Lineare Algebra II

Christopher Frei
LMU München

Skript zur Vorlesung im Sommersemester 2016

Inhaltsverzeichnis

Kapitel 1

Euklidische und unitäre Vektorräume

Bisher wurden Vektorräume über beliebigen Körpern K untersucht. In diesem Kapitel schränken wir uns auf Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$ ein. Wir können dann Längen und Winkel (zumindest über $\mathbb{R}$) von Vektoren definieren und den Begriff der Orthogonalität von Vektoren einführen.

Literatur: das Kapitel basiert hauptsächlich auf den entsprechenden Kapiteln in [1, 3].

Notation 1.0.1.

- $\mathbb{K}$ *bezeichnet einen der Körper* $\mathbb{R}$ *oder* $\mathbb{C}$*, d.h.* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$
- *für* $\alpha \in \mathbb{K}$ *ist* $\overline{\alpha}$ *die komplex Konjugierte, d.h. wenn* $\alpha = a + bi$*, mit* $a, b \in \mathbb{R}$*, dann ist* $\overline{\alpha} = a - bi$*. (Im Fall* $\mathbb{K} = \mathbb{R}$ *gilt also* $\overline{\alpha} = \alpha$*). Weiters schreiben wir* $a = \mathrm{Re}(\alpha)$*,* $b = \mathrm{Im}(\alpha)$ *für Real- und Imaginärteil.*
- *für* $\alpha \in \mathbb{K}$ *ist* $|\alpha|$ *der gewöhnliche Absolutbetrag, d.h.*

$$|\alpha| = \sqrt{\alpha\overline{\alpha}} = \sqrt{\mathrm{Re}(\alpha)^2 + \mathrm{Im}(\alpha)^2}.$$

1.1 Innere Produkte

Sei V ein $\mathbb{K}$-Vektorraum. Wir wollen Begriffe einführen, die die Länge eines Vektors oder den Winkel zwischen zwei Vektoren beschreiben.

Definition 1.1.1. *Sei* V *ein* $\mathbb{K}$*-Vektorraum. Ein* inneres Produkt *oder* Skalarprodukt *auf* V *ist eine Abbildung* $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$*,* $(v, w) \mapsto \langle v, w \rangle$*, die für alle* $v, v_1, v_2, w \in V$ *und* $\alpha \in K$ *folgende Bedingungen erfüllt:*

1. *Linear im ersten Argument:*

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle \text{ und } \langle \alpha v, w\rangle = \alpha\langle v, w\rangle$$

2. *Hermitesch:* $\langle v, w\rangle = \overline{\langle w, v\rangle}$

3. *Positiv definit:* $\langle v, v\rangle \geq 0$ *und* $\langle v, v\rangle = 0 \Leftrightarrow v = 0$.

Das Paar $(V, \langle\cdot,\cdot\rangle)$ *heißt ein* innerer Produktraum*, im Fall* $\mathbb{K} = \mathbb{R}$ *ein* euklidischer Raum*, im Fall* $\mathbb{K} = \mathbb{C}$ *ein* unitärer Raum*. Wir werden oft darauf verzichten, das innere Produkt explizit anzugeben und kurz* V *für den euklidischen oder unitären Raum* $(V, \langle\cdot,\cdot\rangle)$ *schreiben.*

Bemerkung 1.1.2.

1. *Jedes innere Produkt ist* semilinear *im zweiten Argument, das heißt für* $\alpha_1, \alpha_2 \in \mathbb{K}, v, w_1, w_2 \in V$ *gilt*

$$\langle v, \alpha_1 w_1 + \alpha_2 w_2\rangle = \overline{\alpha_1}\langle v, w_1\rangle + \overline{\alpha_2}\langle v, w_2\rangle.$$

   Beweis: aus 1. und 2. folgt

$$\begin{aligned}\langle v, \alpha_1 w_1 + \alpha_2 w_2\rangle &= \overline{\langle \alpha_1 w_1 + \alpha_2 w_2, v\rangle} = \overline{\alpha_1}\overline{\langle w_1, v\rangle} + \overline{\alpha_2}\overline{\langle w_2, v\rangle} \\ &= \overline{\alpha_1}\langle v, w_1\rangle + \overline{\alpha_2}\langle v, w_2\rangle.\end{aligned}$$

2. *Für* $\mathbb{K} = \mathbb{R}$ *bedeuten 1. und 2., dass* $\langle\cdot,\cdot\rangle$ *eine symmetrische Bilinearform ist. Ein inneres Produkt eines* $\mathbb{R}$*-Vektorraums ist also eine symmetrische positiv definite Bilinearform.*

3. *Sei* $\mathbb{K} = \mathbb{C}$*. Eine Sesquilinearform (sesqui = eineinhalb) ist eine Abbildung* $V \times V \to \mathbb{C}$*, die linear im ersten und semilinear im zweiten Argument ist. Ein inneres Produkt ist also eine hermitesche positiv definite Sesquilinearform.*

4. *Für den Nullvektor* $0 \in V$ *und* $v \in V$ *gilt stets* $\langle 0, v\rangle = \langle v, 0\rangle = 0$.

   Beweis: $\langle 0, v\rangle = \langle 0 + 0, v\rangle = \langle 0, v\rangle + \langle 0, v\rangle$*. Also* $\langle 0, v\rangle = 0$*. Analog* $\langle v, 0\rangle = \langle v, 0 + 0\rangle = \langle v, 0\rangle + \langle v, 0\rangle$.

5. *Sei* V *ein endlich-dimensionaler* $\mathbb{K}$*-Vektorraum mit Basis* $\{v_1, \ldots, v_n\}$*. Ein inneres Produkt* $\langle\cdot,\cdot\rangle$ *auf* V *ist durch die Werte* $\langle v_i, v_j\rangle$*,* $1 \leq i, j \leq n$ *eindeutig bestimmt. Für* $v = \sum_{i=1}^n a_i v_i$ *und* $w = \sum_{i=1}^n b_i v_i$ *gilt*

$$\langle v, w\rangle = \sum_{i=1}^n \sum_{j=1}^n a_i \overline{b_j}\langle v_i, v_j\rangle.$$

Beispiel 1.1.3.

1. *Sei $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^n$. Für $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{R}^n$ definieren wir ein inneres Produkt durch*

$$\langle v, w\rangle := \sum_{i=1}^{n} v_i w_i.$$

Wenn wir v, w als Spaltenvektoren betrachten, gilt also $\langle v, w\rangle = v^t w$. Dieses innere Produkt heißt das Standardskalarprodukt *auf $\mathbb{R}^n$.*

2. *Sei $\mathbb{K} = \mathbb{C}$ und $V = \mathbb{C}^n$. Für $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{C}^n$ definieren wir ein inneres Produkt durch*

$$\langle v, w\rangle := \sum_{i=1}^{n} v_i \overline{w_i}.$$

Wenn wir v, w als Spaltenvektoren betrachten, gilt also $\langle v, w\rangle = v^t \overline{w}$, wobei $\overline{w} := (\overline{w_1}, \ldots, \overline{w_n})$. Dieses innere Produkt heißt das Standardskalarprodukt *auf $\mathbb{C}^n$.*

Die Konjugation der w_i ist notwendig, um positive Definitheit zu erreichen. Es gilt

$$\langle v, v\rangle = \sum_{i=1}^{n} v_i \overline{v_i} = \sum_{i=1}^{n} |v_i|^2 \geq 0,$$

mit Gleichheit genau dann, wenn $v = 0$. Aufgrund dieser Konjugation ist $\langle \cdot, \cdot \rangle$ im zweiten Argument nicht linear sondern semilinear.

3. *Sei $V = \mathcal{C}([0, 1], \mathbb{K})$ der $\mathbb{K}$-Vektorraum der stetigen Funktionen von $[0, 1]$ nach $\mathbb{K}$. Dann definiert*

$$\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}\mathrm{d}t$$

ein inneres Produkt auf V.

Beweis: 1. und 2. sind klar. Für 3. sehen wir, dass

$$\langle f, f\rangle = \int_0^1 |f(t)|^2 \mathrm{d}t \geq 0,$$

da $|f(t)|^2 \geq 0$ für alle $t \in [0, 1]$. Wenn $f \neq 0$, gibt es ein t mit $|f(t)|^2 > 0$, und da mit f auch $|f(\cdot)|^2$ stetig ist, folgt $\langle f, f\rangle > 0$.

4. *Sei $L : V \to W$ eine injektive lineare Abbildung zwischen $\mathbb{K}$-Vektorräumen und $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf W. Dann definiert*

$$\langle v, w \rangle_L := \langle L(v), L(w) \rangle$$

ein inneres Produkt auf V. (Injektivität von L wird für den Nachweis der positiven Definitheit benötigt.)

5. *Matrixversion des letzten Beispiels. Sei $A \in M(m, n; \mathbb{K})$ vom Rang n. Sei $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $\mathbb{K}^m$. Dann definiert*

$$\langle v, w \rangle_A := \langle Av, Aw \rangle$$

ein inneres Produkt auf $\mathbb{K}^n$. Wir bemerken, dass

$$\langle v, w \rangle_A = (Av)^t(\overline{Aw}) = v^t(A^t\overline{A})\overline{w}.$$

6. *Konkreter Spezialfall des letzten Beispiels. Sei $V = \mathbb{K}^2$, $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt, und $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Dann ist*

$$\langle v, w \rangle_A = \langle (v_1, v_1 + v_2), (w_1, w_1 + w_2) \rangle = v_1\overline{w_1} + (v_1 + v_2)(\overline{w_1} + \overline{w_2})$$

ein inneres Produkt auf $\mathbb{K}^2$.

7. *Weiterer Spezialfall. Sei V ein euklidischer oder unitärer Raum und $W \subset V$ ein Untervektorraum. Dann ist die Einschränkung des inneren Produkts $\langle \cdot, \cdot \rangle$ auf $W \times W$ ein inneres Produkt auf W.*

Wir können ein inneres Produkt verwenden, um die Länge eines Vektors, sowie den Abstand zwischen Vektoren zu definieren.

Definition 1.1.4. *Sei V ein euklidischer oder unitärer Raum und $v, w \in V$. Dann ist*

$$\|v\| := \sqrt{\langle v, v \rangle}$$

die durch das innere Produkt induzierte Norm *von v. Weiters ist*

$$d(v, w) := \|v - w\|$$

der Abstand *zwischen v und w.*

Beispiel 1.1.5. *Für $V = \mathbb{R}^n$ und $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt erhalten wir*

$$\|v\| = \sqrt{\sum_{i=1}^{n} v_i^2}.$$

Nach dem Satz von Pythagoras ist das der übliche Begriff der Länge eines Vektors.

Lemma 1.1.6 (Cauchy-Schwarz'sche Ungleichung)**.** *Sei V ein euklidischer oder unitärer Raum und $v, w \in V$. Dann gilt*

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

mit Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis. Für $w = 0$ lautet die Ungleichung $0 \leq 0$. Wir nehmen also an, dass $w \neq 0$. Für $\alpha, \beta \in \mathbb{K}$ gilt

$$0 \leq \langle \alpha v + \beta w, \alpha v + \beta w \rangle = \alpha\overline{\alpha}\langle v, v \rangle + \alpha\overline{\beta}\langle v, w \rangle + \beta\overline{\alpha}\langle w, v \rangle + \beta\overline{\beta}\langle w, w \rangle. \quad (1.1)$$

Mit $\alpha := \langle w, w \rangle$ und $\beta := -\langle v, w \rangle$ ist die rechte Seite gleich

$$\begin{aligned} &\|w\|^4\|v\|^2 - \|w\|^2\overline{\langle v, w \rangle}\langle v, w \rangle - \|w\|^2\langle v, w \rangle\overline{\langle v, w \rangle} + \|w\|^2\langle v, w \rangle\overline{\langle v, w \rangle} \\ =&\|w\|^2(\|v\|^2\|w\|^2 - |\langle v, w \rangle|^2). \end{aligned}$$

Wir dividieren durch $\|w\|^2 > 0$ und sehen so, dass (1.1) äquivalent zu

$$0 \leq \|v\|^2\|w\|^2 - |\langle v, w \rangle|^2 \quad (1.2)$$

ist. Wurzelziehen zeigt die geforderte Ungleichung.

Gilt Gleichheit in (1.2), dann auch in (1.1), also $\alpha v + \beta w = 0$. Sind umgekehrt v, w linear abhängig, dann gibt es $\gamma \in \mathbb{K}$ mit $v = \gamma w$ (da $w \neq 0$), also

$$\|v\|\|w\| = \sqrt{\gamma\overline{\gamma}\langle w, w \rangle}\sqrt{\langle w, w \rangle} = |\gamma|\langle w, w \rangle = |\gamma\langle w, w \rangle| = |\langle v, w \rangle|.$$

□

Korollar 1.1.7. *Sei V ein euklidischer oder unitärer Raum. Die Norm $\|\cdot\|$ hat für alle $v, w \in V$, $\alpha \in \mathbb{K}$ die Eigenschaften*

1. *$\|v\| \geq 0$ und $\|v\| = 0 \Leftrightarrow v = 0$*

2. $\|\alpha v\| = |\alpha| \cdot \|v\|$

3. $\|v + w\| \leq \|v\| + \|w\|$.

Weiters hat der Abstand $d(\cdot, \cdot)$ für alle $v, w, z \in V$ die Eigenschaften

4. $d(v, w) \geq 0$ *und* $d(v, w) = 0 \Leftrightarrow v = w$

5. $d(v, w) = d(w, v)$

6. $d(v, z) \leq d(v, w) + d(w, z)$.

Beweis. 1. ist äquivalent zur positiven Definitheit von $\langle \cdot, \cdot \rangle$. 2. gilt, da $\|\alpha v\| = \sqrt{\alpha\overline{\alpha}\langle v, v\rangle} = |\alpha|\|v\|$. Für 3. berechnen wir mit Lemma 1.1.6 und der Tatsache, dass $z + \overline{z} = 2\,\mathrm{Re}(z)$, dass

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle = \|v\|^2 + \langle v, w\rangle + \langle w, v\rangle + \|w\|^2 \\ &= \|v\|^2 + 2\,\mathrm{Re}(\langle v, w\rangle) + \|w\|^2 \leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.\end{aligned}$$

4. folgt aus 1. und 5. folgt aus 2. Weiters folgt 6. aus 3., da

$$d(v, z) = \|v - w + w - z\| \leq \|v - w\| + \|w - z\| = d(v, w) + d(w, z).$$

□

Definition 1.1.8. *Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit den Eigenschaften 1.,2.,3. aus Korollar 1.1.7 heißt* Vektornorm *auf V, und das Paar $(V, \|\cdot\|)$ heißt* normierter Raum.

Sei V jetzt eine beliebige Menge. Eine Abbildung $d : V \times V \to \mathbb{R}$ mit den Eigenschaften 4., 5., 6. aus Korollar 1.1.7 heißt Metrik, und das Paar (V, d) heißt metrischer Raum.

Bemerkung 1.1.9.

1. *Wir haben also gezeigt, dass jedes innere Produkt auf einem $\mathbb{K}$-Vektorraum V eine Norm $\|\cdot\|$ und eine Metrik $d(\cdot, \cdot)$ auf V induziert, und daher jeder euklidische oder unitäre Raum auch ein normierter Raum und ein metrischer Raum ist.*

2. *In den Übungen werden wir sehen, dass nicht jede Norm auf einem $\mathbb{K}$-Vektorraum V durch ein inneres Produkt induziert ist.*

Wir haben gesehen, wie wir mit inneren Produkten Längen (Norm) und Abstände zwischen Vektoren definieren können. Was ist mit Winkeln?

Beispiel 1.1.10. *Sei V ein euklidischer Vektorraum. Für $v, w \in V$ gilt laut der Cauchy-Schwarz'schen Ungleichung, dass*

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\|w\|} \leq 1.$$

Es gibt also ein eindeutiges $\phi \in [0, \pi]$ mit

$$\cos(\phi) = \frac{\langle v, w\rangle}{\|v\|\|w\|}.$$

Wir nennen ϕ den Winkel *zwischen v und w, und schreiben $\phi = \measuredangle(v, w)$.*

Im Fall $V = \mathbb{R}^2$ mit dem Standardskalarprodukt stimmt das mit dem Winkelbegriff aus der Analysis überein: für $v, w \in \mathbb{R}^2$ mit $\|v\| = \|w\| = 1$ (d.h. v, w liegen am Einheitskreis) gibt es eindeutige $\alpha, \beta \in [0, 2\pi)$ mit

$$v = (\cos(\alpha), \sin(\alpha)) \text{ und } w = (\cos(\beta), \sin(\beta)).$$

Wir können annehmen, dass $\beta > \alpha$, denn falls nicht, können wir v, w vertauschen. Der durch v, w eingeschlossene Winkel $\phi \in [0, \pi)$ ist dann

$$\phi = \begin{cases} \beta - \alpha & \text{falls } \beta - \alpha \in [0, \pi] \\ 2\pi - (\beta - \alpha) & \text{falls } \beta - \alpha \in [\pi, 2\pi). \end{cases}$$

Andererseits gilt aufgrund des Additionstheorems für den Cosinus, dass

$$\langle v, w\rangle = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \cos(\beta - \alpha) = \cos(2\pi - (\beta - \alpha)) = \cos(\phi).$$

Für beliebige $v, w \in \mathbb{R}^2$ sei

$$v' := \frac{1}{\|v\|} \cdot v, \quad w' := \frac{1}{\|w\|} \cdot w,$$

dann liegen v', w' am Einheitskreis, und $\measuredangle(v, w) = \measuredangle(v', w')$.

Im letzten Beispiel ist genau dann $\measuredangle(v, w) = \pi/2$ (also $90°$), wenn $\langle v, w\rangle = 0$. Das motiviert folgende Definition von Orthogonalität (die auch im Fall $\mathbb{K} = \mathbb{C}$ gilt).

Definition 1.1.11. *Sei V ein euklidischer oder unitärer Raum.*

1. *Vektoren $v, w \in V$ heißen* orthogonal, *wenn $\langle v, w\rangle = 0$.*
2. *Eine Menge $M \subset V$ von Vektoren heißt* orthogonal, *wenn je zwei verschiedene Vektoren aus M orthogonal sind.*

3. *Eine Menge $M \subset V$ heißt* orthonormal*, wenn M orthogonal ist und $\|v\| = 1$ für jedes $v \in M$ gilt. Ein v mit $\|v\| = 1$ heißt* normiert*.*

4. *Eine* Orthogonalbasis *(bzw.* Orthonormalbasis*) von V ist eine Basis von V die orthogonal (bzw. orthonormal) ist.*

5. *Für $M \subset V$ ist das* orthogonale Komplement *(auch: der Orthogonalraum) zu M definiert als*

$$M^\perp := \{v \in V \mid \langle v, m\rangle = 0 \text{ für alle } m \in M\}.$$

Bemerkung 1.1.12. *Das orthogonale Komplement jeder Teilmenge von V ist ein Untervektorraum von V. Wenn $v \in V$, bezeichnen wir $\{v\}^\perp$ auch als das orthogonale Komplement (oder den Orthogonalraum) zu v und schreiben $v^\perp$ statt $\{v\}^\perp$.*

Beispiel 1.1.13.

1. *Der Nullvektor $0 \in V$ ist orthogonal zu jedem Vektor und ist der einzige Vektor mit dieser Eigenschaft.*

2. *Sei $V = \mathbb{K}^n$ mit dem Standardskalarprodukt. Dann ist die Standardbasis von $\mathbb{K}^n$ eine Orthonormalbasis.*

3. *Sei $V = \mathbb{K}^2$ mit dem Skalarprodukt $\langle v, w\rangle := v_1\overline{w_1} + (v_1 + v_2)(\overline{w_1} + \overline{w_2})$ aus Beispiel 1.1.3. Dann ist die Standardbasis nicht orthogonal, denn es gilt $\langle e_1, e_2\rangle = 1 \cdot 0 + 1 \cdot 1 = 1$.*

   Die Basis $\{(1, 0), (1, -2)\}$ ist eine Orthogonalbasis, aber nicht orthonormal:

$$\begin{aligned}\langle (1,0), (1,-2)\rangle &= 1 \cdot 1 + 1 \cdot (-1) = 0\\ \|(1,0)\| &= \sqrt{1 \cdot 1 + 1 \cdot 1} = \sqrt{2}\\ \|(1,-2)\| &= \sqrt{1 \cdot 1 + (-1) \cdot (-1)} = \sqrt{2}.\end{aligned}$$

   Wir können diese Basis normieren *(d.h. jeden Vektor durch seine Norm dividieren), um die Orthonormalbasis $\{(1/\sqrt{2}, 0), (1/\sqrt{2}, -\sqrt{2})\}$ zu erhalten.*

4. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Wir definieren ein inneres Produkt auf V, für das B eine Orthonormalbasis ist: sei $\{e_1, \ldots, e_n\}$ die Standardbasis auf $\mathbb{K}^n$,*

$\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $\mathbb{K}^n$, und $L : V \to \mathbb{K}^n$ der Vektorraumisomorphismus mit $L(v_i) = e_i$. Dann hat das innere Produkt

$$\langle v, w \rangle_L = \langle L(v), L(w) \rangle$$

die gewünschte Eigenschaft. Tatsächlich gilt $\langle v_i, v_j \rangle_L = \langle e_i, e_j \rangle = \delta_{ij}$.

Lemma 1.1.14. *Sei V ein euklidischer oder unitärer Raum und sei $\{v_1, \ldots, v_n\}$ eine orthogonale Menge von Vektoren $v_i \neq 0$. Sei*

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n \tag{1.3}$$

eine Linearkombination der v_i, dann gilt

$$\lambda_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2}$$

für alle $1 \leq i \leq n$.

Beweis. Wir bilden das Skalarprodukt der Gleichung (1.3) mit v_i:

$$\langle v, v_i \rangle = \sum_{j=1}^{n} \lambda_j \langle v_j, v_i \rangle = \lambda_i \langle v_i, v_i \rangle.$$

□

Korollar 1.1.15. *Sei V ein euklidischer oder unitärer Raum.*

1. *Jede orthogonale Menge $M \subset V$ mit $0 \notin M$ ist linear unabhängig.*
2. *Sei $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V und $v \in V$. Dann gilt*

$$v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$$

Beweis. Zur ersten Aussage: Für jede Linearkombination

$$\lambda_1 v_1 + \cdots + \lambda_l v_l = 0$$

von paarweise verschiedenen Vektoren $v_i \in M$ gilt laut Lemma 1.1.14, dass

$$\lambda_i = \frac{\langle 0, v_i \rangle}{\|v_i\|^2} = 0.$$

Die zweite Aussage folgt unmittelbar aus Lemma 1.1.14. □

In Beispiel 1.1.13 haben wir gesehen, dass es zu jeder Basis B eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V ein inneres Produkt gibt, für das B eine Orthonormalbasis ist. Der nächste Satz zeigt, dass jedes innere Produkt auf V so entsteht.

Satz 1.1.16 (Orthonormalisierungssatz von Gram-Schmidt)**.** *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $W \subset V$ ein Unterraum. Dann kann jede Orthonormalbasis $\{w_1, \ldots, w_l\}$ von W zu einer Orthonormalbasis $\{w_1, \ldots, w_l, w_{l+1}, \ldots, w_n\}$ von V ergänzt werden.*

Beweis. Ergänze $\{w_1, \ldots, w_l\}$ zu einer Basis $\{w_1, \ldots, w_l, v_{l+1}, \ldots, v_n\}$ von V. Wir konstruieren nacheinander Vektoren $w_{l+1}, \ldots, w_n$, sodass für alle $l \leq i \leq n$, gilt, dass $\{w_1, \ldots, w_i\}$ eine Orthonormalbasis von $\mathrm{Spann}_{\mathbb{K}}(w_1, \ldots, w_l, v_{l+1}, \ldots, v_i)$ ist.

Für $i = l$ ist das bereits der Fall. Sei also $i \geq l+1$ und nehme an, dass wir $w_{l+1}, \ldots, w_{i-1}$ bereits konstruiert haben.

Dann konstruieren wir w_i wie folgt: setze

$$\tilde{w_i} := v_i - \langle v_i, w_1\rangle w_1 - \ldots - \langle v_i, w_{i-1}\rangle w_{i-1}.$$

Dann ist $\{w_1, \ldots, w_{i-1}, \tilde{w_i}\}$ wieder eine Basis von $\mathrm{Spann}_{\mathbb{K}}(w_1, \ldots, w_{i-1}, v_i)$, und es gilt $\langle \tilde{w_i}, w_j\rangle = \langle v_i, w_j\rangle - \langle v_i, w_j\rangle\langle w_j, w_j\rangle = 0$ für $1 \leq j \leq i-1$. Wir müssen $\tilde{w_i}$ also nur noch normieren, und wählen daher

$$w_i := \frac{1}{\|\tilde{w_i}\|} \cdot \tilde{w_i}.$$

□

Bemerkung 1.1.17. *Der Beweis von Satz 1.1.16 ist konstruktiv. Er liefert uns ein Verfahren zur Bestimmung einer Orthonormalbasis von V aus einer gegebenen Basis.*

Korollar 1.1.18. *Jeder endlich-dimensionale euklidische oder unitäre Raum V hat eine Orthonormalbasis.*

Beweis. Wende Satz 1.1.16 auf den Unterraum $W = \{0\}$ an, der die leere Menge als Orthonormalbasis hat. □

Beispiel 1.1.19. *Wir betrachten $\mathbb{R}^4$ mit dem Standardskalarprodukt. Gegeben sei der Unterraum $V = \mathrm{Spann}_{\mathbb{K}}\{v_1, v_2, v_3\}$, mit*

$$v_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \; v_2 = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix}, \; v_3 = \begin{pmatrix} 0 \\ 5 \\ 3 \\ 0 \end{pmatrix}.$$

Wir bestimmen eine Orthonormalbasis von V. Dazu starten wir mit der Basis $\{v_1, v_2, v_3\}$.

$i = 1$: $\{v_1\}$ ist bereits eine Orthogonalbasis von $\mathrm{Spann}_{\mathbb{K}}(v_1)$*, also setzen wir*

$$w_1 := \frac{1}{\|v_1\|} \cdot v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

um eine Orthonormalbasis $\{w_1\}$ zu erhalten.

$i = 2$: Wir berechnen

$$\tilde{w_2} := v_2 - \langle v_2, w_1 \rangle w_1 = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix} - \langle \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 4 \end{pmatrix}.$$

Mit

$$w_2 := \frac{1}{\|\tilde{w_2}\|} \cdot \tilde{w_2} = \begin{pmatrix} 0 \\ 3/5 \\ 0 \\ 4/5 \end{pmatrix}$$

ist dann $\{w_1, w_2\}$ eine Orthonormalbasis von $\mathrm{Spann}_{\mathbb{K}}(v_1, v_2)$*.*

$i = 3$: Wir berechnen

$$\begin{aligned} \tilde{w_3} &:= v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2 \\ &= \begin{pmatrix} 0 \\ 5 \\ 3 \\ 0 \end{pmatrix} - \langle \begin{pmatrix} 0 \\ 5 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \langle \begin{pmatrix} 0 \\ 5 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 0 \\ 4/5 \end{pmatrix} \rangle \begin{pmatrix} 0 \\ 3/5 \\ 0 \\ 4/5 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 5 \\ 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 9/5 \\ 0 \\ 12/5 \end{pmatrix} = \begin{pmatrix} 0 \\ 16/5 \\ 3 \\ -12/5 \end{pmatrix}. \end{aligned}$$

Da $\|\tilde{w_3}\| = \sqrt{256/25 + 9 + 144/25} = \sqrt{25} = 5$ gilt, setzen wir

$$w_3 := \frac{1}{5} \cdot \tilde{w_3} = \begin{pmatrix} 0 \\ 16/25 \\ 3/5 \\ -12/25 \end{pmatrix}.$$

Die gesuchte Orthonormalbasis von V ist $\{w_1, w_2, w_3\}$.

Satz 1.1.20. *Sei V ein euklidischer oder unitärer Raum und $W \subset V$ ein endlich-dimensionaler Untervektorraum. Dann gilt*

1. *Sei $\{w_1, \ldots, w_n\}$ ein Erzeugendensystem von W und $v \in V$. Falls $\langle w_i, v\rangle = 0$ für alle $1 \leq i \leq n$ gilt, dann folgt bereits $v \in W^\perp$.*
2. *$W^\perp$ ist ein Komplement von W, d.h. $V = W \oplus W^\perp$.*
3. *$(W^\perp)^\perp = W$.*

Beweis. Zu *1.* Für jedes $w \in W$ gibt es $\alpha_i \in \mathbb{K}$ mit $w = \alpha_1 w_1 + \ldots + \alpha_l w_l$. Also folgt

$$\langle w, v\rangle = \alpha_1\langle w_1, v\rangle + \cdots + \alpha_l\langle w_l, v\rangle = 0.$$

Zu *2.* Sei $v \in W \cap W^\perp$, dann $\langle v, v\rangle = 0$, also $v = 0$. Daher gilt $W \cap W^\perp = \{0\}$.

Wir müssen noch zeigen, dass $V = W + W^\perp$. Sei dazu $v \in V$. Falls $v \in W$, dann gilt insbesondere $v \in W + W^\perp$. Sei also $v \notin W$. Laut Korollar 1.1.18 gibt es eine Orthonormalbasis $\{w_1, \ldots, w_l\}$ von W. Laut Satz 1.1.16 können wir diese zu einer Orthonormalbasis $\{w_1, \ldots, w_l, w\}$ von $\mathrm{Spann}_{\mathbb{K}}(W \cup \{v\})$ ergänzen. Dann gilt $\langle w_i, w\rangle = 0$ für $1 \leq i \leq l$, also $w \in W^\perp$. Insbesondere ist

$$v = \langle v, w_1\rangle w_1 + \ldots + \langle v, w_l\rangle w_l + \langle v, w\rangle w \in W + W^\perp.$$

Wir haben gezeigt, dass $V = W + W^\perp$ und $W \cap W^\perp = \{0\}$, also $V = W \oplus W^\perp$.

Zu *3.* Nach Definition ist

$$(W^\perp)^\perp = \{v \in V \mid \langle v, w\rangle = 0 \text{ für alle } w \in W^\perp\}.$$

Daher ist offensichtlich $W \subset (W^\perp)^\perp$. Sei umgekehrt $v \in (W^\perp)^\perp$. Wegen *2.* gibt es $w \in W$, $\tilde{w} \in W^\perp$ mit $v = w + \tilde{w}$. Dann gilt

$$0 = \langle \tilde{w}, v\rangle = \langle \tilde{w}, w\rangle + \langle \tilde{w}, \tilde{w}\rangle = \|\tilde{w}\|^2,$$

also $\tilde{w} = 0$ und daher $v = w \in W$. □

1.2 Innere Produkte und Dualraum

Der Dualraum eines K-Vektorraums V wurde in der Linearen Algebra 1 definiert als $V^* := L(V, K)$, der Raum aller linearen Abbildungen $V \to K$. Für endlich-dimensionales V wurde gezeigt, dass $V \cong V^*$, allerdings wurde kein *kanonischer* Isomorphismus konstruiert. Für euklidische Räume, also $\mathbb{R}$-Vektorräume mit einem inneren Produkt $\langle \cdot, \cdot \rangle$, kann man einen kanonischen Isomorphismus angeben.

Dazu betrachten wir allgemeiner einen euklidischen oder unitären Raum V. Da das innere Produkt linear im ersten Argument ist, erhalten wir für jedes $w \in V$ eine lineare Abbildung

$$L_w : V \to K, \quad L_w(v) := \langle v, w \rangle.$$

Also gilt $L_w \in V^*$. Wir zeigen, dass im endlich-dimensionalen Fall jedes Element von V^* diese Form hat.

Satz 1.2.1. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum, und $L \in V^*$. Dann gibt es einen eindeutig bestimmten Vektor $w \in V$, sodass $L = L_w$, d.h.*

$$L(v) = \langle v, w \rangle \quad \text{für alle } v \in V.$$

Beweis. Sei $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V, und setze

$$w := \sum_{i=1}^{n} \overline{L(v_i)} v_i.$$

Dann gilt für $1 \leq j \leq n$, dass

$$L_w(v_j) = \langle v_j, w \rangle = \sum_{i=1}^{n} L(v_i) \langle v_j, v_i \rangle = L(v_j) \langle v_j, v_j \rangle = L(v_j).$$

Da L_w und L auf einer Basis übereinstimmen, gilt also $L = L_w$. Zur Eindeutigkeit, sei $L_w = L_{w'}$. Dann gilt für alle $v \in V$, dass

$$0 = L_w(v) - L_{w'}(v) = \langle v, w - w' \rangle.$$

Insbesondere gilt das für $v = w - w'$, also $\|w - w'\| = 0$, also $w = w'$. □

Korollar 1.2.2. *Wir betrachten einen endlich-dimensionalen euklidischen oder unitären Raum V und die Abbildung $\Phi : V \to V^*$, $w \mapsto L_w$.*

1. *Die Abbildung Φ ist ein* Semiisomorphismus*, d.h. eine bijektive semilineare Abbildung.*
2. *Für euklidische Räume (d.h. $\mathbb{K} = \mathbb{R}$) ist Φ ein Isomorphismus.*

Beweis. Laut Satz 1.2.1 ist Φ bijektiv. Weiters gilt für $v, w_1, w_2 \in V$, $\alpha_1, \alpha_2 \in \mathbb{K}$,

$$\begin{aligned} \Phi(\alpha_1 w_1 + \alpha_2 w_2)(v) &= \langle v, \alpha_1 w_1 + \alpha_2 w_2 \rangle = \overline{\alpha}_1 \langle v, w_1 \rangle + \overline{\alpha}_2 \langle v, w_2 \rangle \\ &= \overline{\alpha}_1 \Phi(w_1)(v) + \overline{\alpha}_2 \Phi(w_2)(v), \end{aligned}$$

also $\Phi(\alpha_1 w_1 + \alpha_2 w_2) = \overline{\alpha}_1 \Phi(w_1) + \overline{\alpha}_2 \Phi(w_2)$. Somit ist Φ semilinear, daher im Fall $\mathbb{K} = \mathbb{R}$ linear. □

Beispiel 1.2.3. *Für unendlich-dimensionale Vektorräume stimmt Satz 1.2.1 nicht immer. Sei $V = \mathbb{C}[x]$, der $\mathbb{C}$-Vektorraum der Polynome, mit dem inneren Produkt*

$$\langle p, q\rangle := \int_0^1 p(t)\overline{q(t)}\mathrm{d}t.$$

Sei $z_0 \in \mathbb{C}$. Wir betrachten die lineare Abbildung $L : V \to \mathbb{C}$, $L(p) = p(z_0)$.

Angenommen, es gibt $q \in V$, sodass $L = L_q$, d.h.

$$L(p) = \langle p, q\rangle \quad \textit{für alle } p \in V$$

gilt. Für jedes $p \in V$ gilt $L((x - z_0)p) = (z_0 - z_0)p(z_0) = 0$, also

$$0 = L((x - z_0)p) = \langle (x - z_0)p, q\rangle = \int_0^1 (t - z_0)p(t)\overline{q(t)}\mathrm{d}t.$$

Wir wählen $p = \overline{(t - z_0)}q$ und erhalten

$$0 = \int_0^1 (t - z_0)\overline{(t - z_0)}q(t)\overline{q(t)}\mathrm{d}t = \|(x - z_0)q\|^2.$$

Daher ist $(x - z_0)q = 0$. Da $(x - z_0) \neq 0$, folgt $q = 0$. Dann gilt

$$L(p) = \langle p, 0\rangle = 0 \quad \textit{für alle } p \in V.$$

Das ist ein Widerspruch, da z.B. $L(1) = 1$.

Wir können die Abbildungen L_w verwenden, um zu jeder linearen Abbildung $L : V \to W$ eine *adjungierte* Abbildung $L^* : W \to V$ zu finden.

Definition 1.2.4. *Sei $L : V \to W$ eine lineare Abbildung zwischen euklidischen oder unitären Räumen. Eine Abbildung $L^* : W \to V$ heißt zu L* adjungiert*, wenn*

$$\langle L(v), w\rangle = \langle v, L^*(w)\rangle \quad \textit{für alle } v \in V, w \in W$$

gilt.

Beispiel 1.2.5. *Sei $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ mit den Standardskalarprodukten. Für $A \in M(m, n; \mathbb{K})$ sei $A^* := \overline{A}^t \in M(n, m; \mathbb{K})$ die komplex konjugierte transponierte Matrix, das heißt $A^* = (\overline{a_{ji}})_{\substack{1\leq j\leq n\\ 1\leq 1\leq m}}$. Wir nennen A^* die* adjungierte Matrix *zu A.*

Dann ist die lineare Abbildung $L_{A^*} : W \to V$, $w \mapsto A^*w$ adjungiert zur linearen Abbildung $L_A : V \to W$, $v \mapsto Av$. In der Tat gilt

$$\langle Av, w\rangle = (Av)^t\overline{w} = v^tA^t\overline{w} = v^t\overline{\overline{A}^t w} = v^t\overline{A^*w} = \langle v, A^*w\rangle.$$

Auch im allgemeinen (endlich-dimensionalen) Fall gibt es immer eine adjungierte Abbildung, und diese ist eindeutig durch L bestimmt.

Satz 1.2.6. *Seien V, W endlich-dimensionale euklidische oder unitäre Räume, und sei $L : V \to W$ eine lineare Abbildung. Dann gibt es eine zu L adjungierte Abbildung $L^* : W \to V$. Diese ist eindeutig bestimmt und linear.*

Beweis. Für jedes $w \in W$ ist die Abbildung $f_w : V \to K$, $f_w(v) = \langle L(v), w\rangle$ linear, also $f_w \in V^*$. Es gibt also nach Satz 1.2.1 ein eindeutiges $v_w \in V$ mit $f_w = L_{v_w} = \langle \cdot, v_w\rangle$. Wir setzen $L^*(w) := v_w$. Dann gilt

$$\langle L(v), w\rangle = f_w(v) = \langle v, v_w\rangle = \langle v, L^*(w)\rangle,$$

also ist L^* zu L adjungiert. Falls $\tilde{L} : W \to V$ eine weitere zu L adjungierte Abbildung ist, dann gilt für alle $v \in V, w \in W$, dass

$$\langle v, \tilde{L}(w)\rangle = \langle L(v), w\rangle = \langle v, L^*(w)\rangle,$$

also $\langle v, \tilde{L}(w) - L^*(w)\rangle = 0$. Insbesondere gilt das für $v = \tilde{L}(w) - L^*(w)$, also $\|\tilde{L}(w) - L^*(w)\| = 0$, also $\tilde{L}(w) = L^*(w)$. Das zeigt die Eindeutigkeit.

Wir müssen noch zeigen, dass L^* linear ist. Sei $v \in V$, $w_1, w_2 \in W$, $\alpha_1, \alpha_2 \in K$. Dann ist

$$\langle v, L^*(\alpha_1 w_1 + \alpha_2 w_2)\rangle = \langle L(v), \alpha_1 w_1 + \alpha_2 w_2\rangle = \overline{\alpha}_1\langle L(v), w_1\rangle + \overline{\alpha}_2\langle L(v), w_2\rangle$$
$$= \overline{\alpha}_1\langle v, L^*(w_1)\rangle + \overline{\alpha}_2\langle v, L^*(w_2)\rangle = \langle v, \alpha_1 L^*(w_1) + \alpha_2 L^*(w_2)\rangle.$$

Daher $\langle v, L^*(\alpha_1 w_1 + \alpha_2 w_2) - \alpha_1 L^*(w_1) - \alpha_2 L^*(w_2)\rangle = 0$. Wir wählen

$$v = L^*(\alpha_1 w_1 + \alpha_2 w_2) - \alpha_1 L^*(w_1) - \alpha_2 L^*(w_2),$$

dann $\|v\| = 0$, also $v = 0$. Das zeight die Linearität. □

Seien V, W endlich-dimensionale K-Vektorräume mit Basen $B_V = \{v_1, \ldots, v_n\}$ bzw. $B_W = \{w_1, \ldots, w_m\}$. Wir erinnern uns an die Definition der *darstellenden Matrix* einer linearen Abbildung $L : V \to W$ bezüglich B_V, B_W als

$$[L]_{B_V, B_W} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \quad \text{mit} \quad L(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Wir können die darstellende Matrix der adjungierten Abbildung bezüglich Orthonormalbasen von V und W ähnlich wie in Beispiel 1.2.5 berechnen.

Lemma 1.2.7. *Seien V, W endlich-dimensionale euklidische oder unitäre Räume mit Orthonormalbasen $B_V = \{v_1, \ldots, v_n\}$ bzw. $B_w = \{w_1, \ldots, w_m\}$. Sei $L : V \to W$ eine lineare Abbildung. Dann gilt*

$$[L^*]_{B_W, B_V} = [L]^*_{B_V, B_W}.$$

Beweis. Sei

$$[L]_{B_V, V_W} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in M(m, n; \mathbb{K}),$$
$$[L^*]_{B_W, B_V} = (b_{ji})_{\substack{1 \le j \le n \\ 1 \le i \le m}} \in M(n, m; \mathbb{K}).$$

Dann gilt

$$L(v_j) = \sum_{1=1}^{m} a_{lj} w_l \quad \text{und} \quad L^*(w_i) = \sum_{l=1}^{n} b_{li} v_l,$$

also

$$a_{ij} = \sum_{l=1}^{m} a_{lj} \langle w_l, w_i \rangle = \langle L(v_j), w_i \rangle = \langle v_j, L^*(w_i) \rangle = \sum_{l=1}^{n} \overline{b_{li}} \langle v_j, v_l \rangle = \overline{b_{ji}}.$$

□

Lemma 1.2.8. *Seien U, V, W endlich-dimensionale euklidische oder unitäre Räume. Für lineare Abbildungen $L_0 : U \to V$, $L_1, L_2 : V \to W$ und $\alpha \in \mathbb{K}$ gilt*

1. $(L_1 + L_2)^* = L_1^* + L_2^*$
2. $(\alpha L_1)^* = \overline{\alpha} L_1^*$
3. $(L_1^*)^* = L_1$.
4. $(L_1 \circ L_0)^* = L_0^* \circ L_1^*$.

