$$

Die Drehung eines beliebigen Vektors $v = (v_1; v_2) \in \mathbb{R}^2$ um θ wird also vermittelt durch

$$f(v) = \mathrm{Mat}(f)v = (f(e^{(1)}), f(e^{(2)}))v = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathrm{Rot}(\theta)v.$$

Im Fall $\theta = 0$ ist $f = \mathrm{id}$, d.h. $\lambda = 1$ ist einziger Eigenwert von f und jedes $v \in \mathbb{R}^2 \backslash \{0\}$ ist ein Eigenvektor zu λ. Ist dagegen $\theta = \pi$, dann ist $f = -\mathrm{id}$; $\lambda = -1$ ist einziger Eigenwert und wieder ist jedes $v \in \mathbb{R}^2 \backslash \{0\}$ Eigenvektor. Für jede andere Wahl von θ ist keine Gerade des $\mathbb{R}^2$ invariant unter f und es gibt somit keine Eigenwerte oder Eigenvektoren:

$$\det(f - \lambda) = \lambda^2 - 2\lambda\cos(\theta) + 1 = 0 \qquad \Longleftrightarrow \qquad \lambda = \cos(\theta) \pm \sqrt{\sin^2(\theta)},$$

d.h. es gibt genau dann ein $\lambda \in \mathbb{R}$ mit $\det(f - \lambda) = 0$, wenn $\sin(\theta) = 0$, d.h. $\theta = 0$ oder $\theta = \pi$.

f ist für jedes θ bijektiv, denn nach dem Additionstheorem ist $\det(\mathrm{Mat}(f)) = \cos^2(\theta) + \sin^2(\theta) = 1$.

Die Umkehrabbildung entspricht der Drehung um den Winkel $-\theta$; die zugehörige Darstellungsmatrix ist

$$\mathrm{Mat}(f)^{-1} = \mathrm{Mat}(f^{-1}) = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = \mathrm{Mat}(f)^{\mathrm{T}}.$$

Insbesondere ist die Länge eines Vektors invariant unter f, denn sei $M = \mathrm{Mat}(f)$, dann

$$||f(v)||^2 = \langle f(v), f(v)\rangle = \langle Mv, Mv\rangle = \langle v, M^{\mathrm{T}}Mv\rangle = \langle v, M^{-1}Mv\rangle = \langle v, v\rangle = ||v||^2. \qquad \Diamond$$

Satz 4.8.
Seien $A, B \in \mathfrak{M}(n, n)$ mit $A \approx B$. Dann gilt

$$\lambda \text{ ist ein Eigenwert von } A \qquad \Longleftrightarrow \qquad \lambda \text{ ist ein Eigenwert von B.}$$

Beweis.
Gelte $B = CAC^{-1}$ für ein $C \in \mathfrak{I}(n)$ und sei λ ein Eigenwert von A, dann

$$C(A - \lambda\mathrm{Id})C^{-1} = CAC^{-1} - C(\lambda\mathrm{Id})C^{-1} = CAC^{-1} - \lambda\mathrm{Id} = B - \lambda\mathrm{Id},$$

also

$$0 = \det(A - \lambda\mathrm{Id}) = \det(C(A - \lambda\mathrm{Id})C^{-1}) = \det(B - \lambda\mathrm{Id}).$$

Die Eigenwerte einer Darstellungsmatrix sind also invariant unter Basistransformationen. $\square$

Definition 4.9.
$A \in \mathfrak{M}(n, n)$ heißt diagonalisierbar (über K), falls es ein $\Lambda \approx A$ gibt mit $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$.

Bemerkung 4.10.

Anders ausgedrückt: Eine Matrix A ist genau dann diagonalisierbar, wenn der K^n eine Basis aus zum Endomorphismus $\mathrm{Lin}(A) \in \mathrm{Hom}(K^n, K^n)$ gehörigen Eigenvektoren besitzt. $\Diamond$

Bemerkung 4.11. (Berechnung von Eigenwerten und Eigenvektoren)

1. Die Eigenwerte eines Endomorphismus $f \in \mathrm{Hom}(V, V)$ bzw. einer Matrix $A \in \mathfrak{M}(n, n)$ sind stets als Nullstellen eines Polynoms vom Grad n gegeben, dessen Koeffizienten selbst Polynome in den Koeffizienten von $\mathrm{Mat}_{\mathfrak{B}}^{\mathfrak{B}}(f)$ bzw. von A sind. Sie sind also *implizit* gegeben durch die Bedingung $f(\lambda) = \sum \alpha_i \lambda^i = 0$. Für $n \leq 4$ lässt sich diese Gleichung über $\mathbb{C}$ in Wurzeln auflösen; für $n \geq 5$ existiert keine solche allgemeine Auflösung. Allerdings gibt es nach dem Fundamentalsatz der Algebra stets n komplexe Nullstellen.

2. Sei ein Eigenwert $\lambda \in K$ gegeben, dann erhalten wir die Menge der zugehörigen Eigenvektoren durch Lösen des homogenen Linearen Gleichungssystems

$$(A - \lambda \mathrm{Id})x = 0 \qquad \Longleftrightarrow \qquad \left\{ \begin{array}{ccccccccc} (a_{11} - \lambda)x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & (a_{22} - \lambda)x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\ & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & (a_{nn} - \lambda)x_n & = & 0 \end{array} \right.$$

Wegen der Forderung $\det(A - \lambda \mathrm{Id}) = 0$ besitzt dieses System mindestens eine nicht-triviale Lösung $x = (x_1; ...; x_n) \in K^n$. $\Diamond$

Beispiel 4.12.

Die Matrix $A = (1, 2; -1, 1) \in \mathfrak{M}(2, 2)$ ist über $\mathbb{C}$, aber nicht über $\mathbb{R}$ diagonalisierbar:

$$\det(A - \lambda \mathrm{Id}) = \lambda^2 - 2\lambda + 3 = 0 \qquad \Longleftrightarrow \lambda = 1 \pm \sqrt{-2} = 1 \pm i\sqrt{2} \in \mathbb{C},$$

d.h. A hat keine reellen Eigenwerte und ist damit über $\mathbb{R}$ nicht diagonalisierbar.

Seien v_1 Eigenvektor zu $\lambda_1 = 1 + i\sqrt{2}$ und v_2 Eigenvektor zu $\lambda_2 = 1 - i\sqrt{2}$. Weiter sei $f = \mathrm{Lin}(A)$, d.h. $f(v) = Av$. Dann gilt:

$$A = \mathrm{Mat}(f) = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f) \approx \mathrm{Mat}(\Psi_{\mathfrak{B}} \circ f \circ \Psi_{\mathfrak{B}}^{-1}) = \mathrm{Mat}_{\mathfrak{B}}^{\mathfrak{B}}(f) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Die zugehörigen Eigenvekoren erhalten wir mit dem Gauß-Algorithmus durch

$$\begin{pmatrix} 1 - \lambda_1 & 2 \\ -1 & 1 - \lambda_1 \end{pmatrix} = \begin{pmatrix} -i\sqrt{2} & 2 \\ -1 & -i\sqrt{2} \end{pmatrix} \qquad \longrightarrow \qquad \begin{pmatrix} 1 & +i\sqrt{2} \\ 0 & 0 \end{pmatrix};$$
$$\begin{pmatrix} 1 - \lambda_2 & 2 \\ -1 & 1 - \lambda_2 \end{pmatrix} = \begin{pmatrix} +i\sqrt{2} & 2 \\ -1 & +i\sqrt{2} \end{pmatrix} \qquad \longrightarrow \qquad \begin{pmatrix} 1 & -i\sqrt{2} \\ 0 & 0 \end{pmatrix};$$

eine Basis aus Eigenvektoren ist also gegeben durch $\mathfrak{B} = (v_1, v_2) = \left(\begin{pmatrix} -i\sqrt{2} \\ 1 \end{pmatrix}, \begin{pmatrix} i\sqrt{2} \\ 1 \end{pmatrix} \right)$. $\Diamond$

Beispiel 4.13.

Sei $\theta \in \mathbb{R}$ beliebig. Für die Matrix

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

gilt $\det(A - \lambda \mathrm{Id}) = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$, sie besitzt also die Eigenwerte ± 1 und wir erhalten:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathfrak{B} = \left(\begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2} + \pi) \\ \sin(\frac{\theta}{2} + \pi) \end{pmatrix} \right).$$

$A = A(\theta)$ ist also für jede Wahl von θ diagonalisierbar mit gleicher Diagonalmatrix $\Lambda = \mathrm{diag}(1, -1)$; diese entspricht einer Spiegelung an der x_1-Achse. $\Diamond$

Beispiel 4.14. (Spiegelungen)
Sei $v = (v_1; v_2) \in \mathbb{R}^2 \backslash \{0\}$, dann bildet die Spiegelung $f \in \mathrm{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ v auf sich selbst und den zu v orthogonalen Vektor $w = (v_2; -v_1)$ auf $-w = (-v_2; v_1)$ ab.

Zur Basis $\mathfrak{V} = (v, w)$ ist also $\Lambda = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(1, -1)$ und $\mathrm{Mat}(f) = (v, w)^{-1}\Lambda(v, w)$.

Sei etwa konkret $v = (1; 3)$, dann sind $w = (3, -1)$ und $(v, w)^{-1} = -\frac{1}{10}(-1, -3; -3, 1)$ und wir erhalten $\mathrm{Mat}(f) = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f) = \frac{1}{5}(-4, 3; 3, 4)$.

In der Tat sind $\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f)v = v$, $\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f)w = -w$ und $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)e^{(1)} = e^{(1)}$, $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)e^{(2)} = -e^{(2)}$. ◊

4.2. Das charakteristische Polynom

Wir wollen zu einem Endomorphiusmus f das Polynom χ_f, dessen Nullstellen die Eigenwerte von f sind, näher untersuchen. Wir werden feststellen, dass die Vielfachheit einer Nullstelle von χ_f, d.h. eines Eigenwertes, nicht mit der Dimension des zugehörigen Eigenraumes übereinstimmen muss. Präziser kann die Anzahl an linear unabhängigen Eigenvektoren zu einem Eigenwert kleiner sein als dessen Vielfachheit als Nullstelle von χ_f. Dies hat beispielsweise zur Folge, dass nicht jede Matrix über $\mathbb{C}$ diagonalisierbar ist, obschon χ_f über $\mathbb{C}$ nach dem Fundamentalsatz faktorisieren muss.

Definition 4.15.
Für $A \in \mathfrak{M}(n, n)$ heißt $\chi_A = \det(A - t) = \det(A - t\mathrm{Id})$ das charakteristische Polynom von A.

Bemerkung 4.16.

1. Ist $A = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ die Darstellungsmatrix eines Endomorphismus $f \in \mathrm{Hom}(V, V)$ bzgl. einer Basis $\mathfrak{V}$ von V, so heißt $\chi_f = \det(A - t)$ auch das charakteristische Polynom von f.

   Die Definition ist unabhängig von der gewählten Basis $\mathfrak{V}$, denn ist $\mathfrak{V}'$ eine weitere Basis von V, dann gilt $\det(\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) - t) = \det(\mathrm{Mat}_{\mathfrak{V}'}^{\mathfrak{V}'}(f))$ für alle $t \in K$.

2. Die Eigenwerte von f bzw. von A sind genau die Nullstellen des charakteristischen Polynoms von χ_f bzw. χ_A in K.

3. Jeder Endomorphismus des $\mathbb{R}^{2n+1}$ hat einen reellen Eigenwert (und damit auch einen Eigenvektor).

   χ_f hat nämlich den Grad $2n + 1$, also gemäßdem Zwischenwertsatz eine Nullstelle in $\mathbb{R}$.

4. Sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum, dann besitzt χ_f genau n Nullstellen und f damit n Eigenwerte, die allerdings nicht verschieden sein müssen. Ist dies jedoch der Fall, dann gibt es auch eine Basis aus Eigenvektoren zu f und $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ ist diagnonalisierbar.

5. Für das charakteristische Polynom einer Matrix $A \in \mathfrak{M}(n, n)$ gilt:

   $$\chi_A(t) = \det(A - t) = (-1)^n t^n + (-1)^{n-1}\mathrm{spur}(A)t^{n-1} + \cdots + \det(A),$$

   wobei $\mathrm{spur}(A) = a_{11} + \cdots + a_{nn}$ die Spur von A bezeichnet. Die ersten beiden und der letzte Koeffizient von χ_A sind also stets bekannt.

   Insbesondere ist die Spur einer Matrix invariant unter Ähnlichkeitstransformation. ◊

Beispiel 4.17.
Sei $A \in \mathfrak{M}(2, 2)$ gegeben durch $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$\begin{aligned} \chi_A(t) &= \det(A - t) \\ &= \det \begin{pmatrix} a - t & b \\ c & d - t \end{pmatrix} \\ &= (a - t)(d - t) - bc \\ &= t^2 - (a + d)t + (ad - bc) \\ &= (-1)^2 t^2 + (-1)^1 \mathrm{spur}(A) + \det(A). \end{aligned}$$ ◊

Lemma 4.18.
Sei $B = (b_{ij}) \in \mathfrak{M}_{K[t]}(n,n)$ eine Matrix mit Einträgen b_{ij} aus dem Polynomring $K[t]$.
Dann ist $\det(B) \in K[t]$ und es gilt:

$$\forall \lambda \in K : \det(B)(\lambda) = \det(B(\lambda)),$$

d.h. die Auswertung von Determinante und Polynom vertauscht.

