Skript zur Vorlesung Lineare Algebra II für Informatik am KIT im Sommersemester 2024

Rafael Dahmen

8. August 2024

Inhaltsverzeichnis

1. Skalarprodukte **7**
1.1. Vektoren im $\mathbb{R}^3$ . . . 7
1.2. Skalarprodukte, Längen und Winkel . . . 9
1.3. Orthonormalbasen . . . 19
1.4. Orthogonale Komplemente und Projektionen . . . 25

2. Isometrien und Homomorphismen **29**
2.1. Isometrien und isometrische Isomorphismen . . . 29
2.2. Die orthogonale und die unitäre Gruppe . . . 33
2.3. Unitäre Trigonalisierung und Diagonalisierung . . . 38
2.4. Selbstadjungierte komplexe Matrizen und symmetrische reelle Matrizen . . . 40
2.5. Normale Matrizen . . . 42
2.6. Adjungierte Abbildungen und die Isometrie-Normalform . . . 44

3. Bilinearformen und quadratische Formen **57**
3.1. Der Dualraum . . . 57
3.2. Bilinearformen über beliebigen Körpern . . . 66
3.3. Symmetrische Bilinearformen . . . 77
3.4. Bilinearformen und quadratische Formen über den reellen Zahlen . . . 80
3.5. Quadriken und Hauptachsentransformation . . . 92

A. Anhang **103**
A.1. Das griechische Alphabet . . . 103
Stichwortverzeichnis . . . 106

Dieses Skript ist – nach einigen kleineren Änderungen – während der Vorlesung der Linearen Algebra 2 im Sommersemester 2021 enstanden.

Wenn Sie Fehler finden, so bitte ich Sie, mich darauf hinzuweisen.

Im Skript fehlen bei den allermeisten Sätzen noch die Beweise. Einige davon werde ich im Laufe des Semesters nachtragen. Normalerweise werden die Beweise aber in den Vorlesungen angegeben und lassen sich daher in den Aufzeichnungen finden.

1. Skalarprodukte

1.1. Vektoren im $\mathbb{R}^3$

Wir beginnen mit dem Vektorraum $\mathbb{R}^3$ über dem Grundkörper $\mathbb{K} = \mathbb{R}$.

Die Vektorraumstruktur auf $\mathbb{R}^3$ gibt uns zwei Abbildungen:

- die Vektoraddition $+ : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$, die zwei Vektoren ihre Summe zuordnet
- und die skalare Multiplikation $\cdot : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, die einen Vektor mit einem Skalar multipliziert.

Nur unter Verwendung dieser beiden Abbildungen haben wir Untervektorräume, affine Unterräume, lineare Abbildungen und weiteres definiert und untersucht.

Es gibt aber im Raum $\mathbb{R}^3$ noch viel mehr Struktur, die wir uns bis jetzt ignoriert haben. So gibt es z.B. die *Abstandsfunktion*

$$d : \mathbb{R}^3 \times \mathbb{R}^3 \to [0, +\infty), \quad (x, y) \mapsto d(x, y),$$

die je zwei Elementen in $\mathbb{R}^3$ ihren Abstand zuweist. Ebenso gibt es die *Normabbildung*

$$\|\cdot\| : \mathbb{R}^3 \to [0, +\infty), \quad v \mapsto \|v\|,$$

die jedem Vektor seine Länge zuordnet. Die letzteren beiden Strukturen sind verbunden über die Formeln:

$$d(x, y) = \|x - y\| \quad \text{und} \quad \|v\| = d(v, 0).$$

Man beachte, wie wir hier immer hin- und herwechseln zwischen der Interpretation eines Elements in $\mathbb{R}^3$ als Punkt im Raum oder als Vektorpfeil.

Des Weiteren kann man *Winkel zwischen Vektoren* messen:

$$\sphericalangle : (\mathbb{R}^3 \setminus \{0\}) \times (\mathbb{R}^3 \setminus \{0\}) \to [0, \pi], \quad (v, w) \mapsto \sphericalangle(v, w).$$

Der von zwei Vektoren $v, w \neq 0$ eingeschlossene Winkel $\sphericalangle(v, w)$ liegt immer zwischen $0 = 0°$ und $\pi = 180°$.

Wenn der Winkel zwischen zwei Vektoren genau $\pi/2 = 90°$ beträgt, sind die beiden Vektoren *orthogonal*:

$$\forall v, w \in \mathbb{R}^3 \setminus \{0\} : v \perp w :\Longleftrightarrow \sphericalangle(v, w) = \pi/2.$$

Während in einem allgemeinen Vektorraum kein Produkt zwischen Vektoren definiert ist, gibt es im Raum $\mathbb{R}^3$ gleich zwei verschiedene solche Produkte, die beide eine wichtige Rolle in Physik und Ingenieurwissenschaften haben: Es gibt das *Vektorprodukt* (oder *Kreuzprodukt*)[1]

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3, \quad (v, w) \mapsto v \times w,$$

[1]bekannt aus LA1, Beispiel 5.2.2(b)

das zwei Vektoren wieder auf einen Vektor abbildet und das *Skalarprodukt*[2]

$$\langle\cdot,\cdot\rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \quad (v,w) \mapsto \langle v,w\rangle,$$

das zwei Vektoren auf einen Skalar abbildet (daher der Name).

Das Kreuzprodukt $v \times w$ von $v \neq 0$ und $w \neq 0$ steht senkrecht auf v und w und für die Norm gilt:

$$\|v \times w\| = \|v\|\,\|w\|\sin(\sphericalangle(v,w)).$$

Für das Skalarprodukt $\langle v,w\rangle$ gilt:

$$\langle v,w\rangle = \|v\|\,\|w\|\cos(\sphericalangle(v,w)).$$

Da die Normen $\|v\|$ und $\|w\|$ für Vektoren $v,w \neq 0$ immer positiv sind, richtet sich das Vorzeichen von $\langle v,w\rangle$ nur nach dem Kosinus des eingeschlossenen Winkels, d.h.

$$\begin{aligned}
\forall v,w \in \mathbb{R}^3 \setminus \{0\} : \langle v,w\rangle > 0 &\iff \cos(\sphericalangle(v,w)) > 0 \iff \sphericalangle(v,w) < \pi/2\\
\forall v,w \in \mathbb{R}^3 \setminus \{0\} : \langle v,w\rangle = 0 &\iff \cos(\sphericalangle(v,w)) = 0 \iff \sphericalangle(v,w) = \pi/2\\
\forall v,w \in \mathbb{R}^3 \setminus \{0\} : \langle v,w\rangle < 0 &\iff \cos(\sphericalangle(v,w)) < 0 \iff \sphericalangle(v,w) > \pi/2
\end{aligned}$$

Es gilt also: Zwei Vektoren $v,w \in \mathbb{R}^3 \setminus \{0\}$ sind genau dann orthogonal, wenn ihr Skalarprodukt 0 ist:

$$v \perp w \iff \langle v,w\rangle = 0.$$

Außerdem folgt aus der obigen Formel für das Skalarprodukt folgende Formel für das Skalarprodukt eines Vektors mit sich selbst:

$$\langle v,v\rangle = \|v\|\,\|v\|\cos(\underbrace{\sphericalangle(v,v)}_{=0}) = \|v\|\,\|v\|\underbrace{\cos(0)}_{=1} = \|v\|\,\|v\| = \|v\|^2.$$

Wenn wir also einen Vektor mit sich selbst skalar multiplizieren und daraus die Wurzel ziehen, erhalten wir die Norm des Vektors.

Unser erstes Ziel wird es sein, diese Strukturen zu verallgemeinern, sodass wir auch auf anderen Vektorräumen als $\mathbb{R}^3$ über Abstände, Längen und Winkel sprechen können. Es gibt hier mehrere Ansätze: Man kann entweder die Abstandsfunktion als grundlegend betrachten und versuchen, alles andere daraus abzuleiten. Das führt zum Begriff des *metrischen Raumes*, auf den wir später kurz eingehen werden. Alternativ kann man die Normfunktion als grundlegend betrachten. Das führt dann zum Begriff des *normierten Vektorraums*. Wir werden hier allerdings dem Weg folgen, dass wir das Skalarprodukt als grundlegend betrachten und versuchen, alle anderen Strukturen (Abstände, Längen, Winkel) mit dem Skalarprodukt zu definieren.

Leider ist das Kreuzprodukt eine Struktur, die sich nicht ohne weiteres auf andere Räume als den $\mathbb{R}^3$ übertragen lässt. Es gibt natürlich Operationen auf anderen Räumen (z.B. allgemeinere alternierende Abbildungen (siehe LA1, Kapitel 5.2)), die gewisse Eigenschaften des Kreuzproduktes haben, aber trotz allem ist es möglich zu sagen, dass das Kreuzprodukt eine Struktur ist, die so nur auf dem Raum $\mathbb{R}^3$ existiert.

Wir werden dieses Semester die meiste Zeit nur über den Körpern $\mathbb{R}$ und $\mathbb{C}$ arbeiten, da sich die wichtigen geometrischen Begriffe leider nicht ohne weiteres auf andere Körper übertragen

[2]bekannt aus LA1, Beispiel 5.2.2(c)

lassen und wir des Öfteren besondere Eigenschaft der Körper $\mathbb{R}$ und $\mathbb{C}$ verwenden werden, um Aussagen herzuleiten, die über allgemeinen Körpern entweder nicht wahr sind oder erst gar keinen Sinn ergeben.

Auch wenn wir in den meisten Anwendungen nur an reellen Vektorräumen interessiert sind, so werden wir trotzdem viele Definitionen so formulieren, dass sie in $\mathbb{R}$ und in $\mathbb{C}$ funktionieren. Sie sollten sich also bei allen solchen Definitionen vor allem dafür interessieren, was dies im Falle $\mathbb{K} = \mathbb{R}$ bedeutet.

1.2. Skalarprodukte, Längen und Winkel

Bevor wir die allgemeine Definition eines Skalarproduktes auf einem beliebigen reellen oder komplexen Vektorraum geben, möchten wir an dieser Stelle kurz an einige Fakten und Notationen aus dem Gebiet der komplexen Zahlen erinnern, die alle aus der Linearen Algebra I bzw. der höheren Mathematik I bekannt sein sollten:

Proposition 1.2.1.

(a) *Die komplexen Zahlen $\mathbb{C}$ bilden eine Körpererweiterung der reellen Zahlen, d.h. der Körper der reellen Zahlen $\mathbb{R} \subseteq \mathbb{C}$ ist Unterkörper des Körpers der komplexen Zahlen $\mathbb{C}$.*

(b) *Die komplexe Konjugation $\overline{\cdot} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \overline{z}$ ist ein Körperautomorphismus, d.h.*

$$\forall z, w \in \mathbb{C}: \ \overline{z \cdot w} = \overline{z} \cdot \overline{w} \text{ und } \overline{z + w} = \overline{z} + \overline{w}.$$

(c) *Eine komplexe Zahl $z \in \mathbb{C}$ ist genau dann reell, wenn $\overline{z} = z$. Insbesondere gilt:*

$$\forall t \in \mathbb{R}: \ \overline{t} = t.$$

Wenn wir also wissen, dass eine Zahl reell ist, so können wir das komplexe Konjugieren ignorieren.

(d) *Wenn $a, b \in \mathbb{R}$ und $z = a + \mathrm{i}b$, dann gilt für den Realteil und den Imaginärteil:*

$$\mathrm{Re}(z) = a = \frac{1}{2}(z + \overline{z}).$$
$$\mathrm{Im}(z) = b = \frac{1}{2\mathrm{i}}(z - \overline{z}).$$

Insbesondere gilt also für jede komplexe Zahl $z \in \mathbb{C}$, dass $z + \overline{z} = 2\,\mathrm{Re}(z)$.

(e) *Wenn $a, b \in \mathbb{R}$ und $z = a + \mathrm{i}b$, dann gilt für den Betrag von z:*

$$|z|^2 = z \cdot \overline{z} = a^2 + b^2 = (\mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2.$$

Es gilt also insbesondere auch

$$|\mathrm{Re}(z)| \leq |z| \quad \text{und} \quad |\mathrm{Im}(z)| \leq |z|.$$

(f) *Für die komplexe Exponentialfunktion* $\exp : \mathbb{C} \to \mathbb{C}^\times, \quad z \mapsto e^z$ *gilt:*

$$\forall t \in \mathbb{R}: \ \mathrm{Re}(e^{\mathrm{i}t}) = \cos(t)$$
$$\forall t \in \mathbb{R}: \ \mathrm{Im}(e^{\mathrm{i}t}) = \sin(t)$$
$$\forall z \in \mathbb{C}: \ \left|e^z\right| = e^{\mathrm{Re}(z)}$$

(g) *Für jedes* $z \in \mathbb{C}$ *gibt es ein* $u \in \mathbb{C}$ *mit Betrag* 1*, sodass*

$$|z| = u \cdot z.$$

Auf den komplexen Zahlen $\mathbb{C}$ *gibt es keine Ordnung. Wenn wir also schreiben* $z \geq 0$ *oder* $z > 0$*, so meinen wir damit immer*

$$z \geq 0 :\Longleftrightarrow (z \in \mathbb{R} \textit{ und } z \geq 0).$$

Sollten Ihnen eine oder mehrere dieser Aussagen unklar sein, versuchen Sie, diese selbst zu beweisen (nachzurechnen).

Wir kommen nun zu der wichtigsten Definition dieses Kapitels:

Definition 1.2.2 (Skalarprodukt)**.**
Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein Vektorraum über $\mathbb{K}$. Ein *Skalarprodukt* auf V ist eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}, \quad (v, w) \mapsto \langle v, w \rangle,$$

die die folgenden Eigenschaften erfüllt:

(i) Die Abbildung $\langle \cdot, \cdot \rangle$ ist linear in der ersten Komponente, d.h.

$$\forall v_1, v_2, w \in V : \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$$

und

$$\forall v \in V, \lambda \in \mathbb{K}, w \in V : \langle \lambda v, w \rangle = \lambda \langle v, w \rangle.$$

(ii) Die Abbildung ist konjugiert symmetrisch, d.h.

$$\forall v, w \in V : \overline{\langle v, w \rangle} = \langle w, v \rangle.$$

(iii) Die Abbildung ist positiv definit, d.h.

$$\forall v \in V \setminus \{0\} : \langle v, v \rangle > 0.$$

Ein $\mathbb{K}$-Vektorraum V zusammen mit einem festen Skalarprodukt $\langle \cdot, \cdot \rangle$ nennt man *Vektorraum mit Skalarprodukt*, *Innenproduktraum* oder auch *Prähilbertraum*. Alle diese Begriffe bedeuten das gleiche.

- Falls $\mathbb{K} = \mathbb{R}$ ist, dann nennt man einen Vektorraum mit Skalarprodukt auch *Euklidischen Vektorraum*.

- Falls $\mathbb{K} = \mathbb{C}$ ist, dann nennt man einen Vektorraum mit Skalarprodukt auch *unitären Vektorraum*.

Zwei Vektoren $v, w \in V$ heißen *orthogonal* (bezüglich des gewählten Skalarproduktes), wenn $\langle v, w\rangle = 0$. Hierfür schreiben wir auch $v \perp w$.

Bemerkung 1.2.3. (1) Beachten Sie, dass im Falle $\mathbb{K} = \mathbb{R}$ die komplexe Konjugation in Teil (ii) wegfällt (Proposition 1.2.1(c)).

(2) Beachten Sie, dass im Falle $\mathbb{K} = \mathbb{C}$ die Forderung $\langle v, v\rangle > 0$ in Teil (iii) bereits impliziert, dass $\langle v, v\rangle \in \mathbb{R}$ ist. (Proposition 1.2.1).[3]

(3) Die Bedingung (i), die fordert, dass das Skalarprodukt linear in der *ersten* Komponente ist, ist nicht einheitlich in der Literatur. Gerade in der Physik wird oft die Linearität in der *zweiten* Komponente gefordert. Im Falle $\mathbb{K} = \mathbb{R}$ ist dies aber egal, weil nach (ii) das Skalarprodukt sowieso symmetrisch ist.

(4) Der Nullvektor $0 \in V$ steht auf jedem anderen Vektor orthogonal, weil das Skalarprodukt linear in der ersten Komponente ist.

Für uns wird zumindest am Anfang nur der Fall $\mathbb{K} = \mathbb{R}$ interessant sein, aber es ist sinnvoll, die wichtigsten Eigenschaften von Skalarprodukten auch gleich in dieser Allgemeinheit zu formulieren, sodass sie für $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$ gelten.

Beispiel 1.2.4. (a) Für jedes $n \in \mathbb{N}$ sei auf dem reellen Vektorraum $V = \mathbb{R}^n$ ist das *Standardskalarprodukt* definiert als

$$\langle x, y\rangle = y^\top \cdot x = \sum_{j=1}^{n} x_j y_j.$$

Die Bedingung (i) ist sofort offensichtlich. Da diese Abbildung offenbar symmetrisch in x und y ist, gilt auch (ii).

Für (iii) sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \setminus \{0\}$. Dann gilt:

$$\langle x, x\rangle = \sum_{j=1}^{n} (x_j)^2.$$

Da alle x_j reell sind, sind die Summanden alle nichtnegativ. Das zeigt, dass $\langle x, x\rangle \geq 0$ ist. Nun ist aber nach Voraussetzung der Vektor x nicht der Nullvektor, also gibt es mindestens ein x_j, das nicht 0 ist. Somit ist mindestens ein Summand in der Summe echt größer 0, also ist die gesamte Summe positiv.

Also ist dies ein reelles Skalarprodukt und $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ wird zu einem Euklidischen Vektorraum.

(b) Für jedes $n \in \mathbb{N}$ sei auf dem komplexen Vektorraum $V = \mathbb{C}^n$ das *komplexe Standardskalarprodukt* definiert als

$$\langle x, y\rangle = \overline{y^\top} \cdot x = \sum_{j=1}^{n} x_j \overline{y_j}.$$

Mit Hilfe von Proposition 1.2.1 lässt sich leicht zeigen, dass dies ein komplexes Skalarprodukt ist und somit $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum wird.

[3] Die Aussage, dass $\langle v, v\rangle$ stets reell ist, folgt aber außerdem auch aus Teil (ii) und Proposition 1.2.1(c).

(c) Es seien $a,b \in \mathbb{R}$ mit $a < b$. Dann ist der Raum aller stetigen Funktion

$$V := C([a,b],\mathbb{R}) := \{f : [a,b] \to \mathbb{R} \mid f \text{ ist stetig}\} \subseteq \mathbb{R}^{[a,b]}$$

ein $\mathbb{R}$-Vektorraum. Hierauf definieren wir nun das Skalarprodukt

$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}, \quad (f,g) \mapsto \langle f,g\rangle := \int_a^b f(t)g(t)dt.$$

Dies ist sehr ähnlich zum Standardskalarprodukt auf $\mathbb{R}^n$, wo die beiden Vektoren komponentenweise multipliziert und dann aufsummiert werden, nur dass hier anstelle der Summation das Integral benutzt wird.

Für festes g ist die Abbildung $f \mapsto f \cdot g$ linear. Da Integrieren auch linear ist, folgt, dass diese hier definierte Skalarprodukt Bedingung (i) unserer Definition erfüllt. Da Multiplikation von Funktionen ja punktweiser Multiplikation im Grundkörper $\mathbb{R}$ entspricht, die kommutativ ist, folgt damit (ii).

Kommen wir zu (iii): Es sei $f \in C([a,b],\mathbb{R}) \setminus \{0\}$ eine stetige Funktion. Dann ist

$$\langle f,f\rangle = \int_a^b f(t)\cdot f(t)dt = \int_a^b (f(t))^2\,dt.$$

Da f reellwertig ist, ist $(f(t))^2$ immer größer und gleich 0 und somit ist auch das Integral nichtnegativ.

Da aber außerdem f nicht die Nullfunktion ist, gibt es mindestens einen Punkt $t_0 \in [a,b]$, mit $(f(t_0))^2 > 0$. Aus der Stetigkeit von f folgt nun, dass es eine kleine ε-Umgebung von t_0 geben muss, auf der f^2 echt größer als 0 ist. Somit muss das Integral von f^2 insgesamt auch echt größer als 0 sein.[4]

Also wird $C([a,b],\mathbb{R})$ mit diesem Skalarprodukt ein unendlich dimensionaler Euklidischer Vektorraum.

Man kann dann beispielsweise leicht nachrechnen, dass die Funktionen $f_1(t) = \sin(t)$ und $f_2(t) = \cos(t)$ im Euklidischen Vektorraum $C([0,2\pi],\mathbb{R})$ orthogonal sind.

(d) In Abwandlung des letztes Beispiels lassen sich viele andere Skalarprodukte auf Funktionenräumen definieren:

- Auf dem Raum der komplexwertigen stetigen Funktionen $C([a,b],\mathbb{C})$ ist

  $$\langle f,g\rangle = \int_a^b f(t)\overline{g(t)}dt$$

  ein komplexes Skalarprodukt.

- Auf dem Raum der Polynome $\mathbb{R}[X]$ ist $\langle p,q\rangle := \int_a^b p(t)q(t)dt$ ein Skalarprodukt, das natürlich von a und b abhängt. Es gibt somit – je nach Wahl des Intervalls unendlich viele verschiedene Skalarprodukte auf $\mathbb{R}[X]$.

[4]Die Details dieses Argumentes sind eher analytischer als algebraischer Natur und sind somit dem Leser überlassen.

- Man kann auch unbeschränkte Integrationsgebiete verwenden. Beispielsweise ist
$$\langle p,q\rangle := \int_0^\infty p(t)q(t)e^{-t}dt$$
ein Skalarprodukt auf $\mathbb{R}[X]$. Der Faktor e^{-t} ist dafür da, die Konvergenz des uneigentlichen Integrals zu sichern. Dies muss natürlich zuerst überprüft werden. Analoge Beispiele lassen sich mit anderen Gewichtsfunktionen konstruieren.

Lemma 1.2.5 (Erste Eigenschaften von Skalarprodukten)**.**
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ *und* V *ein* $\mathbb{K}$*-Vektorraum mit Skalarprodukt* $\langle\cdot,\cdot\rangle$.

(a) Für alle $v \in V$ *gilt* $\langle v,v\rangle \geq 0$.

(b) Im Fall $\mathbb{K} = \mathbb{R}$ *ist das Skalarprodukt* bilinear[5]*, d.h. für festes* v *ist* $\langle v,\cdot\rangle$ *linear.*

(c) Im Fall $\mathbb{K} = \mathbb{C}$ *ist das Skalarprodukt* sesquilinear*, d.h. für festes* v *ist* $\langle v,\cdot\rangle$ konjugiert linear*, d.h.*
$$\forall v,w_1,w_2 \in V : \langle v,w_1+w_2\rangle = \langle v,w_1\rangle + \langle v,w_2\rangle$$
und
$$\forall v \in V, \lambda \in \mathbb{C}, w \in V : \langle v,\lambda w\rangle = \overline{\lambda}\,\langle v,w\rangle\,.$$

(d) Im Falle $\mathbb{K} = \mathbb{C}$ *ist die Abbildung* $\mathrm{Re}\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}, \quad (v,w) \mapsto \mathrm{Re}(\langle v,w\rangle)$ *ein* reelles *Skalarprodukt auf dem reellen Vektorraum* V*, den man erhält, wenn man die Skalare des komplexen Vektorraums auf* $\mathbb{R}$ *einschränkt.*

Da nach Lemma 1.2.5 (a) das Skalarprodukt eines Vektors mit sich selbst immer reell und nichtnegativ ist, können wir nun jedem Vektor eine Norm zuweisen:

Definition 1.2.6 (Die Norm eines Vektors)**.**
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann heißt die Abbildung
$$\|\cdot\| : V \to [0,+\infty), \quad v \mapsto \|v\| := \sqrt{\langle v,v\rangle}$$
die zum Skalarprodukt gehörige *Normabbildung*.

Proposition 1.2.7 (Binomische Formeln für Skalarprodukte)**.**
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ *und* V *ein* $\mathbb{K}$*-Vektorraum mit Skalarprodukt* $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$*. Dann gilt:*

(i) $\forall v,w \in V: \ \|v+w\|^2 = \|v\|^2 + 2\,\mathrm{Re}\,\langle v,w\rangle + \|w\|^2$.

(ii) $\forall v,w \in V: \ \|v-w\|^2 = \|v\|^2 - 2\,\mathrm{Re}\,\langle v,w\rangle + \|w\|^2$.

(iii) Falls $\mathbb{K} = \mathbb{R}$*, dann gilt:* $\forall v,w \in V: \ \langle v+w,v-w\rangle = \|v\|^2 - \|w\|^2$.

Im Fall $\mathbb{K} = \mathbb{R}$ *ist der Realteil in Formeln (i) und (ii) natürlich überflüssig.*

Eine direkte Folgerung aus diesen Formeln ist der folgende Satz:

[5] siehe auch LA1, Definition 5.2.1(b)

Satz 1.2.8 (Satz des Pythagoras für Vektorräume mit Skalarprodukt). *Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$:*

$$(\langle v, w\rangle = 0) \implies \left(\|v + w\|^2 = \|v\|^2 + \|w\|^2\right).$$

Für $\mathbb{K} = \mathbb{R}$ gilt auch die Rückimplikation.

Beweis. Wir verwenden die binomischen Formeln für Skalarprodukte (Proposition 1.2.7):

$$\|v + w\|^2 = \|v\|^2 + 2\operatorname{Re}\langle v, w\rangle + \|w\|^2 .$$

Wenn also $\langle v, w\rangle = 0$ ist, dann ist die Aussage des Satzes von Pythagoras gezeigt.

Umgekehrt, wenn die Gleichheit

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

gilt, dann folgt – zusammen mit der oben verwendeten binomischen Formel, dass

$$2\operatorname{Re}\langle v, w\rangle = 0 .$$

Wenn nun schließlich zusätzlich $\mathbb{K} = \mathbb{R}$ gilt, dann fällt der Realteil weg und wir erhalten $\langle v, w\rangle = 0$. □

Bemerkung. Ebenfalls sehr hilfreich für das Rechnen in Vektorräumen mit Skalarprodukt ist die folgende *Parallelogrammgleichung*, die Sie in der Übung bewiesen haben:

$$\forall x, y \in V : 2\|x\|^2 + 2\|y\|^2 = \|x - y\|^2 + \|x + y\|^2 .$$

Diese gilt in jedem (reellen oder komplexen) Vektorraum mit Skalarprodukt. Der Name kommt daher, dass diese Gleichung angewandt auf Vektoren x und y in der euklidischen Ebene $\mathbb{R}^2$ besagt, dass die Summe der Quadrate der vier Seitenlängen eines Parallelogramms gleich ist der Summe der Quadrate der Diagonallängen.

Man kann zeigen, dass eine Norm auf einem Vektorraum (siehe Definition 1.2.18) genau dann die Parallelogrammgleichung erfüllt, wenn sie von einem Skalarprodukt induziert wird. Diese Äquivalenz werden wir in dieser Veranstaltung nicht beweisen, weil der Beweis sehr technisch ist und wir die Aussage nicht weiter benötigen.

Bis jetzt folgte alles, was wir bisher über Skalarprodukte gezeigt haben, mehr oder wenig direkt aus der Definition. Die folgende Aussage ist nun die erste Aussage über Skalarprodukte, die nicht-trivial ist, in dem Sinne, dass der Beweis einen Trick enthält, auf den man sonst vielleicht nicht direkt kommt. Es ist eine der wichtigsten Aussagen über Skalarprodukte überhaupt:

Satz 1.2.9 (Die Ungleichung von Cauchy-Schwarz). *Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$:*

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

Beweis. Wir betrachten zuerst den Fall $\|v\| = 0$.
Hieraus folgt sofort, dass $\langle v, v\rangle = \|v\|^2 = 0$ und mit der positiven Definitheit des Skalarproduktes (Definition 1.2.2 (iii)) dann auch $v = 0$. Für $v = 0$ ist die Ungleichung von Cauchy-Schwarz aber trivialerweise erfüllt, weil beide Seiten der Ungleichung gleich 0 sind (Hierfür verwenden wir die Linearität des Skalarproduktes in der ersten Komponente).

Kommen wir nun zum Fall $\|v\| > 0$.
Der Wert $\langle v, w\rangle$ ist ein Element in $\mathbb{K}$ und somit[6] gibt es ein $u \in \mathbb{K}$ mit $|u| = 1$, sodass

$$|\langle v, w\rangle| = u\,\langle v, w\rangle\,.$$

Definieren wir uns nun die folgende Hilfsfunktion:

$$f : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \|(tu)v + w\|^2\,.$$

Es ist klar, dass für alle $t \in \mathbb{R}$ gilt: $f(t) \geq 0$.

Wir können nun mit den binomischen Formeln (siehe Proposition 1.2.7) ausmultiplizieren und erhalten:

$$\begin{aligned}
f(t) &= \|(tu)v + w\|^2 \\
&= \|(tu)v\|^2 + 2\,\mathrm{Re}(\langle (tu)v, w\rangle) + \|w\|^2 \\
&= |tu|^2\,\|v\|^2 + 2\,\mathrm{Re}(tu\,\langle v, w\rangle) + \|w\|^2 \\
&= \underbrace{|t|^2}_{=t^2}(\underbrace{|u|}_{=1})^2\,\|v\|^2 + 2t\,\mathrm{Re}\big(\underbrace{u\,\langle v, w\rangle}_{=|\langle v,w\rangle|}\big) + \|w\|^2 \\
&= \|v\|^2 \cdot t^2 + 2\,|\langle v, w\rangle| \cdot t + \|w\|^2 \\
&= at^2 + bt + c\,,
\end{aligned}$$

mit $a = \|v\|^2, b = 2\,|\langle v, w\rangle|$ und $c = \|w\|^2$. Hierbei wurde die Homogenität der Norm ($\|\lambda v\| = |\lambda|\,\|v\|$) benutzt, die zwar erst in Satz 1.2.10 offiziell auftaucht, aber direkt aus der Definition der Norm, sowie aus den Eigenschaften des Skalarproduktes (Lemma 1.2.5) folgt.

Wir sehen also: Die Funktion f ist eine Polynomfunktion von Grad 2 mit Leitkoeffizient $\|v\|^2 > 0$. Aus der allgemeinen Lösungstheorie von quadratischen Gleichungen ist bekannt, dass eine solche Gleichung zwei verschiedene Lösungen hat, wenn die Diskriminante $b^2 - 4ac > 0$ ist. In diesem Fall ist die Funktion hat die Funktion an beiden Nullstellen einen Vorzeichenwechsel. Dies ist aber nicht möglich, weil wir aus der Definition der Funktion gesehen haben, dass f immer größer oder gleich 0 ist.

Somit folgt: Die Diskriminante $b^2 - 4ac$ muss immer kleiner gleich 0 sein. Es gilt also:

$$\begin{aligned}
& b^2 - 4ac \leq 0 \\
\Longrightarrow\quad & (2\,|\langle v, w\rangle|)^2 - 4\,\|v\|^2\,\|w\|^2 \leq 0 \\
\Longrightarrow\quad & 4\,|\langle v, w\rangle|^2 \leq 4(\|v\|\,\|w\|)^2 \\
\Longrightarrow\quad & |\langle v, w\rangle| \leq \|v\|\,\|w\|\,.
\end{aligned}$$

□

Wir werden nun mit Hilfe der Cauchy-Schwarz-Ungleichung die Dreiecksungleichung der Normfunktion beweisen. Zusammen mit zwei weiteren (einfacher zu zeigenden) Eigenschaften ist sie im folgenden Satz zusammengefasst:

[6] Das ist Proposition 1.2.1 (g) für $\mathbb{K} = \mathbb{C}$, aber die Aussage gilt natürlich genauso für $\mathbb{K} = \mathbb{R}$.

Satz 1.2.10 (Definitheit, Homogenität und Dreiecksungleichung). *Es sei* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ *und* V *ein* $\mathbb{K}$*-Vektorraum mit Skalarprodukt. Dann gilt:*

(a) $\forall v \in V: \quad \big(\|v\| = 0 \iff v = 0\big).$ (positive Definitheit)

(b) $\forall v \in V, \forall \lambda \in \mathbb{K}: \quad \|\lambda v\| = |\lambda| \, \|v\|.$ (Homogenität)

(c) $\forall v, w \in V: \quad \|v + w\| \leq \|v\| + \|w\|.$ (Dreiecksungleichung)

Beweis. (a)
Es gilt

$$\|v\| = 0 \iff \|v\|^2 = 0 \iff \langle v, v\rangle = 0 \iff v = 0.$$

Die letzte Äquivalenz folgt aus der positiven Definitheit des Skalarproduktes (Definition 1.2.2 (iii)), sowie der Linearität in der ersten Komponente (Definition 1.2.2 (i)).

(b)
Diese Eigenschaft, die wir bereits im Beweis der Cauchy-Schwarz-Ungleichung (Satz 1.2.9) verwendet haben, zeigt man so:

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

Hier haben wir – neben der Definition der Norm – die Linearität des Skalarproduktes in der ersten Komponente (Definition 1.2.2 (i)), die konjugierte Linearität des Skalarproduktes in der zweiten Komponente (Lemma 1.2.5), sowie die Identität $|z|^2 = z\overline{z}$ (Proposition 1.2.1 (e)).

(c)
Der Beweis der Dreiecksungleichung ist nicht mehr schwer, wenn wir die Cauchy-Schwarz-Ungleichung (Satz 1.2.9) zur Verfügung haben:

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

Hier haben zuerst die binomische Formel für Skalarprodukte verwendet (Proposition 1.2.7), dann elementare Umformungen mit Realteil und Betrag (Proposition 1.2.1), dann die Cauchy-Schwarz-Ungleichung (Satz 1.2.9) und schließlich die gewöhnliche binomische Formel für reelle Zahlen. □

Bemerkung 1.2.11. Man ist versucht, die Cauchy-Schwarz- oder die Dreiecksungleichung als trivial anzusehen, wenn man sich von der Geometrie im zwei- oder dreidimensionalen leiten lässt. Bemerkenswert ist aber, dass diese Ungleichungen in beliebigen Vektorräumen mit Skalarprodukt gelten, beispielsweise in $C([a,b], \mathbb{R})$ (versehen mit dem Integral-Skalarprodukt aus Beispiel 1.2.4(c)). Dort sagen die beiden Ungleichungen folgendes aus, dass für stetige Funktionen $f, g : [a,b] \to \mathbb{R}$ gilt:

$$\left| \int_a^b \big(f(t) g(t)\big) dt \right| \leq \sqrt{\int_a^b \big(f(t)\big)^2 dt} \sqrt{\int_a^b \big(g(t)\big)^2 dt}$$

und

$$\sqrt{\int_a^b \big(f(t) + g(t)\big)^2 dt} \leq \sqrt{\int_a^b \big(f(t)\big)^2 dt} + \sqrt{\int_a^b \big(g(t)\big)^2 dt}.$$

Keine dieser beiden Formeln ist offensichtlich, wenn man versucht, sie direkt zu beweisen.

Eine weitere Folgerung aus der Cauchy-Schwarz-Ungleichung ist, dass sie uns ermöglicht, in Euklidischen Vektorräumen Winkel zwischen Vektoren zu definieren:

Definition 1.2.12.
Es sei $\mathbb{K} = \mathbb{R}$ und V ein Euklidischer Vektorraum und es seien $v, w \in V \setminus \{0\}$ gegeben. Dann ist der *Winkel zwischen* v und w definiert als:

$$\sphericalangle(v,w) := \arccos\left(\frac{\langle v,w\rangle}{\|v\|\,\|w\|}\right) \in [0,\pi].$$

Es gilt somit für $v, w \in V \setminus \{0\}$ die Formel

$$\langle v,w\rangle = \|v\|\,\|w\| \cos\big(\sphericalangle(v,w)\big).$$

Bemerkung 1.2.13. (a) Diese Definition ist wohldefiniert, weil – nach der Cauchy-Schwarzschen Ungleichung – das Skalarprodukt $\langle v,w\rangle$ im Intervall $[-\|v\|\,\|w\|, \|v\|\,\|w\|]$ liegt und somit, wenn man durch das Produkt der Normen teilt (was erlaubt ist, weil $v, w \neq 0$ gilt), nur Werte in $[-1,1]$ angenommen werden. Nun ist die Kosinus-Funktion nach entsprechender Einschränkung des Definitionsbereichs auf $[0,\pi]$ und Koeinschränkung des Wertebereichs auf $[-1,1]$ bijektiv mit Umkehrfunktion $\arccos : [-1,1] \to [0,\pi]$ und deshalb ist dieser Winkel wohldefiniert.

(b) Beachten Sie, dass der Winkel zwischen Vektoren nur definiert ist, wenn $\mathbb{K} = \mathbb{R}$ ist und wenn beide Vektoren nicht 0 sind. Die Orthogonalitätsrelation $v \perp w \iff \langle v,w\rangle = 0$ dagegen ist allgemeiner für alle Vektoren in beliebigen reellen oder komplexen Vektorräumen mit Skalarprodukt definiert. Im Spezialfall, dass $\mathbb{K} = \mathbb{R}$ ist und $v, w \neq 0$ sind, gilt aber $v \perp w \iff \sphericalangle(v,w) = \pi/2$.

Definition 1.2.14 (Abstandsfunktion)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann ist der *Abstand* zwischen Punkten $a, b \in V$ definiert als

$$d(a,b) := \|a-b\|.$$

Man zeigt leicht die folgenden Eigenschaften der Abstandsfunktion:

Proposition 1.2.15.

(a) $\forall a,b \in V: \quad d(a,b) = 0 \iff \big(a = b\big).$ *(*positive Definitheit*)*

(b) $\forall a,b \in V: \quad d(a,b) = d(b,a).$ *(*Symmetrie*)*

(c) $\forall a,b,c \in V: \quad d(a,c) \leq d(a,b) + d(b,c).$ *(*Dreiecksungleichung*)*

Wir können also in Vektorräumen mit Skalarprodukt Abstände zwischen Punkten messen. Dies ist wichtig für die Analysis, denn sobald man Abstände messen kann, kann man die gewöhnlichen Definitionen von Folgenkonvergenz oder Stetigkeit von Funktionen verwenden. Es stellt sich heraus, dass es gar nicht notwendig ist, Skalarprodukte zu haben, um Abstände zu messen. Ausreichend ist ein sogenannter metrischer Raum:

Definition 1.2.16.
Es sei X eine beliebige Menge. Dann heißt eine Funktion $d : X \times X \to [0,+\infty)$ eine *Metrik* auf X, wenn sie die drei Eigenschaften aus Proposition 1.2.15 erfüllt. Ein *metrischer Raum* (X,d) ist nun eine Menge, zusammen mit einer Metrik.