Beweis. Seien $A_0 \in M(n, k; \mathbb{K})$, $A_1, A_2 \in M(m, n; \mathbb{K})$. Dann gilt offensichtlich

$$(A_1 + A_2)^* = A_1^* + A_2^*, \quad (\alpha A_1)^* = \overline{\alpha} A_1^*, \quad (A_1^*)^* = A_1.$$

Aus der Linearen Algebra I ist weiters bekannt, dass $(A_1 A_0)^t = A_0^t A_1^t$, also folgt auch $(A_1 A_0)^* = A_0^* A_1^*$. Die Aussagen des Lemmas folgen nun nach Wahl von Orthonormalbasen für U, V, W aus Lemma 1.2.7.

Natürlich lassen sich die Aussagen auch direkt beweisen. Wir zeigen beispielsweise *4.* Für $u \in U, w \in W$ gilt

$$\begin{aligned} \langle (L_1 \circ L_0)(u), w \rangle &= \langle L_1(L_0(u)), w \rangle = \langle L_0(u), L_1^*(w) \rangle \\ &= \langle u, L_0^*(L_1^*(w))) \rangle = \langle u, (L_0^* \circ L_1^*)(w) \rangle, \end{aligned}$$

also $(L_1 \circ L_0)^* = L_0^* \circ L_1^*$. □

Bemerkung 1.2.9. *Sei V ein endlich-dimensionaler unitärer Raum und $L : V \to V$ ein Endomorphismus. Dann ist L^* ebenfalls ein Endomorphismus. Wir betrachten die Endomorphismen*

$$L_1 := \frac{1}{2}(L + L^*) \quad \text{und } L_2 := \frac{1}{2i}(L - L^*).$$

Dann gilt $L_1^* = L_1$, $L_2^* = L_2$, und

$$L = L_1^* + iL_2^*.$$

Die Adjunktion $L \mapsto L^*$ verhält sich also wie die komplexe Konjugation, und die Endomorphismen L_1, L_2 wie Real- und Imaginärteil einer komplexen Zahl. Endomorphismen L mit $L^* = L$ sollten sich zu beliebigen Endomorphismen also wie $\mathbb{R}$ zu $\mathbb{C}$ verhalten.

Definition 1.2.10. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.*

1. *Ein Endomorphismus $L : V \to V$ heißt* selbstadjungiert*, wenn $L^* = L$.*
2. *Eine Matrix $A \in M(n, n; \mathbb{R})$ heißt* symmetrisch*, wenn $A^t = A$.*
3. *Eine Matrix $A \in M(n, n; \mathbb{C})$ heißt* hermitesch*, wenn $A^* = A$.*

Bemerkung 1.2.11.

1. *Ein Endomorphismus $L : V \to V$ ist genau dann selbstadjungiert, wenn*

   $$\langle L(v), w\rangle = \langle v, L(w)\rangle \quad \text{für alle } v, w \in V$$

   gilt.
2. *Ein Endomorphismus $L : V \to V$ ist genau dann selbstadjungiert, wenn die darstellende Matrix $[L]_B$ bezüglich einer Orthonormalbasis B von V symmetrisch (im Fall $\mathbb{K} = \mathbb{R}$) bzw. hermitesch (im Fall $\mathbb{K} = \mathbb{C}$) ist.*

Lemma 1.2.12. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ selbstadjungiert. Dann gilt*

1. *Alle Eigenwerte von L sind reell.*
2. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*
3. *Wenn $V \neq \{0\}$, dann hat L hat mindestens einen Eigenwert.*

Beweis. Zu *1.* Sei $v \in V$ ein Eigenvektor zum Eigenwert λ, d.h. $v \neq 0$ und $L(v) = \lambda v$. Dann gilt

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle L(v), v\rangle = \langle v, L(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle,$$

also $\lambda = \overline{\lambda}$ und daher $\lambda \in \mathbb{R}$.

Zu *2.* Seien $\lambda \neq \mu$ und $v, w \in V$ mit $L(v) = \lambda v$ und $L(w) = \mu w$. Dann gilt

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle L(v), w\rangle = \langle v, L(w)\rangle = \langle v, \mu w\rangle = \overline{\mu}\langle v, w\rangle = \mu\langle v, w\rangle.$$

Da $\lambda \neq \mu$, folgt $\langle v, w\rangle = 0$.

Zu *3.* Sei zuerst $\mathbb{K} = \mathbb{C}$. Da $V \neq \{0\}$, gilt $\dim_{\mathbb{C}}(V) \geq 1$, also ist das charakteristische Polynom von L nicht konstant. Somit hat es eine Nullstelle in $\mathbb{C}$, und diese ist ein Eigenwert von L.

Sei jetzt $\mathbb{K} = \mathbb{R}$. Sei $\dim_{\mathbb{R}}(V) = n \geq 1$. Sei B eine Orthonormalbasis von V und $A = [L]_B \in M(n, n; \mathbb{R})$ die Matrix, die L bezüglich B darstellt. Wir zeigen, dass das charakteristische Polynom $\chi_L = \chi_A = \det(A - X \cdot I_n)$ mindestens eine reelle Nullstelle hat. Da L selbstadjungiert ist, ist A symmetrisch. Wir können A auch als komplexe Matrix betrachten, also als Element von $M(n, n; \mathbb{C})$. Dann ist das charakteristische Polynom immer noch χ_A. Da A symmetrisch ist und nur reelle Einträge hat, ist A hermitesch. Daher ist die lineare Abbildung $L : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto Av$ selbstadjungiert. Sie hat mindestens einen Eigenwert, dieser ist laut *1.* reell. Daher hat χ_A eine reelle Nullstelle. □

Satz 1.2.13 (Spektralsatz)**.** *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ ein selbstadjungierter Endomorphismus. Dann hat V eine Orthonormalbasis aus Eigenvektoren von L.*

Beweis. Wir verwenden Induktion über $n = \dim_{\mathbb{K}}(V)$. Für $n = 0$ ist die Aussage trivial.

Sei also $n \geq 1$. Laut Lemma 1.2.12 hat L einen Eigenwert $\lambda_1 \in \mathbb{K}$. Sei v_1 ein Eigenvektor zu λ_1 mit $\|v_1\| = 1$. Sei $W := \mathrm{Spann}_{\mathbb{K}}(v_1)$. Dann gilt $V = W \oplus W^{\perp}$, also $\dim_{\mathbb{K}}(W^{\perp}) = n - 1$. Weiters gilt $L(W^{\perp}) \subset W^{\perp}$, da für $w \in W^{\perp}$,

$$\langle v_1, L(w)\rangle = \langle L(v_1), w\rangle = \langle \lambda_1 v_1, w\rangle = \lambda_1\langle v_1, w\rangle = 0.$$

Wir wissen bereits, dass $W^{\perp}$ mit der Einschränkung des inneren Produkts $\langle \cdot, \cdot \rangle$ auf V auf $W^{\perp} \times W^{\perp}$ wieder ein euklidischer oder unitärer Raum ist. Da $L(W^{\perp}) \subset W^{\perp}$, definiert die Einschränkung $L|_{W^{\perp}} : W^{\perp} \to W^{\perp}$ einen Endomorphismus von $W^{\perp}$, und dieser ist immer noch selbstadjungiert. Nach Induktionsvoraussetzung gibt es also eine Orthonormalbasis $\{v_2, \ldots, v_n\}$ von

$W^\perp$, die aus Eigenvektoren von $L|_{W^\perp}$ besteht. Natürlich sind $v_2, \ldots, v_n$ auch Eigenvektoren von L, und $\{v_1, v_2, \ldots, v_n\}$ ist eine Orthonormalbasis von $V = W \oplus W^\perp$. □

Bemerkung 1.2.14.

1. *Insbesondere ist L diagonalisierbar.*

2. *Für symmetrische bzw. hermitesche Matrizen $A \in M(n, n; \mathbb{K})$ besagt der Satz folgendes. Es gibt eine Matrix $U \in M(n, n; \mathbb{K})$, deren Spalten eine Orthonormalbasis von $\mathbb{K}^n$ (mit Standardskalarprodukt) bilden, sodass $U^{-1}AU$ eine Diagonalmatrix ist.*

   Beweis: Die durch A bezüglich der Standardbasis dargestellte lineare Abbildung $L_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto Av$ ist selbstadjungiert, daher gibt es eine Orthonormalbasis $\{u_1, \ldots, u_n\}$ von $\mathbb{K}^n$, die aus Eigenvektoren von A besteht. Seien $\lambda_1, \ldots, \lambda_n$ die zugehörigen Eigenwerte und U die Matrix mit Spalten $u_1, \ldots, u_n$. Dann gilt

$$U^{-1}AU = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

   wobei $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *die Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$ ist.*

Matrizen U wie in der letzten Bemerkung haben besondere geometrische Bedeutung, die im folgenden Abschnitt behandelt wird.

1.3 Unitäre Abbildungen

Eine unitäre Abbildung zwischen zwei euklidischen oder unitären Räumen ist eine lineare Abbildung, die mit den Skalarprodukten kompatibel ist. Solche Abbildungen erhalten insbesondere Normen von Vektoren und Winkel zwischen Vektoren. Beispiele sind Drehungen und Spiegelungen im $\mathbb{R}^2$.

Definition 1.3.1. *Seien V, W euklidische oder unitäre Räume. Eine lineare Abbildung $L : V \to W$ heißt* unitär*, wenn*

$$\langle v, w \rangle = \langle L(v), L(w) \rangle \quad \text{für alle } v, w \in V$$

gilt.

Bemerkung 1.3.2. *Sei $L : V \to W$ eine unitäre Abbildung.*

1. *Für $v \in V$ gilt $\|v\| = \|L(v)\|$.*

2. L ist injektiv.

3. Ist L ein Isomorphismus, so ist L^{-1} ebenfalls unitär.

Beweis. 1. ist offensichtlich, 2. folgt, da

$$L(v) = 0 \Leftrightarrow \|L(v)\| = 0 \Leftrightarrow \|v\| = 0 \Leftrightarrow v = 0.$$

Für 3., sei $L(v) = \tilde{v}$, $L(w) = \tilde{w}$. Dann

$$\langle L^{-1}(\tilde{v}), L^{-1}(\tilde{w})\rangle = \langle v, w\rangle = \langle L(v), L(w)\rangle = \langle \tilde{v}, \tilde{w}\rangle.$$

□

Lemma 1.3.3. *Seien V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ ein Endomorphismus. Dann sind folgende Aussagen äquivalent.*

1. L ist unitär

2. $L \circ L^* = \mathrm{id}_V = L^* \circ L$

3. L bildet jede Orthonormalbasis auf eine Orthonormalbasis ab

4. L bildet eine Orthonormalbasis auf eine Orthonormalbasis ab

Beweis. (*1.⇒2.*): L ist injektiv, also ein Isomorphismus. Es gilt für $v, w \in V$

$$\langle L(v), w\rangle = \langle L(v), (L(L^{-1}(w))\rangle = \langle v, L^{-1}(w)\rangle,$$

also $L^{-1} = L^*$.

(*2.⇒3.*): Sei $\{v_1, \ldots, v_n\}$ eine Basis von V. Dann gilt

$$\langle L(v_i), L(v_j)\rangle = \langle v_i, L^*(L(v_j))\rangle = \langle v_i, v_j\rangle.$$

Also ist $\{L(v_1), \ldots, L(v_n)\}$ genau dann eine Orthonormalbasis, wenn $\{v_1, \ldots, v_n\}$ eine ist.

(*3.⇒4.*): trivial.

(*4.⇒1.*): Sei $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V, sodass $\{L(v_1), \ldots, L(v_n)\}$ auch eine Orthonormalbasis ist. Für das innere Produkt $\langle \cdot, \cdot\rangle_L := \langle L(\cdot), L(\cdot)\rangle$ auf V gilt

$$\langle v_i, v_j\rangle_L = \langle L(v_i), L(v_j)\rangle = \delta_{ij} = \langle v_i, v_j\rangle.$$

Da ein inneres Produkt durch seine Werte auf einer Basis vollständig bestimmt ist, folgt $\langle \cdot, \cdot\rangle_L = \langle \cdot, \cdot\rangle$, also

$$\langle L(v), L(w)\rangle = \langle v, w\rangle \quad \text{für alle } v, w \in V.$$

□

Satz 1.3.4. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum. Dann bildet die Menge*

$$\mathrm{U}(V) := \{L : V \to V \mid L \text{ unitär}\}$$

mit der Hintereinanderausführung eine Gruppe, genannt die unitäre Gruppe *von V. Die Menge*

$$\mathrm{SU}(V) := \{L \in \mathrm{U}(V) \mid \det(L) = 1\}$$

ist ein Normalteiler von $\mathrm{U}(V)$, genannt die spezielle unitäre Gruppe *von V.*

Beweis. Jeder unitäre Endomorphismus ist injektiv, also invertierbar. Daher gilt $\mathrm{U}(V) \subset \mathrm{GL}(V)$. Sind $L_1, L_2 \in \mathrm{U}(V)$ dann gilt für alle $v, w \in V$

$$\langle (L_1 \circ L_2)(v), (L_1 \circ L_2)(w)\rangle = \langle L_2(v), L_2(w)\rangle = \langle v, w\rangle,$$

also ist $L_1 \circ L_2 \in \mathrm{U}(V)$. Wir haben bereits gesehen, dass für $L \in \mathrm{U}(V)$ auch $L^{-1} \in \mathrm{U}(V)$ gilt, also ist $\mathrm{U}(V)$ eine Untergruppe von $\mathrm{GL}(V)$. Als Kern des Homomorphismus det ist $\mathrm{SU}(V)$ ein Normalteiler. □

Bemerkung 1.3.5.

1. *Wir werden oft das Verknüpfungssymbol weglassen und L_1L_2 statt $L_1 \circ L_2$ schreiben.*

2. *Für $L \in U(V)$ gilt immer $|\det(L)| = 1$. Denn für eine Orthonormalbasis B gilt*

$$\begin{aligned} 1 = \det(\mathrm{id}_V) &= \det(L^*L) = \det([L^*L]_B) = \det([L^*]_B[L]_B) \\ &= \det([L]_B^*)\det([L]_B) = \overline{\det([L]_B^t)}\det([L]_B) \\ &= \overline{\det([L]_B)}\det([L]_B) = |\det([L]_B)|^2 = |\det(L)|^2. \end{aligned}$$

Jetzt betrachten wir die analogen Begriffe für Matrizen.

Definition 1.3.6. *Wir betrachten den $\mathbb{K}^n$ mit dem Standardskalarprodukt. Eine Matrix $A \in M(n, n; \mathbb{K})$ heißt* orthogonal *(im Fall $\mathbb{K} = \mathbb{R}$) oder* unitär *(im Fall $\mathbb{K} = \mathbb{C}$), falls die lineare Abbildung*

$$L_A : \mathbb{K}^n \to \mathbb{K}^n,\ v \mapsto Av$$

unitär ist.

Lemma 1.3.7. *Wir betrachten $\mathbb{K}^n$ mit dem Standardskalarprodukt. Für eine Matrix $A \in M(n,n;\mathbb{K})$ sind folgende Aussagen äquivalent.*

1. *A ist orthogonal bzw. unitär*
2. *$A^*A = I_n = AA^*$*
3. *Die Spalten von A bilden eine Orthonormalbasis von $\mathbb{K}^n$*

Beweis. (*1.⇒2.*): Wir wissen bereits aus Beispiel 1.2.5, dass $L_{A^*} = L_A^*$. Wenn $L_A : \mathbb{K}^n \to \mathbb{K}^n$ unitär ist, folgt also $L_{A^*A} = L_{A^*} \circ L_A = L_A^* \circ L_A = \mathrm{id}_{\mathbb{K}^n}$, und daher $A^*A = I_n$. Analog $AA^* = I_n$.

(*2.⇒3.*): Sei a_i die i-te Spalte von A, und schreibe $A = (a_{ij})$, $A^* = (b_{ij})$, mit $b_{ij} = \overline{a_{ji}}$. Weiters sei $A^*A = I_n = (c_{ij})$, mit $c_{ij} = \delta_{ij}$. Dann ist

$$\langle a_i, a_j \rangle = a_i^t \overline{a_j} = \sum_{l=1}^{n} a_{li} \overline{a_{lj}} = \sum_{l=1}^{n} \overline{b_{il}} \overline{a_{lj}} = \overline{c_{ij}} = \overline{\delta_{ij}} = \delta_{ij}.$$

(*3.⇒1.*): Seien $\{a_1, \ldots, a_n\}$ die Spalten von A. Dann gilt $L_A(e_i) = a_i$. Daher bildet L_A die Orthonormalbasis $\{e_1, \ldots, e_n\}$ auf eine Orthonormalbasis ab, und ist daher unitär. □

Definition 1.3.8. *Wir definieren die folgenden Matrizengruppen, jeweils mit der Matrizenmultiplikation als Verknüpfung.*

$\mathrm{O}(n) := \{A \in M(n,n;\mathbb{R}) \mid A \textit{ orthogonal}\}$ orthogonale Gruppe *der Ordnung n.*
$\mathrm{U}(n) := \{A \in M(n,n;\mathbb{C}) \mid A \textit{ unitär}\}$ unitäre Gruppe *der Ordnung n.*
$\mathrm{SO}(n) := \{A \in \mathrm{O}(n) \mid \det(A) = 1\}$ spezielle orthogonale Gruppe *der Ordnung n.*
$\mathrm{SU}(n) := \{A \in \mathrm{U}(n) \mid \det(A) = 1\}$ spezielle unitäre Gruppe *der Ordnung n.*

Bemerkung 1.3.9.

1. *$\mathrm{O}(n)$ und $\mathrm{U}(n)$ sind Gruppen, da sie durch die Abbildung $A \mapsto L_A$ mit $U(\mathbb{K}^n)$ identifiziert werden. Genauso werden $\mathrm{SO}(n)$ und $\mathrm{SU}(n)$ mit $\mathrm{SU}(\mathbb{K}^n)$ identifiziert.*
2. *Für $A \in \mathrm{O}(n)$ und $A \in \mathrm{U}(n)$ gilt $|\det(A)| = 1$.*

Lemma 1.3.10. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ ein Endomorphismus. Dann sind folgende Aussagen äquivalent.*

1. *L ist unitär*

2. *Für jede Orthonormalbasis B von V ist die Matrix $[L]_B$ unitär*

3. *Für eine Orthonormalbasis B von V ist die Matrix $[L]_B$ unitär*

Beweis. (*1.⇒2.*): Sei B eine Orthonormalbasis von V. Dann gilt mit Lemma 1.2.7

$$[L]_B^*[L]_B = [L^*]_B[L]_B = [L^*L]_B = [\mathrm{id}_V]_B = I_n,$$

und analog $[L]_B[L]_B^* = I_n$. Also ist $[L]_B$ unitär.

(*2.⇒3.*): trivial.

(*3.⇒1.*): Sei B eine Orthonormalbasis, sodass $[L]_B$ unitär ist. Dann gilt

$$[L^*L]_B = [L]_B[L^*]_B = [L]_B[L]_B^* = I_n,$$

also $L^*L = \mathrm{id}_V$. Analog $LL^* = \mathrm{id}_V$. □

Mit unserem neuen Verständnis unitärer Matrizen können wir die Matrixversion von Satz 1.2.13 wie folgt formulieren.

Korollar 1.3.11 (Matrixversion des Spektralsatzes)**.** *Sei $A \in M(n,n;\mathbb{K})$ eine symmetrische bzw. hermitesche Matrix. Dann gilt*

1. *A hat n reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$ (gezählt mit geometrischer Vielfachheit).*

2. *Wenn $\mathbb{K} = \mathbb{R}$, dann gibt es $U \in \mathrm{SO}(n)$, sodass $U^tAU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

3. *Wenn $\mathbb{K} = \mathbb{C}$, dann gibt es $U \in \mathrm{SU}(n)$ mit $U^*AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

Beweis. Da $A^* = A$, ist der Endomorphismus $L_A : \mathbb{K}^n \to \mathbb{K}^n$ selbstadjungiert. Nach dem Spektralsatz gibt es eine Orthonormalbasis $B = \{u_1, \ldots, u_n\}$ des $\mathbb{K}^n$, die aus Eigenvektoren von A besteht. Sei $U := [\mathrm{id}_{\mathbb{K}^n}]_{B,E}$ die Matrix des Basiswechsels von B zur Standardbasis E, d.h. die Spalten von U sind die Vektoren $u_1, \ldots, u_n$. Dann ist $U \in \mathrm{O}(n)$ (bzw. $U \in \mathrm{U}(n)$), da die Spalten von U eine Orthonormalbasis bilden.

Weiters gilt $|\det(U)| = 1$. Falls $\det(U) \neq 1$, ersetze u_1 durch $u_1' := \det(U)^{-1} \cdot u_1$, dann ist auch $B' := \{u_1', u_2, \ldots, u_n\}$ eine Orthonormalbasis und für die Matrix $U' := [\mathrm{id}_{\mathbb{K}^n}]_{B',E}$ gilt $\det(U') = 1$.

Wir können also annehmen, dass $U \in \mathrm{SO}(n)$, bzw. $U \in \mathrm{SU}(n)$. Sei λ_i der Eigenwert des Eigenvektors u_i. Dann gilt $\lambda_i \in \mathbb{R}$ für $1 \leq i \leq n$, und

$$U^{-1}AU = [\mathrm{id}_{\mathbb{K}^n}]_{E,B} \cdot [L_A]_{E,E} \cdot [\mathrm{id}_{\mathbb{K}^n}]_{B,E} = [L_A]_{B,B} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Wir wissen bereits, dass $U^{-1} = U^*$ für $U \in \mathrm{U}(n)$, also auch $U^{-1} = U^t$ für $U \in \mathrm{O}(n)$ gilt. □

Einfache Eigenschaften unitärer Endomorphismen:

Lemma 1.3.12. *Sei V ein euklidischer oder unitärer Raum und $L : V \to V$ ein unitärer Endomorphismus. Dann gilt:*

1. *Für alle Eigenwerte λ von L gilt $|\lambda| = 1$.*

2. *Eigenvektoren zu verschiedenen Eigenwerten von L sind orthogonal.*

Beweis. Zu *1.* Sei v ein Eigenvektor von L zum Eigenwert λ, dann gilt

$$|\lambda|^2\langle v, v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle L(v), L(v)\rangle = \langle v, v\rangle.$$

Zu *2.* Seien v, w Eigenvektoren zu Eigenwerten $\lambda \neq \mu$ von L. Dann gilt

$$\lambda\overline{\mu}\langle v, w\rangle = \langle L(v), L(w)\rangle = \langle v, w\rangle.$$

Da $\mu\overline{\mu} = |\mu| = 1$, folgt $\overline{\mu} = \mu^{-1} \neq \lambda^{-1}$, also $\lambda\overline{\mu} \neq 1$. Daher $\langle v, w\rangle = 0$. □

Wir bestimmen die unitären Endomorphismen von euklidischen Räumen der Dimensionen 2 und 3.

Lemma 1.3.13. *Die Gruppe* $\mathrm{O}(2)$ *besteht aus den Matrizen*

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \textit{und} \quad \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix},$$

für $\alpha \in [0, 2\pi)$.

Beweis. Sei A eine der gegebenen Matrizen. Dann bilden die Spaltenvektoren von A eine Orthonormalbasis, also $A \in \mathrm{O}(2)$.

Sei jetzt $A \in \mathrm{O}(2)$. Dann gilt $A^tA = I_2$. Wir schreiben

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad A^tA = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

Es gilt genau dann $A^tA = I_2$, wenn

$$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad ab + cd = 0.$$

Aufgrund der ersten beiden Gleichungen gibt es $\alpha, \beta \in [0, 2\pi)$ mit

$$a = \cos\alpha, \quad c = \sin\alpha, \quad b = \sin\beta, \quad d = \cos\beta.$$

Aus der dritten Gleichung folgt

$$0 = \cos\alpha \sin\beta + \sin\alpha \cos\beta = \sin(\alpha + \beta),$$

also $\alpha + \beta = m\pi$, für $m \in \mathbb{Z}$. Wenn m gerade ist, gilt

$$\begin{aligned} b &= \sin\beta = \sin(-\alpha) = -\sin\alpha \\ c &= \cos\beta = \cos(-\alpha) = \cos\alpha, \end{aligned}$$

und A hat die erste Gestalt in der Aussage des Lemmas. Wenn m ungerade ist, gilt

$$\begin{aligned} b &= \sin\beta = \sin(\pi - \alpha) = \sin\alpha \\ c &= \cos\beta = \cos(\pi - \alpha) = -\cos\alpha, \end{aligned}$$

und wir sind im zweiten Fall. □

Satz 1.3.14. *Sei V ein euklidischer Raum mit $\dim_{\mathbb{R}} V = 2$ und $L : V \to V$ ein unitärer Endomorphismus. Dann gilt einer der folgenden Fälle:*

1. *Für jede Orthonormalbasis B von V gibt es $\alpha \in [0, 2\pi)$, sodass*

$$[L]_B = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha. \end{pmatrix} \tag{1.4}$$

   Für Orthonormalbasen B, B' von V und zugehörige $\alpha, \alpha' \in [0, 2\pi)$ gilt $\alpha' = \alpha$ oder $\alpha' = 2\pi - \alpha$.

2. *Es gibt eine Orthonormalbasis B von V mit*

$$[L]_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.5}$$

Bemerkung 1.3.15. *Für $V = \mathbb{R}^2$ haben die Matrizen aus Satz 1.3.14 folgende geometrische Interpretationen:* (1.4) *ist eine Drehung um den Winkel α, und bezüglich einer Orhonormalbasis $B = \{v_1, v_2\}$ ist* (1.5) *eine Spiegelung an der Geraden durch v_1.*

Beweis. (von Satz 1.3.14) Zur Erinnerung an die Lineare Algebra I: die Matrix

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{1.6}$$

beschreibt eine Drehung des $\mathbb{R}^2$ um den Winkel α. Sie hat, als komplexe Matrix betrachtet, die Eigenvektoren

$$\begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ zum Eigenwert } \cos\alpha + i\sin\alpha \text{ und } \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ zum Eigenwert } \cos\alpha - i\sin\alpha.$$

Die Matrix

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \tag{1.7}$$

hat das charakteristische Polynom

$$\chi = (\cos\alpha - X)(-\cos\alpha - X) - \sin(\alpha)^2 = X^2 - 1 = (X+1)(X-1),$$

also die Eigenwerte ± 1.

Sei B eine Orthonormalbasis von V. Dann ist $[L]_B \in \mathrm{O}(2)$, also von der Form (1.6) oder (1.7).

Hat L die Eigenwerte ± 1, dann hat $[L]_B$ die Form (1.7). Seien w_1, w_2 Eigenvektoren zu $1, -1$, mit $\|w_1\| = \|w_2\| = 1$. Wegen Lemma 1.3.12 ist $B' := \{w_1, w_2\}$ dann eine Orthonormalbasis von V, bezüglich der L die Darstellung $[L]_{B'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ hat.

Wenn L nicht die Eigenwerte ± 1 hat, muß $[L]_B$ von der Form (1.6), für $\alpha \in [0, 2\pi)$, sein. Angenommen, $[L]_B$ hat die komplexen Eigenwerte λ_1, λ_2 mit $\{\lambda_1, \lambda_2\} \neq \{-1, 1\}$. Sei jetzt B' eine weitere Orthonormalbasis von $\mathbb{R}^2$. Dann sind $[L]_{B'}$ und $[L]_B$ ähnlich zueinander, also sind sie auch als komplexe Matrizen ähnlich, also hat auch $[L]_{B'}$ die Eigenwerte λ_1, λ_2. Sei $\alpha' \in [0, 2\pi)$, sodass

$$[L]_{B'} = \begin{pmatrix} \cos\alpha' & -\sin\alpha' \\ \sin\alpha' & \cos\alpha' \end{pmatrix}.$$

Dann gilt entweder

$$\begin{aligned} \cos\alpha' + i\sin\alpha' &= \lambda_1 = \cos\alpha + i\sin\alpha \\ \cos\alpha' - i\sin\alpha' &= \lambda_2 = \cos\alpha - i\sin\alpha, \end{aligned}$$

oder

$$\begin{aligned} \cos\alpha' + i\sin\alpha' &= \lambda_2 = \cos\alpha - i\sin\alpha \\ \cos\alpha' - i\sin\alpha' &= \lambda_1 = \cos\alpha + i\sin\alpha. \end{aligned}$$

Im ersten Fall folgt $\alpha' = \alpha$, im zweiten Fall folgt $\cos\alpha' = \cos\alpha$ und $\sin\alpha' = -\sin\alpha$, also $\alpha' = 2\pi - \alpha$. □

Satz 1.3.16. *Sei V ein euklidischer Raum mit $\dim_{\mathbb{R}} V = 3$ und $L : V \to V$ ein unitärer Endomorphismus. Dann gibt es eine Orthonormalbasis B von V und $\alpha \in [0, 2\pi)$, sodass*

$$[L]_B = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Beweis. Das charakteristische Polynom χ_L ist vom Grad 3, also hat es nach dem Zwischenwertsatz eine reelle Nullstelle λ_1. Diese ist ein Eigenwert von L, also $\lambda_1 = \pm 1$. Sei v_1 ein Eigenvektor zu λ_1 mit $\|v_1\| = 1$. Sei $W := \mathrm{Spann}_{\mathbb{R}}(v_1)$, dann ist $V = W \oplus W^\perp$, also $\dim W^\perp = 2$. Es gilt $L(W^\perp) \subset W^\perp$, da, für $w \in W^\perp$,

$$\lambda_1 \langle v_1, L(w)\rangle = \langle L(v_1), L(w)\rangle = \langle v_1, w\rangle = 0,$$

also $\langle v_1, L(w)\rangle = 0$.

Die Einschränkung $L|_{W^\perp}$ von L auf $W^\perp$ ist immer noch unitär, also gibt nach Satz 1.3.14 eine Orthonormalbasis $B' = \{v_2, v_3\}$ von $W^\perp$, sodass $[L]_{B'}$ die Form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

hat. Im ersten Fall setzen wir $B = \{v_1, v_2, v_3\}$. Dann ist B eine Orthonormalbasis von $V = W \oplus W^\perp$, und es gilt

$$[L]_B = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Im zweiten Fall setzen wir, falls $\lambda = 1$, $B = \{v_3, v_1, v_2\}$, und erhalten

$$[L]_B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos 0 & -\sin 0 \\ 0 & \sin 0 & \cos 0 \end{pmatrix}.$$

Im zweiten Fall, falls $\lambda = -1$, setzen wir $B = \{v_2, v_1, v_3\}$ und erhalten

$$[L]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\pi & -\sin\pi \\ 0 & \sin\pi & \cos\pi \end{pmatrix}.$$

□

Bemerkung 1.3.17. *Für $V = \mathbb{R}^3$ haben die Matrizen aus Satz 1.3.16 folgende geometrische Interpretation. Sei $B = \{v_1, v_2, v_3\}$ eine Orthonormalbasis von $\mathbb{R}^3$. Falls*

$$[L]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix},$$

ist L eine Drehung um den Winkel α um die Achse $\mathrm{Spann}_{\mathbb{R}}(v_1)$. Falls

$$[L]_B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix},$$

ist L eine Drehspiegelung, d.h. eine Komposition aus einer Drehung um den Winkel α um die Achse $\mathrm{Spann}_{\mathbb{R}}(v_1)$ und einer Spiegelung an der Ebene $\mathrm{Spann}_{\mathbb{R}}(v_2, v_3)$.

Insbesondere ist L genau dann eine Drehung, wenn $\det(L) = 1$, also $L \in \mathrm{SU}(\mathbb{R}^3)$.

Für Matrizen: eine Matrix $A \in M(3,3;\mathbb{R})$ stellt bezüglich einer Orthonormalbasis genau dann eine Drehung des $\mathbb{R}^3$ dar, wenn $A \in \mathrm{SO}(3)$. Analog stellt $A \in M(2,2;\mathbb{R})$ genau dann eine Drehung des $\mathbb{R}^2$ dar, wenn $A \in \mathrm{SO}(2)$. Allgemein nennt man daher $\mathrm{SO}(n)$ auch die Drehgruppe*, und ihre Elemente* Drehungen*, oder* Drehmatrizen*.*

Korollar 1.3.18 (Satz vom Fußball)**.** *Bei jedem Fußballspiel gibt es zwei Punkte auf der Oberfläche des Balls, die sich zu Beginn der ersten und der zweiten Halbzeit, wenn der Ball genau auf dem Anstoßpunkt liegt, an derselben Stelle im umgebenden Raum befinden.*

Beweis. Jede Drehung des Fußballs hat 1 als Eigenwert. □

1.4 Anwendung: Hauptachsentransformation

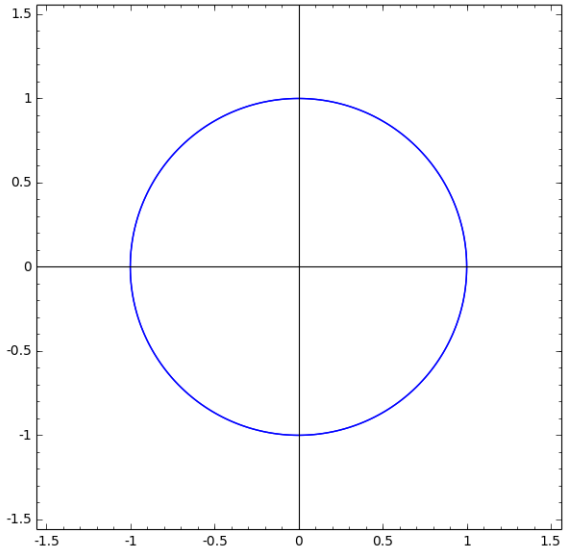
Definition 1.4.1. *Eine (reelle) quadratische Form ist eine Funktion*

$$Q : \mathbb{R}^n \to \mathbb{R}$$
$$Q(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j,$$

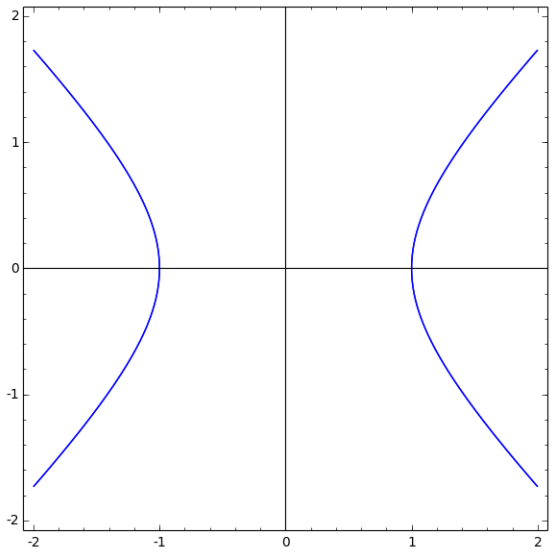
wobei $c_{ij} \in \mathbb{R}$ für $1 \le i \le j \le n$.

Beispiel 1.4.2. *Die Mengen $\{x \in \mathbb{R}^2 \mid Q(x,y) = 1\}$ beschreiben oft Kurven im $\mathbb{R}^2$.*

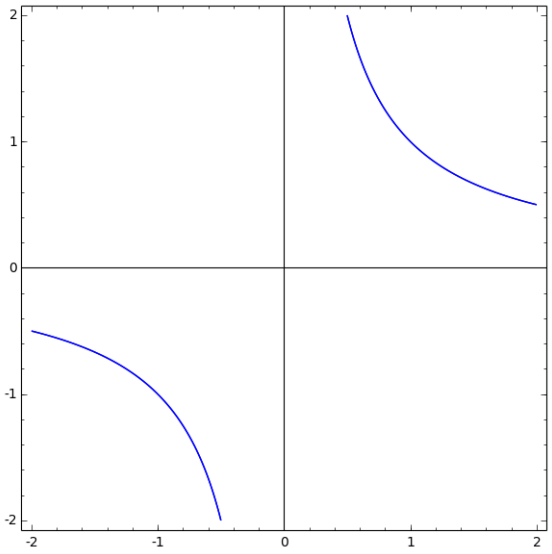
1. $x^2 + y^2 = 1$ *(Kreis)*



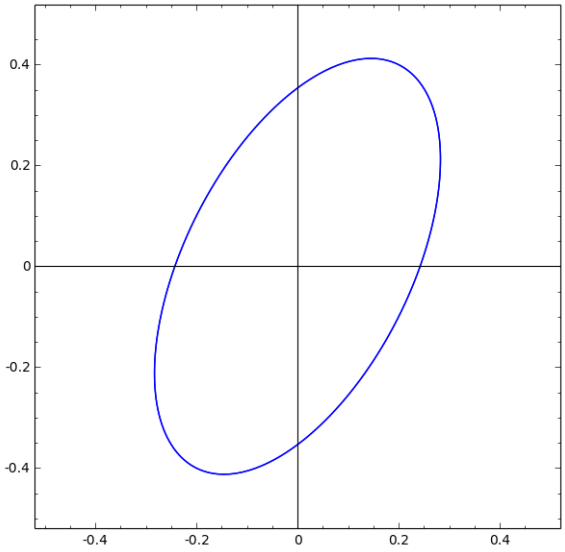
2. $x^2 - y^2 = 1$ *(Hyperbel)*



3. $xy = 1$ *(gedrehte Hyperbel)*



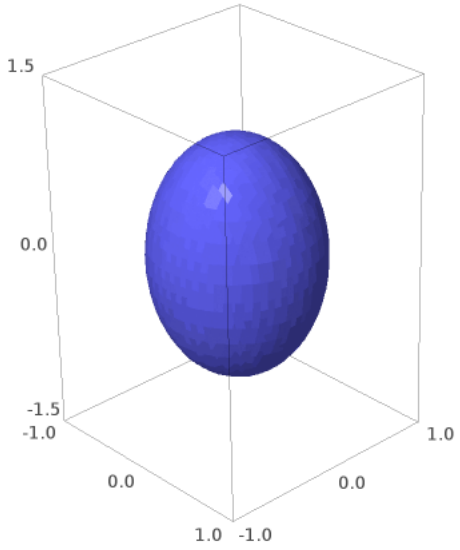
4. $17x^2 - 12xy + 8y^2 = 1$ *(gedrehte Ellipse)*



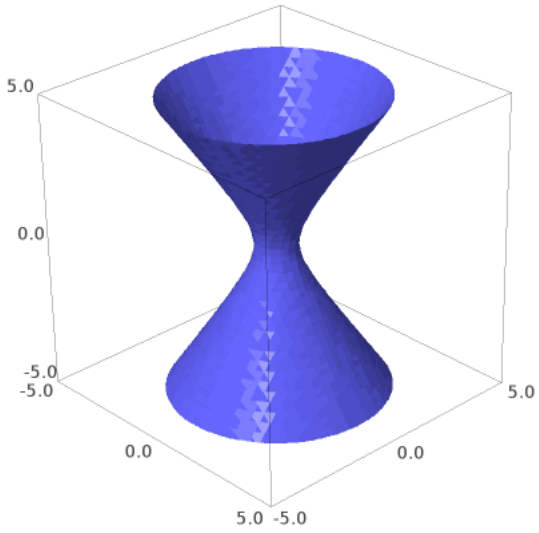
Wie kann man die Form und die Drehung aus der Gleichung ablesen?

Beispiel 1.4.3. *Die Mengen* $\{(x, y, z) \in \mathbb{R}^3 \mid Q(x, y, z) = 1\}$ *beschreiben oft Flächen im* $\mathbb{R}^3$.

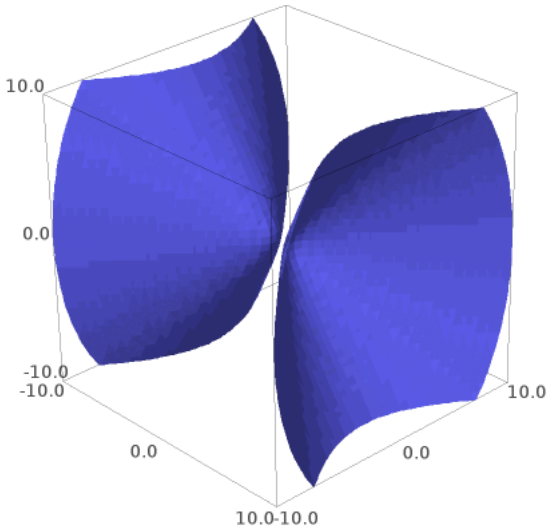
1. $2x^2 + 2y^2 + z^2 = 1$ *(Ellipsoid)*



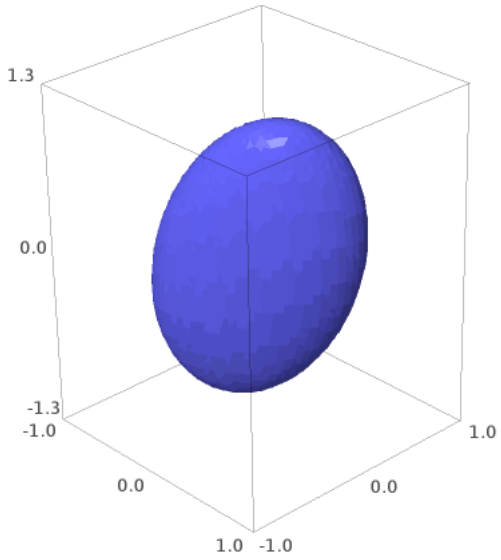
2. $2x^2 + 2y^2 - z^2 = 1$ *(Einschaliges Hyperboloid)*



3. $2x^2 - 2y^2 - z^2 = 1$ *(Zweischaliges Hyperboloid)*



4. $4x^2 - 6xy + 6y^2 - yz + z^2 = 1$ *(gedrehtes Ellipsoid)*



Gleiche Frage: Wie berechnet man Form und Drehung aus der Gleichung?