Beweis.
Per Indultion über n. Der Fall $n = 1$ ist klar, denn $\det(b)(\lambda) = b(\lambda) = \det(b(\lambda))$. Gelte also die Behauptung für $n-1$. Nach der Entwicklungsformel und der Induktionsvoraussetzung gilt:

$$\det(B) = \sum_{k=1}^{n} (-1)^{k-1} \underbrace{b_{1k}}_{\in K[t]} \underbrace{\det B_{1k}}_{\in K[t]} \in K[t]$$

sowie

$$\begin{aligned}
\det(B)(\lambda) &= \left(\sum_{k=1}^{n} (-1)^{k-1} b_{1k} \det B_{1k} \right)(\lambda) \\
&= \sum_{k=1}^{n} (-1)^{k-1} b_{1k}(\lambda)(\det B_{1k})(\lambda) \\
&= \sum_{k=1}^{n} (-1)^{k-1} b_{1k}(\lambda) \underset{n-1}{\det} (B(\lambda))_{1k} \\
&= \det(B(\lambda)).
\end{aligned}$$

□

Satz 4.19.
Seien $f \in \mathrm{Hom}(V,V)$ und $\lambda \in K$ ein Eigenwert von f.
Dann ist der Eigenraum $\mathrm{Eig}(f,\lambda) = \{v \in V \mid f(v) = \lambda v\}$ ein Untervektorraum von V.

Beweis.

1. Wegen $f(0) = 0 = \lambda 0$ ist stets $0 \in \mathrm{Eig}(f,\lambda)$ (obschon 0 kein Eigenvektor ist).
2. Seien $v, w \in \mathrm{Eig}(f,\lambda)$, d.h. $f(v) = \lambda v$ und $f(w) = \lambda w$. Seien $\alpha, \beta \in K$. Dann gilt
$$f(\alpha v + \beta w) = \alpha f(v) + \beta f(w) = \alpha\lambda v + \beta\lambda w = \lambda(\alpha v + \beta w),$$
d.h. auch $\alpha v + \beta w$ liegt in $\mathrm{Eig}(f,\lambda)$.

□

Definition 4.20.
Seien $f \in \mathrm{Hom}(V,V)$ und $\lambda \in K$ ein Eigenwert von f.
Dann heißt die Vielfachheit $\mu_{\mathrm{alg}}(f,\lambda)$ von λ als Nullstelle von χ_f die algebraische Vielfachheit von λ.
$\mu_{\mathrm{geo}}(f,\lambda) = \dim \mathrm{Eig}(f,\lambda) = \dim \mathrm{Kern}(f-\lambda)$ heißt die geometrische Vielfachheit von λ.

Bemerkung 4.21.
Nach dem Fundamentalsatz der Algebra lässt sich jedes Polynom $f \in \mathbb{R}[X]$ oder allgemeiner $f \in \mathbb{C}[X]$ in Linearfaktoren zerlegen, d.h. es gibt paarweise verschiedene Nullstellen $\lambda_1, ..., \lambda_s \in \mathbb{C}$ und Zahlen $\nu_1, ..., \nu_s \in \mathbb{N}$ sowie ein $\gamma \in \mathbb{C}$ mit

$$f(X) = \gamma \prod_{i=1}^{s} (X - \lambda_i)^{\nu_i} = \gamma (X-\lambda_1)^{\nu_1} \cdots (X-\lambda_s)^{\nu_s}$$

und es gilt $\sum \nu_i = \nu_1 + \cdots + \nu_s = \deg(f)$. ◊

Lemma 4.22.
Seien V ein K-Vektorraum und $f \in \operatorname{Hom}(V,V)$. Dann ist $\mu_{\text{alg}}(f,\lambda) \geq \mu_{\text{geo}}(f,\lambda)$.

Beweis.
Seien $\mu_{\text{geo}}(f,\lambda) = r$, d.h. es gibt eine Basis $\mathfrak{V}_\lambda = (v_1, ..., v_r)$ von $\operatorname{Eig}(f,\lambda)$. Wir ergänzen $\mathfrak{V}_\lambda$ zu einer Basis $\mathfrak{V} = (v_1, ..., v_r, v_{r+1}, ..., v_n)$ von V. Dann hat die Darstellungsmatrix von f bzgl. $\mathfrak{V}$ die Gestalt

$$\operatorname{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \begin{pmatrix} M(\lambda) & * \\ 0 & M' \end{pmatrix},$$

wobei $M(\lambda) = \operatorname{diag}(\lambda, ..., \lambda) \in \mathfrak{M}(r,r)$ und $M' \in \mathfrak{M}(n-r, n-r)$, d.h.

$$\begin{aligned} \chi_f(t) &= \det(f - t) \\ &= \det(\operatorname{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) - t\operatorname{Id}(n)) \\ &= \det M(\lambda - t) \det(M' - t\operatorname{Id}(n-r)) \\ &= (\lambda - t)^r \det(M' - t\operatorname{Id}(n-r)). \end{aligned}$$

Also $r \leq \mu_{\text{alg}}(f,\lambda)$. □

Bemerkung 4.23.
Es kann durchaus $\mu_{\text{alg}}(f,\lambda) > \mu_{\text{geo}}(f,\lambda)$ gelten. Betrachte etwa

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad \longrightarrow \qquad \chi_A(\lambda) = \det(A - \lambda\operatorname{Id}) = (1-\lambda)^2,$$

d.h. 1 ist zweifache Nullstelle von A: $\mu_{\text{alg}}(A,\lambda) = 2$. Allerdings gilt

$$A - \lambda\operatorname{Id} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \longrightarrow \qquad \operatorname{rg}(A - \lambda\operatorname{Id}) = 1,$$

d.h. $\dim \operatorname{Bild}(A - \lambda\operatorname{Id}) = 1$ und damit auch $\mu_{\text{geo}}(A,\lambda) = \dim \operatorname{Kern}(A - \lambda\operatorname{Id}) = 1$.

Also ist A weder über $\mathbb{R}$ noch über $\mathbb{C}$ diagonalisierbar.

Die geometrische Vielfachheit und die algebraische Vielfachheit des Eigenwerts einer Matrix $A \in \mathfrak{M}(n,n)$ sind dabei natürlich erklärt über

$$\mu_{\text{alg}}(A,\lambda) = \mu_{\text{alg}}(\operatorname{Lin}(A),\lambda), \qquad \mu_{\text{geo}}(A,\lambda) = \mu_{\text{geo}}(\operatorname{Lin}(A,\lambda). \qquad \Diamond$$

4.3. Diagonalisierung und Trigonalisierung

Wir lernen in diesem Abschnitt zwei Normalformen für die Darstellungsmatrizen von Endomorphismen kennen: Die Diagonal- und die Trigonalmatrizen. Beide haben viel mit den Eigenwerten des Endomorphismus zu tun: Diese stehen auf der Diagonalen der Normalform. Ob eine entsprechende Normalform existiert, hängt von der Reichhaltigkeit der Eigenräume ab, d.h. von μ_{alg} und μ_{geo}.

Definition 4.24.
Seien V ein K-Vektorraum und $f \in \operatorname{Hom}(V,V)$.

f heißt diagonalisierbar, falls es eine Basis $\mathfrak{V}$ von V gibt, so dass $\operatorname{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ Diagonalgestalt hat.

$A \in \mathfrak{M}(n,n)$ heißt diagonalisierbar, falls es ein invertierbares $C \in \mathfrak{I}(n)$ gibt, so dass CAC^{-1} Diagonalgestalt hat, d.h. falls A ähnlich zu einer Diagonalmatrix ist.

Bemerkung 4.25.
$A \in \mathfrak{M}(n,n)$ ist genau dann diagonalisierbar, wenn $\operatorname{Lin}(A) \in \operatorname{Hom}(K^n, K^n)$ diagonalisierbar ist:

1. Gelte $CAC^{-1} = \Lambda$. Wähle die Spalten von C^{-1} als Basis $\mathfrak{V}$, dann ist

$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(\mathrm{Lin}(A)) = \mathrm{Mat}(\Psi_{\mathfrak{V}} \circ \mathrm{Lin}(A) \circ \Psi_{\mathfrak{V}}^{-1}) = \mathrm{Mat}(\Psi_{\mathfrak{V}})\mathrm{Mat}(\mathrm{Lin}(A))\mathrm{Mat}(\Psi_{\mathfrak{V}}^{-1}) = CAC^{-1} = \Lambda.$$

2. Sei umgekehrt $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(\mathrm{Lin(A)}) = \Lambda$. Setze $C = \mathrm{Mat}^{\mathfrak{E}}_{\mathfrak{V}}(\mathrm{id})$, dann ist

$$CAC^{-1} = \mathrm{Mat}^{\mathfrak{E}}_{\mathfrak{V}}(\mathrm{id})\mathrm{Mat}^{\mathfrak{E}}_{\mathfrak{E}}(\mathrm{Lin}(A))\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{E}}(\mathrm{id}) = \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(\mathrm{Lin}(A)) = \Lambda. \qquad \Diamond$$

Satz 4.26.
Seien V ein K-Vektorraum und $f \in \mathrm{Hom}(V, V)$. Dann sind äquivalent:

1. V besitzt eine Basis $\mathfrak{V}$ aus Eigenvektoren zu f.
2. f ist diagonalisierbar.
3. χ_f zerfällt in Linearfaktoren und $\mu_{\mathrm{alg}}(f, \lambda) = \mu_{\mathrm{geo}}(f, \lambda))$.
4. Sind $\lambda_1, ..., \lambda_k$ die verschiedenen Eigenwerte von f, so gilt $V = \mathrm{Eig}(f, \lambda_1) \oplus ... \oplus \mathrm{Eig}(f, \lambda_k)$.

Beweis.

1. Sei $\mathfrak{V}$ eine Basis aus Eigenvektoren zu f, dann besitzt $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ nach dem Hauptsatz der Eigenwerttheorie Diagonalgestalt.
2. Sei $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \Lambda$. Dann ist $\chi_f = \chi_\Lambda) = (\lambda_1 - t)^{\nu_1} \cdots (\lambda_s - t)^{\nu_s}$. Außerdem gilt für $1 \leq i \leq \nu_1$, dass $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)e^{(i)} = \lambda_1 e^{(i)}$, d.h.

$$f(v_i) = (\Psi_{\mathfrak{V}} \circ \mathrm{Lin}(\Lambda) \circ \Psi_{\mathfrak{V}}^{-1})(v_1) = (\Psi_{\mathfrak{V}} \circ \mathrm{Lin}(\Lambda))(e^{(i)}) = \Psi_{\mathfrak{V}}(\lambda_1 e^{(i)}) = \lambda_1 v_i,$$

   d.h. $\mu_{\mathrm{geo}}(f, \lambda_1) = \dim \mathrm{Eig}(f, \lambda) \geq \nu_1 = \mu_{\mathrm{alg}}(f, \lambda_1)$ und damit $\mu_{\mathrm{geo}}(f, \lambda_1) = \mu_{\mathrm{alg}}(f, \lambda_1)$. Selbiges gilt auch für die anderen Eigenwerte $\lambda_2, ..., \lambda_s$.
3. Gelten $\chi_f = \prod(\lambda_i - t)^{\nu_i}$ und $\mu_{\mathrm{alg}}(f, \lambda_i) = \mu_{\mathrm{geo}}(f, \lambda_i)$ für $1 \leq i \leq s$. Es gilt $\mathrm{Eig}(f, \lambda_i) \cap \mathrm{Eig}(f, \lambda_j) = \{0\}$ für $j \neq i$, denn aus $v \in \mathrm{Eig}(f, \lambda_i) \cap \mathrm{Eig}(f, \lambda_j)$ folgt $\lambda_i v = f(v) = \lambda_j v$, d.h. $(\lambda_i - \lambda_j)v = 0$ und damit $v = 0$. Also ist

$$\mathrm{Eig}(f, \lambda_1) + \cdots + \mathrm{Eig}(f, \lambda_s) = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_s).$$

   Außerdem folgt aus $\mu_{\mathrm{alg}}(f, \lambda_i) = \mu_{\mathrm{geo}}(f, \lambda_i)$, dass

$$\dim(\mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_s)) = \dim \mathrm{Eig}(f, \lambda_1) + \cdots + \dim \mathrm{Eig}(f, \lambda_s) = \nu_1 + \cdots + \nu_s = n,$$

   d.h. $\mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_s) = V$.
4. Gelte $V = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_s)$. Seien $(v_1, ..., v_{\nu_1})$ eine Basis von $\mathrm{Eig}(f, \lambda_1)$, selbiges für die anderen Eigenräume. Dann bildet $\mathfrak{V} = (v_1, ..., v_{\nu_1}, ...)$ eine Basis von V aus Eigenvektoren zu f. $\square$

Definition 4.27.
Eine quadratische Matrix $A \in \mathfrak{M}(n, n)$ der Gestalt

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

heißt (obere) Dreiecksmatrix oder Trigonalmatrix.

$f \in \mathrm{Hom}(K^n, K^n)$ heißt trigonalisierbar, falls es eine Basis $\mathfrak{V} = (v_1, ..., v_n)$ von V gibt, so dass $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ Dreiecksgestalt hat.

$A \in \mathfrak{M}(n, n)$ heißt trigonalisierbar, falls es eine invertierbare Matrix $C \in \mathfrak{M}(n, n)$ gibt, so dass CAC^{-1} Dreiecksgestalt hat.

Bemerkung 4.28.

$A \in \mathfrak{M}(n,n)$ ist genau dann trigonalisierbar, wenn $\mathrm{Lin}(A)$ trigonalisierbar ist. ◊

Bemerkung 4.29.