Beispiel 1.2.17. (a) Die Mengen $\mathbb{R}$ oder $\mathbb{C}$ sind mit der jeweiligen Abstandsfunktion $d(x,y) = |x-y|$ metrische Räume.

(b) Jede Teilmenge eines metrischen Raums ist mit der eingeschränkten Abstandsfunktion ein metrischer Raum. Beispielsweise ist das Einheitsintervall $[0,1]$ ein metrischer Raum. Oder die abgeschlossene Kreisscheibe $\{z \in \mathbb{C} \mid |z| \le 1\}$. Ebenso ist ein Dreieck $\Delta \subseteq \mathbb{C}$ ein metrischer Raum. Wir sehen also: Ein metrischer Raum muss kein Vektorraum sein. Es reicht, eine Menge zu haben, auf der es möglich ist, Abstände zu berechnen.

(c) Man kann auch auf der (endlichen) Menge $(\mathbb{F}_2)^n$ eine Metrik einführen, die bei zwei Vektoren zählt, wie viele Einträge (Bits) unterschiedlich sind. Diese Metrik heißt *Hamming-Abstand* und ist in der Informatik von Bedeutung. Für diese Veranstaltung wird diese aber keine große Rolle spielen.

Ein weiterer verwandter Begriff ist der des normierten Vektorraums:

Definition 1.2.18.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein Vektorraum über $\mathbb{K}$. Eine *Norm* auf V ist eine Abbildung

$$\|\cdot\| : V \to [0, +\infty),$$

die die Eigenschaften von Satz 1.2.10 erfüllt.

Ein $\mathbb{K}$-Vektorraum zusammen mit einer Norm heißt *normierter Vektorraum*.

Wir haben in Satz 1.2.10 gesehen, dass jedes Skalarprodukt eine Norm induziert, somit ist jeder Vektorraum mit Skalarprodukt ein normierter Vektorraum. Es gibt aber auch Normen, die nicht von Skalarprodukten kommen und leider spielen auch solche Normen in der Praxis manchmal eine Rolle. In dieser Veranstaltung werden wir uns aber im Wesentlichen nur mit Normen beschäftigen, die von Skalarprodukten kommen.

Wir haben in Proposition 1.2.15 dann gesehen, dass jede Norm eine Metrik impliziert. Somit gilt:

$$\boxed{\text{Skalarprodukt} \rightsquigarrow \text{Norm} \rightsquigarrow \text{Metrik}}.$$

Orthogonalität von Vektoren ist ein Konzept, dass nur in Vektorräumen mit Skalarprodukt Sinn ergibt, die Länge eines Vektors ergibt in jedem normierten Vektorraum Sinn, während Abstände zwischen Punkten (und damit verbunden: Stetigkeit von Funktionen und Konvergenz von Folgen) in jedem metrischen Raum sinnvoll ist.

Zusammenfassung von Abschnitt 1.2

(1) Ein Skalarprodukt auf einem reellen oder komplexen Vektorraum ist eine Abbildung, die ein paar von zwei Vektoren auf einen Skalar abbildet und drei Axiome erfüllen muss.

(2) Ein reeller Vektorraum mit Skalarprodukt heißt Euklidischer Vektorraum, ein komplexer Vektorraum mit Skalarprodukt heißt unitärer Vektorraum.

(3) Zwei Vektoren sind orthogonal, wenn sie Skalarprodukt Null haben. Winkel sind nur in Euklidischen Vektorräumen definiert.

(4) Jedes Skalarprodukt definiert eine Norm. Es gilt die Cauchy-Schwarzsche Ungleichung.

(5) Jede Norm definiert eine Metrik/Abstandsfunktion. ES gilt die Dreiecksungleichung.

1.3. Orthonormalbasen

Definition 1.3.1 (Orthogonalsystem, Orthonormalsystem).
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Eine Teilmenge $M \subseteq V$ heiße

(a) *Orthogonalsystem* (OGS), wenn alle Vektoren ungleich 0 sind und paarweise orthogonal aufeinander stehen.

(b) *Orthonormalsystem* (*ONS*), wenn alle Vektoren Länge 1 haben und paarweise orthogonal aufeinander stehen.

Jedes Orthonormalsystem ist ein Orthogonalsystem, aber nicht umgekehrt.

Bemerkung 1.3.2. Aus jedem Orthogonalsystem M lässt sich durch Skalieren der Vektoren ein Orthonormalsystem herstellen durch

$$\widetilde{M} := \left\{ \frac{1}{\|v\|} v \;\middle|\; v \in M \right\}.$$

Es gilt $\mathrm{LH}_{\mathbb{K}}\left(\widetilde{M}\right) = \mathrm{LH}_{\mathbb{K}}(M)$.

Lemma 1.3.3 (Orthogonalsysteme sind linear unabhängig).
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann ist jedes Orthogonalsystem (und somit natürlich auch jedes Orthonormalsystem) linear unabhängig.

Definition 1.3.4 (Orthonormalbasis).
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein *endlich dimensionaler* $\mathbb{K}$-Vektorraum mit Skalarprodukt. Eine *Orthonormalbasis* (*ONB*) B ist ein (notwendigerweise endliches) Orthonormalsystem von V, das eine Basis von V ist.

Analog ist eine *Orthogonalbasis* (OGB) ein Orthogonalsystem, das eine Basis von V ist.

Wie schon in LA1 definiert man eine *geordnete Orthonormalbasis* (bzw. *Orthogonalbasis*) als ein Tupel $(b_1, \ldots, b_n)$ von n verschiedenen Vektoren, sodass $\{b_1, \ldots, b_n\}$ eine Orthonormal(bzw. gonal)basis ist.

Bemerkung 1.3.5. (a) Da nach Lemma 1.3.3 Orthogonalsysteme (und damit auch Orthonormalsysteme) linear unabhängig sind, ist nur zu zeigen, dass B den ganzen Raum aufspannt, d.h. $\mathrm{LH}_{\mathbb{K}}(B) = V$.

(b) Orthonormalbasen gibt es auch in unendlich dimensionalen Vektorräumen und spielen dort auch eine wichtige Rolle. Leider ist es aber so, dass es dort unterschiedliche und nicht äquivalente Definitionen gibt. Entweder man verlangt wie im Endlichdimensionalen, dass die Teilmenge B den Raum V erzeugt, d.h., dass jeder Vektor in V eine *endliche* Linearkombination von Vektoren aus B ist. Oder man verlangt nur, dass der von B erzeugte Untervektorraum *dicht* liegt, was bedeutet, dass es möglich ist, jeden Vektor aus V beliebig genau durch endliche Linearkombinationen aus B anzunähern. Man kann zeigen (das tun wir hier aber nicht), dass dies bedeutet, dass jeder Vektor in V in unendliche Linearkombination von Vektoren aus B ist. Unendliche Linearkombinationen haben wir in LA1 immer vermieden, weil es in einem beliebigen Vektorraum keinen Sinn ergibt, von Konvergenz von Reihen zu sprechen. In einem Vektorraum mit Skalarprodukt (und noch allgemeiner in einem normierten Vektorraum) ist es aber auch möglich, Grenzwerte zu definieren und somit auch Konvergenz von Reihen. Dieses Konzept, dass es möglich ist, Elemente eines Vektorraums als unendliche Linearkombination von Vektoren einer Orthonormalbasis zu schreiben, haben Sie bereits in der Vorlesung HM1 gesehen, wenn Sie Funktionen als Fourier-Reihen dargestellt haben.

Proposition 1.3.6 (Fourier-Formel)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein n-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und einer Orthonormalbasis $B = \{b_1, \ldots, b_n\}$. Wenn $v \in V$ ein Vektor gegeben ist, dann ist die Darstellung bezüglich B gegeben durch:

$$v = \sum_{j=1}^{n} \langle v, b_j \rangle b_j.$$

In anderen Worten: Die Koordinaten von v bezüglich der geordneten Orthonormalbasis $\mathsf{B} = (b_1, \ldots, b_n)$ ist gegeben durch:

$$(v)_{\mathsf{B}} := \begin{pmatrix} \langle v, b_1 \rangle \\ \vdots \\ \langle v, b_n \rangle \end{pmatrix}.$$

Wir sehen also: Um die Koordinaten eines Vektors bezüglich einer Orthonormalbasis zu bestimmen, reicht es aus, n Skalarprodukte zu berechnen. Es ist nicht notwendig, ein lineares Gleichungssystem zu bestimmen. Darüber hinaus ist es möglich, diese Koordinaten unabhängig voneinander zu berechnen, d.h. wenn nur eine Koordinate benötigt wird, ist auch nur ein Skalarprodukt zu bestimmen.[7]

Es bleibt die Frage: Besitzt jeder endlich dimensionale Vektorraum mit Skalarprodukt eine Orthonormalbasis und falls ja, wie findet man sie?

Satz 1.3.7 (Orthogonalisierung nach Gram-Schmidt)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Gegeben sei eine linear unabhängige Teilmenge $M = \{v_1, v_2, \ldots, v_n\}$ aus n Vektoren.

[7] Außerdem lässt sich die Berechnung von allen n Koordinaten gut parallelisieren, weil die Rechnungen nicht voneinander abhängen.

Wir definieren Vektoren $w_1,\dots,w_n$ rekursiv:

$$w_1 := v_1$$
$$w_j := v_j - \sum_{i=1}^{j-1} \frac{\langle v_j, w_i\rangle}{\langle w_i, w_i\rangle} w_i \qquad \text{für } j \geq 2.$$

Dann gilt: Die n-elementige Menge $\{w_1,\dots,w_n\}$ ist ein Orthogonalsystem.

Außerdem gilt für alle $m \leq n$, dass $\mathrm{LH}(w_1,\dots,w_m) = \mathrm{LH}(v_1,\dots,v_m)$.

Falls $\{v_1,\dots,v_n\}$ eine Basis von V ist, so ist $\{w_1,\dots,w_n\}$ eine Orthogonalbasis von V und $\left\{\frac{1}{\|w_1\|}w_1,\dots,\frac{1}{\|w_n\|}w_n\right\}$ ist eine Orthonormalbasis.

Korollar 1.3.8.
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$. Jeder endlich dimensionale $\mathbb{K}$-Vektorraum mit Skalarprodukt besitzt eine Orthonormalbasis. Diese ist im Allgemeinen nicht eindeutig.

Weiter gilt: Jedes Orthonormalsystem in einem endlich dimensionalen Vektorraum lässt sich zu einer Orthonormalbasis vervollständigen.

Beispiel 1.3.9. (a) Es sei $\mathbb{K} = \mathbb{R}$ und

$$A := \begin{pmatrix} -4 & -3 & -1 \\ 8 & 6 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

Dann ist $V := \ker(A) \subseteq \mathbb{R}^3$ ein Untervektorraum von $\mathbb{R}^3$. Mit dem Standardskalarprodukt von $\mathbb{R}^3$ wird V ein euklidischer Vektorraum. Wir möchten nun eine Orthonormalbasis von V finden.

Dazu berechnen wir zuerst eine Basis von V mit Hilfe des Gauß-Algorithmus:

x_1	x_2	x_3		
-4	-3	-1	0	$\mid \cdot 2$ auf Zeile 2
8	6	2	0	
-4	-3	-1	0	$\mid \cdot \left(\frac{1}{4}\right)$
0	0	0	0	
-1	$-\frac{3}{4}$	$-\frac{1}{4}$	0	
0	0	0	0	

Wir erhalten die Lösungsmenge

$$V := \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix} -\frac{3}{4} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{4} \\ 0 \\ 1 \end{pmatrix}\right) = \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}\right).$$

Der Vektorraum V hat also eine Basis $\mathrm{B} := \left(\begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}\right) = (v_1, v_2)$. Wir wollen nun Gram-Schmidt-Orthogonalisierung auf diese Basis anwenden, um eine Orthogonal- und schließlich eine Orthonormalbasis zu erhalten.

Wir setzen dazu

$$w_1 := v_1 = \begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}$$

und

$$\begin{aligned}
w_2 &:= v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 \\
&= \begin{pmatrix}-1\\0\\4\end{pmatrix} - \frac{\left\langle \begin{pmatrix}-1\\0\\4\end{pmatrix}, \begin{pmatrix}-3\\4\\0\end{pmatrix}\right\rangle}{\left\langle \begin{pmatrix}-3\\4\\0\end{pmatrix}, \begin{pmatrix}-3\\4\\0\end{pmatrix}\right\rangle} \begin{pmatrix}-3\\4\\0\end{pmatrix} \\
&= \begin{pmatrix}-1\\0\\4\end{pmatrix} - \frac{3}{25}\begin{pmatrix}-3\\4\\0\end{pmatrix} \\
&= \frac{1}{25}\left(\begin{pmatrix}-25\\0\\100\end{pmatrix} - 3\begin{pmatrix}-3\\4\\0\end{pmatrix}\right) \\
&= \frac{1}{25}\begin{pmatrix}-25-3(-3)\\-3\cdot 4\\100\end{pmatrix} \\
&= \frac{1}{25}\begin{pmatrix}-16\\-12\\100\end{pmatrix} \\
&= \frac{4}{25}\begin{pmatrix}-4\\-3\\25\end{pmatrix}.
\end{aligned}$$

Also bilden die Vektoren

$$(w_1, w_2) = \left(\begin{pmatrix}-3\\4\\0\end{pmatrix}, \frac{4}{25}\begin{pmatrix}-4\\-3\\25\end{pmatrix}\right)$$

eine geordnete Orthogonalbasis von V. Wir berechnen die Normen der beiden Vektoren:

$$\begin{aligned}
\|w_1\| &= \left\|\begin{pmatrix}-3\\4\\0\end{pmatrix}\right\| = \sqrt{(-3)^2+4^2+0^2} = 5 \\
\|w_2\| &= \left\|\frac{4}{25}\begin{pmatrix}-4\\-3\\25\end{pmatrix}\right\| = \frac{4}{25}\left\|\begin{pmatrix}-4\\-3\\25\end{pmatrix}\right\| = \frac{4}{25}\sqrt{(-4)^2+(-3)^2+25^2} = \frac{4}{25}\sqrt{650}.
\end{aligned}$$

Folglich ist

$$\left(\frac{1}{\|w_1\|}w_1, \frac{1}{\|w_2\|}w_2\right) = \left(\frac{1}{5}\begin{pmatrix}-3\\4\\0\end{pmatrix}, \frac{1}{\sqrt{650}}\begin{pmatrix}-4\\-3\\25\end{pmatrix}\right)$$

eine geordnete Orthonormalbasis für V.

(b) Es sei $\mathbb{K} = \mathbb{C}$ und $V = \mathbb{C}^3$ mit dem komplexen Standardskalarprodukt. Gegeben sei der Vektor

$$w_1 := \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix}.$$

Gesucht ist nun eine Orthonormalbasis von $\mathbb{C}^3$, die w_1 als Vektor enthält.

Zuerst überprüfen wir, ob w_1 wirklich Norm 1 hat:

$$\|w_1\| = \left\| \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \right\| = \frac{1}{2}\left\| \begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \right\| = \frac{1}{2}\sqrt{|1+\mathrm{i}|^2 + |-\mathrm{i}|^2 + |1|^2} = \frac{1}{2}\sqrt{2+1+1} = 1.$$

Um nun mit Gram-Schmidt eine Orthogonalbasis von V zu konstruieren, benötigen wir zuerst eine Basis von V, die w_1 enthält. Dazu müssen wir nur zwei Vektoren angeben, sodass die drei Vektoren linear unabhängig sind. Das geht beispielsweise so:

$$v_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ und } v_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Nun ist

$$\mathsf{B} := \left(\frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

eine geordnete Basis für $\mathbb{C}^3$ und wir können Gram-Schmidt anwenden, um diese orthogonal zu machen: Als zweiter Vektor ergibt sich:

$$\begin{aligned}
w_2 &:= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \\
&= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \right\rangle}{1} \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \left(\frac{1}{2}\mathrm{i}\right) \cdot \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{4}\mathrm{i}\begin{pmatrix} 1+\mathrm{i} \\ -\mathrm{i} \\ 1 \end{pmatrix} \\
&= \frac{1}{4}\left(\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} \mathrm{i}-1 \\ 1 \\ \mathrm{i} \end{pmatrix} \right) \\
&= \frac{1}{4}\begin{pmatrix} 1-\mathrm{i} \\ 3 \\ -\mathrm{i} \end{pmatrix}
\end{aligned}$$

Bevor wir den dritten Vektor berechnen, bietet es sich an, $\langle w_2, w_2\rangle$ zu berechnen, da wir dies gleich benötigen:

$$\|w_2\|^2 = \langle w_2, w_2\rangle = \frac{1}{4^2}\left(|1-\mathrm{i}|^2+|3|^2+|-\mathrm{i}|^2\right) = \frac{1}{16}(2+9+1) = \frac{12}{16} = \frac{3}{4}.$$

Und der der dritte Vektor wird zu:

$$\begin{aligned}
w_3 &:= v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\
&= \begin{pmatrix}0\\0\\1\end{pmatrix} - \underbrace{\frac{\left\langle \begin{pmatrix}0\\0\\1\end{pmatrix}, \frac{1}{2}\begin{pmatrix}1+\mathrm{i}\\-\mathrm{i}\\1\end{pmatrix}\right\rangle}{1}}_{=\frac{1}{2}} \frac{1}{2}\begin{pmatrix}1+\mathrm{i}\\-\mathrm{i}\\1\end{pmatrix} - \underbrace{\frac{\left\langle \begin{pmatrix}0\\0\\1\end{pmatrix}, \frac{1}{4}\begin{pmatrix}1-\mathrm{i}\\3\\-\mathrm{i}\end{pmatrix}\right\rangle}{\frac{3}{4}}}_{=\frac{\mathrm{i}}{3}} \frac{1}{4}\begin{pmatrix}1-\mathrm{i}\\3\\-\mathrm{i}\end{pmatrix} \\
&= \begin{pmatrix}0\\0\\1\end{pmatrix} - \frac{1}{4}\begin{pmatrix}1+\mathrm{i}\\-\mathrm{i}\\1\end{pmatrix} - \frac{\mathrm{i}}{12}\begin{pmatrix}1-\mathrm{i}\\3\\-\mathrm{i}\end{pmatrix} \\
&= \frac{1}{12}\left(\begin{pmatrix}0\\0\\12\end{pmatrix} - \begin{pmatrix}3+3\mathrm{i}\\-3\mathrm{i}\\3\end{pmatrix} - \begin{pmatrix}1+\mathrm{i}\\3\mathrm{i}\\1\end{pmatrix}\right) \\
&= \frac{1}{12}\begin{pmatrix}-4-4\mathrm{i}\\0\\8\end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix}-1-\mathrm{i}\\0\\2\end{pmatrix}
\end{aligned}$$

Also ist folgendes eine Orthogonalbasis von $\mathbb{C}^3$, die den gegebenen Vektor w_1 enthält:

$$\left(\frac{1}{2}\begin{pmatrix}1+\mathrm{i}\\-\mathrm{i}\\1\end{pmatrix}, \frac{1}{4}\begin{pmatrix}1-\mathrm{i}\\3\\-\mathrm{i}\end{pmatrix}, \frac{1}{3}\begin{pmatrix}-1-\mathrm{i}\\0\\2\end{pmatrix}\right).$$

Der erste Vektor hat bereits Norm 1, wenn wir die anderen beiden noch normieren, erhalten wir eine Orthonormalbasis:

$$\left(\frac{1}{2}\begin{pmatrix}1+\mathrm{i}\\-\mathrm{i}\\1\end{pmatrix}, \frac{1}{\sqrt{12}}\begin{pmatrix}1-\mathrm{i}\\3\\-\mathrm{i}\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}-1-\mathrm{i}\\0\\2\end{pmatrix}\right).$$

Zusammenfassung von Abschnitt 1.3

(1) Ein Orthogonalsystem ist eine Menge aus Vektoren, die nicht 0 sind und paarweise orthogonal sind. Ein Orthonormalsystem hat zusätzlich die Eigenschaft, dass alle Vektoren Norm 1 haben.

(2) Orthonormalsysteme sind immer linear unabhängig.

(3) Eine Orthogonalbasis (-orthonormalbasis) ist ein Orthogonalsystem (-normalsystem), dass eine Basis eines (endlich dimensionalen) Vektorraums ist.

(4) Jeder endlich dimensionale Vektorraum mit Skalarprodukt besitzt eine Orthonormalbasis. Jedes Orthonormalsystem lässt sich zu einer Orthonormalbasis fortsetzen. Rechnerisch findet man diese beispielsweise mit dem Gram-Schmidt-Algorithmus.

1.4. Orthogonale Komplemente und Projektionen

Gegeben sei ein Vektorraum V mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und ein Untervektorraum $U \subseteq V$. Häufig trifft man in Anwendungen auf das Problem, dass ein Vektor $v \in V$ kein Element von U ist, aber man wissen möchte, wie weit v von U entfernt ist. Verbunden damit ist die Frage, ob es einen Punkt in U mit minimalem Abstand von V gibt und falls ja, wie man ihn findet. Diese Fragestellung bringt uns zum Begriff der orthogonalen Projektion und damit verbunden auch dem Orthogonalraum eines Punktes.

Definition 1.4.1 (Abstand eines Punktes von einer Teilmenge)**.**
Es sei (X,d) ein metrischer Raum, $A \subseteq X$ eine nichtleere Teilmenge und $x \in X$ ein Punkt.

Der *Abstand von x zu A* ist definiert als

$$d(x,A) := \inf\{d(x,a) \mid a \in A\} \in [0,+\infty).$$

Lemma 1.4.2.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Gegeben sei weiterhin ein Untervektorraum $U \subseteq V$ und ein Punkt $x \in V$.

Wenn es einen Punkt $p \in U$ gibt mit $\|x-p\| = d(x,U)$, dann ist dieser Punkt eindeutig.

Bemerkung 1.4.3. (1) Für normierte Vektorräume gilt dies im Allgemeinen nicht. Es ist wichtig, dass die Norm hier von einem Skalarprodukt induziert wird.

(2) Es ist für die Aussage nicht notwendig, dass U ein Untervektorraum ist. Es reicht, dass U eine *konvexe* Teilmenge ist. Für beliebige Teilmengen ist die Aussage falsch.

Proposition 1.4.4.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Gegeben sei weiterhin ein Untervektorraum $U \subseteq V$ und ein Punkt $x \in V$.

Für einen Punkt $p \in U$ sind die folgenden Aussagen äquivalent:

(i) Der Punkt p hat minimalen Abstand zu x, d.h. $\|x-p\| = d(x,U)$.

(ii) Der Verbindungsvektor $p-x$ steht senkrecht auf U, d.h.

$$\forall u \in U : \langle p-x, u\rangle = 0.$$

Satz 1.4.5 (Orthogonalprojektion auf endlich dimensionale Untervektorräume)**.**
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Gegeben sei weiterhin ein Untervektorraum $U \subseteq V$. Wir nehmen an, dass U endlich dimensional ist.
Dann gilt:

(a) Für jedes $x \in V$ gibt es einen eindeutigen Punkt $p \in U$ mit minimalem Abstand zu x. Für diesen Punkt schreiben wir $\pi_U(x) := p \in U$.

(b) Die Abbildung

$$\pi_U : V \to V, \quad x \mapsto \pi_U(x)$$

ist $\mathbb{K}$-linear. Sie heißt die Orthogonalprojektion *auf U.*[8]

(c) Falls $(b_1,\dots,b_n)$ eine Orthonormalbasis von U ist, dann lässt sich die Orthogonalprojektion auf U folgendermaßen schreiben:

$$\pi_U(x) = \sum_{j=1}^{n} \langle x, b_j\rangle b_j.$$

Definition 1.4.6 (Orthogonalraum)**.**
Es sei V ein Vektorraum mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$.

(a) Für jeden Vektor $v \in V$ definieren wir den *Orthogonalraum* von v als die Menge aller Vektoren, die orthogonal zu v sind. Hierfür schreibt man auch

$$\{v\}^\perp := \{w \in V \mid v \perp w\} \subseteq V.$$

(b) Für eine Teilmenge M von V definieren wir den *Orthogonalraum* von M als den Durchschnitt aller Orthogonalräume der einzelnen Elemente, d.h.

$$M^\perp := \bigcap_{v\in M} \{v\}^\perp = \{w \in V \mid \forall v \in M : v \perp w\}.$$

Es gilt $\emptyset^\perp = V$.

Lemma 1.4.7. *Es sei V ein Vektorraum mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$.*

(a) Für jedes $M \subseteq V$ ist $M^\perp$ ein Untervektorraum von V.

(b) Für jedes $M \subseteq V$ gilt: $M \cap M^\perp \subseteq \{0\}$.

(c) Es gilt: $\forall M_1 \subseteq M_2 : M_2^\perp \subseteq M_1^\perp$.

(d) Es gilt: $\forall M \subseteq V : (\mathrm{LH}_{\mathbb{K}}(M))^\perp = M^\perp$.

(e) Es gilt: $\forall M \subseteq V : M \subseteq (M^\perp)^\perp$ und $M^\perp = ((M^\perp)^\perp)^\perp$.

[8] Meistens wird als Zielbereich der Abbildung π_U der ganze Raum V gewählt, obwohl das Bild nur die Teilmenge U ist. Dies hat den Vorteil, dass π_U auf diese Weise ein Endomorphismus von V ist.

Der Orthogonalraum ist für beliebige Teilmengen definiert, aber im Fall, dass die Teilmenge ein *endlich dimensionaler* Untervektorraum ist, ist er besonders sinnvoll:

Satz 1.4.8 (Orthogonale Zerlegung eines Vektorraums mit Skalarprodukt)**.**
Es sei V ein Vektorraum mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Falls $U \subseteq V$ ein endlich dimensionaler *Untervektorraum von V ist, gilt: Der Untervektorraum $U^\perp$ ist ein komplementärer Vektorraum, d.h.*

$$V = U \oplus U^\perp,$$

jeder Vektor $v \in V$ besitzt also eine eindeutige Zerlegung $v = u + w$ mit $u \in U$ und $w \in U^\perp$.

Man spricht in diesem Fall auch von dem orthogonalen Komplement *von U.*

Für ein $v \in V$ lässt sich diese eindeutige Zerlegung folgendermaßen aufschreiben:

$$v = \underbrace{\pi_U(v)}_{\in U} + \underbrace{(v - \pi_U(v))}_{\in U^\perp}.$$

Außerdem gilt: $U^\perp = \ker(\pi_U)$ und $(U^\perp)^\perp = U$.

Die Zerlegung von V in die direkte Summe von U und $U^\perp$ ist kompatibel mit dem Skalarprodukt:

$$\forall u_1, u_2 \in U; w_1, w_2 \in U^\perp : \langle u_1 + w_1, u_2 + w_2 \rangle = \langle u_1, u_2 \rangle + \langle w_1, w_2 \rangle .$$

Bemerkung 1.4.9. (1) Erinnerung: Wir haben in LA1 gesehen, dass Komplemente nicht eindeutig sind, sondern dass ein Untervektorraum viele verschiedene Komplemente haben kann. Diese Situation wird viel einfacher, sobald wir ein Skalarprodukt gewählt haben. Das *orthogonale* Komplement ist eindeutig.

(2) Wenn der Untervektorraum U unendlich dimensional ist, dann gibt es Fälle, in denen ein orthogonales Komplement existiert und solche, in denen es nicht existiert. Die Bedingung der endlichen Dimension ist also hinreichend für ein orthogonales Komplement, aber nicht notwendig. Eine notwendige und hinreichende Bedingung ist die *Vollständigkeit*, d.h. jede Cauchy-Folge in U konvergiert auch in U. Dies führt uns aber zu sehr in die Analysis und weg von der linearen Algebra, sodass wir dies hier nicht weiter verfolgen wollen.

2. Isometrien und Homomorphismen

2.1. Isometrien und isometrische Isomorphismen

In diesem Kapitel möchten wir nun Abbildungen zwischen Räumen mit Skalarprodukt untersuchen. Wir beginnen mit einer ganz allgemeinen Definition einer Abbildung, die Abstände erhält:

Definition 2.1.1 (Isometrie)**.**
Es sei $f : X \to Y$ eine Abbildung zwischen einem metrischen Raum (X, d_X) und einem metrischen Raum (Y, d_Y). Die Abbildung f heiße *Isometrie* (oder *isometrische Einbettung*), wenn sie Abstände erhält, d.h. wenn gilt:

$$\forall x_1, x_2 \in X : \ d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2).$$

Bemerkung 2.1.2. Es ist leicht zu sehen, dass eine Isometrie immer automatisch injektiv ist, aber im Allgemeinen nicht surjektiv sein muss. Falls aber eine Isometrie surjektiv ist, so ist die Umkehrabbildung auch eine Isometrie.

Achtung: In der Literatur wird der Begriff *Isometrie* manchmal auch so definiert, dass die Abbildung bijektiv ist, d.h. die Surjektivität wird als Teil der Definition genommen. Das, was wir hier *Isometrie* nennen, ist dann dort eine *isometrische Einbettung*. Die Unterscheidung wird in dieser Veranstaltung keine allzu große Rolle spielen, weil viele der Isometrien, die wir uns anschauen werden, automatisch surjektiv sind.

Beispiel 2.1.3. Es sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt und der daraus resultierenden Metrik und $v_0 \in V$ ein Vektor. Die *Translation* um v_0 ist die Abbildung

$$\tau_{v_0} : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x + v_0,$$

die jeden Punkt um den Vektor v_0 verschiebt. Eine solche Translation ist eine Isometrie, weil man leicht nachrechnet:

$$\left\| \tau_{v_0}(x_1) - \tau_{v_0}(x_2) \right\| = \|(x_1 + v_0) - (x_2 + v_0)\| = \|x_1 - x_2\|.$$

Für $v_0 \neq 0$ ist τ_{v_0} keine lineare Abbildung, weil 0 nicht auf 0 abgebildet wird.

Wir wollen im Folgenden vor allem Isometrien untersuchen, die auch linear sind.

Wir benötigen die folgende Aussage über komplexe Skalarprodukte, um die nächste Proposition zu beweisen:

Lemma 2.1.4 (Komplexe Polarisierungsidentität)**.** *Es sei V ein unitärer Vektorraum (also ein komplexer Vektorraum mit Skalarprodukt). Dann gilt:*

$$\langle v, w \rangle = \frac{1}{4} \sum_{k=0}^{3} \mathrm{i}^k \left\| v + \mathrm{i}^k w \right\|^2.$$

Beweis. Wir rechnen – nur unter Verwendung der Linearität in der ersten Komponente und der konjugierten Linearität in der zweiten Komponente – nach:

$$\begin{aligned}
\frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left\| v + \mathrm{i}^k w \right\|^2 &= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left\langle v + \mathrm{i}^k w, v + \mathrm{i}^k w \right\rangle \\
&= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left(\left\langle v, v + \mathrm{i}^k w \right\rangle + \mathrm{i}^k \left\langle w, v + \mathrm{i}^k w \right\rangle \right) \\
&= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left(\langle v, v\rangle + (-\mathrm{i})^k \langle v, w\rangle + \mathrm{i}^k \langle w, v\rangle + \mathrm{i}^k (-\mathrm{i})^k) \langle w, w\rangle \right) \\
&= \frac{1}{4}\sum_{k=0}^{3} \left(\mathrm{i}^k \langle v, v\rangle + \langle v, w\rangle + (-1)^k \langle w, v\rangle + \mathrm{i}^k \langle w, w\rangle \right) \\
&= \underbrace{\frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k}_{=0} \langle v, v\rangle + \frac{1}{4}\sum_{k=0}^{3} \langle v, w\rangle + \underbrace{\frac{1}{4}\sum_{k=0}^{3} (-1)^k}_{=0} \langle w, v\rangle + \underbrace{\frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k}_{=0} \langle w, w\rangle = \langle v, w\rangle .
\end{aligned}$$

□

Bemerkung 2.1.5. Wie man am Beweis erkennen kann, gilt Lemma 2.1.4 analog für beliebige sesquilineare Abbildungen, d.h. die positive Definitheit oder die konjugierte Symmetrie wird nicht benötigt.

Proposition 2.1.6 (Lineare Isometrien)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ Vektorräume über $\mathbb{K}$ mit Skalarprodukt. Die dazugehörigen Normen und Metriken werden mit $\|\cdot\|_V, \|\cdot\|_W$, bzw. d_V, d_W bezeichnet.

Für eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$ sind äquivalent:

(i) φ ist eine Isometrie, also erhält Abstände, d.h.

$$\forall x_1, x_2 \in V : d_W\left(\varphi(x_1), \varphi(x_2)\right) = d_V(x_1, x_2).$$

(ii) φ ist längenerhaltend, d.h.

$$\forall v : \left\|\varphi(v)\right\|_W = \|v\|_V .$$

(iii) φ erhält das Skalarprodukt, d.h.

$$\forall v_1, v_2 \in V : \left\langle \varphi(v_1), \varphi(v_2)\right\rangle_W = \langle v_1, v_2\rangle_V .$$

Beweis. (iii) $\Longrightarrow$ (ii):
Es gilt für jedes $v \in V$:

$$\left\|\varphi(v)\right\|_W^2 = \left\langle \varphi(v), \varphi(v)\right\rangle_W = \langle v, v\rangle_V = \|v\|_V^2 .$$

(ii) $\Longrightarrow$ (i):
Unter Verwendung der Linearität von φ können wir für $x, y \in V$ berechnen:

$$d_W\left(\varphi(x), \varphi(y)\right) = \left\|\varphi(x) - \varphi(y)\right\|_W = \left\|\varphi(x - y)\right\|_W = \|x - y\|_V = d_V(x, y).$$

(i) $\Longrightarrow$ (ii):
Weil φ linear ist, gilt $\varphi(0_V) = 0_W$. Also gilt für jedes $v \in V$:

$$\left\|\varphi(v)\right\|_W = \left\|\varphi(v) - 0\right\|_W = \left\|\varphi(v) - \varphi(0)\right\|_W = d_W\left(\varphi(v), \varphi(0)\right) = d_V(v, 0) = \|v - 0\|_V = \|v\|_V .$$

(ii) $\Longrightarrow$ (iii):
Wir unterscheiden zwei Fälle: Falls $\mathbb{K} = \mathbb{R}$ ist, gilt mit der binomischen Formel (1.2.7(i)), dass

$$\langle v_1, v_2 \rangle = \frac{1}{2}\left(\|v_1 + v_2\|^2 - \|v_1\|^2 - \|v_2\|^2\right).$$

Diese Formel wenden wir nun einmal im Zielbereich W und einmal im Definitionsbereich V an und verwenden außerdem die Linearität von φ:

$$\begin{aligned}
\langle \varphi(v_1), \varphi(v_2) \rangle_W &= \frac{1}{2}\left(\left\|\varphi(v_1) + \varphi(v_2)\right\|_W^2 - \left\|\varphi(v_1)\right\|_W^2 - \left\|\varphi(v_2)\right\|_W^2\right) \\
&= \frac{1}{2}\left(\left\|\varphi(v_1 + v_2)\right\|_W^2 - \left\|\varphi(v_1)\right\|_W^2 - \left\|\varphi(v_2)\right\|_W^2\right) \\
&= \frac{1}{2}\left(\|v_1 + v_2\|_V^2 - \|v_1\|_V^2 - \|v_2\|_V^2\right) \\
&= \langle v_1, v_2 \rangle_V .
\end{aligned}$$

Im Fall, dass $\mathbb{K} = \mathbb{C}$ ist, verwenden wir nicht die binomische Formel, sondern stattdessen die Polarisierungsidentität (Lemma 2.1.4):

$$\begin{aligned}
\langle \varphi(v_1), \varphi(v_2) \rangle_W &= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left\|\varphi(v_1) + \mathrm{i}^k \varphi(v_2)\right\|_W^2 \\
&= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left\|\varphi(v_1 + \mathrm{i}^k v_2)\right\|_W^2 \\
&= \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \left\|v_1 + \mathrm{i}^k v_2\right\|_V^2 \\
&= \langle v_1, v_2 \rangle_V .
\end{aligned}$$

□

Bemerkung 2.1.7. In Proposition 2.1.6 und dem dazugehörigen Beweis haben wir immer fein säuberlich zwischen dem Skalarprodukt $\langle , \rangle_V$ im Definitionsbereich und dem Skalarprodukt $\langle \cdot, \cdot \rangle_W$ im Zielbereich unterschieden. Dies ist normalerweise unüblich. Meistens schreibt man einfach nur $\langle \cdot, \cdot \rangle$ und es ist aus dem Kontext klar, was gemeint ist. Das gleiche gilt für Normen und Metriken.

Definition 2.1.8.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Es seien weiterhin Vektorräume $(V, \langle \cdot, \cdot \rangle)$ und $(W, \langle \cdot, \cdot \rangle)$ mit Skalarprodukt gegeben.

Eine lineare Abbildung $\varphi : V \to W$ heißt *lineare Isometrie*, wenn sie die äquivalenten Eigenschaften von Proposition 2.1.6 erfüllt. Insbesondere erhält sie das Skalarprodukt.

Ist φ zusätzlich bijektiv, also ein Isomorphismus von Vektorräumen, so nennt man φ einen *isometrischen Isomorphismus*.

Die Vektorräume mit Skalarprodukt $(V, \langle \cdot, \cdot \rangle)$ und $(W, \langle \cdot, \cdot \rangle)$ heißen *isometrisch isomorph*, wenn es einen isometrischen Isomorphismus zwischen ihnen gibt.