Bemerkung 1.4.4. *Jede quadratische Form lässt sich durch eine symmetrische Matrix darstellen: sei $x = (x_1, \dots, x_n)^t$. Es gilt*

$$Q(x) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j = (x_1, \dots, x_n) A_Q \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x^t A_Q x,$$

für

$$A_Q := \begin{pmatrix} c_{11} & \frac{c_{12}}{2} & \cdots & \frac{c_{1n}}{2} \\ \frac{c_{12}}{2} & c_{22} & & \\ \vdots & & \ddots & \\ \frac{c_{1n}}{2} & & & c_{nn} \end{pmatrix}, \qquad \text{d.h. } a_{ij} := \begin{cases} \frac{c_{ij}}{2} & \text{wenn } i < j \\ c_{ij} & \text{wenn } i = j \\ \frac{c_{ji}}{2} & \text{wenn } i > j. \end{cases}$$

Die Matrix $A_Q \in M(n, n; \mathbb{R})$ ist symmetrisch.

Satz 1.4.5 (Hauptachsentransformation)**.** *Sei $Q : \mathbb{R}^n \to \mathbb{R}$ eine quadratische Form mit Matrix A_Q. Dann gibt es $L \in \mathrm{SU}(\mathbb{R}^n)$, sodass $Q \circ L : \mathbb{R}^n \to \mathbb{R}$ die Form*

$$(x_1, \dots, x_n) \mapsto \sum_{i=1}^{n} \lambda_i x_i^2$$

hat, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ die Eigenwerte von A_Q (mit geometrischer Vielfachheit) sind.

Beweis. Laut Spektralsatz hat A_Q Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (mit geometrischer Vielfachheit), und es gibt eine Matrix $U \in \mathrm{SO}(n)$, sodass $U^t A U = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Setze $L := L_U : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto Ux$. Dann ist $L \in \mathrm{SU}(\mathbb{R}^n)$ und es gilt

$$\begin{aligned}(Q \circ L)(x) = Q(L(x)) = Q(Ux) = (Ux)^t A_Q (Ux) = x^t (U^t A_Q U) x \\ = x^t \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} x = \sum_{i=1}^{n} \lambda_i x_i^2.\end{aligned}$$

□

Bemerkung 1.4.6. *Für $n = 2$ heißt das: die Drehung L des $\mathbb{R}^2$ transformiert*

$$\{(x, y) \in \mathbb{R}^2 \mid \lambda_1 x^2 + \lambda_2 y^2 = 1\} \quad \text{zu} \quad C_Q := \{(x, y) \in \mathbb{R}^2 \mid Q(x, y) = 1\}.$$

In anderen Worten: L dreht die Koordinatenachsen von $\mathbb{R}^2$ auf die Hauptachsen von C_Q.

Daher hat C_Q eine der folgenden Gestalten:

1. *$\lambda_1, \lambda_2 \leq 0$: leere Menge*
2. *$\lambda_1, \lambda_2 > 0$: Ellipse mit Hauptachsen $1/\sqrt{\lambda_1}, 1/\sqrt{\lambda_2}$*
3. *$\lambda_1 \lambda_2 < 0$: Hyperbel mit Hauptachsen $1/\sqrt{\lambda_1}, 1/\sqrt{\lambda_2}$*
4. *$\lambda_1 > 0$ und $\lambda_2 = 0$ oder $\lambda_1 = 0$ und $\lambda_2 > 0$: Zwei parallele Geraden*

Ähnlich lassen sich auch im Fall $n = 3$ die Mengen $\{(x, y, z) \mid Q(x, y, z) = 1\}$ in Normalformen drehen.

Beispiel 1.4.7. *Wir rechnen nach, dass die Kurve*

$$C = \{(x, y) \mid 17x^2 - 12xy + 8y^2 = 1\}$$

tatsächlich eine gedrehte Ellipse ist, wie das Bild in Beispiel 1.4.2 vermuten lässt. Für $Q(x, y) = 17x^2 - 12xy + 8y^2$ gilt

$$A_Q = \begin{pmatrix} 17 & -6 \\ -6 & 8 \end{pmatrix}.$$

Das charakteristische Polynom von A_Q ist

$$\chi_{A_Q} = (17 - X)(8 - X) - 36 = X^2 - 25X + 100 = (X - 5)(X - 20).$$

Die Eigenwerte sind also $\lambda_1 = 5$, $\lambda_2 = 20$, daher handelt es sich um eine Ellipse. Die Drehung, die $5x^2 + 20y^2 = 1$ zu C transformiert, ist jenes $U \in \mathrm{SO}(2)$ mit $U^t A_Q U = \begin{pmatrix} 5 & 0 \\ 0 & 20 \end{pmatrix}$. Die Spalten u_1, u_2 von U sind normierte Eigenvektoren von A_Q zu λ_1, λ_2.

$$0 = (A_Q - 5I_2)u_1 = \begin{pmatrix} 12 & -6 \\ -6 & 3 \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix} \Rightarrow u_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

$$0 = (A_Q - 20I_2)u_2 = \begin{pmatrix} -3 & -6 \\ -6 & -12 \end{pmatrix} \begin{pmatrix} u_{12} \\ u_{22} \end{pmatrix} \Rightarrow u_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Daher

$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

Der Drehwinkel ist also $\phi = \arctan(2) \approx 1.11 \approx 63.43°$.

1.5 Anwendung: QR-Zerlegung

Sei $A \in M(m,n;\mathbb{K})$, $b \in \mathbb{K}^m$. Wenn $m > n$, ist das lineare Gleichungssystem $Ax = b$ nicht für jedes $b \in \mathbb{K}^m$ lösbar. Mann will oft eine Näherungslösung finden, für die $Ax - b$ möglichst klein ist.

Definition 1.5.1. *Wir betrachten $\mathbb{K}^m$ mit dem Standardskalarprodukt. Sei $A \in M(m,n;\mathbb{K})$, $b \in \mathbb{K}^m$. Der Vektor $x_0 \in \mathbb{K}^n$ heißt* Lösung zu $Ax = b$ im Sinne der kleinsten Fehlerquadrate*, wenn*

$$\|Ax_0 - b\| = \min_{x \in \mathbb{K}^n} \|Ax - b\|.$$

Beispiel 1.5.2 (Regression)**.** *Gegeben seien (viele) Punkte $(x_0, y_0), \ldots, (x_m, y_m) \in \mathbb{R}^2$. Wir wollen eine Funktion der Form, z.B.,*

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 e^x + c_5 \sin(x) + c_6 \cos(x)$$

finden, deren Graph diese Punkte möglichst gut approximiert. Das heißt, wir wollen $\sum_{i=1}^m (f(x_i) - y_i)^2$ minimieren. Die Koeffizienten $c_0, \ldots, c_6 \in \mathbb{R}$ lassen sich als Lösung im Sinne der kleinsten Fehlerquadrate des folgenden linearen Gleichungssystems bestimmen:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & e^{x_1} & \sin(x_1) & \cos(x_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & x_m^3 & e^{x_m} & \sin(x_m) & \cos(x_m) \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Wie bestimmt man so eine Lösung? Zum Beispiel über die QR-Zerlegung der Matrix A.

Definition 1.5.3. *Eine* rechte obere Dreiecksmatrix *ist eine Matrix $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in M(m,n;K)$ mit $a_{ij} = 0$ für alle Indizes $i > j$.*

Satz 1.5.4 (QR-Zerlegung)**.** *Sei $A \in M(m,n;\mathbb{K})$ mit $\operatorname{Rang} A = n$. Dann gibt es eine orthogonale bzw. unitäre Matrix $Q \in M(m,m;\mathbb{K})$ und eine rechte obere Dreiecksmatrix $R \in M(m,n;\mathbb{K})$ mit $A = QR$.*

Beweis. Seien $a_1, \ldots, a_n \in \mathbb{K}^m$ die Spalten von A. Diese bilden eine Basis des Spaltenraums $\operatorname{im} A$ von A. Der Orthonormalisierungssatz von Gram-Schmidt liefert uns $w_1, \ldots, w_n \in \mathbb{K}^m$, sodass, für $1 \le j \le n$,

$$\{w_1, \ldots, w_j\} \text{ eine Orthonormalbasis von } \operatorname{Spann}_{\mathbb{K}}(a_1, \ldots, a_j) \text{ ist.}$$

Insbesondere gibt es $r_{ij} \in \mathbb{K}$, sodass

$$a_j = \sum_{i=1}^{j} r_{ij} w_i.$$

Wir ergänzen $\{w_1, \ldots, w_n\}$ zu einer Orthonormalbasis $\{w_1, \ldots, w_n, w_{n+1}, \ldots, w_m\}$ von $\mathbb{K}^m$. Dann ist die Matrix $Q \in M(m, m; \mathbb{K})$ mit Spalten $w_1, \ldots, w_m$ orthogonal bzw. unitär.

Sei $R \in M(m, n; \mathbb{K})$ mit Einträgen r_{ij} für $i \leq j$ und 0 sonst. Dann ist R eine rechte obere Dreiecksmatrix, und $A = QR$. Zur Veranschaulichung:

$$(a_1, \ldots, a_n) = (w_1, \ldots, w_n, w_{n+1}, \ldots, w_m) \cdot \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & & & r_{nn} \\ 0 & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

□

Bemerkung 1.5.5. *Der Beweis liefert eine Methode zum Bestimmen einer QR-Zerlegung von A. In der Praxis werden auch andere Methoden verwendet.*

Satz 1.5.6. *Sei $A \in M(m, n; \mathbb{K})$ mit $\operatorname{Rang} A = n$ und $b \in \mathbb{K}^m$. Sei $Q \in M(m, m; \mathbb{K})$ unitär und $R \in M(m, n; \mathbb{K})$ eine rechte obere Dreiecksmatrix, sodass $A = QR$. Schreibe*

$$Q^* b = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{K}^m, \quad R = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} \in M(m, n; \mathbb{K}),$$

mit $c \in \mathbb{K}^n$, $d \in \mathbb{K}^{m-n}$, $\tilde{R} \in M(n, n; \mathbb{K})$.

Dann hat das lineare Gleichungssystem $\tilde{R}x = c$ eine Lösung $x_0 \in \mathbb{K}^n$, und x_0 ist eine Lösung im Sinne der kleinsten Fehlerquadrate von $Ax = b$.

Beweis. Da Q invertierbar ist, gilt

$$\operatorname{Rang} \tilde{R} = \operatorname{Rang} R = \operatorname{Rang} QR = \operatorname{Rang} A = n,$$

also gibt es $x_0 \in \mathbb{K}^n$ mit $\tilde{R} x_0 = c$.

Sei $x \in \mathbb{K}^n$. Dann gilt

$$\begin{aligned}\|Ax-b\|^2 &= \|QRx - QQ^*b\|^2 = \|Q(Rx-Q^*b)\|^2 = \|Rx-Q^*b\|^2 \\ &= \|\begin{pmatrix}\tilde{R}\\0\end{pmatrix}x - \begin{pmatrix}c\\d\end{pmatrix}\|^2 = \|\tilde{R}x-c\|^2 + \|-d\|^2 \geq \|d\|^2.\end{aligned}$$

Für $x = x_0$ gilt $\|Ax_0 - b\|^2 = \|d\|^2$. □

Bemerkung 1.5.7.

1. *Das Gleichungssystem $\tilde{R}x_0 = c$ lässt sich schnell durch Rückeinsetzen lösen, da R bereits in oberer Dreiecksform ist.*

2. *Die Methode funktioniert auch wenn $m = n$ und liefert dann eine exakte Lösung. Die Lösung von $Ax = b$ wird als Lösung von $Rx = Q^*b$ bestimmt.*

Beispiel 1.5.8. *Bestimme eine Lösung im Sinne der kleinsten Fehlerquadrate zu $Ax = b$, mit*

$$A = \begin{pmatrix}1 & 2\\0 & 1\\1 & 0\end{pmatrix}, \quad b = \begin{pmatrix}1\\0\\0\end{pmatrix}.$$

Zuerst bestimmen wir eine QR-Zerlegung von A, dazu orthonormalisieren wir die Spalten mit Gram-Schmidt. Es ist

$$w_1 = \frac{1}{\|a_1\|}a_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix},$$

das gibt uns die erste Spalte von Q. Da $a_1 = \sqrt{2}w_1$, ist die erste Spalte von R gleich $(\sqrt{2}, 0, 0)^t$. Weiters

$$\tilde{w_2} = \begin{pmatrix}2\\1\\0\end{pmatrix} - \langle \begin{pmatrix}2\\1\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}\rangle \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}2\\1\\0\end{pmatrix} - \sqrt{2}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

Die zweite Spalte von Q ist also

$$w_2 = \frac{1}{\|\tilde{w_2}\|}\tilde{w_2} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

Da

$$a_2 = \tilde{w_2} + \sqrt{2}w_1 = \sqrt{3}w_2 + \sqrt{2}w_1,$$

ist die zweite Spalte von R gleich $(\sqrt{2}, \sqrt{3}, 0)^t$*, also*

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \\ 0 & 0 \end{pmatrix}.$$

Zur Bestimmung der dritten Spalte von Q ergänzen wir $\{w_1, w_2\}$ *zu einer Orthonormalbasis von* $\mathbb{K}^3$*. Sei* $a_3 := (0, 1, 0)$*, dann ist* $\{w_1, w_2, a_3\}$ *eine Basis von* $\mathbb{K}^3$*. Dann*

$$\begin{aligned} \tilde{w_3} &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \rangle \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \end{aligned}$$

also

$$w_3 = \frac{1}{\|\tilde{w_3}\|} w_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Daher ist

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix},$$

und wir haben eine QR-Zerlegung von A bestimmt. Nun gilt

$$Q^* b = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} \end{pmatrix}, \qquad \textit{also } c = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}}. \end{pmatrix}$$

Wir lösen das System $\tilde{R}x = c$*, also*

$$\begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

und erhalten

$$x_2 = \frac{1}{3}, \quad x_1 = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} - \frac{\sqrt{2}}{3} \right) = \frac{1}{6}.$$

Die Lösung im Sinne der kleinsten Fehlerquadrate zu $Ax = b$ ist also $x_0 = (1/6, 1/3)^t$. Es gilt

$$Ax_0 - b = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{6} \\ \frac{1}{3} \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix},$$

also $\|Ax_0 - b\|^2 = \frac{1}{6} = \|d\|^2$, wie erwartet.

1.6 Normale Endomorphismen

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum. Laut Spektralsatz gibt es für jeden selbstadjungierten Endomorphismus $L : V \to V$ eine Orthonormalbasis von V aus Eigenvektoren von L. Gilt auch die Umkehrung?

Korollar 1.6.1. *Sei V ein endlich-dimensionaler euklidischer Raum und $L : V \to V$ ein Endomorphismus. Dann ist L genau dann selbstadjungiert, wenn V eine Orthonormalbasis aus Eigenvektoren von L hat.*

Beweis. $\Rightarrow$: Spektralsatz. $\Leftarrow$: Sei $B = \{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V, sodass v_i ein Eigenvektor zum Eigenwert $\lambda_i \in \mathbb{R}$ von L ist. Dann gilt $[L]_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Es folgt $[L^*]_B = [L]_B^t = [L]_B$, und daher $L^* = L$. $\square$

Für unitäre Räume gilt die Umkehrung nicht unbedingt, denn für $\lambda_i \in \mathbb{C}$ ist $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ nicht selbstadjungiert, wenn nicht $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Es gilt aber zumindest

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} \overline{\lambda}_1 & & \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda}_1 & & \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Daher folgt: wenn V eine Orthonormalbasis aus Eigenvektoren von L hat, gilt $L \circ L^* = L^* \circ L$.

Definition 1.6.2.

1. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum. Ein Endomorphismus $L : V \to V$ heißt* normal*, falls*

$$L^* \circ L = L \circ L^*.$$

2. *Eine Matrix $A \in M(n, n; \mathbb{K})$ heißt* normal*, falls*

$$A^* A = A A^*.$$

Bemerkung 1.6.3.

1. *Selbstadjungierte und unitäre Endomorphismen sind normal.*
2. *Sei B eine Orthonormalbasis von V. Dann ist $L : V \to V$ genau dann normal, wenn $[L]_B$ normal ist.*

Lemma 1.6.4. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ ein normaler Endomorphismus.*

1. *Für alle $v \in V$ ist $\|L(v)\| = \|L^*(v)\|$.*
2. $\ker L = \ker L^*$
3. *Sei $v \in V$ ein Eigenvektor zum Eigenwert λ von L. Dann ist v ein Eigenvektor zum Eigenwert $\overline{\lambda}$ von L^*.*
4. *Eigenvektoren zu verschiedenen Eigenwerten von L sind orthogonal.*

Beweis. Zu *1.*:

$$\langle L(v), L(v)\rangle = \langle v, (L^* \circ L)(v)\rangle = \langle v, (L \circ L^*)(v)\rangle = \langle L^*(v), L^*(v)\rangle.$$

Zu *2.*:

$$L(v) = 0 \Leftrightarrow \|L(v)\| = 0 \Leftrightarrow \|L^*(v)\| = 0 \Leftrightarrow L^*(v) = 0.$$

Zu *3.*:

$$\begin{aligned}\|L^*(v) - \overline{\lambda}v\|^2 &= \langle L^*(v) - \overline{\lambda}v, L^*(v) - \overline{\lambda}v\rangle \\ &= \langle L^*(v), L^*(v)\rangle - \lambda\langle L^*(v), v\rangle - \overline{\lambda}\langle v, L^*(v)\rangle + \overline{\lambda}\lambda\langle v, v\rangle \\ &= \langle L(v), L(v)\rangle - \lambda\langle v, L(v)\rangle - \overline{\lambda}\langle L(v), v\rangle + \overline{\lambda}\lambda\langle v, v\rangle \\ &= \langle L(v) - \lambda v, L(v) - \lambda v\rangle = \|L(v) - \lambda v\|^2 = 0.\end{aligned}$$

Zu *4.*: Seien $\lambda \neq \mu$ Eigenwerte von L mit Eigenvektoren v, w. Dann

$$\lambda\langle v, w\rangle = \langle L(v), w\rangle = \langle v, L^*(w)\rangle = \langle v, \overline{\mu}w\rangle = \mu\langle v, w\rangle.$$

Da $\lambda \neq \mu$, folgt $\langle v, w\rangle = 0$. □

Satz 1.6.5 (Spektralsatz)**.** *Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $L : V \to V$ ein Endomorphismus. Dann sind folgende Aussagen äquivalent:*

1. *L ist normal und das charakteristische Polynom χ_L zerfällt in Linearfaktoren.*

2. *V besitzt eine Orthonormalbasis aus Eigenvektoren von L.*

Beweis. 1.⇒2.: Induktion nach $n = \dim_{\mathbb{K}} V$. Der Fall $n = 0$ ist trivial. Sei $n \geq 1$ und gelte die Aussage für alle V, L mit $\dim_{\mathbb{K}} V \leq n-1$.

Dann hat χ_L eine Nullstelle $\lambda_1 \in \mathbb{K}$, diese ist ein Eigenwert von L. Sei v_1 ein zugehöriger Eigenvektor mit $\|v_1\| = 1$. Dann ist $\{v_1\}$ eine Orthonormalbasis von $W := \text{Spann}_{\mathbb{K}}(v_1)$. Da $V = W \oplus W^\perp$, gilt $\dim_{\mathbb{K}} W^\perp = n-1$. Weiters gilt $L(W^\perp) \subset W^\perp$, da für $w \in W^\perp$,

$$\langle L(w), v_1\rangle = \langle w, L^*(v_1)\rangle = \langle w, \overline{\lambda}_1 v_1\rangle = \lambda_1\langle w, v_1\rangle = 0.$$

Ähnlich gilt auch $L^*(W^\perp) \subset W^\perp$, da für $w \in W^\perp$,

$$\langle L^*(w), v_1\rangle = \langle w, L(v_1)\rangle = \overline{\lambda}_1\langle w, v_1\rangle = 0.$$

Also sind die Einschränkungen $L|_{W^\perp}$, $L^*|_{W^\perp}$ Endomorphismen von $W^\perp$, und $(L|_{W^\perp})^* = L^*|_{W^\perp}$. Insbesondere ist auch $L^*|_{W^\perp}$ normal. Für jede beliebige Basis $B' = \{w_2, \ldots, w_n\}$ von $W^\perp$ ist $B := \{v_1, w_2, \ldots, w_n\}$ eine Basis von V, und die darstellende Matrix von L hat die Form

$$[L]_B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & [L|_{W^\perp}]_{B'} \end{pmatrix}.$$

Daher folgt $\chi_L = (\lambda_1 - X)\cdot\chi_{L|_{W^\perp}}$. Mit χ_L zerfällt also auch $\chi_{L|_{W^\perp}}$ in Linearfaktoren. Daher erfüllt $L|_{W^\perp}$ alle Bedingungen in *1.*, und laut Induktionsvoraussetzung besitzt $W^\perp$ eine Orthonormalbasis $\{v_2, \ldots, v_n\}$ aus Eigenvektoren von $L|_{W^\perp}$. Diese sind auch Eigenvektoren von L, und $\{v_1, \ldots, v_n\}$ ist eine Orthonormalbasis von V.

2.⇒1.: L ist diagonalisierbar, also zerfällt χ_L in Linearfaktoren. Sei B eine Orthonormalbasis aus Eigenvektoren von L zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Dann gilt, wie schon zu Beginn des Kapitls bemerkt,

$$[L \circ L^*]_B = [L]_B[L]_B^* = \begin{pmatrix} \lambda_1\overline{\lambda}_1 & & \\ & \ddots & \\ & & \lambda_n\overline{\lambda}_n \end{pmatrix} = [L]_B^*[L]_B = [L^* \circ L]_B,$$

also ist L normal. □

Für Matrizen folgt sofort folgende Version.

Korollar 1.6.6 (Spektralsatz für Matrizen)**.** *Sei $A \in M(n,n;\mathbb{K})$ dann sind äquivalent:*

1. *A ist normal und χ_A zerfällt in Linearfaktoren.*
2. *Es gibt $U \in \mathrm{O}(n)$ bzw. $U \in \mathrm{U}(n)$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass*

$$U^*AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Beweis. 2. ist äquivalent dazu, dass V eine Orthonormalbasis aus Eigenvektoren von A hat: für U wie in *2.* bilden die Spalten von U so eine Orthonormalbasis, und wenn $\{u_1, \ldots, u_n\}$ eine Orthonormalbasis aus Eigenvektoren von A ist, dann ist die Matrix U mit Spalten $u_1, \ldots, u_n$ orthogonal bzw. unitär und, mit den zugehörigen Eigenwerten $\lambda_1, \ldots, \lambda_n$, gilt $U^*AU = U^{-1}AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. (Details wie im Beweis von Korollar 1.3.11) □

Kapitel 2

Bilinearformen und quadratische Formen

Wie betrachten wieder Vektorräume über einem beliebigen Körper K.

Literatur: das Kapitel basiert hauptsächlich auf [2].

2.1 Bilinearformen

Definition 2.1.1. *Sei V ein K-Vektorraum. Eine* Bilinearform *auf V ist eine Funktion $\beta : V \times V \to K$, die linear in beiden Argumenten ist. Das heißt, für alle $v_1, v_2, w \in V$ und $\alpha \in K$ gilt*

$$\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w), \qquad \beta(\alpha v_1, w) = \alpha\,\beta(v_1, w)$$
$$\beta(w, v_1 + v_2) = \beta(w, v_1) + \beta(w, v_2), \qquad \beta(w, \alpha v_1) = \alpha\,\beta(w, v_1).$$

Die Bilinearform β heißt

- symmetrisch*, wenn $\beta(v, w) = \beta(w, v)$ für alle $v, w \in V$*
- schiefsymmetrisch *(bzw.* antisymmetrisch*), wenn $\beta(v, w) = -\beta(w, v)$ für alle $v, w \in V$*
- alternierend*, wenn $\beta(v, v) = 0$ für alle $v \in V$.*

Bemerkung 2.1.2.

1. *Sei V ein euklidischer Raum, dann ist das innere Produkt $\langle \cdot, \cdot \rangle$ eine Bilinearform. Unsere Hauptmotivation zur Betrachtung allgemeiner Bilinearformen ist eine Verallgemeinerung von inneren Produkten in zwei Richtungen:*

   a) *auf andere Körper, wichtig z.B. in Zahlentheorie, Algebra, algebraischer Geometrie*
   b) *über $\mathbb{R}$, aber mit schwächeren Voraussetzungen (z.B. keine positive Definitheit), wichtig z.B. in der Physik (Relativitätstheorie)*

2. *Aus der Linearen Algebra I ist bereits bekannt:*
   a) *β alternierend $\Rightarrow$ β schiefsymmetrisch,*
   b) *falls* $\operatorname{char} K \neq 2$*: β schiefsymmetrisch $\Leftrightarrow$ β alternierend,*
   c) *falls* $\operatorname{char} K = 2$*: β schiefsymmetrisch $\Leftrightarrow$ β symmetrisch.*

3. *Sei V ein endlich-dimensionaler K-Vektorraum mit Basis $\{v_1, \ldots, v_n\}$. Eine Bilinearform β auf V ist durch die Werte $\beta(v_i, v_j)$, $1 \leq i, j \leq n$ eindeutig bestimmt. Für $v = \sum_{i=1}^n a_i v_i$ und $w = \sum_{i=1}^n b_i v_i$ gilt*

$$\beta(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \, \beta(v_i, v_j).$$

Beispiel 2.1.3.

1. *Sei V ein euklidischer Raum und $L : V \to V$ ein Endomorphismus. Dann ist $\beta(v, w) := \langle v, L(w) \rangle$ eine Bilinearform.*

2. *Sei $A \in M(m, n; K)$ und $V = K^n$. Für Spaltenvektoren v, w ist $\beta(v, w) := v^t A w$ eine Bilinearform.*

3. *In der speziellen Relativitätstheorie modelliert man die Raumzeit unter Anderem als $\mathbb{R}^4$ mit der symmetrischen Bilinearform*

$$\beta(x, y) = x_1 y_1 - x_2 y_2 - x_3 y_3 - x_4 y_4.$$

4. *Sei V endlich-dimensional. Auf dem Vektorraum $\operatorname{End}_K(V)$ der Endomorphismen von V ist eine Bilinearform durch*

$$\beta(L_1, L_2) := \operatorname{Spur}(L_1 \circ L_2)$$

   definiert. Diese ist symmetrisch, da $\operatorname{Spur}(L_1 \circ L_2) = \operatorname{Spur}(L_2 \circ L_1)$.

5. *Sei $V = \mathcal{C}(\mathbb{R}, \mathbb{R})$, der Raum der stetigen Funktionen auf $\mathbb{R}$. Dann ist*

$$\beta(f, g) := \int_0^1 f(t) g(t) \mathrm{d}t$$

   eine symmetrische Bilinearform. Diese ist kein inneres Produkt, da nicht positiv definit.

6. *Seien $L_1, L_2 : V \to W$ lineare Abbildungen und β_W eine Bilinearform auf W. Dann ist*
$$\beta(v, w) := \beta_W(L_1(v), L_2(w))$$
eine Bilinearform auf V.

7. *Spezialfall $W = K$: Seien $L_1, L_2 \in V^*$. Dann ist $\beta(v, w) := L_1(v)L_2(w)$ eine Bilinearform.*

8. *Spezialfall $V \subset W$: Sei β eine Bilinearform auf W, dann ist $\beta|_V := \beta\,|_{V\times V}$ eine Bilinearform auf V.*

Lemma 2.1.4. *Sei* $\operatorname{char} K \neq 2$.

1. *Jede Bilinearform β kann eindeutig als $\beta = \beta_1 + \beta_2$ geschrieben werden, mit β_1 symmetrisch und β_2 schiefsymmetrisch.*

2. *Eine symmetrische Bilinearform β ist vollständig durch die Werte $\beta(v, v)$, $v \in V$, bestimmt.*

Beweis. Zu *1.* Angenommen,
$$\beta(v, w) = \beta_1(v, w) + \beta_2(v, w),$$
mit β_1 symmetrisch, β_2 schiefsymmetrisch. Dann
$$\beta(w, v) = \beta_1(v, w) - \beta_2(v, w),$$
also
$$\beta_1(v, w) = \frac{1}{2}(\beta(v, w) + \beta(w, v)), \quad \beta_2(v, w) = \frac{1}{2}(\beta(v, w) - \beta(w, v)). \quad (2.1)$$
Daher sind β_1, β_2 eindeutig durch β bestimmt. Zur Existenz: definiere β_1, β_2 durch (2.1)

Zu *2.* Wie für euklidische innere Produkte gilt für symmetrisches β, dass
$$\beta(v, w) = \frac{1}{2}(\beta(v + w, v + w) - \beta(v, v) - \beta(w, w)).$$

□

2.2 Bilinearformen und Matrizen

Sei V ein endlich-dimensionaler Vektorraum. Lineare Abbildungen können nach Wahl einer Basis von V durch Matrizen dargestellt werden. Auch Bilinearformen lassen sich durch Matrizen beschreiben.

Wir haben bereits gesehen, dass jede Matrix $A = (a_{ij})_{1\leq i,j\leq n} \in M(n,n;K)$ eine Bilinearform β auf K^n wie folgt definiert: für Spaltenvektoren $v = (v_1,\ldots,v_n)^t$, $w = (w_1,\ldots,w_n)^t \in K^n$, setze

$$\beta(v,w) = v^t A w = \sum_{i=1}^{n} v_i \sum_{j=1}^{n} a_{ij} w_j = \sum_{1\leq i,j\leq n} a_{ij} v_i w_j.$$

Nach Wahl einer Basis hat jede Bilinearform auf einem endlich-dimensionalen K-Vektorraum diese Form.

Definition 2.2.1. *Sei V ein endlich-dimensionaler K-Vektorraum mit Basis $B = \{v_1,\ldots,v_n\}$.*

1. *Sei β eine Bilinearform auf V. Die* Strukturmatrix *von β bezüglich der Basis B ist die Matrix*

$$[\beta]_B := (\beta(v_i,v_j))_{1\leq i,j\leq n} \in M(n,n;K).$$

2. *Sei $A = (a_{ij})_{1\leq i,j\leq n} \in M(n,n;k)$. Die* durch A bezüglich der Basis B dargestellte *Bilinearform auf V ist definiert durch*

$$\beta_A^B(v_i,v_j) := a_{ij} \quad \text{für } 1 \leq i,j \leq n.$$

Bemerkung 2.2.2. *Für $v = x_1v_1 + \cdots + x_nv_n$, $w = y_1v_1 + \cdots + y_nv_n \in V$ gilt also*

$$\beta_A^B(v,w) = \sum_{1\leq i,j\leq n} x_i y_j\, \beta(v_i,v_j) = \sum_{1\leq i,j\leq n} x_i y_j a_{ij} = (x_1,\ldots,x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Beispiel 2.2.3. *Die alternierende und schiefsymmetrische Bilinearform $\beta((x_1,x_2),(y_1,y_2)) = x_1y_2 - x_2y_1$ auf $\mathbb{R}^2$ hat bezüglich der Standardbasis E die Strukturmatrix*

$$[\beta]_E = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Satz 2.2.4. *Sei V ein K-Vektorraum und* $\mathrm{Bil}(V,K)$ *die Menge der Bilinearformen auf V.*

1. *Mit punktweiser Addition und Skalarmultiplikation ist* $\mathrm{Bil}(V,K)$ *ein* K*-Vektorraum.*

2. *Ist* V *endlich-dimensional mit Basis* $B = \{v_1, \ldots, v_n\}$*, dann sind die Abbildungen*

$$\begin{aligned} \Phi_B : \mathrm{Bil}(V,K) &\to M(n,n;K), & \Psi_B : M(n,n;K) &\to \mathrm{Bil}(V,K) \\ \beta &\mapsto [\beta]_B & A &\mapsto \beta_A^B \end{aligned}$$

zueinander inverse Vektorraumisomorphismen.

Beweis. Zu *1.* Für $\beta_1, \beta_2 \in \mathrm{Bil}(V,K)$ und $\alpha \in K$ sind auch

$$\begin{aligned} \beta_1 + \beta_2 &: (v,w) \mapsto \beta_1(v,w) + \beta_2(v,w) \\ \alpha\,\beta_1 &: (v,w) \mapsto \alpha\,\beta_1(v,w) \end{aligned}$$

wieder Bilinearformen auf V. Diese Addition und Skalarmultiplikation erfüllen offensichtlich die Vektorraumaxiome.

Zu *2.* Beide Abbildungen sind offensichtlich K-linear. Weiters gilt, für $\beta \in \mathrm{Bil}(V,K)$ und $1 \le i,j \le n$,

$$(\Psi_B \circ \Phi_B)(\beta)(v_i,v_j) = \beta^B_{[\beta]_B}(v_i,v_j) = (i,j)\text{-Eintrag von } [\beta]_B = \beta(v_i,v_j),$$

also $(\Psi_B \circ \Phi_B)(\beta) = \beta$, und daher $\Psi_B \circ \Phi_B = \mathrm{id}_{\mathrm{Bil}(V,K)}$. Ähnlich gilt, für $A \in M(n,n;K)$,

$$(\Phi_B \circ \Psi_B)(A) = [\beta^B_A]_B = ((\beta^B_A)(v_i,v_j))_{1\le i,j\le n} = (a_{i,j})_{1\le i,j\le n} = A,$$

also $\Phi_B \circ \Psi_B = \mathrm{id}_{M(n,n;K)}$. □

Definition 2.2.5. *Sei* $A = (a_{ij})_{1\le i,j\le n} \in M(n,n;K)$.

1. *A heißt* schiefsymmetrisch*, wenn* $A^t = -A$.

2. *A heißt* alternierend*, wenn* $A^t = -A$ *und* $a_{ii} = 0$ *für alle* $1 \le i \le n$.

Lemma 2.2.6. *Sei* V *ein endlich-dimensionale* K*-Vektorraum mit einer Bilinearform* β*. Sei* B *eine Basis von* V*. Dann gilt*

1. β *ist genau dann symmetrisch, wenn* $[\beta]_B$ *symmetrisch ist.*

2. β *ist genau dann schiefsymmetrisch, wenn* $[\beta]_B$ *schiefsymmetrisch ist.*

3. β *ist genau dann alternierend, wenn* $[\beta]_B$ *alternierend ist.*

Beweis. Sei $B = \{v_1, \ldots, v_n\}$ und $[\beta]_B = (a_{ij})_{1 \le i,j \le n}$.

Zu *2.* Wenn β schiefsymmetrisch ist, gilt $a_{ij} = \beta(v_i, v_j) = -\beta(v_j, v_i) = -a_{ji}$, also ist auch $[\beta]_B$ schiefsymmetrisch. Sei umgekehrt $[\beta]_B$ schiefsymmetrisch und $v = x_1 v_1 + \cdots + x_n v_n$, $w = y_1 v_1 + \cdots + y_n v_n \in V$. Dann gilt

$$\begin{aligned}\beta(v,w) &= \sum_{1 \le i,j \le n} x_i y_j \, \beta(v_j, v_j) = \sum_{1 \le i,j \le n} x_i y_j a_{ij} = -\sum_{1 \le i,j \le n} x_i y_j a_{ji} \\ &= -\sum_{1 \le i,j \le n} y_j x_i \, \beta(v_j, v_i) = -\beta(w,v).\end{aligned}$$

Der Beweis zu *1.* verläuft genau gleich, nur ohne die Minuszeichen.

Zu *3.* Wenn β alternierend ist, ist β, und damit $[\beta]_B$ auch schiefsymmetrisch. Weiters gilt $a_{ii} = \beta(v_i, v_i) = 0$ für $1 \le i \le n$. Sei umgekehrt A alternierend und $v = x_1 v_1 + \cdots + x_n v_n \in V$. Dann gilt

$$\begin{aligned}\beta(v,v) &= \sum_{1 \le i,j \le n} x_i x_j a_{ij} = \sum_{\substack{1 \le i,j \le n \\ i<j}} x_i x_j a_{ij} + \sum_{\substack{1 \le i,j \le n \\ i=j}} x_i x_j a_{ij} + \sum_{\substack{1 \le i,j \le n \\ i>j}} x_i x_j a_{ij} \\ &= \sum_{\substack{1 \le i,j \le n \\ i<j}} x_i x_j a_{ij} - \sum_{\substack{1 \le i,j \le n \\ i>j}} x_j x_i a_{ji} = 0.\end{aligned}$$

□

Nach Wahl einer Basis von V sind also swowhl $\mathrm{End}_K(V)$, der Raum der Endomorphismen von V, als auch $\mathrm{Bil}(V,K)$, der Raum der Bilinearformen auf V, isomorph zu $M(n,n;K)$. Was geschieht bei Basiswechsel. Seien B_1, B_2 Basen von V. Für Endomorphismen wissen wir bereits, dass

$$[L]_{B_2} = [\mathrm{id}_V]^{-1}_{B_2,B_1} \cdot [L]_{B_1} \cdot [\mathrm{id}_V]_{B_2,B_1}.$$

Satz 2.2.7. *Sei V ein endlich-dimensionaler K-Vektorraum mit einer Bilinearform β. Seien B_1, B_2 Basen von V. Dann gilt*

$$[\beta]_{B_2} = [\mathrm{id}_V]^{t}_{B_2,B_1} \cdot [\beta]_{B_1} \cdot [\mathrm{id}_V]_{B_2,B_1}.$$

In anderen Worten: sei C die Matrix, deren j-te Spalte die Koordinaten des j-ten Vektors in B_2 bezüglich der Basis B_1 enthält, dann ist

$$[\beta]_{B_2} = C^t [\beta]_{B_1} C.$$

Beweis. Sei $B_1 = \{v_1, \dots, v_n\}$ und $B_2 = \{w_i, \dots, w_n\}$. Wir schreiben

$$\begin{aligned} [\beta]_{B_1} &= (a_{ij})_{1\le i,j\le n}, & [\beta]_{B_2} &= (b_{ij})_{1\le i,j\le n},\\ [\mathrm{id}_V]_{B_2,B_1} &= (c_{ij})_{1\le i,j\le n}, & [\beta]_{B_1}[\mathrm{id}_K]_{B_2,B_1} &= (d_{ij})_{1\le i,j\le n}\\ & & [\mathrm{id}_V]^t_{B_2,B_1}[\beta]_{B_1}[\mathrm{id}_V]_{B_2,B_1} &= (f_{ij})_{1\le i,j\le n}. \end{aligned}$$

Dann gilt

$$\begin{aligned} b_{ij} = \beta(w_i, w_j) &= \beta(\sum_{k=1}^n c_{ki}v_k, \sum_{l=1}^n c_{lj}v_l) = \sum_{1\le k,l\le n} c_{ki}c_{lj}\,\beta(v_k, v_l)\\ &= \sum_{1\le k,l\le n} c_{ki}c_{lj}a_{kl} = \sum_{k=1}^n c_{ki}\sum_{l=1}^n a_{kl}c_{lj} = \sum_{k=1}^n c_{ki}d_{kj} = f_{ij}. \end{aligned}$$

Alternativ mit Koordinaten: Sei $C = [\mathrm{id}_V]_{B_2,B_1}$ die Matrix des Basiswechsels und

$$\begin{aligned} v &= x_1v_1 + \cdots + x_nv_n = x'_1w_1 + \cdots + x'_nw_n\\ w &= y_1v_1 + \cdots + y_nv_n = y'_1w_1 + \cdots + y'_nw_n. \end{aligned}$$

Dann gilt $x = Cx'$, $y = Cy'$, also

$$x'^t[\beta]_{B_2}y' = \beta(v, w) = x^t[\beta]_{B_1}y = (Cx')^t[\beta]_{B_1}(Cy') = x'^t(C^t[\beta]_{B_1}C)y',$$

also $[\beta]_{B_2} = C^t[\beta]_{B_1}C$. □

Definition 2.2.8. *Zwei Bilinearformen β_1, β_2 auf K-Vektorräumen V_1, V_2 heißen* äquivalent*, wenn es einen Isomorphismus $L : V_1 \to V_2$ gibt, sodass*

$$\beta_2(L(v), L(w)) = \beta_1(v, w) \quad \textit{für alle } v, w \in V_1.$$

Bemerkung 2.2.9.