Nicht alle Matrizen sind diagonalisierbar: Drehmatrizen um den Winkel θ haben zum Beispiel für $\theta \neq n\pi$ keine Eigenvektoren, also auch keine Eigenwerte, und lassen sich damit weder über $\mathbb{R}$ noch über $\mathbb{C}$ in Diagonalgestalt überführen.

Allerdings sind alle Matrizen über $\mathbb{C}$ trigonalisierbar: ◊

Satz 4.30.

$f \in \mathrm{Hom}(V,V)$ ist genau dann trigonalisierbar, wenn χ_f in Linearfaktoren zerfällt.

Beweis.

1. Sei $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ eine Trigonalmatrix. Dann gilt für das charakteristische Polynom χ_f:

$$\chi_f = \det \begin{pmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} = \prod_{i=1}^{n} (\alpha_i - t) = (-1)^n \prod_{i=1}^{n} (t - \alpha_i),$$

   d.h. χ_f zerfällt in Linearfaktoren.

2. Wir führen eine Induktion über n. Der Fall $n = 1$ ist trivial. Gelte die Behauptung also für $n-1$.

   Seien $\chi_f = \prod(\alpha_i - t)$ mit $\alpha_i \in K$, $1 \leq i \leq n$, und sei v_1 ein Eigenvektor zum Eigenwert α_1. Wir ergänzen zu einer Basis $\mathfrak{V} = (v_1, ..., v_n)$ von V.

   Wegen $f(v_1) = \alpha_1 v_1 + 0v_2 + ... + 0v_n$ besitzt $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ die Darstellung

$$A = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \begin{pmatrix} \alpha_1 & * \\ 0 & A_1 \end{pmatrix}$$

   für eine Matrix $A_1 \in \mathfrak{M}(n-1, n-1)$. Weiter ist

$$\chi_A = \det(A - t\mathrm{Id}(n)) = (\alpha_1 - t)\det(A_1 - t\mathrm{Id}(n-1)) = (\alpha_1 - t)\chi_{A_1} = (\alpha_1 - t)\prod_{i=2}^{n}(\alpha_i - t),$$

   d.h. $\chi_{A_1} = \prod_{i=2}^{n}(\alpha_i - t)$ faktorisiert über K. Nach Induktionsvoraussetzung gibt es ein $C \in \mathfrak{I}(n-1)$, so dass $\Lambda_1 = CA_1C^{-1}$ Dreiecksgestalt hat. Also ist

$$A \approx \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} \alpha_1 & * \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}^{-1} = \begin{pmatrix} \alpha_1 & * \\ 0 & CA_1C^{-1} \end{pmatrix} = \begin{pmatrix} \alpha_1 & * \\ 0 & \Lambda_1 \end{pmatrix},$$

   d.h. auch A kann auf Diagonalgestalt transformiert werden. □

4.4. Der Satz von Hamilton-Cayley

Wir beweisen in diesem Abschnitt den Satz von Hamilton-Cayley, welcher besagt, dass das charakteristische Polynom χ_f eines Endomorphismus f, ausgewertet in f selbst, das Nullpolynom ergibt. Wir werden daraus später eine allgemeine Normalform über $\mathbb{R}$ ableiten: Jede Matrix $A \in \mathfrak{M}_{\mathbb{R}}(n,n)$ mit $A^{\mathrm{T}} = A^{-1}$ lässt sich in Drehungen in gewissen Ebenen und Streckungen längs gewisser Achsen zerlegen.

Satz 4.31. (Hamilton-Cayley)

Seien V ein K-Vektorraum, $f \in \mathrm{Hom}(V,V)$ und χ_f das charakteristische Polynom von f.

Dann gilt $\chi_f(f) = 0$.

Beweis.

1. Sei zunächst K *algebraisch abgeschlossen*, d.h. jedes $f \in K[t]$ zerfalle über K in Linearfaktoren.

   Dann gibt es eine Basis $\mathfrak{V} = (v_1, ..., v_n)$ von V, so dass $A = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ eine Dreiecksmatrix ist.

   Seien $\alpha_1, ..., \alpha_n \in K$ die Eigenwerte von f, d.h. $\alpha_i = a_{ii}$. Setze $W_i = \mathrm{span}(v_1, ..., v_i)$ für $0 \leq i \leq n$, d.h. $\{0\} = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_n = V$. Für $f_i = f - \alpha_i \mathrm{id}$ gilt dann $f_i : W_i \to W_{i-1}$, denn für $j \leq i$ ist

$$\begin{aligned} f_i(v_j) &= (\Psi_{\mathfrak{V}}^{-1} \circ \mathrm{Lin}(A) \circ \Psi_{\mathfrak{V}} - \alpha_j \Psi_{\mathfrak{V}}^{-1} \circ \mathrm{id} \circ \Psi_{\mathfrak{V}})(v_j) \\ &= (\Psi_{\mathfrak{V}}^{-1} \circ \mathrm{Lin}(A) - \alpha_j \Psi_{\mathfrak{V}}^{-1} \circ \mathrm{id}) e^{(j)} \\ &= \Psi_{\mathfrak{V}}^{-1} A^{(j)} - \alpha_j \Psi_{\mathfrak{V}}^{-1} e^{(j)} \\ &= a_{1j} v_1 + \cdots + \underbrace{a_{jj} v_j - \alpha_j v_j}_{=0} \end{aligned}$$

   Also ist $f_1 \circ \cdots \circ f_n : W_n \to W_0$ die Nullabbildung, d.h.

$$\chi_f(f) = \det(A - \mathrm{id} f) = \prod_{i=1}^{n} (\alpha_i - f) = (-1)^n \prod_{i=1}^{n} (f - \alpha_i) = (-1)^n (f_1 \circ \cdots \circ f_n) = 0.$$

2. Sei nun K ein beliebiger Körper. Dann besitzt K einen algebraisch abgeschlossenen Oberkörper $\bar{K}$. Die Abbildung $\bar{f} = \mathrm{Lin}(A) \in \mathrm{Hom}(\bar{K}^n, \bar{K}^n)$ ist trigonalisierbar, d.h. $\chi_{\bar{f}}(\bar{f}) = 0$. Speziell gilt für alle $v \in K^n \subseteq \bar{K}^n$, dass $\chi_{\bar{f}}(\bar{f})(v) = 0$, wegen $\chi_f(f)(v) = \chi_{\bar{f}}(\bar{f})(v) = 0$ folgt also $\chi_f(f) = 0$. □

Beispiel 4.32.

Für $A = (1, 2; -1, 1)$ ist $\chi_A = (1 - t)^2 + 2 = t^2 - 2t + 3$ und damit

$$\chi_A(A) = A^2 - 2A + 3 = \begin{pmatrix} -1 & 4 \\ -2 & -2 \end{pmatrix} - \begin{pmatrix} 2 & 4 \\ -2 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Für den zugehörigen Endomorphismus $f = \mathrm{Lin}(A)$ gelten $f(v) = (v_1 + 2v_2; -v_1 + v_2)$ und

$$f(f(v)) - 2f(v) + 3v = \begin{pmatrix} (v_1 + 2v_2) + 2(-v_1 + v_2) \\ -(v_1 + 2v_2) + (-v_1 + v_2) \end{pmatrix} - 2 \begin{pmatrix} v_1 + 2v_2 \\ -v_1 + v_2 \end{pmatrix} + 3 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad \Diamond$$

Korollar 4.33.

Seien V ein $\mathbb{R}$-Vektorraum und $f \in \mathrm{Hom}(V, V)$.

Dann gibt es einen f-invarianten Unterraum $W \subseteq V$ der Dimension 1 oder 2.

Beweis.

1. Habe χ_f eine reelle Nullstelle $\lambda \in \mathbb{R}$ mit zugehörigem Eigenvektor v. Setze $W = \mathbb{R}v$, dann gelten $f(W) = W$ und $\dim W = 1$.

2. Habe χ_f nun keine Nullstelle in $\mathbb{R}$. Dann ist n gerade und χ_f zerfällt über $\mathbb{R}$ in quadratische Faktoren, denn χ_f faktorisiert über $\mathbb{C}$ und zu jedem rein komplexen Eigenwert $\lambda \in \mathbb{C}$ ist auch sein komplex konjugiertes $\bar{\lambda} \in \mathbb{C}$ ein Eigenwert von χ_f mit der gleichen algebraischen Vielfachheit, so dass das Produkt der zugehörigen komplexen Linearfaktoren $(X - \lambda), (X - \bar{\lambda}) \in \mathbb{C}[X]$ wieder in $\mathbb{R}[X]$ liegt: $(X - \lambda)(X - \bar{\lambda}) = X^2 - 2\mathrm{Re}\lambda + |\lambda|^2 \in \mathbb{R}[X]$.

   Also lässt sich χ_f zerlegen in $\chi_f = \pm\chi_1 \cdots \chi_k \in \mathbb{R}[t]$ mit $\deg(\chi_i) = 2$. Nach dem Satz von Hamilton-Cayley ist $\chi_f(f) = \chi_k(f) \circ \cdots \circ \chi_1(f) = 0$. Wir setzen $\chi_0 = \mathrm{id}$.

   Wegen $\chi_k(f) \circ \cdots \circ \chi_0(f) = 0$ und $\chi_0(f) = f \neq 0$ muss es ein $i \leq k$ geben mit $\chi_i(f) \circ \cdots \circ \chi_0(f) = 0$ und $\chi_{i-1}(f) \circ \cdots \circ \chi_0(f) \neq 0$, etwa $(\chi_{i-1}(f) \circ \cdots \circ \chi_0(f))(w) \neq 0$ mit $w \in V \backslash \{0\}$. Habe χ_i die Gestalt $\chi_i = t^2 + \alpha t + \beta \mathrm{id}$ mit Koeffizienten $\alpha, \beta \in \mathbb{R}$. Setze $v = (\chi_{i-1}(f) \circ \cdots \circ \chi_0(f))(w)$.

   Dann ist $0 = \chi_i(f)(v) = f(f(v)) + \alpha f(v) + \beta v$, d.h. $f(f(v)) = -\alpha f(v) - \beta v \in \mathbb{R}f(v) + \mathbb{R}v$. Wir setzen $W = \mathrm{span}(v, f(v))$, dann ist W ein f-invarianter Unterraum der Dimension eins oder zwei. □

5. Euklidische und unitäre Räume

Das letzte Kapitel beschäftigt sich mit Vektorräumen über den Körpern $\mathbb{R}$ und $\mathbb{C}$, auf denen über ein Skalarprodukt eine geometrische Struktur erklärt ist. Hier können Längen, Abstände und Winkel auf abstrakte Weise definiert werden. Wir werden dann Endomorphismen untersuchen, die in gewisser Weise mit dem Skalarprodukt veräglich sind, beispielsweise solche, die Längen- oder Winkel-treu sind.

5.1. Räume mit Skalarprodukt

Im ersten Abschnitt übertragen wir die geometrischen Konzepte, die wir für den $\mathbb{R}^n$ und den $\mathbb{C}^n$ eingeführt haben, auf abstrakte endlich-dimensionale $\mathbb{R}$- bzw. $\mathbb{C}$-Vektorräume, d.h. wir führen mit Hilfe einer symmetrischen, positiv definiten Bilinear- oder Sesquilinearform Normen, Metriken und Winkel ein und leiten Aussagen wie die Cauchy-Schwarzsche Ungleichung, die Dreiecksungleichung und die Parallelogrammregel her.

Sei ab jetzt stets $\mathbb{K}$ einer der Körper $\mathbb{R}, \mathbb{C}$.

Definition 5.1.

$\mathrm{s} : V \times V \to \mathbb{R}$ heißt eine symmetrische Bilinearform, falls für alle $v_1, v_2, v_3 \in V$ und alle $\alpha, \beta \in \mathbb{K}$ gelten:

$$\begin{aligned} \mathrm{s}(\alpha v_1 + \beta v_2, v_3) &= \alpha \mathrm{s}(v_1, v_3) + \beta \mathrm{s}(v_2, v_3), \\ \mathrm{s}(v_1, \alpha v_2 + \beta v_3) &= \alpha \mathrm{s}(v_1, v_2) + \beta \mathrm{s}(v_1, v_3), \\ \mathrm{s}(v_1, v_2) &= \mathrm{s}(v_2, v_1). \end{aligned}$$

$\mathrm{s} : V \times V \to \mathbb{C}$ heißt eine Hermitesche Form, falls für alle $v_1, v_2, v_3 \in V$ und alle $\alpha, \beta \in \mathbb{K}$ gelten:

$$\begin{aligned} \mathrm{s}(\alpha v_1 + \beta v_2, v_3) &= \alpha \mathrm{s}(v_1, v_2) + \beta \mathrm{s}(v_2, v_3), \\ \mathrm{s}(v_1, \alpha v_2 + \beta v_3) &= \overline{\alpha} \mathrm{s}(v_1, v_2) + \overline{\beta} \mathrm{s}(v_1, v_3), \\ \mathrm{s}(v_1, v_2) &= \overline{\mathrm{s}(v_2, v_1)}. \end{aligned}$$

Bemerkung 5.2.

Hermitesche Formen sind also linear in der ersten und sesquilinear in der zweiten Komponenten.

Wir hätten genau so gut Linearität in der zweiten und Sesquilinearität in der ersten Komponenten fordern können. $\diamondsuit$

Definition 5.3.