Bemerkung 2.1.9. Es sei $\mathbb{K}$ ein beliebiger Körper, z.B. $\mathbb{R}$ oder $\mathbb{C}$. In LA1 haben wir gesehen, dass man alle endlich dimensionalen $\mathbb{K}$-Vektorräume bis auf Isomorphie klassifizieren kann: Jeder endlich dimensionale Vektorraum V ist isomorph zu genau einem Vektorraum von der Form $\mathbb{K}^n$.

Wir sagen: *Bis auf Isomorphie* ist jeder endlich dimensionale Vektorraum ein Raum von Spaltenvektoren.

Dieser Isomorphismus ist allerdings im Allgemeinen nicht eindeutig. Außerdem erhält dieser Isomorphismus erst einmal nur die Vektorraumstruktur, also Vektoraddition und skalare Multiplikation und im Allgemeinen keine andere Struktur, die eventuell vorhanden ist. Wir wollen nun eine analoge Klassifikation für Vektorräume mit Skalarprodukt beweisen:

Satz 2.1.10 (Klassifikation aller endlich dimensionaler Vektorräume mit Skalarprodukt)**.** *Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}$. Dann gilt: Jeder n-dimensionale Vektorraum V mit Skalarprodukt ist* isometrisch isomorph *zu $\mathbb{K}^n$ mit dem Standardskalarprodukt.*

Man kann also sagen: Bis auf isometrische Isomorphie *ist jeder endlich dimensionale Vektorraum mit Skalarprodukt ein Raum von Spaltenvektoren mit dem Standardskalarprodukt.*

Beweis. Weil V endlich dimensional ist, können wir mit Gram-Schmidt eine Orthonormalbasis

$$\mathsf{B} = (b_1, \ldots, b_n)$$

finden (Korollar 1.3.8).

Weil B eine Basis ist, folgt, dass die Koordinatenabbildung

$$(\cdot)_{\mathsf{B}} : V \to \mathbb{K}^n, \quad v = \sum_{j=1}^{n} \alpha_j b_j \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

wohldefiniert und ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist.

Es bleibt zu zeigen, dass es sich um einen isometrischen Isomorphismus handelt, wenn wir im Definitionsbereich V das gegebene Skalarprodukt wählen und im Zielbereich das Standardskalarprodukt. Es seien dazu $v_1, v_2 \in V$ gegeben mit Koordinaten

$$v_1 = \sum_{j=1}^{n} \alpha_j b_j \quad \text{und} \quad v_2 = \sum_{k=1}^{n} \beta_k b_k$$

gegeben. Dann gilt:

$$\begin{aligned} \langle v_1, v_2 \rangle &= \left\langle \sum_{j=1}^{n} \alpha_j b_j, \sum_{k=1}^{n} \beta_k b_k \right\rangle \\ &= \sum_{j=1}^{n} \alpha_j \sum_{k=1}^{n} \overline{\beta_k} \underbrace{\langle b_j, b_k \rangle}_{=0 \text{ für } k \neq j} \\ &= \sum_{j=1}^{n} \alpha_j \overline{\beta_j} \\ &= \left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle. \end{aligned}$$ □

In gewisser Weise können wir uns also beim Studium von endlich dimensionalen Vektorräumen mit Skalarprodukt auf den Raum $\mathbb{K}^n$ mit dem Standardskalarprodukt zurückziehen.

Wir haben in Beispiel 2.1.3 gesehen, dass Translationen Isometrien sind, die nicht linear sind (außer wir verschieben um den Nullvektor).

Weitere Beispiele erhält durch Verketten einer Translation und einer linearen Isometrie. Man kann sich fragen, ob es noch weitere nichtlineare Isometrien gibt und es stellt sich heraus, dass zumindest über den reellen Zahlen keine weiteren existieren:

Satz 2.1.11.
Wenn V und W euklidische Vektorräume sind, dann ist jede Isometrie $f: V \to W$ eine Verkettung einer linearen Isometrie $\varphi: V \to W$ und einer Verschiebung $\tau_w: W \to W$. Insbesondere ist eine Isometrie, die 0_V auf 0_W abbildet, immer eine lineare Abbildung.

Um diesen Satz zu beweisen, wollen wir das folgende Lemma verwenden:

Lemma 2.1.12.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(V, \langle\cdot,\cdot\rangle)$ sei ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Weiterhin seien $a, b, x \in V$ gegeben. Dann sind die folgenden Aussagen äquivalent:

(i) $x = \frac{1}{2}(a+b)$.

(ii) $d(x,a) = d(x,b) = \frac{1}{2}d(a,b)$.

Man beachte, dass die erste Bedingung ausschließlich die Vektorraumstruktur und nicht das Skalarprodukt verwendet, während die zweite Bedingung ausschließlich die Metrik verwendet und keine Vektoraddition oder skalare Multiplikation.

Mit Lemma 2.1.12 können wir nun Satz 2.1.11 beweisen.

2.2. Die orthogonale und die unitäre Gruppe

Notation 2.2.1. Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}$. Gegeben sei eine Matrix

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \in \mathbb{K}^{m\times n}.$$

Die zu A gehörige *komplex konjugierte Matrix* $\overline{A}$ ist einfach die Matrix, die man erhält, wenn man jeden einzelnen Eintrag komplex konjugiert, d.h.

$$\overline{A} = \begin{pmatrix} \overline{a_{1,1}} & \cdots & \overline{a_{1,n}} \\ \vdots & \ddots & \vdots \\ \overline{a_{m,1}} & \cdots & \overline{a_{m,n}} \end{pmatrix} \in \mathbb{K}^{m\times n}.$$

Im Falle $\mathbb{K} = \mathbb{R}$ gilt natürlich $\overline{A} = A$.

Die zu A gehörige *adjungierte Matrix* $A^* := \overline{A^\top} = \overline{A}^\top$ ist einfach die Matrix, die man erhält, wenn man die Matrix zuerst transponiert und dann jeden einzelnen Eintrag komplex konjugiert, d.h.

$$A^* = \begin{pmatrix} \overline{a_{1,1}} & \cdots & \overline{a_{m,1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1,n}} & \cdots & \overline{a_{m,n}} \end{pmatrix} \in \mathbb{K}^{n\times m}.$$

Im Falle $\mathbb{K} = \mathbb{R}$ gilt natürlich $A^* = A^\top$.

Achtung: In LA1 haben wir im Zusammenhang mit der Cramerschen Regel die *Adjunkte* einer Matrix kennengelernt. Dies hat **nichts** zu tun mit dieser hier eingeführten *Adjungierten*. Dies ist einfach nur ein unglückliches Zusammentreffen von Bezeichnungen, die sich leider beide eingebürgert haben.

Für Vektoren $v, w \in \mathbb{K}^n$ lässt sich das Standardskalarprodukt schreiben als:

$$\langle v, w \rangle = w^* v,$$

wobei das Produkt zwischen $w^* \in \mathbb{K}^{1\times n}$ und $v \in \mathbb{K}^n = \mathbb{K}^{n\times 1}$ das gewöhnliche Matrixprodukt ist.

Proposition 2.2.2.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}$. Für eine Matrix $A \in \mathbb{K}^{n\times n}$ sind äquivalent:

(i) Die Abbildung $\mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ ist ein isometrischer Isomorphismus bezüglich des Standardskalarproduktes auf $\mathbb{K}^n$.

(ii) Die Abbildung $\mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ ist eine Isometrie bezüglich des Standardskalarproduktes auf $\mathbb{K}^n$.

(iii) Die Spalten der Matrix A bilden eine Orthonormalbasis von $\mathbb{K}^n$.

(iv) Die Spalten der Matrix A bilden ein Orthonormalsystem von $\mathbb{K}^n$.

(v) Es gilt $A^*A = \mathbb{1}_n$.

(vi) Die Matrix A ist invertierbar und $A^{-1} = A^*$.

Definition 2.2.3.
Es sei $\mathbb{K} = \mathbb{R}$. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heiße *orthogonale Matrix*, wenn sie die Eigenschaften aus Proposition 2.2.2 erfüllt, insbesondere ist eine Matrix orthogonal, wenn $A^\top A = \mathbb{1}_n$ oder wenn die Spalten eine Orthonormalbasis von $\mathbb{R}^n$ bilden.

Die Menge aller orthogonalen $(n \times n)$-Matrizen bezeichnet man mit

$$\mathrm{O}(n) := \left\{A \in \mathbb{R}^{n\times n} \;\middle|\; A^\top A = \mathbb{1}_n\right\}$$

und nennt sie die *orthogonale Gruppe*.

Oft interessiert man sich nicht für alle orthogonalen Matrizen, sondern nur für solche mit Determinante 1. Diese Teilmenge erhält ihre eigene Bezeichnung:

$$\mathrm{SO}(n) := \{A \in \mathrm{O}(n) \mid \det(A) = 1\}$$

und wird *spezielle orthogonale Gruppe* genannt.

Bemerkung 2.2.4. (i) Beachten Sie: Eine Matrix A ist genau dann eine ortho*gonale* Matrix, wenn die Spalten der Matrix eine Ortho*normal*basis bilden. Diese Bezeichnung ist von daher ein bisschen ungünstig. Man würde vielleicht eher erwarten, dass man eine solche Matrix ortho*normale* Matrix nennt, aber diesen Begriff gibt es nicht. Es reicht also nicht aus, dass die Spalten einer Matrix ein Orthogonalsystem bilden, es ist immer auch notwendig, dass alle Spalten Norm 1 haben.

(ii) Die orthogonale Gruppe $\mathrm{O}(n)$ ist auch wirklich eine Gruppe. Genau gesagt ist sie eine Untergruppe der *allgemeinen linearen Gruppe* $\mathrm{GL}(n, \mathbb{R})$, also der Menge aller invertierbaren Matrizen. Die Gruppe $\mathrm{SO}(n)$ ist der Durchschnitt der Gruppe $\mathrm{SL}(n, \mathbb{R})$ der Matrizen mit Determinante 1 und der Gruppe $\mathrm{O}(n)$, also auch eine Untergruppe.

Bevor wir den komplexen Fall betrachten, wollen wir nun ganz speziell den einfachsten reellen Fall untersuchen, der interessant ist[1]:

Beispiel 2.2.5. Wir wollen versuchen, alle orthogonalen (2×2)-Matrizen zu finden und uns überlegen, was diese geometrisch bedeuten:

Gegeben sei eine Matrix $A \in \mathrm{O}(2)$. Dann gilt:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ mit } a, b, c, d \in \mathbb{R}.$$

Es gilt: die Spalten $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ bilden eine Orthonormalbasis von $\mathbb{R}^2$ mit dem Standardskalarprodukt. Insbesondere hat $\begin{pmatrix} a \\ c \end{pmatrix}$ Norm 1, es gilt also:

$$a^2 + c^2 = 1.$$

Somit muss der Vektor von der Form $\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$ sein[2].

Der zweite Spaltenvektor $\begin{pmatrix} b \\ d \end{pmatrix}$ steht nun senkrecht auf dem ersten und liegt somit im Orthogonalraum $\left\{ \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \right\}^{\perp}$.

Der Vektor $\begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$ ist orthogonal zum ersten Spaltenvektor und spannt somit den Orthogonalraum auf. Der gesuchte Vektor $\begin{pmatrix} b \\ d \end{pmatrix}$ ist somit ein Vielfaches von $\begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$. Da die zweite Spalte außerdem noch Länge 1 hat, bleiben genau zwei Fälle:

Fall I: $\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$:

In diesem Fall ist die Matrix A gegeben durch:

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

In den Spalten der Matrix stehen die Bilder der Basisvektoren. Somit sieht man, dass die durch A gegebene Abbildung jeden Basisvektor um den Winkel α gegen den Uhrzeigersinn dreht.

Eine solche Matrix nennt man auch ein *Drehkästchen*. Wir werden später sehen, dass sich jede orthogonale Matrix durch die Wahl einer geeigneten Orthonormalbasis von $\mathbb{R}^n$ in eine Form bringen lässt, die im Wesentlichen aus solchen Drehkästchen besteht (Satz 2.6.11).

Die Determinante eines solchen Drehkästchens ist immer

$$\det \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} = (\cos(\alpha))^2 + (\sin(\alpha))^2 = 1.$$

Die Matrix A gehört somit zur speziellen orthogonalen Gruppe, es gilt also: $A \in \mathrm{SO}(2)$.

[1]Die Gruppen $\mathrm{O}(1)$ und $\mathrm{SO}(1)$ sind zwar noch einfacher zu verstehen, aber nicht so interessant…

[2]Dies sieht man entweder direkt geometrisch, rechnet es von Hand nach oder geht den Umweg über die komplexen Zahlen: Da $a^2 + b^2 = 1$ ist, hat die Zahl $z = a + ib$ Betrag 1, liegt also auf dem Einheitskreis und somit gilt: $a + \mathrm{i}b = e^{\mathrm{i}\alpha} = \cos(\alpha) + \mathrm{i}\sin(\alpha)$.

Fall II: $\begin{pmatrix} b \\ d \end{pmatrix} = -\begin{pmatrix} \sin(\alpha) \\ -\cos(\alpha) \end{pmatrix}$:
In diesem Fall ist die Matrix A gegeben durch:

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Dies ist nun keine Drehung mehr. Wir werden in Übung oder Tutorium sehen, dass es sich hierbei um eine Spiegelung an einer Ursprungsgerade handelt.

Die Determinante in diesem Fall ist nun nicht mehr 1, sondern

$$\det \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} = -\left((\cos(\alpha))^2 + (\sin(\alpha))^2\right) = -1.$$

Zusammengefasst kann man also sagen: Eine Matrix $A \in \mathrm{O}(2)$ hat entweder Determinante 1 oder -1. Hat sie Determinante 1, so beschreibt sie eine Drehung um den Ursprung, hat diese spezielle Kosinus-Sinus-Form und heißt Drehkästchen. Hat sie Determinante -1, so beschreibt sie eine Spiegelung an einer Ursprungsgeraden.

Lemma 2.2.6.
Die Determinante einer reellen orthogonalen Matrix ist immer 1 *oder* -1*. Somit ist die Menge* $\mathrm{O}(n)$ *die disjunkte Vereinigung von*

$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det(A) = 1\} \quad \textit{und} \quad \mathrm{O}(n) \setminus \mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det(A) = -1\}.$$

Die Matrizen in $\mathrm{SO}(n)$ *entsprechen den* orientierungserhaltenden Isometrien*, die anderen den* orientierungsumkehrenden Isometrien*.*

Beweis. Es gilt:

$$(\det(A))^2 = \det(A)\det(A) = \det\left(A^\top\right)\det(A) = \det\left(A^\top A\right) = \det(\mathbb{1}_n) = 1.$$

Also ist $\det(A)$ eine reelle Nullstelle des Polynoms $X^2 - 1$ und muss somit 1 oder -1 sein. □

Definition 2.2.7.
Es sei $\mathbb{K} = \mathbb{C}$. Eine Matrix $A \in \mathbb{C}^{n\times n}$ heiße *unitäre Matrix*, wenn sie die Eigenschaften aus Proposition 2.2.2 erfüllt, insbesondere ist eine Matrix unitär, wenn $A^*A = \mathbb{1}_n$ oder wenn die Spalten eine Orthonormalbasis von $\mathbb{C}^n$ bilden.

Die Menge aller unitären $(n \times n)$-Matrizen bezeichnet man mit

$$\mathrm{U}(n) := \left\{A \in \mathbb{C}^{n\times n} \,\middle|\, A^*A = \mathbb{1}_n\right\}$$

und nennt sie die *unitäre Gruppe*.

Wie im reellen Fall ist man oft nur an solchen mit Determinante 1 interessiert:

$$\mathrm{SU}(n) := \{A \in \mathrm{U}(n) \mid \det(A) = 1\}.$$

Dies ist die *spezielle unitäre Gruppe*.

Direkt aus der Definition folgt:

$$\mathrm{O}(n) = \mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R}).$$

Analog zu Lemma 2.2.6 zeigt man folgendes Lemma:

Lemma 2.2.8.
Die Determinante einer komplexen unitären Matrix hat immer Betrag 1*. Somit ist die Determinantenabbildung auf den unitären Matrizen ein Gruppenhomomorphismus von der Gruppe* $\mathrm{U}(n)$ *in die Gruppe* $\mathrm{U}(1)$*:*

$$\mathrm{U}(n) \to \mathrm{U}(1), \quad A \mapsto \det(A).$$

Man beachte, dass $\mathrm{U}(1) = \{z \in \mathbb{C} \mid \overline{z}z = 1\}$ *nichts anderes als der Einheitskreis in* $\mathbb{C}$ *ist. Die Gruppe* $\mathrm{SU}(n)$ *ist der Kern dieses Gruppenhomomorphismus und somit eine Untergruppe von* $\mathrm{U}(n)$*.*

Bemerkung 2.2.9 (Ein würfelförmiges Diagramm mit Matrizengruppen)**.** Mit den aus den aus der LA1 bekannten Notationen:

$$\mathrm{GL}(n,\mathbb{K}) = \left\{A \in \mathbb{K}^{n\times n} \,\middle|\, A \text{ ist invertierbar}\right\}$$

$$\mathrm{SL}(n,\mathbb{K}) = \left\{A \in \mathbb{K}^{n\times n} \,\middle|\, \det(A) = 1\right\}$$

ergeben sich nun für jede natürliche Zahl $n \in \mathbb{N}$ acht unterschiedliche Gruppen, die man in folgendem Diagramm anordnen kann:

$$\begin{tikzcd}
 & & \mathrm{GL}(n,\mathbb{C}) & \\
\mathrm{GL}(n,\mathbb{R}) \arrow[urr] & & & \\
 & & & \mathrm{SL}(n,\mathbb{C}) \arrow[uul] \\
 & \mathrm{SL}(n,\mathbb{R}) \arrow[uul] \arrow[urr] & & \\
 & & \mathrm{U}(n) \arrow[uuuu] & \\
\mathrm{O}(n) \arrow[uuuu] \arrow[urr] & & & \\
 & & & \mathrm{SU}(n) \arrow[uul] \arrow[uuuu] \\
 & \mathrm{SO}(n) \arrow[uul] \arrow[uuuu] \arrow[urr] & &
\end{tikzcd}$$

Dieses Diagramm ist so zu lesen, dass A in B enthalten ist, wenn ein Pfeil von A nach B zeigt. Alle diese Gruppen sind also Untergruppen von $\mathrm{GL}(n,\mathbb{C})$ und enthalten $\mathrm{SO}(n)$ als Untergruppe.

Die linke Seite des „Würfels“ steht für die Matrizengruppen mit reellen Einträgen, die rechte Seite für die mit komplexen Einträgen.

Die „Vorderseite“ enthält die vier Gruppen, die nur Matrizen mit Determinante 1 enthalten, während die „Hinterseite“ keine Determinanteneinschränkung hat.

Die oberen vier Gruppen bestehen aus Matrizen, die nicht nur Isometrien enthalten, während die unteren vier Gruppen nur Isometrien enthalten.

Orthogonale und unitäre Matrizen spielen nicht nur beim Studium von $\mathbb{K}^n$ mit dem Standardskalarprodukt eine Rolle, sondern sind essenziell für das Verständnis von Vektorräumen mit Skalarprodukt im Allgemeinen:

Proposition 2.2.10.
Für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ seien zwei endlich dimensionale $\mathbb{K}$-Vektorräume V und W mit Skalarprodukt gegeben. Wir nehmen an, V habe eine Orthonormalbasis B und W eine Orthonormalbasis C. Dann gilt für eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$, dass die folgenden Aussagen äquivalent sind:

(i) Die Abbildung $\varphi : V \to W$ ist ein isometrischer Isomorphismus.

(ii) Die Matrix $\mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi)$ ist eine unitäre Matrix (im Fall $\mathbb{K} = \mathbb{R}$ also eine orthogonale).

Wir haben in LA1 gesehen, dass Basiswechsel immer einer Matrixmultiplikation aus der Gruppe $\mathrm{GL}(n, \mathbb{K})$ entsprechen. Wenn wir uns auf Vektorräume mit Skalarprodukt beschränken und nur Orthonormalbasen betrachten, so sind die Basiswechselmatrizen immer aus der Gruppe $\mathrm{U}(n)$ (im komplexen Fall), bzw. aus der Gruppe $\mathrm{O}(n)$ (im reellen Fall).

2.3. Unitäre Trigonalisierung und Diagonalisierung

Wir haben in LA1 definiert, dass zwei quadratische Matrizen $A, B \in \mathbb{K}^{n\times n}$ *ähnlich* sind, wenn sie den selben Endomorphismus darstellen, nur bezüglich unterschiedlicher Basen. In Formeln bedeutet dies, dass A und B ähnlich sind, wenn es eine Basiswechselmatrix $S \in \mathrm{GL}(n, \mathbb{K})$ gibt mit

$$B = SAS^{-1}.$$

Wenn wir nun fordern, dass die Basiswechselmatrix nicht aus der $\mathrm{GL}(n, \mathbb{K})$, sondern aus der $\mathrm{O}(n)$, bzw. der $\mathrm{U}(n)$ stammt, erhalten wir die Begriffe *orthogonal ähnlich* und *unitär ähnlich*. Dies entspricht eines Basiswechsels von einer Orthonormalbasis in eine andere.

Eine Matrix, die orthogonal (unitär) ähnlich zu einer Diagonalmatrix ist heißt *orthogonal diagonalisierbar* (*unitär diagonalisierbar*). Analog nennt man eine Matrix, die orthogonal (unitär) ähnlich zu einer oberen Dreiecksmatrix ist *orthogonal trigonalisierbar* (*unitär trigonalisierbar*).

Mittelfristiges Ziel wird es nun sein, Kriterien aufzustellen, welche Matrizen orthogonal/unitär diagonalisierbar sind.

Beispiel 2.3.1. Es sei

$$A := \begin{pmatrix} 1 & 42 \\ 0 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Wir können die Eigenwerte direkt ablesen, es sind 1 und 2. Somit hat eine (2×2)-Matrix zwei verschiedene Eigenwerte. Aus LA1 wissen wir, dass dies hinreichend ist, dass die Matrix diagonalisierbar ist. Genauer: Es gibt eine Basis aus Eigenvektoren $\mathsf{B} = (b_1, b_2)$ von $\mathbb{R}^2$, sodass die Matrix bezüglich dieser Basis

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

ist. Die Frage ist nun: Ist A auch orthogonal diagonalisierbar, d.h. gibt es eine Orthonormalbasis aus Eigenvektoren?

Nehmen wir einmal an, es gäbe so eine Orthonormalbasis $\mathsf{C} = (b_1, b_2)$. Ohne Beschränkung der Allgemeinheit nehmen wir an, b_1 sei ein Eigenvektor zum Eigenwert $\lambda = 1$. Dann muss b_1

ein Vielfaches der ersten Standardbasisvektors $e_1 = \begin{pmatrix}1\\0\end{pmatrix}$ sein, da der Eigenraum zum Eigenwert $\lambda = 1$ aus genau den Vielfachen von e_1 besteht. Da b_2 senkrecht auf b_1 steht, muss b_2 im Orthogonalraum zu e_1 liegen und somit ein Vielfaches von e_2 sein. Nun ist aber e_2 kein Eigenvektor von A, weil e_2 nicht auf ein Vielfaches von sich selbst abgebildet wird.

Wir sehen also: Es gibt Matrizen, die sind diagonalisierbar, aber nicht orthogonal diagonalisierbar.

Die Überlegungen aus diesem Beispiel können wir allgemein festhalten:

Lemma 2.3.2 (Orthogonale/Unitäre Diagonalisierung)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Für einen Endomorphismus $\varphi : V \to V$ sind äquivalent:

(i) *Der Endomorphismus ist* orthogonal/unitär diagonalisierbar*, d.h. Es gibt eine Orthonormalbasis B von V, sodass $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ eine Diagonalmatrix ist.*

(ii) *Es gibt eine Orthonormalbasis B, die aus Eigenvektoren von φ besteht.*

(iii) *Der Endomorphismus φ ist diagonalisierbar und je zwei Eigenräume sind orthogonal zueinander.*

Beispiel 2.3.1 zeigt uns, dass orthogonale Diagonalisierbarkeit (unitäre Diagonalisierbarkeit) echt stärker ist als nur Diagonalisierbarkeit. Deshalb mag es verwundern, dass dies für Trigonalisierung nicht gilt:

Satz 2.3.3 (Orthogonale/Unitäre Trigonalisierung)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Für einen Endomorphismus $\varphi : V \to V$ sind äquivalent:

(i) *Das charakteristische Polynom von φ zerfällt über $\mathbb{K}$ in Linearfaktoren.*

(ii) *Der Endomorphismus ist trigonalisierbar, d.h. es gibt eine Basis B von V, sodass $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ eine obere Dreiecksmatrix ist.*

(iii) *Der Endomorphismus ist* orthogonal/unitär trigonalisierbar*, d.h. es gibt eine Orthonormalbasis B von V, sodass $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ eine obere Dreiecksmatrix ist.*

Beweis. Die Äquivalenz (i) $\iff$ (ii) wurde in der LA1 gezeigt und es ist klar, dass (iii) die Bedingung (ii) impliziert. Aber wie kommt man von (i) oder (ii) nach (iii)?

Man kann hier entweder zeigen, dass (iii) aus (i) folgt, indem man den Beweis der entsprechenden Aussage in LA1 wörtlich kopieren und nur jedes Mal, wo der Begriff Basis verwendet wird, Orthonormalbasis schreiben. Essenziell ist hier wieder einmal das Gram-Schmidt-Verfahren, das uns ermöglicht, Orthonormalsysteme zu Orthonormalbasen fortzusetzen. Dies gibt einem sozusagen auch gleich einen Algorithmus, wie man eine solche Orthonormalbasis finden kann.

Alternativ dazu kann man aber auch zeigen, dass (iii) aus (ii) folgt. Nehmen wir an, es gibt eine geordnete Basis $\mathsf{B} = (v_1, \dots, v_n)$, sodass

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix}$$

obere Dreiecksgestalt hat. Da in den Spalten der Matrix die Bilder der Basisvektoren (ausgedrückt als Linearkombinationen eben jeder Basisvektoren) stehen, bedeutet dies, dass gilt:

$$\forall j \in \{1,\dots,n\} : \varphi(v_j) \in \mathrm{LH}\left(v_1,\dots,v_j\right).$$

Nun wenden wir Gram-Schmidt (Satz 1.3.7) auf diese geordnete Basis an und erhalten eine geordnete Orthonormalbasis $\widetilde{\mathsf{B}} = (w_1,\dots,w_n)$. Nach Satz 1.3.7 gilt, dass für jedes $j \in \{1,\dots,n\}$ gilt, dass $\mathrm{LH}\left(w_1,\dots,w_j\right) = \mathrm{LH}\left(v_1,\dots,v_j\right)$ gilt.

Wir wollen nun zeigen, dass die Matrix von φ bezüglich $\widetilde{\mathsf{B}}$ ebenfalls obere Dreiecksgestalt hat. Sei dazu $j \in \{1,\dots,n\}$ beliebig.

Da $w_j \in \mathrm{LH}\left(w_1,\dots,w_j\right) = \mathrm{LH}\left(v_1,\dots,v_j\right)$ gilt, und φ linear ist, können wir folgern:

$$\varphi(w_j) \in \mathrm{LH}\left(\varphi(v_1),\dots,\varphi(v_j)\right).$$

Weil aber $\mathrm{M}_{\mathsf{B},\mathsf{B}}\left(\varphi\right)$ obere Dreiecksgestalt hat, sind $\varphi(v_1),\dots,\varphi(v_j)$ alle in $\mathrm{LH}\left(v_1,\dots,v_j\right)$ enthalten und wir können insgesamt folgern:

$$\varphi(w_j) \in \mathrm{LH}\left(\varphi(v_1),\dots,\varphi(v_j)\right) \subseteq \mathrm{LH}\left(v_1,\dots,v_j\right) = \mathrm{LH}\left(w_1,\dots,w_j\right).$$

Das zeigt, dass in der j-ten Spalte der Matrix $\mathrm{M}_{\widetilde{\mathsf{B}},\widetilde{\mathsf{B}}}\left(\varphi\right)$ nur in den ersten j Einträgen von 0 verschiedene Zahlen stehen können. Somit ist $\mathrm{M}_{\widetilde{\mathsf{B}},\widetilde{\mathsf{B}}}\left(\varphi\right)$ eine obere Dreiecksmatrix. □

Bemerkung 2.3.4 (Trigonalisierung vs. Jordanscher Normalform)**.** Wir sehen also: Jede komplexe Matrix ist komplex unitär trigonalisierbar. Dies ist eine bemerkenswerte und sehr hilfreiche Aussage. Wir hatten am Ende von LA1 außerdem behauptet (aber nicht bewiesen), dass jede komplexe Matrix ähnlich ist zu einer Matrix in Jordan-Normalform. Man kann sich also fragen, ob es auch immer möglich ist, eine komplexe Matrix unitär in eine Jordan-Normalform zu überführen. Dies ist aber nicht möglich. Wir können wieder Beispiel 2.3.1 verwenden. Diese Matrix ist trigonalisierbar (sie ist ja bereits eine obere Dreiecksmatrix), aber ihre Jordan-Normalform wäre eine Diagonalmatrix und wir haben gesehen, dass sie nicht orthogonal diagonalisierbar ist, da die Eigenräume nicht orthogonal aufeinander stehen.

Wenn man eine komplexe Matrix $A \in \mathbb{C}^{n\times n}$ gegeben hat, muss sich also entscheiden: Entweder man kann eine Orthonormalbasis finden, sodass die Matrix in oberer Dreiecksform ist (aber nicht notwendigerweise in Jordan-Normalform), oder man kann eine Basis finden, sodass die Matrix in Jordan-Normalform ist (aber diese Basis ist nicht notwendigerweise eine Orthonormalbasis).

2.4. Selbstadjungierte komplexe Matrizen und symmetrische reelle Matrizen

Wir kommen nun zu einer Klasse von Matrizen, mit der sich besonders leicht arbeiten lässt:

Definition 2.4.1 (Selbstadjungierte Matrix)**.**
Es sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ und $n \in \mathbb{N}$. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *selbstadjungiert* oder *hermitesch*, falls sie gleich ihrer Adjungierten ist, falls also gilt:

$$A^* = A.$$

Bemerkung 2.4.2. Falls A nur reelle Einträge hat, so ist selbstadjungiert das selbe wie symmetrisch.

Wir listen nun einige Eigenschaften der Klasse selbstadjungierter Matrizen auf:

Lemma 2.4.3 (Selbstadjungierte Matrizen)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ *und* $n \in \mathbb{N}$.

(a) Die Menge $\{A \in \mathbb{K}^{n\times n} \mid A^* = A\}$ *ist ein* $\mathbb{R}$*-Vektorraum.*

(b) Wenn $A, B \in \mathbb{K}^{n\times n}$ *selbstadjungiert sind, dann ist das Produkt* AB *nur dann selbstadjungiert, falls* A *und* B *kommutieren.*

(c) Die Diagonaleinträge einer selbstadjungierten Matrix sind reell.

(d) Jeder komplexe Eigenwert λ *einer selbstadjungierten Matrix ist reell.*

(e) Für jede komplexe Matrix $A \in \mathbb{C}^{n\times n}$ *gibt es selbstadjungierte komplexe Matrizen* X, Y *mit*

$$A = X + \mathrm{i}Y.$$

Diese Darstellung ist eindeutig.

Beachten Sie, dass die Menge der selbstadjungierten Matrizen (auch im Falle $\mathbb{K} = \mathbb{C}$) nur ein *reeller* Vektorraum ist.

Lemma 2.4.4.
Es sei $n \in \mathbb{N}$ *und* $A \in \mathbb{C}^{n\times n}$*. Dann sind äquivalent:*

(i) A *ist eine obere Dreiecksmatrix und* A *ist selbstadjungiert.*

(ii) A *ist eine Diagonalmatrix mit reellen Einträgen.*

Lemma 2.4.5.
Es sei $n \in \mathbb{N}$ *und* $A, B \in \mathbb{C}^{n\times n}$*. Falls* A *unitär ähnlich zu* B *ist und* A *selbstadjungiert ist, dann ist auch* B *selbstadjungiert.*

Beweis. Da A unitär ähnlich zu B ist, muss es eine Basiswechselmatrix $U \in \mathrm{U}(n)$ geben mit

$$B = U^{-1}AU.$$

Da aber U unitär ist, gilt $U^{-1} = U^*$, es gilt also:

$$B = U^*AU.$$

Daraus folt nun:

$$\begin{aligned} B^* &= (U^*AU)^* \\ &= U^*A^*(U^*)^* \\ &= U^*AU \\ &= B. \end{aligned}$$

□

Wenn wir in Lemma 2.4.5 „unitär ähnlich“ durch „ähnlich“ ersetzen, wird die Aussage falsch. Es kann also vorkommen, dass eine selbstadjungierte Matrix ähnlich ist zu einer nicht selbstadjungierten Matrix, falls der Basiswechsel keine unitäre Matrix ist.

Durch Kombination von Lemma 2.4.4 und Lemma 2.4.5 erhalten wir nun den folgenden Satz, der auch unter dem Namen *Spektralsatz* bekannt ist[3]:

[3] Genau genommen gibt es mehrere miteinander verwandte Sätze, die alle *Spektralsatz* heißen, siehe auch Wikipedia

Satz 2.4.6 (Unitäre Diagonalisierbarkeit komplexer selbstadjungierter Matrizen – Spektralsatz für selbstadjungierte Matrizen)**.**
Es sei $n \in \mathbb{N}$ und $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen (i) und (ii) äquivalent:

(i) Die Matrix A ist selbst adjungiert, d.h. $A = A^*$.

(ii) Die Matrix A ist unitär diagonalisierbar und alle Eigenwerte von A sind reell.

Ebenso erhält man den reellen Spektralsatz:

Satz 2.4.7 (Orthogonale Diagonalisierbarkeit reeller symmetrischer Matrizen – Spektralsatz für reelle symmetrische Matrizen)**.**
Es sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$. Dann sind äquivalent:

(i) Die Matrix A ist symmetrisch, d.h. $A = A^\top$.

(ii) Die Matrix A ist orthogonal diagonalisierbar, d.h.

$$\exists U \in \mathrm{O}(n) : A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\top.$$

2.5. Normale Matrizen

Wir haben in Satz 2.4.7 gesehen, dass es sehr leicht ist, bei einer *reellen* quadratischen Matrix zu entscheiden, ob sie orthogonal diagonalisierbar ist. Wir müssen nur schauen, ob sie symmetrisch ist. Das ist notwendig und hinreichend für orthogonale Diagonalisierbarkeit.

Für komplexe Matrizen haben wir bis jetzt nur eine hinreichende Bedingungen für unitäre Diagonalisierbarkeit, nämlich selbstadjungiert zu sein (Satz 2.4.6), aber es gibt auch unitär diagonalisierbare Matrizen, die nicht selbstadjungiert, aber trotzdem unitär diagonalisierbar sind, wie zum Beispiel

$$A := \begin{pmatrix} \mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix}.$$

Was wir also als nächstens versuchen wollen, ist eine Bedingung an eine komplexe Matrix zu finden, die leicht zu überprüfen ist und die äquivalent zu unitären Diagonalisierbarkeit ist.

Definition 2.5.1 (Normale Matrix)**.**
Es sei $n \in \mathbb{N}$. Eine komplexe Matrix $A \in \mathbb{C}^{n\times n}$ heißt *normal*, falls sie mit ihrer Adjungierten kommutiert, falls also gilt:

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

Bemerkung 2.5.2.

- Jede selbstadjungierte Matrix ist offenbar auch normal, aber nicht jede normale Matrix ist selbstadjungiert.
- Während die selbstadjungierten Matrizen einen reellen Vektorraum bilden, ist die Summe von zwei normalen Matrizen im Allgemeinen nicht normal. Es ist sogar möglich, jede beliebige Matrix $A \in \mathbb{C}^{n\times n}$ als Summe von zwei normalen Matrizen zu schreiben. (Übung/Tutorium)

Wir versuchen nun zu verstehen, was es für eine Matrix bedeutet, normal zu sein:

Lemma 2.5.3.
Es sei $n \in \mathbb{N}$ und $A \in \mathbb{C}^{n \times n}$. Wir werden die Spalten von A mit $a_1, \dots, a_n \in \mathbb{C}^n$ und die Spalten von A^* mit $b_1, \dots, b_n \in \mathbb{C}^n$ bezeichnen.[4] *Die folgenden Aussagen sind äquivalent:*

(i) Die Matrix A ist normal, d.h. $A^* A = AA^*$.

(ii) Für alle $i, j \in \{1, \dots, n\}$ gilt $\langle a_i, a_j \rangle = \langle b_i, b_j \rangle$.

Insbesondere gilt für eine normale Matrix, dass $\forall i \in \{1, \dots, n\} : \|a_i\|^2 = \|b_i\|^2$.

Das nächste Lemma ist eine Variante von Lemma 2.4.4 für normale Matrizen:

Lemma 2.5.4.
Es sei $n \in \mathbb{N}$ und $A \in \mathbb{C}^{n \times n}$. Dann sind äquivalent:

(i) A ist eine obere Dreiecksmatrix und A ist normal.

(ii) A ist eine Diagonalmatrix.

Beweis. „(ii) $\Longrightarrow$ (i)“:
Eine Diagonalmatrix ist eine obere Dreiecksmatrix und eine Diagonalmatrix kommutiert mit ihrer adjungierten Matrix, da diese auch eine Diagonalmatrix ist und zwei Diagonalmatrizen immer kommutieren. Also ist jede Diagonalmatrix normal.

„(i) $\Longrightarrow$ (ii)“:
Nehmen wir nun an, die Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix}$$

sei eine normale obere Dreiecksmatrix. Die erste Spalte hat Länge $|a_{1,1}|$. Nach Lemma 2.5.3 muss also auch die erste Zeile Länge $|a_{1,1}|$ haben, es gilt also:

$$|a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2 = |a_{1,1}|^2 .$$

Also sind alle Summanden bis auf den ersten Null, d.h. $a_{1,2} = a_{1,3} = \cdots = a_{1,n} = 0$. Also besteht die erste Zeile von A mit Ausnahme des ersten Eintrags nur aus Nullen:

$$A = \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix}.$$

Das impliziert nun direkt, dass die zweite Spalte von A die Länge $|a_{2,2}|$ hat. Nochmaliges Anwenden von Lemma 2.5.3 liefert, dass die zweite Zeile die selbe Norm haben muss. Somit gilt $a_{2,3} = a_{2,4} = \cdots = a_{2,n} = 0$.