1. *Äquivalenz von Bilinearformen ist eine Äquivalenzrelation.*
2. *Zwei Bilinearformen β_1, β_2 auf K^n sind genau dann äquivalent, wenn Sie durch einen linearen Koordinatenwechsel ineinander übergehen, d.h. es gibt $A \in \mathrm{GL}_n(K)$, sodass*
$$\beta_2(Ax, Ay) = \beta_1(x, y).$$
3. *Sei* $\mathrm{char}\, K \neq 2$. *Dann ist jede symmetrische Bilinearform β auf V durch $\beta(v, v)$, $v \in V$, bestimmt. Es folgt: symmetrische Bilinearformen β_1, β_2 auf V_1, V_2 sind genau dann äquivalent, wenn es einen Isomorphismus $L : V_1 \to V_2$ gibt, sodass $\beta_2(L(v), L(v)) = \beta_1(v, v)$.*

Beispiel 2.2.10. *Betrachte auf* $\mathbb{R}^2$ *die Bilinearformen*

$$\beta_1(x,y) = x^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} y, \quad \beta_2(x,y) = x^t \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} y.$$

Diese sind symmetrisch, und

$$\beta_1(x,x) = x_1^2 - x_2^2 = (x_1+x_2)(x_1-x_2), \quad \beta_2(x,x) = \frac{1}{2}(x_1x_2 + x_2x_1) = x_1x_2.$$

Setze $x_1' = x_1 + x_2$, $x_2' = x_1 - x_2$, *dann gilt*

$$\beta_2(x',x') = x_1'x_2' = (x_1+x_2)(x_1-x_2) = \beta_1(x,x),$$

also, für $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$, $\beta_2(Ax,Ax) = \beta_1(x,x)$. *Daher sind* β_1 *und* β_2 *äquivalent.*

Wie erkennt man Äquivalenz an den darstellenden Matrizen?

Satz 2.2.11. *Seien* β_1, β_2 *Bilinearformen auf* n*-dimensionalen* K*-Vektorräumen* V_1, V_2 *mit Basen* B_1, B_2.

Dann sind β_1, β_2 *genau dann äquivalent, wenn es eine Matrix* $C \in \mathrm{GL}_n(K)$ *gibt, sodass* $[\beta_2]_{B_2} = C^t[\beta_1]_{B_1}C$.

Beweis. Sei $L : V_1 \to V_2$, sodass $\beta_2(L(v), L(w)) = \beta_1(v,w)$ für alle $v, w \in V_1$ gilt, d.h. die Bilinearformen β_1 und $\beta_{2,L} : (v,w) \mapsto \beta_2(L(v), L(w))$ auf V stimmen überein. Sei $B_2 = \{w_1, \ldots, w_n\}$ und setze $B_1' := \{L^{-1}(w_1), \ldots, L^{-1}(w_n)\}$. Dann gilt

$$[\beta_1]_{B_1'} = [\beta_{2,L}]_{B_1'} = [\beta_2]_{B_2},$$

da $\beta_{2,L}(L^{-1}(w_i), L^{-1}(w_j)) = \beta_2(w_i, w_j)$. Wir wählen $C = [\mathrm{id}_{V_1}]_{B_1', B_1}$, dann

$$C^t[\beta_1]_{B_1}C = [\beta_1]_{B_1'} = [\beta_2]_{B_2}.$$

Sei umgekehrt $C \in \mathrm{GL}_n(K)$, sodass $C^t[\beta_1]_{B_1}C = [\beta_2]_{B_2}$. Sei B_1' jene Basis von V_1, sodass $C = [\mathrm{id}_{V_1}]_{B_1', B_1}$, d.h. für $B_1 = \{v_1, \ldots, v_n\}$ ist $B_1' = \{w_1, \ldots, w_n\}$, mit $w_j = \sum_{i=1}^n c_{ij} v_i$.

Dann gilt $[\beta_2]_{B_2} = [\beta_1]_{B_1'}$, also wählen wir $L : V_1 \to V_2$ als den Isomorphismus, der die Basis B_1' von V_1 auf die Basis B_2 von V_2 abbildet. Es folgt

$$\beta_2(L(w_i), L(w_j)) = \beta_1(w_i, w_j).$$

□

Bemerkung 2.2.12. *Sind zwei Bilinearformen β_1, β_2 auf V_1, V_2 äquivalent, dann unterscheiden sich die Determinanten von $[\beta_1]_{B_1}$ und $[\beta_2]_{B_1}$ nur um einen quadratischen Faktor in K:*

$$\det(C^t[\beta_1]_{B_1}C) = \det(C)^2 \det[\beta_1]_{B_1}.$$

Jede Bilinearform auf V definiert Homomorphismen $V \to V^*$.

Definition 2.2.13. *Sei V ein K-Vektorraum mit einer Bilinearform β. Wir definieren die Abbildungen*

$$\begin{aligned} R_\beta : V &\to V^* & \qquad L_\beta : V &\to V^* \\ v &\mapsto \beta(\cdot, v), & v &\mapsto \beta(v, \cdot). \end{aligned}$$

Bemerkung 2.2.14.

1. *R_β, L_β sind lineare Abbildungen $V \to V^*$, da β bilinear ist.*
2. *Die Abbildungen $\mathrm{Bil}(V, K) \to \mathrm{Hom}(V, V^*)$, $\beta \mapsto R_\beta$, $\beta \mapsto L_\beta$ sind Isomorphismen von K-Vektorräumen (siehe Tutoriumsblatt 5 für den Fall L_β).*

Die Strukturmatrix $[\beta]_B$ beschreibt auch die linearen Abbildungen R_β, L_β.

Satz 2.2.15. *Sei V ein endlich-dimensionaler K-Vektorraum mit Basis B, sei B^* die zu B duale Basis von V^*, und sei β eine Bilinearform auf V. Dann gilt*

$$[\beta]_B = [R_\beta]_{B,B^*}, \quad \textit{und} \quad [\beta]_B^t = [L_\beta]_{B,B^*}.$$

D.h. die Strukturmatrix $[\beta]_B$ stellt $R_\beta : V \to V^*$ bezüglich der Basen B und B^* dar.

Beweis. Sei $B = \{v_1, \ldots, v_n\}$, dann ist $B^* = \{L_1, \ldots, L_n\}$, wobei $L_i : V \to K$ durch $L_i(v_j) = \delta_{ij}$ gegeben ist. Sei $[\beta]_B = (a_{ij})_{1 \le i,j \le n}$, $[R_\beta]_{B,B^*} = (b_{ij})_{1 \le i,j \le n}$, und $[L_\beta]_{B,B^*} = (c_{ij})_{1 \le i,j \le n}$. Dann gilt

$$\begin{aligned} \beta(\cdot, v_j) = R_\beta(v_j) &= \sum_{l=1}^{n} b_{lj} L_l, \\ \beta(v_j, \cdot) = L_\beta(v_j) &= \sum_{l=1}^{n} c_{lj} L_l. \end{aligned}$$

Um b_{ij} zu bestimmen, setzen wir v_i in beide Seiten ein. Es gilt

$$a_{ij} = \beta(v_i, v_j) = \sum_{l=1}^{n} b_{lj} L_l(v_i) = b_{ij},$$
$$a_{ji} = \beta(v_j, v_i) = \sum_{l=1}^{n} c_{lj} L_l(v_i) = c_{ij}.$$

□

Bemerkung 2.2.16.

1. *Insbesondere gilt also: β ist genau dann symmetrisch, wenn $R_\beta = L_\beta$ und genau dann schiefsymmetrisch, wenn $R_\beta = -L_\beta$.*

Wir wollen jene Bilinearformen β auf V charakterisieren, für die R_β und L_β Isomorphismen sind.

2.3 Nichtdegenerierte Bilinearformen und Orthogonalität

Definition 2.3.1. *Eine Bilinearform β auf einem K-Vektorraum V heißt* nichtdegeneriert*, wenn*

für alle $v \in V$ gilt: wenn $\beta(v, w) = 0$ für alle $w \in V$, dann $v = 0$.

Anderenfalls heißt β degeneriert*.*

Bemerkung 2.3.2. *Sei V ein Euklidischer Raum. Dann ist $\langle \cdot, \cdot \rangle$ eine nichtdegenerierte Bilinearform, da $\langle v, v \rangle > 0$ für $v \in V \setminus \{0\}$.*

Nichtdegeneriertheit ist eine Verallgemeinerung der positiven Definitheit von inneren Produkten.

Das innere Produkt auf einem endlich-dimensionalen Euklidischen Raum V induziert einen kanonischen Isomorphismus $V \to V^*$. Gleiches gilt für nichtdegenerierte Bilinearformen.

Satz 2.3.3. *Sei V ein endlich-dimensionaler K-Vektorraum mit einer Bilinearform β. Dann sind folgende Aussagen äquivalent:*

1. *β ist nichtdegeneriert*
2. *$L_\beta : V \to V^*$ ist ein Isomorphismus*

3. für jede Basis B von V ist $[\beta]_B$ invertierbar

Beweis. 1.⇔2. Da $\dim_K V = \dim_K V^*$, ist L_β genau dann ein Isomorphismus, wenn L_β injektiv ist.

Sei $v \in V$. Da $\beta(v,w) = L_\beta(v)(w)$, gilt genau dann $\beta(v,w) = 0$ für alle $w \in V$, wenn $v \in \ker L_\beta$. Daher ist β genau dann nichtdegeneriert, wenn L_β injektiv ist.

2.⇔3. Sei B eine beliebige Basis von V. Da $[\beta]_B = [L_\beta]^t_{B,B^*}$ gilt

$$\begin{aligned}[\beta]_B \text{ invertierbar } &\Leftrightarrow [L_\beta]^t_{B,B^*} \text{ invertierbar } \Leftrightarrow [L_\beta]_{B,B^*} \text{ invertierbar}\\ &\Leftrightarrow L_\beta \text{ invertierbar.}\end{aligned}$$

□

Korollar 2.3.4. *Weitere äquivalente Bedingungen zu 1., 2., 3. aus Satz 2.3.3 sind:*

4. $R_\beta : V \to V^*$ ist ein Isomorphismus

5. für alle $w \in V$ gilt: wenn $\beta(v,w) = 0$ für alle v, dann ist $w = 0$

6. für eine Basis B von V ist $[\beta]_B$ invertierbar

Beweis. 4.⇔3. Für jede Basis B von V gilt $[R_\beta]_{B,B^*} = [\beta]_B$.

4.⇔5. Analog zu *1.⇔2.*

6.⇔3. Die Richtung *3.⇒6.* ist trivial. Für die Gegenrichtung, seien B, B' Basen von V und sei $[\beta]_B$ invertierbar. Dann ist auch

$$[\beta]_{B'} = [\mathrm{id}_V]^t_{B',B}[\beta]_B[\mathrm{id}_V]_{B',B}$$

invertierbar. □

Bemerkung 2.3.5. *Da $\mathrm{Bil}(V,K) \cong \mathrm{Hom}(V,V^*)$, ist die Wahl einer nichtdegenerierten Bilinearform β auf V nach Satz 2.3.3 äquivalent zur Wahl eines Isomorphismus $L_\beta : V \to V^*$.*

Beispiel 2.3.6. *Seien $p, q \in \mathbb{N}_0$ mit $p+q = n$. Auf $\mathbb{R}^n$ ist die symmetrische Bilinearform*

$$\beta_{p,q}(x,y) = y_1y_1 + \cdots + x_py_p - x_{p+1}y_{p+1} - \cdots - x_{p+q}y_{p+q}$$

definiert. Wenn $q > 0$, ist $\beta_{p,q}$ kein inneres Produkt, da $\beta_{p,q}(e_n, e_n) = -1 \leq 0$. Für die Standardbasis $E = \{e_1, \ldots, e_n\}$ gilt allerdings

$$[\beta_{p,q}]_E = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Diese Matrix ist invertierbar, also ist $\beta_{p,q}$ nichtdegeneriert. Der $\mathbb{R}^n$ mit einer Bilinearform $\beta_{p,q}$ wird oft pseudoeuklidischer Raum *genannt. Wir werden sehen, dass jede nichtdegenerierte symmetrische Bilinearform auf $\mathbb{R}^n$ äquivalent zu genau einem $\beta_{p,q}$ ist.*

Wir verallgemeinern den Begriff der Orthogonalität von inneren Produkten auf beliebige Bilinearformen.

Definition 2.3.7. *Sei β eine Bilinearform auf einem Vektorraum V und $v, w \in V$. Dann heißt v* orthogonal *zu w, wenn $\beta(v, w) = 0$.*

Bemerkung 2.3.8. *Meistens wollen wir, dass Orthogonalität symmetrisch in v und w ist, also $\beta(v, w) = 0$ genau dann, wenn $\beta(w, v) = 0$. Das gilt jedenfalls für symmetrische Bilinearformen, da $\beta(v, w) = \beta(w, v)$ und für schiefsymmetrische Bilinearformen, da dann $\beta(v, w) = -\beta(w, v)$.*

Tatsächlich ist Orthogonalität sogar genau dann symmetrisch, wenn β symmetrisch oder schiefsymmetrisch ist (siehe Übung).

Definition 2.3.9. *Sei β eine symmetrische oder schiefsymmetrische Bilinearform auf einem Vektorraum V.*

1. *Eine Menge $M \subset V$ heißt* orthogonal*, wenn je zwei verschiedene Vektoren $v, w \in M$ orthogonal sind.*

2. *Eine* Orthogonalbasis *von V ist eine Basis von V, die orthogonal ist.*

3. *Für $M \subset V$ heißt die Menge*

$$M^\perp := \{v \in V \mid \beta(v, w) = 0 \text{ für alle } w \in M\}$$

der Orthogonalraum *von M.*

Bemerkung 2.3.10.

1. *Der Orthogonalraum $M^\perp$ ist ein Untervektorraum von V. Für einen Untervektorraum $W \subset V$ ist jedoch auch im endlich-dimensionalen Fall $W^\perp$ nicht immer ein Komplement von W.*

2. *Sei $\beta = 0$ die triviale Bilinearform. Dann $M^\perp = V$ für alle $M \subset V$.*

3. *Wenn mehr als eine Bilinearform zugleich betrachtet wird, schreiben wir $M^{\perp_{\beta_1}}, M^{\perp_{\beta_2}}, \ldots$, um die Orthogonalräume zu unterscheiden.*

Beispiel 2.3.11. *Sei $V = \mathbb{R}^3$ und $W = \mathrm{Spann}_{\mathbb{R}}(v_1, v_2)$, wobei $v_1 = (1,0,1)$, $v_2 = (0,1,0) \in V$.*

1. *Die Bilinearform $\beta_{2,1}(x,y) = x_1y_1 + x_2y_2 - x_3y_3$ auf V ist nichtdegeneriert, aber die Einschränkung $\beta_{2,1}|_W$ ist degeneriert: Es gilt*

$$\beta_{2,1}(v_1, v_1) = 1 + 0 - 1 = 0 \quad \text{und } \beta_{2,1}(v_1, v_2) = 0 + 0 + 0 = 0,$$

also, für $w = a_1v_1 + a_2v_2 \in W$,

$$\beta_{2,1}(v_1, w) = a_1\beta_{2,1}(v_1, v_1) + a_2\,\beta_{2,1}(v_1, v_2) = 0.$$

Wir sehen: die Einschränkung einer nichtdegenerierten Bilinearform auf einen Untervektorraum kann degeneriert sein. Genauere Untersuchung: sei $B_1 = \{v_1, v_2\}$ die Basis von W. Wir ergänzen B_1 zu einer Basis $B_2 = \{v_1, v_2, v_3\}$ von V, wobei $v_3 = (0,0,1)$. Dann gilt

$$[\beta_{2,1}]_{B_2} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(K), \quad [\beta_{2,1}|_W]_{B_1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \notin \mathrm{GL}_2(K).$$

2. *Seien B_1, B_2 wie oben, und $\beta(x,y) := x_1x_2 + x_2y_2$ auf $\mathbb{R}^3$. Dann*

$$[\beta]_{B_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \notin \mathrm{GL}_3(K), \quad [\beta|_W]_{B_2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_2(K).$$

Daher ist β degeneriert, aber $\beta|_W$ nichtdegeneriert.

Satz 2.3.12. *Sei V ein endlich-dimensionaler K-Vektorraum mit einer symmetrischen oder schiefsymmetrischen Bilinearform β. Sei $W \subset V$ ein Untervektorraum. Dann sind folgende Aussagen äquivalent:*

1. *Die Einschränkung $\beta|_W$ ist nichtdegeneriert,*

2. $W \cap W^\perp = \{0\}$,

3. $V = W \oplus W^\perp$.

Beweis. 1.⇔2. Es gilt

$$W \cap W^\perp = \{v \in W \mid \beta(v, w) = 0 \text{ für alle } w \in W\}.$$

Daher ist $\beta|_W$ genau dann nichtdegeneriert, wenn $W \cap W^\perp = \{0\}$.

3.⇒2. trivial, da genau dann $V = W \oplus W^\perp$, wenn $V = W + W^\perp$ und $W \cap W^\perp = \{0\}$.

2.⇒3. Wir müssen noch zeigen, dass $V = W + W^\perp$. Betrachte die lineare Abbildung

$$\begin{aligned} \Phi : W &\to W^* \\ w &\mapsto L_\beta(w)|_W = \beta(w, \cdot)|_W : W \to K. \end{aligned}$$

Sei $w \in W$. Dann ist $w \in \ker \Phi$ genau dann, wenn $\beta(w, v) = 0$ für alle $v \in W$, also $\ker \Phi = W \cap W^\perp = \{0\}$. Daher ist Φ ein Isomorphismus $W \to W^*$.

Sei $v \in V$. Dann ist auch

$$L_\beta(v)|_W = \beta(v, \cdot)|_W : W \to K,$$

eine lineare Abbildung, also $L_\beta(v)|_W \in W^*$, und somit $L_\beta(v)|_W = \Phi(w)$ für ein $w \in W$. D.h., für alle $w' \in W$ gilt

$$\beta(v, w') = L_\beta(v)|_W(w') = \Phi(w)(w') = L_\beta(w)|_W(w') = \beta(w, w'),$$

und daher $\beta(v - w, w') = 0$ für alle $w' \in W$. Daher ist $v - w \in W^\perp$, und $v = w + (v - w) \in W + W^\perp$. □

Satz 2.3.13. *Sei β eine nichtdegenerierte symmetrische oder schiefsymmetrische Bilinearform auf einem K-Vektorraum V, und sei W ein Unterraum von V. Dann gilt*

1. $\dim_K V = \dim_K W + \dim_K W^\perp$,

2. $(W^\perp)^\perp = W$,

3. $\beta|_W$ *ist genau dann nichtdegeneriert, wenn* $\beta|_{W^\perp}$ *nichtdegeneriert ist.*

Beweis. Zu *1.* Wir betrachten die lineare Abbildung

$$\begin{aligned}\Phi : V &\to W^* \\ v &\mapsto L_\beta(v)|_W = \beta(v,\cdot)|_W : W \to K.\end{aligned}$$

Es gilt

$$\ker \Phi = \{v \in V \mid \beta(v,w) = 0 \text{ für alle } w \in W\} = W^\perp.$$

Weiters ist Φ surjektiv: sei $\tilde{L} \in W^*$ und $\tilde{B} = \{v_1, \dots, v_m\}$ eine Basis von W. Wir ergänzen $\tilde{B}$ zu einer Basis $B = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ von V. Definiere $L \in V^*$ durch

$$L(v_i) = \begin{cases} \tilde{L}(v_i) & \text{für } 1 \le i \le m \\ 0 & \text{für } m+1 \le i \le n. \end{cases}$$

Dann gilt $L|_W = \tilde{L}$. Da β nichtdegeneriert ist, ist $L_\beta : V \to V^*$ ein Isomorphismus, also $L = L_\beta(v)$ für ein $v \in V$. Es folgt $\Phi(v) = L_\beta(v)|_W = \tilde{L}$.

Nach dem Homomorphiesatz induziert Φ einen Isomorphismus

$$V/W^\perp \cong W^*,$$

also $\dim_K V - \dim_K W^\perp = \dim_K W^* = \dim_K W$.

Zu *2.* Da

$$(W^\perp)^\perp = \{v \in V \mid \beta(v,w) = 0 \text{ für alle } w \in W^\perp\},$$

folgt sofort $W \subset (W^\perp)^\perp$. Weiters gilt

$$\dim_K (W^\perp)^\perp = \dim_K V - \dim_K W^\perp = \dim_K W,$$

also folgt $W = (W^\perp)^\perp$.

Zu *3.* Da $W = (W^\perp)^\perp$, gilt $W \cap W^\perp = W^\perp \cap (W^\perp)^\perp$, und die Aussage folgt aus Satz 2.3.12. □

Beispiel 2.3.14. *Wir betrachten wieder die Bilinearform $\beta_{2,1}$ auf $\mathbb{R}^3$, gegeben durch $\beta_{2,1}(x,y) = x_1y_1 + x_2y_2 - x_3y_3$. Sei $W = \mathrm{Spann}_{\mathbb{R}}(v_1, v_2)$ mit $v_1 = (1,0,1)$, $v_2 = (0,1,0)$, dann haben wir bereits gesehen, dass $\beta_{2,1}$ nichtdegeneriert, aber $\beta_{2,1}|_W$ degeneriert ist.*

Wir berechnen $W^\perp$: Für $x \in \mathbb{R}$ gilt genau dann $x \in W^\perp$, wenn

$$\begin{aligned} 0 &= \beta_{2,1}(x, v_1) = x_1 - x_3 \\ 0 &= \beta_{2,1}(x, v_2) = x_2. \end{aligned}$$

Daher folgt $W^\perp = \mathrm{Spann}_{\mathbb{R}}(v_1)$. Es gilt (wie erwartet) $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W + \dim_{\mathbb{R}} W^\perp$, aber nicht $V = W \oplus W^\perp$, da $W^\perp \subset W$.

Für alternierende Bilinearformen β, die nicht konstant 0 sind, kann es keine Orthogonalbasen geben: wäre $\{v_1, \ldots, v_n\}$ so eine Orthogonalbasis, dann gilt $\beta(v_i, v_j) = 0$ für alle $i \neq j$. Da β alternierend ist, folgt außerdem $\beta(v_i, v_i) = 0$ für alle $1 \leq i \leq n$, also $\beta = 0$.

Im symmetrischen Fall gilt jedoch folgendes.

Satz 2.3.15. *Sei* $\operatorname{char} K \neq 2$*, sei* V *ein endlich-dimensionaler* K*-Vektorraum und* β *eine symmetrische Bilinearform auf* V*. Dann gibt es eine Orthogonalbasis* B *von* V*.*

Beweis. Beweis durch Induktion über $n = \dim_K V$. Für $n = 0$ ist die leere Menge eine Orthogonalbasis, für $n = 1$ ist jede Basis eine Orthogonalbasis. Sei $n \geq 2$ und gelte die Aussage für alle V, β mit $\dim_K V \leq n - 1$.

Falls $\beta = 0$, die Nullabbildung, ist, ist jede Basis von V eine Orthogonalbasis. Sei $\beta \neq 0$. Da β vollständig durch die Werte $\beta(v, v)$ für $v \in V$ bestimmt ist (Lemma 2.1.4), gibt es $v_1 \in V$ mit $\beta(v_1, v_1) \neq 0$. Setze $W := \operatorname{Spann}_K(v_1)$. Da $\beta(v_1, v_1) \neq 0$, ist $\beta\,|_W$ nichtdegeneriert, also $V = W \oplus W^\perp$. Da $\dim W^\perp = n-1$, und da $\beta\,|_{W^\perp}$ auch symmetrisch ist, gibt es nach Induktionsvoraussetzung eine Orthogonalbasis $\{v_2, \ldots, v_n\}$ von $W^\perp$. Dann ist $\{v_1, \ldots, v_n\}$ eine Orthogonalbasis von V. □

Bemerkung 2.3.16.

1. *Sei* $B = \{v_1, \ldots, v_n\}$ *eine Orthogonalbasis von* V*. Dann gilt*
$$[\beta]_B = \begin{pmatrix} \beta(v_1, v_1) & & \\ & \ddots & \\ & & \beta(v_n, v_n) \end{pmatrix}.$$

2. *Insbesondere folgt auch die Umkehrung von Satz 2.3.15: wenn* V *eine Orthogonalbasis hat, ist* β *symmetrisch.*

3. *Insbesondere ist* β *genau dann nichdegeneriert, wenn* $\beta(v_i, v_i) \neq 0$ *für* $1 \leq i \leq n$*.*

4. *Im Gegensatz zu euklidischen Räumen, existieren nicht immer* Orthonormalbasen*, also Orthogonalbasen* B *von* V *mit* $\beta(v_i, v_i) = 1$ *für* $1 \leq i \leq n$ *(siehe Übung). Um eine Orthogonalbasis zu normieren, müssen in* K *Quadratwurzeln aus* $\beta(v_i, v_i)$ *existieren, also Lösungen zu* $X^2 = \beta(v_i, v_i)$*.*

Korollar 2.3.17. *Sei* $\operatorname{char} K \neq 2$ *und* $A \in M(n, n; K)$ *symmetrisch. Dann gibt es eine Matrix* $C \in \mathrm{GL}_n(K)$*, sodass* $C^t AC$ *eine Diagonalmatrix ist.*

Beweis. Sei E die Standardbasis von K^n und B eine Orthogonalbasis von K^n bezüglich der Bilinearform $\beta_A^E(x,y) = x^tAy$. Dann ist $[\beta_A^E]_B$ eine Diagonalmatrix, und

$$[\beta_A^E]_B = [\mathrm{id}_{K^n}]_{B,E}^t \cdot [\beta_A^E]_E \cdot [\mathrm{id}_{K^n}]_{B,E} = [\mathrm{id}_{K^n}]_{B,E}^t \cdot A \cdot [\mathrm{id}_{K^n}]_{B,E}.$$

□

Bemerkung 2.3.18. *Um C und die Diagonalmatrix $D = C^tAC$ zu berechnen, kann man gleichzeitig Zeilen- und Spaltenumformungen anwenden. Zur Erinnerung: elementare Zeilenumformungen von A ergeben sich durch Multiplikation von links mit Elementarmatrizen E.*

1. *Addition des λ-fachen der j-ten Zeile zur i-ten Zeile: $I_n + \lambda E_{ij}$ (I_n mit Eintrag λ an Stelle (i,j))*
2. *Vertauschen der i-ten und j-ten Zeile: Multiplikation mit T_{ij} (I_n mit i-ter und j-ter Zeile vertauscht)*

Elementare Spaltenumformungen sind Zeilenumformungen von A^t, und da $(EA^t)^t = AE^t$, ergeben sich diese durch Multiplikation von links mit E^t. Wenn wir A durch simultane Anwendung derselben Zeilen- und Spaltenumformungen diagonalisieren können, ergibt sich also

$$D = E_k^t \cdots E_1^t A E_1 \cdots E_k = (E_1 \cdots E_k)^t A (E_1 \cdots E_k), \qquad \text{also } C = E_1 \cdots E_k.$$

Wir erhalten folgendes Verfahren: wende simultan Zeilen- und Spaltenumformungen auf A an, um eine Diagonalmatrix D zu erreichen (dabei ist unerheblich, ob zuerst die Zeilen- oder die Spaltenumformung durchgeführt wird). Wende die Spaltenumformungen außerdem auf I_n an, um C zu erhalten.

Beispiel 2.3.19. *Wir suchen $C \in \mathrm{GL}_2(\mathbb{Q})$, sodass C^tAC eine Diagonalmatrix ist, wobei*

$$A = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}.$$

Wir formen um

$$\begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & -1/4 \end{pmatrix}, \begin{pmatrix} 1 & -1/2 \\ 1 & 1/2 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Also gilt $C^tAC = D$*, für*

$$C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Siehe auch Beispiel 2.2.10.

2.4 Quadratische Formen

Wir haben reelle quadratische Formen auf $\mathbb{R}^n$ als Funktionen $Q : \mathbb{R}^n \to \mathbb{R}$ der Form

$$Q(x) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j, \quad c_{ij} \in \mathbb{R},$$

definiert. Nun betrachten wir quadratische Formen auf beliebigen K-Vektorräumen, für $\operatorname{char} K \neq 2$.

Definition 2.4.1. *Sei* $\operatorname{char} K \neq 2$ *und* V *ein* K*-Vektorraum. Eine* quadratische Form *auf* V *ist eine Funktion* $Q : V \to K$*, sodass*

1. $Q(av) = a^2 Q(v)$ *für alle* $v \in V$*,* $a \in K$*,*
2. *Die Funktion* $\beta_Q : V \times V \to K$*,*

$$\beta_Q(v, w) := \frac{1}{2}(Q(v + w) - Q(v) - Q(w))$$

ist eine Bilinearform auf V*.*

Bemerkung 2.4.2.

1. *Die Bilinearform* β_Q *ist symmetrisch.*
2. *Es gilt*

$$Q(v + w) = Q(v) + Q(w) + 2\,\beta_Q(v, w). \tag{2.2}$$

Lemma 2.4.3. *Sei* $Q : V \to K$ *eine quadratische Form auf* V *und* $v_1, \ldots, v_n \in V$*. Dann gilt*

$$Q(v_1 + \cdots + v_n) = Q(v_1) + \cdots + Q(v_n) + 2 \sum_{1 \le i < j \le n} \beta_Q(v_i, v_j).$$

Beweis. Induktion nach n. Für $n = 1$ klar. Gelte die Aussage für $n-1$ Summanden. Dann gilt mit (2.2) und der Induktionsvoraussetzung

$$\begin{aligned} &Q(v_1 + \cdots + v_n) = Q(v_1 + \cdots + v_{n-1}) + Q(v_n) + 2\,\beta_Q(v_1 + \cdots + v_{n-1}, v_n) \\ &= Q(v_1) + \cdots + Q(v_{n-1}) + 2 \sum_{1 \le i < j \le n-1} \beta_Q(v_i, v_j) + Q(v_n) + 2 \sum_{1 \le i \le n} \beta_Q(v_i, v_n) \\ &= Q(v_1) + \cdots + Q(v_n) + 2 \sum_{1 \le i, j \le n} \beta_Q(v_i, v_j). \end{aligned}$$

□

Wenn $\operatorname{char} K \neq 2$, sind symmetrische Bilinearformen und quadratische Formen auf einem K-Vektorraum V im Wesentlichen dasselbe.

Definition 2.4.4. *Sei V ein K-Vektorraum.*

1. *Wir bezeichnen den Vektorraum der symmetrischen Bilinearformen auf V mit* $\mathrm{Bil}_{\mathrm{sym}}(V, K)$.
2. *Sei* $\operatorname{char} K \neq 2$. *Wir bezeichnen den Vektorraum der quadratischen Formen auf V mit* $\mathrm{Quad}(V)$.

Bemerkung 2.4.5. *Beide dieser Mengen sind tatsächlich Vektorräume mit punktweiser Addition und Skalarmultiplikation. In den Übungen haben wir im Fall* $\dim_K V = n < \infty$ *gezeigt, dass* $\dim_K \mathrm{Bil}_{\mathrm{sym}}(V, K) = n(n+1)/2$.

Satz 2.4.6. *Sei* $\operatorname{char} K \neq 2$ *und V ein K-Vektorraum.*

1. *Sei $\beta \in \mathrm{Bil}_{\mathrm{sym}}(V, K)$. Dann ist*

$$\begin{aligned} Q_\beta : V &\to K \\ v &\mapsto \beta(v, v) \end{aligned}$$

   eine quadratische Form auf V.

2. *Die Abbildungen*

$$\begin{aligned} \Phi : \mathrm{Quad}(V) &\to \mathrm{Bil}_{\mathrm{sym}}(V, K) & \qquad \Psi : \mathrm{Bil}_{\mathrm{sym}}(V, K) &\to \mathrm{Quad}(V) \\ Q &\mapsto \beta_Q & \beta &\mapsto Q_\beta \end{aligned}$$

   sind zueinander inverse Isomorphismen.

Beweis. Zu *1.* Für $\beta \in \mathrm{Bil}_{\mathrm{sym}}(V,K)$ gilt

$$Q_\beta(av) = \beta(av,av) = a^2\,\beta(v,v) = a^2 Q_\beta(v).$$

Weiters gilt

$$\begin{aligned}\beta(v,w) &= \frac{1}{2}(\beta(v+w,v+w) - \beta(v,v) - \beta(w,w)) \\ &= \frac{1}{2}(Q_\beta(v+w) - Q_\beta(v) - Q_\beta(w)) = \beta_{Q_\beta}(v,w),\end{aligned}$$

also ist $\beta_{Q_\beta} = \beta$ bilinear.

Zu *2.* Die Abbildungen Φ, Ψ sind K-linear (nachrechnen). Wir haben soeben gesehen, dass $(\Phi \circ \Psi)(\beta) = \beta_{Q_\beta} = \beta$ für alle $\beta \in \mathrm{Bil}_{\mathrm{sym}}(V,K)$ gilt, also $\Phi \circ \Psi = \mathrm{id}_{\mathrm{Bil}_{\mathrm{sym}}(V,K)}$.

Sei umgekehrt $Q \in \mathrm{Quad}(V)$, dann gilt

$$Q_{\beta_Q}(v) = \beta_Q(v,v) = \frac{1}{2}(Q(2v) - 2Q(v)) = Q(v),$$

also $Q_{\beta_Q} = Q$, und somit $(\Psi \circ \Phi)(Q) = Q$, also $\Psi \circ \Phi = \mathrm{id}_{\mathrm{Quad}(V)}$. □

Dadurch, dass wir quadratische Formen mit Bilinearformen identifizieren, können wir sie bezüglich einer Basis B von V auch wieder durch Matrizen darstellen.

Definition 2.4.7. *Sei Q eine quadratische Form auf einem endlichdimensionalen K-Vektorraum V. Sei B eine Basis von V. Die Q bezüglich B* darstellende Matrix $[Q]_B \in M(n,n;K)$ *ist die Strukturmatrix von β_Q bezüglich B, das heißt*

$$[Q]_B := [\beta_Q]_B \in M(n,n;K).$$

Bemerkung 2.4.8.

1. *Da β_Q symmetrisch ist, ist auch $[Q]_B$ symmetrisch.*

2. *Sei $B = \{v_1, \ldots, v_n\}$ und $v = x_1v_1 + \cdots + x_nv_n \in V$. Dann gilt*

$$Q(v) = x^t[Q]_B x.$$

Beweis: $Q(v) = \beta_Q(v,v) = x^t[\beta_Q]_B x = x^t[Q]_B x.$

3. *Sei* $B = \{v_1, \dots, v_n\}$. *Dann gilt*

   $[Q]_B = (a_{ij})_{1 \le i,j \le n}$ *wobei* $a_{ii} = Q(v_i)$ *und* $a_{ij} = \beta_Q(v_i, v_j)$ *für* $i \neq j$.

   Beweis $a_{ij} = \beta_Q(v_i, v_j)$ *für alle* i, j, *da* $[Q]_B = [\beta_Q]_B$, *und* $\beta_Q(v_i, v_i) = Q(v_i)$.

In Koordinaten bezüglich einer fixen Basis sehen alle quadratischen Formen wie quadratische Polynomfunktionen aus.

Satz 2.4.9. *Sei* V *ein endlich-dimensionaler* K*-Vektorraum und* $Q : V \to K$ *eine Funktion. Sei* $B = \{v_1, \dots, v_n\}$ *eine Basis von* V. *Dann sind folgende Aussagen äquivalent:*

1. Q *ist eine quadratische Form*
2. *Es gibt* $c_{ij} \in K$, *sodass*

$$Q(x_1 v_1 + \cdots + x_n v_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j$$

   für alle $x_1, \dots, x_n \in K$ *gilt.*

Beweis. *1.⇒2.* Sei $[Q]_B = (a_{ij})_{i \le i,j \le n}$. Sei $v = x_1 v_1 + \cdots + x_n v_n$. Dann gilt

$$Q(v) = x^t [Q]_B x = \sum_{1 \le i,j \le n} a_{ij} x_i x_j.$$

Da $x_i x_j = x_j x_i$, und $a_{ij} = a_{ji}$, setzen wir

$$c_{ij} := \begin{cases} a_{ij} & \text{für } i = j \\ 2a_{ij} & \text{für } i < j. \end{cases}$$

Dann folgt $Q(v) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j$.

2.⇒1. Sei $v = x_1 v_1 + \cdots + x_n v_n$. Dann gilt

$$Q(v) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j = \sum_{1 \le i,j \le n} a_{ij} x_i x_j = x^t A_Q x,$$

wobei

$$A_Q = (a_{ij}) := \begin{pmatrix} c_{11} & \frac{c_{12}}{2} & \cdots & \frac{c_{1n}}{2} \\ \frac{c_{12}}{2} & c_{22} & & \\ \vdots & & \ddots & \\ \frac{c_{1n}}{2} & & & c_{nn} \end{pmatrix}, \quad \text{d.h. } a_{ij} := \begin{cases} \frac{c_{ij}}{2} & \text{wenn } i < j \\ c_{ij} & \text{wenn } i = j \\ \frac{c_{ji}}{2} & \text{wenn } i > j. \end{cases} \tag{2.3}$$

Definiere die Bilinearform β auf V wie folgt: für $x = x_1v_1 + \cdots + x_nv_n$, sei

$$\beta(v, w) := x^t A_Q y.$$

Dann gilt $Q(v, v) = x^t A_Q x = \beta(v, v)$, also ist Q eine quadratische Form und $\beta_Q = \beta$. □

Bemerkung 2.4.10. *Der Fall $V = \mathbb{R}^n$, und B die Standardbasis, zeigt, dass unsere beiden Definitionen von quadratischen Formen auf $\mathbb{R}^n$ übereinstimmen.*

Beispiel 2.4.11. *Auf K^3 betrachte die quadratische Form*

$$Q(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_3^2.$$

Bezüglich der Standardbasis E gilt

$$[Q]_E = A_Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

und β_Q ist gegeben durch

$$\beta_Q(x, y) = x^t[Q]_E y = x_1y_1 + x_1y_2 + x_2y_1 + x_3y_3.$$

Viele Aussagen über quadratische Formen werden nun direkt aus bereits bewiesenen Aussagen über Bilinearformen folgen. Zum Beispiel wissen wir bereits, wie sich die Strukturmatrizen von Bilinearformen bezüglich verschiedener Basen verhalten.

Satz 2.4.12. *Sei Q eine quadratische Form auf einem endlich-dimensionalen K-Vektorraum V. Seien B_1, B_2 zwei Basen von V. Dann gilt*

$$[Q]_{B_2} = [\mathrm{id}_V]^t_{B_2,B_1}[Q]_{B_1}[\mathrm{id}_V]_{B_2,B_1}.$$

Beweis. Es gilt $[Q]_{B_i} = [\beta_Q]_{B_i}$ und

$$[\beta_Q]_{B_2} = [\mathrm{id}_V]^t_{B_2,B_1}[\beta_Q]_{B_1}[\mathrm{id}_V]_{B_2,B_1}.$$

□

Definition 2.4.13. *Seien K-Vektorräume V_1, V_2 mit quadratischen Formen Q_1, Q_2 gegeben. Dann heißen Q_1 und Q_2* äquivalent*, wenn es einen Isomorphismus $L : V_1 \to V_2$ gibt, sodass $Q_2 \circ L = Q_1$.*

Satz 2.4.14. *Zwei quadratische Formen Q_1, Q_2 auf K-Vektorräumen V_1, V_2 mit $\dim_K V_1 = \dim_K V_2 < \infty$ und Basen B_1, B_2 sind genau dann äquivalent, wenn es $C \in \mathrm{GL}_n(K)$ gibt, sodass*

$$[Q_2]_{B_1} = C^t [Q_1]_{B_1} C.$$

Beweis. Die Gleichung $Q_2 \circ L = Q_1$ gilt genau dann, wenn

$$\beta_{Q_2}(L(v), L(v)) = \beta_{Q_1}(v, v).$$

Da β_{Q_1}, β_{Q_2} symmetrisch sind, ist das äquivalent dazu, dass β_{Q_1} und β_{Q_2} äquivalent sind. Da $[Q_i]_{B_i} = [\beta_{Q_i}]_{B_i}$, folgt der Satz sofort aus Satz 2.2.11. □

Bemerkung 2.4.15. *Im Beweis haben wir insbesondere gezeigt: Quadratische Formen Q_1, Q_2 sind genau dann äquivalent, wenn die Bilinearformen β_{Q_1}, β_{Q_2} äquivalent sind.*

Definition 2.4.16. *Sei $Q : V \to K$ eine quadratische Form auf einem endlich-dimensionalen K-Vektorraum V. Die* Diskriminante *von Q, geschrieben* $\mathrm{discr}(Q)$*, ist die Determinante einer darstellenden Matrix, betrachtet modulo Quadrate in $K \setminus \{0\}$. (Das heißt, Zahlen c und cd^2, die sich nur um einen quadratischen Faktor d^2, mit $d \in K \setminus \{0\}$, unterscheiden, werden als gleich betrachtet.)*

Bemerkung 2.4.17.

1. *Für Basen B_1, B_2 von V gilt $[Q]_{B_2} = C^t [Q]_{B_1} C$, also $\det[Q]_{B_2} = (\det C)^2 \det[Q]_{B_2}$. Daher hängt die Definition der Diskriminante nicht von der Wahl der Basis ab.*

2. *Äquivalente quadratische Formen haben dieselbe Diskriminante. Das folgt aus Satz 2.4.14*

Beispiel 2.4.18. *Wir betrachten auf K^2 die quadratische Form*

$$Q(x, y) = ax^2 + bxy + cy^2, \qquad \text{mit } a, b, c \in K.$$

Für die Standardbasis E von K^2 gilt also

$$[Q]_E = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix},$$

also folgt $\mathrm{discr}(Q) = ac - b^2/4 = 4ac - b^2$.