Eine symmetrische Bilinearform (Hermitesche Form) $\mathrm{s} : V \times V \to \mathbb{K}$ heißt ein Skalarprodukt, falls s positiv definit ist, d.h. falls für alle $v \in V \backslash \{0\}$ gilt

$$\mathrm{s}(v, v) > 0.$$

(V, s) heißt ein euklidischer Raum, falls V ein $\mathbb{R}$-Vektorraum und s ein Skalarprodukt auf V ist.

(V, s) heißt ein unitärer Raum, falls V ein $\mathbb{C}$-Vektorraum und s ein Skalarprodukt auf V ist.

Bemerkung 5.4.

1. Sei (V, s) ein euklidischer oder unitärer Raum. Dann definiert $||\cdot|| : V \times V \to \mathbb{R}$,
$$||v|| := \sqrt{\mathrm{s}(v, v)}$$
eine Norm auf (V, s), d.h. $(V, ||\cdot||)$ ist ein normierter Raum.
2. Insbesondere gilt die Dreiecksungleichung, d.h. für alle $v, w \in V$ ist
$$||v + w|| \leq ||v|| + ||w||.$$

3. Für $||\cdot||$ gilt die Cauchy-Schwarzsche Ungleichung, d.h. für alle $v, w \in V$ ist
$$|\mathrm{s}(v,w)| \leq ||v|| ||w||,$$
wobei Gleichheit genau dann auftritt, wenn v und w linear abhängig sind.

4. Es gilt die Parallelogrammregel, d.h. für alle $v, w \in V$ ist
$$||v+w||^2 + ||v-w||^2 = 2(||v||^2 + ||w||^2).$$

5. Es gelten die Binomischen Formeln bzw. der Satz des Pythagoras, d.h. für alle $v, w \in V$ ist
$$||v+w||^2 = ||v||^2 + 2\mathrm{Re}\langle v,w\rangle + ||w||^2,$$
$$||v-w||^2 = ||v||^2 - 2\mathrm{Re}\langle v,w\rangle + ||w||^2.$$

6. $||\cdot||$ wiederum definiert eine Metrik $\mathrm{d}: V \times V \to \mathbb{R}$ auf (V, s) via
$$\mathrm{d}(v,w) = ||v-w||,$$
d.h. (V, d) ist ein metrischer Raum.

7. Zwei Vektoren $v, w \in V$ heißen orthogonal oder senkrecht – wir schreiben $v \perp w$ – wenn gilt
$$\mathrm{s}(v,w) = 0.$$
◊

5.2. Orthonormalisierung und orthonormale Summen

Wir lernen hier ein Verfahren kennen, aus einer gegebenen Basis eine solche zu konstruieren, deren Elemente normiert und paarweise orthogonal zueinander sind. Eine solche Orthonormalbasis hat den Vorteil, den Vektorraum gleichmäßig zu erzeugen. Die Norm eines Vektors entspricht dann beispielsweise der Summe seiner Koeffizienten in der Orthonormalbasis. Dies ist insbesondere für die Numerik von Bedeutung, da in solchen Basen stabil gerechnet werden kann.

Definition 5.5.

Seien (V, s) ein euklidischer oder unitärer Raum und $v_1, ..., v_m \in V \backslash \{0\}$.

$(v_1, ..., v_m)$ heißt ein orthogonales System, falls $v_i \perp v_j$ für $i \neq j$ erfüllt ist.

$(v_1, ..., v_m)$ heißt ein orthonormales System, falls zusätzlich $||v_i|| = 1$ für alle i gilt, d.h. falls
$$\mathrm{s}(v_i, v_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

δ_{ij} heißt das Kronecker-Delta.

$(v_1, ..., v_m)$ ist eine Orthogonalbasis bzw. eine Orthonormalbasis, falls $(v_1, ..., v_n)$ ein Orthogonalsystem bzw. ein Orthonormalsystem und zusätzlich eine Basis von V ist.

Bemerkung 5.6.

1. Ist $(v_1, ..., v_m)$ ein Orthogonalsystem, so sind $v_1, ..., v_m$ linear unabhängig:

   Gelte $\alpha_1 v_1 + ... + \alpha_m v_m = 0$ für gewisse $\alpha_1, ..., \alpha_m \in \mathbb{K}$, dann gilt für beliebiges i:
$$0 = \mathrm{s}(v_i, 0) = \mathrm{s}(v_i, \alpha_1 v_1 + ... + \alpha_n v_n) = \alpha_1 \mathrm{s}(v_i, v_1) + ... + \alpha_n \mathrm{s}(v_i, v_n) = \alpha_i \underbrace{\mathrm{s}(v_i, v_i)}_{=||v_i||^2 \neq 0},$$
   also $\alpha_i = 0$. ◊

2. Ist $\mathfrak{V} = (v_1, ..., v_n)$ Orthonormalbasis eines euklidischen oder unitären Raums (V, s) und sind $v, w \in V$ mit Koeffizientenvektoren $x, y \in \mathbb{K}^n$, dann stimmt das Skalarprodukt von v, w in V mit dem Skalarprodukt von x, y in $\mathbb{K}^n$ überein:
$$\mathrm{s}(v,w) = \langle x, y\rangle.$$
   Es gilt nämlich
$$\mathrm{s}(v,w) = \mathrm{s}\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i,j=1}^n \mathrm{s}(x_i v_i, y_j v_j) = \sum_{i,j=1}^n \overline{x}_i y_j \mathrm{s}(v_i, v_j) = \sum_{i=1}^n \overline{x}_i y_i = \langle x, y\rangle.$$
   ◊

Algorithmus 5.7. (Orthonormalisierungsverfahren von Gram und Schmidt)

Seien $(v_1, ..., v_m)$ ein Orthonormalsystem und $v \notin \text{span}(v_1, ..., v_m)$. Dann bildet $(v_1, ..., v_m, v_{m+1})$ ein Orthonormalsystem mit $\text{span}(v_1, ..., v_{m+1}) = \text{span}(v_1, ..., v_m, v)$, wenn man setzt

$$v_{m+1} = \frac{v'_{m+1}}{||v'_{m+1}||}, \qquad v'_{m+1} = v - \sum_{i=1}^{m} \mathrm{s}(v_i, v)v_i.$$

$(v_1, ..., v_m, v_{m+1})$ heißt die Orthonormalisierung von $(v_1, ..., v_m, v)$.

Beweis.

Offenbar ist v_{m+1} normiert und nicht das Nullelement, da aus $v_{m+1} = 0$, d.h. $v'_{m+1} = 0$, folgen würde, dass v eine Linearkombination der $v_1, ..., v_m$ ist. Weiter gilt $v_{m+1} \perp v_j$ für alle j, denn

$$\mathrm{s}(v_j, v'_{m+1}) = \mathrm{s}(v_j, v) - \sum_{i=1}^{m} \mathrm{s}(v_j, v)\mathrm{s}(v_j, v_i) = \mathrm{s}(v_j, v) - \mathrm{s}(v_j, v)\underbrace{\mathrm{s}(v_j, v_j)}_{=1} = 0.$$

Schließlich sind

$$v = v'_{m+1} + \sum_{i=1}^{m} \mathrm{s}(v_i, v)v_i \in \text{span}(v_1, ..., v_m, v_{m+1}), \quad v'_{m+1} = v - \sum_{i=1}^{m} \mathrm{s}(v_i, v)v_i \in \text{span}(v_1, ..., v_m, v). \quad \square$$

Definition 5.8.

Seien (V, s) ein euklidischer oder unitärer Raum und $V_1, ..., V_k$ Untervektorräume von V.

Dann heißt V orthogonale Summe von $V_1, ..., V_k$, falls $V = V_1 + ... + V_k$ und $V_i \perp V_j$ für $i \neq j$.

Wir schreiben dafür $V = V_1 \circledast \cdots \circledast V_k$, wobei $U \perp W$ für zwei Untervektorräume $U, W \subseteq V$ gelte, falls $u \perp w$ für alle $u \in U$ und alle $w \in W$.

Bemerkung 5.9.

Eine orthogonale Summe ist immer direkt, denn per Definition gibt es zu $v \in V$ stets $v_1 \in V_1, ..., v_k \in V_k$ mit $v = v_1 + ... + v_k$ und wegen der linearen Unabhängigkeit von $v_1, ..., v_k$ ist $0 = v_1 + ... + v_k$ eindeutig, d.h. $v_1 = ... = v_k = 0$. Also gilt:

$$V = V_1 \circledast \cdots \circledast V_k \qquad \Longrightarrow \qquad V = V_1 \oplus \cdots \oplus V_k. \qquad \Diamond$$

Definition 5.10.

Ist W ein Untervektorraum von V, dann setzen wir

$$W^\perp = \{v \in V \mid v \perp_s w \text{ für alle } w \in W\}.$$

$W^\perp$ heißt der Orthogonalraum oder das orthogonale Komplement von W.

Bemerkung 5.11.

Bei Linearen Gleichungssystemen bezog sich $\perp$ auf das Zeilen-Spalten-Produkt $\circ$, das im Reellen mit dem euklidischen Skalarprodukt $\mathrm{s} = \langle \cdot, \cdot \rangle$ übereinstimmt. Im Komplexen ist dies nicht der Fall, da $\circ$ bilinear und nicht sesquilinear ist; außerdem ist $\circ$ nicht definit.

Das folgende Lemma gilt für die von s induzierte Orthogonalitätsrelation $\perp$, aber nicht für $\circ$: $\Diamond$

Lemma 5.12.

Ist W ein Untervektorraum eines euklidischen oder unitären Raumes (V, s), so ist $V = W \circledast W^\perp$.

Beweis.

Seien $(w_1, ..., w_m)$ eine Orthonormalbasis von W und $w_{m+1}, ..., w_n$ eine Ergänzung zu einer Orthonormalbasis von V. Weiter sei $W' = \operatorname{span}(w_{m+1}, ..., w_n)$, d.h. $V = W \oplus W'$. Wir müssen zeigen: $W' = W^\perp$. Dabei ist die Inklusion $W' \subseteq W^\perp$ klar nach Konstruktion.

Sei also $w = \alpha_1 w_1 + ... + \alpha_n w_n \in W^\perp$, d.h. $w \perp w_i$ für alle $1 \leq i \leq m$. Dann $\alpha_i = 0$ für $i \leq m$ und damit $w = \alpha_{m+1} w_{m+1} + ... + \alpha_n w_n \in W'$, d.h. $W^\perp \subseteq W'$. Also insgesamt $V = W \circledast W^\perp$. □

5.3. Orthogonale und unitäre Endomorphismen

Wir führen in diesem Abschnitt Endomorphismen ein, die die Längen- und Orthogonalitätsrelation respektieren, d.h. die *verträglich* mit dem Skalarprodukt des euklidischen bzw. unitären Raumes sind. Diese Endomorphismen sind stets invertierbar mit Determinante Eins und die Inverse einer Darstellungsmatrix ist die (im unitären Fall konjugierte) transponierte Matrix. Wir haben solche Matrizen eben kennengelernt: Die Transformationsmatrizen von Orthonormalbasen.

Anschließend zeigen wir, dass unitäre Endomorphismen sich immer diagonalisieren lassen und dass orthogonale Endomorphismen mit Hilfe des Satzes von Hamilton-Cayley auf eine Block-Diagonalgestalt gebracht werden können, wobei die Blöcke aus (1×1)-Matrizen, d.h. Streckungen längs einer Achse, und (2×2)-Drehmatrizen, d.h. Rotationen in einer Ebene, bestehen. Zentrales Argument dabei ist, den Vektorraum soweit möglich als orthogonale Summe von Eigenräumen zu schreiben.

Definition 5.13.

Seien (V, s) ein euklidischer bzw. unitärer Raum und $f \in \operatorname{Hom}(V, V)$.

f heißt orthogonal bzw. unitär, falls für alle $v, w \in V$ gilt:

$$\mathrm{s}(f(v), f(w)) = \mathrm{s}(v, w).$$

Bemerkung 5.14.

Insbesondere respektieren orthogonale und unitäre Endomorphismen die Orthogonalitätsrelation:

$$v \perp w = 0 \qquad \Longrightarrow \qquad f(v) \perp f(w).$$ ♢

Lemma 5.15.

Seien (V, s) ein euklidischer oder unitärer Raum und $f \in \operatorname{Hom}(V, V)$. Dann sind äquivalent:

1. f ist orthogonal oder unitär.
2. $||f(v)|| = ||v||$ für alle $v \in V$.
3. $||v|| = 1 \Rightarrow ||f(v)|| = 1$.

Beweis.

Die Implikationen (1) $\Rightarrow$ (2) und (2) $\Rightarrow$ (3) sind klar.