So fahren wir fort, bis wir sehen, dass alle Einträge über der Diagonalen 0 sein müssen und A eine Diagonalmatrix ist. □

[4] Die Zeilen von A sind also $b_1^*, \dots, b_n^*$.

Analog zu Lemma 2.4.5 zeigt man:

Lemma 2.5.5.
Es sei $n \in \mathbb{N}$ und $A, B \in \mathbb{C}^{n\times n}$. Falls A unitär ähnlich zu B ist und A normal ist, dann ist auch B normal.

Wie schon bei Selbstadjungiertheit stellt man fest, dass man „unitär ähnlich" hier nicht durch „ähnlich" ersetzen kann. Es kann also vorkommen, dass eine normale Matrix ähnlich ist zu einer nicht normalen Matrix, falls der Basiswechsel keine unitäre Matrix ist.

Nun können wir die Charakterisierung unitärer Diagonalisierbarkeit für komplexe Matrizen beweisen; auch diese Aussage ist unter dem Namen *Spektralsatz* bekannt:

Satz 2.5.6 (Unitäre Diagonalisierbarkeit normaler Matrizen – Spektralsatz für normale komplexe Matrizen)**.**
Es sei $n \in \mathbb{N}$ und $A \in \mathbb{C}^{n\times n}$. Dann sind äquivalent:

(i) Die Matrix A ist normal, d.h. $A^*A = AA^*$.

(ii) Die Matrix A ist unitär diagonalisierbar, d.h.

$$\exists U \in \mathrm{U}(n) : A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*.$$

2.6. Adjungierte Abbildungen und die Isometrie-Normalform

Gegeben sei ein endlich dimensionaler Vektorraum V mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und ein $\mathbb{K}$-linearer Endomorphismus φ von V. Wir wollen wissen, ob φ orthogonal diagonalisierbar ist, d.h. ob eine geordnete Orthonormalbasis B von V gibt, sodass $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ eine Diagonalmatrix ist.

Wir haben diese Fragen in den Sätzen 2.5.6 (komplex) und 2.4.7 (reell) beantwortet. Allerdings nur für $V = \mathbb{K}^n$ mit dem Standardskalarprodukt. Die Eigenschaften einer Matrix *selbstadjungiert* oder *normal* zu sein, hingen essentiell von der adjungierten Matrix ab und übertragen sich nicht direkt auf Endomorphismen, da wir nicht definiert haben, was die adjungierte Abbildung eines Endomorphismus ist. Das wollen wir jetzt nachholen, um dann die Sätze 2.5.6 (komplex) und 2.4.7 auf allgemeine (endlich dimensionale) Vektorräume mit Skalarprodukt übertragen zu können.

Bevor wir das tun, benötigen wir allerdings noch ein kurzes Lemma über Skalarprodukte:

Lemma 2.6.1.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(V, \langle\cdot,\cdot\rangle)$ ein Vektorraum mit Skalarprodukt. Es seien $v_1, v_2 \in V$. Dann sind äquivalent:

(i) $\forall u \in V : \langle v_1, u\rangle = \langle v_2, u\rangle$.

(ii) $\forall u \in V : \langle u, v_1\rangle = \langle u, v_2\rangle$.

(iii) $v_1 = v_2$.

Beweis. Wir zeigen die Implikation „(i) $\Longrightarrow$ (iii)". Die anderen Implikationen sind einfach zu zeigen.

Gegeben seien also $v_1, v_2 \in V$. Wir setzen $u := v_2 - v_1 \in V$. Nach (i) gilt nun:

$$\|v_1 - v_2\|^2 = \langle v_1 - v_2, v_1 - v_2\rangle = \langle v_1 - v_2, u\rangle = \langle v_1, u\rangle - \langle v_2, u\rangle = 0.$$

Also ist $\|v_1 - v_2\| = 0$ und somit gilt: $v_1 = v_2$. □

Definition 2.6.2 (Adjungierte Abbildungen)**.**
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ seien Vektorräume mit Skalarprodukt. Weiterhin seien zwei Abbildungen $\varphi : V \to W$ und $\psi : W \to V$ gegeben.

Wir sagen ψ ist *adjungiert* zu φ, falls gilt:

$$\forall v \in V; \forall w \in W : \ \left\langle \varphi(v), w \right\rangle_W = \left\langle v, \psi(w) \right\rangle_V .$$

Wenn eine Abbildung ψ existiert, die adjungiert zu φ ist, so ist diese eindeutig (Lemma 2.6.3) und wir schreiben

$$\varphi^* := \psi$$

für diese eindeutige adjungierte Abbildung.

Aus dieser Definition folgen nun ein paar Eigenschaften:

Lemma 2.6.3.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ seien Vektorräume mit Skalarprodukt.

(a) Wenn $\psi : W \to V$ zu $\varphi : V \to W$ adjungiert ist, dann sind φ und ψ linear.

(b) Wenn $\psi : W \to V$ zu $\varphi : V \to W$ adjungiert ist, dann ist φ zu ψ adjungiert.

(c) Wenn $\varphi : V \to W$ eine adjungierte Abbildung $\psi : W \to V$ besitzt, so ist diese eindeutig.

Lemma 2.6.3 erlaubt es uns also, von *der* Adjungierten einer linearen Abbildung $\varphi : V \to W$ zu sprechen, wenn sie existiert. Es ist allerdings nicht wahr, dass jede lineare Abbildung eine solche Adjungierte besitzt, siehe Beispiel 2.6.5(c):

Bemerkung 2.6.4. Die Notation für die *adjungierte Abbildung* φ^* ist Standard in der Literatur, kollidiert aber mit der *dualen Abbildung* $\varphi^* : W^* \to V^*$, die wir im Kapitel über Dualräume einführen werden (siehe Bemerkung 3.1.4). Aus dem Kontext sollte aber immer klar sein, ob mit φ^* die duale Abbildung oder die adjungierte Abbildung gemeint ist.

Beispiel 2.6.5. (a) Es sei V ein (endlich oder unendlich dimensionaler) Vektorraum mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und U ein *endlich dimensionaler* Untervektorraum. Dann ist die Orthogonalprojektion $\pi_U : V \to V$ (siehe Satz 1.4.5) auf U ein Endomorphismus von V. Wir behaupten nun, π_U ist adjungiert zu sich selbst.[5]. Es seien $v, w \in V$. Wir werden nun folgende Aussagen zeigen:

- $\langle \pi_U(v), w\rangle = \langle \pi_U(v), \pi_U(w)\rangle$
- $\langle v, \pi_U(w)\rangle = \langle \pi_U(v), \pi_U(w)\rangle$

[5] Hieraus würde dann auch folgen, dass die Orthogonalprojektion linear ist, aber das wissen wir ja schon...

Sobald wir dies gezeigt haben, folgt damit natürlich, dass

$$\langle \pi_U(v), w\rangle = \langle v, \pi_U(w)\rangle$$

und der Beweis ist beendet.

Um die erste Aussage zu zeigen, nutzen wir aus, dass sich der Raum V orthogonal zerlegen lässt in $V = U \oplus U^\perp$ (Satz 1.4.8). Insbesondere lässt sich also der Vektor $w \in V$ schreiben als:

$$w = \underbrace{\pi_U(w)}_{\in U} + \underbrace{(w - \pi_U(w))}_{\in U^\perp}.$$

Somit gilt:

$$\langle \pi_U(v), w\rangle = \langle \pi_U(v), \pi_U(w) + (w - \pi_U(w))\rangle = \langle \pi_U(v), \pi_U(w)\rangle + \underbrace{\langle \pi_U(v), w - \pi_U(w)\rangle}_{=0},$$

wobei der letzte Summand verschwindet, weil hier ein Element aus U mit einem aus $U^\perp$ skalar multipliziert wird. Mit der selben Idee angewandt auf den Vektor $v \in V$ zeigt man die andere Identität und somit ist $\pi_U : V \to V$ selbstadjungiert.

Achtung: Wenn man die Orthogonalprojektion auf ihr Bild koeinschränkt, d.h. wenn man die Abbildung

$$V \to U, \quad v \mapsto \pi_U(v)$$

betrachtet, dann ist diese Abbildung nicht gleich ihrer Adjungierten, weil die Adjungierte[6] eine Abbildung von U nach V sein muss und somit schon aus rein formalen Gründen eine andere Abbildung sein muss.

(b) Es seien $a, b \in \mathbb{R}$ mit $a < b$. Wir betrachten den Vektorraum $V = C([a,b],\mathbb{C})$ der stetigen Funktionen vom kompakten Intervall $[a,b] \subseteq \mathbb{R}$ in die komplexen Zahlen. Das Skalarprodukt sei gegeben durch

$$\langle f, g\rangle := \int_a^b f(t)\overline{g(t)}dt.$$

Für eine Funktion $h \in C([a,b],\mathbb{C})$ ist der Multiplikationsoperator

$$M_h : V \to V, \quad f \mapsto h \cdot f$$

definiert als die Abbildung, die eine Funktion f punktweise mit h multipliziert. Da das punktweise Produkt von stetigen Funktionen wieder stetig ist, ist dies eine wohldefinierte Abbildung von V nach V.

Wenn wir – wie üblich – mit $\overline{h}$ die punktweise komplex konjugierte von h bezeichnen, dann ist der Multiplikationsoperator $M_{\overline{h}}$ mit $\overline{h}$ die adjungierte Abbildung zu M_h, wie

[6] Man muss natürlich zuerst zeigen, dass sie existiert, was hier aber der Fall ist.

man leicht nachrechnet:

$$\begin{aligned}\langle M_h(f), g\rangle &= \langle h\cdot f, g\rangle \\ &= \int_a^b h(t)f(t)\overline{g(t)}dt \\ &= \int_a^b f(t)\overline{\overline{h(t)}g(t)}dt \\ &= \left\langle f, \overline{h}\cdot g\right\rangle \\ &= \left\langle f, M_{\overline{h}}(g)\right\rangle.\end{aligned}$$

Wenn die Funktion h also nur reelle Werte annimmt, so ist die Abbildung M_h gleich ihrer Adjungierten (also *selbstadjungiert* (siehe Definition 2.6.7(a))). In jedem Fall gilt aber, dass M_h mit $M_{\overline{h}}$ kommutiert (die Abbildung ist also *normal* (siehe Definition 2.6.7(b))).

(c) Gegeben $m, n \in \mathbb{N}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{m\times n}$. Wir versehen $\mathbb{K}^n$ und $\mathbb{K}^m$ mit ihrem jeweiligen Standardskalarprodukt.

Dann besitzt die lineare Abbildung

$$\mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax$$

eine adjungierte Abbildung (im Sinne der Definition 2.6.2) und diese ist

$$\mathbb{K}^m \to \mathbb{K}^n, \quad y \mapsto A^* y.$$

Hierbei ist $A^* = \overline{A^\top}$ die adjungierte Matrix (im Sinne der Notation 2.2.1).

(d) Zum Abschluss noch ein Beispiel für eine lineare Abbildung, die keine adjungierte Abbildung besitzt. Wir betrachten den (unendlich dimensionalen) Vektorraum aller Polynome $V := \mathbb{K}[X]$ mit Koeffizienten aus $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Auf diesem betrachten wir das folgende Skalarprodukt:

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

Als lineare Abbildung nehmen wir den folgenden *Ableitungsoperator*:

$$\varphi : V \to V, \quad p \mapsto p',$$

der ein Polynom auf seine erste Ableitung abbildet, d.h.

$$\varphi(X^k) = kX^{k-1} \text{ für } k \in \mathbb{N} \quad \text{und} \quad \varphi(X^0) = 0.$$

Wir wollen nun zeigen, dass es keine Abbildung $\psi : V \to V$ gibt, die adjungiert zu φ ist.

Per Widerspruch nehmen wir an, es existiere so eine Abbildung $\psi : V \to V$. In diese Abbildung können wir nun das Element $X^0 \in V$ einsetzen und erhalten ein Polynom

$$q_0 := \psi(X^0) \in V = \mathbb{K}[X].$$

Nun gilt für jedes $n \in \mathbb{N}$:

$$\begin{aligned}\langle X^n, q_0 \rangle &= \langle X^n, \psi(X^0) \rangle \\ &= \langle \varphi(X^n), X^0 \rangle \\ &= \langle nX^{n-1}, X^0 \rangle \\ &= \int_0^1 nt^{n-1} t^0 dt \\ &= \left[t^n\right]_0^1 \\ &= 1 - 0 = 1.\end{aligned}$$

Andererseits können wir aber mit der Cauchy-Schwarz-Ungleichung (Satz 1.2.9) abschätzen:

$$\begin{aligned}\left|\langle X^n, q_0 \rangle\right| &\leq \left\|X^n\right\| \|q_0\| \\ &= \sqrt{\langle X^n, X^n \rangle}\, \|q_0\| \\ &= \sqrt{\int_0^1 t^{2n} dt}\, \|q_0\| \\ &= \sqrt{\left[\frac{t^{2n+1}}{2n+1}\right]_0^1}\, \|q_0\| \\ &= \frac{1}{\sqrt{2n+1}} \|q_0\|.\end{aligned}$$

Kombinieren wir diese Ergebnisse, so sehen wir, dass wir gerade gezeigt haben:

$$\forall n \in \mathbb{N} : 1 \leq \frac{1}{\sqrt{2n+1}} \|q_0\|.$$

Da die linke Seite konstant 1 ist und die rechte Seite für $n \to \infty$ gegen 0 konvergiert, ist dies ein Widerspruch.

Es gilt also, dass es das Polynom q_0 nicht geben kann und somit auch nicht die adjungierte Abbildung $\psi : V \to V$.

Es ist also möglich, eine lineare Abbildung zu konstruieren, die keine Adjungierte besitzt.

Das Gegenbeispiel 2.6.5 (d) zeigt, dass es nicht immer sicher ist, ob es eine adjungierte Abbildung gibt. Glücklicherweise ist dies im Endlichdimensionalen immer der Fall:

Proposition 2.6.6.
Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Gegeben seien weiterhin $\mathbb{K}$-Vektorräume mit Skalarprodukt V und W.

Falls $\dim(V) = n \in \mathbb{N}$ und $\dim(W) = m \in \mathbb{N}$ endlich sind, dann hat jede lineare Abbildung $\varphi : V \to W$ eine adjungierte Abbildung $\varphi^* : W \to V$.

Weiterhin gilt: Falls B eine Orthonormalbasis von V und C eine Orthonormalbasis von W ist, dann gilt:

$$\mathrm{M}_{\mathsf{B},\mathsf{C}}\left(\varphi^*\right) = \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^*.$$

Nun können wir die Begriffe *selbstadjungiert* und *normal* direkt auf Endomorphismen von allgemeinen Vektorräumen mit Skalarprodukt übertragen:

Definition 2.6.7.
Es V ein Vektorraum mit Skalarprodukt über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Ein Endomorphismus $\varphi \in \operatorname{End}(V)$ heiße

(a) *selbstadjungiert*, falls φ adjungiert zu sich selbst ist, d.h.

$$\forall v, w \in V : \langle \varphi(v), w \rangle = \langle v, \varphi(w) \rangle .$$

(b) *normal*, falls φ eine Adjungierte φ^* besitzt und mit dieser kommutiert, d.h.

$$\varphi \circ \varphi^* = \varphi^* \circ \varphi.$$

Nun können wir die Sätze 2.5.6 und 2.4.7 direkt auf Endomorphismen von endlich dimensionalen Vektorräumen mit Skalarprodukt übertragen:

Satz 2.6.8 (Charakterisierung der komplex unitär diagonalisierbaren Endomorphismen (*Spektralsatz* für normale Abbildungen))**.**
Es sei $n \in \mathbb{N}$ und V ein unitärer Vektorraum der Dimension n. Ein Endomorphismus $\varphi \in \operatorname{End}_{\mathbb{C}}(V)$ ist genau dann unitär diagonalisierbar, wenn er eine normale Abbildung ist.

Satz 2.6.9 (Charakterisierung der reell orthogonal diagonalisierbaren Endomorphismen (*Spektralsatz* für selbstadjungierte Abbildungen))**.**
Es sei $n \in \mathbb{N}$ und V ein Euklidischer Vektorraum der Dimension n. Ein Endomorphismus $\varphi \in \operatorname{End}_{\mathbb{R}}(V)$ ist genau dann orthogonal diagonalisierbar, wenn er eine selbstadjungierte Abbildung ist.

Satz 2.6.10 (Komplexe Isometrienormalform)**.**
Es sei V ein n-dimensionaler unitärer Vektorraum ($n \in \mathbb{N}$). Für einen Endomorphismus $\varphi : V \to V$ sind die folgenden Aussagen (i) und (ii) äquivalent:

(i) φ ist eine lineare Isometrie.

(ii) φ ist unitär diagonalisierbar und alle Eigenwerte haben Betrag 1, d.h.

$$\exists \mathsf{B} \text{ Orthonormalbasis von } V : \quad \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} e^{\mathrm{i}\alpha_1} & & & \\ & e^{\mathrm{i}\alpha_2} & & \\ & & \ddots & \\ & & & e^{\mathrm{i}\alpha_n} \end{pmatrix} \text{ mit } \alpha_j \in \mathbb{R}.$$

Satz 2.6.11 (Reelle Isometrie-Normalform)**.**
Es sei V ein n-dimensionaler Euklidischer Vektorraum ($n \in \mathbb{N}$). Für einen Endomorphismus $\varphi : V \to V$ sind äquivalent:

(i) φ ist eine lineare Isometrie.

(ii) Es gibt eine geordnete Orthonormalbasis B von V, sodass die Darstellungsmatrix die folgende Blockstruktur hat:

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_r} \end{pmatrix}.$$

Jeder Block A_j ist entweder

- *eine* (1×1)*-Matrix mit dem Eintrag* $A_j = 1 \in \mathbb{R}^{1\times 1}$,
- *eine* (1×1)*-Matrix mit dem Eintrag* $A_j = -1 \in \mathbb{R}^{1\times 1}$
- *oder ein Drehkästchen* $A_j = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$, *wobei* α *ein Winkel ist mit* $0 < \alpha < \pi$.

Die Drehkästchen *sind uns in Beispiel 2.2.5 schon einmal begegnet.*

Lemma 2.6.12.
Es sei $n \in \mathbb{N}$ *und*

$$v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n.$$

Wir setzen

$$\overline{v} := \begin{pmatrix} \overline{z_1} \\ \vdots \\ \overline{z_n} \end{pmatrix} \in \mathbb{C}^n \quad \textit{und} \quad \mathrm{Re}(v) := \begin{pmatrix} \mathrm{Re}(z_1) \\ \vdots \\ \mathrm{Re}(z_n) \end{pmatrix} \in \mathbb{R}^n \quad \textit{und} \quad \mathrm{Im}(v) := \begin{pmatrix} \mathrm{Im}(z_1) \\ \vdots \\ \mathrm{Im}(z_n) \end{pmatrix} \in \mathbb{R}^n.$$

Dann gilt:

$$\mathrm{LH}_{\mathbb{C}}(v, \overline{v}) = \mathrm{LH}_{\mathbb{C}}(\mathrm{Im}(v), \mathrm{Re}(v)).$$

Außerdem sind äquivalent:

- *Die Vektoren* $\mathrm{Im}(v)$ *und* $\mathrm{Re}(v)$ *sind orthogonal in* $\mathbb{R}^n$ *und haben die selbe Norm (bezüglich des reellen Standardskalarproduktes auf* $\mathbb{R}^n$.
- *Die Vektoren* v *und* $\overline{v}$ *sind orthogonal in* $\mathbb{C}^n$ *bezüglich des komplexen Standardskalarproduktes.*

Lemma 2.6.13.
Es sei $n \in \mathbb{N}$ *und* $A \in \mathbb{C}^{n\times n}$. *Dann gilt:*

(a) *Wenn* $v \in \mathbb{C}^n$ *ein Eigenvektor von* A *zum Eigenwert* $\lambda \in \mathbb{C}$ *ist, dann ist* $\overline{v} \in \mathbb{C}^n$ *ein Eigenvektor von* $\overline{A}$ *zum Eigenwert* $\overline{\lambda}$.

(b) *Falls* A *nur reelle Einträge besitzt, dann ist zu jedem komplexen Eigenwert* λ *auch die komplex konjugierte Zahl* $\overline{\lambda}$ *ein Eigenwert. Insbesondere treten nicht-reelle Eigenwerte immer in Paaren auf.*

Beweis. Beweisskizze:
Dies folgt aus der Eigenschaft, dass komplexe Konjugation ein Körperautomorphismus ist, sich also in alle Summen und Produkte hineinziehen lässt.

Aus $Av = \lambda v$ folgt durch Konjugieren auf beiden Seiten sofort: $\overline{A}\,\overline{v} = \overline{\lambda}\overline{v}$.

Falls A nur reelle Einträge besitzt, gilt $\overline{A} = A$. Für jeden Eigenwert $\lambda \in \mathbb{C}$ ist somit $\overline{\lambda}$ auch ein Eigenwert von A. Und falls λ nicht reell ist, dann ist $\overline{\lambda} \neq \lambda$. □

Beweis von Satz 2.6.11. Wir werden nur die Implikation „(i) $\Longrightarrow$ (ii)“ zeigen:

Nach Gram-Schmidt (Korollar 1.3.8) besitzt V eine Orthonormalbasis und somit ist V isometrisch isomorph zu $\mathbb{R}^n$ mit dem Standardskalarprodukt (Satz 2.1.10).

Wir können (und werden) im Folgenden also annehmen, dass $V = \mathbb{R}^n$ mit dem Standardskalarprodukt ist und die Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto Ax$ durch Multiplikation mit der Matrix

$A = \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi) \in \mathbb{R}^{n\times n}$ gegeben ist. Da nach Voraussetzung (i) die Abbildung eine Isometrie ist, gilt mit Proposition 2.2.2, dass A eine orthogonale Matrix ist, d.h. $A \in \mathrm{O}(n)$.

Nun ist aber $\mathrm{O}(n) = \mathrm{U}(n) \cap \mathbb{R}^{n\times n}$ (siehe Definition 2.2.7), d.h. die Matrix A kann auch als unitäre Matrix aufgefasst werden.

So interpretiert gehört die selbe Matrix A auch zu der folgenden $\mathbb{C}$-linearen Abbildung auf dem unitären Vektorraum $\mathbb{C}^n$:

$$\varphi_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n, \quad x \mapsto Ax.$$

Nach Satz 2.6.10 können wir diese Matrix – aufgefasst als komplexe Matrix – komplex unitär diagonalisieren. Es ist also möglich, eine Orthonormalbasis des unitären Vektorraums $\mathbb{C}^n$ zu finden, die aus Eigenvektoren der $\mathbb{C}$-linearen Abbildung $\varphi_{\mathbb{C}}$ besteht. Anders gesagt zerfällt der Raum $\mathbb{C}^n$ in eine direkte Summe von Eigenräumen

$$\mathbb{C}^n = \bigoplus_{j=1}^{r} E_{\lambda_j}(\varphi_{\mathbb{C}}),$$

wobei $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen komplexen Eigenwerte von $\varphi_{\mathbb{C}}$ sind und je zwei Eigenräume orthogonal aufeinander stehen (Lemma 2.3.2). Weil die Matrix A unitär ist, haben alle diese Eigenwerte Betrag 1, liegen also auf dem Einheitskreis

$$\forall j \in \{1, \ldots, r\} : \mathrm{U}(1) = \{z \in \mathbb{C} \mid |z| = 1\} = \left\{ e^{\mathrm{i}\alpha} \,\middle|\, \alpha \in \mathbb{R} \right\},$$

wir können also jeden Eigenwert $\lambda_j = e^{\mathrm{i}\alpha_j}$ schreiben und wegen der Periodizität können wir das Argument α_j im halboffenen Intervall von $-\pi$ bis π wählen:

$$\forall j \in \{1, \ldots, r\} : \alpha_j \in (-\pi, \pi].$$

Die beiden Fälle $\alpha_j = 0$ und $\alpha_j = \pi$ verhalten sich hierbei sehr anders als alle anderen:

Betrachten wir zuerst $\alpha_j = 0$. Dann ist $\lambda_j = e^{\mathrm{i}0} = 1$, d.h. 1 ist ein Eigenwert von $\varphi_{\mathbb{C}}$. Wir möchten nun zeigen, dass der komplexe Eigenraum $E_1(\varphi_{\mathbb{C}}) \subseteq \mathbb{C}^n$ eine Orthonormalbasis besitzt, deren Vektoren alle nur reelle Einträge haben.

Die Dimension des komplexen Eigenraums $E_1(\varphi_{\mathbb{C}})$ ist gleich der geometrischen Vielfachheit von $\lambda_j = 1$ und diese ist gleich $n - \mathrm{rg}(\mathbb{1}_n - A)$. Diese Zahl hängt nun aber nicht davon ab, ob man die Matrix als reelle oder als komplexe Matrix betrachtet, d.h. wir sehen, dass $\lambda_j = 1$ auch ein Eigenwert der reellen Abbildung φ ist und der reelle Eigenraum die selbe Dimension hat. Wir wählen nun also mit Gram-Schmidt eine reelle Orthonormalbasis von $E_1(\varphi)$. Da $E_1(\varphi) \subseteq E_1(\varphi_{\mathbb{C}})$ ist und die reelle Dimension des ersten Raums die komplexe Dimension des zweiten Raums ist, folgt daraus, dass die gewählte reelle Orthonormalbasis auch eine Orthonormalbasis von $E_1(\varphi_{\mathbb{C}})$ ist.

Betrachten wir als nächstes $\alpha_j = \pi$. Dann ist $\lambda_j = e^{\mathrm{i}\pi} = -1$, d.h. -1 ist ein Eigenwert von $\varphi_{\mathbb{C}}$. Mit genau der selben Argumentation wie oben sehen wir, dass es möglich ist, eine Orthonormalbasis des Eigenraums $E_{-1}(\varphi_{\mathbb{C}})$ zu finden, die nur aus Vektoren aus dem $\mathbb{R}^n$ besteht.

Falls die Matrix nur Eigenwerte 1 und -1 haben sollte, sind wir nun schon fertig und können die Orthonormalbasen zu einer Basis des $\mathbb{R}^n$ zusammensetzen, sodass A bezüglich dieser Basis Diagonalgestalt hat. In diesem Fall hätten wir die Isometrie-Normalform ganz ohne Drehkästchen.

Kommen wir nun zu dem Fall, dass es Eigenwerte $\lambda_j = e^{\mathrm{i}\alpha_j}$ gibt mit $\alpha_j \in (-\pi, 0) \cup (0, \pi)$.

Ein solcher Eigenwert ist also keine reelle Zahl. Nach Lemma 2.6.13(b) treten solche Eigenwerte immer in Paaren auf. Wir werden uns auf positive α_j, d.h. auf solche im Intervall $(0, \pi)$

beschränken. In der komplexen Zahlenebene sind dies die Punkte über der reellen Achse, d.h. solche mit positivem Imaginärteil.

Zu jedem solchen $\alpha_j \in (0,\pi)$ betrachten wir den Untervektorraum

$$U_j := E_{\lambda_j}\left(\varphi_{\mathbb{C}}\right) \oplus E_{\overline{\lambda_j}}\left(\varphi_{\mathbb{C}}\right) = E_{e^{\mathrm{i}\alpha_j}}\left(\varphi_{\mathbb{C}}\right) \oplus E_{e^{-\mathrm{i}\alpha_j}}\left(\varphi_{\mathbb{C}}\right) \subseteq \mathbb{C}^n$$

Die Summe ist tatsächlich direkt, da bei unitär diagonalisierbaren Matrizen Eigenräume zu unterschiedlichen Eigenwerten immer orthogonal aufeinander stehen (Lemma 2.3.2).

Ziel ist es nun, eine Orthonormalbasis von $U_j \subseteq \mathbb{C}^n$ zu finden, die nur aus Vektoren mit reellen Einträgen besteht. Wir wählen zuerst mit Gram-Schmidt eine komplexe Orthonormalbasis von $E_{\lambda_j}\left(\varphi_{\mathbb{C}}\right) = E_{e^{\mathrm{i}\alpha_j}}\left(\varphi_{\mathbb{C}}\right)$:

$$(v_1, \ldots, v_t).$$

Wenn wir nun jeden einzelnen dieser Vektoren komponentenweise komplex konjugieren (siehe Lemma 2.6.13), erhalten wir eine Orthonormalbasis von $E_{\overline{\lambda_j}}\left(\varphi_{\mathbb{C}}\right) = E_{e^{-\mathrm{i}\alpha_j}}\left(\varphi_{\mathbb{C}}\right)$

$$(\overline{v_1}, \ldots, \overline{v_t}).$$

Da die Vektoren aus unterschiedlichen Eigenräume orthogonal aufeinander stehen, können wir diese beiden Orthonormalsysteme zu einer Orthonormalbasis von U_j zusammensetzen:

$$(v_1, \overline{v_1}, v_2, \overline{v_2}, \ldots, v_t, \overline{v_t}).$$

Die beiden Vektoren $v_1, \overline{v_1}$ erzeugen einen komplexen zweidimensionalen Untervektorraum. Nach Lemma 2.6.12 wird dieser ebenfalls erzeugt von a_1, b_1, wobei $a_1 := \mathrm{Re}(v_1)$ der komponentenweise Realteil und $b_1 := \mathrm{Im}(v_1)$ der komponentenweise Imaginärteil von v_1 ist. Da v_1 und $\overline{v}_1$ orthogonal aufeinander stehen, gilt – ebenfalls nach Lemma 2.6.12 –, dass a_1 und b_1 ebenfalls orthogonal aufeinander stehen und die selbe Länge $r_1 := \|a_1\| = \|b_1\|$ haben. Hierbei kann r_1 nicht 0 sein, weil ansonsten $a_1 = b_1 = 0$ wären. Wir können also (a_1, b_1) normieren und erhalten mit

$$\left(\widetilde{b_1} := \frac{1}{r_1} b_1 \text{ , } \widetilde{a_1} := \frac{1}{r_1} a_1\right)$$

eine Orthonormalbasis für $\mathrm{LH}_{\mathbb{C}}(v_1, \overline{v_1})$, die aus Vektoren aus $\mathbb{R}^n$ besteht.

Was passiert nun, wenn wir die Matrix A auf $\frac{1}{r_1}a_1$ und $\frac{1}{r_1}b_1$ anwenden?

$$\begin{aligned}
A\,\widetilde{a_1} + \mathrm{i}A\,\widetilde{b_1} &= A\frac{1}{r_1}a_1 + \mathrm{i}A\frac{1}{r_1}b_1 \\
&= \frac{1}{r_1}A(a_1 + \mathrm{i}b_1) \\
&= \frac{1}{r_1}Av_1 \\
&= \frac{1}{r_1}\lambda_1 v_1 \\
&= \frac{1}{r_1}e^{\mathrm{i}\alpha_j}v_1 \\
&= \frac{1}{r_1}\big(\cos(\alpha_j) + \mathrm{i}\sin(\alpha_j)\big)(a_1 + \mathrm{i}b_1) \\
&= \frac{1}{r_1}\big(\cos(\alpha_j)a_1 + \mathrm{i}\cos(\alpha_j)b_1 + \mathrm{i}\sin(\alpha_j)a_1 - \sin(\alpha_j)b_1\big) \\
&= \frac{1}{r_1}\big(\cos(\alpha_j)a_1 - \sin(\alpha_j)b_1\big) + \mathrm{i}\frac{1}{r_1}\big(\sin(\alpha_j)a_1 + \cos(\alpha_j)b_1\big) \\
&= \Big(-\sin(\alpha_j)\widetilde{b_1} + \cos(\alpha_j)\widetilde{a_1}\Big) + \mathrm{i}\Big(\cos(\alpha_j)\widetilde{b_1} + \sin(\alpha_j)\widetilde{a_1}\Big)
\end{aligned}$$

Wenn wir also auf beiden Seiten den Imaginärteil betrachten, erhalten wir:

$$A\widetilde{b_1} = \cos(\alpha_j)\widetilde{b_1} + \sin(\alpha_j)\widetilde{a_1}$$

und wenn wir den Realteil nehmen, erhalten wir:

$$A\widetilde{a_1} = -\sin(\alpha_j)\widetilde{b_1} + \cos(\alpha_j)\widetilde{a_1}.$$

Die Wirkung der Matrix A auf dem von $\left(\widetilde{b_1},\widetilde{a_1}\right)$ erzeugten Untervektorraum wird also bezüglich der Orthonormalbasis $\left(\widetilde{b_1},\widetilde{a_1}\right)$ durch das Drehkästchen

$$\begin{pmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix}$$

beschreiben.[7]

Dieses Verfahren werden wir nun der Reihe nach auch noch auf $(v_2,\overline{v_2})$, $(v_3,\overline{v_3})$... an, bis wir eine neue geordnete Orthonormalbasis von U_j gefunden haben:

$$\left(\widetilde{b_1},\widetilde{a_1},\dots,\widetilde{b_t},\widetilde{a_t}\right).$$

Diese Vektoren sind alle aus dem $\mathbb{R}^n$ und die Darstellungsmatrix ist nun eine Blockmatrix mit lauter Drehkästchen mit dem Winkel α_j.

Wenn wir dies nun – wie oben gesagt – für alle Eigenwerte oberhalb der reellen Achse durchführen und mit den (falls vorhandenen) Eigenräumen für 1 beziehungsweise -1 kombinieren, erhalten wir schließlich eine Orthonormalbasis von $\mathbb{R}^n$, sodass die Matrix bezüglich dieser Basis die gewünschte Form hat. □

[7] Es mag irritierend wirken, dass wir in der geordneten Basis zuerst $\widetilde{b_1}$ und dann $\widetilde{a_1}$ nehmen. Man könnte es auch anders herum machen, aber so ist sichergestellt, dass das Drehkästchen, das entsteht genau die Form hat, die wir im Satz postuliert haben.

Bemerkung 2.6.14 (Die Spur einer Isometrie)**.** Es sei V ein endlich-dimensionaler Euklidischer Vektorraum und $\varphi : V \to V$ eine lineare Isometrie.

Da $\varphi : V \to V$ ein Endomorphismus eines endlich dimensionalen Vektorraums ist, besitzt φ eine Spur $\operatorname{tr}\varphi \in \mathbb{R}$. Was sagt uns diese Spur über die Isometrie aus?

Aus Satz 2.6.11 folgt, dass φ bezüglich der richtigen Basis eine Matrix in Isometrie-Normalform hat. Die Spur von φ ist also nichts anderes als die Summe der Diagonaleinträge der Isometrie-Normalform. Für jede 1 oder -1 summiert man eben diese Zahl auf und für jedes Drehkästchen mit Winkel α_j ergibt dies $2\cos(\alpha_j)$.

Insofern kann einem die Spur oft schon eine Auskunft über die Struktur der Isometrie-Normalform geben, ohne dass man irgendwelche Eigenwerte berechnen muss.

Beispiel 2.6.15. Es sei die folgende Matrix gegeben:

$$A := \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix}.$$

Man rechnet leicht nach, dass $A^\top A = \mathbb{1}_3$ gilt. Also ist A eine orthogonale Matrix und die dazugehörige lineare Abbildung

$$\varphi : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto Ax$$

ist eine lineare Isometrie. Nach Satz 2.6.11 ist die Matrix orthogonal ähnlich zu einer Matrix $\widetilde{A}$ in Isometrie-Normalform. Man rechnet ebenfalls leicht nach, dass $\det(A) = 1$ gilt, d.h. $A \in \mathrm{SO}(3)$.[8]

Da A in $\mathrm{SO}(3)$ liegt, muss die Isometrie-Normalform wie folgt aussehen:

$$\widetilde{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{mit } \alpha \in [0,\pi].$$

Also ist A eine Drehung um den Winkel α um eine Achse im $\mathbb{R}^3$. Da wir hier $\alpha = 0$ und $\alpha = \pi$ ebenfalls erlauben, ist die Identität (Drehung um $\alpha = 0$) und die Spiegelung an einer Geraden (Drehung um $\alpha = \pi$) auch theoretisch erlaubt.

Wie kann man nun mit möglichst wenig Rechnung den Winkel α bestimmen, wenn man nur die Matrix A oben gegeben hat?

Da sich die Spur bei einem Basiswechsel nicht ändert (Bemerkung 2.6.14), muss gelten:

$$\operatorname{tr}(A) = \operatorname{tr}(\widetilde{A}).$$

Die Spur von A ist $\frac{1}{3} + \frac{1}{3} - \frac{1}{3} = \frac{1}{3}$; die Spur von $\widetilde{A}$ ist $1 + 2\cos(\alpha)$, es muss also gelten:

$$\frac{1}{3} = 1 + 2\cos(\alpha).$$

Dies können wir (weil α im Intervall $[0,\pi]$ liegt, auf dem der Kosinus injektiv ist) eindeutig nach α auflösen und erhalten:

$$\alpha = \arccos(-\frac{1}{3})$$

[8]Beachten Sie auch Bemerkung 2.2.9 über die wichtigsten acht Matrizengruppen und ihre Beziehung untereinander.

als Drehwinkel.[9]

Wollen wir auch die Drehachse berechnen, so müssen wir nur einen Eigenvektor zum Eigenwert 1 bestimmen (Gauß).

Bemerkung 2.6.16 (Berechnung der Isometrie-Normalform)**.** Theoretisch ist es möglich, zu gegebenem Endomorphismus die Isometrie-Normalform und die Orthonormalbasis, bezüglich derer die Matrix die gegebene Form hat, wie im Beweis oben zu bestimmen. Oft hilft aber auch die folgende Beobachtung:

Gegeben sei ein n-dimensionaler Euklidischer Vektorraum und eine lineare Isometrie $\varphi : V \to V$. Nach Satz 2.6.11 gibt es nun eine Orthonormalbasis B von V mit

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_r} \end{pmatrix},$$

wobei jeder Block entweder ein Drehkästen, oder 1 oder -1 ist.