Satz 2.4.19. *Sei* $\operatorname{char} K \neq 2$ *und* V *ein endlich-dimensionaler* K*-Vektorraum. Sei* Q *eine quadratische Form auf* V*. Dann gibt es eine Basis* $B = \{v_1, \ldots, v_n\}$ *von* V*, sodass* Q *die Form*

$$Q(x_1v_1 + \cdots + x_nv_n) = \sum_{1=1}^{n} a_i x_i^2, \qquad \text{mit } a_i = Q(v_i) \in K,$$

hat.

Beweis. Nach Satz 2.3.15 hat V eine Orthogonalbasis $B = \{v_1, \ldots, v_n\}$ bezüglich der Bilinearform β_Q. Dann gilt

$$[Q]_B = [\beta_Q]_B = \begin{pmatrix} \beta_Q(v_1, v_1) & & \\ & \ddots & \\ & & \beta_Q(v_n, v_n) \end{pmatrix} = \begin{pmatrix} Q(v_1) & & \\ & \ddots & \\ & & Q(v_n) \end{pmatrix}.$$

Es folgt

$$Q(x_1v_1 + \cdots + x_nv_n) = x^t[Q]_B x = \sum_{i=1}^{n} Q(v_i)x_i^2.$$

□

Bemerkung 2.4.20.

1. *Wir nennen eine Basis* B *wie im Satz auch eine* Orthogonalbasis *von* V *bezüglich* Q*. Wir nennen eine Darstellung von* Q *wie im Satz eine Diagonalisierung von* Q*.*

2. *Wenn* $B = \{v_1, \ldots, v_n\}$ *eine Orthogonalbasis von* V *bezüglich* Q *ist, gilt*

$$\operatorname{discr}(Q) = Q(v_1) \cdots Q(v_n).$$

Für quadratische Formen auf K^n ergibt sich folgendes.

Korollar 2.4.21. *Sei* $\operatorname{char} K \neq 2$ *und*

$$Q(x) = \sum_{1 \leq i \leq j \leq n} c_{ij} x_i x_j$$

eine quadratische Form auf K^n*. Dann gibt es* $a_1, \ldots, a_n \in K$ *mit* $\operatorname{discr}(Q) = a_1 \cdots a_n$*, sodass* Q *äquivalent zur quadratischen Form*

$$\tilde{Q}(x) = \sum_{i=1}^{n} a_i x_i^2$$

ist.

Beweis. Sei $B = \{v_1, \dots, v_n\}$ eine Orthogonalbasis von K^n bezüglich Q und $A = [\mathrm{id}_{K^n}]_{B,E}$ die Matrix des Basiswechsels von B zur Standardbasis E. Dann folgt

$$Q(Ax) = Q(x_1v_1 + \cdots + x_nv_n),$$

und die Aussage folgt aus Satz 2.4.19. □

Beispiel 2.4.22. *Wir diagonalisieren die quadratische Form*

$$Q(x, y, z) = xy + xz + yz$$

auf K^3. Bezüglich der Standardbasis E gilt

$$[Q]_E = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

Wir bestimmen $C \in \mathrm{GL}_n(Q)$ und eine Diagonalmatrix $D \in M(n, n; K)$, sodass $C^t[Q]_E C = D$ durch simultane Zeilen- und Spaltenumformungen, und erhalten

$$C = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Siehe Tutorium 6 für ein ähnliches Beispiel. Daher folgt

$$Q(Cx) = x^t(C^t[Q]_E C)x = x^t Dx,$$

also

$$Q(x - y - z, x + y - z, z) = x^2 - y^2 - z^2.$$

Insbesondere gilt $\mathrm{discr}(Q) = 1$.

Definition 2.4.23. *Sei Q eine quadratische Form Q auf einem K-Vektorraum V.*

1. *Q heißt* nichtdegeneriert*, wenn die Bilinearform β_Q nichtdegeneriert ist. Ansonsten heißt Q degeneriert.*

2. *Q heißt* isotrop*, wenn es $v \in V \setminus \{0\}$ gibt, sodass $Q(v) = 0$.*

Bemerkung 2.4.24.

1. *Natürlich gilt immer $Q(0) = 0$.*

2. *Isotropie von Q hängt nur von der Äquivalenzklasse von Q ab.*

   Beweis: sei $Q_2 \circ L = Q_1$ und Q_1 isotrop. Dann gibt es $x \neq 0$ mit $Q_1(x) = 0$. Da L ein Isomorphismus ist, gilt auch $L(x) \neq 0$, und $Q_2(L(x)) = 0$. Wir haben gezeigt: Q_1 isotrop $\Rightarrow$ Q_2 isotrop. Durch Vertauschen der Rollen von Q_1 und Q_2 folgt die Äquivalenz.

3. *Sei $\dim_K V < \infty$. Dann ist Q genau dann nichtdegeneriert, wenn $[Q]_B$ für eine Basis B invertierbar ist.*

4. *Sei $\dim_K V < \infty$. Dann ist Q genau dann nichtdegeneriert, wenn $\operatorname{discr} Q \neq 0$.*

5. *Nichtdegeneriertheit hängt auch nur von der Äquivalenzklasse von Q ab.*

6. *Jede degenerierte quadratische Form ist isotrop.*

   Beweis: Sei β_Q degeneriert, dann gibt es $v \in V \setminus \{0\}$ mit $Q(v) = \beta_Q(v, v) = 0$.

Satz 2.4.25. *Sei $\operatorname{char} K \neq 2$ und Q eine quadratische Form auf einem K-Vektorraum V mit $\dim_K V = 2$. Dann sind die folgenden Aussagen äquivalent:*

1. *Q ist nichtdegeneriert und isotrop*

2. *$\operatorname{discr} Q = -1$ (modulo Quadrate $\neq 0$)*

3. *Es gibt eine Basis B von V, sodass*

$$[Q]_B = \begin{pmatrix} 1 & 0 \\ 0 & -1. \end{pmatrix}$$

Beweis. *1.⇒2.* Sei $B = \{v_1, v_2\}$ eine Orthogonalbasis von V bezüglich Q, also

$$Q(xv_1 + yv_2) = ax^2 + by^2,$$

mit $a, b \neq 0$, da Q nichtdegeneriert ist. Da Q isotrop ist, gibt es $x_0, y_0 \in K$, $(x_0, y_0) \neq (0, 0)$, sodass

$$ax_0^2 + by_0^2 = Q(x_0v_1 + y_0v_2) = 0. \tag{2.4}$$

Es folgt $x_0, y_0 \neq 0$, und $b = -a(x_0/y_0)^2$. Also

$$\operatorname{discr} Q = ab = -a^2 \left(\frac{x_0}{y_0}\right)^2 = -1.$$

2.⇒3. Sei wieder $B = \{v_1, v_2\}$ eine Orthogonalbasis, also gilt (2.4) für $a, b \in K$. Da $\operatorname{discr} Q = -1$, gibt es $u \in K \smallsetminus \{0\}$, sodass $abu^2 = -1$, also

$$Q(xv_1 + yv_2) = ax^2 - \frac{1}{au^2}x^2.$$

Wir ersetzen B durch die Orthogonalbasis $B' = \{v_1, uv_2\}$, sodass

$$[Q]_{B'} = \begin{pmatrix} a & 0 \\ 0 & -1/a \end{pmatrix}.$$

$$\begin{aligned} Q(xv_1 + yuv_2) &= ax^2 - \frac{1}{a}y^2 = a(x^2 - \frac{1}{a^2}y^2) = a(x - \frac{1}{a}y)(x + \frac{1}{a}y) \\ &= (ax - y)(x + \frac{1}{a}y) = (x' - y')(x' + y') = x'^2 - y'^2, \end{aligned}$$

wobei

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x' - y' \\ x' + y' \end{pmatrix} = \begin{pmatrix} ax - y \\ x + (1/a)y \end{pmatrix} = \begin{pmatrix} a & -1 \\ 1 & 1/a \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

Beide Matrizen haben Determinante 2, sind also invertierbar. Sei

$$A = (a_{ij})_{1 \le i,j \le 2} := \begin{pmatrix} 1 & -1 \\ 1 & 1/a \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

und $B'' = \{w_1, w_2\}$, mit $w_1 = a_{11}v_1 + a_{21}uv_2$, $w_2 = a_{12}v_2 + a_{22}uv_2$, d.h. $[\mathrm{id}_V]_{B'',B'} = A$. Dann gilt $x'w_1 + y'w_2 = xv_1 + yuv_2$, also

$$Q(x'w_1 + y'w_2) = Q(xv_1 + yuv_2) = x'^2 - y'^2.$$

Daher folgt

$$[Q]_{B''} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

3.⇒1. Sei $B = \{v_1, v_2\}$. Da $[Q]_B$ invertierbar ist, ist Q nichtdegeneriert, und da z.B.

$$Q(v_1 + v_2) = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1^2 - 1^2 = 0,$$

ist Q isotrop. □

Beispiel 2.4.26. *Die Quadratische Form $Q(x, y) = ax^2 + bxy + cy^2$ auf K^2 hat Diskriminante* $\operatorname{discr} Q = ac - b^2/4 = 4ac - b^2$. *Daher ist Q*

1. *genau dann nichtdegeneriert, wenn* $b^2 - 4ac \neq 0$,
2. *genau dann isotrop, wenn* $b^2 - 4ac$ *ein Quadrat in* K *ist,*
3. *genau dann nichtdegeneriert und isotrop, wenn* $b^2 - 4ac$ *ein Quadrat in* $K \setminus \{0\}$ *ist.*

Über $\mathbb{R}$ haben Sie in der Analysis Ableitungen definiert. Mann kann Ableitungen auch für Polynome über beliebigen Körpern rein formal definieren.

Definition 2.4.27. *Die* formale Ableitung *auf* $K[X]$ *ist die* K*-lineare Abbildung*

$$\begin{aligned} \frac{\partial}{\partial X} &: K[X] \to K[X] \\ p = \sum_{i=0}^{n} a_i X^i &\mapsto \frac{\partial p}{\partial X} = \sum_{i=1}^{n} i a_i X^{i-1}. \end{aligned}$$

Bemerkung 2.4.28.

1. *Diese Definition hat nichts mit Grenzwerten zu tun, stimmt aber natürlich für Polynomfunktionen über* $K = \mathbb{R}$ *mit der Ableitung aus der Analysis überein.*
2. *Sei* V *ein endlich-dimensionaler* K*-Vektorraum und* $B = \{v_1, \ldots, v_n\}$ *eine Basis von* V*. Wir können eine quadratische Form* $Q : V \to K$ *nach allen Koordinaten partiell ableiten und erhalten so lineare Abbildungen* $V \to K$*: sei*

$$Q(x_1 v_1 + \cdots + x_n v_n) = \sum_{1 \leq i \leq j \leq n} c_{ij} x_i x_j.$$

Dann

$$\begin{aligned} \frac{\partial Q(x_1 v_1 + \cdots + x_n v_n)}{\partial x_k} &= \frac{\partial}{\partial x_k} \sum_{1 \leq i \leq j \leq n} c_{ij} x_i x_j \\ &= \sum_{1 \leq i < k} c_{ik} x_i + 2 c_{kk} x_k + \sum_{k < j \leq n} c_{kj} x_j. \end{aligned}$$

Beispiel 2.4.29. *Sei* $Q(x, y) = ax^2 + bxy + cy^2$ *auf* K^2*. Dann gilt*

$$\begin{aligned} \frac{\partial Q(x,y)}{\partial x} &= 2ax + by \\ \frac{\partial Q(x,y)}{\partial y} &= bx + 2cy. \end{aligned}$$

Satz 2.4.30. *Sei* $\operatorname{char} K \neq 2$ *und* V *ein endlich-dimensionaler* K*-Vektorraum mit einer quadratischen Form* Q*. Dann sind folgende Aussagen äquivalent:*

1. *Q ist nichtdegeneriert*
2. *Es gibt keine Basis $B = \{v_1, \dots, v_n\}$ von V, bezüglich der sich Q als Polynom in $n-1$ Variablen schreiben lässt, also*
$$Q(x_1v_1 + \cdots + x_nv_n) = \sum_{1 \leq i \leq j \leq n-1} c_{ij}x_ix_j, \tag{2.5}$$
3. *Für jede Basis $B = \{v_1, \dots, v_n\}$ von V ist $x = 0 \in K^n$ die einzige Lösung des linearen Gleichungssystems*
$$\frac{\partial Q(x_1v_1 + \cdots + x_nv_n)}{\partial x_1}(x) = \cdots = \frac{\partial Q(x_1v_1 + \cdots + x_nv_n)}{\partial x_n}(x) = 0. \tag{2.6}$$

Beweis. *2.⇒1.* Sei $B = \{v_1, \dots, v_n\}$ eine Orthogonalbasis von V. Dann gilt

$$[Q]_B = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

mit $a_i \in K$ und $Q(x_1v_1 + \cdots + x_nv_n) = \sum_{i=1}^{n} a_ix_i^2$. Wenn $a_i = 0$ für ein i gilt, dann können wir (durch Vertauschen von v_i und v_n) annehmen, dass $i = n$. Also ist $Q(x_1v_1 + \cdots + x_nv_n) = \sum_{i=1}^{n-1} a_ix_i^2$ ein Polynom in $n-1$ Variablen, was ausgeschlossen war. Daher folgt $a_i \neq 0$ für alle i, und daher $\operatorname{discr} Q = a_1 \cdots a_n \neq 0$.

1.⇒2. Wir nehmen an, dass *2.* nicht gilt, und zeigen, dass Q dann degeneriert ist. Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V für die (2.5) gilt. Sei $v = x_1v_1 + \cdots + x_nv_n \in V$. Dann gilt $Q(v_n) = 0$ und

$$Q(v + v_n) = \sum_{1 \leq i \leq j \leq n-1} c_{ij}x_ix_j = Q(v),$$

also

$$\beta_Q(v_n, v) = \frac{1}{2}\left(Q(v + v_n) - Q(v_n) - Q(v)\right) = 0.$$

Wir haben gezeigt, dass $\beta_Q(v_n, v) = 0$ für alle $v \in V$, also ist β_Q degeneriert.

1.⇔3. Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V und schreibe

$$Q(x_1v_1 + \cdots + x_nv_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j,$$

also

$$\frac{\partial Q(x_1v_1 + \cdots + x_nv_n)}{\partial x_k}(x) = \begin{pmatrix} c_{1k} & \cdots & c_{k-1,k} & 2c_{kk} & c_{k+1,k} & \cdots & c_{nk} \end{pmatrix} \cdot x.$$

Also ist $x = 0$ genau dann die einzige Lösung zu (2.6), wenn die Matrix

$$C := \begin{pmatrix} 2c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & 2c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & 2c_{nn} \end{pmatrix}$$

invertierbar ist. Allerdings gilt, für die Matrix A_Q aus (2.3), dass $C = 2A_Q = 2[Q]_B$, und daher $\det C = 2^n \operatorname{discr} Q$. □

2.5 Symmetrische Bilinearformen und quadratische Formen über $\mathbb{R}$

Definition 2.5.1. *Sei V ein $\mathbb{R}$-Vektorraum und β eine symmetrische Bilinearform auf V.*

1. *β heißt* positiv definit*, wenn $\beta(v, v) > 0$ für alle $v \in V \setminus \{0\}$.*
2. *β heißt* negativ definit*, wenn $\beta(v, v) < 0$ für alle $v \in V \setminus \{0\}$.*

Sei $A \in M(n, n; \mathbb{R})$ eine symmetrische Matrix.

3. *A heißt* positiv definit*, wenn $x^t Ax > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.*
4. *A heißt* negativ definit*, wenn $x^t Ax < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.*

Bemerkung 2.5.2.

1. *Sei B eine Basis von V. Dann ist β genau dann positiv (bzw. negativ) definit, wenn $[\beta]_B$ positiv (bzw. negativ) definit ist.*
2. *Eine symmetrische Bilinearform auf V ist genau dann positiv definit, wenn sie ein inneres Produkt auf V ist.*

Lemma 2.5.3. *Sei V ein $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform β, und sei $\{v_1, \ldots, v_n\}$ eine Orthogonalbasis von V bezüglich β. Dann gilt*

$$\beta \text{ positiv definit } \Leftrightarrow \beta(v_i, v_i) > 0 \text{ für alle } 1 \leq i \leq n.$$

Beweis. $\Rightarrow$: trivial

$\Leftarrow$: Sei $v = x_1 v_1 + \cdots + x_n v_n \in V \setminus \{0\}$, also $x_i \neq 0$ für ein i. Dann gilt

$$\beta(v, v) = x_1^2 \beta(v_1, v_1) + \cdots + x_n^2 \beta(v_n, v_n) > 0.$$

□

Lemma 2.5.4. *Sei $A \in M(n, n; \mathbb{R})$ symmetrisch, $C \in \mathrm{GL}_n(\mathbb{R})$, und $D = \mathrm{diag}(d_1, \ldots, d_n)$ eine Diagonalmatrix, sodass $C^t A C = D$. Dann gilt*

$$A \text{ positiv definit } \Leftrightarrow d_i > 0 \text{ für alle } 1 \leq i \leq n.$$

Beweis. Die Spalten $c_1, \ldots, c_n$ von C bilden eine Orthogonalbasis von $\mathbb{R}^n$ bezüglich der Bilinearform $\beta_A(x, y) = x^t A y$. In der Tat gilt $c_i = C e_i$, wobei e_i der i-te Einheitsvektor ist. Daher

$$\beta_A(c_i, c_j) = c_i^t A c_j = e_i^t D e_j = \begin{cases} d_i & \text{wenn } i = j \\ 0 & \text{sonst.} \end{cases}$$

Die Aussage folgt sofort aus Lemma 2.5.3. □

Wir zeigen weitere Kriterien für die positive Definitheit von symmetrischen Matrizen.

Definition 2.5.5. *Sei $A = (a_{ij})_{1 \leq i,j \leq n} \in M(n, n; K)$ und $1 \leq k \leq n$. Der k-te* führende Hauptminor *von A ist die Determinante der Teilmatrix $A^{(k)} := (a_{ij})_{1 \leq i,j \leq k} \in M(k, k; K)$.*

Satz 2.5.6. *Sei $A \in M(n, n; \mathbb{R})$ symmetrisch. Dann sind folgende Aussagen äquivalent:*

1. *A ist positiv definit,*
2. *Alle Eigenwerte von A sind > 0,*
3. *Alle führenden Hauptminoren von A sind > 0.*

Beispiel 2.5.7. *Wir weisen nach, dass die Matrix*

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

positiv definit ist, die Bilinearform

$$(x, y) \mapsto x^t Ay = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2 - x_1y_3 - x_3y_1 + 2x_3y_3$$

also ein inneres Produkt auf $\mathbb{R}^3$ *ist. Es gilt*

$$\begin{aligned} \det A^{(1)} &= 2 > 0 \\ \det A^{(2)} &= \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 2 - 1 = 1 > 0 \\ \det A^{(3)} &= \det A = -1 \cdot \begin{vmatrix} -1 & -1 \\ 1 & 0 \end{vmatrix} + 2 \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = -1 + 2 = 1 > 0. \end{aligned}$$

Beweis. (von Satz 2.5.6) *1.⇔2.* Laut Matrixversion des Spektralsatzes gibt es $U \in \mathrm{SO}(n) \subset \mathrm{GL}_n(\mathbb{R})$, sodass

$$U^t AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} =: D,$$

wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A (mit geometrischer Vielfachheit) sind. Damit folgt die Aussage aus Lemma 2.5.4.

1.⇒3. Sei $A^{(k)} = (a_{ij})_{1 \le i,j \le k} \in M(k, k; \mathbb{R})$. Sei A positiv definit. Für $x \in \mathbb{R}^k \setminus \{0\}$ gilt dann

$$x^t A^{(k)} x = \begin{pmatrix} x^t & 0 \end{pmatrix} A \begin{pmatrix} x \\ 0 \end{pmatrix} > 0.$$

Also ist $A^{(k)}$ positiv definit. Da *1.⇔2.* bereits bewiesen ist, sind alle Eigenwerte von $A^{(k)}$ positiv, also auch $\det A^{(k)} > 0$.

3.⇒1. Für $1 \le k \le n$, sei $V_k := \mathrm{Spann}_{\mathbb{R}}(e_1, \ldots, e_k)$, wobei $E = \{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{R}^n$ ist.

Sei $\beta = \beta_A^E$ die durch A dargestellte Bilinearform, d.h. $\beta(e_i, e_j) = a_{ij}$. Dann stellt $A^{(k)}$ die Einschränkung $\beta|_{V_k}$ dar, d.h.

$$A^{(k)} = [\beta|_{V_k}]_{\{e_1, \ldots, e_k\}}.$$

Wir zeigen per Induktion über k, dass alle Einschränkungen $\beta|_{V_k}$ positiv definit sind, also auch β. Für $k = 1$ gilt $A^{(1)} = a_{11} = \det a_{11} > 0$, also ist $\beta|_{V_1}$ positiv definit.

Sei also $\beta|_{V_{k-1}}$ positiv definit. Dann ist $\beta|_{V_{k-1}}$ insbesondere nichtdegeneriert, und daher

$$V_k = V_{k-1} \oplus V_{k-1}^{\perp}.$$

(Hier bezeichnet $V_{k-1}^{\perp}$ den Orthogonalraum von $V_{k-1} \subset V_k$ in V_k bezüglich der Bilinearform $\beta|_{V_k}$.)

Sei $B_{k-1} = \{v_1, \ldots, v_{k-1}\}$ eine Orthogonalbasis von V_{k-1} und $v_k \in V_{k-1}^{\perp} \setminus \{0\}$. Dann ist $B_k = \{v_1, \ldots, v_{k-1}, v_k\}$ eine Orthogonalbasis von V_k. Da $\beta|_{V_{k-1}}$ positiv definit ist, folgt

$$\beta(v_i, v_i) > 0 \text{ für } 1 \leq i \leq k-1.$$

Weiters gilt

$$[\beta|_{V_k}]_{B_k} = \begin{pmatrix} \beta(v_1, v_1) & & \\ & \ddots & \\ & & \beta(v_k, v_k) \end{pmatrix},$$

also $\operatorname{discr} \beta|_{V_k} = \beta(v_1, v_1) \cdots \beta(v_k, v_k)$. Da außerdem $\operatorname{discr} \beta|_{V_k} = \det A^{(k)} > 0$, folgt

$$0 < \det A^{(k)} = c^2\, \beta(v_1, v_1) \cdots \beta(v_k, v_k),$$

für $0 \neq c \in \mathbb{R}$. Da

$$c^2\, \beta(v_1, v_1) \cdots \beta(v_{k-1}, v_{k-1}) > 0,$$

folgt auch $\beta(v_k, v_k) > 0$, also ist $\beta|_{V_k}$ positiv definit. □

Bemerkung 2.5.8. *Sei β eine symmetrische Bilinearform auf einem $\mathbb{R}$-Vektorraum V.*

1. *β ist genau dann negativ definit, wenn $-\beta$ positiv definit ist.*
2. *Sei $\{v_1, \ldots, v_n\}$ eine Orthogonalbasis von V bezüglich β. Dann gilt*

$$\beta \text{ negativ definit} \iff \beta(v_i, v_i) < 0 \text{ für alle } 1 \leq i \leq n.$$

Korollar 2.5.9. *Sei $A \in M(n, n; \mathbb{R})$ symmetrisch. Dann sind folgende Aussagen äquivalent:*

1. *A ist negativ definit*

2. *Wenn* $C^tAC = D$, *mit* $C \in \mathrm{GL}_n(\mathbb{R})$ *und einer Diagonalmatrix* $D = \mathrm{diag}(d_1, \ldots, d_n)$, *dann gilt* $d_i < 0$ *für alle* $1 \leq i \leq n$

3. *Alle Eigenwerte von* A *sind* < 0.

4. *Die führenden Hauptminoren von* A *sind abwechselnd negativ und positiv:* $(-1)^k \det A^{(k)} > 0$ *für* $1 \leq k \leq n$.

Beweis. 1. ist äquivalent zur positiven Definitheit von $-A$.

1.$\Leftrightarrow$*2.* folgt aus Lemma 2.5.4, da $C^tAC = D \Leftrightarrow C^t(-A)C = -D$.

1.$\Leftrightarrow$*3.* folgt aus Satz 2.5.6, da die Eigenwerte von $-A$ genau $-\lambda_i$, für λ_i die Eigenwerte von A, sind.

1.$\Leftrightarrow$*4.* folgt auch aus Satz 2.5.6, da $\det(-A)^{(k)} = (-1)^k \det A^{(k)}$. □

Symmetrische Bilinearformen über $\mathbb{C}$ haben eine sehr einfache Normalform.

Satz 2.5.10. *Sei* V *ein endlich-dimensionaler* $\mathbb{C}$*-Vektorraum und* β *eine symmetrische Bilinearform auf* V. *Dann gibt es eine Basis* B *von* V, *sodass*

$$[\beta]_B = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_k & \\ & 0 \end{pmatrix}. \tag{2.7}$$

Für jede solche Basis gilt $n-k = \dim V^\perp$, *d.h. die Anzahl der* 1*-en ist eindeutig durch* β *bestimmt.*

Beweis. Sei $B = \{v_1, \ldots, v_n\}$ eine Orthogonalbasis von V. Wenn $\beta(v_i, v_i) \neq 0$, können wir v_i durch

$$\frac{1}{\sqrt{\beta(v_i, v_i)}} \cdot v_i$$

ersetzen, und erhalten so $\beta(v_i, v_i) = 1$. Durch Umordnen von B können wir annehmen, dass $\beta(v_i, v_i) = 1$ für $1 \leq i \leq k$, und $\beta(v_i, v_i) = 0$ für $k+1 \leq i \leq n$, mit $k \leq n$. Daher folgt (2.7).

Für $k + 1 \leq i \leq n$ gilt $\beta(v_i, v_j) = 0$ für alle $1 \leq j \leq n$, also $\mathrm{Spann}_{\mathbb{C}}(v_{k+1}, \ldots, v_n) \subset V^\perp$. Sei Umgekehrt $v = a_1v_1 + \cdots + a_nv_n \in V$, mit $a_i \neq 0$ für $1 \leq i \leq k$. Dann $\beta(v, v_i) = a_i \neq 0$, also $v \notin V^\perp$. Daher folgt $\mathrm{Spann}_{\mathbb{C}}(v_{k+1}, \ldots, v_n) = V^\perp$, und daher $n - k = \dim_{\mathbb{C}} V^\perp$. □

Korollar 2.5.11. *Sei V ein $\mathbb{C}$-Vektorraum mit $\dim_{\mathbb{C}} V = n < \infty$, und sei Q eine quadratische Form auf V. Dann gibt es ein eindeutiges $k \in \{1, \dots, n\}$, sodass Q äquivalent zu der quadratischen Form*

$$Q_k(x) := x_1^2 + \cdots + x_k^2$$

auf $\mathbb{C}^n$ ist.

Beweis. Sei $k = n - \dim_{\mathbb{C}} V^{\perp}$, bezüglich der Bilinearform β_Q. Nach Satz 2.5.10 gibt es eine Orthogonalbasis $B = \{v_1, \dots, v_n\}$ von V bezüglich β_Q, sodass $\beta_Q(v_i, v_i) = 1$ für $1 \leq i \leq k$, und $\beta_Q(v_i, v_i) = 0$ für $k+1 \leq i \leq n$, daher gilt

$$Q(x_1 v_1 + \cdots + x_n v_n) = \sum_{i=1}^{n} x_i^2 \, \beta_Q(v_i, v_i) = \sum_{i=1}^{k} x_i^2 = Q_k(x_1, \dots, x_n).$$

Sei also $L : \mathbb{C}^n \to V$ der Isomorphismus gegeben durch $e_i \mapsto v_i$, dann folgt $Q \circ L = Q_k$.

Für $k \neq l$ sind Q_k und Q_l nicht äquivalent. Denn für die Standardbasis E von $\mathbb{R}^n$ gilt,

$$[Q_k]_E = \begin{pmatrix} I_k & \\ & 0 \end{pmatrix}, \quad [Q_l]_E = \begin{pmatrix} I_l & \\ & 0 \end{pmatrix}.$$

Nach Satz 2.5.10 gibt es keine Basis von V, bezüglich der Q_k durch $[Q_l]_E$ dargestellt wird, also kein $C \in \mathrm{GL}_n(\mathbb{C})$, sodass $C^t [Q_k]_E C = [Q_l]_E$. Daher sind Q_k und Q_l nicht äquivalent. □

Über $\mathbb{R}$ ist die Normalform etwas komplizierter, da wir nur aus positiven Zahlen die Wurzel ziehen können.

Lemma 2.5.12. *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und β eine symmetrische Bilinearform auf V. Dann gibt es eine Basis $B = \{v_1, \dots, v_n\}$ von V, sodass*

$$[\beta]_B = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_k & & \\ & -I_l & \\ & & 0 \end{pmatrix} \tag{2.8}$$

wobei $k + l \leq n$ *und* $n - l - k = \dim V^\perp$.

Beweis. Sei $B = \{v_1, \ldots, v_n\}$ eine Orthogonalbasis von V. Wenn $\beta(v_i, v_i) \neq 0$, ersetzen wir v_i durch

$$\frac{1}{\sqrt{|\,\beta(v_i, v_i)|}} \cdot v_i,$$

erhalten also $\beta(v_i, v_i) = \pm 1$. Nach Umordnen folgt also (2.8). Mit demselben Argument wie in Satz 2.5.10 gilt $\mathrm{Spann}_{\mathbb{R}}(v_{k+l+1}, \ldots, v_n) = V^\perp$, also $n-k-l = \dim_{\mathbb{R}} V^\perp$. □

Bemerkung 2.5.13.

1. *Wir haben noch nicht gezeigt, dass* k *und* l *eindeutig sind.*
2. *Setze* $W_+ := \mathrm{Spann}(v_1, \ldots, v_k)$, $W_- := \mathrm{Spann}(v_{k+1}, \ldots, v_{k+l})$. *Dann gilt*
   a) $\beta\,|_{W_+}$ *ist positiv definit,*
   b) $\beta\,|_{W_-}$ *ist negativ definit,*
   c) $V = W_+ \oplus W_- \oplus V^\perp$.

Satz 2.5.14. *(Trägheitssatz von Sylvester) Sei* V *ein endlich-dimensionaler* $\mathbb{R}$*-Vektorraum mit einer symmetrischen Bilinearform* β. *Seien* W_+ *und* W_- *Unterräume von* V, *sodass*

1. $\beta\,|_{W_+}$ *positiv definit ist,*
2. $\beta\,|_{W_-}$ *negativ definit ist, und*
3. $V = W_+ \oplus W_- \oplus V^\perp$ *gilt.*

Sei $r_+ := \dim_{\mathbb{R}} W_+$ *und* $r_- := \dim_{\mathbb{R}} W_-$. *Dann sind* r_+ *und* r_- *eindeutig durch* β *bestimmt, und zwar*

$$r_+ = \max\{\dim_{\mathbb{R}} W \mid W \subset V \text{ Unterraum und } \beta\,|_W \text{ positiv definit}\}$$
$$r_- = \max\{\dim_{\mathbb{R}} W \mid W \subset V \text{ Unterraum und } \beta\,|_W \text{ negativ definit}\}.$$

Beweis. Sei $W \subset V$ ein Unterraum, sodass $\beta\,|_W$ positiv definit ist. Dann folgt

$$W \cap (W_- \oplus V^\perp) = \{0\}.$$

Tatsächlich, sei $v_- \in W_-$ und $v_0 \in V^\perp$, und $w := v_- + v_0 \in W$. Dann gilt

$$\begin{aligned}\beta(w, w) &= \beta(v_- + v_0, v_- + v_0) = \beta(v_-, v_-) + \beta(v_-, v_0) + \beta(v_0, v_-) + \beta(v_0, v_0) \\ &= \beta(v_-, v_-) \leq 0.\end{aligned}$$

Da $\beta|_W$ positiv definit ist, folgt $w = 0$. Es folgt also

$$\dim_{\mathbb{R}} W \leq n - (\dim_{\mathbb{R}} V_- + \dim_{\mathbb{R}} V^\perp) = r_+.$$

Andererseits gilt für $W = W_+$, dass $\beta|_W$ positiv definit ist, und $\dim_{\mathbb{R}} W = r_+$. Wir haben gezeigt, dass

$$r_+ = \max\{\dim_{\mathbb{R}} W \mid W \subset V \text{ Unterraum und } \beta|_W \text{ positiv definit}\}.$$

Die Aussage über r_- folgt analog. □

Definition 2.5.15. *Folgende Bezeichnungen sind für die Invarianten in Satz 2.5.14 gebräuchlich:*

1. *Das Tripel $\sigma(\beta) = (r_+, r_-, r_0)$, wobei $r_0 := \dim_{\mathbb{R}} V^\perp$, heißt die* Signatur *von β.*
2. *Wenn $r_0 = 0$, also wenn β nichtdegeneriert ist, heißt auch (r_+, r_-) die* Signatur *von β.*
3. *$r_+ + r_- = n - r_0$ heißt der* Rang *von β.*
4. *r_- heißt (manchmal) der* Index *von β.*
5. *$r_+ - r_-$ heißt (manchmal) auch die* Signatur *von β.*

Bemerkung 2.5.16.

1. *Die Unterräume W_+, W_- in einer Zerlegung von V wie im Satz sind nicht eindeutig, nur ihre Dimensionen r_+, r_- sind eindeutig.*
2. *Äquivalente symmetrische Bilinearformen auf endlich-dimensionalen $\mathbb{R}$-Vektorräumen haben dieselbe Signatur.*

   Beweis: Sei $\beta_2(L(v), L(w)) = \beta_1(v, w)$, für einen Isomorphismus $L : V_1 \to V_2$. Dann ist $W \mapsto L(W)$ eine bijektive Abbildung zwischen den Unterräumen von V_1 und den Unterräumen von V_2, und $\beta_1|_W$ ist genau dann positiv (bzw. negativ) definit, wenn $\beta_2|_{L(W)}$ positiv (bzw. negativ) definit ist. Daher sind die maximalen Dimensionen dieser Unterräume gleich.

Korollar 2.5.17. *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und β eine symmetrische Bilinearform auf V, mit Signatur $\sigma(\beta) = (r_+, r_-, r_0)$. Dann gilt:*

1. Es gibt eine Basis B von V, sodass

$$[\beta]_B = \begin{pmatrix} I_{r_+} & & \\ & -I_{r_-} & \\ & & 0 \end{pmatrix}.$$

2. Für jede Basis B von V wie in Lemma 2.5.12 gilt $k = r_+$, $l = r_-$.

Beweis. Laut Lemma 2.5.12 gibt es eine Basis B von V, sodass

$$[\beta]_B = \begin{pmatrix} I_k & & \\ & -I_l & \\ & & 0 \end{pmatrix}.$$

Laut Satz 2.5.14 folgt für jede solche Zerlegung, dass

$$\begin{aligned} k &= \dim_{\mathbb{R}} \operatorname{Spann}(w_1, \ldots, w_k) = r_+ \\ l &= \dim_{\mathbb{R}} \operatorname{Spann}(w_{k+1}, \ldots, w_{k+l}) = r_-. \end{aligned}$$

□

Korollar 2.5.18. *Sei V ein $\mathbb{R}$-Vektorraum mit $\dim_{\mathbb{R}} V = n < \infty$, und sei Q eine quadratische Form auf V. Dann gibt es eindeutig bestimmte $k, l \in \mathbb{N}_0$ mit $k + l \leq n$, sodass Q äquivalent zu der quadratischen Form*

$$Q_{k,l}(x) = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2$$

auf $\mathbb{R}^n$ ist.

Beweis. Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V wie in Lemma 2.5.12 für die Bilinearform β_Q. Sei $L : \mathbb{R}^n \to V$ der Isomorphismus mit $e_i \mapsto v_i$. Dann folgt $Q \circ L = Q_{k,l}$.

Die Bilinearform $\beta_{k,l} = \beta_{Q_{k,l}}$ auf $\mathbb{R}^n$ hat Signatur $(k, l, n-k-l)$. Für $(k', l') \neq (k, l)$ sind $Q_{k,l}$ und $Q_{k',l'}$ also nicht äquivalent. □

2.6 Anwendung: Spezielle Relativitätstheorie

Dieser Abschnitt wird bei der Klausur und der Nachklausur nicht geprüft werden.

Wir betrachten ein mathematisches Modell der speziellen Relativitätstheorie, das, aufbauend auf Einsteins Arbeit, von Minkowski entwickelt wurde.

Definition 2.6.1.

1. *Eine* (flache) Raumzeit *ist Tripel* (V, β, v_0)*, wobei* V *ein* 4*-dimensionaler* $\mathbb{R}$*-Vektorraum,* β *eine symmetrische Bilinearform auf* V *mit Signatur* $\sigma(\beta) = (1, 3, 0)$*, und* $v_0 \in V$ *mit* $\beta(v_0, v_0) = 1$ *ist.*

2. $v \in V$ *heißt* zeitartig*, wenn* $\beta(v, v) > 0$*,*

3. $v \in V$ *heißt* raumartig*, wenn* $\beta(v, v) < 0$*,*

4. $v \in V$ *heißt* lichtartig*, wenn* $\beta(v, v) = 0$*,*

5. *Sei* $v \in V$ zeitartig*. Dann heißt* v zukunftsgerichtet*, wenn* $\beta(v, v_0) > 0$*, und sonst* vergangenheitsgerichtet*.*

6. *Wir definieren eine Abbildung* $\|\cdot\| : V \to [0, \infty)$*,* $\|v\| := \sqrt{|\beta(v, v)|}$*.*

Beispiel 2.6.2. *Sei* $V = \mathbb{R}^4$*, wobei die Vektoren* $(t, x, y, z) \in V$ *als Ereignisse mit einer Zeitkoordinate* t *und drei Raumkoordinaten* (x, y, z) *betrachtet werden. Wir wählen die Bilinearform*

$$\beta(v_1, v_2) = \beta_{1,3}(v_1, v_2) = t_1t_2 - x_1x_2 - y_1y_2 - z_1z_2,$$

und $v_0 = (1, 0, 0, 0)$*. Die Einheiten seien so gewählt, dass die Lichtgeschwindigkeit gleich* 1 *ist.*

Zeitartige Vektoren sind jene (t, x, y, z)*, für die* $t^2 > x^2 + y^2 + z^2$ *ist. Das heißt, der räumliche Abstand von* 0 *zu* (x, y, z) *ist kleiner als die Zeit von* 0 *bis* t*, und daher können Ereignisse bei* (t, x, y, z) *kausal mit Ereignissen bei* 0 *zusammenhängen. Die zukunftsgerichteten Vektoren sind genau jene zeitartigen* (t, x, y, z) *mit* $t > 0$*, die also von* 0 *aus in der Zukunft liegen.*

Ereignisse bei raumartigen Vektoren können in keinem kausalen Zusammenhang mir Ereignissen bei 0 *stehen, denn dazu müsste Informationsübertragung mit Überlichtgeschwindigkeit möglich sein.*

Definition 2.6.3.

1. *Ein* Beobachter *in einer Raumzeit* (V, β, v_0) *ist eine stetige Abbildung* $\gamma : I \to V$*, wobei* $I \subset \mathbb{R}$ *ein offenes Intervall ist, sodass:*

   a) γ *ist stückweise stetig differenzierbar.*

   b) *Für jedes* $t \in I$*, sodass* γ *bei* t *stetig differenzierbar ist, ist* $\gamma'(t)$ *zeitartig und zukunftsgerichtet, und es gilt* $\|\gamma'(t)\| = 1$*.*

Für $t_1 < t_2 \in I$ heißt $t_2 - t_1$ die für den Beobachter γ zwischen den Ereignissen $\gamma(t_1)$ und $\gamma(t_2)$ verstrichene Eigenzeit.

2. *Ein Beobachter $\gamma : I \to V$ heißt* gleichförmig bewegt, *wenn γ auf ganz I stetig differenzierbar, und γ' auf I konstant ist.*

Beispiel 2.6.4. *Wir betrachten wieder die Raumzeit (V, β, v_0) aus Beispiel 2.6.2. Hier sind zwei gleichförmig bewegte Beobachter.*

1. *Ruhender Beobachter: γ_1 steht still am (räumlichen) Ursprung:*

$$\gamma_1(t) = (t, 0, 0, 0) \quad \text{für } t \in \mathbb{R}.$$

Sei $t \in \mathbb{R}$. Für γ_1 geschehen alle Ereignisse (t, x, y, t), für $(x, y, z) \in \mathbb{R}^3$ gleichzeitig zum Zeitpunkt t.

2. *Sei $v \in V$ zeitartig und zukunftsgerichtet mit $\|v\| = 1$. Dann ist*

$$\gamma(t) = tv \quad \text{für } t \in \mathbb{R}$$

ein gleichförmig bewegter Beobachter. In seinem eigenen Inertialsystem ist γ ruhend. Mathematisch bedeutet das folgendes: da $\beta|_{\mathbb{R}v}$ nichtdegeneriert und positiv definit, schreibe $V = \mathbb{R}v \oplus (\mathbb{R}v)^\perp$. Dann ist $\beta|_{(\mathbb{R}v)^\perp}$ negativ definit (das war eine Übungsaufgabe). Sei $\{b_1, b_2, b_3\}$ eine Orthogonalbasis von $(\mathbb{R}v)^\perp$ mit $\beta(b_1, b_1) = \beta(b_2, b_2) = \beta(b_3, b_3) = -1$. Dann ist $B = \{v, b_1, b_2, b_3\}$ eine Orthogonalbasis von V, und

$$[\beta]_B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

In Koordinaten bezüglich B ist γ ein ruhender Beobachter. Das heißt, für $t \in \mathbb{R}$ geschehen für γ alle Ereignisse im affinen Unterraum $tv + (\mathbb{R}v)^\perp$ gleichzeitig zum Zeitpunkt t.