1. (3) $\Rightarrow$ (2): Sei $v \in V$, dann gilt:

$$\begin{aligned} \left\| \frac{1}{||v||} v \right\| &= \frac{1}{||v||} ||v|| = 1 \\ \Longrightarrow \qquad 1 &= \left\| f\left(\frac{1}{||v||} v\right) \right\| = \frac{1}{||v||} ||f(v)|| \\ \Longrightarrow \qquad & ||f(v)|| = ||v||. \end{aligned}$$

2. (2) $\Rightarrow$ (1): Seien $v, w \in V$. Wir betrachten zunächst den euklidischen Fall:

$$2\mathrm{s}(v, w) = \mathrm{s}(v + w, v + w) - \mathrm{s}(v, v) - \mathrm{s}(w, w)$$

$$
\begin{aligned}
&= ||v+w||^2 - ||v||^2 - ||w||^2 \\
&= ||f(v+w)||^2 - ||f(v)||^2 - ||f(w)||^2 \\
&= ||f(v)+f(w)||^2 - ||f(v)||^2 - ||f(w)||^2 \\
&= \mathrm{s}(f(v)+f(w), f(v)+f(w)) - \mathrm{s}(f(v),f(v)) - \mathrm{s}(f(w),f(w)) \\
&= 2\mathrm{s}(f(v),f(w)).
\end{aligned}
$$

Ebenso gilt im unitären Fall:

$$
\begin{aligned}
2\mathrm{s}(v,w) &= \mathrm{s}(v,w) + \overline{\mathrm{s}(v,w)} + (\mathrm{s}(v,w) - \overline{\mathrm{s}(v,w)}) \\
&= \mathrm{s}(v,w) + \overline{\mathrm{s}(v,w)} + i(\mathrm{s}(v,iw) + \overline{\mathrm{s}(v,iw)}) \\
&= \mathrm{s}(v+w,v+w) - \mathrm{s}(v,v) - \mathrm{s}(w,w) + i(\mathrm{s}(v+iw,v+iw) - \mathrm{s}(v,v) - \mathrm{s}(iw,iw)) \\
&= ||v+w||^2 - ||v||^2 - ||w||^2 + i(||v+iw||^2 - ||v||^2 - ||iw||^2) \\
&= ||f(v)+f(w)||^2 - ||f(v)||^2 - ||f(w)||^2 + i(||f(v)+if(w)||^2 - ||f(v)||^2 - ||if(w)||^2) \\
&= \mathrm{s}(f(v)+f(w), f(v)+f(w)) - \mathrm{s}(f(v),f(v)) - \mathrm{s}(f(w),f(w)) \\
&\quad + i(\mathrm{s}(f(v)+if(w), f(v)+if(w)) - \mathrm{s}(f(v),f(v)) - \mathrm{s}(if(w),if(w))) \\
&= \mathrm{s}(f(v),f(w)) + \overline{\mathrm{s}(f(v),f(W))} + i(\mathrm{s}(f(v),if(w)) + \overline{\mathrm{s}(f(v),if(w))}) \\
&= \mathrm{s}(f(v),f(w)) + \overline{\mathrm{s}(f(v),f(W))} + (\mathrm{s}(f(v),f(w)) - \overline{\mathrm{s}(f(v),f(w))}) \\
&= 2\mathrm{s}(f(v),f(w)).
\end{aligned}
$$

□

Korollar 5.16.

Seien (V,s) ein euklidischer (unitärer) Raum und $f \in \mathrm{Hom}(V,V)$ orthogonal (unitär). Dann gelten:

1. f ist injektiv und damit bijektiv.
2. f^{-1} ist ebenfalls orthogonal oder unitär.
3. Ist $g \in \mathrm{Hom}(V,V)$ orthogonal oder unitär, dann ist auch $f \circ g \in \mathrm{Hom}(V,V)$ orthogonal oder unitär.
4. Ist $\lambda \in \mathbb{K}$ ein Eigenwert von f, dann gilt $|\lambda| = 1$.

Beweis.

1. Ist $f(v) = 0$, dann $0 = ||f(v)|| = ||v||$, d.h. auch $v = 0$. Also ist $\mathrm{Kern}(f) = \{0\}$ und damit $f \in \mathfrak{I}(V)$. Hierbei bezeichnet $\mathfrak{I}(V)$ natürlich die Menge der invertierbaren Endomorphismen über V, welche via $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}$ für eine beliebige Basis $\mathfrak{V}$ von V isomorph zu $\mathfrak{I}(n)$ ist.
2. Es gilt $||v|| = ||f(f^{-1}(v))|| = ||f^{-1}(v)||$. Also ist mit f auch f^{-1} orthogonal bzw. unitär.
3. Es ist $||(f \circ g)(v)|| = ||f(g(v))|| = ||g(v)|| = ||v||$.
4. Ist $f(v) = \lambda v$ mit $v \neq 0$, dann $0 \neq ||v|| = ||f(v)|| = ||\lambda v|| = |\lambda|||v||$, d.h. $|\lambda| = 1$. □

Bemerkung 5.17.

1. Die orthogonalen bzw. unitären Endomorphismen eines euklidischen bzw. unitären Raumes (V,s) bilden eine Gruppe, die orthogonale Gruppe $(\mathfrak{O},\mathrm{s})$ bzw. die unitäre Gruppe $(\mathfrak{U},\mathrm{s})$. Diese sind Untergruppen der Allgemeinen Linearen Gruppe $\mathfrak{I}(V)$.
2. Die orientierten orthogonalen bzw. unitären Endomorphismen sind diejenigen mit Determinante $+1$. Sie bilden eine Untergruppe von $(\mathfrak{O},\mathrm{s})$ bzw. $(\mathfrak{U},\mathrm{s})$ und auch von der Speziellen Linearen Gruppe $\mathfrak{S}(V) = \{f \in \mathrm{Hom}(V,V) \mid \det f = 1\}$, welche ebenfalls eine Untergruppe von $\mathfrak{I}(V)$ ist.
3. Die Determinante einer Matrix $A \in \mathfrak{M}_{\mathbb{R}}(n,n)$ kann als Maßdes von den Spalten von A aufgespannten Spats verstanden werden. Wegen $\det(A) = \det(A^{\mathrm{T}})$ kann man alternativ auch den von den Zeilen aufgespannten Spat betrachten. Speziell im $\mathbb{R}^2$ ist $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ der Flächeninhalt des von $(a_{11}; a_{21})$ und $(a_{12}; a_{22})$ aufgespannten Parallelogramms. In diesem Sinne erhalten Endomorhismen aus $\mathfrak{S}(V)$ das Volumen $\det(A)$ des von A aufgespannten Spats. ◊

Beispiel 5.18.

1. Drehungen f_θ um einen Winkel θ in der Ebene $\mathbb{R}^2$ erhalten stets Orientierung und Volumen, denn $\det f_\theta = \sin^2\theta + \cos^2\theta = 1$:
$$\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(f_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

2. Spiegelungen f_v an einer Achse $v \in \mathbb{R}^2$ sind dagegen stets volumenerhaltend, aber orientierungsumkehrend: Es gilt $\det f = -1$:
$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f_v) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
mit $\mathfrak{V} = ((v_1; v_2), (-v_2; v_1))$. ◊

Lemma 5.19.

Sei $\mathfrak{V} := (v_1, ..., v_n)$ eine Orthonormalbasis des euklidischen bzw. unitären Raumes (V, s) und für $f \in \mathrm{Hom}(V, V)$ sei $\mathfrak{W} := (f(v_1), ..., f(v_n))$. Dann gilt:

$$f \text{ ist orthogonal bzw. unitär} \quad \Longleftrightarrow \quad \mathfrak{W} \text{ ist eine Orthonormalbasis von } V.$$

Beweis.

1. Ist $\mathfrak{V}$ eine Orthonormalbasis von (V, s), dann gilt $\delta_{ij} = \mathrm{s}(v_i, v_j) = \mathrm{s}(f(v_i), f(v_j))$, d.h. auch $\mathfrak{W}$ ist eine Orthonormalbasis von V.

2. Sei $v = \alpha_1 v_1 + ... + \alpha_n v_n$, dann gilt mit dem Satz des Pythagoras für $f(v) = \alpha_1 f(v_1) + ... + \alpha_n f(v_n)$:
$$\begin{aligned} ||f(v)||^2 &= ||\alpha_1 f(v_1) + ... + \alpha_n f(v_n)||^2 \\ &= ||\alpha_1 f(v_1)||^2 + ... + ||\alpha_n f(v_n)||^2 \\ &= |\alpha_1|^2 + ... + |\alpha_n|^2 \\ &= ||\alpha_1 v_1||^2 + ... + ||\alpha_n v_n||^2 \\ &= ||\alpha_1 v_1 + ... + \alpha_n v_n||^2 \\ &= ||v||^2. \end{aligned}$$
Also ist f orthogonal bzw. unitär. □

Satz 5.20.

Seien (V, s) ein euklidischer oder unitärer Raum, $\mathfrak{V} = (v_1, ..., v_n)$ eine Orthonormalbasis von V und $f \in \mathrm{Hom}(V, V)$. Dann gilt:

$$f \text{ ist orthogonal bzw. unitär} \quad \Longleftrightarrow \quad \overline{(M_{\mathfrak{V}}^{\mathfrak{V}}(f))}^{\mathrm{T}} = (M_{\mathfrak{V}}^{\mathfrak{V}}(f))^{-1}.$$

Beweis.

Seien $\mathfrak{W} = (f(v_1), ..., f(v_n))$ und $A = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$. Dann gilt:

$$\mathrm{s}(w_i, w_j) = \mathrm{s}\left(\sum_{l=1}^n a_{li} v_l, \sum_{k=1}^n a_{kj} v_j\right) = \sum_{l=1}^n \sum_{k=1}^n a_{li} \overline{a_{kj}} \mathrm{s}(v_i, v_j) = \sum_{l=1}^n a_{li} \overline{a_{lj}} = \left(A^{\mathrm{T}} \overline{A}\right)_{ij},$$

d.h. genau dann gilt $\mathrm{s}(w_i, w_j) = \delta_{ij}$, wenn $A^{\mathrm{T}} \overline{A} = \mathrm{Id}$ bzw. $\overline{A}^{\mathrm{T}} A = \mathrm{Id}$ erfüllt ist. □

Definition 5.21.

Sei $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis des euklidischen bzw. unitären Raumes (V, s). Dann heißt

$$\mathrm{Mat}_{\mathfrak{V}}(\mathrm{s}) = (\mathrm{s}(v_i, v_j))_{1\le j\le n}^{1\le i\le n} = \begin{pmatrix} \mathrm{s}(v_1, v_1) & \cdots & \mathrm{s}(v_1, v_n) \\ \vdots & \ddots & \vdots \\ \mathrm{s}(v_n, v_1) & \cdots & \mathrm{s}(v_n, v_n) \end{pmatrix}$$

die Darstellungsmatrix von s bzgl. $\mathfrak{V}$.

Bemerkung 5.22.

1. Eine Basis $\mathfrak{V}$ von V ist genau dann eine Orthonormalbasis, wenn $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{s}) = \mathrm{Id}$.
2. Ist V euklidisch, dann gilt $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{s}) = \mathrm{Mat}_{\mathfrak{V}}^{\mathrm{T}}(\mathrm{s})$. Ist V unitär, dann $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{s}) = \overline{\mathrm{Mat}_{\mathfrak{V}}^{\mathrm{T}}(\mathrm{s})}$. ◊

Definition 5.23.

Sei $A \in \mathfrak{M}_{\mathbb{R}}(n, n)$. Dann heißt A^{T} die zu A adjungierte Matrix.

Ist dagegen $A \in \mathfrak{M}_{\mathbb{C}}(n, n)$, dann heißt $\overline{A^{\mathrm{T}}}$ die zu A adjungierte Matrix.

$A \in \mathfrak{I}_{\mathbb{R}}(n)$ heißt orthogonal, falls $A^{\mathrm{T}} = A^{-1}$. $A \in \mathfrak{I}_{\mathbb{C}}(n)$ heißt unitär, falls $\overline{A^{\mathrm{T}}} = A^{-1}$.

Bemerkung 5.24.

1. $A \in \mathfrak{M}(n, n)$ ist also genau dann orthogonal bzw. unitär, wenn $\mathrm{Lin}(A)$ orthogonal bzw. unitär ist.
2. A ist genau dann orthogonal bzw. unitär, wenn für alle $x, y \in \mathbb{K}^n$ gilt: $\langle Ax, Ay\rangle = \langle x, y\rangle$.
3. A ist genau dann orthogonal bzw. unitär, wenn die Spalten von A eine Orthonormalbasis des $\mathbb{K}^n$ bilden.

   Setze nämlich $f = \mathrm{Lin}(A)$, dann sind die Spalten von A gegeben durch $\mathfrak{W} = (f(e^{(1)}), ..., f(e^{(n)}))$.
4. Ist A orthogonal bzw. unitär, dann ist $|\det A| = 1$, denn

   $$|\det A|^2 = \overline{\det A} \det A = \det \overline{A} \det A = \det \overline{A^{\mathrm{T}}} \det A = \det \overline{A^{\mathrm{T}}} A = \det \mathrm{Id} = 1.$$

   Also gilt auch jeden orthogonalen bzw. unitären Endomorphismus $f \in \mathrm{Hom}(V, V)$, dass $|\det f| = 1$.
5. Die orthogonalen bzw. unitären Matrizen aus $\mathfrak{M}(n, n)$ bilden eine Untergruppe von $\mathfrak{I}(n)$, die orthogonale bzw. unitäre Gruppe, bezeichnet mit $\mathfrak{O}(n)$ bzw. $\mathfrak{U}(n)$. Diese Gruppen sind kanonisch isomorph zu $\mathfrak{O}(V, \mathrm{s})$ bzw. $\mathfrak{U}(V, \mathrm{s})$ via Lin. ◊

Satz 5.25. (Spektralsatz für unitäre Endomorphismen)

Seien (V, s) ein unitärer Raum und $f \in \mathrm{Hom}(V, V)$ unitär.

Dann besitzt V eine Orthonormalbasis aus Eigenvektoren zu f.

Beweis.