Nun hat aber φ auch eine adjungierte Abbildung φ^*. Die Darstellungsmatrix von φ^* ist nun einfach die konjugiert transponierte Matrix von der von φ. Und weil alle Einträge reell sind, können wir uns das Konjugieren sparen und erhalten die Darstellung

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi^*) = \begin{pmatrix} \boxed{A_1^\top} & & & \\ & \boxed{A_2^\top} & & \\ & & \ddots & \\ & & & \boxed{A_r^\top} \end{pmatrix},$$

Das Transponieren der Blöcke, die 1 oder -1 sind, ändert nichts, aber wenn ein Drehkästchen transponiert wird, dann hat dies den folgenden Effekt:

$$\begin{pmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix}^\top = \begin{pmatrix} \cos(\alpha_j) & \sin(\alpha_j) \\ -\sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix}$$

Was passiert nun, wenn wir φ und φ^* addieren? In LA1 haben wir gelernt, dass sich dann einfach nur die Darstellungsmatrizen addieren. In jedem Block, in dem vorher eine 1 oder eine -1 stand, steht nun eine 2 oder -2, aber bei den Drehkästchen passiert etwas bemerkenswertes:

$$\begin{pmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix}^\top + \begin{pmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix} = \begin{pmatrix} 2\cos(\alpha_j) & 0 \\ 0 & 2\cos(\alpha_j) \end{pmatrix}.$$

Wir sehen also: Wenn φ bezüglich B in Isometrie-Normalform ist, dann ist φ^* bezüglich B ebenfalls in Isometrie-Normalform und $\varphi^* + \varphi$ ist in Diagonalform, wobei die Diagonaleinträge 2, -2, bzw. $2\cos(\alpha_j)$ sind.

Dies kann man nun rückwärts anwenden: Man bestimmt die Eigenwerte von $\varphi^* + \varphi$. Diese sind nach Satz 2.6.9 immer reell, weil $\varphi^* + \varphi$ selbstadjungiert ist. Nun kann man anhand der Eigenwerte mit ihren entsprechenden Vielfachheiten (algebraische Vielfachheit ist ja gleich der geometrischen Vielfachheit) die Isometrie-Normalform von φ ablesen.

[9] Diesen Ausdruck kann man nicht mehr wirklich weiter vereinfachen und lässt ihn am Besten so stehen. Man kann höchstens noch $\arccos(-\frac{1}{3}) = \pi - \arccos\frac{1}{3}$ draus machen.

Dieser „Trick“ lässt sich nicht nur verwenden, um die Struktur der Isometrie-Normalform zu bestimmen, sondern auch um die dazugehörige Orthonormalbasis zu finden. Hier muss man aber aufpassen, sobald es mehrere Drehkästchen zum selben Winkel gibt.

In solchen Fällen kann man wie folgt vorgehen: Angenommen, für eine Zahl $\alpha \in (0,\pi)$ hat der Eigenwert $\lambda = 2\cos(\alpha)$ der Abbildung $\varphi^* + \varphi$ eine (algebraische=geometrische) Vielfachheit größer als 2. Dann wählt man einen Eigenvektor v von $\varphi^* + \varphi$ zum Eigenwert λ und wendet dann den ursprünglichen Endomorphismus φ darauf an. Da wir schon wissen, wie eine Isometrie-Normalform im Reellen grundsätzlich aussieht, wissen wir, dass $\varphi(v)$ so aussehen muss:

$$\varphi(v) = \cos(\alpha)v + \sin(\alpha)w$$

für einen Eigenvektor w von $\varphi^* + \varphi$, der senkrecht auf v steht und können somit w ablesen. Damit haben wir die ersten beiden Vektoren unserer Basis; nun können wir mit den verbliebenen Eigenvektoren von $\varphi^* + \varphi$ fortfahren.

3. Bilinearformen und quadratische Formen

3.1. Der Dualraum

Wir wollen uns nun mit der Theorie der Bilinearformen beschäftigen. Um diese systematisch zu untersuchen, ist es hilfreich, zuerst den Begriff des Dualraums einzuführen, der eigentlich schon in der LA1 hätte vorkommen sollen, aber dann aus Zeitgründen nicht durchgenommen wurde. Hier und im Folgenden ist $\mathbb{K}$ wieder ein beliebiger Körper, wir beschränken uns also nicht mehr nur auf $\mathbb{R}$ oder $\mathbb{C}$.

Wir haben in der LA1 (Proposition 4.1.17) gesehen, dass die Menge $\mathrm{Hom}_{\mathbb{K}}(V,W)$, bestehend aus $\mathbb{K}$-linearen Abbildungen von einem $\mathbb{K}$-Vektorraum V in einen $\mathbb{K}$-Vektorraum W mit punktweiser Addition und skalarem Vielfachen selbst ein Vektorraum wird. Interessant sind zwei Spezialfälle: Mit dem Fall $W = V$, also mit dem Raum aller Endomorphismen eines Vektorraums V haben wir uns in der LA1 in Kapitel 5 beschäftigt. Das führte uns zu den Begriffen Eigenwert, Eigenvektor, Determinante, Spur, Diagonalisierbarkeit u.s.w.

Jetzt interessiert uns der Spezialfall $W = \mathbb{K}$, also der Raum aller linearen Abbildungen von V in den Grundkörper $\mathbb{K}$:

Definition 3.1.1.
Es sei $\mathbb{K}$ ein Körper und V ein Vektorraum. Eine *Linearform* (manchmal auch *lineares Funktional*) σ auf V ist ein Element in $\mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$, also eine lineare Abbildung $\sigma : V \to \mathbb{K}$.

Der Vektorraum $V^* := \mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$ aller Linearformen auf V heißt der *Dualraum* von V.

Bemerkung 3.1.2. Dualräume spielen vor allem in der Funktionalanalysis und der Differentialgeometrie eine wichtige Rolle.

Es sei $V := C^\infty([-1,1],\mathbb{R})$ der $\mathbb{R}$-Vektorraum aller beliebig oft differenzierbaren Funktionen auf $[-1,1]$. Weiterhin sei $f : [-1,1] \to \mathbb{R}$ eine stetige Funktion. Dann ist die Abbildung

$$\sigma_f : V \to \mathbb{R}, \quad h \mapsto \int_{-1}^{1} f(t)h(t)dt$$

linear, also ein Element in $V^* = \mathrm{Hom}_{\mathbb{K}}(V,\mathbb{R})$.

Man kann zeigen, dass beim Übergang der Funktion f zu der Linearform σ_f keine Information verloren geht, dass man also f rekonstruieren kann, wenn man nur σ_f kennt. Man kann also eine nichtlineare Funktion f ersetzen durch eine lineare Abbildung, die dann allerdings auf einem unendlichdimensionalen Vektorraum definiert ist. Auch wenn es auf den ersten Blick nicht klar ist, warum man das tun sollte, so hat sich diese Sichtweise oft als hilfreich erwiesen, weil es viele Funktionale auf V gibt, die nicht von dieser Form sind, die man aber als „verallgemeinerte Funktionen" auffassen kann.

Die bekannteste solche „verallgemeinerten Funktionen", die in der Elektrotechnik und der Physik immer wieder auftaucht, ist die „Dirac-Delta-Funktion" δ. Sie ist überall gleich 0, nur an der Stelle 0 ist sie unendlich groß und das Integral über diese „Funktion" ist gleich 1. Obgleich sehr hilfreich in den Anwendungen, war es für die Mathematik lange Zeit nicht klar, wie man diese Idee formalisieren kann. Denn eine „echte" Funktion, die überall 0 ist bis auf einen Punkt, kann niemals Integral 1 haben – das lässt die Integrationstheorie nicht zu. Allerdings wird diese Delta-„Funktion“ auch nie wie eine normale Funktion benutzt, sondern immer nur als Faktor in einem Integral der Form $\int_{-1}^{1}\delta(t)h(t)dt$, sodass man das Problem dann dadurch lösen konnte, indem man sagt, die Dirac-Delta-„Funktion“ und andere „verallgemeinerte Funktionen“ sind gar keine richtigen Funktionen, sondern nur Linearformen auf V (sogenannte *Distributionen*).

Dies ist z.B. in der Theorie der partiellen Differentialgleichungen von Vorteil, wo es wesentlich einfacher ist, eine Lösung im Raum der „verallgemeinerten Funktionen“ (Distributionen) zu finden als in dem kleineren Raum der echten Funktionen.

Proposition 3.1.3 (Dualisierung von linearen Abbildungen)**.**
Es seien V und W Vektorräume über demselben Grundkörper $\mathbb{K}$.

(a) Gegeben sei eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$. Dann ist die Abbildung

$$\varphi^* : W^* \to V^*, \quad \sigma \mapsto \sigma \circ \varphi,$$

die eine Linearform auf W mit φ verkettet, selbst eine $\mathbb{K}$-lineare Abbildung.

(b) Ist $\varphi : V \to W$ surjektiv, so ist $\varphi^* : W^* \to V^*$ injektiv.

(c) Ist $\varphi : V \to W$ bijektiv, so ist $\varphi^* : W^* \to V^*$ bijektiv.

(d) Es gilt

$$(\mathrm{id}_V)^* = \mathrm{id}_{V^*}.$$

(e) Es sei U ein Vektorraum über $\mathbb{K}$ und $\varphi : V \to W$ und $\psi : U \to V$ lineare Abbildungen. Dann gilt:

$$\big(\varphi \circ \psi\big)^* = \psi^* \circ \varphi^*.$$

Beweis. (a)
Wir müssen zeigen:

$$\begin{aligned} &\forall \sigma, \tau \in W^* : \quad \varphi^*(\sigma + \tau) = \varphi^*(\sigma) + \varphi^*(\tau) \\ \text{und} \quad &\forall \sigma \in W^*, \lambda \in \mathbb{K} : \quad \varphi^*(\lambda\sigma) = \lambda\varphi^*(\sigma). \end{aligned}$$

Da beide Seiten der ersten zu zeigenden Gleichung Elemente in V^* sind, zeigt man die Gleichheit, indem man zeigt, dass für jeden beliebigen Vektor $v \in V$ den gleichen Wert erhält, wenn man ihn sowohl in die linke als auch die rechte Seite einsetzt. Dies zeigen wir nun:

$$\begin{aligned} \big(\varphi^*(\sigma + \tau)\big)(v) &= \big((\sigma + \tau) \circ \varphi\big)(v) \\ &= (\sigma + \tau)\big(\varphi(v)\big) \\ &= \sigma\big(\varphi(v)\big) + \tau\big(\varphi(v)\big) \\ &= (\sigma \circ \varphi)(v) + (\tau \circ \varphi)(v) \\ &= \big(\varphi^*(\sigma)\big)(v) + \big(\varphi^*(\tau)\big)(v) \\ &= \big(\varphi^*(\sigma) + \varphi^*(\tau)\big)(v). \end{aligned}$$

Analog gehen wir bei der zweiten zu zeigenden Gleichheit vor:

$$\begin{aligned}\big(\varphi^*(\lambda\sigma)\big)(v) &= \big((\lambda\sigma)\circ\varphi\big)(v)\\ &= (\lambda\sigma)\big(\varphi(v)\big)\\ &= \lambda\sigma\big(\varphi(v)\big)\\ &= \lambda(\sigma\circ\varphi)(v)\\ &= \lambda\left(\varphi^*(\sigma)\right)(v)\\ &= \left(\lambda\varphi^*(\sigma)\right)(v).\end{aligned}$$

(b)
Es sei $\sigma \in \ker(\varphi^*)$. Wir wollen zeigen, dass $\sigma = 0$ ist.

Dazu sei $w \in W$ ein beliebiger Vektor. Wir werden zeigen, dass $\sigma(w) = 0$ ist. Weil w beliebig gewählt war, folgt daraus dann, dass die Funktion σ die Nullfunktion ist.

Aus der Surjektivität von $\varphi : V \to W$ folgt nun, dass es ein $v \in V$ gibt mit $\varphi(v) = w$. Daraus folgt nun:

$$\sigma(w) = \sigma(\varphi(v)) = (\sigma\circ\varphi)(v) = (\underbrace{\varphi^*(\sigma)}_{=0})(v) = 0\,.$$

(c)
Wir wissen bereits aus (b), dass $\varphi^* : W^* \to V^*$ injektiv ist. Es bleibt zu zeigen, dass φ^* auch surjektiv ist. Es sei dazu $\sigma \in V^*$ gegeben. Wir setzen $\tau := \sigma\circ\varphi^{-1}$.

Dann ist

$$\varphi^*(\tau) = \tau\circ\varphi = \left(\sigma\circ\varphi^{-1}\right)\circ\varphi = \sigma\circ\big(\underbrace{\varphi^{-1}\circ\varphi}_{=\mathrm{id}_V}\big) = \sigma\,.$$

(d)
Es sei $\sigma \in V^*$ gegeben. Dann gilt:

$$(\mathrm{id}_V)^*(\sigma) = \sigma\circ\mathrm{id}_V = \sigma = \mathrm{id}_{V^*}(\sigma)\,.$$

(e)
Für jedes $\sigma \in W^*$ gilt:

$$(\varphi\circ\psi)^*(\sigma) = \sigma\circ(\varphi\circ\psi) = (\sigma\circ\varphi)\circ\psi = (\psi^*)(\sigma\circ\varphi) = (\psi^*)(\varphi^*(\sigma)) = (\psi^*\circ\varphi^*)(\sigma)\,.$$ □

Bemerkung 3.1.4. Die Abbildung $\varphi^* : W^* \to V^*$, die in Proposition 3.1.3 eingeführt wurde, nennt man auch die *duale Abbildung* zu φ. Die Notation φ^* ist – wie schon in Bemerkung 2.6.4 angemerkt – nicht eindeutig, weil sie (zumindest wenn V und W Vektorräume mit Skalarprodukt sind) sowohl für die duale Abbildung, als auch für die adjungierte Abbildung stehen könnte. Aus dem Kontext sollte es aber immer klar sein, was gemeint ist.

Proposition 3.1.5.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Die Menge aller Linearformen auf V, die auf U verschwinden

$$\left\{\sigma \in V^* \;\middle|\; \sigma|_U = 0\right\} \subseteq V^*$$

ist ein Untervektorraum von V^* und ist isomorph zum Dualraum des Faktorraums $(V/U)^*$ über den Isomorphismus

$$(V/U)^* \to \left\{\rho \in V^* \;\middle|\; \rho|_U = 0\right\}, \quad \sigma \mapsto q^*(\sigma) = \sigma\circ q,$$

wobei $q : V \to V/U$ die Quotientenabbildung ist.

Beweis. Die Quotientenabbildung $q : V \to V/U$ ist linear und surjektiv (LA1, Satz 4.5.1). Nach Proposition 3.1.3 ist die dualisierte Abbildung

$$q^* : (V/U)^* \to V^*, \quad \sigma \mapsto \sigma \circ q$$

injektiv.

Wie jede andere injektive Abbildung auch kann man diese Abbildung bijektiv machen, indem wir sie auf ihr Bild koeinschränkt (LA1, Bemerkung 1.3.14), das folgende ist also ein Isomorphismus von $\mathbb{K}$-Vektorräumen:

$$(V/U)^* \to \mathrm{Bild}(q^*), \quad \sigma \mapsto \sigma \circ q.$$

Was ist aber nun das Bild(q^*)? Eine Linearform $\rho \in V^*$ lässt sich genau dann als $\sigma \circ q$ schreiben (mit $\sigma \in (V/U)^*$), wenn $U = \ker(q) \subseteq \ker \rho$ (LA1, Satz 4.5.5) und das ist gleichbedeutend mit $\rho|_U = 0$. □

Lemma 3.1.6 (Auswertungsabbildung)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $v \in V$ ein Element. Dann ist die Auswertungsabbildung

$$\eta_V(v) : V^* \to \mathbb{K}, \quad \sigma \mapsto \sigma(v)$$

eine $\mathbb{K}$-Linearform definiert auf dem Dualraum V.

Proposition 3.1.7 (Bidual)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$.

(a) Die Abbildung $\eta_V : V \to (V^*)^*, \quad v \mapsto \eta_V(v)$, die jeden Vektor auf die dazugehörige Auswertungsabbildung abbildet, ist $\mathbb{K}$-linear.

(b) Es sei W ein weiterer Vektorraum über $\mathbb{K}$ und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Dann ist das folgende Diagramm kommutativ:

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\eta_V"'] & W \arrow[d, "\eta_W"] \\ (V^*)^* \arrow[r, "(\varphi^*)^*"'] & (W^*)^*, \end{tikzcd}$$

es gilt also:

$$(\varphi^*)^* \circ \eta_V = \eta_W \circ \varphi.$$

(c) Es gilt:

$$\left(\eta_V\right)^* \circ \eta_{V^*} = \mathrm{id}_{V^*}.$$

Der Raum $(V^*)^*$ wird auch als der Bidualraum *von V bezeichnet.*

Beweis. (a)

Wir wollen die folgende Gleichheit zeigen:

$$(\varphi^*)^* \circ \eta_V = \eta_W \circ \varphi.$$

Da beide Seiten Abbildungen von V nach $(W^*)^*$ sind, nehmen wir ein beliebiges $v \in V$ und zeigen, dass

$$\big((\varphi^*)^* \circ \eta_V\big)(v) = \big(\eta_W \circ \varphi\big)(v).$$

Beide Seiten dieser Gleichung sind nun Elemente in $(W^*)^* = \mathrm{Hom}_{\mathbb{K}}(W^*, \mathbb{K})$. Um zu zeigen, dass beide Seiten gleich sind, nehmen wir ein beliebiges $\sigma \in W^*$ und zeigen, dass

$$\Big(\big((\varphi^*)^* \circ \eta_V\big)(v)\Big)(\sigma) = \Big(\big(\eta_W \circ \varphi\big)(v)\Big)(\sigma).$$

Dies wollen wir nun zeigen, indem wir vorsichtig die linke Seite umformen:

$$\begin{aligned}
\Big(\big((\varphi^*)^* \circ \eta_V\big)(v)\Big)(\sigma) &= \Big(\big((\varphi^*)^*(\eta_V(v))\Big)(\sigma) \\
&= \Big(\eta_V(v) \circ (\varphi^*)\Big)(\sigma) \\
&= \Big(\eta_V(v) \circ (\varphi^*)\Big)(\sigma) \\
&= \eta_V(v)\big(\varphi^*(\sigma)\big) \\
&= \big(\varphi^*(\sigma)\big)(v) \\
&= \big(\sigma \circ \varphi\big)(v) \\
&= \sigma\big(\varphi(v)\big).
\end{aligned}$$

Nun werden wir die rechte Seite umformen, mit dem Ziel, bei dem gleichen Ausdruck $\sigma\big(\varphi(v)\big)$ anzukommen:

$$\begin{aligned}
\Big(\big(\eta_W \circ \varphi\big)(v)\Big)(\sigma) &= \Big(\eta_W\big(\varphi(v)\big)\Big)(\sigma) \\
&= \sigma\big(\varphi(v)\big).
\end{aligned}$$

(c)

Wir wollen nun noch die Gleichung

$$\big(\eta_V\big)^* \circ \eta_{V^*} = \mathrm{id}_{V^*}$$

beweisen. Beide Seiten sind Elemente in $\mathrm{Hom}_{\mathbb{K}}(V^*, V^*)$. Um zu zeigen, dass sie gleich sind, sei $\sigma \in V^*$ gegeben. Wir werden zeigen, dass

$$\Big(\big(\eta_V\big)^* \circ \eta_{V^*}\Big)(\sigma) = \sigma.$$

Beide Seiten dieser Gleichung sind nun Elemente in $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. Es sei also $v \in V$ gegeben. Wir werden zeigen, dass

$$\bigg(\Big(\big(\eta_V\big)^* \circ \eta_{V^*}\Big)(\sigma)\bigg)(v) = \sigma(v).$$

Dies zeigen wir mit folgenden Umformungen:

$$\begin{aligned}
\bigg(\Big(\big(\eta_V\big)^* \circ \eta_{V^*}\Big)(\sigma)\bigg)(v) &= \Big(\big(\eta_V\big)^*\big(\eta_{V^*}(\sigma)\big)\Big)(v) \\
&= \Big(\big(\eta_{V^*}(\sigma)\big) \circ \eta_V\Big)(v) \\
&= \big(\eta_{V^*}(\sigma)\big)\big(\eta_V(v)\big) \\
&= \eta_V(v)(\sigma) \\
&= \sigma(v).
\end{aligned}$$

□

Satz 3.1.8.
Es sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum mit Basis $B = \{b_i \mid i \in I\}$. Wir nehmen an, $b_i \neq b_j$ für $i \neq j$.

(a) *Für jedes $j \in I$ ist die Koordinatenabbildung*

$$b_j^* : V \to \mathbb{K}, \quad \sum_{i \in I} \lambda_i b_i \mapsto \lambda_j,$$

die einen Vektor auf seine b_j-Koordinate abbildet, linear. Für unendliches I ist diese Summe ist so zu verstehen, dass immer nur endlich viele Skalare ungleich 0 sind.

(b) *Die Menge $\{b_i^* \mid i \in I\}$ ist eine linear unabhängige Teilmenge von V^*.*

(c) *Der Dualraum V^* ist isomorph zu $\mathbb{K}^B$, dem Raum aller Funktionen von B nach $\mathbb{K}$ (und damit auch isomorph zu $\mathbb{K}^I$).*

(d) *Die Abbildung*

$$\eta_V : V \to (V^*)^*, \quad v \mapsto \sigma \mapsto \sigma(v),$$

ist injektiv.

Beweis. (a)
Für ein $j \in I$ betrachten wir die Funktion

$$f_j : B \to \mathbb{K}, \quad b_i \mapsto \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst.} \end{cases}$$

Nach dem Fortsetzungssatz für lineare Abbildungen (LA1, Satz 4.2.14) gibt es eine lineare Fortsetzung $\varphi : V \to \mathbb{K}$ und man sieht leicht, dass dies genau die Koordinatenabbildung b_j^* ist, die somit wohldefiniert und linear ist.

(b)
Eine Menge ist linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.

Es sei deshalb

$$\sum_{j \in I} \lambda_j b_j^* = 0, \tag{$*$}$$

wobei nur endlich viele λ_j nicht 0 sind. Wir müssen zeigen, dass alle $\lambda_j = 0$ sind.

Sei dazu $i \in I$. Es ist zu zeigen, dass $\lambda_i = 0$ gilt.

Wir wenden nun die Gleichung $(*)$ an auf den Vektor $b_j \in V$:

$$\sum_{j \in I} \lambda_j b_j^*(b_i) = 0.$$

Die linke Seite der Gleichung wird nun einfach zu λ_i und damit folgt die Behauptung.

(c)
Dies folgt durch direkte Anwendung von LA1, Satz 4.2.14.

(d)
Es sei $v \in \ker(\eta_V)$, d.h. $\eta_V(v) = 0$. Wir wollen zeigen, dass $v = 0$ ist.

Weil $V = \mathrm{LH}_{\mathbb{K}}(B)$, gilt:

$$v = \sum_{i \in I} \lambda_i b_i,$$

wobei wieder nur endlich viele λ_i ungleich 0 sind. Wir wollen nun zeigen, dass alle λ_i null sind.

Es sei dazu $j \in I$ gegeben. Dann ist $b_j^* \in V^*$ und $\eta_V(v) \in (V^*)^*$. Es ist somit möglich, $\eta_V(v)$ auf b_j^* anzuwenden:

$$(\eta_V(v))(b_j^*) = b_j^*(v) = b_j^*\left(\sum_{i\in I} \lambda_i b_i\right) = \lambda_i.$$

Andererseits ist aber $(\eta_V(v))(b_j^*) = 0$, weil $\eta_V(v)$ die konstante Nullabbildung ist. Daraus folgt die Behauptung. □

Kommen wir nun zu dem endlichdimensionalen Fall:

Satz 3.1.9.
Es sei $\mathbb{K}$ ein Körper und V ein endlichdimensionaler Vektorraum über $\mathbb{K}$ mit geordneter Basis

$$\mathsf{B} = (b_1, \dots, b_n).$$

Dann bilden die Koordinatenabbildungen aus Satz 3.1.8

$$\mathsf{B}^* := \left(b_1^*, \dots, b_n^*\right)$$

eine geordnete Basis des Dualraums V^*, genannt die duale Basis. *Insbesondere gilt also* $\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$ *und* $V^* \cong_{\mathbb{K}} V$.

Für eine Linearform $\sigma : V \to \mathbb{K}$ gilt:

$$(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top.$$

Beweis. Nach LA1, Satz 4.3.11 (e) ist die Abbildung

$$V^* \to \mathbb{K}^{1\times n}, \quad \sigma \mapsto \mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)$$

ein $\mathbb{K}$-Vektorraumisomorphismus. Insbesondere gilt:

$$\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(\mathbb{K}^{1\times n}) = n.$$

Die Koordinatenabbildungen $b_1^*, \dots, b_n^*$ sind linear unabhängig nach Satz 3.1.8. Eine linear unabhängige n-elementige Menge in einem n-elementigen Vektorraum ist eine Basis.

Es bleibt zu zeigen, dass für ein $\sigma \in V^*$ die folgende Formel gilt:

$$(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top.$$

Wir setzen $\alpha_j := \sigma(b_j) \in \mathbb{K}$ für jedes $j \in \{1, \dots, n\}$.

Da in den Spalten der Darstellungsmatrix die Bilder der Basisvektoren stehen, folgt somit:

$$\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma) = \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix} \in \mathbb{K}^{1\times n}.$$

Definieren wir nun die Linearform $\rho := \alpha_1 b_1^* + \cdots + \alpha_n b_n^* \in V^*$, dann sehen wir durch direktes Einsetzen, dass

$$\rho(b_j) = \alpha_j \quad \text{für alle } j \in \{1, \dots, n\}.$$

Die linearen Abbildungen σ und ρ stimmen somit auf einer Basis von V überein, somit müssen sie gleich sein (siehe LA1, Satz 4.2.14). Es gilt also:

$$\sigma = \rho = \alpha_1 b_1^* + \cdots + \alpha_n b_n^*.$$

Also sind die Koordinaten von σ aufgefasst als Element im Vektorraum V^* bezüglich der Basis $B^* = \left(b_1^*, \ldots, b_n^*\right)$ genau die Skalare $\alpha_1, \ldots, \alpha_n$ und der Darstellungsvektor von σ bezüglich B^* ist also

$$(\sigma)_{\mathsf{B}^*} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

und dies endet den Beweis. □

Proposition 3.1.10 (Transponieren entspricht Dualisieren)**.**
Es seien V und W endlichdimensionale Vektorräume über einem Körper $\mathbb{K}$. Weiterhin sei eine geordnete Basis B von V, sowie eine geordnete Basis C von W gegeben. Die dazugehörigen dualen Basen von V^* und W^* werden mit B^* und C^* bezeichnet.

Für jede lineare Abbildung $\varphi : V \to W$ und ihre duale Abbildung $\varphi^* : W^* \to V^*$ gilt:

$$\mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right) = \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top .$$

Das Dualisieren einer Abbildung entspricht also dem Transponieren der Matrix.

Beweis. Es sei $\sigma \in W^*$ ein beliebiges Element im Dualraum von W^*. Dann gilt:

$$\begin{aligned} \mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right)(\sigma)_{C^*} &= \left(\varphi^*(\sigma)\right)_{\mathsf{B}^*} \\ &= \left(\sigma \circ \varphi\right)_{\mathsf{B}^*} \\ &= \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}\left(\sigma \circ \varphi\right)\right)^\top \\ &= \left(\mathrm{M}_{\mathsf{E},\mathsf{C}}(\sigma) \cdot \mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \\ &= \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \cdot \left(\mathrm{M}_{\mathsf{E},\mathsf{C}}(\sigma)\right)^\top \\ &= \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \cdot (\sigma)_{C^*} . \end{aligned}$$

Da $\sigma \in W^*$ beliebig war, gilt somit:

$$\mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right) = \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top .$$ □

Proposition 3.1.11 (Bidual endlichdimensional)**.**
Es sei V ein endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$. Dann ist die Abbildung

$$\eta_V : V \to (V^*)^*, \quad v \mapsto \eta(v) : \sigma \mapsto \sigma(v)$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Man sagt: Der Bidualraum eines endlichdimensionalen Vektorraum V ist natürlich isomorph zu V.

Beweis. Nach Satz 3.1.8(d) ist $\eta_V : V \to (V^*)^*$ injektiv, d.h. Bild(η_V) ist isomorph zu V.

Nach Satz 3.1.9 gilt: $\dim_{\mathbb{K}}((V^*)^*) = \dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$.

Also ist Bild(η_V) ein Untervektorraum von $(V^*)^*$ mit der gleichen (endlichen) Dimension. Demnach muss Bild(η_V) bereits der ganze Raum $(V^*)^*$ sein und η_V ist auch surjektiv. □

Bemerkung 3.1.12. Für einen endlichdimensionalen Vektorraum gilt auch $V \cong_{\mathbb{K}} V^*$. Trotzdem ist der Dualraum V^* nicht *natürlich isomorph* zu V. Dies liegt daran, dass ein solcher Isomorphismus immer von der Wahl einer Basis abhängt. Der Isomorphismus zwischen V und dem

Bidualraum $(V^*)^*$ ist dagegen *natürlich*, weil man ihn explizit ohne willkürliche Wahl einer Basis schreiben kann.

Der Versuch, eine formal korrekte Definition dieser intuitiven Idee einer *natürlichen* Abbildung zu geben führte zu einem Gebiet der Mathematik, das man als *Kategorientheorie* nennt. Dort werden Begriffe wie *Objekte*, *Morphismen* und *Funktoren* eingeführt, um schließlich eine korrekte Definition einer *natürlichen Transformation* geben zu können. All dies würde für diese Veranstaltung zu weit führen. Es ist aber durchaus möglich, dass Ihnen die Kategorientheorie früher oder später über den Weg läuft: Viele Gebiete der Mathematik werden in der Sprache der Kategorientheorie formuliert; außerdem hat die Kategorientheorie Gebiete der theoretischen Informatik und der funktionalen Programmierung[1] stark beeinflusst.

Beispiel 3.1.13.
Es sei $\mathbb{K} = \mathbb{R}$ und $V := \mathbb{K}^{(\mathbb{N})}$ der Raum der abbrechenden Folgen (siehe LA1, Beispiel 4.2.5(e) oder LA1, Beispiel 4.2.11(d)). Dann ist $\{e_k \mid k \in \mathbb{N}\}$ eine unendliche Basis für diesen Vektorraum.

Die Koordinatenabbildungen

$$e_k^* : V \to \mathbb{R}, \quad (x_j)_{j\in\mathbb{N}} \mapsto x_k$$

bilden nun eine abbrechende Folge $(x_j)_{j\in\mathbb{N}}$ auf den k-ten Folgenterm ab. Nach Satz 3.1.8(b) ist die Menge $L := \{e_k^*\}\, k \in \mathbb{N}$ linear unabhängig in V^*. Allerdings ist L *keine* Basis für V^*, weil L kein Erzeugendensystem ist.

Beispielsweise ist die Linearform

$$\Sigma : V \to \mathbb{R}, \quad (x_j)_j \mapsto \sum_{j=1}^{\infty} x_j,$$

die eine abbrechende Folge auf ihre Summe abbildet, keine Linearkombination der Linearformen aus L.

Allgemein gibt es sehr viele Linearformen in V^*, die nicht in $\mathrm{LH}_{\mathbb{R}}(L)$ sind. Für jede beliebige (nicht notwendigerweise abbrechende) Folge $(y_k)_{k\in\mathbb{N}}$ ist das folgende eine Linearform auf V

$$\Sigma_{(y_j)_j} : V \to \mathbb{R}, \quad (x_j)_j \mapsto \sum_{j=1}^{\infty} y_j x_j.$$

Man beachte, dass alle auftauchenden Summen immer endliche Summen sind, weil immer nur endlich viele Terme nicht null sind. Man kann relativ leicht nachweisen, dass alle Linearformen auf V von diesem Typ sind, es gilt:

$$\mathbb{R}^{\mathbb{N}} \to V^*, \quad (y_j)_j \mapsto \Sigma_{(y_j)_j}$$

ist ein Isomorphismus.

Der Dualraum von $\mathbb{R}^{(\mathbb{N})}$ ist also isomorph zu $\mathbb{R}^{\mathbb{N}}$.

Allgemeiner gilt: Für jede Menge J und jeden Körper $\mathbb{K}$ ist der Dualraum von $\mathbb{K}^{(J)}$ der Raum $\mathbb{K}^J$, wenn wir also den Raum der Funktionen, die nur an endlich vielen Stellen ungleich 0 sind, dualisieren, erhalten wir den Raum *aller* Funktionen von J nach $\mathbb{K}$. Falls J unendlich viele Elemente hat, gibt es somit immer Elemente im Dualraum, die keine Linearkombination der Koordinatenabbildungen sind.

Kurz gesagt: Das Konzept einer *dualen Basis* existiert nur im Endlichdimensionalen.

[1] Wenn Sie mal ganz viel Zeit haben, lernen Sie HASKELL … (siehe auch Fußnote 4 auf Seite 69)

Zusammenfassung von Abschnitt 3.1

(1) Zu jedem Vektorraum V ist der Dualraum V^* der Raum aller linearen Abbildungen von V in den Grundkörper $\mathbb{K}$.

(2) Zu jeder geordneten Basis B eines endlichdimensionalen Vektorraums V gibt es eine duale Basis B^* des Dualraums V^*. Es gilt: $V \cong V^*$ für endlichdimensionale Vektorräume.

(3) Es gilt die Formel: $(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top$

(4) Im Unendlichdimensionalen gibt es keine duale Basis.

(5) Jede lineare Abbildung $\varphi : V \to W$ lässt sich dualisieren zu einer dualen linearen Abbildung $\varphi^* : W^* \to V^*$.

(6) Wenn wir für V und W geordnete Basen wählen und die Dualräume mit den entsprechenden dualen Basen versehen, so ist die Darstellungsmatrix der dualen Abbildung φ^* genau die Transponierte der Darstellungsmatrix von φ.

(7) Es gibt eine natürliche lineare Abbildung $\eta_V : V \to (V^*)^*$ von V in seinen Bidual $(V^*)^*$.

(8) $(\mathbb{K}^{(J)})^* \cong \mathbb{K}^J$.

3.2. Bilinearformen über beliebigen Körpern

In LA1 (Definition 5.2.1 im Kapitel über Determinanten) haben wir bereits multilineare Abbildungen kennengelernt. Zur Erinnerung: Es seien V und W Vektorräume über einem Körper $\mathbb{K}$ und es sei $n \in \mathbb{N}$. Eine Abbildung $\omega : V^n = V \times \cdots \times V \to W$ heiße *multilinear*, falls sie in jeder einzelnen Komponente linear ist.

Wir möchten uns nun auf in zweierlei Hinsicht einschränken auf $n = 2$ und $W = \mathbb{K}$. Den Grundkörper $\mathbb{K}$ werden wir – im Gegensatz zu allen bisherigen Kapiteln in der Linearen Algebra II – vorerst noch allgemein lassen, d.h. es sind nun nicht nur $\mathbb{R}$ und $\mathbb{C}$ erlaubt, sondern auch $\mathbb{Q}$, endliche Körper und sämtliche andere Körper, die in der Vorlesung bis jetzt nicht erwähnt wurden.

Definition 3.2.1.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine *Bilinearform* ist eine Abbildung

$$\beta : V \times V \to \mathbb{K},$$

die in jeder Komponente linear ist.

Die Menge aller Bilinearformen auf V bezeichnen wir mit

$$\mathrm{Bil}(V,\mathbb{K}) := \left\{\beta : V \times V \to \mathbb{K} \;\middle|\; \beta \text{ ist bilinear}\right\}.$$

Man rechnet leicht nach, dass $\mathrm{Bil}(V,\mathbb{K})$ ein Untervektorraum von $\mathbb{K}^{V\times V}$ ist und somit ein $\mathbb{K}$-Vektorraum.

Für jede Bilinearform $\beta \in \mathrm{Bil}(V,\mathbb{K})$ ist

$$q_\beta : V \to \mathbb{K}, \quad v \mapsto \beta(v,v)$$

die zu β dazugehörige *quadratische Form*.

Beispiel 3.2.2. (a) Für $V = \mathbb{K}^2$ ist die *Determinantenform*

$$\omega : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}, \quad \left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \right) \mapsto \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

eine Bilinearform. Falls $\mathbb{K} = \mathbb{R}$ gibt sie den orientierten Flächeninhalt des Parallelogramms an, das von den beiden Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ aufgespannt wird. Das war in LA1 unsere „Motivation“, uns mit alternierenden Abbildungen zu beschäftigen.

Beachten Sie aber, dass für $n > 2$ die Determinantenform nicht mehr bilinear, also kein Beispiel mehr für eine Bilinearform ist.

Die zur Determinantenform dazugehörige quadratische Form ist die konstante Nullabbildung, weil $\omega(v, v) = 0$ ist.

Genaugenommen ist eine Bilinearform genau dann alternierend, wenn die dazugehörige quadratische Form konstant 0 ist.

(b) Es sei $\mathbb{K} = \mathbb{R}$ und V ein Euklidischer Vektorraum. Dann ist das Skalarprodukt

$$\langle \cdot, \cdot \rangle : V \times V \to V$$

eine Bilinearform auf V nach Lemma 1.2.5 (b).

Die dazugehörige quadratische Form ist gegeben durch das Quadrat der Norm.

Beachten Sie aber, dass für $\mathbb{K} = \mathbb{C}$ ein Skalarprodukt *nicht* bilinear ist, sondern nur sesquilinear, d.h. in der zweiten Komponente ist das Skalarprodukt nur $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

(c) Für einen beliebigen Körper $\mathbb{K}$ können wir die Konstruktion des reellen Standardskalarprodukt imitieren und definieren:

$$\beta : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}, \quad \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \mapsto \sum_{j=1}^{n} x_j y_j.$$

Für $\mathbb{K} = \mathbb{R}$ ist dies gerade das Standardskalarprodukt und die dazugehörige quadratische Form ist das Quadrat der Standradnorm, aber für andere Körper hat diese Bilinearform und ihre dazugehörige quadratische Form unerwartete Eigenschaften, beispielsweise gilt für $\mathbb{K} = \mathbb{F}_2$ und $n = 2$:

$$q_\beta \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \beta \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = 0.$$

Das gleiche passiert für $\mathbb{K} = \mathbb{C}$ und den Vektor $\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} \in \mathbb{C}^2$.