Definition 2.6.5. *Sei $\gamma : I \to V$ ein Beobachter. Der* Gleichzeitigkeitsraum *von γ zum Zeitpunkt $t \in I$ ist der affine Unterraum $\gamma(t) + \gamma'(t)^\perp$.*

Beispiel 2.6.6.

1. *Der Gleichzeitigkeitsraum für den ruhenden Beobachter γ_1 zum Zeitpunkt t ist*

$$\gamma_1(t) + \gamma_1'(t)^\perp = (t, 0, 0, 0) + (1, 0, 0, 0)^\perp = \{(t, x, y, z) \mid (x, y, z) \in \mathbb{R}^3\}.$$

2. *Gleichförmig mit Geschwindigkeit $v \in [0,1)$ in x-Richtung bewegter Beobachter:*

$$\gamma_2(t) = \frac{1}{\sqrt{1-v^2}}(t, vt, 0, 0) \quad \text{für } t \in \mathbb{R}.$$

Wir bestimmen den Gleichzeitigkeitsraum für γ_2 zum Zeitpunkt t: es gilt

$$\gamma_2'(t)^\perp = (1, v, 0, 0)^\perp = \{(vx, x, y, z) \mid (x, y, z) \in \mathbb{R}^3\},$$

also

$$\gamma_2(t) + \gamma_2'(t) = \{\frac{1}{\sqrt{1-v^2}}(t, vt, 0, 0) + (vx, x, y, z) \mid (x, y, z) \in \mathbb{R}^3\}.$$

Lemma 2.6.7. *Sei $\gamma : I \to V$ ein gleichförmig bewegter Beobachter und $t_1 < t_2 \in I$. Dann gilt $t_2 - t_1 = \|\gamma(t_2) - \gamma(t_1)\|$.*

Beweis. Sei $v = \gamma'(t)$ für alle $t \in I$. Dann ist v zeitartig, $\|v\| = 1$, und $\gamma(t) = \gamma(t_1) + (t - t_1)v$ für alle $t \in I$. Daher gilt $\gamma(t_2) - \gamma(t_1) = (t_2 - t_1)v$, und

$$\begin{aligned}\beta(\gamma(t_2) - \gamma(t_1), \gamma(t_2) - \gamma(t_1)) &= \beta((t_2 - t_1)v, (t_2 - t_1)v) = (t_2 - t_1)^2 \beta(v, v) \\ &= (t_2 - t_1)^2 \|v\|^2 = (t_2 - t_1)^2.\end{aligned}$$

□

Wir werden nun in unserer Raumzeit $(\mathbb{R}^4, \beta, (1,0,0,0))$ aus Beispiel 2.6.2 einige klassische Relativistische Phänomene beobachten.

Beispiel 2.6.8 (Zeitdilatation)**.** *Wir betrachten zwei Ereignisse $v_1 = (1,0,0,0)$, $v_2 = (2,0,0,0)$ (zwei Lichtblitze am Standort von γ_1), und werden zeigen, dass für γ_2 zwischen diesen Ereignissen mehr Zeit vergeht, als für γ_1.*

Es gilt $v_1 \in \gamma_1(1) + \gamma_1'(1)^\perp$, $v_2 \in \gamma_1(2) + \gamma_1'(2)^\perp$, also finden v_1, v_2 für γ_1 zu den Zeitpunkten $1, 2$ statt. Die verstrichene Eigenzeit beträgt also $2 - 1 = 1$.

Für γ_2 gilt genau dann $(a,0,0,0) \in \gamma_2(t) + \gamma_2'(t)^\perp$, wenn

$$(a, 0, 0, 0) = \frac{1}{\sqrt{1-v^2}}(t, vt, 0, 0) + (vx, x, 0, 0),$$

also $x = -vt/\sqrt{1-v^2}$ und $a = \frac{1}{\sqrt{1-v^2}}(t - v^2 t) = \sqrt{1-v^2}t$. Also geschehen v_1, v_2 für γ_2 zu den Zeitpunkten $1/\sqrt{1-v^2}$ und $2/\sqrt{1-v^2}$. Die verstrichene Eigenzeit beträgt $1/\sqrt{1-v^2} > 1$ und geht für $v \to 1$ sogar gegen ∞.

Beispiel 2.6.9 (Relativität der Gleichzeitigkeit)**.** *Jetzt betrachten wir die Ereignisse $v_1 = (0, -1, 0, 0)$ und $v_2 = (0, 1, 0, 0)$. Wir werden sehen, dass diese für γ_1 gleichzeitig stattfinden, aber nicht für γ_2.*

Tatsächlich gilt $v_1, v_2 \in \gamma_1(0) + \gamma_1'(0)^\perp$, finden beide Ereignisse für γ_1 zum Zeitpunkt 0 statt.

Andererseits gilt genau dann $(0, a, 0, 0) \in \gamma_2(0) + \gamma_2'(0)^\perp$, wenn

$$(0, a, 0, 0) = \frac{1}{\sqrt{1-v^2}}(t, vt, 0, 0) + (vx, x, 0, 0),$$

also $x = -t/(v\sqrt{1-v^2})$ und

$$a = \frac{1}{\sqrt{1-v^2}}vt - \frac{1}{v\sqrt{1-v^2}}t = \frac{-t}{v}\sqrt{1-v^2}.$$

Für γ_2 finden die Ereignisse v_1, v_2 also zu den Zeitpunkten $v/\sqrt{1-v^2}$ und $-v/\sqrt{1-v^2}$ statt. Wenn $v \neq 0$, tritt v_2 für γ_2 also früher ein, als v_1.

Beispiel 2.6.10 (Zwillingsparadoxon)**.** *Wir betrachten einen weiteren Beobachter γ_3, der mit konstanter Geschwindigkeit v bis zu einem Zeitpunkt α in x-Richtung (z.B. nach Alpha Centauri) reist, dann umkehrt und wieder zurück reist. Das heißt,*

$$\gamma_3(t) = \begin{cases} \frac{1}{\sqrt{1-v^2}}(t, vt, 0, 0) & \text{für } 0 \leq t \leq \alpha, \\ \frac{1}{\sqrt{1-v^2}}(t, v(2\alpha - t), 0, 0) & \text{für } \alpha < t \leq 2\alpha, \\ (t, 0, 0, 0) & \text{sonst.} \end{cases}$$

Wir werden sehen, dass zwischen Abreise und Rückkehr von γ_3 für γ_3 weniger Zeit vergeht, als für γ_1. Die Abreise von γ_3 ist das Ereignis

$$v_1 = (0, 0, 0, 0) = \gamma_1(0) = \gamma_3(0).$$

Die Unkehr von γ_3 ist das Ereignis

$$v_2 = \gamma_3(\alpha) = \frac{1}{\sqrt{1-v^2}}(\alpha, v\alpha, 0, 0),$$

und die Rückkehr von γ_3 zur Erde ist das Ereignis

$$v_3 = \gamma_3(2\alpha) = \frac{1}{\sqrt{1-v^2}}(2\alpha, 0, 0, 0) = \gamma_1(\frac{2\alpha}{\sqrt{1-v^2}}).$$

Für γ_1 ist also um den Faktor $1/\sqrt{1-v^2}$ mehr Zeit vergangen, als für γ_3.

Die Eigenzeit für γ_1 zwischen v_1 und v_3 ist nach Lemma 2.6.7 gleich $\|v_3 - v_1\|$. Der Beobachter γ_3 ist aber nur in den Intervallen $(0, \alpha)$ und $(\alpha, 2\alpha)$ gleichförmig bewegt (denn bei α kehrt er um). Die Eigenzeit für γ_3 zwischen v_1 und v_3 ist also $\|v_2 - v_1\| + \|v_3 - v_2\|$. Es gilt

$$\|v_3 - v_1\| = \frac{2\alpha}{\sqrt{1 - v^2}} > 2\alpha = \|v_2 - v_1\| + \|v_3 - v_2\|.$$

Für zeitartige Vektoren in einer Raumzeit gilt so eine umgekehrte Dreiecksungleichung immer.

Kapitel 3

Ringe und Moduln

Wenn man den Körper K in der Definition eines Vektorraums durch einen Ring R ersetzt, erhält man einen Modul über R. Moduln über dem Ring $\mathbb{Z}$ sind genau die abelschen Gruppen, d.h. Moduln verallgemeinern gleichzeitig Vektorräume und abelsche Gruppen. Bevor wir Moduln behandeln, brauchen wir etwas Ringtheorie. In der Vorlesung Algebra werden Sie noch mehr über Ringe lernen.

Literatur: [4].

3.1 Ringhomomorphismen und Ideale

Definition 3.1.1. *Ein* Ring *ist eine nichtleere Menge R mit zwei Verknüpfungen $+ : R \times R \to R$, genannt Addition, und $\cdot : R \times R \to R$, genannt Multiplikation, für die folgendes gilt.*

1. *$(R, +)$ ist eine abelsche Gruppe (mit neutralem Element $0_R = 0$),*
2. *$r \cdot (s \cdot t) = (r \cdot s) \cdot t$ für alle $r, s, t \in R$ (Assoziativität von $\cdot$),*
3. *$r \cdot (s + t) = r \cdot s + r \cdot t$ und $(r + s) \cdot t = r \cdot t + s \cdot t$ für alle $r, s, t \in R$ (Distributivität).*

Der Ring R heißt kommutativ*, wenn*

4. *$r \cdot s = s \cdot r$ für alle $r, s \in R$,*

und ein Ring mit Eins*, falls es ein neutrales Element, genannt $1_R = 1$, bezüglich der Multiplikation gibt, d.h.*

5. *$1r = r = r1$ für alle $r \in R$.*

Beispiel 3.1.2.

1. *Jeder Körper ist ein kommutativer Ring mit Eins. Genauer: Körper sind genau jene kommutativen Ringe $R \neq \{0\}$ mit Eins, in denen jedes Element $r \neq 0$ invertierbar ist.*

2. *$\mathbb{Z}$ ist ein kommutativer Ring mit 1.*

3. *Der Nullring $R = \{0\}$ ist ein kommutativer Ring mit Eins. Er ist der einzige solche Ring, in dem $0 = 1$ gilt.*

4. *Seien $R_1, \ldots, R_n$ Ringe. Dann ist die Menge*

$$R_1 \times \cdots \times R_n := \{(r_1, \cdots, r_n) \mid r_i \in R_i\}$$

   mit komponentenweiser Addition und Multiplikation wieder ein Ring. Das neutrale Element bezüglich der Addition ist $(0, \ldots, 0)$. Wenn jedes R_i ein Ring mit Eins ist, dann ist $(1, \ldots, 1)$ das Einselement von $R_1 \times \cdots \times R_n$. Wenn jedes R_i kommutativ ist, dann auch $R_1 \times \cdots \times R_n$. Wir schreiben auch $R^n := R \times \cdots \times R$.

5. *Sei R kommutativer Ring mit Eins. Dann ist der Polynomring $R[X]$ wieder ein kommutativer Ring mit Eins.*

Definition 3.1.3. *Seien R, S Ringe.*

1. *Ein* Ringhomomorphismus *ist eine Abbildung $\varphi : R \to S$, sodass*

   a) $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$, *und*

   b) $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$

   für alle $r_1, r_2 \in R$ gelten.

2. *Sei $\varphi : R \to S$ ein Ringhomomorphismus. Der* Kern *von φ ist die Menge*

$$\ker \varphi := \{r \in R \mid \varphi(r) = 0\} \subset R.$$

Kerne von Gruppenhomomorphismen sind Normalteiler, Kerne von Vektorraumhomomorphismen sind Untervektorräume. Was sind Kerne von Ringhomomorphismen?

Definition 3.1.4. *Sei R ein Ring. Ein* Ideal *von R ist eine Teilmenge I von R, sodass*

1. *I ist Untergruppe von $(R,+)$, und*
2. *Für alle $i \in I$ und $r \in R$ gilt: $r \cdot i \in I$ und $i \cdot r \in I$.*

Bemerkung 3.1.5.

1. *Die Teilmengen R und $\{0\}$ sind immer Ideale von R. Wir sagen, dass I ein* echtes Ideal *ist, wenn $I \neq R$.*
2. *Wenn R kommutativ ist, dann ist eine additive Untergruppe I von R genau dann ein Ideal, wenn*

$$r \cdot i \in I \quad \text{für alle } r \in R,\ i \in I.$$

Beweis: Wenn R kommutativ ist, gilt $i \cdot r = r \cdot i$.

Lemma 3.1.6. *Die Ideale von $\mathbb{Z}$ sind genau die Teilmengen*

$$m\mathbb{Z} = \{m \cdot n \mid n \in \mathbb{Z}\},$$

für $m \in \mathbb{N}_0$.

Beweis. Wir wissen bereits, dass die Mengen $m\mathbb{Z}$ Untergruppen von $\mathbb{Z}$ sind. Wenn $i = m \cdot n \in m\mathbb{Z}$ und $r \in \mathbb{Z}$, dann ist auch $ri = rmn = m(rn) \in m\mathbb{Z}$, also ist $m\mathbb{Z}$ ein Ideal.

Jede Untergruppe, und damit auch jedes Ideal, hat diese Form: das ist klar für $\{0\}$. Sei $U \neq \{0\}$ eine Untergruppe, dann gibt es $u \in U$, $u > 0$. Sei u minimal mit dieser Eigenschaft. Dann gilt $U = u\mathbb{Z}$. Tatsächlich, sei $x \in U$. Division mit Rest ergibt $x = qu + r$, mit $0 \leq r < u$. Da $r = x - qu \in U$, folgt daher $r = 0$, da u minimal gewählt war. Also gilt $x = qu = uq \in u\mathbb{Z}$. □

Bemerkung 3.1.7.

1. *Für die Ideale $n\mathbb{Z}$ von $\mathbb{Z}$ gilt:*

$$n\mathbb{Z} \subset m\mathbb{Z} \Leftrightarrow n \in m\mathbb{Z} \Leftrightarrow m \mid n.$$

2. *Sei R ein kommutativer Ring mit* 1*. Mengen der Form*

$$(r) = rR := \{r \cdot s \mid s \in R\}$$

sind Ideale von R und werden Hauptideale *genannt.*

Beweis: für $rs_1, rs_2 \in rR$ ist $rs_1 - rs_2 = r(s_1 - s_2) \in rR$, und für $rs \in rR$, $t \in R$ ist $trs = rts \in rR$.

3. *Nicht jedes Ideal jedes kommutativen Rings mit* 1 *ist ein Hauptideal. Z.B. ist das Ideal*

$$(2, x) = 2\mathbb{Z}[X] + X\mathbb{Z}[X] := \{\sum_{i=0}^{n} a_i X^i \mid n \in \mathbb{N}_0,\ a_i \in \mathbb{Z},\ 2 \mid a_0\}$$

kein Hauptideal (siehe Übung).

Lemma 3.1.8. *Sei R ein Ring und I ein Ideal von R.*

1. *Die Menge*
$$R/I = \{r + I \mid r \in R\}$$
ist ein Ring, mit den Verknüpfungen
$$(r + I) + (s + I) = r + s + I$$
$$(r + I) \cdot (s + I) = r \cdot s + I,$$
und Nullelement $0 + I = I$.

2. *Wenn R kommutativ ist, dann ist auch R/I kommutativ.*

3. *Wenn R ein Einselement* 1 *hat, dann hat R/I das Einselement* $1 + I$.

4. *Die natürliche Abbildung $\pi : R \to R/I$, $r \mapsto r + I$, ist ein surjektiver Ringhomomorphismus mit* $\ker \pi = I$.

Beweis. Zu *1.* Da $(I, +)$ eine Untergruppe von $(R, +)$ ist, und $(R, +)$ abelsch, ist $(I, +)$ ein Normalteiler. Es ist bereits aus der Linearen Algebra I bekannt, dass dann R/I mit der angegebenen Addition wieder eine Gruppe ist.

Wir zeigen, dass die angegebene Multiplikation wohldefiniert ist, also nicht von der Wahl der Repräsentanten r, s von $r+I$, $s+I$ abhängt: sei $r+I = r'+I$, d.h. $r - r' \in I$, und $s + I = s' + I$. Dann folgt

$$rs - r's' = r(s - s') + (r - r')s' \in I,$$

da $r, s' \in R$ und $r - r', s - s' \in I$. Daher gilt $rs + I = r's' + I$, und die Multiplikation ist wohldefiniert.

Alle Rechenregeln übertragen sich von R auf R/I: z.B. Assoziativität,

$$\begin{aligned}(r + I)((s + I)(t + I)) &= (r + I)(st + I) = r(st) + I = (rs)t + I \\ &= (rs + I)(t + I) = ((r + I)(s + I))(t + I),\end{aligned}$$

und ein Distributivgesetz,

$$\begin{aligned}(r+I)((s+I)+(t+I)) &= (r+I)(s+t+I) = r(s+t)+I = rs+rt+I \\ &= rs+I+rt+I = (r+I)(s+I)+(r+I)(t+I).\end{aligned}$$

Analog folgt das zweite Distributivgesetz.

Zu *2.* Kommutativität überträgt sich ebenfalls von R auf R/I.

Zu *3.* $(1+I)(r+I) = 1r+I = r+I = r1+I = (r+I)(1+I)$

Zu *4.* Es ist bereits bekannt, dass π ein Gruppenhomomorphismus ist. Offensichtlich ist π surjektiv und auch ein Ringhomomorphismus. Weiters gilt

$$\pi(r) = r+I = I \iff r \in I,$$

also $\ker \pi = I$. □

Lemma 3.1.9. *Sei $\varphi : R \to S$ ein Ringhomomorphismus.*

1. *$\varphi(R)$ ist ein Unterring von S (d.h. $\varphi(R) \subset S$ ist mit den Einschränkungen der Verknüpfungen von S wieder ein Ring).*
2. *Sei I ein Ideal von R, dann ist $\varphi(I)$ ein Ideal von $\varphi(R)$.*
3. *Sei J ein Ideal von S. Dann ist $\varphi^{-1}(J)$ ein Ideal von R und $\ker\varphi \subset \varphi^{-1}(J)$.*

Beweis. Zu *1.* $\varphi(R)$ ist nicht leer und für $r, s \in R$ gilt $\varphi(r)-\varphi(s) = \varphi(r-s) \in \varphi(R)$, also ist $\varphi(R)$ eine Untergruppe von S. Weiters gilt $\varphi(r)\varphi(s) = \varphi(rs)$, also definiert die Einschränkung der Multiplikation auf S eine Verknüpfung

$$\cdot|_{\varphi(R)} : \varphi(R) \times \varphi(R) \to \varphi(R).$$

Assoziativität und Distributivität gelten für $\cdot$, also auch für $\cdot|_{\varphi(R)}$.

Zu *2.* Mit demselben Argument wie in *1.* ist $\varphi(I)$ eine Untergruppe von $\varphi(R)$. Weiters gilt für $r \in R$ und $i \in I$, dass $\varphi(r)\varphi(i) = \varphi(ri) \in \varphi(I)$ und $\varphi(i)\varphi(r) = \varphi(ir) \in \varphi(I)$. Also ist $\varphi(I)$ ein Ideal von $\varphi(R)$.

Zu *3.* Es gilt $0 \in \varphi^{-1}(J)$. Weiters seien $i_1, i_2 \in \varphi^{-1}(J)$, dann gilt $\varphi(i_1 - i_2) = \varphi(i_1) - \varphi(i_2) \in J$, also ist $\varphi^{-1}(J)$ eine additive Untergruppe von R. Für $r \in R$ und $i \in \varphi^{-1}(J)$ gilt $\varphi(ri) = \varphi(r)\varphi(i) \in J$, also $ri \in \varphi^{-1}(J)$. Da $0 \in J$, folgt sofort $\ker\varphi = \varphi^{-1}(\{0\}) \subset \varphi^{-1}(J)$. □

Bemerkung 3.1.10. *$\varphi(I)$ ist nicht unbedingt ein Ideal von S.*

Die Ideale von R/I lassen sich durch Ideale von R beschreiben.

Satz 3.1.11. *Sei R ein Ring, I ein Ideal von R, und $\pi : R \to R/I$ der natürliche Homomorphismus. Dann sind die Abbildungen*

$$\begin{aligned} \{J \mid J \text{ Ideal von } R/I\} &\to \{L \mid L \text{ Ideal von } R \text{ und } I \subset L\} \\ J &\mapsto \pi^{-1}(J) \\ \pi(L) &\hookleftarrow L \end{aligned}$$

zueinander inverse Bijektionen.

Beweis. Da π surjektiv ist, gilt

$$\pi(\pi^{-1}(J)) = J.$$

Wir zeigen umgekehrt, dass

$$\pi^{-1}(\pi(L)) = L.$$

Offensichtlich gilt $L \subset \pi^{-1}(\pi(L))$. Sei $r \in \pi^{-1}(\pi(L))$. Dann gilt $\pi(r) = \pi(l)$, für ein $l \in L$. Also $\pi(r - l) = 0$, und daher $r - l \in \ker \pi \subset L$. Daher auch $r = r - l + l \in L$. □

Beispiel 3.1.12. *Die Ideale von $\mathbb{Z}/n\mathbb{Z}$ sind genau die Mengen*

$$m\mathbb{Z}/n\mathbb{Z} := \pi(m\mathbb{Z}) = \{x + n\mathbb{Z} \mid x \in m\mathbb{Z}\} \subset \mathbb{Z}/n\mathbb{Z},$$

für $m \mid n$.

Wie für Gruppen und Vektorräume, gibt es auch für Ringe einen Homomorphiesatz.

Satz 3.1.13. *(Homomorphiesatz für Ringe) Sei $\varphi : R \to S$ ein Ringhomomorphismus. Dann gilt:*

1. $\ker \varphi$ *ist ein Ideal von R.*

2. *Es gibt einen injektiven Ringhomomorphismus $\widetilde{\varphi} : R/\ker \varphi \to S$ mit*

$$\widetilde{\varphi}(r + \ker \varphi) = \varphi(r)$$

für alle $r \in R$, d.h. $\widetilde{\varphi} \circ \pi = \varphi$, wobei $\pi : R \to R/(\ker \varphi)$ der natürliche Homomorphismus ist.

Beweis. Zu *1.* Wir wissen bereits, dass $\ker\varphi$ eine Untergruppe von $(R,+)$ ist. Sei $r \in R$ und $k \in \ker\varphi$. Dann gilt

$$\varphi(rk) = \varphi(r)\varphi(k) = \varphi(r)\cdot 0 = 0,$$

also $rk \in \ker\varphi$. Analog folgt $kr \in \ker\varphi$. Daher ist $\ker\varphi$ ein Ideal von R.

Zu *2.* Nach dem Homomorphiesatz für Gruppen, gibt es einen Gruppenhomomorphismus mit den angegebenen Eigenschaften. Dieser ist auch ein Ringhomomorphismus, da

$$\begin{aligned}\widetilde{\varphi}((r+\ker\varphi)(s+\ker\varphi)) &= \widetilde{\varphi}(rs+\ker\varphi) = \varphi(rs)\\ &= \varphi(r)\varphi(s) = \widetilde{\varphi}(r+\ker\varphi)\widetilde{\varphi}(s+\ker\varphi).\end{aligned}$$

□

Jeder Körper ist ein kommutativer Ring mit Eins. Wir können die Körper unter diesen Ringen anhand ihrer Ideale charakterisieren.

Lemma 3.1.14. *Sei $R \neq \{0\}$ ein kommutativer Ring mit Eins. Dann sind folgende Aussagen äquivalent:*

1. *R ist ein Körper,*
2. *Die einzigen Ideale von R sind $\{0\}$ und R.*

Beweis. *1.⇒2.* Sei $I \neq \{0\}$ ein Ideal von R und $i \in I \setminus \{0\}$. Da R ein Körper ist, ist i in R invertierbar, und für $r \in R$ gilt $r = (ri^{-1})i \in I$. Also $I = R$.

2.⇒1. Sei $r \in R \setminus \{0\}$. Wir zeigen, dass r invertierbar ist. Betrachte das Hauptideal

$$rR = \{rs \mid s \in R\}.$$

Da $r = r1 \in rR$, gilt $rR \neq \{0\}$. Also folgt $rR = R$. Insbesondere folgt $1 = rs$, für ein $s \in R$, also ist r invertierbar. □

3.2 Maximale Ideale

Definition 3.2.1. *Sei R ein Ring. Ein Ideal I von R heißt* maximal*, wenn es ein maximales Element der Menge aller echten Ideale von R ist. Das heißt, $I \neq R$ und für jedes Ideal J von R gilt*

$$I \subset J \Rightarrow J = I \text{ oder } J = R.$$

Beispiel 3.2.2.

1. *Der Nullring* $\{0\}$ *hat keine maximalen Ideale.*

2. *Sei* R *ein kommutativer Ring mit Eins. Dann gilt:*

$$\{0\} \subset R \text{ ist ein maximales Ideal } \Leftrightarrow R \text{ ist ein Körper.}$$

3. *Die maximalen Ideale von* $\mathbb{Z}$ *sind die Ideale* $p\mathbb{Z}$*, für* p *prim.*

Beweis. Sei p eine Primzahl. Falls $p\mathbb{Z} \subset m\mathbb{Z}$, folgt $m \mid p$, also $m = 1$ oder $m = p$, und daher $m\mathbb{Z} = \mathbb{Z}$ oder $m\mathbb{Z} = p\mathbb{Z}$.

Sei umgekehrt $n \in \mathbb{N}$ zusammengesetzt, d.h. $n = m_1 \cdot m_2$, mit $1 < m_1, m_2 < n$. Dann folgt $1 \mid m_1 \mid n$, aber $n \nmid m_1$, $m_1 \nmid 1$, also

$$n\mathbb{Z} \subsetneq m_1\mathbb{Z} \subsetneq \mathbb{Z}.$$

□

Satz 3.2.3. *Sei* R *ein kommutativer Ring mit Eins und* I *ein Ideal von* R*. Dann sind die folgenden Aussagen äquivalent:*

1. I *ist ein maximales Ideal,*

2. R/I *ist ein Körper.*

Beweis. Die Ideale von R/I stehen in Bijektion zu den Idealen von R, die I enthalten.

1. $\Longleftrightarrow$ die einzigen solchen Ideale sind I und R
2. $\Longleftrightarrow$ die einzigen Ideale von R/I sind $\{0\}$ und R/I. □

Beispiel 3.2.4. *Wir haben (erneut) gezeigt:* $\mathbb{Z}/n\mathbb{Z}$ *ist genau dann ein Körper, wenn* n *prim ist.*

Wir werden zeigen, dass jeder Ring $R \neq \{0\}$ mit Eins maximale Ideale hat. Der Beweis verwendet das Lemma von Zorn.

Satz 3.2.5. *Sei* R *ein Ring mit Eins und* $I \neq R$ *ein echtes Ideal von* R*. Dann hat* R *ein maximales Ideal* M*, sodass* $I \subset M$*.*

Beweis. Sei

$$S := \{J \mid J \text{ Ideal von } R,\ I \subset J \subsetneq R\}.$$

Die Enthaltensrelation $\subset$ ist eine Ordnungsrelation auf S, und maximale Elemente von S sind genau maximale Ideale von R, die I enthalten.

Wir zeigen, dass jede Kette (d.h. jede totalgeordnete Teilmenge) in S eine obere Schranke in S besitzt, dann folgt mit dem Lemma von Zorn die Existenz solcher maximalen Elemente.

Sei also $K \subset S$ eine Kette. Wir setzen $O := \bigcup\{J \mid J \in K\}$, die Vereinigung aller Ideale in K.

Offensichtlich gilt $I \subset O$. Weiters gilt $1 \notin O$, denn sonst gäbe es $J \in K$ mit $1 \in J$, aber dann $r = r1 \in J$ für alle $r \in R$, also $J = R$. Es gilt also $I \subset O \subsetneq R$.

Wir müssen noch zeigen, dass O ein Ideal ist. Wir wissen bereits, dass $O \neq \emptyset$, da $I \subset O$. Seien $x, y \in O$, dann gibt es $J_1, J_2 \in K$ mit $x \in J_1$, $y \in J_2$. Da K eine Kette ist, folgt (ohne Einschränkung der Allgemeinheit) $J_1 \subset J_2$, also $x, y \in J_2$, und daher $x - y \in J_2 \subset O$. Daher ist O eine additive Untergruppe von R. Für $r \in R$ gilt außerdem $rx, xr \in J_1 \subset O$, also ist O ein Ideal. Daher ist O in S, und nach Konstruktion eine obere Schranke der Kette K. □

3.3 Chinesischer Restsatz

Bemerkung 3.3.1.

1. *Seien I, J Ideale eines Rings R. Dann sind auch*

$$I + J := \{x + y \mid x \in I,\ y \in J\},$$
$$IJ := \{x_1y_1 + \cdots + x_ny_n \mid n \in \mathbb{N},\ x_i \in I,\ y_i \in J\},$$

   Ideale von R.

2. *Für Ideale I, J, K von R gilt*

$$I + (J + K) = (I + J) + K,$$
$$I(JK) = (IJ)K,$$
$$I(J + K) = IJ + IK,$$
$$(I + J)K = IK + JK.$$

   Insbesondere definieren, für Ideale $I_1, \ldots, I_n$ von R, auch $I_1 + \cdots + I_n$ und $I_1 \cdots I_n$ Ideale von R.

   Beweis: Übung.

3. *Sei $(I_j)_{j \in M}$, eine Familie von Idealen, indiziert durch eine (endliche oder unendliche) Indexmenge M. Dann ist auch*

$$\bigcap_{j \in M} I_j$$

ein Ideal von R.

Beweis: Der Durchschnitt ist nicht leer, da $0 \in I_j$ für alle $j \in M$ gilt. Seien $i_1, i_2 \in I_j$ für alle $j \in M$, dann auch $i_1 - i_2$, also ist der Durchschnitt eine additive Untergruppe von R. Für $r \in R$ und $i \in I_j$ für alle $j \in M$ sind auch $ri, ir \in I_j$ für alle $j \in M$, also ist der Durchschnitt ein Ideal.

4. *Für Ideale I, J von R gilt*

$$I + I = I, \quad IJ \subset I \cap J.$$

Wenn R ein Ring mit Eins ist, gilt außerdem $RI = I = IR$.

Beispiel 3.3.2.

1. *Seien $m, n \in \mathbb{N}$. Dann gilt*

$$\begin{aligned}(n\mathbb{Z})(m\mathbb{Z}) &= (nm)\mathbb{Z},\\ n\mathbb{Z} \cap m\mathbb{Z} &= \mathrm{kgV}(n,m)\mathbb{Z},\\ m\mathbb{Z} + n\mathbb{Z} &= \mathrm{ggT}(n,m)\mathbb{Z}.\end{aligned}$$

Beweis der dritten Aussage: $m\mathbb{Z} + n\mathbb{Z}$ ist ein Ideal von $\mathbb{Z}$, hat also die Form $k\mathbb{Z}$ für $k \in \mathbb{N}_0$. Da $m = m + 0 \in k\mathbb{Z}$ und $n = 0 + n \in k\mathbb{Z}$, folgt $k \mid m$ und $k \mid n$, also $k \mid \mathrm{ggT}(m,n)$. Da $k \in m\mathbb{Z} + n\mathbb{Z}$, gilt $k = ma + nb$, mit $a, b \in \mathbb{Z}$. Sei d ein gemeinsamer Teiler von m und n, dann $m = dm'$, $n = dn'$, und $k = d(m'a + n'b)$, also ist d auch ein Teiler von k. Daher ist k der größte gemeinsame Teiler von m und n.

2. *Es gilt genau dann $nm = \mathrm{kgV}(n,m)$, wenn $\mathrm{ggT}(n,m) = 1$. Das heißt*

$$n\mathbb{Z} + m\mathbb{Z} = \mathbb{Z} \Leftrightarrow (n\mathbb{Z})(m\mathbb{Z}) = n\mathbb{Z} \cap m\mathbb{Z}.$$

Wir verallgemeinern das letzte Beispiel.

Lemma 3.3.3. *Sei R ein kommutativer Ring mit Eins und I, J Ideale von R mit $I + J = R$. Dann gilt $IJ = I \cap J$.*

Beweis. Es genügt, zu zeigen, dass $I \cap J \subset IJ$. Da R ein Ring mit Eins ist, gilt

$$I \cap J = (I \cap J)R = (I \cap J)(I + J) = (I \cap J)I + (I \cap J)J \subset JI + IJ = IJ.$$

Wir haben verwendet, dass $(I \cap J) \subset J \Rightarrow (I \cap J)I \subset JI$, und dass $IJ = JI$, da R kommutativ ist. □

Lemma 3.3.4. *Sei R ein Ring mit Eins und $I_1, \ldots, I_n$ Ideale von R mit $I_i + I_j = R$ für alle $i \neq j$. Dann gilt für alle $1 \leq i \leq n$, dass*

$$I_i + I_1 \cdots I_{i-1} I_{i+1} \cdots I_n = R.$$

Insbesondere gilt auch

$$I_i + \bigcap_{j \neq i} I_j = R.$$

Beweis. Schreibe $\{J_1, \ldots, J_{n-1}\} = \{I_j \mid j \neq i\}$. Wir zeigen per Induktion über k, dass $I_i + J_1 \cdots J_k = R$ für alle $1 \leq k \leq n-1$ gilt.

$k = 1$: $I_i + J_1 = R$ nach Voraussetzung.

Gelte $I_i + J_1 \cdots J_{k-1} = R$. Dann folgt

$$\begin{aligned} R = RR &= (I_i + J_1 \cdots J_{k-1})R = (I_i + J_1 \cdots J_{k-1})(I_i + J_k) \\ &= I_i^2 + I_i J_k + J_1 \cdots J_{k-1} I_i + J_1 \cdots J_k \subset I_i + J_1 \cdots J_k. \end{aligned}$$

Für $k = n-1$ gilt also $I_i + J_1 \cdots J_{n-1} = R$, also insbesondere

$$I_i + \bigcap_{j \neq i} I_j \supset I_i + \prod_{j \neq i} I_j = R.$$

□

Definition 3.3.5. *Sei R ein Ring, I ein Ideal von R, und $a, b \in R$. Dann ist a* kongruent zu b modulo I*, geschrieben $a \equiv b \bmod I$, wenn $b - a \in I$. Es gilt also*

$$a \equiv b \bmod I \Leftrightarrow a + I = b + I \text{ in } R/I.$$

Bemerkung 3.3.6. *Seien $a_1, b_1, a_2, b_2 \in R$ mit $a_i \equiv b_i \bmod I$. Wir wissen bereits, dass dann auch $a_1 + a_2 \equiv b_1 + b_2 \bmod I$ und $a_1 a_2 \equiv b_1 b_2 \bmod I$ gelten. (Denn Addition und Multiplikation in R/I sind wohldefiniert.)*

Satz 3.3.7 (Chinesischer Restsatz)**.** *Sei R ein Ring mit Eins und $I_1, \ldots, I_n$ Ideale von R, sodass $I_i + I_j = R$ für $i \neq j$ gilt. Seien $a_1, \ldots, a_n \in R$. Dann gibt es $a \in R$, sodass $a \equiv a_i \bmod I_i$ für alle $1 \leq i \leq n$ gilt.*

Für jedes weitere $b \in R$ mit $b \equiv a_i \bmod I_i$ für $1 \leq i \leq n$ gilt $b \equiv a \bmod \bigcap_{1 \leq i \leq n} I_i$.

Beweis. Für $1 \leq i \leq n$ gilt $I_i + \bigcap_{j \neq i} I_j = R$. Seien $b_i \in I_i$, $c_i \in \bigcap_{j \neq i} I_j$, sodass $a_i = b_i + c_i$.

Dann gilt $c_i - a_i = b_i \in I_i$, und $c_i \in I_j$ für $j \neq i$, also

$$c_i \equiv \begin{cases} a_i \bmod I_i, \\ 0 \bmod I_j \text{ für } j \neq i. \end{cases}$$

Für $a = c_1 + \cdots + c_n$ gilt somit

$$a = c_i + \sum_{j \neq i} c_j \equiv a_i + \sum_{j \neq i} 0 = a_i \bmod I_i,$$

für $1 \leq i \leq n$.

Gelte auch $b \equiv a_i \bmod I_i$ für $1 \leq i \leq n$. Dann folgt $b - a \equiv a_i - a_i = 0 \bmod I_i$ für $1 \leq i \leq n$, also $b - a \in \bigcap_{1 \leq i \leq n} I_i$. □

Korollar 3.3.8 (Chinesischer Restsatz für $\mathbb{Z}$)**.** *Seien $m_1, \ldots, m_n \in \mathbb{Z}$ mit $\mathrm{ggT}(m_i, m_j) = 1$ für alle $i \neq j$. Seien $a_1, \ldots, a_n \in \mathbb{Z}$. Dann gibt es $a \in \mathbb{Z}$, sodass $a \equiv a_i \bmod m_i$ für $1 \leq i \leq n$ gilt. Dieses a ist eindeutig modulo $m_1 \cdots m_n$.*

Beweis. Da $\mathrm{ggT}(m_i, m_j) = 1$, folgt $m_i\mathbb{Z} + m_j\mathbb{Z} = \mathbb{Z}$, und wir können den Chinesischen Restsatz mit den Idealen $m_i\mathbb{Z}$ anwenden. Es gibt also ein eindeutiges $a \bmod \bigcap_{1 \leq i \leq n} m_i\mathbb{Z} = m_1 \cdots m_n\mathbb{Z}$, mit $a \equiv a_i \bmod m_i$. □

Beispiel 3.3.9. *Gesucht ist $a \in \mathbb{Z}$ mit*

$$\begin{aligned} a &\equiv 2 \bmod 3 \\ a &\equiv 3 \bmod 5 \\ a &\equiv 5 \bmod 7. \end{aligned}$$

Da $3, 5, 7$ paarweise relativ prim sind, existiert laut Chinesischem Restsatz eine Lösung. Wir suchen c_1, c_2, c_3 wie im Beweis von Satz 3.3.7. Diese erfüllen

$$c_1 \equiv \begin{cases} 2 \mod 3 \\ 0 \mod 35 \end{cases}, \quad c_2 \equiv \begin{cases} 3 \mod 5 \\ 0 \mod 21 \end{cases}, \quad c_3 \equiv \begin{cases} 5 \mod 7 \\ 0 \mod 15. \end{cases}$$

Da $35 \equiv 2 \bmod 3$, wählen wir $c_1 = 35$.
Da $21 \equiv 1 \bmod 5$, wählen wir $c_2 = 3 \cdot 21 = 63$.
Da $15 \equiv 1 \bmod 7$, wählen wir $c_3 = 5 \cdot 15 = 75$.

Wir erhalten also $a = c_1 + c_2 + c_3 = 35 + 63 + 75 \equiv 173$. Um eine minimale Lösung zu erhalten, können wir a noch modulo $3 \cdot 5 \cdot 7 = 105$ reduzieren und erhalten die Lösung $a = 68 \equiv 173 \bmod 105$. Tatsächlich gilt $68 \equiv 2 \bmod 3$, $68 \equiv 3 \bmod 5$, $68 \equiv 5 \bmod 7$.

Korollar 3.3.10 (Chinesischer Restsatz - abstrakte Version)**.** *Sei R ein Ring mit Eins und $I_1, \dots, I_n$ Ideale von R, sodass $I_i + I_j = R$ für alle $i \neq j$ gilt. Dann gibt es einen Isomorphismus von Ringen*

$$\varphi : R/(I_1 \cap \cdots \cap I_n) \to R/I_1 \times \cdots \times R/I_n,$$

sodass

$$\varphi(r + (I_1 \cap \cdots \cap I_n)) = (r + I_1, \dots, r + I_n)$$

für alle $r \in R$ gilt.

Beweis. Betrachte die Abbildung

$$\psi : R \to R/I_1 \times \cdots \times R/I_n, \quad r \mapsto (r + I_1, \dots, r + I_n).$$

Diese ist offensichtlich ein Ringhomomorphismus und nach dem Chinesischen Restsatz surjektiv. Weiters gilt $\ker \psi = I_1 \cap \cdots \cap I_n$, und die Aussage folgt aus dem Homomorphiesatz für Ringe. □

Beispiel 3.3.11. *Da $3 \cdot 5 \cdot 7 = 105$ und $3, 5, 7$ paarweise relativ prim sind, gilt $\mathbb{Z}/105\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}$.*

3.4 Arithmetik in kommutativen Ringen mit Eins

Der Ring $\mathbb{Z}$ hat einige besondere Eigenschaften. Zum Beispiel gilt $m \cdot n = 0$ nur dann, wenn $m = 0$ oder $n = 0$ gilt. Weiters hat jede natürliche Zahl eine eindeutige Faktorisierung in Primzahlen, jedes Ideal von $\mathbb{Z}$ ist ein Hauptideal, und es gibt eine Division mit Rest. Wir werden all diesen Eigenschaften Namen geben und ihre Zusammenhänge untersuchen.

Definition 3.4.1. *Ein kommutativer Ring R mit Eins ist ein Integritätsbereich, wenn für alle $r, s \in R$ gilt*

$$r \cdot s = 0 \Rightarrow r = 0 \text{ oder } s = 0.$$

Beispiel 3.4.2.