Das charakteristische Polynom von f zerfällt in Linearfaktoren: $\chi_f = (-1)^n (t - \lambda_1) \cdots (t - \lambda_n)$ mit $\lambda_1, ..., \lambda_n \in \mathbb{C}$ (nicht unbedingt paarweise verschieden). Sei v ein normierter Eigenvektor zum Eigenwert λ. Dann gilt: $V = \mathrm{span}(v) \circledast \mathrm{span}(v)^{\perp}$. Setze $W = \mathrm{span}(v)$ und $U = W^{\perp}$, dann gilt $f(W) \subseteq W$. Wir zeigen: $f(U) \subseteq U$. Da außerdem $\dim(U) = n - 1$, folgt induktiv, dass sich V als orthogonale Summe aus Eigenräumen zu f schreiben lässt.

Sei $u \in U$, dann gilt:

$$0 = \mathrm{s}(v, u) = \mathrm{s}(f(v), f(u)) = \mathrm{s}(\lambda v, f(u)) = \overline{\lambda} \mathrm{s}(v, f(u)),$$

also $f(u) \perp v$, d.h. $f(u) \in W^{\perp} = U$. Also ist $f(U) \subseteq U$. □

Bemerkung 5.26. (Normalform für unitäre Endomorphismen)

Sei $f \in \mathfrak{U}(V)$. Wir wissen bereits, dass die Eigenwerte dann ± 1 sind, d.h. es gibt eine Basis $\mathfrak{V} = (v_1, ..., v_n)$ von f mit

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \begin{pmatrix} \mathrm{diag}(1) & 0 \\ 0 & \mathrm{diag}(-1) \end{pmatrix}. \qquad \Diamond$$

Bemerkung 5.27.

1. Sei $\mathfrak{V}$ eine Orthonormalbasis des $\mathbb{K}^n$. Dann ist $\Psi_{\mathfrak{V}}$ orthogonal:
$$\mathrm{Mat}(\Psi_{\mathfrak{V}}) = (\Psi_{\mathfrak{V}}(e^{(1)}), ..., \Psi_{\mathfrak{V}}(e^{(n)})) = (v_1, ..., v_n)^{-1}.$$
Da die Spalten von $A = (v_1, ..., v_n)$ eine Orthonormalbasis bilden, ist A orthogonal, d.h. es gilt $\mathrm{Mat}(\Psi_{\mathfrak{V}})^{-1} = A = \overline{A^{\mathrm{T}}}^{-1} = \overline{\mathrm{Mat}(\Psi_{\mathfrak{V}})^{\mathrm{T}}}$.
2. Damit gilt insbesondere: Ist $\mathfrak{V}$ eine Orthonormalbasis des $\mathbb{K}^n$, dann kann die Darstellungsmatrix eines Endomorphismus auf $\mathfrak{V}$ transformiert werden, ohne die Inverse der Transformationsmatrix berechnen zu müssen:
$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{E}}(\mathrm{id})\mathrm{Mat}(f)\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{id}) = (v_1, ..., v_n)^{\mathrm{T}}\mathrm{Mat}(f)(v_1, ..., v_n).$$
3. Zu jedem $A \in \mathfrak{U}(n)$ gibt es ein $P \in \mathfrak{U}(n)$ mit $\overline{P^{\mathrm{T}}}AP = \mathrm{diag}(\lambda_1, ..., \lambda_n)$. Dabei gilt $AP^{(i)} = \lambda_i P^{(i)}$.
4. Zu $f \in \mathfrak{U}(V, \mathrm{s})$ seien $\lambda_1, ..., \lambda_k$ die verschiedenen Eigenwerte. Dann lässt sich V als orthogonale Summe der zugehörigen Eigenräume darstellen: $V = \mathrm{Eig}(f, \lambda_1) \circledast \cdots \circledast \mathrm{Eig}(f, \lambda_k)$. $\Diamond$

Bemerkung 5.28.

Orthogonale Endomorphismen $f \in \mathfrak{O}(V)$ erhalten Längen und auch Winkel:

$$\sphericalangle(f(v), f(w)) = \cos^{-1}\frac{\mathrm{s}(f(v), f(w))}{||f(v)||||f(w)||} = \cos^{-1}\frac{\mathrm{s}(v, w)}{||v||||w||} = \sphericalangle(v, w).$$

Sie sind also Kongruenzabbildungen.

Wir werden im Folgenden sehen, dass diese sich mittels Rotationen in Ebenen um feste Winkel und Streckungen längs Streckachsen um gewisse Faktoren ausdrücken lassen. $\Diamond$

Beispiel 5.29.

Sei (V, s) ein euklidischer Raum.

1. Wir betrachten eine Spiegelung an einer Hyperebene von (V, s), d.h. an einem Untervektorraum der Dimension $n - 1$.

   Sei $v \in V$ mit $||v|| = 1$. Dann definiert $H = \mathrm{span}(v)^{\perp}$ eine Hyperebene. v heißt der Normaleneinheitsvektor von H. Definiere $f : V \to V$ durch $f(v) = -v$ und $f|_H = \mathrm{id}$. Sei weiter $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V mit $v_1 = v$. Dann ist
$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \begin{pmatrix} -1 & 0 \\ 0 & \mathrm{Id}(n-1) \end{pmatrix}$$
und es gilt $\det \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = -1$.
2. Wir betrachten eine Drehung in einer Ebene von V um den Winkel θ.

   Sei U ein zweidimensionaler Unterraum von V mit Orthonormalbasis (v_1, v_2). Definiere $f : V \to V$ durch $f(v_1) = \cos\theta v_1 + \sin\theta v_2$, $f(v_2) = -\sin\theta v_1 + \cos\theta v_2$ und $f|_{U^{\perp}} = \mathrm{id}$. Sei weiter $(v_3, ..., v_n)$ eine Basis von $U^{\perp}$, dann ist $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V und es gilt
$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \begin{pmatrix} \begin{matrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{matrix} & 0 \\ 0 & \mathrm{Id}(n-2) \end{pmatrix} = \begin{pmatrix} \mathrm{Rot}(\theta) & 0 \\ 0 & \mathrm{Id}(n-2) \end{pmatrix}$$
mit $\det \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \cos^2(\theta) + \sin^2(\theta) = 1$.

   Speziell im Fall $\dim_{\mathbb{R}}(V) = 3$ heißen θ der Rotationswinkel und $U^{\perp}$ die Rotationsachse von f. $\Diamond$

Satz 5.30. (Spektralsatz für orthogonale Endomorphismen)
Seien (V, s) ein euklidischer Raum und $f \in \mathrm{Hom}(V, V)$ orthogonal.
Dann gibt es eine orthogonale Zerlegung $V = U_1 \circledast \cdots \circledast U_m$, so dass $\dim U_i \in \{1, 2\}$ und alle U_i f-invariant sind.

Beweis.
Nach dem Korollar zum Satz von Hamilton-Cayley gibt es einen f-invarianten Untervektorraum W von V mit $1 \le \dim W \le 2$. Wir zeigen: $f(W^\perp) \subseteq W^\perp$, dann ist $V = W \circledast W^\perp$ und per Induktion folgt, dass sich f in eine orthogonale Summe f-invarianter Unterräume der Dimensionen eins oder zwei zerlegen lässt.

Da f orthogonal, ist f injektiv, d.h. $f(W) = W$, und invertierbar mit f^{-1} orthogonal. Seien $w \in W$ und $v \in W^\perp$. Dann gilt

$$\mathrm{s}(f(v), w) = \mathrm{s}(f^{-1}(f(v)), f^{-1}(w)) = \mathrm{s}(f(v), \underbrace{f^{-1}(w)}_{\in W}) = 0,$$

d.h. $f(v) \in W^\perp$, also $f(W^\perp) \perp W$ und damit $f(W^\perp) \subseteq W^\perp$. $\square$

Bemerkung 5.31. (Normalform orthogonaler Endomorphismen)
Seien (V, s) ein euklidischer Raum und f ein orthogonaler Endomorphismus von V. Dann gibt es eine Orthonormalbasis $\mathfrak{V}$ von V und Rotationswinkel $\theta_1, ..., \theta_k$ mit

$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \begin{pmatrix} \mathrm{Id} & & & & \\ & -\mathrm{Id} & & & 0 \\ & & \mathrm{Rot}(\theta_1) & & \\ & 0 & & \ddots & \\ & & & & \mathrm{Rot}(\theta_k) \end{pmatrix}.$$

Wir wissen bereits, dass sich V als orthogonale Summe f-invarianter, ein- oder zweidimensionaler Unterräume schreiben lässt. Die eindimensionalen Unterräume $U = \lambda\mathbb{R}$ werden wegen $f(U) = \lambda\mathbb{R}$ von Eigenvektoren aufgespannt. Wir müssen also nur zeigen, dass sich f, eingeschränkt auf einen der zweidimensionalen, f-invarianten Unterräume U, als Rotationsmatrix schreiben lässt.

Seien also $g = f|_U$ und $\mathfrak{W} = (v_1, v_2)$ eine Orthonormalbasis von U. Dann ist $R = \mathrm{Mat}^{\mathfrak{W}}_{\mathfrak{W}}(g)$ orthogonal:

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Longrightarrow \quad R^{\mathrm{T}} R = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Wegen $a^2 + c^2 = 1$ und $b^2 + d^2 = 1$ gibt es $\alpha, \beta \in [0, 2\pi)$ mit $a = \sin\alpha,\ b = \sin\beta,\ c = \cos\alpha,\ d = \cos\beta$, d.h.

$$R = \begin{pmatrix} \sin\alpha & \sin\beta \\ \cos\alpha & \cos\beta \end{pmatrix}.$$

Weiter folgt aus $ab + cd = 0$, dass

$$0 = \cos\alpha\sin\alpha + \cos\beta\sin\beta = \sin(\alpha + \beta) \qquad \Longrightarrow \qquad \alpha + \beta \in \mathbb{Z}\pi.$$

Im Fall $\alpha + \beta \in \pi 2\mathbb{Z}$ sind $b = \sin(-\alpha) = \sin\alpha$ und $d = \cos(-\alpha) = -\cos\alpha$, d.h. $R = \mathrm{Rot}(\alpha)$:

$$R = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Ist dagegen $\alpha + \beta \in \pi(2\mathbb{Z} + 1)$, dann $b = \sin(\pi - \alpha) = \sin\alpha$ und $d = \cos(\pi - \alpha) = -\cos\alpha$, d.h.

$$R = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Seien dann w_1, w_2 die beiden Eigenvektoren von R, dann besitzt g bzgl. der Basis $\mathfrak{W}' = (w_1, w_2)$ die Darstellung $\mathrm{Mat}^{\mathfrak{W}'}_{\mathfrak{W}'}(g) = \mathrm{diag}(1, -1)$. $\Diamond$

Bemerkung 5.32.
Zu jedem $A \in \mathfrak{O}(n)$ gibt es ein $P \in \mathfrak{O}(n)$, so dass $\overline{P^{\mathrm{T}}} AP$ obige Normalform hat. $\Diamond$

5.4. Selbstadjungierte Endomorphismen

Eine weitere Klasse von linearen Abbildungen, die mit dem Skalarprodukt interagieren, sind die selbstadjungierten Endomorphismen. Hier ist es erlaubt, den Endomorphismus im Skalarprodukt von der linken auf die rechte Seite zu "schieben". Die zugehörigen Darstellungsmatrizen sind im euklidischen wie im unitären Fall durch orthogonale Transformation diagonalisierbar; auch über $\mathbb{C}$ sind alle Eigenwerte sind reell.

Bemerkung 5.33.

1. Sei $A \in \mathfrak{M}(n,n)$. Dann existiert genau eine Matrix $A^\star \in \mathfrak{M}(n,n)$, so dass für alle $x, y \in \mathbb{K}^n$ gilt $\langle Ax, y\rangle = \langle x, A^\star y\rangle$: Im Fall $\mathbb{K} = \mathbb{R}$ definiere $A^\star = A^{\mathrm{T}}$; im komplexen Fall $\mathbb{K} = \mathbb{C}$ setze $A = \overline{A}^{\mathrm{T}} = \overline{A^{\mathrm{T}}}$. $A^\star$ heißt die adjungierte Matrix zu A.

   Zur Eindeutigkeit: Speziell für $x = e^{(i)}$ und $y = e^{(j)}$ erhalten wir
   $$a_{ij} = \langle Ae^{(i)}, e^{(j)}\rangle = \langle e^{(i)}, A^\star e^{(j)}\rangle = \overline{\langle A^\star e^{(j)}, e^{(i)}\rangle} = \overline{a^\star_{ji}}.$$

2. Sei nun (V, s) ein euklidischer oder unitärer Raum, dann existiert auch zu jedem $f \in \mathrm{Hom}(V,V)$ genau ein $f^\star \in \mathrm{Hom}(V,V)$, so dass für alle $v, w \in V$ gilt $\mathrm{s}(f(v), w) = \mathrm{s}(v, f^\star(w))$. $f^\star$ heißt der adjungierte Endomorphismus zu f. $\Diamond$

Definition 5.34.

Seien (V, s) ein euklidischer oder unitärer Raum und $f \in \mathrm{Hom}(V,V)$.

f heißt selbstadjungiert, falls für alle $v, w \in V$ gilt: $\mathrm{s}(f(v), w) = \mathrm{s}(v, f(w))$.

Im euklidischen Fall nennt man f auch symmetrisch, im euklidischen Fall Hermitesch

Bemerkung 5.35.

Für die Darstellungsmatrizen selbstadjungierter Abbildungen gilt:
$$\langle \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}} e^{(i)}, e^{(j)}\rangle = \mathrm{s}(f(v_i), v_j) = \mathrm{s}(v_i, f(v_j)) = \overline{\mathrm{s}(f(v_j), v_i)} = \langle \overline{\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)}^{\mathrm{T}} e^{(i)}, e^{(j)}\rangle,$$
d.h. $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)^\star$. $\Diamond$

Definition 5.36.