(d) In der Relativitätstheorie spielt der sogenannte *Minkowski-Raum* $(\mathbb{R}^4, \beta)$ eine besondere Rolle. Hierbei handelt es sich um den reellen Vektorraum

$$\mathbb{R}^4 = \left\{ \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \; t, x_1, x_2, x_3 \in \mathbb{R} \right\},$$

der mit der folgenden Bilinearform versehen ist:

$$\beta\left(\begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} u \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) := -tu + x_1y_1 + x_2y_2 + x_3y_3 \in \mathbb{R}.$$

Dies ist kein Skalarprodukt, weil $\beta\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right) < 0$ ist.

Die erste Komponente des Minkowski-Raums steht für die Zeit und die restlichen drei stehen für die drei Ortskoordinaten, sodass ein Element im Minkowski-Raum ein *Ereignis* mit Zeit- und Raumkomponente darstellt. Die Bilinearform wird eingeschränkt auf den Raum der drei Raumdimensionen das Euklidische Standardskalarprodukt.

Die zu β gehörige quadratische Form nimmt sowohl positive als auch negative Werte an: Vektoren $v \in \mathbb{R}^4$ mit $q_\beta(v) > 0$ heißen *raumartige* Vektoren, Vektoren $v \in \mathbb{R}^4$ mit $q_\beta(v) < 0$ heißen *zeitartige* Vektoren, Vektoren $v \in \mathbb{R}^4$ mit $q_\beta(v) = 0$ heißen *lichtartige Vektoren*.

(e) Eine wichtige Anwendung von Bilinearformen kommt aus der Analysis.[2] Gegeben sei eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$, die zweimal stetig differenzierbar ist und ein Punkt $x_0 \in \mathbb{R}^n$. Die „erste Ableitung“ von f am Punkt x_0 ist eine Linearform $\varphi : \mathbb{R}^n \to \mathbb{R}$, also ein Element im Dualraum von $\mathbb{R}^n$.

$$\varphi(v) := \frac{\partial f}{\partial v}(x_0) := \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}$$

Die „zweite Ableitung“ von f am Punkt x_0 ist eine Bilinearform $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:

$$\beta(v,w) := \lim_{s\to 0} \frac{\frac{\partial f}{\partial v}(x_0 + sw) - \frac{\partial f}{\partial v}(x_0)}{s}.$$

Dass diese Abbildung wirklich bilinear ist, ist nicht vollkommen offensichtlich, folgt aber aus der zweimaligen stetigen Differenzierbarkeit. Man kann weiterhin zeigen, dass diese Bilinearform sogar symmetrisch (siehe Definition 3.3.1) ist (Satz von Schwarz). Die Fundamentalmatrix (siehe Definition 3.2.6) der Bilinearform ist die sogenannte *Hesse-Matrix*.

Die dazugehörige quadratische Form spielt nun eine wichtige Rolle im 2. Taylor-Polynom von f, also der quadratischen Näherung an die Funktion in der Nähe des Punktes x_0. Es gilt nämlich:

$$f(x_0 + v) \approx f(x_0) + \varphi(v) + \frac{1}{2} q_\beta(v), \quad \text{falls } \|v\| \text{ klein ist.}$$

Bemerkung 3.2.3 (Currying)**.** Bilinearformen $\beta : V \times V \to \mathbb{K}$ sind keine lineare Abbildungen und von daher nicht direkt mit Methoden der linearen Algebra zu untersuchen. Es gibt aber einen „Trick“: Wenn Sie eine Funktion

$$f : X \times Y \to Z$$

von zwei Variablen gegeben haben, erhalten Sie – wenn Sie eine der Komponenten, z.B. $y \in Y$ festhalten – eine Funktion einer Variablen $f(\cdot, y) : X \to Z$, die einfacher zu handhaben ist. Sie

[2] Diese Überlegungen hier sind selbstverständlich nicht relevant für diese LA-2-Veranstaltung, sondern sollen nur die Motivation für die Beschäftigung mit den hier eingeführten Begriffen erhöhen.

haben also eine Möglichkeit, jedem Wert $y \in Y$ eine Funktion von X nach Z zuzuweisen. Dafür schreibt man auch

$$f^\vee : Y \to Z^X, \quad y \mapsto f(\cdot, y).$$

Hierbei steht Z^X die Menge aller Funktionen von X nach Z und es stellt sich heraus, dass hierbei keine Information verloren geht. Man kann also eine Funktion $f \in Z^{X\times Y}$ auffassen als eine Funktion $f^\vee \in (Z^X)^Y$, die jedem Wert aus y eine Funktion von X nach Z zuweist. Dieses Vorgehen wird *Currying*[3] genannt und spielt in der funktionalen Programmierung[4] eine größere Rolle.

Für uns hat das *Currying* den Vorteil eine bilineare Abbildung umzuwandeln in eine lineare Abbildung, die als Rückgabewert wieder eine lineare Abbildung hat:

Lemma 3.2.4.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Wie Defintion 3.1.1 eingeführt bezeichnen wir den Dualraum von V mit $V^* := \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$.

(a) *Es sei $\beta : V \times V \to \mathbb{K}$ eine Bilinearform auf V. Dann ist für jeden Vektor $w \in V$ die Abbildung*

$$\beta(\cdot, w) : V \to \mathbb{K}, \quad v \mapsto \beta(v, w)$$

linear, d.h. $\beta(\cdot, w) \in V^*$.

(b) *Es sei $\beta : V \times V \to \mathbb{K}$ eine Bilinearform auf V. Dann ist*

$$\beta^\vee : V \to V^*, \quad w \mapsto \beta(\cdot, w)$$

linear. Jede Bilinearform auf V lässt sich also natürlich auf eine lineare Abbildung von V nach V^* abbilden.

(c) *Die Abbildung, die eine Bilinearform β auf $\beta^\vee$ abbildet, ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen:*

$$\mathrm{Bil}(V, \mathbb{K}) \to \mathrm{Hom}(V, V^*), \quad \beta \mapsto \beta^\vee.$$

Beweis. Vorbemerkung: Auch wenn dieser Beweis relativ lang ist, so sollte das nicht darüber hinweg täuschen, dass eigentlich alles direktes Einsetzen ist und alle Aussagen ziemlich direkt aus den Definitionen folgen. Wenn der Beweis teilweise schwer zu verstehen ist, so liegt das nicht daran, dass irgendwo geniale Umformungen, Abschätzungen und Wahlen vorkommen, sondern nur daran, dass die verwendeten Objekte relativ abstrakt sind.

(a)
Dies folgt direkt aus der Definition der Bilinearität (Definition 3.2.1).

(b)
Wir zeigen zuerst die Additivität. Es seien dazu $w_1, w_2 \in V$ gegeben. Wir werden zeigen:

$$\beta^\vee(w_1 + w_2) = \beta^\vee(w_1) + \beta^\vee(w_2).$$

Da beide Seiten der zu zeigenden Gleichungen Elemente in $V^* = \mathrm{Hom}(V, \mathbb{K})$ sind, zeigen wir die Gleichheit, indem wir ein beliebiges Element $v \in V$ vorgeben und zeigen, dass folgendes gilt:

$$\beta^\vee(w_1 + w_2)(v) = \big(\beta^\vee(w_1) + \beta^\vee(w_2)\big(v).$$

[3] In der Linguistik wohl auch *Schönfinkeln* genannt (Nein, ich habe mir das Wort nicht ausgedacht)
[4] Lernen Sie HASKELL, die schönste Programmiersprache aller Zeiten (siehe auch Fußnote 1 auf Seite 65) :-)

Dies folgt aus der folgenden Rechnung, in der wir die Additivität von β in der zweiten Komponente verwenden:

$$\begin{aligned}\beta^\vee(w_1+w_2)(v) &= \beta(v,w_1+w_2)\\ &= \beta(v,w_1)+\beta(v,w_2)\\ &= \beta^\vee(w_1)(v)+\beta^\vee(w_2)(v)\\ &= \big(\beta^\vee(w_1)+\beta^\vee(w_2)\big(v\big).\end{aligned}$$

Ebenso zeigt man für $w \in V$ und $\lambda \in \mathbb{K}$, dass für jedes $v \in V$ gilt:

$$\begin{aligned}\big(\beta^\vee(\lambda w)\big)(v) &= \beta(v,\lambda w)\\ &= \lambda\beta(v,w)\\ &= \lambda\big(\beta^\vee(w)(v)\big)\\ &= \big(\lambda\beta^\vee(w)\big)(v).\end{aligned}$$

Da $v \in V$ beliebig war, folgt $\beta^\vee(\lambda w) = \lambda\beta^\vee(w)$ und die Linearität von $\beta^\vee$ ist bewiesen.

(c)

Wir haben in Teil (b) gesehen, dass für jedes $\beta \in \mathrm{Bil}(V,\mathbb{K})$ gilt, dass $\beta^\vee \in \mathrm{Hom}(V,V^*)$ ist. Die folgende Abbildung ist also wohldefiniert:

$$\Phi : \mathrm{Bil}(V,\mathbb{K}) \to \mathrm{Hom}(V,V^*), \quad \beta \mapsto \beta^\vee.$$

Es bleibt zu zeigen, dass Φ linear und bijektiv ist.

Wir zeigen zuerst die Linearität. Dazu seien Bilinearformenen $\beta,\gamma \in \mathrm{Bil}(V,\mathbb{K})$ gegeben. Wir werden zeigen:

$$\Phi(\beta+\gamma) = \Phi(\beta)+\Phi(\gamma).$$

Da beiden Seiten in $\mathrm{Hom}(V,V^*)$ sind, sei $w \in V$ gegeben. Wir werden zeigen:

$$\big(\Phi(\beta+\gamma)\big)(w) = \big(\Phi(\beta)+\Phi(\gamma)\big)(w).$$

Beide Seiten dieser Gleichung sind nun in V^*. Deshalb sei $v \in V$ gegeben.

$$\begin{aligned}\Big(\big(\Phi(\beta+\gamma)\big)(w)\Big)(v) &= \Big((\beta+\gamma)^\vee(w)\Big)(v)\\ &= (\beta+\gamma)(v,w)\\ &= \beta(v,w)+\gamma(v,w)\\ &= \big(\beta^\vee(w)\big)(v)+\big(\gamma^\vee(w)\big)(v)\\ &= \Big(\beta^\vee(w)+\gamma^\vee(w)\Big)(v)\\ &= \Big(\big(\Phi(\beta)\big)(w)+\big(\Phi(\gamma)\big)(w)\Big)(v)\\ &= \Big(\big(\Phi(\beta)+\Phi(\gamma)\big)(w)\Big)(v).\end{aligned}$$

Dies zeigt die Additivität von Φ. Die Homogenität, also $\Phi(\lambda\beta) = \beta\Phi(\lambda)$ zeigt man analog.

Es bleibt zu zeigen, dass die lineare Abbildung

$$\Phi : \mathrm{Bil}(V,\mathbb{K}) \to \mathrm{Hom}(V,V^*), \quad \beta \mapsto \beta^\vee$$

bijektiv ist.

Zeigen wir zuerst die Injektivität. Es sei dazu $\beta \in \ker \Phi$ gegeben. Wir wollen zeigen, dass $\beta = 0$ ist. Da $\beta : V \times V \to \mathbb{K}$ eine Funktion auf der Menge $V \times V$ ist, nehmen wir ein beliebiges Paar $(v,w) \in V \times V$ und zeigen, dass $\beta(v,w) = 0$ ist:

$$\beta(v,w) = \left(\beta^{\vee}(w)\right)^{\vee}(v) = \Big(\underbrace{\big(\Phi(\beta)\big)}_{=0}(w)\Big)(v) = 0\,.$$

Es bleibt, die Surjektivität zu zeigen. Es sei dazu $\varphi \in \operatorname{Hom}(V,V^*)$ gegeben. Wir wollen zeigen, dass es eine Bilinearform $\beta \in \operatorname{Bil}(V,\mathbb{K})$ gibt mit $\Phi(\beta) = \varphi$.

Da φ im Raum $\operatorname{Hom}(V,V^*)$ liegt, gilt, dass für jedes $w \in V$ der Wert $\varphi(w) \in V^*$ eine Linearform auf V ist. Wir definieren:

$$\beta : V \times V \to \mathbb{K}, \quad (v,w) \mapsto \big(\varphi(w)\big)(v)\,.$$

Es ist zu zeigen, dass β bilinear ist und dass $\Phi(\beta) = \varphi$.

Beginnen wir mit der Linearität in der ersten Komponente, diese folgt aus der Tatsache, dass für jedes feste $w \in V$ die Abbildung $\varphi(w) \in V^*$ eine Linearform ist:

$$\begin{aligned}\beta(\lambda_1 v_1 + \lambda_2 v_2, w) &= \big(\varphi(w)\big)(\lambda_1 v_1 + \lambda_2 v_2)\\ &= \lambda_1\big(\varphi(w)\big)(v_1) + \lambda_2\big(\varphi(w)\big)(v_2)\\ &= \lambda_1 \beta(v_1,w) + \lambda_2 \beta(v_2,w)\,.\end{aligned}$$

Zeigen wir nun die Linearität in der zweiten Komponente, diese folgt aus der Tatsache, dass die Abbildung $\varphi : V \to V^*$ linear ist:

$$\begin{aligned}\beta(v, \lambda_1 w_1 + \lambda_2 w_2) &= \big(\varphi(\lambda_1 w_1 + \lambda_2 w_2)\big)(v)\\ &= \big(\lambda_1 \varphi(w_1) + \lambda_2 \varphi(w_2)\big)(v)\\ &= \lambda_1\big(\varphi(w_1)\big)(v) + \lambda_2\big(\varphi(w_2)\big)(v)\\ &= \lambda_1 \beta(v,w_1) + \lambda_2 \beta(v,w_2)\,.\end{aligned}$$

Also ist $\beta : V \times V \to \mathbb{K}$ bilinear, also $\beta \in \operatorname{Bil}(V,\mathbb{K})$. Es bleibt zu zeigen, dass

$$\Phi(\beta) = \varphi\,.$$

Dazu sei nun $w \in V$ gegeben. Wir wollen zeigen, dass

$$\big(\Phi(\beta)\big)(w) = \varphi(w)\,.$$

Dazu sei nun $v \in V$ gegeben. Es bleibt zu zeigen, dass

$$\Big(\big(\Phi(\beta)\big)(w)\Big)(v) = \big(\varphi(w)\big)(v)\,.$$

Dies folgt nun aber mehr oder weniger direkt aus der Definition der Abbildung β:

$$\begin{aligned}\Big(\big(\Phi(\beta)\big)(w)\Big)(v) &= \Big(\beta^{\vee}(w)\Big)(v)\\ &= \beta(v,w)\\ &= \big(\varphi(w)\big)(v)\,.\end{aligned}$$

Das beendet den Beweis. □

Bemerkung 3.2.5. Achtung: Die Abbildung $\beta^\vee : V \to V^*$ ist zwar eine lineare Abbildung, aber *kein* Endomorphismus, weil V^* nicht der selbe Raum ist wie V. Wir können also unsere Erkenntnisse und „Werkzeuge“ wie Determinante, Spur, Eigenwerte nicht zur Untersuchung von Bilinearformen verwenden... Zumindest nicht direkt[5]...

Was hingegen problemlos möglich ist, ist einer Bilinearform β (beziehungsweise der dazugehörigen linearen Abbildung $\beta^\vee : V \to V^*$) eine Matrix zuzuweisen.

Definition 3.2.6 (Fundamentalmatrix)**.**
Es sei $n \in \mathbb{N}$ und β eine Bilinearform auf einem n-dimensionalen Vektorraum V über einem Körper $\mathbb{K}$.

Es sei weiterhin B eine geordnete Basis von V und B^* die dazugehörige duale Basis (siehe Definition 3.1.9) von V^*. Dann ist die *Fundamentalmatrix* von β definiert als

$$\mathrm{FM}_{\mathsf{B}}\left(\beta\right) := \mathrm{M}_{\mathsf{B}^*,\mathsf{B}}\left(\beta^\vee\right) \in \mathbb{K}^{n\times n}.$$

Proposition 3.2.7.
Gegeben sei eine natürliche Zahl $n \in \mathbb{N}$ und ein Körper $\mathbb{K}$.

(a) Es sei β eine Bilinearform auf einem n-dimensionalen $\mathbb{K}$-Vektorraum V mit geordneter Basis $\mathsf{B} = (b_1, \ldots, b_n)$. Es sei $A := \mathrm{FM}_{\mathsf{B}}\left(\beta\right) = \mathrm{M}_{\mathsf{B}^*,\mathsf{B}}\left(\beta^\vee\right)$ die Fundamentalmatrix zu β. Dann gilt für Vektoren $v, w \in V$:

$$\beta(v,w) = \beta^\vee(w)(v) = (v)_{\mathsf{B}}^\top \mathrm{FM}_{\mathsf{B}}\left(\beta\right)(w)_{\mathsf{B}}\,..$$

Der (i,j)-Eintrag von A ist gegeben durch $\beta(b_i, b_j) \in \mathbb{K}$.

(b) Es sei $\beta : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ eine Bilinearform auf dem Raum $\mathbb{K}^n$. Dann ist die Fundamentalmatrix von β bezüglich der Standardbasis gegeben durch

$$A = \mathrm{FM}_{\mathsf{E}}\left(\beta\right) = \left(\beta(e_i, e_j)\right)_{i,j\in\{1,\ldots,n\}}.$$

Für $x, y \in \mathbb{K}^n$ gilt:

$$\beta(x,y) = x^\top A y.$$

Beweis. Nach Satz 3.1.9 ist der Koordinatenvektor von einem Element des Dualraums $\sigma \in V^*$ die Transponierte der Darstellungsmatrix:

$$(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top.$$

Wenn nun ein $w \in V$ gegeben ist, dann ist $\beta^\vee(w)$ ein Element im Dualraum, es gilt also:

$$\left(\beta^\vee(w)\right)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}\left(\beta^\vee(w)\right)\right)^\top.$$

Diese Identität werden wir weiter unten verwenden.

Es sei nun ein weiterer Vektor $v \in V$ gegeben. Wir möchten $\beta(v,w)$ berechnen. Da $\beta(v,w) \in \mathbb{K}$ ein Skalar ist – und somit eine (1×1)-Matrix, können wir diesen Skalar transponieren, ohne

[5]Ein Beispiel, wo man Eigenwerte dann doch verwenden kann, um Bilinearformen zu untersuchen ist Satz 3.4.10.

den Wert zu ändern:

$$\begin{aligned}\beta(v,w) &= \Big(\beta(v,w)\Big)^\top \\ &= \Big((\beta^\vee(w))(v)\Big)^\top \\ &= \Big(\mathrm{M}_{\mathsf{E},\mathsf{B}}\left(\beta^\vee(w)\right)(v)_{\mathsf{B}}\Big)^\top \\ &= (v)_{\mathsf{B}}^\top \Big(\mathrm{M}_{\mathsf{E},\mathsf{B}}\left(\beta^\vee(w)\right)\Big)^\top \\ &= (v)_{\mathsf{B}}^\top \left(\beta^\vee(w)\right)_{\mathsf{B}^*} \\ &= (v)_{\mathsf{B}}^\top \mathrm{M}_{\mathsf{B}^*,\mathsf{B}}\left(\beta^\vee\right)(w)_{\mathsf{B}} \\ &= (v)_{\mathsf{B}}^\top \mathrm{FM}_{\mathsf{B}}\left(\beta\right)(w)_{\mathsf{B}}\,.\end{aligned}$$

Wenn wir nun diese Formel anwenden auf das Paar (b_i, b_j), so erhalten wir:

$$\beta(b_i,b_j) = (b_i)_{\mathsf{B}}^\top \mathrm{FM}_{\mathsf{B}}\left(\beta\right)\left(b_j\right)_{\mathsf{B}} = e_i^\top \mathrm{FM}_{\mathsf{B}}\left(\beta\right) e_j\,,$$

wobei e_i, e_j der i-te, bzw. der j-te Standardbasisvektor des $\mathbb{K}^n$ sind. Das Multiplizieren mit e_j von rechts gibt die j-te Spalte zurück (in den Spalten der Matrix stehen die Bilder der Basisvektoren) und das Multiplizeiren mit dem Zeilenvektor $e_i^\top$ gibt die i-te Zeile zurück.

(b)

Dies ist ein Spezialfall von Teil (a). □

Proposition 3.2.8 (Basiswechsel bei Fundamentalmatrizen)**.**
Es sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum der Dimension $n \in \mathbb{N}$. Gegeben sei eine Bilinearform $\beta \in \mathrm{Bil}(V,\mathbb{K})$, sowie zwei geordnete Basen B und C von V. Dann stehen die Fundamentalmatrizen von β bezüglich B und die bezüglich C in folgender Beziehung zueinander:

$$\mathrm{FM}_{\mathsf{C}}\left(\beta\right) = \left(\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\right)^\top \cdot \mathrm{FM}_{\mathsf{B}}\left(\beta\right) \cdot \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\,.$$

Achtung: Beachten Sie, dass die Matrix auf der linken Seite transponiert *und nicht* invertiert *ist!*

Fundamentalmatrizen von Bilinearformen transformieren sich also anders als Darstellungsmatrizen von Endomorphismen!!!

Beweis. Wir setzen:

$$A := \left(\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\right)^\top \cdot \mathrm{FM}_{\mathsf{B}}\left(\beta\right) \cdot \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}) \quad \text{und} \quad B := \mathrm{FM}_{\mathsf{C}}\left(\beta\right).$$

Wir wollen die Gleichung $A = B$ zeigen.

Beide Seiten der zu zeigenden Gleichung sind Elemente in $\mathbb{K}^{n\times n}$. Es reicht also zu zeigen, dass für alle $x, y \in \mathbb{K}^n$ gilt:

$$x^\top A y = x^\top B y\,.$$

Wir wählen $v, w \in V$ so, dass

$$x := (v)_{\mathsf{C}} \in \mathbb{K}^n \quad \text{und} \quad y := (w)_{\mathsf{C}} \in \mathbb{K}^n\,.$$

Dann gilt:

$$\begin{aligned}
x^\top A y &= (v)_{\mathsf{C}}^\top \left(\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\right)^\top \cdot \mathrm{FM}_{\mathsf{B}}\left(\beta\right) \cdot \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})(w)_{\mathsf{C}}\\
&= \left(\left(\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\right)(v)_{\mathsf{C}}\right)^\top \cdot \mathrm{FM}_{\mathsf{B}}\left(\beta\right) \cdot \left(\left(\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id})\right)(w)_{\mathsf{C}}\right)\\
&= (v)_{\mathsf{B}}^\top \cdot \mathrm{FM}_{\mathsf{B}}\left(\beta\right) \cdot \left((w)_{\mathsf{B}}\right)\\
&= \beta(v,w)\\
&= (v)_{\mathsf{C}}^\top \cdot \mathrm{FM}_{\mathsf{C}}\left(\beta\right) \cdot \left((w)_{\mathsf{C}}\right)\\
&= x^\top B y. \qquad \square
\end{aligned}$$

Bemerkung 3.2.9 (Fundamentalmatrizen transformieren mit Gauß-Schritten)**.** In der LA1 (genaugenommen in Lemma 2.5.4 LA1-Skript) haben wir gesehen, dass die drei *elementaren Zeilenumformungen* des Gauß-Algorithmus Matrixmultiplikationen von links entsprechen.

Analog dazu entspricht die Multiplikation der entsprechenden Transponierten von rechts der dazugehörigen Spaltenumformung.

Das Addieren des μ-fachen der j-ten Zeile auf die i-te Zeile wird beispielsweise erreicht durch die Multiplikation von links mit der $(m \times m)$-Matrix

$$\begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & & \ddots & & & \\
& & \mu & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{pmatrix},$$

wobei der Eintrag μ in der i-ten Zeile und in der j-ten Spalte steht.

Wenn Sie nun das μ-fache der j-ten Spalte auf die i-te Spalte addieren wollen[6], so erreichen Sie dies, in dem Sie die Matrix von rechts mit der folgenden Matrix multiplizieren:

$$\begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & \mu & & \\
& & & \ddots & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{pmatrix},$$

Beachten Sie, dass dies nach Proposition 3.2.8 genau die Operation ist, die beim Basiswechsel von Fundamentalmatrizen auftaucht.

Wenn Sie also die Fundamentalmatrix einer Bilinearform gegeben haben und dann eine elementare Zeilenumformung durchführen **und dabei gleichzeitig die entsprechende Spaltenumformung** anwenden, so erhalten Sie eine neue Fundamentalmatrix der selben Bilinearform,

[6] Beachten Sie aber, dass das im Gauß-Algorithmus nicht erlaubt ist, weil sich hierdurch die Lösungsmenge eines dazugehörigen linearen Gleichungssystems ändern kann!

nur bezüglich einer anderen Basis. Auf diese Weise ist es also möglich, mit einer Variante des Gauß-Algorithmus Fundamentalmatrizen zu vereinfachen.

Wie ändert sich aber nun die Basis, bezüglich derer wir die Matrix beschreiben?

Hier gilt folgendes: Es sei $\mathsf{B} = (b_1, \ldots, b_n)$ eine geordnete Basis für V ist und $A = \mathrm{FM}_{\mathsf{B}}(\beta)$ die Fundamentalmatrix einer Bilinearform $\beta \in \mathrm{Bil}(V, \mathbb{K})$. Wenn wir nun das μ-fache der j-ten Zeile auf die i-Zeile addieren und anschließend das μ-fache der j-ten Spalte auf die i-te Spalte addieren, dann erhalten wir die Fundamentalmatrix $\mathrm{FM}_{\mathsf{B}'}(\beta)$ von β bezüglich der Basis

$$\mathsf{B}' = (b_1, \ldots, b_j, \ldots, b_i + \mu b_j, \ldots, b_n),$$

die man erhält, indem man das μ-fache des j-ten Basisvektors auf den i-ten Basisvektor addieren. Analog geht man bei Zeilentausch-Operationen oder Skalierungen vor. Beispielsweise können wir die i-te Spalte mit $\lambda \neq 0$ multiplizieren, wenn wir anschließend auch die i-te Spalte mit der selben Zahl λ multiplizieren und in der geordneten Basis den i-ten Basisvektor mit λ multiplizieren.

Beispiel 3.2.10. Es sei $\mathbb{K} = \mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$ und $V = \mathbb{K}^3$. Es sei die Bilinearform

$$\beta : \mathbb{K}^3 \to \mathbb{K}^3, \qquad \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) \mapsto x_3 y_1 + x_2 y_2 + x_3 y_2 + x_1 y_3 + x_2 y_3.$$

gegeben. Die Frage ist nun: Gibt es eine Basis B von V, sodass $\mathrm{FM}_{\mathsf{B}}(\beta)$ eine besonders einfache Gestalt hat?

Zuerst einmal bestimmen wir mit Proposition 3.2.7 die Fundamentalmatrix $\mathrm{FM}_{\mathsf{E}}(\beta)$ bezüglich der Standardbasis:

$$\mathrm{FM}_{\mathsf{E}}(\beta) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Wenn wir nun den Gauß-Algorithmus auf diese Matrix anwenden würden, so wäre ein erster Schritt, dass wir die dritte Zeile mit der ersten vertauschen, um ein Pivot-Element in die erste Zeile zu bekommen. Bemerkung 3.2.9 sagt uns nun, dass dies erlaubt ist, wenn wir anschließend die dritte Spalte mit der ersten Spalte vertauschen:

$$\mathrm{FM}_{\mathsf{B}_1}(\beta) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Dies ist nun die Fundamentalmatrix bezüglich der Basis $\mathsf{B}_1 = (e_3, e_2, e_1)$, die aus der Basis $\mathsf{E} = (e_1, e_2, e_3)$ durch vertauschen des ersten und dritten Vektors hervorgegangen ist.

Leider hat uns dies nicht weitergebracht, weil wir immer noch eine 0 oben links in der Ecke haben.

Probieren wir etwas anderes: Wir beginnen wieder mit der Ursprungsmatrix $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, aber diesmal vertauschen wir die dritte Zeile nicht mit der ersten, sondern wir addieren die dritte Zeile auf die erste Zeile. Das ist erlaubt, wenn wir anschließend die dritte Spalte auf die erste Spalte addieren. Tun wir dies, erhalten wir:

$$\mathrm{FM}_{\mathsf{B}_2}(\beta) = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Dies ist die Fundamentalmatrix bezüglich der geordneten Basis

$$\mathsf{B}_2 = (e_1 + e_3, e_2, e_3) = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right),$$

die wir aus der ursprünglichen Basis dadurch erhalten haben, dass wir den dritten Basisvektor auf den ersten addiert haben.

Nun haben wir oben links ein Nicht-Null-Element ($2 \neq 0$ in $\mathbb{F}_3$). Weiter im Gauß-Algorithmus: Wir addieren nun die erste Zeile auf die zweite Zeile. Das gleiche tun wir anschließend mit der ersten Spalte und der zweiten:

$$\mathrm{FM}_{\mathsf{B}_3}(\beta) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{pmatrix}.$$

Nun die erste Zeile auf die dritte – und die erste Spalte auf die dritte:

$$\mathrm{FM}_{\mathsf{B}_4}(\beta) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Wenn wir nun noch das 2-fache der zweiten Zeile auf die dritte addieren – und danach das 2-fache der zweiten Spalte auf die dritte Spalte addieren, sollte unsere Matrix in Stufenform sein:

$$\mathrm{FM}_{\mathsf{B}_5}(\beta) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Wenn wir jeden einzelnen Schritt, den wir auf Spalten und Zeilen angewendet haben, auch parallel auf die Basis anwenden, erhalten wir die Basis

$$\mathsf{B}_5 = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right),$$

bezüglich derer die Bilinearform β diese besonders einfache Fundamentalmatrix hat, in diesem Falle sogar eine Diagonalmatrix. Wir werden uns im Folgenden mit der Frage beschäftigen, für welche Bilinearformen dies möglich ist.

Achtung: An dieser Stelle sei noch einmal darauf hingewiesen, dass diese neue Matrix *nicht* ähnlich zur Originalmatrix ist, d.h. wäre die Matrix die Darstellungsmatrix eines Endomorphismus, so wären diese Operationen nicht erlaubt gewesen! Zur Verdeutlichung: Die Originalmatrix hatte Spur $1 \in \mathbb{F}_3$, während die Diagonalmatrix Spur $2+2+2=0 \in \mathbb{F}_3$ hat. Da die Spur immer die Summe der Eigenwerte (in einem entsprechenden Erweiterungskörper) ist, sieht man daran auch schon, dass sich die Eigenwerte verändert haben müssen.[7] Die Determinante ist in diesem Falle gleichgeblieben. Das liegt daran, dass alle Basiswechselmatrizen Determinante 1 hatten. Sobald man Zeilen und Spalten mit Skalaren multipliziert, kann sich auch die Determinante der Fundamentalmatrix ändern.[8]

[7] Man kann auch direkt nachrechnen, dass die Originalmatrix im Körper $\mathbb{F}_3$ überhaupt keine Eigenwerte hat, während die Diagonalmatrix offensichtlicherweise den Eigenwert 2 besitzt.

[8] Im Körper $\mathbb{F}_3$ ändert sich die Determinante der Fundamentalmatrix bei einem Basiswechsel allerdings nie – das ist ein Spezialität von $\mathbb{F}_3$. Können Sie herausfinden, wieso das so ist?

3.3. Symmetrische Bilinearformen

Definition 3.3.1. Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine Bilinearform

$$\beta : V \times V \to \mathbb{K}$$

heiße *symmetrisch*, falls

$$\forall v, w \in V : \ \beta(v,w) = \beta(w,v).$$

Proposition 3.3.2. *Es sei V ein n-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ mit $n \in \mathbb{N}$. Es sei B eine geordnete Basis von V und $\beta : V \times V \to \mathbb{K}$ eine Bilinearform. Dann sind die folgenden Aussagen äquivalent:*

- *Die Bilinearform β ist symmetrisch.*
- *Die Fundamentalmatrix $\mathrm{FM}_{\mathsf{B}}(\beta)$ ist symmetrisch.*

Wir wollen nun der Frage nachgehen, wie man eine Basis eines endlich dimensionalen Vektorraums bestimmen kann, bezüglich derer die Fundamentalmatrix einer gegebenen Bilinearform Diagonalgestalt hat. Dies führt zu dem folgenden Begriff:

Definition 3.3.3. Es sei V ein n-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ und es sei $\beta : V \times V \to \mathbb{K}$ eine Bilinearform. Eine geordnete Basis $\mathsf{B} = (b_1, \dots, b_n)$ heiße *Orthogonalbasis* für β, falls für alle $i \neq j$ gilt, dass $\beta(b_i, b_j) = 0$.

Bemerkung 3.3.4. Falls $\mathbb{K} = \mathbb{R}$ und β ein Skalarprodukt ist, dann stimmt dieser Begriff der Orthogonalbasis mit dem aus Definition 1.3.4 überein.

In Definition 3.3.3 ist es erlaubt, dass $\beta(b_i, b_i) = 0$ sind. Insbesondere ist für die Nullabbildung $\beta = 0 : V \times V \to \mathbb{K}, \quad (v,w) \mapsto 0$ jede Basis eine Orthogonalbasis.

Lemma 3.3.5. *Es sei V ein n-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ und es sei $\beta : V \times V \to \mathbb{K}$ eine Bilinearform. Für eine geordnete Basis $\mathsf{B} = (b_1, \dots, b_n)$ sind äquivalent:*

(i) B *ist eine Orthogonalbasis für β.*

(ii) $\mathrm{FM}_{\mathsf{B}}(\beta)$ *ist eine Diagonalmatrix.*

Lemma 3.3.6. *Es sei V ein n-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ und es sei $\beta : V \times V \to \mathbb{K}$ eine Bilinearform. Wenn V eine Orthogonalbasis bezüglich β besitzt, dann ist β symmetrisch.*

Beweis. Da eine Diagonalmatrix immer symmetrisch ist, folgt diese Aussage direkt aus Lemma 3.3.5 und Proposition 3.3.2. □

Bemerkung 3.3.7. Es wäre schön, wenn wir zeigen könnten, dass die Bedingung symmetrisch zu sein nicht nur notwendig, sondern auch hinreichend ist. Leider ist dies nicht korrekt: In den Übungen werden Sie sehen, dass die symmetrische Bilinearform $\beta : \mathbb{F}_2^2 \times \mathbb{F}_2^2 \to \mathbb{F}_2$ mit der Fundamentalmatrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

nicht in Diagonalform gebracht werden kann, egal bezüglich welcher Basis.

Es stellt sich heraus, dass dieses Problem aber nur auftauchen kann, weil in $\mathbb{F}_2$ die ungewöhnliche Gleichheit $1+1=0$ gilt. [9] Wenn wir Körper mit dieser Eigenschaft ausschließen, können wir die Umkehrung beweisen.

Definition 3.3.8 (Charakteristik 2)**.** Ein Körper $\mathbb{K}$ habe *Charakteristik 2*, geschrieben $\operatorname{char}(\mathbb{K}) = 2$, falls gilt $1+1=0$.

Für den allgemeineren Begriff der Charakteristik eines Körpers oder Rings schauen Sie in den Anhang des LA1-Skriptes. Für uns im Moment reicht aber diese Definition.

Satz 3.3.9. *Es sei $\mathbb{K}$ ein Körper mit Charakteristik ungleich* 2. *Es sei V ein Vektorraum über $\mathbb{K}$ der Dimension $n \in \mathbb{N}$ und es sei*

$$\beta : V \times V \to \mathbb{K}$$

eine Bilinearform gegeben. Dann sind äquivalent:

(i) β ist symmetrisch.

(ii) Es gibt eine Orthogonalbasis B für β, d.h.

$$\mathrm{FM}_{\mathsf{B}}(\beta) \textit{ hat Diagonalform.}$$

Wir haben in Definition 3.2.1 bereits das Konzept einer quadratischen Form q_β eingeführt, die zu einer Bilinearform β gehört. Nun wollen wir der Frage nachgehen, ob es immer möglich ist, aus einer gegebenen quadratischen Form die zugrundeliegende Bilinearform zurückzubekommen. In Beispiel 3.2.2 (a) haben wir gesehen, dass eine Bilinearform, die nicht konstant 0 ist (nämlich die Determinantenform in 2 Dimensionen) als zugehörige quadratische Form die konstante 0-Funktion zugeordnet bekommt. In diesem Falle geht also beim Übergang von bilinearer Form zu quadratischer Form Information verloren. Wir werden nun zeigen, dass dies nicht passiert, wenn die Bilinearform symmetrisch ist und der Grundkörper Charakteristik ungleich 2 hat:

Satz 3.3.10 (Polarisierungsidentität für Bilinearformen)**.**
Es sei $\mathbb{K}$ ein Körper mit Charakteristik ungleich 2 *und V ein Vektorraum über $\mathbb{K}$. Zu jeder quadratischen Form $q : V \to \mathbb{K}$ gibt es genau eine symmetrische Bilinearform $\beta : V \times V \to \mathbb{K}$ mit $q = q_\beta$:*

$$\beta(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big).$$

Hierbei steht $\frac{1}{2} = 2^{-1} \in \mathbb{K}$ für das multiplikative Inverse von $2 = 1+1$ in $\mathbb{K}$.
Die Abbildung

$$\big\{\beta \in \operatorname{Bil}(V,\mathbb{K}) \,\big|\, \beta \textit{ ist symmetrisch}\big\} \to \{q : V \to \mathbb{K} \mid q \textit{ ist eine quadratische Form}\}, \quad \beta \mapsto q_\beta$$

ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen.