1. *Jeder Körper ist ein Integritätsbereich.*
2. *$\mathbb{Z}$ ist ein Integritätsbereich.*
3. *Sei K ein Körper, dann ist der Polynomring $K[X]$ ein Integritätsbereich.*

   Beweis: Seien $f, g \in K[X] \setminus \{0\}$. Schreibe $f = aX^n +$ niedrigere Terme, $g = bX^m +$ niedrigere Terme, mit $a, b \neq 0$. Dann gilt $fg = abX^{n+m} +$ niedrigere Terme. Da $a, b \neq 0$, ist auch $ab \neq 0$. Somit folgt $fg \neq 0$.

4. *Der Ring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Integritätsbereich, wenn n prim ist.*

   Beweis: Wenn n prim ist, ist $\mathbb{Z}/n\mathbb{Z}$ ein Körper, also auch ein Integritätsbereich. Wenn n nicht prim ist, gilt $n = m_1 m_2$, mit $1 < m_1, m_2 < n$. Dann gilt $m_1 + n\mathbb{Z} \neq 0$, $m_2 + n\mathbb{Z} \neq 0$, aber

$$(m_1 + n\mathbb{Z})(m_2 + n\mathbb{Z}) = m_1 m_2 + n\mathbb{Z} = n + n\mathbb{Z} = n\mathbb{Z} = 0.$$

Definition 3.4.3. *Sei R ein Integritätsbereich.*

1. *$r \in R$ heißt* invertierbar*, oder eine* Einheit*, wenn es $s \in R$ gibt, sodass $rs = 1$. Wenn so ein s existiert, ist es eindeutig. Wir schreiben dann $r^{-1} := s$ und bezeichnen r^{-1} als das* inverse Element *zu r.*

2. *$p \in R$ heißt* irreduzibel*, wenn $p \neq 0$, p keine Einheit, und für $r, s \in R$ gilt*

$$p = r \cdot s \implies r \text{ ist Einheit oder } s \text{ ist Einheit.}$$

Beispiel 3.4.4.

1. *In $\mathbb{Z}$ sind die Einheiten genau $\{1, -1\}$, und die irreduziblen Elemente genau $\{\pm p \mid p \text{ Primzahl }\}$.*

2. *In $K[X]$ sind die Einheiten genau die konstanten Polynome in $K^\times$. Jedes lineare Polynom ist irreduzibel, doch es gibt noch viele weitere irreduzible Polynome.*

Definition 3.4.5. *Sei R ein Integritätsbereich.*

1. *Sei $r \in R$, $r \neq 0$. Eine* Zerlegung von r in irreduzible Elemente *(oder* Primzerlegung *oder* Faktorisierung*) von R ist eine Darstellung*

$$r = u \cdot p_1 \cdots p_n,$$

   wobei u eine Einheit in R, $n \in \mathbb{N}_0$, und $p_i \in R$ irreduzibel für $1 \leq i \leq n$.

2. *Die Primzerlegung von r ist* eindeutig*, wenn aus*

$$u \cdot p_1 \cdots p_n = r = v \cdot q_1 \cdots q_m$$

   folgt, dass $m = n$, und dass es eine Permutation $\pi \in S_n$ und Einheiten $u_1, \ldots, u_n$ gibt, sodass $q_i = u_i p_{\pi(i)}$ für alle $1 \leq i \leq n$. Das heißt, die Primzerlegung ist eindeutig bis auf Umordnen der irreduziblen Faktoren und Multiplikation der irreduziblen Faktoren mit Einheiten.

3. *R heißt ein* faktorieller Ring *(oder* ZPE-*Ring, „**Z**erlegung in **P**rimfaktoren **e**indeutig“), wenn jedes $r \in R$, $r \neq 0$, eine eindeutige Zerlegung in irreduzible Elemente hat.*

Beispiel 3.4.6. *$\mathbb{Z}$ ist ein faktorieller Ring.*

Bemerkung 3.4.7. *Sei R ein faktorieller Ring. Die Relation*

$$p \sim q \Leftrightarrow p = uq \text{ für eine Einheit } u \in R$$

ist eine Äquivalenzrelation auf der Menge der irreduziblen Elemente von R. Sei $\mathcal{P}$ ein Repräsentantensystem der Äquivalenzklassen, d.h. $\mathcal{P}$ enthält genau ein Element aus jeder Äquivalenzklasse. Dann hat jedes $r \in R$, $r \neq 0$ eine eindeutige Darstellung

$$r = u \prod_{p \in \mathcal{P}} p^{e_p},$$

mit $e_p \in \mathbb{N}_0$ und $e_p = 0$ für alle bis auf endlich viele $p \in \mathcal{P}$.

Beispiel 3.4.8. *Für $R = \mathbb{Z}$ sei $\mathcal{P} := \{p \mid p \text{ Primzahl}\}$. Dann hat jedes $n \in \mathbb{Z}$, $n \neq 0$ die eindeutige Darstellung*

$$n = \pm \prod_{p \in \mathcal{P}} p^{e_p},$$

mit $e_p \in \mathbb{N}_0$ und $e_p = 0$ für alle bis auf endlich viele $p \in \mathcal{P}$.

Definition 3.4.9. *Sei R ein Integritätsbereich.*

1. *Seien $r, s \in R$. Wir sagen r* teilt *s, geschrieben $r \mid s$, wenn es $t \in R$ gibt, sodass $rt = s$.*

2. *Seien $r_1, \ldots, r_n \in R$ und $r_i \neq 0$ für ein i. Ein Element $d \in R$ heißt* größter gemeinsamer Teiler *von $r_1, \ldots, r_n$, wenn*

   a) *$d \mid r_i$ für $1 \leq i \leq n$, und*

   b) *für jedes $e \in R$ mit $e \mid r_i$ für $1 \leq i \leq n$, folgt $e \mid d$.*

Bemerkung 3.4.10.

1. *Für jedes $r \in R$ gilt $r \mid 0$, da $r \cdot 0 = 0$.*

2. *$0 \mid r \Rightarrow r = 0$.*

3. *Ein Element $r \in R$ ist genau dann eine Einheit, wenn $r \mid 1$.*

4. *Sei d ein größter gemeinsamer Teiler von $r_1, \ldots, r_n$ und $u \in R$ eine Einheit. Dann ist auch ud ein größter gemeinsamer Teiler von $r_1, \ldots, r_n$.*

5. *Seien d, d' größte gemeinsame Teiler von $r_1, \ldots, r_n$. Dann folgt $d \mid d'$ und $d' \mid d$. Für die Hauptideale $(d) = dR$ und $(d') = d'R$ gilt also $(d) = (d')$.*

6. *Aus $(d) = (d')$ folgt insbesondere $d = d'u$ und $d' = dv$, mit $u, v \in R$. Daher $d(1 - uv) = 0$. Da $d \neq 0$, folgt $uv = 1$, also sind u, v Einheiten. Zwei größte gemeinsame Teiler von $r_1, \ldots, r_n$ unterscheiden sich also nur durch Multiplikation mit einer Einheit.*

7. *Sei d ein größter gemeinsamer Teiler von $r_1, \ldots, r_n$. Dann schreiben wir $d = \mathrm{ggT}(r_1, \ldots, r_n)$, obwohl d nur bis auf Multiplikation mit Einheiten bestimmt ist.*

8. *Größte gemeinsame Teiler müssen nicht in jedem Integritätsbereich existieren.*

9. *Es gilt $\mathrm{ggT}(r_1, \ldots, r_n) = \mathrm{ggT}(\mathrm{ggT}(r_1, \ldots, r_{n-1}), r_n)$, wenn ein $r_i \neq 0$ für $1 \leq i \leq n-1$.*

10. *Für $r \neq 0$ gilt $\mathrm{ggT}(r, 0) = r$.*

Lemma 3.4.11. *Sei R ein faktorieller Ring und $r_1, \ldots, r_n \in R \setminus \{0\}$. Dann gibt es einen größten gemeinsamen Teiler von $r_1, \ldots, r_n$ in R. Sei $\mathcal{P}$ ein Repräsentantensystem der Äquivalenzklassen von irreduziblen Elementen in R, und*

$$r_i = u_i \cdot \prod_{p \in \mathcal{P}} p^{e_{i,p}}$$

die Primfaktorzerlegung von r_i, mit einer Einheit u_i, $e_{i,p} \in \mathbb{N}_0$ für alle $p \in \mathcal{P}$, und $e_{i,p} = 0$ für alle bis auf endlich viele $p \in \mathcal{P}$. Dann gilt

$$\mathrm{ggT}(r_1, \ldots, r_n) = \prod_{p \in \mathcal{P}} p^{\min\{e_{i,p} \mid 1 \leq i \leq n\}}.$$

Beweis. Übung. □

Definition 3.4.12. *Sei R ein Integritätsbereich.*

1. *R heißt ein* Hauptidealbereich*, wenn jedes Ideal von R ein Hauptideal ist.*

2. *R heißt ein* Euklidischer Ring*, wenn es eine Abbildung $\phi : R \setminus \{0\} \to \mathbb{N}_0$ gibt, sodass gilt: für alle $x, y \in R$, $y \neq 0$, gibt es $q, r \in R$ mit*

$$\begin{array}{l} x = qy + r, \\ r = 0 \text{ oder } \phi(r) < \phi(y). \end{array} \quad \text{„Division mit Rest“}$$

Beispiel 3.4.13.

1. *$\mathbb{Z}$ ist ein Hauptidealbereich und ein euklidischer Ring ($\phi(n) = |n|$, Division mit Rest).*
2. *Sei K ein Körper. Der Polynomring $K[X]$ ist ein Hauptidealbereich (Übung) und ein euklidischer Ring ($\phi(f) = \operatorname{grad} f$, Polynomdivision).*
3. *$\mathbb{Z}[X]$ ist kein Hauptidealbereich (Übung).*

Lemma 3.4.14. *Sei R ein Hauptidealbereich und $r_1, \ldots, r_n \in R$, nicht alle gleich 0. Dann existiert $\operatorname{ggT}(r_1, \ldots, r_n)$, und*

$$d = \operatorname{ggT}(r_1, \ldots, r_n) \Longleftrightarrow dR = r_1R + \cdots + r_nR.$$

Beweis. Sei $dR = r_1R + \cdots + r_nR$. Dann gilt $r_i \in dR$, also $d \mid r_i$ für $1 \leq i \leq r$. Falls $e \mid r_i$ für $1 \leq i \leq r$, dann $r_1, \ldots, r_n \in eR$, also $r_1R + \cdots + r_nR \subset eR$, also $dR \subset eR$, also $e \mid d$. Daher ist $d = \operatorname{ggT}(r_1, \ldots, r_n)$. Da R ein Hauptidealbereich ist, ist $r_1R + \cdots + r_nR$ ein Hauptideal, also existiert so ein d.

Sei $d = \operatorname{ggT}(r_1, \ldots, r_n)$. Dann $d \mid r_i$ für $1 \leq i \leq n$, also folgt $r_1R + \cdots + r_nR \subset dR$. Sei $e \in R$ mit $eR = r_1R + \cdots + r_nR$. Dann gilt $eR \subset dR$.

Da $e \mid r_i$ für $1 \leq i \leq n$, folgt $e \mid d$, also auch $dR \subset eR$. Es folgt $dR = eR$, und daher auch $dR = r_1R + \cdots + r_nR$. □

Wir zeigen nun Inklusionen zwischen den bisher beschriebenen Typen von Integritätsbereichen.

Satz 3.4.15. *Sei R ein Euklidischer Ring. Dann ist R ein Hauptidealbereich.*

Beweis. Der Beweis ist analog zum Fall $R = \mathbb{Z}$, siehe Lemma 3.1.6. Es gilt $\{0\} = 0R$, sei also $I \neq \{0\}$ und $0 \neq i \in I$, sodass $\phi(i)$ minimal ist. Dann gilt $I = iR$. Tatsächlich, sei $j \in I$. Dann gibt es $q, r \in R$, sodass $j = qi + r$ und $r = 0$ oder $\phi(r) < \phi(i)$. Da $r = j - qi \in I$, und da $\phi(i)$ minimal ist, folgt $r = 0$, also $j = qi \in iR$. □

Lemma 3.4.16. *Sei R ein Hauptidealbereich und $p \in R$, $p \neq 0$.*

1. *p ist genau dann irreduzibel, wenn pR ein maximales ideal ist.*

2. Sei p irreduzibel. Dann gilt für $r_1, r_2 \in R$:

$$p \mid r_1 r_2 \Longrightarrow p \mid r_1 \text{ oder } p \mid r_2.$$

Beweis. Zu *1.* Sei p irreduzibel. Dann ist p nicht invertierbar, also $pR \subsetneq R$. Sei $pR \subset rR \subset R$, dann $p = rs$ für ein $s \in R$. Da p irreduzibel ist, ist r oder s eine Einheit. Im ersten Fall folgt $rR = R$, im zweiten Fall folgt $pR = rsR = rR$. Also ist pR maximal.

Sei umgekehrt pR maximal. Da $pR \neq R$, ist p keine Einheit. Sei $p = rs$. Dann folgt $pR \subset rR \subset R$, also $pR = rR$ oder $rR = R$. Im zweiten Fall ist r eine Einheit. Im ersten Fall folgt $rsR = rR$, also $r = rst$ für ein $t \in R$. Dann $r(1 - st) = 0$, aber $r \neq 0$ (da $r \mid p$), also $st = 1$, und s ist eine Einheit.

Zu *2.* Der Faktorring R/pR ist ein Körper, also insbesondere auch ein Integritätsbereich. Daher folgt

$$\begin{aligned} p \mid r_1 r_2 &\Rightarrow (r_1 + pR)(r_2 + pR) = 0 \Rightarrow r_1 + pR = 0 \text{ oder } r_2 + pR = 0 \\ &\Rightarrow p \mid r_1 \text{ oder } p \mid r_1. \end{aligned}$$

□

Lemma 3.4.17. *Sei R ein Hauptidealbereich und*

$$I_1 \subset I_2 \subset I_3 \subset \cdots$$

eine aufsteigende Kette von Idealen*. Dann wird die Kette stationär, d.h. es gibt $n \in \mathbb{N}$, sodass $I_k = I_n$ für alle $k \geq n$.*

Beweis. Sei $I := \bigcup\{I_i \mid i \in \mathbb{N}\}$. Wir haben bereits im Beweis von Satz 3.2.5 gesehen, dass die Vereinigung einer Kette von Idealen ein Ideal ist. Daher gibt es $r \in I$, sodass $I = rR$. Sei $n \in \mathbb{N}$, sodass $r \in I_n$. Dann folgt für $k \geq n$:

$$rR \subset I_n \subset I_k \subset I = rR,$$

also $I_k = rR$. □

Bemerkung 3.4.18. *Eine aufsteigende Kette von Idealen*

$$r_1 R \subset r_2 R \subset \cdots \subset r_{n-1} R \subset r_n R \subset \cdots$$

ist äquivalent zu einer Teilerkette

$$\cdots \; r_n \mid r_{n-1} \mid \cdots \mid r_2 \mid r_1.$$

Sei $0 \neq r_i = r_{i+1}s_{i+1}$, *mit* $s_{i+1} \in R$. *Dann gilt genau dann* $r_iR \subsetneq r_{i+1}R$, *wenn* s_i *keine Einheit ist.*

Beweis: Ist s_i *eine Einheit, folgt sofort* $r_{i+1} = s_{i+1}^{-1}r_i \in r_iR$, *also* $r_{i+1}R = r_iR$. *Gilt* $r_{i+1}R = r_iR = r_{i+1}s_{i+1}R$, *dann folgt* $r_{i+1} = r_{i+1}s_{i+1}t_{i+1}$, *und daher* $s_{i+1}t_{i+1} = 1$.

Satz 3.4.19. *Sei* R *ein Hauptidealbereich. Dann ist* R *faktoriell.*

Beweis. Wir zeigen zuerst, dass jedes $r_1 \in R \setminus \{0\}$ eine Zerlegung in irreduzible Elemente hat. Das ist klar für Einheiten. Sei r_1 keine Einheit, und nehmen wir an, dass r_1 keine Zerlegung in irreduzible Elemente hat. Wir konstruieren eine Kette

$$r_1R \subsetneq r_2R \subsetneq \cdots$$

von Idealen, die nicht stationär wird, im Widerspruch zu Lemma 3.4.17. Seien $r_1, \ldots, r_{n-1}$ bereits konstruiert, sodass $r_{n-1}R$ keine Faktorisierung hat, und

$$r_1R \subsetneq r_2R \subsetneq \cdots \subsetneq r_{n-1}R.$$

Dann ist r_{n-1} nicht irreduzibel und keine Einheit, also $r_{n-1} = r_ns_n$, wobei r_n und s_n keine Einheiten sind. Wenn r_n und s_n Zerlegungen in irreduzible Elemente hätten, dann auch r_{n-1}. Daher hat, ohne Einschränkung der Allgemeinheit, r_n keine Faktorisierung, und da s_n keine Einheit ist, folgt

$$r_{n-1}R \subsetneq r_nR.$$

Wir müssen noch zeigen, dass die Faktorisierung eindeutig ist. Sei also

$$up_1 \cdots p_n = vq_1 \cdots q_m,$$

mit Einheiten u, v, $n, m \in \mathbb{N}_0$, und irreduziblen Elementen $p_1, \ldots, p_n, q_1, \ldots, q_m \in R$.

Wir zeigen per Induktion über $\max\{n, m\}$, dass $n = m$, und dass es eine Permutation $\pi \in S_n$ und Einheiten $u_1, \ldots, u_n$ gibt, sodass $q_i = u_ip_{\pi(i)}$ für alle $1 \leq i \leq n$ gilt.

Die Aussage ist klar für $\max\{n, m\} = 0$. Sei, ohne Einschränkung der Allgemeinheit, $n \geq 1$. Da $p_n \mid vq_1 \cdots q_m$, und da p_n keine Einheit ist, folgt auch $m \geq 1$. Wegen Lemma 3.4.16 (und Induktion), und da $p_n \nmid v$, folgt $p_n \mid q_j$ für ein $j \in \{1, \ldots, m\}$. Durch Umnumerierung können wir annehmen, dass $p_n \mid q_m$. Da q_m irreduzibel ist, folgt $q_m = u_mp_n$, mit einer Einheit $u_m \in R$, also

$$0 = (up_1 \cdots p_{n-1} - vu_mq_1 \cdots q_{m-1})p_n.$$

Da $p_n \neq 0$, folgt

$$up_1 \cdots p_{n-1} = (vu_m)q_1 \cdots q_{m-1}.$$

Beide Produkte haben einen Faktor weniger, nach Induktionsvoraussetzung gilt also $n = m$, und es gibt eine Permutation $\tilde{\pi} \in S_{n-1}$, sodass $q_i = u_i p_{\pi(i)}$ für $1 \leq i \leq n-1$. Wähle $\pi \in S_n$ als

$$\pi(i) = \begin{cases} \tilde{\pi}(i) & \text{wenn } 1 \leq i \leq n-1 \\ n & \text{wenn } i = n. \end{cases}$$

□

Beispiel 3.4.20. *Wir wissen bereits, dass $K[X]$, der Polynomring über dem Körper K, ein Hauptidealbereich ist. Daher ist $K[X]$ auch faktoriell.*

3.5 Moduln

Definition 3.5.1. *Sei R ein Ring mit Eins. Ein R-*(Links)modul *ist eine abelsche Gruppe M mit einer Abbildung $\cdot : R \times M \to M$ (genannt* Skalarmultiplikation*), sodass für alle $a, b \in R$ und $m, n \in M$ gilt:*

1. $a \cdot (b \cdot m) = (ab) \cdot m$
2. $(a + b) \cdot m = a \cdot m + b \cdot m$
3. $a \cdot (m + n) = a \cdot m + a \cdot n$
4. $1 \cdot m = m$.

Beispiel 3.5.2. *Sei K ein Körper. Dann sind K-Moduln genau die K-Vektorräume.*

Viele Eigenschaften, die für Vektorräume selbstverständlich sind, gelten für Moduln im allgemeinen nicht, wie wir bald sehen werden.

Beispiel 3.5.3.

1. *Jedes Ideal von R ist ein R-Modul.*
2. *Die rationalen Zahlen $\mathbb{Q}$ sind ein $\mathbb{Z}$-Modul. Die Skalarmultiplikation ist die gewöhnliche Multiplikation $\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ eingeschränkt auf $\mathbb{Z} \times \mathbb{Q}$.*

3. *Allgemeiner: sei G eine abelsche Gruppe. Dann ist G ein $\mathbb{Z}$-Modul mit der Skalarmultiplikation*

$$\cdot : \mathbb{Z} \times G \to G$$
$$(n, g) \mapsto n \cdot g := \begin{cases} g + \cdots + g \ (n \text{ mal}) & \text{wenn } n > 0 \\ 0 & \text{wenn } n = 0 \\ (-g) + \cdots + (-g) \ (-n \text{ mal}) & \text{wenn } n < 0. \end{cases}$$

Nach Definition ist auch jeder $\mathbb{Z}$-Modul eine abelsche Gruppe, und die Skalarmultiplikation ist die hier angegebene. Die $\mathbb{Z}$-Moduln sind also genau die abelschen Gruppen.

4. *Sei V ein endlich-dimensionaler K-Vektorraum und $L : V \to V$ ein Endomorphismus. Dann ist V ein $K[X]$-Modul mit der Skalarmultiplikation*

$$f \cdot v = f(L)(v).$$

Nach Cayley-Hamilton gilt $\chi_L \cdot v = 0$ für alle $v \in V$.

5. *Seien $M_1, \ldots, M_n$ R-Moduln. Dann ist das kartesische Produkt*

$$M_1 \times \cdots \times M_n := \{(m_1, \ldots, m_n) \mid m_i \in M_i\}$$

ein R-Modul mit komponentenweiser Addition und Skalarmultiplikation, d.h.

$$(m_1, \ldots, m_n) + (m'_1, \ldots, m'_n) = (m_1 + m'_1, \ldots, m_n + m'_n)$$
$$a \cdot (m_1, \ldots, m_n) = (a \cdot m_1, \ldots, a \cdot m_n).$$

Dieser R-Modul heißt die (äußere) direkte Summe *von $M_1, \ldots, M_n$, und wird oft als $M_1 \oplus \cdots \oplus M_n$ geschrieben.*

6. *Spezialfall: $M = R^n$ ist ein R-Modul.*

Definition 3.5.4. *Sei M ein R-Modul. Eine Teilmenge $N \subset M$ ist ein R-*Untermodul *von M, wenn N eine Untergruppe von M ist und für alle $r \in R$ und $n \in N$ gilt $r \cdot n \in N$.*

Bemerkung 3.5.5. *Ein R-Untermodul N von M bildet mit der Einschränkung der Skalarmultiplikation von M auf $R \times N$ wieder einen R-Modul.*

Definition 3.5.6. *Gegeben seien R-Moduln M, N. Eine Abbildung $L : M \to N$ ist ein* Homomorphismus von R-Moduln *(oder eine R-*lineare Abbildung*), wenn L ein Gruppenhomomorphismus und mit der Skalarmultiplikation verträglich ist. Das heißt*

$$L(m_1 + m_2) = L(m_1) + L(m_2) \quad \text{und} \quad L(a \cdot m) = a \cdot L(m)$$

gelten für alle $m_1, m_2, m \in M$ und $a \in R$.

Lemma 3.5.7. *Sei M ein R-Modul und $N \subset M$ ein Untermodul.*

1. *Die Faktorgruppe*
$$M/N := \{m + N \mid m \in M\}$$
bildet mit der Skalarmultiplikation
$$r \cdot (m + N) := rm + N, \text{ für } r \in R \text{ und } m + N \in M/N,$$
wieder einen R-Modul.

2. *Die natürliche Abbildung $\pi : M \to M/N$, $\pi(m) = m + N$ ist ein surjektiver R-Modul-Homomorphismus mit* $\ker \pi = N$.

Beweis. Zu *1.* Die Skalarmultiplikation ist wohldefiniert: sei $m+N = m'+N$, dann $m - m' \in N$, also auch $a(m - m') \in N$, da N ein Untermodul ist. Daher folgt $am + N = am' + N$. Die R-Modul-Axiome übertragen sich von M auf M/N, z.B. gilt

$$a \cdot (b \cdot (m + N)) = a \cdot (bm + N) = a(bm) + N = (ab)m + N = (ab) \cdot (m + N).$$

Zu *2.* Offensichtlich, z.B. gilt

$$\pi(a \cdot m) = am + N = a \cdot (m + N).$$

□

Satz 3.5.8 (Homomorphiesatz für Moduln)**.** *Sei $L : M \to N$ ein Homomorphismus von R-Moduln. Dann gilt:*

1. $\ker L := \{m \in M \mid L(m) = 0\}$ *ist ein Untermodul von M.*

2. *Es gibt einen injektiven R-Modul-Homomorphismus $\widetilde{L} : M/\ker L \to N$, sodass*
$$\widetilde{L}(m + \ker L) = L(m)$$
für alle $m \in M$ gilt. D.h., $\widetilde{L} \circ \pi = L$, wobei $\pi : M \to M/\ker L$ der natürliche Homomorphismus ist.

Beweis. Zu *1.* Wir wissen bereits, dass $\ker L$ eine Untergruppe von M ist. Sei $m \in \ker L$ und $a \in R$. Dann gilt $L(am) = aL(m) = a \cdot 0 = 0$, also $am \in \ker L$.

Zu *2.* Nach dem Homomorphiesatz für Gruppen gibt es einen Grupenhomomorphismus mit den angegebenen Eigenschaften. Dieser ist auch ein R-Modul-Homomorphismus, denn es gilt

$$\widetilde{L}(r(m + \ker L)) = \widetilde{L}(rm + \ker L) = L(rm) = rL(m) = r\widetilde{L}(m + \ker L).$$

□

3.6 Erzeugendensysteme, lineare Unabhängigkeit, Basen

Hier werden wir erste Unterschiede zwischen Vektorräumen und allgemeinen Moduln feststellen.

Lemma 3.6.1. *Sei M ein R-Modul und sei (N_i), $i \in I$, eine Familie von Untermoduln von M. Dann ist auch $\bigcap\{N_i \mid i \in I\}$ ein Untermodul von M.*

Beweis. Da $0 \in N_i$ für alle $i \in I$, ist $0 \in \bigcap\{N_i \mid i \in I\}$. Seien $m_1, m_2 \in N_i$ für alle $i \in I$, dann auch $m_1 - m_2$. Sei $r \in R$. Dann ist $rm_i \in N_i$ für alle $i \in I$. □

Definition 3.6.2. *Sei M ein R-Modul und $S \subset M$.*

1. *Der* von S erzeugte Untermodul $\langle S \rangle$ von M *ist der kleinste Untermodul von M, der S enthält. Das heißt,*

$$\langle S \rangle := \bigcap\{N \mid N \subset M \text{ Untermodul und } S \subset N\}. \tag{3.1}$$

2. *S heißt* Erzeugendensystem *von M, wenn $\langle S \rangle = M$.*

Bemerkung 3.6.3.

1. *Es gilt*

$$\langle S \rangle = \{\sum_{i=1}^{n} a_i s_i \mid n \in \mathbb{N}, a_i \in R, s_i \in S\}.$$

Beweis: Bezeichne die Menge auf der rechten Seite mit U. Dann ist U offensichtlich ein Untermodul von M, der S enthält. Also $\langle S \rangle \subset U$. Andererseits enthält jeder Untermodul von M, der S enthält, auch Summen von skalaren Vielfachen von Elementen aus S, also U. Daher ist U in jedem der Untermoduln in (3.1) *enthalten, also $U \subset \langle S \rangle$.*

2. *Wenn M als Modul über verschiedenen Ringen R_1, R_2 betrachtet wird, schreiben wir $\langle S\rangle_{R_1}$, $\langle S\rangle_{R_2}$, um die erzeugten Untermoduln zu unterscheiden.*

3. *Sei V ein K-Vektorraum und $S \subset V$. Dann gilt $\langle S\rangle = \mathrm{Spann}_K(S)$.*

Beispiel 3.6.4.

1. *Sei V ein K-Vektorraum mit Basis $B = \{v_1, \ldots, v_m\}$. Dann gilt $\langle B\rangle_K = V$. Weiters ist V als abelsche Gruppe auch ein $\mathbb{Z}$-Modul, und*

$$\langle S\rangle_{\mathbb{Z}} = \{\sum_{i=1}^{m} a_i v_i \mid a_i \in \mathbb{Z}\}.$$

   Anschaulich: $\langle S\rangle_{\mathbb{Z}}$ ist das durch die Vektoren $v_1, \ldots, v_m$ aufgespannte Gitter.

2. *Sei $V = \mathbb{R}^2$ und $B = \{e_1, e_2\}$. Dann gilt $\langle B\rangle_{\mathbb{Z}} = \mathbb{Z}^2 \subset \mathbb{R}^2$.*

Definition 3.6.5. *Sei M ein R-Modul und $S \subset M$.*

1. *S heißt* linear unabhängig, *wenn die einzige Linearkombination von 0 aus Elementen in S die triviale Linearkombination ist. Das heißt, für alle $n \in \mathbb{N}$, $a_1, \ldots, a_n \in R$, und $s_1, \ldots, s_n \in S$ mit $s_i \neq s_j$ für $i \neq j$ gilt*

$$a_1 s_1 + \cdots + a_n s_n = 0 \Longrightarrow a_1 = \cdots = a_n = 0.$$

   Ansonsten heißt S linear abhängig.

2. *S heißt* Basis *von M, wenn S ein linear unabhängiges Erzeugendensystem von M ist.*

3. *M heißt* frei, *wenn M eine Basis hat.*

Beispiel 3.6.6.

1. *Jeder K-Vektorraum ist ein freier K-Modul.*

2. *R^n ist ein freier R-Modul mit Basis $E = \{e_1, \ldots, e_n\}$, wobei $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, mit der 1 an i-ter Stelle.*

Nicht jeder R-Modul ist frei.

Beispiel 3.6.7.

1. *Betrachte den $\mathbb{Z}$-Modul $M = \mathbb{Z}/3\mathbb{Z}$. Für jedes Element $m = a + 3\mathbb{Z} \in M$ gilt $3 \cdot m = 3a + 3\mathbb{Z} = 3\mathbb{Z} = 0$, aber $3 \neq 0$. Daher ist die Menge $\{m\}$ linear abhängig.*

2. *Für K-Vektorräume V gilt: seien $v_1, \ldots, v_n$ linear abhängig, also $a_1v_1 + \cdots + a_nv_n = 0$. Für jedes i mit $a_i \neq 0$ ist dann v_i eine Linearkombination der restlichen Vektoren, z.B.,*

$$v_1 = -\frac{a_2}{a_1}v_2 - \cdots - \frac{a_n}{a_1}v_n.$$

   Das ist falsch für Moduln.

3. *Sei $M = \mathbb{Z}^2$ als $\mathbb{Z}$-Modul, und $m_1 = (2,0)$, $m_2 = (0,3)$, $m_3 = (1,1)$. Dann gilt*

$$3 \cdot m_1 + 2 \cdot m_2 - 6 \cdot m_3 = (6,0) + (0,6) - (6,6) = (0,0),$$

   also ist $\{m_1, m_2, m_2\}$ linear abhängig. Jedoch kann kein m_i als Linearkombination der beiden anderen geschrieben werden. Das lässt sich z.B. so beweisen: es gilt $\mathbb{Z}^2 \subset \mathbb{Q}^2$, und $B = \{m_1, m_2\}$ ist eine Basis von $\mathbb{Q}^2$. Daher ist die Darstellung

$$m_3 = \frac{1}{2} \cdot m_1 + \frac{1}{3} \cdot m_2$$

   eindeutig. Doch die Koeffizienten sind nicht in $\mathbb{Z}$, also ist m_3 keine $\mathbb{Z}$-Linearkombination von m_1, m_2. Analog gilt

$$m_1 = -\frac{2}{3} \cdot m_2 + 2 \cdot m_3, \quad m_2 = -\frac{3}{2} \cdot m_1 + 3 \cdot m_3.$$

Der Rest dieses Kapitels dient hauptsächlich dazu, diese Phänomene in den Griff zu bekommen, zumindest über „netten“ Ringen. Wir werden nur endlich erzeugte Moduln betrachten, d.h. Moduln mit einem endlichen Erzeugendensystem, obwohl einige der Aussagen auch für nicht endlich erzeugte Moduln gelten.

Wir wollen zuerst den Rang eines freien R-Moduls, analog zur Dimension eines Vektorraums, als die Kardinalität einer Basis definieren. Dazu müssen wir wissen, dass diese nicht von der Wahl der Basis abhängt. (Das ist im Allgemeinen falsch!)

Satz 3.6.8. *Sei M ein freier R-Modul mit endlicher Basis $B = \{m_1, \ldots, m_n\}$.*

1. *Jedes $m \in M$ hat eine eindeutige Darstellung*

$$m = a_1 m_1 + \cdots + a_n m_n.$$

2. *Sei N ein weiterer R-Modul und $f : B \to N$ eine beliebige Funktion. Dann gibt es genau einen R-Modul-Homomorphismus $L : M \to N$ mit $L(m_i) = f(m_i)$ für $1 \leq i \leq n$. Dieser ist gegeben durch*

$$L(a_1 m_1 + \cdots + a_n m_n) = a_1 f(m_1) + \cdots + a_n f(m_n).$$

Beweis. Zu *1.* Da B ein Erzeugendensystem ist, hat jedes $m \in M$ so eine Darstellung. Sei

$$a_1 m_1 + \cdots + a_n m_n = b_1 m_1 + \cdots + b_n m_n,$$

dann

$$(a_1 - b_1)m_1 + \cdots + (a_n - b_n)m_n = 0.$$

Da B linear unabhängig ist, folgt $a_i = b_i$ für $1 \leq i \leq n$.

Zu *2.* Das angegebene L ist eine wohldefinierte Funktion $M \to N$, da jedes $m \in M$ wegen *1.* eine eindeutige Darstellung $m = a_1 m_1 + \cdots + a_n m_n$ hat. Offensichtlich ist L R-linear. □

Korollar 3.6.9. *Sei M ein endlich erzeugter R-Modul.*

1. *Es gibt $n \in N$ und einen surjektiven R-Modul-Homomorphismus $R^n \to M$.*

2. *Sei M frei mit einer Basis B, sodass $|B| = n < \infty$. Dann gibt es einen R-Modul-Isomorphismus $L : R^n \to M$.*

Beweis. Zu *1.* Sei $M = \langle S \rangle$, mit $S = \{s_1, \ldots, s_n\}$ endlich. Sei $E = \{e_1, \ldots, e_n\}$ die Standardbasis von R^n, dann gibt es nach Satz 3.6.8 einen eindeutigen R-Modul-Homomorphismus $L : R^n \to M$ mit $L(e_i) = s_i$. Dieser ist surjektiv: sei $m \in M$, dann gibt es (nicht ungedingt eindeutige) $a_1, \ldots, a_n \in R$, sodass

$$m = a_1 s_1 + \cdots + a_n s_n = L(a_1 e_1 + \cdots + a_n e_n). \tag{3.2}$$

Zu *2.* Sei $S = B$ in *1.* Dann ist die Darstellung (3.2) eindeutig, also ist L auch injektiv. □

Um den Rang eines R-Moduls mit endlicher Basis definieren zu können, muß also $R^n \cong R^m \Rightarrow n = m$ gelten.

Lemma 3.6.10. *Sei M ein R-Modul und I ein Ideal von R. Sei*

$$IM := \{\sum_{k=1}^{n} i_k m_k \mid n \in \mathbb{N},\ i_k \in I,\ m_k \in M\}$$

1. *IM ist ein R-Untermodul von M.*
2. *Der Faktormodul M/IM ist auch ein R/I-Modul, mit der Skalarmultiplikation*
$$(a+I)\cdot(m+IM) = am + IM.$$
3. *Sei $L : M \to N$ ein R-Modul-Homomorphismus. Dann induziert L einen R/I-Modul-Homomorphismus $\tilde{L} : M/IM \to N/IN$, sodass*
$$\tilde{L}(m+IM) = L(m) + IN.$$
4. *Wenn L ein R-Modul-Isomorphismus ist, dann ist $\tilde{L}$ ein R/I-Modul-Isomorphismus.*
5. *Für $M = R^n$ gilt $IM = I^n$ und es gibt einen R-Modul-Isomorphismus*
$$L : R^n/I^n \to (R/I)^n \text{ mit } (r_1,\dots,r_n) + I^n = (r_1+I,\dots,r_n+I).$$
Dieses L ist auch ein R/I-Modul-Isomorphismus.

Beweis. Zu *1.* Es gilt $0 \in IM$, die Differenz zweier Elemente von IM ist in IM, und $r \cdot \sum_{k=1}^{n} i_k m_k = \sum_{k=1}^{n} (r i_k) m_k \in IM$.

Zu *2.* Die Skalarmultiplikation ist wohldefiniert: sei $a+I = b+I$, $m+IM = n+IM$. Dann $m-n \in IM$, $b-a \in I$, also

$$am - bn = a(m-n) - (b-a)n \in IM.$$

Daher folgt $am + IM = bn + IM$. Die Axiome übertragen sich direkt von M, z.B.

$$\begin{aligned}(a+I)((b+I)(m+IM)) &= (a+I)(bm+IM) = a(bm)+IM = (ab)m+IM \\ &= (ab+I)(m+IM) = ((a+I)(b+I))(m+IM).\end{aligned}$$

Zu *3.* $\tilde{L}$ ist wohldefiniert: sei $m+IM = n+IM$. Dann $m-n \in IM$. Sei $m-n = \sum_{k=1}^{n} i_k m_k$. Dann

$$L(m) - L(n) = L(m-n) = L(\sum_{k=1}^{n} i_k m_k) = \sum_{k=1}^{n} i_k L(m_k) \in IN.$$

$\tilde{L}$ ist ein R/I-Modul-Homomorphismus, da L ein R-Modul-Homomorphismus ist. Z.B.

$$\begin{aligned}\tilde{L}((a+I)(m+IM)) &= \tilde{L}(am+IM) = L(am)+IN = aL(m)+IN \\ &= (a+I)(L(m)+IN) = (a+I)\tilde{L}(m+IM).\end{aligned}$$

Zu *4.* $\tilde{L}$ ist injektiv: sei $\tilde{L}(m+IM)=0$, dann $L(m)\in IN$, also $L(m)=\sum_{k=1}^{n} i_k n_k$. Für $i\le k\le n$, sei $m_k\in M$ mit $L(m_k)=n_k$. Dann

$$m = L^{-1}(L(m)) = \sum_{k=1}^{n} i_k L^{-1}(L(m_k)) = \sum_{k=1}^{n} i_k m_k \in IM,$$

also $m+IM=0$.

$\tilde{L}$ ist surjektiv: sei $n+IN\in N/IN$ und $n=L(m)$. Dann $n+IN=\tilde{L}(m+IM)$.

Zu *5.* Für $i\in I$ und $m=(r_1,\dots,r_n)\in R^n$ gilt $im=(ir_n,\cdots,ir_n)\in I^n$, also $IM\subset I^n$. Umgekehrt sei $i=(i_1,\dots,i_n)\in I^n$. Dann $i=i_1e_1+\cdots+i_ne_n\in IM$, wobei $\{e_1,\dots,e_n\}$ die Standardbasis von $M=R^n$ ist. Daher folgt $IM=I^n$.

Betrachte die Abbildung

$$L_1 : R^n \to (R/I)^n,\ (r_1,\dots,r_n)\mapsto (r_1+I,\dots,r_n+I).$$

L_1 ist offensichtlich ein surjektiver R-Modul-Homomorphismus mit $\ker L_1 = I^n$. Laut Homomorphiesatz gibt es einen R-Modul-Isomorphismus $L : R^n/I^n \to (R/I)^n$ mit $L((r_1,\dots,r_n)+I^n)=(r_1+I,\dots,r_n+I)$. Dieser ist auch R/I-linear:

$$\begin{aligned}L((a+I)((r_1,\dots,r_n)+I^n)) &= L(a(r_1,\dots,r_n)+I^n) \\ &= L((ar_1,\dots,ar_n)+I^n) \\ &= (ar_1+I,\dots,ar_n+I) \\ &= ((a+I)(r_1+I),\dots,(a+I)(r_n+I)) \\ &= (a+I)(r_1+I,\dots,r_n+I) \\ &= (a+I)L((r_1,\dots,r_n)+I^n).\end{aligned}$$

□

Bemerkung 3.6.11. *Der Beweis der Injektivität in* 4. *hat verwendet dass L ein Isomorphismus (also auch surjektiv) ist. Es gilt im allgemeinen nicht, dass L injektiv $\Rightarrow$ $\tilde{L}$ injektiv! Betrachte z.B. $M=N=\mathbb{Z}$ als $\mathbb{Z}$-Moduln, $L:M\to N$, $L(m)=3\cdot m$, und $I=3\mathbb{Z}$. Dann ist L injektiv, aber $\tilde{L}=0$.*

Satz 3.6.12. *Sei $R \neq \{0\}$ ein kommutativer Ring mit Eins und $m, n \in \mathbb{N}$. Falls $R^m \cong R^n$ als R-Moduln, dann ist $m = n$.*

Beweis. Da $R \neq \{0\}$ ein Ring mit Eins ist, hat R ein maximales Ideal I. Nach Lemma 3.6.10 erhalten wir R/I-Modul-Isomorphismen

$$(R/I)^m \cong R^m/IR^m \cong R^n/IR^n \cong (R/I)^n.$$

Sei $K = R/I$. Da I maximal ist, ist K ein Körper, und $K^m \cong K^n$ als K-Vektorräume. Daraus folgt $m = n$. □

Korollar 3.6.13. *Sei $R \neq \{0\}$ ein kommutativer Ring mit Eins und M ein R-Modul. Seien B_1, B_2 zwei Basen von M mit $|B_1| = m$, $|B_2| = n$, für $m, n \in \mathbb{N}$. Dann gilt $m = n$.*

Beweis. Nach Korollar 3.6.9 gibt es R-Modul-Isomorphismen $R^m \to M$ und $R^n \to M$, also $R^n \cong R^m$. Nach Korollar 3.6.12 folgt $m = n$. □

Definition 3.6.14. *Sei R ein kommutativer Ring mit Eins und $n \in \mathbb{N}$. Ein R-Modul M heißt* frei vom Rang n, *wenn M eine Basis B mit $|B| = n$ hat. Per Definition ist der Nullmodul $M = \{0\}$ frei vom Rang* 0.