$A \in \mathfrak{M}(n,n)$ heißt selbstadjungiert, falls $A = A^\star$.

Im euklidischen Fall heißt A auch symmetrisch, im unitären Fall Hermitesch.

Bemerkung 5.37.

1. Genau dann ist A selbstadjungiert, wenn $\mathrm{Lin}(A)$ selbstadjungiert ist.
2. Genau dann ist A selbstadjungiert, wenn für alle $x, y \in \mathbb{K}^n$ gilt: $\langle Ax, y\rangle = \langle x, Ay\rangle$. $\Diamond$

Lemma 5.38.

Das charakteristische Polynom χ_A einer selbstadjungierten Matrix $A \in \mathfrak{M}(n,n)$ hat nur reelle Nullstellen, faktorisiert also über $\mathbb{R}$.

Beweis.

Seien $\lambda \in \mathbb{C}$ eine Nullstelle von χ_A und $x \in \mathbb{C}^n \backslash \{0\}$ mit $(A - \lambda \mathrm{Id})x = 0$. Dann gilt:
$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle,$$
also $\lambda = \overline{\lambda}$ und damit $\lambda \in \mathbb{R}$. $\Diamond$

Satz 5.39. (Spektralsatz für selbstadjungierte Endomorphismen)
Seien (V, s) ein euklidischer oder unitärer Raum und $f \in \mathrm{Hom}(V, V)$ selbstadjungiert.
Dann besitzt V eine Orthonormalbasis aus Eigenvektoren zu f.

Beweis.
Wir führen eine Induktion über n. Der Fall $n = 1$ ist trivial. Gelte die Behauptung also für $n - 1$ und sei $\lambda \in \mathbb{R}$ ein Eigenwert von f ist. Sei $v \in V$ ein normierter Eigenvektor zu λ und sei $U = \mathrm{span}(v)$, dann gilt $f(U) \subseteq U$. Weiter gilt $f(U^\perp) \subseteq U^\perp$, denn sind $v \in U^\perp$ & $u \in U$, dann $\mathrm{s}(f(v), u) = \mathrm{s}(v, f(u)) = 0$, d.h. $f(v) \perp u$. Damit ist $f|_{U^\perp} \in \mathrm{Hom}(U^\perp, U^\perp)$ und nach Induktionsvoraussetzung hat $f_{U^\perp}$ eine Orthonormalbasis $(v_2, ..., v_n)$ aus Eigenvektoren zu f. Wir ergänzen mit v zu einer Orthonormalbasis $\mathfrak{V} = (v_1, ..., v_n)$ von V aus Eigenvektoren von f. □

Bemerkung 5.40. (Normalform für selbstadjungierte Endomorphismen)
Sei $f \in \mathrm{Hom}(V, V)$ selbstadjungiert, dann existieren Eigenwerte $\lambda_1, ..., \lambda_n \in \mathbb{R}$ zu f und eine Basis aus zugehörigen Eigenvektoren $\mathfrak{V} = (v_1, ..., v_n)$ mit

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda. \qquad \lozenge$$

Bemerkung 5.41.

1. Seien (V, s) ein euklidischer bzw. unitärer Raum und $A \in \mathfrak{M}(n, n)$ selbstadjungiert. Dann gibt es ein $P \in \mathfrak{O}(n)$ bzw. $P \in \mathfrak{U}(n)$ mit $P^\star AP = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ und $AP^{(i)} = \lambda_i P^{(i)}$, wobei alle λ_i reell sind.
2. Seien (V, s) ein euklidischer oder Raum und $f \in \mathrm{Hom}(V, V)$ selbstadjungiet. Dann lässt sich V zerlegen in $V = \mathrm{Eig}(f, \lambda_1) \circledast \cdots \circledast \mathrm{Eig}(f, \lambda_m)$, wobei $\lambda_1, ..., \lambda_m \in \mathbb{R}$ die paarweise verschiedenen Eigenwerte von f sind. $\lozenge$

5.5. Die Hauptachsentransformation

In diesem Abschnitt behandeln wir Darstellungsmatrizen für Bilinearformen bzw. Hermitesche Formen. Wir werden sehen, dass symmetrische Formen sich stets mittels geeigneter Basen diagonalisieren lassen; diese werden Hauptachsen genannt. Darstellungsmatrizen zu Bilinearformen entscheiden sich in einigen Punkten erheblich von denen zu Endomorphismen; beispielsweise sind die Eigenwerte nicht invariant unter Transformationen. Der Sylvesterschen Trägheitssatz besagt aber zumindest, dass sich bei einem Basiswechsel die Anzahl der positiven und der negativen Eigenwerte einer Bilinearform nicht ändert.

Wir setzen ab jetzt stets $V = \mathbb{K}^n$ und $\mathrm{s} = \langle \cdot, \cdot \rangle$. Weiter sei b eine symmetrische Bilinearform bzw. eine Hermitesche Form.

Definition 5.42.
Sei $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V. Dann heißt

$$\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) = (\mathrm{b}(v_i, v_j))_{1 \le j \le n}^{1 \le i \le n} = \begin{pmatrix} \mathrm{b}(v_1, v_1) & \cdots & \mathrm{b}(v_1, v_n) \\ \vdots & \ddots & \vdots \\ \mathrm{b}(v_n, v_1) & \cdots & \mathrm{b}(v_n, v_n) \end{pmatrix}$$

die Darstellungsmatrix von b bzgl. $\mathfrak{V}$.

Bemerkung 5.43.

1. $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ bedeutet gerade , dass $\mathfrak{V}$ ein b-Orthogonalsystem bildet.
2. Seien $v, w \in V$ mit zugehörigen Koeffizientenvektoren $x, y \in \mathbb{K}^n$. Dann gilt:

$$\mathrm{b}(v, w) = \mathrm{b}\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i,j=1}^n x_i \overline{y_j} \mathrm{b}(v_i, v_j) = \langle x, \mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) y \rangle.$$

3. Seien $\mathfrak{W}$ eine weitere Basis von V und seien u, z die zugehörigen Koeffizientenvektoren. Dann gilt:

$$\begin{aligned} \mathrm{b}(v, w) &= \langle x, \mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) y\rangle \\ &= \langle u, \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b}) z\rangle \\ &= \langle \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) x, \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b}) \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) y\rangle \\ &= \langle x, \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id})^\star \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b}) \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) y\rangle, \end{aligned}$$

   d.h. $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id})^\star \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b}) \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) = \mathrm{Mat}_{\mathfrak{V}}(\mathrm{b})$.

   Beachte: Die Rücktransformation wird bei Bilinearformen also durch die Adjungierte vermittelt, nicht wie bei linearen Abbildungen durch die Inverse.

4. $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b})$ ist selbstadjungiert, da b symmetrisch bzw. Hermitesch ist. $\Diamond$

Korollar 5.44. (Hauptachsentransformation)
Sei $\mathfrak{W}$ eine Basis von V und seien $\lambda_1, ..., \lambda_n$ die Eigenwerte von $\mathrm{Mat}_{\mathfrak{W}}(\mathrm{b})$.

Dann existiert eine Basis $\mathfrak{V}$ von V, so dass die Transformation $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id})$ orthogonal bzw. unitär ist und $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ gilt.

Die Basiselemente $v_1, ..., v_n$ von $\mathfrak{V}$ heißen Hauptachsen, die Eigenwerte $\lambda_1, ..., \lambda_n$ Hauptträgheitsmomente.

Beweis.
Sei $A = \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b})$ da A selbstadjungiert ist, gibt es eine Transformationsmatrix $P \in \mathfrak{O}(n)$ bzw. $P \in \mathfrak{U}(n)$ mit $P^{-1}AP = \mathrm{diag}(\lambda_1, ..., \lambda_n) = \Lambda$, wobei alle $\lambda_i \in \mathbb{R}$. Wähle $v_i = \sum p_{ji} w_j$, dann ist $P = \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id})$:

$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) e^{(i)} = \Psi_{\mathfrak{W}}(v_i) = \Psi_{\mathfrak{W}}\left(\sum_{j=1}^n p_{ji} w_j\right) = \sum_{j=1}^n p_{ji} \Psi_{\mathfrak{W}}(w_j) = \sum_{j=1}^n p_{ji} e^{(j)} = P e^{(i)}.$$

Wegen $P^{-1} = P^\star$ ist schließlich

$$\Lambda = P^{-1}AP = \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id})^\star \mathrm{Mat}_{\mathfrak{W}}(\mathrm{b}) \mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{W}}(\mathrm{id}) = \mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}).$$ $\square$

Bemerkung 5.45.
Die Hauptträgheitsmomente $\lambda_1, ..., \lambda_n$ sind zu gegebenem Bezugssystem $\mathfrak{W}$ eindeutig bestimmt. Für eine andere Ausgangsbasis $\mathfrak{W}$ ergeben sich aber in der Regel andere Eigenwerte, man sollte daher nicht von Eigenwerten der Bilinearform bzw. Hermiteschen Form b sprechen, sondern nur von Eigenwerten bzgl. des Bezugssystems $\mathfrak{W}$. $\Diamond$

Beispiel 5.46. (aus der Theorie starrer Körper in der Physik)
Sei $B \in \mathfrak{M}_{\mathbb{R}}(3, 3)$ symmetrisch, der *Trägheitstensor* eines *starren Körpers* bzgl. einer fixierten Orthonormalbasis $\mathfrak{W}$. Weiter sei B *positiv definit*, d.h. für alle $x \in \mathbb{R}^3 \backslash \{0\}$ gelte $\langle x, Bx\rangle > 0$. Dann ist $\mathrm{b}(x, y) = \langle x, By\rangle = x^{\mathrm{T}} By$ ein Skalarprodukt. Wir definieren eine *quadratische Form* $\mathrm{q} : \mathbb{R}^3 \to \mathbb{R}$ durch $\mathrm{q}(x) = \mathrm{b}(x, x)$, das sog. *Trägheitsmoment* bei einer Drehung um die Achse x.

$\mathfrak{W}$ lässt sich mittels Haupsachsentransformation in eine Orthonormalbasis $\mathfrak{V}$ *drehen*, so dass für alle $x = \sum \alpha_i v_i$ gilt:

$$\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) = \Lambda, \qquad \mathrm{q}(x) = x^{\mathrm{T}} Bx = \sum_{i=1}^3 \lambda_i \alpha_i^2 = \alpha^{\mathrm{T}} \Lambda \alpha,$$

wobei $\lambda_i = \mathrm{q}(v_i)$, die Hauptträgheitsmomente und die Basiselemente v_1, v_2, v_3 von $\mathfrak{V}$ die Hauptachsen sind. Die Vektoren $x \in \mathbb{R}^3$ mit $\mathrm{q}(x) = 1$ bilden einen *Trägheitselipsoid*. Dieser wird durch Drehungen um Hauptachsen nicht deformiert. $\Diamond$

Korollar 5.47.

Seien b eine symmetrisch Bilinearform oder Hermitesche Form auf dem $(\mathbb{K}^n, \langle\cdot,\cdot\rangle)$ und $\mathfrak{W} = (w_1, ..., w_n)$ eine Basis des $\mathbb{K}^n$.

Dann gilt: b ist genau dann positiv definit, wenn es eine Basis $\mathfrak{V}$ des $\mathbb{K}^n$ gibt, so dass alle Eigenwerte von $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b})$ positiv sind.

Beweis.

Sei $\mathrm{Mat}_{\mathfrak{V}}(\mathrm{b}) = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ mit $\lambda_i \in \mathbb{R}$ und $v = \sum x_i v_i \in V$ beliebig. Dann gilt:

$$\mathrm{b}(v, v) = \overline{x}^{\mathrm{T}} \Lambda x = \sum_{i=1}^{n} \lambda_i \overline{x}_i x_i = \sum_{i=1}^{n} \lambda_i |x_i|^2,$$

Also gilt $\mathrm{b}(v, v) > 0$ für alle $v \neq 0$ genau dann, wenn alle $\lambda_i > 0$ sind. □

Bemerkung 5.48.

Ist also b ein Skalarprodukt, dann sind die Eigenwerte von b bzgl. jeder Ausgangsbasis $\mathfrak{W}$ positiv.

Allgemeiner gilt: ◊

Satz 5.49. (Sylvesterscher Trägheitssatz)

Sei b eine symmetrische Bilibearform oder Hermitesche Form auf dem $\mathbb{K}^n$ und seien $\mathfrak{V}_1, \mathfrak{V}_2$ zwei Basen des $\mathbb{K}^n$. Seien p_k die Anzahl der positiven und n_k die der negativen Eigenwerte von $\mathrm{Mat}_{\mathfrak{V}_k}(\mathrm{b}),\ k = 1, 2$.

Dann gelten $p_1 = p_2,\ n_1 = n_2$ und $\mathrm{rg}(\mathrm{Mat}_{\mathfrak{V}_1}(\mathrm{b})) = \mathrm{rg}(\mathrm{Mat}_{\mathfrak{V}_2}(\mathrm{b}))$.

Beweis.