[9] In LA1 haben wir außerdem auch den Körper $\mathbb{F}_4$ kennengelernt (der nicht isomorph zum Ring $\mathbb{Z}/4\mathbb{Z}$ ist). Der Körper $\mathbb{F}_4$ hat ebenfalls Charakteristik 2. Man kann mit noch mehr Algebra zeigen, dass es noch viele weitere Körper der Charakteristik 2 gibt, es gibt sogar unendliche (sogar überabzählbare) Körper mit Charakteristik 2.

Definition 3.3.11. Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine symmetrische Bilinearform $\beta : V \times V \to \mathbb{K}$ heiße *entartet* (oder auch *ausgeartet* oder *degeneriert*), falls es ein $w \in V \setminus \{0\}$ gibt mit

$$\forall v \in V : \beta(v,w) = 0.$$

Andernfalls heißt sie *nicht entartet* (oder dementsprechend *nicht ausgeartet* oder *nicht degeneriert*).

Lemma 3.3.12. *Es sei $\beta : V \times V \to \mathbb{K}$ eine symmetrische Bilinearform auf einem Vektorraum über einem Körper $\mathbb{K}$. Dann sind die folgenden Aussagen äquivalent:*

(i) *β ist nicht entartet.*

(ii) $\forall v \neq 0 : \exists w \neq 0 : \beta(v,w) \neq 0$.

(iii) *Die Abbildung $\beta^\vee : V \to V^*$ ist injektiv.*

Falls V endlich dimensional ist, so sind die genannten Aussagen außerdem äquivalent zu

(iv) *Die Abbildung $\beta^\vee : V \to V^*$ ist bijektiv.*

(v) *Es gibt eine geordnete Basis B von V mit* $\det \mathrm{FM}_{\mathsf{B}}(\beta) \neq 0$

(vi) *Für jede geordnete Basis B von V gilt:* $\det \mathrm{FM}_{\mathsf{B}}(\beta) \neq 0$

Satz 3.3.13. *Es sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper mit* $\mathrm{char}(\mathbb{K}) \neq 2$ *(beispielsweise der Körper $\mathbb{C}$). Es sei weiterhin $\beta : V \times V \to \mathbb{K}$ eine nicht entartete symmetrische Bilinearform auf einem n-dimensionalen Vektorraum V. Dann gibt es immer eine geordnete Basis B von V mit*

$$\mathrm{FM}_{\mathsf{B}}(\beta) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = \mathbb{1}_n.$$

Es sind also alle nicht entarteten symmetrischen Bilinearformen durch Basiswechsel ineinander überführbar.

Beweis. Die Bilinearform β ist symmetrisch und definiert auf einem endlich-dimensionalen Vektorraum über einem Körper der Charakteristik ungleich 2. Somit können wir Satz 3.3.9 anwenden und erhalten eine Orthogonalbasis $\mathsf{B}_0 = (b_1, \ldots, b_n)$ für β. Bezüglich dieser Basis gilt:

$$\mathrm{FM}_{\mathsf{B}_0}(\beta) = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}.$$

Da die Bilinearform nicht entartet ist, muss nach Lemma 3.3.12, diese Matrix eine Determinante ungleich 0 haben. Dies bedeutet, dass alle $\alpha_1, \ldots, \alpha_n$ ungleich 0 sind.

Der Körper $\mathbb{K}$ ist algebraisch abgeschlossen, insbesondere hat jedes α_j eine Quadratwurzel, d.h. es gibt $\beta_1, \ldots, \beta_n$ mit $\alpha_j^2 = \beta_j$. Da die Skalare α_j alle ungleich 0 sind, gilt dies auch für die Skalare β_j.

Wenn wir nun die Basis $\mathsf{B}_0 = (b_1, \ldots, b_n)$ durch

$$\mathsf{B} := \left(\frac{1}{\beta_1} b_1, \ldots, \frac{1}{\beta_n} b_n \right)$$

ersetzen (das entspricht dem Multiplizieren jeder Zeile und Spalte mit $1/b_j$ (symmetrischer Gauß)), dann erhalten wir

$$\mathrm{FM}_{\mathsf{B}}(\beta) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

und das war zu zeigen. □

Über dem Körper $\mathbb{R}$, der nicht algebraisch abgeschlossen ist, gilt dies so nicht, beispielsweise ist die nicht entartete Bilinearform des Minkowski-Raums (siehe Beispiel 3.2.2(d)) nicht in das reelle Standardskalarprodukt überführbar, weil sie nicht positiv definit ist (siehe Definition 3.4.1).

Analog zu Satz 3.3.13 zeigt man die folgende allgemeinere Version:

Satz 3.3.14 (Klassifikation aller Bilinearformen auf einem n-dimensionalen komplexen Vektorraum)**.**
Es sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper mit $\operatorname{char}(\mathbb{K}) \neq 2$ *(beispielsweise der Körper $\mathbb{K} = \mathbb{C}$) und V ein n-dimensionaler $\mathbb{K}$-Vektorraum ($n \in \mathbb{N}$). Dann gibt es für jede symmetrische Bilinearform $\beta : V \times V \to \mathbb{K}$ eine geordnete Basis B von V und eine natürliche Zahl $k \in \{0, \ldots, n\}$ mit*

$$\mathrm{FM}_{\mathsf{B}}(\beta) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

wobei k die Anzahl der Einsen ist. Die Zahl k ist der Rang von $\beta^\vee : V \to V^*$ und somit eindeutig durch β bestimmt, die Basis B ist nicht eindeutig.

Da nach Satz 3.3.10 quadratische Formen und symmetrische Bilinearformen in Bijektion stehen (vorausgesetzt, die Charakteristik ist nicht gleich 2), haben wir mit Satz 3.3.14 auch alle quadratischen Formen auf einem endlich dimensionalen komplexen Vektorraum klassifiziert.

3.4. Bilinearformen und quadratische Formen über den reellen Zahlen

Wir haben in Satz 3.3.14 gesehen, dass jede symmetrische Bilinearform (und damit jede quadratische Form) über einem endlich dimensionalen *komplexen* Vektorraum durch eine einzige Zahl, nämlich den Rang klassifiziert ist. Da hier aber explizit eingegangen ist, dass der Körper $\mathbb{C}$ algebraisch abgeschlossen ist[10], können wir dieses Resultat nicht verwenden, um symmetrische Bilinearformen über $\mathbb{R}$ zu klassifizieren. Beispielsweise können Sie die Bilinearform auf $\mathbb{R}^4$ aus Beispiel 3.2.2 (d) nicht durch einen Basiswechsel in das Standardskalarprodukt (Beispiel 3.2.2 (b)) umwandeln, obwohl beide Bilinearformen symmetrisch sind und Rang 4 haben – das sieht

[10] Genaugenommen reicht es zu fordern, dass jedes quadratische Polynom eine Nullstelle hat, das ist eine echt schwächere Forderung an den Grundkörper als zu fordern, dass jedes nicht konstante Polynom eine Nullstelle besitzt.

man direkt an den entsprechenden Fundamentalmatrizen. Der Grund, warum die beiden echt verschieden sind, ist dass bei Beispiel 3.2.2 (d) Vektoren v existieren, sodass die dazugehörige quadratische Form q_β angewandt auf v negativ ist. Somit ist dieses β kein Skalarprodukt. Dies führt uns wieder auf den Begriff der *Definitheit*, den wir schon in Definition 1.2.2 kennengelernt haben und den wir hier ein wenig ausweiten wollen:

Definition 3.4.1 (Definitheit einer reellen Bilinearform)**.**
Es sei V ein reeller Vektorraum und $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform mit dazugehöriger quadratischer Form $q_\beta : V \to \mathbb{R}$. Dann nennen wir β (und auch q_β)

- *positiv definit*, falls $\forall v \in V \setminus \{0\} : q_\beta(v) > 0$
- *negativ definit*, falls $\forall v \in V \setminus \{0\} : q_\beta(v) < 0$
- *positiv semidefinit*, falls $\forall v \in V : q_\beta(v) \geq 0$
- *negativ semidefinit*, falls $\forall v \in V : q_\beta(v) \leq 0$
- *indefinit*, falls es Vektoren $v_1, v_2 \in V$ gibt mit $q_\beta(v_1) < 0$ und $q_\beta(v_2) > 0$.

Bemerkung 3.4.2. (i) Es sei V ein reeller Vektorraum. Durch Vergleichen Definition 1.2.2 und Definition 3.4.1 sehen wir also: Eine Funktion $\beta : V \times V \to \mathbb{R}$ ist genau dann ein Skalarprodukt auf V, wenn β eine symmetrische positiv definite Bilinearform ist. Beachten Sie aber, dass komplexe Skalarprodukte auf diese Weise nicht charakterisiert werden können, weil sie keine Bilinearformen sind.[11]

(ii) Es ist leicht zu sehen, dass aus positiver bzw. negativer Definitheit immer die dazugehörige Semidefinitheit folgt und dass die Nullabbildung die einzige symmetrische Bilinearform ist, die sowohl positiv semidefinit als auch negativ semidefinit ist.

(iii) Man sieht an der Definition direkt, dass Definitheit eigentlich keine Eigenschaft der symmetrischen Bilinearform, sondern der quadratischen Form ist, aber da nach Satz 3.3.10 zwischen den beiden problemlos hin- und herwechseln kann (Es gilt ja $\mathrm{char}(\mathbb{R}) \neq 2$), ist dies kein Problem.

(iv) Die Definitheit einer quadratischen Form ist wichtig für die Optimierung von reellwertigen Funktionen, also für das Finden von lokalen und globalen Minima und Maxima. Es gilt: Eine quadratische Form q_β hat genau dann in 0 ein Minimum, wenn q_β positiv semidefinit ist und genau dann ein Maximum, wenn q_β negativ semidefinit ist.

Funktionen, die keine quadratischen Formen sind, aber genügend oft stetig differenzierbar sind, kann man durch ihr quadratisches Taylor-Polynom annähern, sodass sie sich lokal annähern lassen durch die Summe eines konstanten Terms, einer Linearform und einer quadratischen Form. Die Abbildungsmatrix der Linearform entspricht dann dem Gradienten und die Fundamentalmatrix der zur quadratischen Form gehörigen Bilinearform nennt man die Hesse-Matrix.

Im Endlichdimensionalen liegt es nahe, die Definitheit einer symmetrsichen Bilinearform mit Hilfe ihrer Fundamentalmatrix zu bestimmen:

[11] Man kann die Begriffe aus Definition 3.4.1 allerdings analog für komplexe *Sesquilinearformen* einführen, was wir hier aber nicht tun werden. Dann kann man sagen: Ein Skalarprodukt auf einem komplexen Vektorraum ist eine positiv definite Sesquilinearform.

Definition 3.4.3. Es sei $n \in \mathbb{N}$. Eine symmetrische[12] Matrix $A \in \mathbb{R}^{n\times n}$ heiße

- *positiv definit*, falls $\forall x \in \mathbb{R}^n \setminus \{0\} : x^\top Ax > 0$
- *negativ definit*, falls $\forall x \in \mathbb{R}^n \setminus \{0\} : x^\top Ax < 0$
- *positiv semidefinit*, falls $\forall x \in \mathbb{R}^n : x^\top Ax \geq 0$
- *negativ semidefinit*, falls $\forall x \in \mathbb{R}^n : x^\top Ax \leq 0$
- *indefinit*, falls es Vektoren $x, y \in \mathbb{R}^n$ gibt mit $x^\top Ax < 0$ und $y^\top Ay > 0$.

Es gilt jetzt direkt:

Lemma 3.4.4. *Es sei V ein n-dimensionaler reeller Vektorraum mit geordneter Basis B und $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gilt: Die Bilinearform β ist genau dann positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit, wenn die Matrix $\mathrm{FM}_{\mathsf{B}}(\beta)$ positiv definit / positiv semidefinit / negativ definit / negativ semidefinit / indefinit ist.*

Das reduziert jetzt das Problem, die Definitheit einer quadratischen Form zu bestimmen auf das entsprechende Problem bei der Fundamentalmatrix. Aber wie macht man das?

Wir werden im Folgenden vier unterschiedliche Kriterien vorstellen, um die Definitheit einer symmetrischen Bilinearform bzw. Matrix zu bestimmen: Satz 3.4.5, Satz 3.4.8, Satz 3.4.10 und Satz 3.4.12.

Satz 3.4.5 (Erstes Kriterium für Definitheit)**.**
Es sei V ein n-dimensionaler reeller Vektorraum und $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Es sei weiterhin B eine Orthogonalbasis[13] für β. Dann gilt:

- positiv definit*, falls alle Diagonaleinträge von* $\mathrm{FM}_{\mathsf{B}}(\beta)$ *größer als* 0 *sind.*
- negativ definit*, falls alle Diagonaleinträge von* $\mathrm{FM}_{\mathsf{B}}(\beta)$ *kleiner als* 0 *sind.*
- positiv semidefinit*, falls alle Diagonaleinträge von* $\mathrm{FM}_{\mathsf{B}}(\beta)$ *größer als oder gleich* 0 *sind.*
- negativ semidefinit*, falls alle Diagonaleinträge von* $\mathrm{FM}_{\mathsf{B}}(\beta)$ *kleiner als oder gleich* 0 *sind.*
- indefinit*, falls es sowohl positive als auch negative Zahlen auf der Diagonalen gibt.*

Zusammen mit Bemerkung 3.2.9 gibt uns dies einen Algorithmus, um die Definitheit zu bestimmen. Man bringt die Matrix mit Gauß-Umformungen (die man immer auf Zeilen *und* Spalten anwenden muss!) in Diagonalform und schaut sich dann die Vorzeichen der Diagonaleinträge an.

Lemma 3.4.6.
Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische $(n \times n)$-Matrix mit $n \geq 2$ und $B \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix, die man erhält, wenn man die letzte Zeile und die letzte Spalte streicht. Dann sind die folgenden Aussagen äquivalent:

[12] Es gibt Autoren, die (positive) Definitheit auch für nichtsymmetrische reelle Matrizen definieren. Das werden wir hier nicht tun.

[13] Siehe Definition 3.3.3; so eine existiert immer nach 3.3.9

(i) Die Matrix A ist positiv definit.

(ii) Die Matrix B ist positiv definit und $\det(A) > 0$.

Insbesondere ist die Determinante einer positiv definiten Matrix immer positiv.

Bemerkung 3.4.7. Achtung: Die Determinante einer negativ definiten $(n \times n)$- Matrix ist *nicht* immer negativ. Es gilt: Falls A negativ definit ist, ist $-A$ positiv definit. Also ist:

$$\det(A) = \det(-(-A)) = (-1)^n \underbrace{\det(-A)}_{>0}.$$

Falls also n gerade ist, dann ist $\det(A) > 0$, aber für ungerade n gilt: $\det(A) < 0$.

Beweis von Lemma 3.4.6. „(i) $\Longrightarrow$ (ii)“
Wir zeigen, dass B positiv definit ist. Sei dazu $x \in \mathbb{R}^{n-1} \setminus \{0\}$. Wir ergänzen diesen Vektor zu einem Vektor $y \in \mathbb{R}^n$, indem wir in der letzten Komponente eine 0 hinzufügen:

$$y = \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{R}^n .$$

Da x nicht der Nullvektor ist, ist auch $y \neq 0$. Aus der positiven Definitheit von A folgt nun, dass

$$y^\top A y > 0 ,$$

aber Ausmultiplizieren ergibt direkt

$$y^\top A y = \begin{pmatrix} x^\top & 0 \end{pmatrix} \left(\begin{array}{c|c} B & * \\ \hline * & * \end{array}\right) \begin{pmatrix} x \\ 0 \end{pmatrix} = x^\top B x .$$

Also ist $x^\top B x > 0$ und B ist positiv definit.

Es bleibt zu zeigen, dass $\det(A) > 0$ ist. Nach Proposition 3.2.7 gibt es eine Orthogonalbasis für die zu A gehörige Bilinearform, d.h. zusammen mit Proposition 3.2.8 erhalten wir eine Basiswechselmatrix $S \in \mathrm{GL}(n, \mathbb{R})$ mit

$$A = S^\top D S,$$

wobei D eine Diagonalmatrix ist. Nach Satz 3.4.5 sind nun alle Diagonaleinträge in D positiv und somit gilt:

$$\det(A) = \det\left(S^\top D S\right) = \underbrace{\det(S^\top)}_{=\det(S)} \det(D) \det(S) = \left(\det(S)\right)^2 \det(D).$$

Da S invertierbar ist, gilt $\det(S) \neq 0$ und somit $(\det(S))^2 > 0$. Da D eine Diagonalmatrix mit positiven Diagonaleinträgen ist, gilt $\det(D)$ größer als 0. Also ist auch das Produkt und damit $\det(A) > 0$.

Achtung: Es ist im Allgemeinen *nicht* so, dass $\det(A) = \det(D)$ ist. Alles, was wir sagen können, ist, dass beide Determinanten das gleiche Vorzeichen haben.

„(ii) $\Longrightarrow$ (i)"

$$A = \left(\begin{array}{c|c} B & v \\ \hline v^\top & * \end{array}\right).$$

Wir nehmen nun an, dass B positiv definit ist und $\det(A) > 0$. Wir wollen zeigen, dass A positiv definit ist.

Wir wissen aus der Implikation, die wir bereits gezeigt haben, dass jede positiv definite Matrix eine positive Determinante hat. Insbesondere gilt also $\det(B) > 0$ und B ist somit invertierbar. Das LGS

$$Bx = -v$$

hat somit eine Lösung $x_0 \in \mathbb{R}^{n-1}$, d.h. es gibt ein $x_0 \in \mathbb{R}^{n-1}$ mit $Bx_0 + v = 0$.

Betrachten wir nun die Matrix

$$S := \left(\begin{array}{c|c} \mathbb{1}_{n-1} & x_0 \\ \hline 0 & 1 \end{array}\right).$$

Man sieht direkt, dass $\det(S) = 1$, also ist S invertierbar. Wenn wir nun die Matrix A von rechts mit S und von links mit $S^\top$ multiplizieren, so erhalten wir die Fundamentalmatrix der gleichen Bilinearform, nur zu einer anderen Basis:

$$\begin{aligned} \widetilde{A} &:= S^\top A S \\ &= \left(\begin{array}{c|c} \mathbb{1}_{n-1} & 0 \\ \hline x_0^\top & 1 \end{array}\right) \left(\begin{array}{c|c} B & v \\ \hline v^\top & * \end{array}\right) \left(\begin{array}{c|c} \mathbb{1}_{n-1} & x_0 \\ \hline 0 & 1 \end{array}\right) \\ &= \left(\begin{array}{c|c} B & Bx_0 + v \\ \hline (Bx_0 + v)^\top & * \end{array}\right) = \left(\begin{array}{c|c} B & 0 \\ \hline 0 & c \end{array}\right) \end{aligned}$$

mit einer Zahl $c \in \mathbb{R}$. Wir werden nun zeigen, dass diese Zahl positiv ist. Wichtig ist: Die Matrix B oben links hat sich bei dieser Aktion nicht verändert.

Da die Matrix A nach Voraussetzung eine positive Determinante hat, gilt nun:

$$\begin{aligned}
0 &< \det A \\
&= \underbrace{\det(S^\top)}_{=1} \det A \underbrace{\det(S)}_{=1} \\
&= \det(S^\top A S) \\
&= \det(\widetilde{A}) \\
&= \det \left(\begin{array}{c|c} B & 0 \\ \hline 0^\top & c \end{array}\right) \\
&= \underbrace{\det(B)}_{>0} \cdot c \,.
\end{aligned}$$

Also muss die Zahl c größer als 0 sein.

Nun zeigen wir, dass A positiv definit ist. Da positive Definitheit eine Eigenschaft der Bilinearform ist, die nicht von der gewählten Basis abhängt, reicht es, zu zeigen, dass die Matrix

$$\left(\begin{array}{c|c} B & 0 \\ \hline 0 & c \end{array}\right)$$

positiv definit ist. Sei dazu $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^n$ ein beliebiger Vektor ungleich 0, mit $x \in \mathbb{R}^{n-1}$ und $y \in \mathbb{R}$. Dann gilt:

$$\begin{pmatrix} x^\top & y \end{pmatrix} \left(\begin{array}{c|c} B & 0 \\ \hline 0 & c \end{array}\right) \begin{pmatrix} x \\ y \end{pmatrix} = x^\top B x + c y^2 \,.$$

Der erste Summand ist immer nichtnegativ, weil B positiv definit ist, der zweite Summand ist nichtnegativ, weil $c > 0$ ist. Das zeigt schon einmal, dass A positiv semidefinit ist. Da aber der Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \neq 0$ ist, muss entweder $x \neq 0$ sein (dann wird der erste Summand positiv) oder die Zahl y ist ungleich 0, dann wird der zweite Summand positiv.

Also ist A positiv definit. □

Eine positive Determinante ist eine notwendige Bedingung für positive Definitheit, aber leider nicht hinreichend. Es gibt aber nichtsdestotrotz eine Möglichkeit, nur durch Berechnung von Determinanten eine Matrix auf positive Definitheit zu testen:

Satz 3.4.8 (Hurwitz-Kriterium).
Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix und für jedes $k \in \{1,\dots,n\}$ sei $F_k \in \mathbb{R}^{k\times k}$ die Matrix, die aus den obersten k Zeilen und den ersten k Spalten besteht.

(a) Die Matrix A ist genau dann positiv definit, wenn für alle $k \in \{1,\dots,n\}$ gilt:

$$\det(F_k) > 0.$$

(b) Die Matrix A ist genau dann negativ definit, wenn für alle $k \in \{1,\dots,n\}$ gilt:

$$(-1)^k \cdot \det(F_k) > 0.$$

(c) Falls $\det(A) \neq 0$ *ist und A weder positiv noch negativ definit ist, dann ist A indefinit.*

Beweis. Wir zeigen nur Teil (a). Aussage (b) folgt dann direkt, wenn man Teil A auf die Matrix $-A$ anwendet und $\det(-A) = (-1)^n \det(A)$ verwendet (siehe auch Bemerkung 3.4.7). Aussage (c) folgt dann aus (a) und (b).

Zeigen wir nun Teil (a) per vollständiger Induktion nach $n \in \mathbb{N}$:

Für $n = 1$ ist $F_1 = A$ eine reelle Zahl und eine reelle Zahl, aufgefasst als (1×1)-Matrix ist genau dann positiv definit, wenn sie eine positive Zahl ist.

Nehmen wir nun an, $n \in \mathbb{N}$ sei so gewählt, dass die Aussage stimmt. Zeigen wir die Aussage für $n+1$. Sei also $A \in \mathbb{R}^{(n+1)\times(n+1)}$ eine Matrix. Dann gilt nach Lemma 3.4.6, dass A genau dann positiv definit ist, wenn

$$\det(A) > 0 \quad \text{und} \quad F_n \text{ positiv definit ist.}$$

Die Matrix F_{n+1} ist genau die Matrix A und die Matrix F_n ist nun – nach Induktionsvoraussetzung – genau dann positiv definit, wenn die Zahlen $\det(F_1), \det(F_2), \dots, \det(F_n)$ alle positiv sind.

Damit haben wir also, dass die $(n+1)\times(n+1)$-Matrix A genau dann positiv definit ist, wenn

$$\det(F_{n+1}) > 0 \quad \text{und} \quad \det(F_1), \det(F_2), \dots, \det(F_n) > 0\,.$$

Das war zu zeigen. □

Bemerkung 3.4.9. (i) Die Determinanten $\det(F_k)$ in Satz 3.4.8 werden auch *Hauptabschnittsunterdeterminanten* oder *Hauptminoren* genannt.

(ii) Da die Aussagen (b) und (c) direkt aus Teil (a) folgen, wird oft nur die Aussage von Teil (a) Hurwitz-Kriterium genannt.

(iii) Falls $\det(A) = 0$ ist, folgt mit Teil (a), dass A nicht positiv definit ist und mit Teil (b), dass A nicht negativ definit ist. Ob A aber indefinit oder semidefinit ist, lässt sich mit dem Hurwitz-Kriterium in diesem Fall nicht entscheiden.

Kommen wir nun zu unserem dritten Kriterium, um zu entscheiden, ob eine gegebene Matrix positiv definit ist. Hier stellen wir eine Verbindung her zu dem, was wir in Kapitel 2.4 gemacht haben.

Satz 3.4.10 (Definitheit und Eigenwerte).
Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gilt:

- positiv definit*, falls alle Eigenwerte von A größer als* 0 *sind.*
- negativ definit*, falls alle Eigenwerte von A kleiner als* 0 *sind.*
- positiv semidefinit*, falls alle Eigenwerte von A größer als oder gleich* 0 *sind.*
- negativ semidefinit*, falls alle Eigenwerte von A kleiner als oder gleich* 0 *sind.*
- indefinit*, falls A sowohl positive als auch negative Eigenwerte besitzt.*

Beweis. Da die Matrix A reell und symmetrisch ist, lässt sie sich nach dem reellen Spektralsatz (Satz 2.4.7) orthogonal diagonalisieren, d.h. es gibt eine Diagonalmatrix D und eine orthogonale Basiswechselmatrix $U \in \mathrm{O}(n)$ mit

$$A = U^{-1} D U$$

Man beachte, dass U orthogonal ist und somit $U^{-1} = U^\top$ gilt. Somit ist

$$A = U^\top D U$$

auch ein Basiswechsel als Fundamentalmatrizen. Die Aussage folgt jetzt direkt aus Satz 3.4.5. □

Bemerkung 3.4.11. (i) Da A reell und symmetrisch ist, sind alle komplexen Eigenwerte von A reell (Lemma 2.4.3(d)) und es ergibt Sinn, von positiven und negativen Eigenwerten zu sprechen.

(ii) Satz 3.4.10 ist beeindruckend, weil positive Definitheit eine Eigenschaft einer Bilinearform ist und die Eigenwerte der Fundamentalmatrix überhaupt nicht invariant sind unter Basistransformationen. Es stellt sich aber heraus, dass die Eigenschaft nur positive oder nur negative Eigenwerte zu besitzen, doch invariant ist.

(iii) Die abstrakte Begründung, warum es plötzlich möglich ist, Eigenwerte (die ja zu Endomorphismen gehören) mit positiver Definitheit (was ja eine Eigenschaft einer Bilinearform, bzw. quadratischer Form ist) in Verbindung zu bringen ist die folgende:

Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich dimensionaler Euklidischer Vektorraum, d.h. ein reeller Vektorraum mit einem festen Skalarprodukt $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$. Dieses Skalarprodukt ist positiv definit, also insbesondere nicht entartet. Also ist die Abbildung

$$\Phi := (\langle\cdot,\cdot\rangle)^\vee : V \to V^*, \quad v \mapsto \langle\cdot, v\rangle$$

injektiv. Weil V und V^* beide endlich dimensional sind und die selbe Dimension haben, ist diese Abbildung dann sogar bijektiv. Es gibt also – wenn ein Skalarprodukt gegeben ist – einen natürlichen Isomorphismus[14] von einem Vektorraum V und seinem Dualraum V^*.

Wenn nun eine andere Bilinearform $\beta : V \times V \to \mathbb{R}$ auf dem Vektorraum gegeben ist, so lässt sich diese zu einer linearen Abbildung $\beta^\vee : V \to V^*$ umwandeln und dann mit der Abbildung $\Phi^{-1} : V^* \to V$ verketten und wir erhalten einen Endomorphismus

$$\Phi^{-1} \circ \beta^\vee \in \mathrm{End}(V).$$

[14] Dies ist im direkten Gegensatz zu Bemerkung 3.1.12, die gesagt hat, dass dies nicht möglich ist, wenn man Vektorräume betrachtet, die kein Skalarprodukt mitbringen. Dies zeigt wieder einmal, dass die Theorie von Euklidischen Vektorräumen schöner ist als die Theorie von Vektorräumen ohne zusätzliche Struktur.

Umgekehrt kann jeder Endomorphismus $\varphi \in \mathrm{End}(V)$ mit Φ verkettet werden und wir erhalten eine Abbildung von V nach V^*:

$$\Phi \circ \varphi : V \to V^*,$$

die man als Bilinearform auf V interpretieren kann.

Kurz gesagt gilt also: Auf einem Euklidischen Vektorraum – mit einem festen Skalarprodukt – kann man eine 1-zu-1-Beziehung zwischen Bilinearformen und Endomorphismen herstellen. Symmetrische Bilinearformen werden dabei auf selbstadjungierte Endomorphismen abgebildet und nach Wahl einer Orthonormalbasis für das Skalarprodukt kann man Fundamentalmatrizen der Bilinearformen mit Darstellungsmatrizen der dazugehörigen Endomorphismen gleichsetzen.

Wichtig für all das ist aber, dass es ein fest gewähltes Skalarprodukt auf dem Raum gibt.

(iv) Man beachte, dass es für große Matrizen aufwendig sein kann, die Eigenwerte zu berechnen (immerhin müsste man ja die Nullstellen eines Polynoms mit einem sehr hohen Grad bestimmen). Ob eine Matrix positiv definit ist, kann aber viel einfacher bestimmt werden (entweder mit symmetrischem Gauß (Satz 3.4.5) oder dem Hurwitz-Kriterium (Satz 3.4.8)). Somit lässt sich die Äquivalenz aus Satz 3.4.10 auch in die andere Richtung verwenden. Wir rechnen mit einem anderen Kriterium nach, dass die Matrix positiv definit ist und wissen dann, dass alle Eigenwerte positiv sind. Dann wissen wir zwar immer noch nicht, wie die Eigenwerte aussehen, aber wir kennen ihre Vorzeichen, was je nach Anwendung auch schon sehr interessant sein kann.

Satz 3.4.12. *Es sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix.*

(a) Es sind äquivalent:

(i) A ist symmetrisch und positiv definit.

(ii) Es gibt ein $S \in \mathrm{GL}(n,\mathbb{R})$ mit $S^\top S = A$.

(iii) Es gibt eine symmetrische positiv definite Matrix $S \in \mathbb{R}^{n\times n}$ mit $S^2 = A$.

(b) Es sind äquivalent:

(i) A ist symmetrisch und positiv semidefinit.

(ii) Es gibt ein $m \in \mathbb{N}$ und ein $S \in \mathbb{R}^{m\times n}$ mit $S^\top S = A$.

(iii) Es gibt eine symmetrische positiv semidefinite Matrix $S \in \mathbb{R}^{n\times n}$ mit $S^2 = A$.

Beweis. Wir zeigen nur Teil (a). Teil (b) geht ganz analog, die Argumente müssen nur an wenigen Stellen leicht abgeändert werden.
(i) $\Longrightarrow$ (iii):
Da A reell und symmetrisch ist, gilt nach dem reellen Spektralsatz (Satz 2.4.7): Wir können A orthogonal diagonalisieren, d.h. es gibt eine Matrix $U \in \mathrm{O}(n)\mathbb{R}$ mit

$$A = U^\top \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U,$$

wobei die Diagonaleinträge $\lambda_1,\dots,\lambda_n$ die Eigenwerte von A sind. Da A positiv definit ist, sind alle Diagonaleinträge von A echt größer als 0 (Satz 3.4.5 oder Satz 3.4.10) und wir können aus jedem dieser Einträge die Quadratwurzel ziehen:

$$\mu_j := \sqrt{\lambda_j} > 0\,.$$

Dann setzen wir:

$$S := U^\top \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} U\,.$$

Offensichtlich ist S symmetrisch und positiv definit.

Außerdem ist $S^2 = A$.

(iii) $\Longrightarrow$ (ii):

Da die Matrix aus Teil (iii) symmetrisch ist, gilt $S^\top = S$ und somit

$$A = S^2 = S^\top S\,.$$

Außerdem folgt aus positiver Definitheit von S, dass S invertierbar ist, also ist $S \in \mathrm{GL}(n,\mathbb{R})$.

(ii) $\Longrightarrow$ (i):

Wenn $A = S^\top S$ ist, dann gilt:

$$A^\top = (S^\top S)^\top = S^\top (S^\top)^\top = S^\top S = A$$

und somit ist A symmetrisch. Es bleibt zu zeigen, dass A positiv definit ist (Definition 3.4.3). Sei dazu $x \in \mathbb{R}^n \setminus \{0\}$. Dann gilt:

$$x^\top A x = x^\top S^\top S x = (Sx)^\top S x = \langle Sx, Sx\rangle = \|Sx\|^2\,.$$

Da die Norm eines Vektors immer nichtnegativ ist, folgt somit $x^\top A x \geq 0$. Da die Matrix S zusätzlich als invertierbar angenommen wurde und $x \neq 0$ ist, folgt damit, dass $Sx \neq 0$ und somit $\|Sx\| > 0$ ist.

Also ist A positiv definit. $\square$

Nun wollen wir – analog zu Satz 3.3.14 – alle symmetrischen Bilinearformen (und damit auch alle quadratischen Formen) auf einem endlich dimensionalen reellen Vektorraum klassifizieren. Da der Körper $\mathbb{R}$ leider nicht algebraisch abgeschlossen ist, ist eine einzige Zahl (nämlich der Rang) nicht mehr ausreichend, um alle zu beschreiben:

Satz 3.4.13 (Trägheitssatz von Sylvester)**.** *Es sei $n \in \mathbb{N}$ und V ein n-dimensionaler reeller Vektorraum. Es sei $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform.*

(a) Dann gibt es eine Zerlegung von V als eine direkte Summe

$$V = V_+ \oplus V_- \oplus V_0,$$

sodass $\beta|_{V_+\times V_+}$ positiv definit, $\beta|_{V_-\times V_-}$ negativ definit und $\beta|_{V_0\times V_0}$ die Nullabbildung ist. Weiterhin gilt, dass für alle Vektoren u, die aus verschiedenen *der drei direkten Summanden stammen, orthogonal bezüglich β sind, d.h.*

$$\beta(u,v) = 0.$$

Das Tripel

$$\left(\dim(V_+)\ ,\ \dim(V_-)\ ,\ \dim(V_0)\right) \in \mathbb{N}_0^3$$

hängt nicht von der Zerlegung ab und ist eindeutig bestimmt.

Die Zahl $\dim(V_-)$ *wird die* Signatur[15] *von* β *genannt. Wenn die Signatur von* β *und der Rang* $\dim(V_+)+\dim(V_-)$ *bekannt sind, ist die Bilinearform bis auf Basiswechsel eindeutig bestimmt.*

(b) Es gibt eine geordnete Basis B *von* V*, sodass*

$$\mathrm{FM}_{\mathsf{B}}\left(\beta\right)$$

eine Diagonalmatrix ist, die nur die Zahlen $1,-1$ *und* 0 *auf der Diagonale hat. Die Signatur entspricht der Anzahl der* -1 *auf der Diagonale. Der Rang ist die Summe der Anzahl aller* 1 *und aller* -1.

Beweis. Der Anfang des Beweises funktioniert genauso wie im komplexen Fall (Satz 3.3.13 und Satz 3.3.14):

Nach Satz 3.3.9 gibt es eine Basis B_0, sodass $\mathrm{FM}_{\mathsf{B}_0}\left(\beta\right)$ Diagonalform hat. Durch Permutieren der Vektoren können wir erreichen, dass

$$\mathrm{FM}_{\mathsf{B}_0}\left(\beta\right) = \begin{pmatrix} \alpha_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \alpha_{k_+} & & & & & & \\ & & & \beta_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \beta_{k_-} & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix},$$

wobei $\alpha_1,\ldots\alpha_{k_+} > 0$ und $\beta_1,\ldots,\beta_{k_-} < 0$ sind.

Jede positive reelle Zahl α_j ist von der Form $(\alpha'_j)^2$, jede negative reelle Zahl ist von der Form $-(\beta'_j)^2$. Wir können also durch Skalieren der einzelnen Basiselemente (wie im Beweis von Satz 3.3.13) eine Basis $\mathsf{B} = \left(v_1,\ldots,v_{k_+},w_1,\ldots,w_{k_-},u_1,\ldots,u_{n-k_+-k_-}\right)$ erhalten mit

$$\mathrm{FM}_{\mathsf{B}_0}\left(\beta\right) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix},$$

[15] Die Bezeichnung ist hier nicht eindeutig; manchmal werden auch alle drei Zahlen als *Signatur* bezeichnet.

wobei genau k_+ viele Diagonaleinträge den Wert 1, genau k_- viele Diagonaleinträge den Wert -1 und $n-k_+-k_-$ viele Diagonaleinträge den Wert 0 haben. Hierbei, sind k_+, k_- nicht-negative ganze Zahlen.

Wir setzen nun

$$\begin{aligned} V_+ &:= \mathrm{LH}\left(v_1,\ldots,v_{k_+}\right), \\ V_- &:= \mathrm{LH}\left(w_1,\ldots,w_{k_-}\right), \\ V_0 &:= \mathrm{LH}\left(u_1,\ldots,u_{n-k_+-k_-}\right). \end{aligned}$$

Die Bilinearform ist positiv definit auf V_+, negativ definit auf V_- und konstant 0 auf V_0.

Die Summe $k = k_+ + k_- = \dim(V_+) + \dim(V_-)$ ist die Anzahl aller Einträge auf der Diagonalen, die nicht Null sind und somit ist k der Rang der Fundamentalmatrix.

Die Fundamentalmatrix ist per Definition 3.2.6 die Abbildungsmatrix der Abbildung $\beta^\vee : V \to V^*$ und somit ist

$$k = \mathrm{rg}(\beta^\vee)$$

direkt durch die Abbildung β gegeben und nicht abhängig von irgendeiner Wahl. Der Kern von $\beta^\vee$ ist genau der Raum V_0.

Wenn wir nun noch zeigen können, dass k_- unabhängig von der Wahl der Zerlegung ist, ist der Satz bewiesen.

Nehmen wir dazu an, wir haben Zerlegungen:

$$\begin{aligned} V &= V_+ \oplus V_- \oplus V_0 \\ V &= U_+ \oplus U_- \oplus U_0, \end{aligned}$$

die beide die Bedingung aus (a) erfüllen, d.h.

- β ist positiv definit auf V_+ und U_+,
- β ist negativ definit auf V_- und U_-,
- β ist 0 auf V_0 und U_0.

Es bleibt zu zeigen, dass $\dim(V_-) = \dim(U_-)$. Dann folgt der Rest.

