



Karlsruher Institut für Technologie

LINEARE ALGEBRA II FÜR INFORMATIK

SOMMERSEMESTER 2012

Dr. WOLFGANG GLOBKE
INSTITUT FÜR ALGEBRA UND GEOMETRIE
KARLSRUHER INSTITUT FÜR TECHNOLOGIE (KIT)

Version vom 18. Juli 2012.

Vorwort

Das vorliegende Skriptum entsteht im Sommersemester 2012 für den zweiten Teil der Vorlesung *Lineare Algebra für die Fachrichtung Informatik*. Es baut auf dem ersten Teil der Vorlesung auf, wie er nachzulesen ist im Skriptum von Prof. Enrico Leuzinger [10].

In Teil II der Vorlesung wird der Stoff stärker auf das Wesentliche reduziert als in Teil I. Dennoch soll verstärkt ein Einblick in Anwendungen im Bereich der Informatik gegeben werden. Insbesondere die algorithmischen Aspekte werden stärker betont.

Der Anhang bietet weiterführende Informationen, die nicht unmittelbar zum Stoff der Vorlesung gehören, aber das Verständnis an einigen Stellen erleichtert und vertieft.

Was mich nicht um bringt, macht mich stärker.

FRIEDRICH NIETZSCHE

Notation

Wir werden in diesem Skript die folgenden Schreibweisen für Konzepte aus dem ersten Teil der Vorlesung verwenden:

- Das Kronecker-Symbol bezeichnen wir mit δ_{ij} (es gilt $\delta_{ii} = 1$ und $\delta_{ij} = 0$ falls $i \neq j$).
- Für die Äquivalenzklasse eines Elementes x aus einer Menge (Gruppe, Ring, Vektorraum...) X schreiben wir $[x]$ oder \bar{x} , je nachdem, was der Lesbarkeit förderlicher ist.
- Der Ring der $m \times n$ -Matrizen mit Einträgen aus einem Ring R wird mit $R^{m \times n}$ bezeichnet. Ist \mathbb{K} ein Körper, so bezeichnet $\mathbf{GL}_n(\mathbb{K})$ die (multiplikative) Gruppe der invertierbaren Matrizen.
- Die Vektoren der Standardbasis im \mathbb{K}^n bezeichnen wir mit e_1, \dots, e_n . Die $n \times n$ -Einheitsmatrix wird mit I_n oder I bezeichnet.
- Mit $R[X]$ bezeichnen wir den Ring der Polynome in der Variablen X mit Koeffizienten aus einem Ring R . Für den Grad eines Polynoms $f \in R[X]$ schreiben wir $\deg(f)$. Das Nullpolynom hat nach Definition den Grad $-\infty$.
- Für \mathbb{K} -Vektorräume V, W bezeichnet $\text{Hom}(V, W)$ die Menge der Homomorphismen (\mathbb{K} -lineare Abbildungen) von V nach W . Wenn wir die Abhängigkeit vom Skalarkörper \mathbb{K} betonen, schreiben wir auch $\text{Hom}_{\mathbb{K}}(V, W)$. Für die Menge der Endomorphismen $\text{Hom}(V, V)$ schreiben wir auch $\text{End}(V)$. Die Gruppe der Automorphismen (invertierbare Endomorphismen) wird mit $\mathbf{Aut}(V)$ oder $\mathbf{GL}(V)$ bezeichnet. (Analoge Notationen verwenden wir für Homomorphismen von Gruppen, Ringen und Körpern.)
- Der Dualraum $\text{Hom}(V, \mathbb{K})$ zu einem \mathbb{K} -Vektorraum V wird mit V^* bezeichnet. Für eine lineare Abbildung $\Phi : V \rightarrow W$ wird die duale Abbildung mit $\Phi^* : W^* \rightarrow V^*$ bezeichnet.
- Die Darstellung eines Vektors $x \in V$ bzgl. einer Basis B durch einen Spaltenvektor in \mathbb{K}^n schreiben wir $\Theta_B(x)$. Die Abbildungsmatrix einer linearen Abbildung Φ bzgl. zweier Basen C und B schreiben wir $M_B^C(\Phi)$. Für Basiswechselmatrizen schreiben wir kurz M_B^C statt $M_B^C(\text{id}_V)$.
- Der Eigenraum zum Eigenwert λ eines Endomorphismus Φ wird mit E_λ oder $E_\lambda(\Phi)$ bezeichnet. Die Menge aller Eigenwerte (das Spektrum von Φ) schreiben wir als $\text{Spec } \Phi$.

Weitere Notation wird im Laufe der Vorlesung eingeführt.

Inhaltsverzeichnis

Notation	vii
I Polynome und Endomorphismen	1
1 Teilbarkeit in Ringen	1
1.1 Einheiten, Ideale und Teilbarkeit	1
1.2 Euklidische Ringe	9
1.3 Nullstellen von Polynomen	17
1.4 Quotientenringe	18
1.5 Der chinesische Restsatz	23
2 Die Jordansche Normalform	33
2.1 Invariante Untervektorräume	33
2.2 Nilpotente Endomorphismen	36
2.3 Das Programm	39
2.4 Die Primärzerlegung	42
2.5 Die Normalform nilpotenter Endomorphismen	46
2.6 Die Jordansche Normalform	52
3 Kryptographie	57
3.1 Der Satz von Lagrange	58
3.2 Einheiten in $\mathbb{Z}/n\mathbb{Z}$	59
3.3 Schlüsselaustausch nach Diffie und Hellman	60
3.4 El Gamal-Verschlüsselung	61
3.5 RSA-Verschlüsselung	62
II Geometrie in Vektorräumen	65
4 Vektorräume mit Skalarprodukt	65
4.1 Das Standardskalarprodukt	65