$\mathrm{rg}(\mathrm{Mat}_{\mathfrak{V}_1}(\mathrm{b})) = \mathrm{rg}(\mathrm{Mat}_{\mathfrak{V}_2}(\mathrm{b}))$ ist klar, da für die invertierbare Transformationsmatrix $\mathrm{Mat}^{\mathfrak{V}_1}_{\mathfrak{V}_2}(\mathrm{id})$ gilt:

$$\mathrm{Mat}^{\mathfrak{V}_1}_{\mathfrak{V}_2}(\mathrm{id})^\star \mathrm{Mat}_{\mathfrak{V}_2}(\mathrm{b}) \mathrm{Mat}^{\mathfrak{V}_1}_{\mathfrak{V}_2}(\mathrm{id}) = \mathrm{Mat}_{\mathfrak{V}_1}(\mathrm{b}).$$

Für $k = 1, 2$ seien $P_k \in \mathfrak{O}(n)$ bzw. $\mathfrak{U}(n)$ mit $P_k^\star \mathrm{Mat}_{\mathfrak{V}_k}(\mathrm{b}) P_k = \Lambda_k$. Setze $w_{k,i} = \sum p_{k,ji} v_{k,i}$, dann bildet $\mathfrak{W}_k$ eine Basis des $\mathbb{K}^n$ aus Eigenvektoren zu $\mathrm{Mat}_{\mathfrak{V}_k}$ mit $P_k = \mathrm{Mat}^{\mathfrak{W}_k}_{\mathfrak{V}_k}(\mathrm{id})$.

Bezeichne $w^+_{k,i}$ diejenigen Eigenvektoren mit zugehörigem positivem Eigenwert $\lambda^+_{k,i} > 0$, entsprechend $w^-_{k,i}$ mit $\lambda^-_{k,i} < 0$ und $w^0_{k,i}$ mit $\lambda^0_{k,i} = 0$. Seien V^+_k, V^-_k und V^0_k die von den entsprechenden Eigenvektoren aufgespannten Eigenräume. Dann sind $p_k = \dim V^+_k$, $n_k = \dim V^-_k$ und $V = V^+_k \circledast V^-_k \circledast V^0_k$.

Setze $V^0 = \{v \in V \mid \forall w \in V : \mathrm{b}(v, w) = 0\}$, dann ist $V^0_1 = V^0_2 = V^0$: Seien zunächst $v = \sum x_i w_{k,i} \in V^0_k$ und $w = \sum y_i w_{k,i} \in \mathbb{K}^n$ beliebig, dann ist $\mathrm{b}(v, w) = \overline{x}^{\mathrm{T}} \Lambda y = \sum \lambda_{k,i} x_{k,i} y_{k,i} = 0$ wegen $x^+_{k,i} = x^-_{k,i} = 0$ und $\lambda^0_{k,i} = 0$, d.h. $v \in V^0$ und damit $V^0_k \subseteq V^0$. Sei umgekehrt $v = \sum x_i w_{k,i} \in V^0$. Angenommen, $v \notin V^0_k$, dann gibt es ein $x_j \neq 0$ mit $\lambda_{k,j} \neq 0$. Wähle $y = e^{(j)}$ und $w = \sum y_i w_{k,i}$, dann $\mathrm{b}(v, w) = \overline{x}^{\mathrm{T}} \Lambda y = \lambda_j x_j \neq 0$, ein Widerspruch. Also ist auch $V^0 \subseteq V^0_k$, d.h. $V^0 = V^0_k$. Setze $o = \dim V^0$.

Es gilt $V^+_2 \cap (V^-_1 \oplus V^0) = \{0\}$, denn seien $v = \sum x_i w_{2,i} \in V^+_2 \backslash \{0\}$ und $w = \sum y_i w_{1,i} \in V^-_1 \oplus V^0$, dann $\mathrm{b}(v, v) = \sum |x_i|^2 \lambda_i > 0$ und $\mathrm{b}(w, w) = \sum |y_i|^2 \lambda_i \leq 0$, d.h. $x \neq y$. Also ist

$$p_2 + n_1 + o = \dim(V^+_2 \oplus (V^0 \oplus V^-_1)) \leq n = \dim(V^+_1 \oplus (V^0 \oplus V^-_1)) = p_1 + n_1 + o,$$

d.h. $p_2 \leq p_1$. Entsprechend ist $V^+_1 \cap (V^-_2 \oplus V^0) = \{0\}$, also

$$p_1 + n_2 + o = \dim(V^+_1 \oplus (V^0 \oplus V^-_2)) \leq n = \dim(V^+_2 \oplus (V^0 \oplus V^-_2)) = p_2 + n_2 + o,$$

d.h. $p_1 \leq p_2$. Damit ist insgesamt $p_1 = p_2$ und folglich auch $n_1 = n_2$. □

Bemerkung 5.50.

Seien b eine symmetrische Bilinearform oder Hermitesche Form und $\mathrm{q} : x \mapsto \mathrm{b}(x, x)$ die zugehörige quadratische Form.

Dann lässt sich V in eine orthogonale Summe

$$V = V^+ \circledast V^- \circledast V^0$$

zerlegen, wobei $\mathrm{q}(v) > 0$ für alle $V \in V^+ \backslash \{0\}$, $\mathrm{q}(v) < 0$ für alle $v \in V^- \backslash \{0\}$ und $\mathrm{q}(v) = 0$ für alle $v \in V^0$.

Der Raum V^0 sowie die Dimensionen von V^+, V^- sind dabei eindeutig bestimmt (nicht jedoch die Räume V^+, V^- selbst). Die Elemente aus $V^0 \backslash \{0\}$ heißen b-isotrop. ◊

Satz 5.51. (Hauptminorenkriterium)

Sei $B = (b_{ij})_{1 \leq j \leq n}^{1 \leq i \leq n} \in \mathfrak{M}(n, n)$ selbstadjungiert. Dann sind äquivalent:

1. B ist positiv definit, d.h. für alle $v \in \mathbb{K}^n \backslash \{0\}$ gilt $\overline{v}^{\mathrm{T}} B v = \langle v, Bv \rangle > 0$.
2. Alle Hauptminoren $\det B_k$, $B_k = (b_{ij})_{1 \leq j \leq k}^{1 \leq i \leq k}$, sind positiv.

Beweis.

1. Sei B positiv definit. Wir setzen $\mathrm{b}(v, w) = \langle v, Bw \rangle$. Da B selbstadjungiert, definiert dies eine symmetrische Bilinearform bzw. Hermitesche Form mit $\mathrm{Mat}_{\mathfrak{E}(n)}(\mathrm{b}) = B$. Dann ist auch der k-te Hauptminor von B selbstadjungiert:
$$B_k = (\mathrm{b}(e^{(i)}, e^{(j)}))_{1 \leq j \leq k}^{1 \leq i \leq k} = \mathrm{Mat}_{\mathfrak{E}(k)}(\mathrm{b}).$$
Hauptachsentransformation liefert ein $P \in \mathfrak{O}(n)$ bzw. $\mathfrak{U}(n)$ mit $P^\star B_k P = \Lambda$ und $\lambda_1, ..., \lambda_k > 0$, also auch
$$\det B_k = \det(P^\star B P) = \prod_{i=1}^{k} \lambda_i > 0.$$

2. Sei umgekehrt $k \geq 1$ maximal gewählt, so dass $\mathrm{Mat}_{\mathfrak{E}(k)}(\mathrm{b}) = B_k$ positiv definit ist. Angenommen, $k < n$. Wähle bzgl. b eine Orthogonalbasis $\mathfrak{V}(k+1) = (v_1, ..., v_{k+1})$, so dass
$$\mathrm{Mat}_{\mathfrak{V}(k+1)}(\mathrm{b}) = P^\star \mathrm{Mat}_{\mathfrak{E}(k+1)}(\mathrm{b}) P = \Lambda(k+1), \qquad P = \mathrm{Mat}_{\mathfrak{V}(k+1)}^{\mathfrak{E}(k+1)}(\mathrm{id}).$$
Nach Voraussetzung ist $\det B_{k+1} = \prod \lambda_i > 0$, d.h. es gibt $i, j \in \{1, ..., k+1\}$, $i \neq j$, mit $\lambda_i < 0$, $\lambda_j < 0$. Seien $\alpha, \beta \in \mathbb{K}^{k+1}$ die Koeffizientenvektoren der zugehörigen Eigenvektoren. Da k maximal gewählt, sind $v_i, v_j \notin \mathrm{span}(\mathfrak{E}(k))$, d.h. $\alpha_{k+1}, \beta_{k+1} \neq 0$. Setze $w = \beta_{m+1} v_i - \alpha_{m+1} v_j$, dann ist $w \in \mathrm{span}(\mathfrak{E}(k))$, d.h.
$$0 < \mathrm{b}(w, w) = |\beta_{m+1}|^2 \lambda_i + |\alpha_{m+1}|^2 \lambda_j < 0,$$
ein Widerspruch. □

Index

Literatur

[1] Balser, W.: *Lineare Algebra I.* Vorlesungsskript Universität Ulm, p. 90, 2008.

[2] Balser, W.: *Lineare Algebra II.* Vorlesungsskript Universität Ulm, p. 60, 2008.

[3] Beutelspacher, A.: *Lineare Algebra – Eine Einführung.* Vierweg+Teubner Verlag, 7th ed., 2010.

[4] Bosch, S.: *Lineare Algebra.* Springer-Verlag, 4th ed., 2008.

[5] Bruns, W.: *Lineare Algebra.* Osnabrücker Schriften zur Mathematik, p. 137, 2001.

[6] Dieck, T. t.: *Lineare Algebra.* Vorlesungsskript Uni Göttingen, p. 134, 2004.

[7] Ferus, D.: *Lineare Algebra I.* Vorlesungsskript Technische Universität Berlin, p. 87, 2001.

[8] Ferus, D.: *Lineare Algebra II.* Vorlesungsskript Technische Universität Berlin, p. 140, 2002.

[9] Fischer, G.: *Lineare Algebra und analytische Geometrie.* Vierweg+Teubner Verlag, 2011.

[10] Gräter, J. & Rosenberger, E.: *Lineare Algebra.* Vorlesungsskript Universität Potsdam, p. 119, 2004.

[11] Gubisch, M.: *Lineare Algebra II.* Vorlesungsskript Universität Konstanz, p. 58, 2006.

[12] Gubisch, M.: *Tutorium zur Linearen Algebra I.* Vorlesungsskript Universität Konstanz, p. 44, 2010.

[13] Hilgert, J.: *Lineare Algebra I & II.* Vorlesungsskript Universität Paderborn, p. 291, 2011.

[14] Huber-Klawitter, A.: *Lineare Algebra.* Vorlesungsskript Universität Ulm, p. 100, 2009.

[15] Johannson, K.: *Lineare Algebra I.* Vorlesungsskript Universität Frankfurt, p. 204, 2003.

[16] Johannson, K.: *Lineare Algebra II.* Vorlesungsskript Universität Frankfurt, p. 111, 2003.

[17] Kappel, F.: *Lineare Algebra I & II.* Vorlesungsskript Universität Graz, p. 241, 2006.

[18] Kersten, I.: *Analytische Geometrie & Lineare Algebra.* Vorlesungsskript Uni Göttingen, p. 255, 2001.

[19] Knauf, A.: *Lineare Algebra & Analytische Geometrie I.* Vorlesungsskript Uni Nürnberg, p. 109, 2001.

[20] Knauf, A.: *Lineare Algebra & Analytische Geometrie II.* Vorlesungsskript Uni Nürnberg, p. 92, 2001.

[21] Kowalsky, H.-J.: *Lineare Algebra.* De Gruyter Lehrbuch, 7th ed., 1975.

[22] Lau, D.: *Übungsbuch zur Linearen Algebra & analytischen Geometrie.* Springer, 2nd ed., 2011.

[23] Menth, M.: *Lineare Algebra.* Vorlesungsskript Universität Würzburg, p. 412, 1997.

[24] Mutzbauer, O.: *Lineare Algebra.* Vorlesungsskript Universität Würzburg, p. 141, 2007.

[25] Nieper-Wißkirchen, M. A.: *Lineare Algebra I.* Vorlesungsskript Universität Augsburg, p. 168, 2009.

[26] Schaal, H.: *Lineare Algebra & Analytische Geometrie I.* Vieweg & Sohn Verlagsgesellschaft, 1976.

[27] Schaal, H.: *Lineare Algebra & Analytische Geometrie II.* Vieweg & Sohn Verlagsgesellschaft, 1976.

[28] Schaal, H.: *Lineare Algebra & Analytische Geometrie III.* Vieweg & Sohn Verlagsgesellschaft, 1977.

[29] Scheiderer, C.: *Lineare Algebra I und II.* Vorlesungsskript Universität Duisburg, p. 248, 2001.

[30] Schnürer, O. C.: *Lineare Algebra I.* Vorlesungsskript Universität Konstanz, p. 79, 2011.

[31] Schnürer, O. C.: *Lineare Algebra II.* Vorlesungsskript Universität Konstanz, p. 52, 2011.

[32] Schulz, R.-H.: *Lineare Algebra I.* Vorlesungsskript Freie Universität Berlin, p. 204, 2008.

[33] Spindler, H.: *Lineare Algebra.* Vorlesungsskript Universität Osnabrück, p. 236, 2004.

[34] Stammbach, U.: *Lineare Algebra.* Vorlesungsskript ETH Zürich, p. 208, 2002.

[35] Stoppel, H. & Griese, B.: *Übungsbuch zur Linearen Algebra.* Vieweg Studium, 5th ed., 2005.

[36] Vogt, R.: *Lineare Algebra.* Vorlesungsskript Universität Osnabrück, p. 170, 2003.

[37] Wille, D.: *Repetitorium der Linearen Algebra 1.* Binomi Verlag, 2003.

[38] Wille, D. & Holz, M.: *Repetitorium der Linearen Algebra 2.* Binomi Verlag, 2002.