Wir zeigen zuerst, dass $(U_+ + U_0) \cap V_- = \{0\}$.
Sei dazu $v \in (U_+ + U_0) \cap V_-$. Wir wollen zeigen, dass $v = 0$ ist. Da v in $U_+ + U_0$ liegt, gibt es $u_+ \in U_+$ und $u_0 \in U_0$ mit $v = u_+ + u_0$. Dann gilt:

$$\begin{aligned} q_\beta(v) &= \beta(v,v) \\ &= \beta(u_+ + u_0, u_+ + u_0) \\ &= \underbrace{\beta(u_+,u_+)}_{\geq 0} + 2\underbrace{\beta(u_+,u_0)}_{=0} + \underbrace{\beta(u_0,u_0)}_{=0} \\ &\geq 0. \end{aligned}$$

Da aber $v \in V_-$ liegt, gilt auch $q_\beta(v) < 0$, falls $v \neq 0$. Widerspruch. Also muss $v = 0$ sein.

Wir betrachten nun den Untervektorraum $E = (U_+ + U_0) + V_- \subseteq V$.

Da wir gerade gezeigt haben, dass $(U_+ + U_0) \cap V_- = \{0\}$ gilt, wissen wir, dass diese Summe direkt ist:

$$E = (U_+ + U_0) \oplus V_-.$$

Da U_- und U_0 sich nach Voraussetzung nur trivial schneiden, gilt sowieso $U_+ + U_0 = U_+ \oplus U_0$. Somit erhalten wir:

$$E = U_+ \oplus U_0 \oplus V_- .$$

Für die Dimension von E gilt nun:

$$n \geq \dim(E) = \dim(U_+) + \underbrace{\dim(U_0)}_{n-k} + \dim(V_-),$$

woraus durch Umstellen folgt:

$$\dim(U_+) + \dim(V_-) \leq k .$$

Erinnern wir uns, dass $k = \dim(U_+) + \dim(U_-)$ ist, so erhalten wir

$$\dim(U_+) + \dim(V_-) \leq \dim(U_+) + \dim(U_-) .$$

woraus schließlich

$$\dim(V_-) \leq \dim(U_-)$$

folgt.

Wenn wir nun die Rolle von (U_+, U_-, U_0) und $(V_+, V_-, V_0$ vertauschen, so erhalten wir aus Symmetriegründen ebenfalls

$$\dim(U_-) \leq \dim(U_+),$$

woraus die Gleichheit

$$\dim(U_-) = \dim(V_-)$$

folgt. □

3.5. Quadriken und Hauptachsentransformation

Für die lineare Algebra sind Lösungsmengen von linearen Gleichungen und linearen Gleichungssystemen sehr wichtig.

Trotzdem wollen wir uns in diesem Kapitel der Lösungsmenge von quadratischen Gleichungen widmen.

Als Motivation für alles was folgt, beginnen wir mit einem Fall, der hoffentlich schon sehr gut beherrscht wird, nämlich die Lösung von quadratischen Gleichungen einer Variablen:

Beispiel 3.5.1 (Quadratische Gleichungen in einer und mehreren Veränderlichen).

(a) Es sei $\mathbb{K}$ ein Körper mit $\mathrm{char}(\mathbb{K}) \neq 2$. Wir untersuchen die Gleichung

$$x^2 = s \quad \text{mit } s \in \mathbb{K}.$$

Die Frage lautet nun: Wie sieht die Lösungsmenge dieser Gleichung in $\mathbb{K}$ aus? Hier gibt es drei Fälle zu untersuchen:

(i) Falls $s = 0$, so gibt es genau eine Lösung. Die Lösungsmenge besteht also aus nur einem Punkt: $\{0\}$.

(ii) Falls $s \neq 0$ und s ein Quadrat in $\mathbb{K}$ ist (also $\exists w \in \mathbb{K} : w^2 = s$), so gibt es genau zwei Lösungen: $\{w, -w\}$.

(iii) Falls $s \neq 0$ und s ist kein Quadrat in $\mathbb{K}$, so ist die Lösungsmenge leer: $\emptyset$.

Beachten Sie, dass wir in Fall (ii) verwendet haben, dass $\mathrm{char}(\mathbb{K}) \neq 2$ ist. Wenn nämlich $\mathrm{char}(\mathbb{K}) = 2$ wäre, so hätten wir auch in Fall (ii) nur genau eine Lösung, da dann nämlich $w = -w$ wäre.

Wie kann man entscheiden, ob ein gegebenes $s \in \mathbb{K}$ ein Quadrat ist oder nicht?

- Die komplexen Zahlen $\mathbb{C}$ sind algebraisch abgeschlossen (siehe Fundamentalsatz der Algebra, LA1) und somit ist jedes s ein Quadrat und es gibt immer zwei verschiedene Lösungen (für $s \neq 0$). Der Fall (iii) kann also über $\mathbb{C}$ nicht eintreten.
- In den reellen Zahlen $\mathbb{R}$ gilt: Für $s > 0$ ist s ein Quadrat und für $s < 0$ ist s kein Quadrat. Somit ist das Vorzeichen von s entscheidend über die Anzahl der Lösungen. Unter der Verwendung der *Signum*-Funktion

$$\mathrm{sgn} : \mathbb{R} \to \{-1,0,1\}, \quad t \mapsto \begin{cases} 1 & \text{falls } t > 0 \\ 0 & \text{falls } t = 0\,. \\ -1 & \text{falls } t < 0 \end{cases}$$

gilt dann also:

$$\begin{aligned} x^2 = s \text{ hat genau zwei Lösungen} &\iff \mathrm{sgn}(s) = 1 \\ x^2 = s \text{ hat genau eine Lösung} &\iff \mathrm{sgn}(s) = 0 \\ x^2 = s \text{ hat keine Lösungen} &\iff \mathrm{sgn}(s) = -1 \end{aligned}$$

- Im Körper $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ gibt es analog dazu das sogenannte *Legendre-Symbol*, das aussieht wie ein gewöhnlicher Bruch mit Klammern, aber keiner ist, sondern für ein $k \in \mathbb{Z}$ und eine Primzahl $p > 2$ folgendermaßen definiert ist:

$$\begin{aligned} \left(\frac{k}{p}\right) := 1 &\iff x^2 = [k] \text{ hat genau zwei Lösungen in } \mathbb{Z}/p\mathbb{Z}. \\ \left(\frac{k}{p}\right) := 0 &\iff x^2 = [k] \text{ hat genau eine Lösung in } \mathbb{Z}/p\mathbb{Z}, \text{ d.h. } p \text{ ist ein Teiler von } k. \\ \left(\frac{k}{p}\right) := -1 &\iff x^2 = [k] \text{ hat keine Lösung in } \mathbb{Z}/p\mathbb{Z} \end{aligned}$$

.

Für kleine Primzahlen kann man durch Ausprobieren aller (endlich vielen) Kandidaten leicht überprüfen, ob ein gegebenes $[k] \in \mathbb{Z}/p\mathbb{Z}$ ein Quadrat ist oder nicht – zur effizienten Berechnung des Legendre-Symbols für große Primzahlen gibt es das sogenannte *quadratische Reziprozitätsgesetz*, auf das wir in dieser Veranstaltung aber nicht eingehen wollen.

(b) Es sei $\mathbb{K}$ ein Körper mit $\mathrm{char}(\mathbb{K}) \neq 2$. Eine allgemeine quadratische Gleichung

$$ax^2 + bx + c = 0 \quad \text{mit } a, b, c \in \mathbb{K}; a \neq 0$$

lässt sich durch Variablensubstitution (*quadratische Ergänzung*) in die Form aus (a) bringen. Hierzu setzen wir

$$u := x + \frac{b}{2a} \quad x = u - \frac{b}{2a}.$$

Dann wird die obige Gleichung zu

$$a\left(u - \frac{b}{2a}\right)^2 + b\left(u - \frac{b}{2a}\right) + c = 0,$$

was sich durch Ausmultiplizieren und Dividieren durch a zu

$$u^2 = \frac{b^2 - 4ac}{(2a)^2}$$

vereinfacht. Dies lässt sich nun mit den Methoden aus (a) lösen und dann wieder mit $x = u - \frac{b}{2a}$ zurück transformieren, womit wir dann auch die Lösungsformel für quadratische Gleichungen wiederentdeckt hätten.[16]

Diese Variablensubstitution entspricht nun einer bijektiven Abbildung:

$$\Phi : \mathbb{K} \to \mathbb{K}, \quad x \mapsto x + \frac{b}{2a}$$

Diese Abbildung ist keine lineare Abbildung, sondern eine Verschiebung/Translation. Sie „verschiebt die Gleichung“ so, dass die Lösungsmenge symmetrisch zur 0 ist und somit der lineare Term verschwindet.

Genau dieses Vorgehen funktioniert auch in mehr als einer Variablen, wie wir später sehen werden.

(c) Es sei nun $\mathbb{K} = \mathbb{R}$. Die wichtigste quadratische Gleichung in zwei Variablen ist die folgende:

$$x_1^2 + x_2^2 = 1.$$

Die Lösungsmenge besteht aus allen Punkten, die vom Ursprung Abstand 1 haben. Die Lösungsmenge ist also der Einheitskreis. Analog ist

$$x_1^2 + x_2^2 + x_3^2 = 1$$

die Gleichung für die Kugeloberfläche im Raum $\mathbb{R}^3$. Allgemein ist die Einheitssphäre im $\mathbb{R}^n$ gegeben als Lösungsmenge der quadratischen Gleichung

$$\sum_{j=1}^{n} x_j^2 = 1.$$

Die Einheitssphäre im n-dimensionalen Raum ist übrigens nur ein $(n-1)$-dimensionales Objekt, weil sie lokal nur wie eine gekrümmte $(n-1)$-dimensionale Ebene aussieht. Man schreibt für die Einheitssphäre im $\mathbb{R}^n$ auch

$$S^{n-1} := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \;\middle|\; \sum_{j=1}^{n} x_j^2 = 1 \right\}$$

und nennt dieses Objekt die $(n-1)$-Sphäre. Die n-Sphäre ist somit dann eine Teilmenge von $\mathbb{R}^{n+1}$.

[16] Wo haben wir in dieser Rechnung verwendet, dass $\mathrm{char}(\mathbb{K}) \neq 2$?

(d) Betrachten wir noch einmal den Einheitskreis

$$x_1^2 + x_2^2 = 1.$$

Gegeben seien $a_1, a_2 > 0$. Wenn wir nun x_1 durch x_1/a_1 und x_2 durch x_2/a_2 ersetzen, so erhalten wir die Gleichung

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1.$$

Geometrisch entspricht das einer Streckung um den Faktor a_1 in x_1-Richtung und einer Streckung um den Faktor a_2 in x_2-Richtung, die Lösungsmenge dieser so entstandenen Gleichung nennt man eine *Ellipse* mit den *Halbachsen* a_1 und a_2.

Falls $a_1 = a_2 = a$, so ist diese Ellipse einfach nur ein Kreis mit Radius a.

Wenn wir analog dazu eine höherdimensionale Sphäre in Richtung der einzelnen Koordinatenachsen strecken oder stauchen, erhalten wir ein *Ellipsoid*.

(e) Die Gleichung

$$x_1^2 - x_2^2 = 1$$

beschreibt eine *Hyperbel* mit den beiden Winkelhalbierenden als Asymptoten. Im Gegensatz zum Einheitskreis ist diese Menge unbeschränkt, weil es auf der Hyperbel Punkte gibt, die beliebig weit vom Ursprung entfernt sind. Ein weiterer Unterschied ist, dass die Hyperbel zwei *Zusammenhangskomponenten* besitzt: Ein Hyperbelast rechts von der x_2-Achse und einen damit nicht zusammenhängenden links von der x_2-Achse.

(f) Die Gleichung

$$x_1^2 - x_2 = 0$$

ist nichts anderes als der Graph der Funktion

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2$$

und somit ist die Lösungsmenge dieser Gleichung eine Standardparabel.

(g) Die Gleichung

$$x_1^2 + x_2^2 - x_3^2 = 0$$

beschreibt den Mantel eines *Doppelkegels*. Man kann zeigen, dass jede nichtleere Lösungsmenge einer quadratischen Gleichung in *zwei* Variablen durch den Schnitt dieses Doppelkegels mit einer entsprechenden (zweidimensionalen) Ebene im $\mathbb{R}^3$ entsteht. Deswegen nennt man Lösungsmengen von quadratischen Gleichungen in zwei Variablen auch *Kegelschnitte*.

Versuchen wir uns nun an einer formalen Definition, was eine quadratische Gleichung in n Variablen sein soll. Wir werden einfach die Gleichung

$$ax^2 + bx + c = 0$$

in einer Variablen entsprechend verallgemeinern:

Definition 3.5.2 (Quadrik)**.**
Es sei $\mathbb{K}$ ein Körper mit $\operatorname{char}(\mathbb{K}) \neq 2$ und $n \in \mathbb{N}$. Eine *Quadrik* ist eine Teilmenge von $\mathbb{K}^n$, die sich schreiben lässt als die Lösungsmenge einer Gleichung der Form

$$q(x) + 2\varphi(x) + c = 0,$$

wobei $q : \mathbb{K}^n \to \mathbb{K}$ eine quadratische Form mit $q \neq 0$, $\varphi : \mathbb{K}^n \to \mathbb{K}$ eine Linearform und $c \in \mathbb{K}$ eine Konstante ist.

Eine Gleichung dieser Form nennt man *quadratische Gleichung* in n Variablen.

Falls $q = 0$ ist, so liegt keine quadratische Gleichung vor, sondern eine lineare Gleichung.

Bemerkung 3.5.3. Dank Satz 3.3.10 wissen wir, dass zu einer quadratischen Form immer eine symmetrische Bilinearform β gehört und somit können wir eine Quadrik immer beschreiben als

$$\left\{x \in \mathbb{K}^n \;\middle|\; x^\top A x + 2b^\top x + c = 0\right\},$$

wobei $A = \mathrm{FM}_{\mathsf{E}}(\beta) \in \mathbb{K}^{n\times n}$ eine symmetrische Matrix ungleich 0 ist, $b = (\varphi)_{\mathsf{E}^*} = (\mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi))^\top$ der Koordinatenvektor von φ bezüglich der dualen Standardbasis und $c \in \mathbb{K}$ eine Konstante ist.

Durch Ausmultiplizieren erhält man auch folgende Darstellung:

$$\sum_{i,j=1}^{n} a_{i,j} x_i x_j + 2\sum_{i=1}^{n} b_i x_i + c = 0$$

mit Skalaren $a_{i,j}, b_i, c \in \mathbb{K}$, sodass nicht alle $a_{i,j}$ gleich null sind und $a_{i,j} = a_{j,i}$ gilt.

Das Ziel wird es nun sein, durch lineare Transformationen und Translationen eine gegebene Quadrik in eine besonders schöne Form zu bringen. Der erste Schritt ist hierbei, die Matrix A als Fundamentalmatrix zu diagonalisieren. Damit lässt sich erreichen, dass die gemischten Terme x_1x_2 verschwinden:

Proposition 3.5.4.
Es sei $\mathbb{K}$ *ein Körper mit* $\operatorname{char}(\mathbb{K}) \neq 2$*,* $n \in \mathbb{N}$ *und*

$$Q \subseteq \mathbb{K}^n$$

eine Quadrik. Dann ist es immer möglich, einen Vektorraumautomorphismus

$$\psi : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Sx \quad \textit{mit } S \in \mathrm{GL}(n, \mathbb{K})$$

und Skalare $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n, c \in \mathbb{K}$ *zu finden, sodass*

$$\psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n \;\middle|\; \sum_{i=1}^{n} \alpha_i u_i^2 + 2\sum_{i=1}^{n} \beta_i u_i + c = 0 \right\}.$$

Falls $\mathbb{K} = \mathbb{R}$ *ist, so kann man* ψ *als isometrischen Isomorphismus des Standardskalarproduktes wählen, sodass bei der Transformation Längen und Winkel erhalten bleiben.*

Beweis. Die Quadrik Q ist gegeben durch

$$Q = \left\{x \in \mathbb{K}^n \;\middle|\; x^\top A x + 2b^\top x + c = 0\right\},$$

wobei $A \in \mathbb{K}^{n\times n}$ eine symmetrische Matrix ungleich 0 ist, $b \in \mathbb{K}^n$ ein Vektor und $c \in \mathbb{K}$ ein Skalar.

Nach Satz 3.3.9 gibt es eine Basiswechselmatrix $S \in \mathrm{GL}(n,\mathbb{K})$, sodass

$$A = S^\top DS \quad \text{mit } D = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}$$

gilt. Falls $\mathbb{K} = \mathbb{R}$, so können wir Satz 2.4.7 verwenden, um sogar $S \in \mathrm{O}(n)$ zu bekommen.

Wir setzen:

$$\psi : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Sx$$

Nun können wir schreiben:

$$\begin{aligned}
u \in \psi(Q) &\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } x \in Q \\
&\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } x^\top Ax + 2b^\top x + c = 0 \\
&\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } x^\top S^\top DSx + 2b^\top x + c = 0 \\
&\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } (Sx)^\top D(Sx) + 2b^\top S^{-1}Sx + c = 0 \\
&\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } (Sx)^\top D(Sx) + 2\left((S^{-1})^\top b\right)^\top (Sx) + c = 0 \\
&\iff \exists x \in \mathbb{K}^n, u = \psi(x) \text{ und } \psi(x)^\top D\psi(x) + 2\left((S^{-1})^\top b\right)^\top \psi(x) + c = 0 \\
&\iff u^\top Du + 2\left((S^{-1})^\top b\right)^\top u + c = 0
\end{aligned}$$

Wenn wir nun den Vektor $(S^{-1})^\top b \in \mathbb{K}^n$ schreiben als

$$(S^{-1})^\top b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

erhalten wir damit die gesuchte Form. □

In einem nächsten Schritt werden wir versuchen durch eine Translation (Quadratische Ergänzung), die linearen Terme in der quadratischen Gleichung auch noch loszuwerden, obwohl wir bereits in Beispiel 3.5.1 (f) gesehen haben, dass dies nicht in allen Fällen möglich sein wird.

Solche Translationen sind leider keine linearen Abbildungen mehr (sie bilden ja insbesondere nicht 0 auf 0 ab, was eine notwendige Bedingung für Linearität ist). Deshalb erweitern wir den Begriff der linearen Abbildung:

Definition 3.5.5 (Affine Abbildungen)**.** Es sei $\mathbb{K}$ ein Körper und V und W seien Vektorräume über $\mathbb{K}$.

Eine Abbildung

$$\Psi : V \to W$$

heiße *affine Abbildung*, wenn sich Ψ schreiben lässt als die Summe einer linearen und einer konstanten Abbildung, d.h. Ψ ist genau dann affin, wenn es eine lineare Abbildung $\varphi : V \to W$ und eine Konstante $w_0 \in W$ gibt, sodass

$$\forall v \in V : \ \Psi(v) = \varphi(v) + w_0.$$

Eine bijektive affine Abbildung $\Psi : V \to W$ heißt auch *affiner Isomorphismus* oder *Affinität* zwischen V und W.

Ein *affiner Automorphismus* von V oder eine *Affinität* auf V ist nun ein affiner Isomorphismus zwischen V und sich selbst. Die Menge aller affinen Automorphismen eines Vektorraums V bezeichnen wir mit $\mathrm{Aff}(V)$.

Lemma 3.5.6.

- *Die Zerlegung $\Psi = \varphi + w_0$ ist eindeutig.*
- *Linearkombinationen affiner Abbildungen sind wieder affin.*[17]
- *Verkettungen von affinen Abbildungen sind wieder affin.*
- *Die Umkehrabbildung eines affinen Isomorphismus ist wieder ein affiner Isomorphismus.*
- *Das Bild eines affinen Unterraums unter einer affinen Abbildung ist wieder ein affiner Unterraum.*
- *Das Urbild eines affinen Unterraums unter einer affinen Abbildung ist entweder leer oder wieder ein affiner Unterraum.*
- *Die Menge $\mathrm{Aff}(V)$ ist eine Gruppe mit der Verkettung.*

Mit diesen neuen Begriffen wollen wir nun versuchen, eine gegebene Quadrik einfacher zu machen, indem wir affine Isomorphismen auf sie anwenden. Wir beginnen mit dem Fall, dass die zur Quadrik gehörende symmetrische Bilinearform nicht entartet ist, also vollen Rang hat:

Proposition 3.5.7.
Es sei $\mathbb{K}$ ein Körper mit $\mathrm{char}(\mathbb{K}) \neq 2$, $n \in \mathbb{N}$ und

$$Q = \left\{x \in \mathbb{K}^n \;\middle|\; q_\beta(x) + 2\varphi(x) + c = 0\right\} \subseteq \mathbb{K}^n$$

eine Quadrik, wobei die zugrundeliegende symmetrsiche Bilinearform $\beta : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ als nicht entartet angenommen wird.

Dann ist es immer möglich, einen affinen Automorphismus

$$\Psi : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Sx + v_0 \quad \text{mit } S \in \mathrm{GL}(n,\mathbb{K}) \text{ und } v_0 \in \mathbb{K}^n$$

und Skalare $\alpha_1, \dots, \alpha_n \neq 0$, sowie ein $d \in \mathbb{K}$ zu finden, sodass

$$\forall u \in \mathbb{K}^n : \; \Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n \;\middle|\; \sum_{i=1}^n \alpha_i u_i^2 + d = 0 \right\}.$$

Im entarteten Fall müssen wir nun noch etwas arbeiten.

Wir nehmen im Folgenden an, $\mathbb{K}$ sei ein beliebiger Körper mit Charakteristik ungleich 2. Wir werden später noch genauer darauf eingehen, was dies für den Spezialfall $\mathbb{K} = \mathbb{R}$ bedeutet.

Beim Diagonalisieren unserer Quadrik können wir in jedem Fall dafür sorgen, dass die Nichtnulleinträge (sofern vorhanden) oben links in der Diagonalmatrix stehen und die Nullen (sofern vorhanden) unten rechts.

Nehmen wir also von nun an, unsere Quadrik hat die folgende Form:

$$x^\top A x + 2b^\top x + c = 0$$

[17] Die Menge aller affinen Abbildungen von V nach W ist also ein Untervektorraum von dem Vektorraum W^V von *allen* Funktionen von V nach W.

mit

$$A = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \\ b_{m+1} \\ \vdots \\ b_n \end{pmatrix}, \quad c \in \mathbb{K},$$

wobei $m \in \{0, \dots, n\}$ der Rang von A ist.

Mittels quadratischer Ergänzung können wir durch einen affinen Koordinatenwechsel die Zahlen $b_1, \dots, b_m$ zu Null machen, für die Zahlen $b_{m+1}, \dots, b_n$ geht dies leider nicht, weil kein entsprechender quadratischer Term existiert.

Was aber geht: Durch einen geschickten Basiswechsel, also eine lineare Transformation, können wir nun alle Einträge $b_{m+2}, \dots, b_n$ zu Null zu bringen, sodass in unserem Vektor b anschließend höchstens ein Nicht-Null-Term übrig bleibt.

Der Nullraum $\mathrm{Null}(\beta)$ der Bilinearform β ist gerade der Kern von A und wird – wenn die Matrix A in der oben angegebenen Form ist – von den letzten $(n-m)$ Basisvektoren aufgespannt.

Betrachten wir nun die Abbildung

$$\varphi|_{\mathrm{Null}(\beta)} : \mathrm{Null}(\beta) \to \mathbb{K}, \quad x \mapsto \varphi(x) = b^\top x.$$

Falls diese Abbildung gleich 0 ist, gibt es nichts zu tun, denn dann sind alle Terme $b_{m+1}, \dots, b_n$ bereits gleich 0. Falls die Abbildung $\varphi|_{\mathrm{Null}(\beta)}$ nicht die Nullabbildung ist, so können wir für den Kern dieser Abbildung, der dann $(n-m-1)$-dimensional ist, eine Basis $w_{m+2}, \dots, w_n$ finden und dann mit einem Vektor w_{m+1} zu einer Basis von $\mathrm{Null}(\beta)$ ergänzen.

Bei diesem Basiswechsel ändert sich die Matrix A nicht, aber der Vektor b hat anschließend ab der $(m+2)$-ten Komponente nur noch Nullen.

Satz 3.5.8 (Klassifikation der Quadriken (affiner Fall))**.** *Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein beliebiger Körper $\mathbb{K}$ mit* $\mathrm{char}(\mathbb{K}) \neq 2$ *und $Q \subseteq \mathbb{K}^n$ eine Quadrik. Dann gibt es immer einen affinen Automorphismus*

$$\Psi : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Sx + v_0 \quad \text{mit } S \in \mathrm{GL}(n, \mathbb{K}) \text{ und } v_0 \in \mathbb{K}^n,$$

eine Zahl $m \in \{1, \dots, n\}$ und Skalare $\alpha_1, \dots, \alpha_m \neq 0$, sodass genau einer der drei folgenden Fälle zutrifft:

(i) Q ist eine kegelartige Quadrik *(oder auch* kegelige Quadrik*), d.h.*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n \;\middle|\; \sum_{j=1}^m \alpha_j u_j^2 = 0 \right\}$$

(ii) Q ist eine Mittelpunktsquadrik*, d.h.*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n \;\middle|\; \sum_{j=1}^m \alpha_j u_j^2 = 1 \right\}$$

(iii) Q ist eine parabolische Quadrik*, d.h. $m \leq n-1$ und*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n \;\middle|\; \sum_{j=1}^{m} \alpha_j u_j^2 + 2u_{m+1} = 0 \right\}.$$

Die Zahl m entspricht hier dem Rang der zur Quadrik gehörenden symmetrischen Bilinearform.

Beachten Sie, dass in Satz 3.5.8 die Quadriken bis auf *affine Transformationen* klassifiziert werden, d.h. die Abbildung Ψ kann (im Falle von $\mathbb{K} = \mathbb{R}$ auch Längen oder Winkel ändern, so kann aus einer Ellipse beispielsweise ein Kreis werden). Will man die Geometrie einer Quadrik genauer bestimmen und auch metrische Größen wie Halbachsen einer Ellipse berechnen, so ist eine beliebige affine Transformation zu allgemein.

Glücklicherweise gibt es einen entsprechenden Satz in $\mathbb{K} = \mathbb{R}$, der eine Klassifikation bis auf Isometrie erlaubt:

Satz 3.5.9 (Klassifikation der Quadriken (isometrischer Fall; Hauptachsentransformation))**.** *Es sei $n \in \mathbb{N}$ und $Q \subseteq \mathbb{R}^n$ eine Quadrik. Dann gibt es immer einen* isometrischen *affinen Automorphismus*

$$\Psi : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto Sx + v_0 \quad \text{mit } S \in \mathrm{O}(n) \text{ und } v_0 \in \mathbb{R}^n,$$

eine Zahl $m \in \{1, \ldots, n\}$ und Skalare $\alpha_1, \ldots, \alpha_m \neq 0$, sodass genau einer der drei folgenden Fälle zutrifft:

(i) Q ist eine kegelartige Quadrik *(oder auch* kegelige Quadrik*), d.h.*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n \;\middle|\; \sum_{j=1}^{m} \alpha_j u_j^2 = 0 \right\}$$

(ii) Q ist eine Mittelpunktsquadrik*, d.h.*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n \;\middle|\; \sum_{j=1}^{m} \alpha_j u_j^2 = 1 \right\}$$

(iii) Q ist eine parabolische Quadrik*, d.h. $m \leq n-1$ und*

$$\Psi(Q) = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n \;\middle|\; \sum_{j=1}^{m} \alpha_j u_j^2 + 2u_{m+1} = 0 \right\}.$$

Die Zahl m entspricht hier dem Rang der zur Quadrik gehörenden symmetrsichen Bilinearform.

Das Verfahren, eine Quadrik durch eine affine Isometrie Ψ in eine dieser Formen zu bringen, nennt man auch Hauptachsentransformation*.*

Beispiel 3.5.10 (Quadriken im $\mathbb{R}^2$)**.** Für den Fall $\mathbb{K} = \mathbb{R}$ und $n = 2$ lassen sich alle Quadriken wie folgt klassifizieren:

Es sei $Q \subseteq \mathbb{R}^2$ eine Quadrik. Dann gibt es mehrere Fälle zu untersuchen

(i) Q ist *kegelartig*: Der Rang m ist entweder 1 oder 2.

- Falls $m = 1$ ist:
$$\alpha_1 u_1^2 = 0.$$
Dies ist äquivalent zu
$$u_1 = 0.$$
Dies ist eine Geradengleichung. Unsere Quadrik beschreibt also eine **Gerade in der Ebene**.
- Falls $m = 2$ ist:
$$\alpha_1 u_1^2 + \alpha_2 u_2^2 = 0.$$
Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $\alpha_1 > 0$ ist.
Falls $\alpha_2 > 0$ ist, so erhalten wir
$$\alpha_1 u_1^2 + \alpha_2 u_2^2 = 0,$$
was sich umschreiben lässt zu:
$$\left\| \begin{pmatrix} \sqrt{\alpha_1} u_1 \\ \sqrt{\alpha_2} u_2 \end{pmatrix} \right\|^2 = 0.$$
und dies ist nur möglich, wenn $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ gilt. Wir erhalten in diesem Fall also einen **Punkt in der Ebene**.
Falls aber $\alpha_2 < 0$ ist, dann gilt $\alpha_2 = -|\alpha_2|$ und so erhalten wir die Gleichung
$$\sqrt{\alpha_1}^2 u_1^2 - \sqrt{|\alpha_2|}^2 u_2^2 = 0,$$
die sich wie folgt umschreiben lässt:
$$\left(\sqrt{\alpha_1} u_1 + \sqrt{|\alpha_2|} u_2\right)\left(\sqrt{\alpha_1} u_1 - \sqrt{|\alpha_2|} u_2\right) = 0.$$
Ein Punkt $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ liegt also in $\Psi(Q)$, wenn er entweder auf der Geraden
$$\sqrt{\alpha_1} u_1 + \sqrt{|\alpha_2|} u_2 = 0$$
oder auf der Geraden
$$\sqrt{\alpha_1} u_1 - \sqrt{|\alpha_2|} u_2 = 0$$
liegt. Wir erhalten in diesem Fall also **zwei sich schneidende Geraden**.

(ii) Q ist eine *Mittelpunktsquadrik*: Der Rang m ist entweder 1 oder 2.

- Falls $m = 1$ ist:
$$\alpha_1 u_1^2 = 1.$$
Falls $\alpha_1 < 0$ ist, ist dies die **leere Menge**.
Falls $\alpha_1 > 0$ ist, erhalten wir die Gleichung
$$\left(u_1 - \frac{1}{\sqrt{\alpha_1}}\right)\left(u_1 + \frac{1}{\sqrt{\alpha_1}}\right) = 0.$$
Wir erhalten in diesem Fall also **zwei parallele Geraden**.

– Falls $m = 2$ ist:
$$\alpha_1 u_1^2 + \alpha_2 u_2^2 = 1.$$
Auch hier gilt es, die Vorzeichen von α_1 und α_2 zu untersuchen:

Falls $\alpha_1 < 0$ und $\alpha_2 < 0$ sind, ist dies wieder die **leere Menge**.

Falls $\alpha_1 > 0$ und $\alpha_2 > 0$ sind, so können wir die Gleichung umformen zu
$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1,$$
und erhalten genau die Gleichung einer **Ellipse** (siehe Beispiel 3.5.1(d)), wobei
$$a_1 := \frac{1}{\sqrt{\alpha_1}} \quad \text{und} \quad a_2 := \frac{1}{\sqrt{\alpha_2}}$$
die Halbachsen der Ellipse sind.[18]

Im Falle, dass α_1 positiv und α_2 negativ ist, erhält man die Gleichung einer **Hyperbel** (siehe Beispiel 3.5.1(e)):
$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1$$
mit
$$a_1 := \frac{1}{\sqrt{\alpha_1}} \quad \text{und} \quad a_2 := \frac{1}{\sqrt{|\alpha_2|}}.$$

(iii) Q ist *parabolisch*: Für den Rang m gilt in diesem Fall: $m = 1$ und die Gleichung lautet:
$$\alpha_1 u_1^2 + 2u_2 = 0.$$
Dies ist eine **Parabel** (siehe Beispiel 3.5.1(f)).

[18]Und falls α_1 und α_2 gleich sind, erhalten wir sogar einen Kreis mit Radius $r = \frac{1}{\sqrt{\alpha_1}}$.

A. Anhang

A.1. Das griechische Alphabet

α			A	Alpha
β			B	Beta
γ			Γ	Gamma
δ			Δ	Delta
ε	oder	ϵ	E	Epsilon
ζ			Z	Zeta
η			H	Eta
ϑ	oder	θ	Θ	Theta
ι			I	Iota
κ			K	Kappa
λ			Λ	Lambda
μ			M	My
ν			N	Ny
ξ			Ξ	Xi
o			O	Omikron
π			Π	Pi
ρ			P	Rho
σ			Σ	Sigma
τ			T	Tau
υ			Y	Ypsilon
φ	oder	ϕ	Φ	Phi
χ			X	Chi
ψ			Ψ	Psi
ω			Ω	Omega

Stichwortverzeichnis

Ableitungsoperator, 47
Abstand eines Punktes von einer Teilmenge, 25
Abstandsfunktion, 7, 17
adjungierte Abbildung, 45
adjungierte Matrix, 33
affine Abbildung, 97
affiner Automorphismus, 98
affiner Isomorphismus, 97
Affinität, 97, 98
allgemeine lineare Gruppe, 34
ausgeartete Bilinearform, 79

Bidualraum, 60
bilinear, 13
Bilinearform, 66
binomische Formeln für Skalarprodukte, 13

Cauchy-Schwarzsche Ungleichung, 14
Charakteristik Zwei, 78
Currying, 69

Definitheit, 81
degenerierte Bilinearform, 79
Determinantenform, 67
Distributionen, 58
Doppelkegel, 95
Drehkästchen, 35, 50
Dreiecksungleichung, 16, 17
duale Abbildung, 59
duale Basis, 63
Dualraum, 57

Ellipse, 95, 102
Ellipsoid, 95
entartete Bilinearform, 79
Euklidischer Vektorraum, 10

Fundamentalmatrix, 72

geordnete Orthogonalbasis, 19
geordnete Orthonormalbasis, 19
Gram-Schmidt-Algorithmus, 20

Halbachse einer Ellipse, 95
Hamming-Abstand, 18
Hauptabschnittsunterdeterminanten, 86
Hauptachsentransformation, 100
Hauptminoren, 86
hermitesche Matrix, 40
Hesse-Matrix, 68
Homogenität der Norm, 16
Hurwitz-Kriterium, 86
Hyperbel, 95, 102

indefinit, 81, 82, 87
Innenproduktraum, 10
Isometrie, 29
isometrisch isomorph, 31
isometrische Einbettung, 29
isometrischer Isomorphismus, 31

Kategorientheorie, 65
kegelartige Quadrik, 99, 100
kegelige Quadrik, 99, 100
Kegelschnitt, 95
komplex konjugierte Matrix, 33
komplexes Standardskalarprodukt, 11
konjugiert linear, 13
Kreuzprodukt, 7

Legendre-Symbol, 93
lichtartige Vektoren, 68
lineare Isometrie, 31
lineares Funktional, 57
Linearform, 57

Metrik, 17
metrischer Raum, 8, 17
Minkowski-Raum, 67
Mittelpunktsquadrik, 99, 100
multilineare Abbildung, 66

negativ definit, 81, 82, 87
negativ semidefinit, 81, 82, 87
nicht ausgeartete Bilinearform, 79
nicht degenerierte Bilinearform, 79
nicht entartete Bilinearform, 79
Norm, 7, 13, 18
normale Abbildung, 49
normale Matrix, 42
normierter Vektorraum, 8, 18

ONB, 19
ONS, 19
orientierungserhaltende Isometrie, 36
orientierungsumkehrende Isometrie, 36
orthogonal diagonalisierbar, 38
orthogonal diagonalisierbarer Endomorphismus, 39
orthogonal trigonalisierbar, 38
orthogonal trigonalisierbarer Endomorphismus, 39
orthogonal ähnlich, 38
Orthogonalbasis, 19
Orthogonalbasis für eine Bilinearform, 77
orthogonale Gruppe, 34
orthogonale Matrix, 34
orthogonales Komplement, 27
Orthogonalität, 7
Orthogonalität bezüglich eines Skalarproduktes, 11
Orthogonalprojektion, 26
Orthogonalraum, 26
Orthogonalsystem, 19
Orthonormalbasis, 19
Orthonormalsystem, 19

Parabel, 95, 102
parabolische Quadrik, 100
Parallelogrammgleichung, 14
Polarisierungsformel, 78
Polarisierungsidentität, 78
positiv definit, 81, 82, 87
positiv semidefinit, 81, 82, 87
positive Definitheit der Metrik, 17
positive Definitheit der Norm, 16
Prähilbertraum, 10

quadratische Ergänzung, 94
quadratische Form, 67
quadratische Gleichung, 96
quadratisches Reziprozitätsgesetz, 93
Quadrik, 96

raumartige Vektoren, 68

Satz des Pythagoras für Vektorräume mit Skalarprodukt, 14
Schleichwerbung, 65, 69
selbstadjungierte Abbildung, 49
selbstadjungierte Matrix, 40
sesquilinear, 13
Signatur, 90
Signum, 93
Skalarprodukt, 10
Spektralsatz, 41, 44, 49
spezielle orthogonale Gruppe, 34
spezielle unitäre Gruppe, 36
Standardskalarprodukt, 8, 11
Symmetrie der Metrik, 17
symmetrische Bilinearform, 77

Translation, 29

Ungleichung von Cauchy-Schwarz, 14
unitär diagonalisierbar, 38
unitär trigonalisierbar, 38
unitär ähnlich, 38
unitäre Gruppe, 36
unitäre Matrix, 36
unitärer Vektorraum, 11

Vektorprodukt, 7
Vektorraum mit Skalarprodukt, 10

Winkel zwischen Vektoren, 7, 17

zeitartige Vektoren, 68
Zusammenhangskomponente, 95