Bemerkung 3.6.15. *M ist genau dann frei vom Rang n, wenn $M \cong R^n$.*

Definition 3.6.16. *Sei M ein R-Modul und $N_1, \ldots, N_k$ Untermoduln von M. Der Modul M ist die* direkte Summe *von $N_1, \ldots, N_k$, geschrieben*

$$M = N_1 \oplus \cdots \oplus N_k \text{ oder } M = \bigoplus_{i=1}^{k} N_i,$$

wenn der R-Modul-Homomorphismus

$$N_1 \times \cdots \times N_k \to M,\ (n_1, \ldots, n_k) \mapsto n_1 + \cdots + n_k$$

ein Isomorphismus ist.

Bemerkung 3.6.17. *Für $k = 2$ gilt: Sei*

$$N_1 + N_2 := \{m_1 + m_2 \mid m_1 \in N_1,\ m_2 \in N_2\} = \langle N_1 \cup N_2 \rangle$$

(Die letzte Gleichheit wird in der Übung gezeigt.) Dann gilt

$$M = N_1 \oplus N_2 \iff M = N_1 + N_2 \text{ und } N_1 \cap N_2 = \{0\}.$$

3.7 Moduln über Hauptidealbereichen

Definition 3.7.1. *Sei R ein Integritätsbereich und M ein R-Modul.*

1. *Ein Element $m \in M$ heißt* Torsionselement*, wenn es $r \in R$, $r \neq 0$ gibt, sodass $r \cdot m = 0$.*

2. *Die Menge*

$$T(M) := \{m \in M \mid m \text{ is Torsionselement}\}$$

   heißt der Torsionsuntermodul *oder die* Torsion *von M.*

3. *M heißt* Torsionsmodul*, wenn $T(M) = M$.*

4. *M heißt* torsionsfrei*, wenn $T(M) = \{0\}$.*

Bemerkung 3.7.2. *$T(M)$ ist ein R-Untermodul von M: offensichtlich gilt $0 \in T(M)$. Sei $r_1 m_1 = 0$, $r_2 m_2 = 0$, mit $r_1, r_2 \neq 0$. Dann $r_1 r_2 \neq 0$, da R ein Integritätsbereich ist, und*

$$r_1 r_2 (m_1 - m_2) = r_2(r_1 m_1) - r_1(r_2 m_2) = r_2 0 - r_1 0 = 0.$$

Für $r \in R$ gilt weiters $r_1(rm) = r(r_1 m) = r0 = 0$.

Beispiel 3.7.3.

1. *Sei V ein K-Vektorraum. Dann $T(V) = \{0\}$, d.h., alle Vektorräume sind torsionsfrei.*

2. *Allgemeiner: jeder freie R-Modul ist torsionsfrei. Beweis: sei B eine Basis von M und $m = \sum_{i=1}^{m} r_i b_i \in M$, mit $r_i \in R$, $b_i \in B$. Sei $r \in R$, $r \neq 0$, mit $rm = 0$. Dann*

$$0 = rm = \sum_{i=1}^{m} (r r_i) b_i.$$

   Da B eine Basis ist, folgt $r r_i = 0$ für $1 \leq i \leq n$. Da $r \neq 0$ und R ein Integritätsbereich ist, folgt $r_i = 0$ für $1 \leq i \leq n$, also $m = 0$.

3. *Sei $n \in \mathbb{N}$. Der $\mathbb{Z}$-Modul $\mathbb{Z}/n\mathbb{Z}$ ist ein Torsionsmodul: für $m \in \mathbb{Z}/n\mathbb{Z}$ gilt $nm = 0$, also $T(\mathbb{Z}/n\mathbb{Z}) = \mathbb{Z}/n\mathbb{Z}$.*

4. *Der $\mathbb{Z}/n\mathbb{Z}$-Modul $\mathbb{Z}/n\mathbb{Z}$ ist torsionsfrei (da frei).*

5. *Für den $\mathbb{Z}$-Modul $M = \mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ gilt $T(M) = \{0\} \times \mathbb{Z}/n\mathbb{Z} \subset M$. Daher ist M weder torsionsfrei, noch ein Torsionsmodul.*

Wir werden sehen, dass jeder endlich erzeugte Modul über einem Hauptidealbereich direkte Summe von $T(M)$ und einem freien Untermodul ist.

Lemma 3.7.4. *Sei M ein R-Modul und $m_1, \dots, m_n \in M \setminus \{0\}$.*

1. *Für $1 \leq i \leq n$ ist*
$$Rm_i := \{rm_i \mid r \in R\}$$
ein Untermodul von M.

2. *Sei M torsionsfrei. Dann gilt*
$$M = \bigoplus_{i=1}^{n} Rm_i \Leftrightarrow \{m_1, \dots, m_n\} \text{ ist eine Basis von } M$$

Beweis. 1. ist offensichtlich. Zu *2.*: sei $\{m_1, \dots, m_n\}$ eine Basis. Dann ist

$$L : Rm_1 \times \cdots \times Rm_n \to M,\ (r_1m_1, \dots, r_nm_n) \mapsto r_1m_1 + \cdots + r_nm_n,$$

surjektiv. Weiters gilt $r_1m_1 + \cdots + r_nm_n = 0$ genau dann, wenn $r_1 = \cdots = r_n = 0$, also ist L auch injektiv.

Sei umgekehrt L ein Isomorphismus. Da L surjektiv ist, folgt $M = \langle m_1, \dots, m_n \rangle$. Sei $r_1m_1 + \cdots + r_nm_n = 0$. Dann gilt $(r_1m_1, \dots, r_nm_n) \in \ker L$, also $(r_1m_1, \dots, r_nm_n) = (0, \dots, 0)$. Aus $r_im_i = 0$ folgt aber $r_i = 0$, da m_i kein Torsionselement ist. Daher sind $\{m_1, \dots, m_n\}$ linear unabhängig. □

Lemma 3.7.5. *Sei $L : M \to F$ ein surjektiver R-Modul-Homomorphismus und F frei mit Basis $\{n_1, \dots, n_k\}$. Dann gibt es einen freien Untermodul $F_1 \subset M$, sodass*
$$F \cong F_1 \text{ und } M = \ker L \oplus F_1.$$

Beweis. Seien $m_1, \dots, m_k \in M$ mit $L(m_i) = n_i$ für alle $1 \leq i \leq k$. Setze $F_1 := \langle m_1, \dots, m_k \rangle$. Falls $a_1m_1 + \cdots + a_km_k = 0$, dann folgt $a_1n_1 + \cdots + a_kn_k = L(0) = 0$, also $a_1 = \cdots = a_k = 0$. Daher ist F_1 frei mit Basis $\{m_1, \dots, m_k\}$, und $F_1 \cong F$.

Sei $m \in M$, dann gibt es $b_1, \dots, b_k \in R$ mit

$$L(m) = \sum_{i=1}^{k} b_in_i = \sum_{i=1}^{k} b_iL(m_i) = L(\sum_{i=1}^{k} b_im_i).$$

Daher $m - \sum_{i=1}^k b_i m_i \in \ker L$, und $m = (m - \sum_{i=1}^k b_i m_i) + \sum_{i=1}^k b_i m_i \in \ker L + F_1$. Es folgt $M = \ker L + F_1$.

Sei $m = \sum_{i=1}^k r_i m_i \in F_1 \cap \ker L$. Dann

$$0 = L(m) = \sum_{i=1}^k r_i n_i,$$

also $r_i = 0$ für alle $1 \leq i \leq k$, und daher $m = 0$. Wir haben gezeigt, dass $M = \ker L \oplus F_1$. □

Satz 3.7.6. *Sei R ein Hauptidealbereich und M ein freier R-Modul vom Rang $n \in \mathbb{N}$. Sei N ein Untermodul von M. Dann ist auch N frei, vom Rang $\leq n$.*

Beweis. Für $N = \{0\}$ ist die Aussge klar, also nehmen wir an, dass $N \neq \{0\}$. Induktion nach n.

Für $n = 1$, sei $\{m_1\}$ eine Basis von M. Dann gilt $M = Rm_1$. Sei

$$I := \{r \in R \mid rm_1 \in N\}.$$

Dann ist $I \neq \{0\}$ ein Ideal von R, also $I = Rx$ für $x \in R \setminus \{0\}$. Es folgt $N = R(xm_1)$. Da $rxm_1 = 0 \Rightarrow rx = 0 \Rightarrow r = 0$, ist $\{xm_1\}$ eine Basis von N.

Sei jetzt $n \geq 2$ und $\{m_1, \ldots, m_n\}$ eine Basis von M. Setze $F := \langle m_2, \ldots, m_n \rangle$ und

$$L : M \to F, \ \sum_{i=1}^n a_i m_i \mapsto \sum_{i=2}^n a_i m_i.$$

Dann ist F frei mit Basis $\{m_2, \ldots, m_n\}$, und L ist ein surektiver R-Modul-Homomorphismus mit $\ker L = Rm_1$ frei vom Rang 1. Weiters ist $L(N)$ ein Untermodul von F, also nach Induktionsvoraussetzung frei vom Rang $\leq n-1$. Wir wenden Lemma 3.7.5 auf den surjektiven Homomorphismus $L|_N : N \to L(N)$ an. Da $\ker(L|_N) = \ker L \cap N$, gibt es einen freien Untermodul N_1 von N vom Rang $\leq n - 1$, sodass

$$N = (\ker L \cap N) \oplus N_1.$$

Da $\ker L \cap N \subset \ker L = Rm_1$, liefert die Induktionsvoraussetzung im Fall $n = 1$, dass $\ker L \cap N$ frei vom Rang ≤ 1 ist.

Sei B_1 eine Basis von $\ker L \cap N$ und B_2 eine Basis von N_1. Dann ist $B := B_1 \cup B_2$ eine Basis von N, und $|B| \leq 1 + (n-1) = n$. □

Bemerkung 3.7.7. *Sei R ein Integritätsbereich und I ein Ideal von R. Dann ist I ein Untermodul des freien R-Moduls R. I ist genau dann ein freier R-Modul, wenn I ein Hauptideal ist. Daher ist die Aussage von Satz 3.7.6 falsch, wenn R kein Hauptidealbereich ist.*

Satz 3.7.8. *Sei R ein Hauptidealbereich und M ein endlich erzeugter R-Modul. Dann gilt*

$$M \text{ ist torsionsfrei} \Rightarrow M \text{ ist frei von endlichem Rang.}$$

Beweis. Sei $M = \langle m_1, \dots, m_n \rangle$. Da R torsionsfrei ist, ist $\langle m_i \rangle$ frei für alle $1 \leq i \leq n$. Sei $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ eine maximale Teilmenge, sodass $\langle m_{i_1}, \dots, m_{i_k} \rangle$ frei ist. Ohne Einschränkung der Allgemeinheit gelte $\{i_1, \dots, i_k\} = \{1, \dots, k\}$, sonst ordnen wir die m_i um. Also ist $F := \langle m_1, \dots, m_k \rangle$ ein freier Untermodul von M, und für $j > k$ sind $m_1, \dots, m_k, m_j$ linear abhängig. Das heißt,

$$\sum_{i=1}^{k} r_{ij} m_i - r_j m_j = 0,$$

mit $r_{ij}, r_j \in R$, $r_j \neq 0$. Insbesondere, $r_j m_j \in F$. Sei $r := r_{k+1} \cdots r_n \in R \setminus \{0\}$, dann $r m_j \in F$ für $1 \leq j \leq n$, und da $M = \langle m_1, \dots, m_n \rangle$ folgt $rM \subset F$. Als Untermodul des freien Moduls F vom Rang k ist rM frei vom Rang $\leq k$.

Weiters definiert $L(m) := rm$ einen surjektiven Homomorphismus $L : M \to rM$. Nach Lemma 3.7.5, ist $M = \ker L \oplus F_1$, für einen freien R-Modul $F_1 \cong rM$. Jedoch gilt

$$\ker L = \{m \in M \mid rm = 0\} \subset T(M) = \{0\},$$

also $\ker L = \{0\}$ und $M = F_1$ ist frei vom Rang $\leq k$. □

Bemerkung 3.7.9.

1. *Sei M endlich erzeugt und frei. Dann ist M offensichtlich torsionsfrei, also laut Satz frei von endlichem Rang.*

2. *Sei R ein Integritätsbereich und $I \subset R$ ein Ideal, das kein Hauptideal ist. Dann ist der R-Modul I torsionsfrei, aber nicht frei.*

3. *Der $\mathbb{Z}$-Modul $\mathbb{Q}$ ist torsionsfrei, aber nicht frei. Die Bedingung, dass M endlich erzeugt ist, ist also notwendig in Satz 3.7.8.*

Satz 3.7.10. *Sei R ein Hauptidealbereich und M ein endlich erzeugter R-Modul. Dann gibt es einen freien Untermodul F von M von endlichem Rang, sodass*

$$M = T(M) \oplus F.$$

Beweis. Wir betrachten den R-Modul $M/T(M)$. Sei $M = \langle m_1, \ldots, m_n \rangle$, dann ist auch

$$M/T(M) = \langle m_1 + T(M), \ldots, m_n + T(M) \rangle$$

endlich erzeugt.

Weiters ist $M/T(M)$ torsionsfrei: sei $m + T(M)$ ein Torsionselement von $M/T(M)$. Dann gibt es $r \in R$, $r \neq 0$, sodass

$$0 = r(m + T(M)) = rm + T(M),$$

also $rm + T(M) = T(M)$, und $rm \in T(M)$. Das heißt, $srm = 0$ für ein $s \in R \setminus \{0\}$, aber dann $sr \neq 0$, und $m \in T(M)$. Es folgt $m + T(M) = 0$, also $T(M/T(M)) = \{0\}$.

Laut Satz 3.7.8 ist $M/T(M)$ also frei von endlichem Rang. Sei $\pi : M \to M/T(M)$ die natürliche Projektion. Dann ist π surjektiv und $\ker \pi = T(M)$. Nach Lemma 3.7.5 gibt es einen freien Untermodul F von M mit $F \cong M/T(M)$, sodass

$$M = T(M) \oplus F.$$

□

Wir wissen also, dass $M = T(M) \oplus F$, und $F \cong R^k$, für ein $k \in \mathbb{N}$. Um die Struktur von M vollständig zu verstehen, müssen wir noch $T(M)$ genauer beschreiben.

3.8 Matrixumformungen

Für einen freien R-Modul M und einen freien Untermodul N, beide von endlichem Rang, suchen wir Basen von N und M, bezüglich derer die Inklusion $N \subset M$ möglichst einfach dargestellt wird.

Definition 3.8.1. *Seien M, N freie R-Moduln mit Basen $B_M = \{m_1, \ldots, m_k\}$ und $B_N = \{n_1 \ldots, n_l\}$, $k, l \geq 1$. Sei $L : M \to N$ ein R-Modul-Homomorphismus und $a_{ij} \in R$ die eindeutigen Elemente, sodass*

$$L(m_j) = \sum_{i=1}^{l} a_{ij} n_i, \text{ für } 1 \leq j \leq k.$$

Die darstellende Matrix *von L bezüglich der Basen B_M, B_N ist die Matrix*

$$[L]_{B_M,B_N} := (a_{ij})_{\substack{1\le i\le l\\ 1\le j\le k}} \in M(l,k;R).$$

Bestimmte elementare Zeilen- und Spaltenoperationen an $[L]_{B_M,B_N}$ entsprechen Modifikationen der Basen B_M, B_N.

Lemma 3.8.2. *Sei R ein Integritätsbereich und $M, N \neq \{0\}$ zwei freie R-Moduln mit endlichen Basen B_M, B_N. Sei $L : M \to N$ ein R-Modul-Homomorphismus und $A = [L]_{B_M,B_N} \in M(l,k;R)$ die darstellende Matrix. Sei $A' \in M(l,k;R)$ die Matrix, die aus A durch eine der folgenden Operationen hervorgeht:*

1. *Vertauschen der i-ten Zeile mit der j-ten Zeile, für $1 \le i, j \le l$.*
2. *Vertauschen der i-ten Spalte mit der j-ten Spalte, für $1 \le i, j \le k$.*
3. *Addition des λ-fachen der j-ten Zeile zur i-ten Zeile, für $\lambda \in R$, $1 \le i \neq j \le l$*
4. *Addition des λ-fachen der j-ten Spalte zur i-ten Spalte, für $\lambda \in R$, $1 \le i \neq j \le k$*

Dann gibt es Basen B'_M, B'_N von M, N, sodass $A' = [L]_{B'_M,B'_N}$.

Beweis. Sei $B_M = \{m_1, \dots, m_k\}$, $B_N = \{n_1, \dots, n_l\}$. Wir zeigen in jedem der vier Fälle, wie die Basis B_N oder B_M zu modifizieren ist.

1. $B'_M := B_M$ und B'_N ist B_N mit n_i und n_j vertauscht.
2. B'_M ist B_M mit m_i und m_j vertauscht, und $B'_N := B_N$.
3. $B'_M := B_M$, und B'_N ist B_N mit n_j ersetzt durch $n_j - \lambda n_i$. Tatsächlich gilt

$$L(m_h) = \sum_{g=1}^{l} a_{gh} n_g = \sum_{g\neq i,j} a_{gh} n_g + (a_{ih} + \lambda a_{jh}) n_i + a_{jh}(n_j - \lambda n_i).$$

B'_N ist wieder eine Basis: es gilt $\langle B'_N \rangle = N$, da $n_j = (n_j - \lambda n_i) + \lambda n_i \in \langle B'_N \rangle$. Sei

$$0 = a_j(n_j - \lambda n_i) + a_i n_i + \sum_{h\neq i,j} a_h n_h = a_j n_j + (a_i - \lambda a_j) n_i + \sum_{h\neq i,j} a_h n_h,$$

dann $a_j, a_i - \lambda a_j = 0$ und $a_h = 0$ für $h \neq i, j$, da B_N linear unabhängig ist. Daher $a_h = 0$ für alle $1 \leq h \leq l$, und B'_N ist ebenfalls linear unabhängig.

4. B'_M ist B_M, mit m_i ersetzt durch $m_i + \lambda m_j$, und $B'_N := B_N$. Tatsächlich gilt

$$m_i + \lambda m_j = \sum_{g=1}^{l} a_{gi} n_g + \lambda \sum_{g=1}^{l} a_{gj} n_g = \sum_{g=1}^{l} (a_{gi} + \lambda a_{gj}) n_g.$$

Analog wie in 3. ist B'_M wieder eine Basis von M.

□

Bemerkung 3.8.3.

1. *Im Gegensatz zu Vektorräumen, erlauben wir* nicht *die Multiplikation einer Zeile/Spalte mit einem Skalar.*

2. *Wir nennen die Operationen* 1. *und* 3. elementare Zeilenoperationen *über R, und die Operationen* 2. *und* 4. elementare Spaltenoperationen *über R.*

Definition 3.8.4. *Sei R ein Integritätsbereich und $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in M(m, n; R)$. Sei $r = \min\{m, n\}$. Wir sagen, dass A in* Smith-Normalform *ist, wenn es $d_1, \ldots, d_r \in R$ gibt, sodass $d_1 \mid d_2 \mid \cdots \mid d_r$, und*

$$a_{ij} = \begin{cases} d_i, & \text{wenn } i = j \\ 0, & \text{sonst.} \end{cases}$$

Anschaulich, z.B. für $n > m$,

$$A = \begin{pmatrix} d_1 & & & 0 & \cdots & 0 \\ & d_2 & & 0 & \cdots & 0 \\ & & \ddots & \vdots & & \vdots \\ & & & d_r & 0 & \cdots & 0 \end{pmatrix}.$$

Bemerkung 3.8.5. *Wir schließen nicht aus, dass $d_i = 0$. In diesem Fall gilt auch $d_j = 0$ für alle $j \geq i$.*

Lemma 3.8.6. *Sei R ein Euklidischer Ring mit Rangfunktion $\phi : R \setminus \{0\} \to \mathbb{N}_0$. Sei $A \in M(m, n; R)$. Dann kann A durch elementare Zeilen- und Spaltenumformungen zu einer Matrix $A' = (a'_{ij})$ transformiert werden, sodass $a'_{11} \mid a'_{ij}$ für alle i, j.*

Beweis. Die Aussage gilt für die Nullmatrix, sei also $A \neq 0$. Durch Zeilen- und Spaltenvertauschungen erreichen wir, dass $a_{11} \neq 0$. Angenommen, a_{11} teilt nicht alle anderen Einträge von A.

Behauptung: Dann kann A durch elementare Zeilen- und Spaltenumformungen zu einer Matrix $B = (b_{ij})$ transformiert werden, sodass $b_{11} \neq 0$ und $\phi(b_{11}) < \phi(a_{11})$.

Da die absteigende Folge natürlicher Zahlen $\phi(a_{11}) > \phi(b_{11}) > \cdots$ nach endlich vielen Schritten abbricht, muß die dann erreichte Matrix A' die gewünschte Eigenschaft haben.

Wir müssen also nur noch die Behauptung beweisen. Dazu unterscheiden wir drei Fälle.

1. Fall: Es gibt $2 \leq j \leq n$, sodass $a_{11} \nmid a_{1j}$. Division mit Rest liefert $a_{1j} = qa_{11} + r$, mit $r \neq 0$ und $\phi(r) < \phi(a_{11})$. Wir ziehen das q-Fache der ersten Spalte von der j-ten Spalte ab und vertauschen danach die erste mit der j-ten Spalte. Für die dadurch erhaltene Matrix B gilt $b_{11} = r$, also $\phi(b_{11}) < \phi(a_{11})$.

2. Fall: Es gibt $2 \leq i \leq m$, sodass $a_{11} \nmid a_{i1}$. In diesem Fall gehen wir gleich vor, wie im ersten Fall, nur mit Zeilen statt Spalten.

3. Fall: a_{11} teilt alle Einträge der ersten Zeile und der ersten Spalte. Durch Abziehen geeigneter vielfacher der ersten Zeile/Spalte von allen anderen Zeilen/Spalten erreichen wir, dass alle anderen Einträge der ersten Zeile und Spalte gleich 0 sind. Die erhaltene Matrix B hat immer noch einen Eintrag, der nicht durch $b_{11} = a_{11}$ geteilt wird, da von allen Einträgen nur Vielfache von a_{11} abgezogen wurden. Gelte $b_{11} \nmid a_{ij}$ mit $i \neq 1$. Wir addieren die i-te Zeile zur ersten, um a_{ij} in die erste Zeile zu bringen. Das lässt $a_{11} = b_{11}$ unverändert, und wir können im 1. Fall fortfahren.

□

Satz 3.8.7 (Smith-Normalform)**.** *Sei R ein Hauptidealbereich und $A \in M(m, n; R)$. Dann kann A durch elementare Zeilen- und Spaltenumformungen in Smith-Normalform gebracht werden.*

Beweis. Wir führen den Beweis hier nur in dem Spezialfallfall, dass R ein euklidischer Ring ist. Der Beweis für allgemeine Hauptidealbereiche benötigt etwas mehr Vorbereitung (siehe z.B. [4, Satz 11.5.7]). Sei also R euklidisch.

Falls $\min\{m, n\} = 1$, besteht A nur aus einer Zeile oder Spalte. Mit Lemma 3.8.6 bringen wir A in eine Form A', sodass a'_{11} alle anderen Einträge teilt.

Durch Subtraktion geeigneter Vielfacher von a'_{11} von allen anderen Einträgen erreichen wir Smith-Normalform $\begin{pmatrix} a'_{11} & 0 & \cdots & 0 \end{pmatrix}$ oder $\begin{pmatrix} a'_{11} & 0 & \cdots & 0 \end{pmatrix}^t$.

Sei $\min\{m, n\} > 1$. Mit Lemma 3.8.6 bringen wir A in eine Form A', sodass a'_{11} alle anderen Einträge von A' teilt. Durch Abziehen geeigneter Vielfacher der ersten Zeile/Spalte von allen anderen Zeilen/Spalten, bringen wir A' in die Form

$$A'' = \begin{pmatrix} a'_{11} & 0 \\ 0 & B \end{pmatrix},$$

wobei $B \in M(m-1, n-1; R)$. Da von allen Einträgen nur Vielfache von a'_{11} abgezogen wurden, teilt a'_{11} immer noch alle Einträge von B.

Nach Induktionsvoraussetzung kann B durch elementare Zeilen- und Spaltenoperationen in Smith-Normalform gebracht werden. Diese Operationen ändern nichts an der Tatsache, dass a'_{11} alle Einträge von B teilt, also bringen sie auch A'' in Smith-Normalform. □

3.9 Elementarteiler und invariante Faktoren

Satz 3.9.1. *Sei R ein Hauptidealbereich und M ein freier R-Modul von endlichem Rang $k \in \mathbb{N}$. Sei N ein Untermodul von M. Dann gibt es eine Basis $\{m_1, \ldots, m_k\}$ von M, $0 \leq l \leq k$, und $d_1, \ldots, d_l \in R \setminus \{0\}$, mit $d_1 \mid d_2 \mid \cdots \mid d_l$, sodass $\{d_1 m_1, \ldots, d_l m_l\}$ eine Basis von N ist.*

Beweis. Nach Satz 3.7.6 ist N frei vom Rang $l \leq k$. Falls $l = 0$, ist die Aussage trivial, sei also $l > 0$. Seien B_M, B_N Basen von M, N, und $L : N \to M$ die Inklusion von N in M, d.h. $L(n) = n$. Wir bringen die darstellende Matrix $[L]_{B_N, B_M}$ durch elementare Zeilen- und Spaltenumformungen in Smith-Normalform $D = \operatorname{diag}(d_1, \ldots, d_l)$, mit $d_1 \mid d_2 \mid \cdots \mid d_l$. Die Zeilen- und Spaltenumformungen entsprechen nach Lemma 3.8.2 Modifikationen der Basen B_N, B_M. Daher gibt es Basen $B'_M = \{m_1, \ldots, m_k\}$ von M und $B'_N = \{n_1, \ldots, n_l\}$ von N, sodass $[L]_{B'_N, B'_M} = D$. Das heißt, $n_i = d_i m_i$ für $1 \leq i \leq l$. Insbesondere gilt also $d_i \neq 0$ für $1 \leq i \leq l$. □

Lemma 3.9.2. *Sei R ein Integritätsbereich, $n \in \mathbb{N}_0$, und $d_1, \ldots, d_n \in R$, sodass d_1 keine Einheit ist, und $d_1 \mid d_2 \mid \cdots \mid d_n$. Sei*

$$N = R/d_1R \times \cdots \times R/d_nR.$$

Dann ist

$$n = \min\{i \in \mathbb{N}_0 \mid N \text{ hat ein Erzeugendensystem aus } i \text{ Elementen}\}.$$

(Wir erlauben hier, dass $d_i = 0$. In dem Fall ist $R/d_iR = R$.)

Beweis. Tatsächlich hat N das Erzeugendensystem $\{e_1, \ldots, e_n\}$, mit

$$e_i = (0, \ldots, 0, 1 + d_i R, 0, \ldots, 0).$$

Angenommen, $N = \langle m_1, \ldots, m_{n-1} \rangle$. Da d_1 keine Einheit ist, gilt $d_1 R \subsetneq R$, also gibt es ein maximales Ideal I von R mit $d_1 R \subset I$. Da $d_1 \in I$, folgt auch $d_j \in I$ für alle $1 \leq j \leq n$, und daher ist

$$R/d_j R \to R/I, \quad r + d_j R \mapsto r + I$$

ein wohldefinierter surjektiver R-Modul-Homomorphismus. Wir erhalten einen surjektiven R-Modul-Homomorphismus

$$\begin{aligned} L : N = R/d_1 R \times \cdots \times R/d_n R &\to (R/I)^n, \\ (r_1 + d_1 R, \ldots, r_n + d_n R) &\mapsto (r_1 + I, \ldots, r_n + I). \end{aligned}$$

Die Bilder $L(m_1), \ldots, L(m_{n-1})$ erzeugen dann $(R/I)^n$ als R-Modul, und auch als R/I-Modul: sei $v \in (R/I)^n$, dann $v = L(m)$, für ein $m \in N$. Sei $m = a_1 m_1 + \ldots + a_{n-1} m_{n-1}$, dann

$$v = a_1 L(m_1) + \cdots + a_{n-1} L(m_{n-1}) = (a_1 + I) L(m_1) + \cdots + (a_{n-1} + I) L(m_{n-1}).$$

Da I maximal ist, ist $K = R/I$ ein Körper, und $L(m_1), \ldots, L(m_{n-1})$ erzeugen den K-Vektorraum K^n, ein Widerspruch. □

Lemma 3.9.3. *Sei R faktoriell und $d, r \in R$, $r \neq 0$. Fixiere einen $\mathrm{ggT}(d, r)$, und schreibe $d' := d/\mathrm{ggT}(d, r)$. Dann gibt es einen R-Modul-Isomorphismus*

$$L : R/d'R \to r \cdot R/dR, \quad s + d'R \mapsto rs + dR.$$

Beweis. Sei $r' := r/\mathrm{ggT}(d, r)$, dann folgt $\mathrm{ggT}(d', r') = 1$. Die Abbildung L ist wohldefiniert und injektiv, da

$$d \mid rs - rt \Leftrightarrow \mathrm{ggT}(d, r) d' \mid \mathrm{ggT}(d, r) r'(s - t) \Leftrightarrow d' \mid r'(s - t) \Leftrightarrow d' \mid (s - t).$$

Die letzte Äquivalenz gilt, da $\mathrm{ggT}(d', r') = 1$. Die Abbildung ist offensichtlich surjektiv und R-linear, also insgesamt ein Isomorphismus. □

Satz 3.9.4 (Klassifikationssatz für endlich erzeugte Moduln über Hauptidealbereichen)**.** *Sei R ein Hauptidealbereich und M ein endlich erzeugter R-Modul. Dann gibt es $r, l \in \mathbb{N}_0$ und $d_1, \ldots, d_l \in R \setminus \{0\}$, sodass d_1 keine Einheit ist und $d_1 \mid d_2 \mid \cdots \mid d_l$ mit*

$$M \cong R/d_1 R \times \cdots \times R/d_l R \times R^r.$$

Hierbei sind r, l eindeutig und $d_1, \ldots, d_l$ eindeutig bis auf Multiplikation mit Einheiten bestimmt.

Beweis. Existenz: sei $M = \langle m_1, \ldots, m_n \rangle$ und $L : R^n \to M$ der surjektive R-Modul-Homomorphismus mit $L(e_i) = m_i$, für $1 \leq i \leq n$. Sei $U := \ker L$. Dann ist U ein Untermodul von R^n und $R^n/U \cong M$.

Laut Satz 3.9.1 gibt es eine Basis $\{b_1, \ldots, b_n\}$ von R^n und $d_1 \mid \cdots \mid d_l \in R$, sodass $\{d_1 b_1, \ldots, d_l b_l\}$ eine Basis von U ist.

Betrachte den Homomorphismus

$$\begin{aligned} &L_2 : R^n \to R/d_1R \times \cdots \times R/d_lR \times R^{n-l}, \\ &\sum_{i=1}^{n} a_i b_i \mapsto (a_1 + d_1R, \cdots, a_l + d_lR, a_{l+1}, \ldots, a_n). \end{aligned}$$

Dann ist L_2 surjektiv und $\ker L_2 = \langle d_1 b_1, \ldots, d_l b_l \rangle = U$. Mit dem Homomorphiesatz folgt also

$$M \cong R^n/U \cong R/d_1R \times \cdots \times R/d_lR \times R^{n-l}.$$

Wenn d_1 eine Einheit ist, folgt $R/d_1R = \{0\}$, also kann der Faktor weggelassen werden.

Zur Eindeutigkeit: wegen Lemma 3.9.2 ist $n := l + r$ eindeutig als die minimale Anzahl an Elementen in einem Erzeugendensystem von N bestimmt. Setze $d_j = 0$ für $l + 1 \leq j \leq n$ (also $R/d_jR = R$). Dann ist das Ideal d_jR, für $1 \leq j \leq n$ eindeutig bestimmt als

$$d_jR = \{r \in R \mid rM \text{ hat ein Erzeugendensystem aus } \leq n - j \text{ Elementen}\}. \tag{3.3}$$

Sei $0 \neq r \in R$ mit $d_j \mid r$, also $d_j = \mathrm{ggT}(d_j, r)$. Da $d_i \mid d_j$ für $1 \leq i \leq j$, folgt auch $d_i = \mathrm{ggT}(d_i, r)$ für $1 \leq i \leq j$. Mit Lemma 3.9.3 folgt

$$r \cdot R/d_iR \cong R/R = \{0\},$$

für alle $1 \leq i \leq j$, und daher

$$rM \cong \{0\} \times \cdots \times \{0\} \times r \cdot R/d_{j+1}R \times \cdots \times r \cdot R/d_nR.$$

Der letzte dieser Moduln wird von $n - j$ Elementen $\{re_{j+1}, \ldots, re_n\}$ erzeugt. Gelte umgekehrt $d_j \nmid r$, und sei $j' \leq j$ minimal mit $d_{j'} \nmid r$. Für $1 \leq i \leq n$, schreibe $d_i = \mathrm{ggT}(d_i, r) d_i'$. Dann gilt

$$d_i' \text{ ist Einheit in } R \iff d_i \mid r \iff i < j',$$

also

$$r \cdot R/d_iR \cong R/d_i'R \begin{cases} \cong \{0\} & \text{für } 1 \leq i < j' \\ \ncong \{0\} & \text{für } j' \leq i \leq n. \end{cases}$$

Insbesondere folgt

$$\begin{aligned} rM &\cong r \cdot R/d_1R \times \cdots r \cdot R/d_nR \cong R/d_1'R \times \cdots \times R/d_n'R \\ &\cong R/d_{j'}'R \times \cdots \times R/d_n'R. \end{aligned}$$

Laut Lemma 3.9.2 hat der letzte R-Modul kein Erzeugendensystem mit weniger als $n - j' + 1 \geq n - j + 1$ Elementen.

Da d_jR eindeutig durch (3.3) bestimmt ist, ist d_j bis auf Multiplikation mit Einheiten eindeutig bestimmt. □

Definition 3.9.5. *Die bis auf Multiplikation mit Einheiten eindeutig bestimmten Elemente $d_1, \ldots, d_l$ aus Satz 3.9.4 heißen die* Elementarteiler *von M. (Manche Autoren nennen diese Elemente auch die* invarianten Faktoren *von M.)*

Bemerkung 3.9.6.

1. *Für M wie im Satz 3.9.4 gilt $T(M) \cong R/d_1R \times \cdots \times R/d_lR$.*
2. *M ist torsionsfrei $\Leftrightarrow l = 0$*
3. *M ist Torsionsmodul $\Leftrightarrow r = 0$*

Korollar 3.9.7 (Klassifikationssatz für endlich erzeugte abelsche Gruppen)**.** *Sei G eine endlich erzeugte abelsche Gruppe. Dann gibt es eindeutig bestimmte $r, l \in \mathbb{N}_0$ und $d_1, \ldots, d_l \in \mathbb{N}$, mit $d_1 > 1$ und $d_1 \mid d_2 \mid \cdots \mid d_l$, sodass*

$$G \cong \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_l\mathbb{Z} \times \mathbb{Z}^r.$$

Die Gruppe G ist genau dann endlich, wenn $r = 0$. In diesem Fall gilt $|G| = d_1 \cdots d_l$.

Beispiel 3.9.8. *Jede Abelsche Gruppe mit 12 Elementen ist isomorph zu*

$$\mathbb{Z}/12\mathbb{Z} \text{ oder } \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}.$$

In der Tat sind (12) und $(2, 6)$ die einzigen Möglichkeiten für $(d_1, \ldots, d_l)$ mit $d_1 > 1$, $d_1 \mid \cdots \mid d_l$ und $d_1 \cdots d_l = 12$.

Bemerkung 3.9.9. *Sei R ein Hauptidealbereich.*

1. *R-Moduln mit einem einelementigen Erzeugendensystem werden* zyklische *R-Moduln genannt. Sie sind isomorph zu R/d_iR für ein $d_i \in R$.*

2. *Die Darstellung von M im Satz 3.9.4 stellt M als Produkt von möglichst wenigen zyklischen R-Moduln dar.*

3. *Sei $d = p_1^{e_1} \cdots p_n^{e_n}$ eine Zerlegung von d in irreduzible Elemente, wobei $p_i R \neq p_j R$ für $i \neq j$ (das heißt, p_i und p_j unterscheiden sich nicht nur durch Multiplikation mit einer Einheit). Dann gilt $p_i^{e_i} R + p_j^{e_j} R = R$, denn wenn $p_i^{e_i} R + p_j^{e_j} R = rR$, dann folgt $r \mid p_i^{e_i}$ und $r \mid p_j^{e_j}$, also ist r wegen der Eindeutigkeit der Faktorisierung in R eine Einheit, und $rR = R$. Nach dem Chinesischen Restsatz gibt es einen Isomorphismus von Ringen*

$$\begin{aligned} R/dR &\to R/p_1^{e_1} R \times \cdots \times R/p_n^{e_n} R \\ x + dR &\mapsto (x + p_1^{e_1} R, \ldots, x + p_n^{e_n} R). \end{aligned}$$

Dieser ist auch ein Isomorphismus von R-Moduln.

Korollar 3.9.10. *Sei R ein Hauptidealbereich und M ein endlich erzeugter R-Torsionsmodul. Sei $\mathcal{P}$ ein Repräsentantensystem der Äquivalenzklassen irreduzibler Elemente in R. Dann gibt es (bis auf Reihenfolge) eindeutige $p_1, \ldots, p_n \in \mathcal{P}$ und $e_{ij} \in \mathbb{N}$, sodass*

$$M \cong \prod_{i=1}^{n} \prod_{j=1}^{l_i} R/p_i^{e_{ij}} R.$$

Beweis. Seien $d_1, \ldots, d_l$ die Elementarteiler von M. Dann gilt

$$M \cong R/d_1 R \times \cdots \times R/d_l R.$$

Durch Multiplikation mit Einheiten kann man erreichen, dass jedes d_i ein Produkt von Elementen aus $\mathcal{P}$ ist. Wende nun Punkt 3. der Bemerkung auf alle Elementarteiler $d_1, \ldots, d_n$ an. □

Definition 3.9.11. *Die Potenzen $p_i^{e_{ij}}$ von irreduziblen Elementen in Korollar 3.9.10 heißen die* invarianten Faktoren *von M.*

Im Fall $R = \mathbb{Z}$ wählen wir immer

$$\mathcal{P} = \{p \mid p \text{ Primzahl}\},$$

die invarianten Faktoren sind also Primzahlpotenzen.

Beispiel 3.9.12. *Wir berechnen die Elementarteiler und invarianten Faktoren des $\mathbb{Z}$-Moduls $M = \mathbb{Z}/16\mathbb{Z} \times \mathbb{Z}/63\mathbb{Z} \times \mathbb{Z}/14\mathbb{Z}$.*

Da $63 = 9 \cdot 7$ *und* $14 = 2 \cdot 7$*, folgt aus dem Chinesischen Restsatz, dass*

$$\begin{aligned} M &\cong \mathbb{Z}/16\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \\ &\cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/16\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}. \end{aligned}$$

Die invarianten Faktoren sind also die Primzahlpotenzen $2, 16, 9, 7, 7$*. Diese lassen sich eindeutig zu den Elementarteilern* $d_1, \ldots, d_l$ *mit* $d_1 \mid \cdots \mid d_l$ *kombinieren: die höchsten Potenzen aller vorkommenden Primzahlen ergeben* d_l*, die zweithöchsten* d_{l-1}*, und so weiter. Wir erhalten*

$$d_2 = 16 \cdot 9 \cdot 7 = 1008 \text{ und } d_1 = 2 \cdot 7 = 14,$$

also sind $14, 1008$ *die Elementarteiler, und*

$$M \cong \mathbb{Z}/14\mathbb{Z} \times \mathbb{Z}/1008\mathbb{Z}.$$

Bemerkung 3.9.13. *Die Zerlegung in invariante Faktoren* $M \cong \prod_{i,j} R/p_j^{e_{ij}} R$ *stellt* R *als Produkt von möglichst vielen zyklischen* R*-Moduln dar.*

Literaturverzeichnis

[1] S. Bosch. *Lineare Algebra*. Springer-Lehrbuch. Springer, 2003.

[2] K. Conrad. Bilinear forms.

[3] G. Fischer. *Lineare Algebra: Eine Einführung für Studienanfänger*. vieweg studium; Grundkurs Mathematik. Vieweg+Teubner Verlag, 2013.

[4] G. Michler and H.J. Kowalsky. *Lineare Algebra*. De Gruyter Lehrbuch. De Gruyter, 2003.