4.2	Bilinearformen	68
4.3	Euklidische Vektorräume	72
4.4	Normen, Winkel und Orthogonalität	75
4.5	Unitäre Vektorräume	82
5	Orthogonalsysteme	87
5.1	Mengen orthogonaler Vektoren	87
5.2	Das Gram-Schmidt-Verfahren	91
5.3	Orthogonalprojektionen	96
5.4	Vollständige Orthogonalsysteme und Hilbert-Räume	99
5.5	Positiv definite Matrizen	104
5.6	Eine Anomalie im \mathbb{R}^3 : Das Vektorprodukt	109
6	Isometriegruppen	113
6.1	Isometrien	113
6.2	Orthogonale und unitäre Gruppen	115
6.3	Die Normalform für lineare Isometrien	125
6.4	Euler-Winkel	133
7	Selbstadjungierte Endomorphismen	137
7.1	Die adjugierte Abbildung	137
7.2	Selbstadjungierte Endomorphismen	140
7.3	Der Spektralsatz	141
7.4	Normale Endomorphismen	143
III	Anhang	147
A	Geometrie der komplexen Zahlen	147
A.1	Die Gaußsche Zahlenebene	147
A.2	Komplexe Multiplikation	148
A.3	Komplexe Konjugation	150
A.4	Einheitswurzeln	150

B Trigonometrische Funktionen	153
B.1 Funktionsgraphen	153
B.2 Rechenregeln	153
B.3 Wertetabelle	155
Literatur	156
Index	157

Teil I

Polynome und Endomorphismen

Unser Ziel im ersten Teil der Vorlesung ist es, eine vollständige Klassifikation der Konjugationsklassen von komplexen Matrizen herzuleiten. Das bedeutet, dass jede Matrix $A \in \mathbb{C}^{n \times n}$ zu einer eindeutigen Matrix \tilde{A} ähnlich ist, die eine besonders einfache Gestalt annimmt und nur von der Konjugationsklasse von A abhängt. Dies ist die sogenannte *Jordansche Normalform* von A . Sie verallgemeinert die bereits bekannte Diagonalform für diagonalisierbare Matrizen.

Die genaue Gestalt der Jordanschen Normalform von A ist mit der Gestalt des charakteristischen Polynoms f_A von A eng verknüpft. Das werden wir zum Anlass nehmen, im ersten Kapitel die Struktur der Polynomringe $\mathbb{K}[X]$ genauer zu studieren. Dabei werden wir Analogien zu den Ringen der ganzen Zahlen \mathbb{Z} und den Restklassenringen $\mathbb{Z}/n\mathbb{Z}$ besonders herausarbeiten. Ein erfreulicher Nebeneffekt dieser Untersuchungen wird sein, dass wir bereits einige einfache Verfahren aus der Kryptographie verstehen können.

1 Teilbarkeit in Ringen

Sofern nicht anders angegeben, sei in diesem Kapitel stets R ein Ring mit Eins.

1.1 Einheiten, Ideale und Teilbarkeit

Einer der einfachsten uns bekannten Ringe ist der Ring \mathbb{Z} der ganzen Zahlen. Die einzigen multiplikativ invertierbaren Elemente von \mathbb{Z} sind -1 und $+1$, und jedes Element $n \in \mathbb{Z}$ lässt sich als Produkt

$$n = p_1^{v_1} \cdots p_k^{v_k}$$

von verschiedenen Primzahlen p_1, \dots, p_k darstellen. Diese Darstellung ist eindeutig bis auf die Reihenfolge und das Multiplizieren einiger p_i mit einem Faktor -1 ; so ist etwa

$$60 = 2^2 \cdot 3 \cdot 5 = (-5) \cdot 2^2 \cdot (-3).$$

Die Vielfachen einer Zahl a bilden die Menge $a\mathbb{Z}$, und ist n ein Vielfaches sowohl von a als auch von b , so gilt

$$n \in a\mathbb{Z} \cap b\mathbb{Z}.$$

Insbesondere gilt mit der obigen Primfaktorzerlegung:

$$n \in p_1^{v_1} \mathbb{Z} \cap \dots \cap p_k^{v_k} \mathbb{Z}.$$

Gilt $|a| > |b|$ für zwei nicht-invertierbare Elemente $a, b \in \mathbb{Z} \setminus \{\pm 1\}$, so können wir Division mit Rest durchführen. Dies liefert

$$a = q \cdot b + r,$$

wobei für den Rest r stets $|r| < |b|$ gilt. In der Sprache der Kongruenzen bedeutet dies

$$a \equiv r \pmod{b} \quad \text{bzw.} \quad \bar{a} = \bar{r} \in \mathbb{Z}/b\mathbb{Z}$$

und a ist genau dann durch b teilbar, wenn $r = 0$ gilt. Durch wiederholtes Anwenden der Division mit Rest (diesmal auf b und r) können wir den größten gemeinsamen Teiler von a und b ermitteln (Algorithmus 1.28).

Im Ring \mathbb{Z} lässt es sich also sehr angenehm rechnen. In diesem Kapitel wollen wir uns mit Ringen R beschäftigen, die ähnlich gutartige Eigenschaften haben wie \mathbb{Z} . Dazu werden wir alle der oben angeführten Eigenschaften von \mathbb{Z} für abstrakte Ringe verallgemeinern und dann die Klasse derjenigen Ringe studieren, die diese verallgemeinerten Eigenschaften besitzen (die wir als *euklidische Ringe* kennenlernen werden). Besonders interessant ist dabei der Ring $\mathbb{K}[X]$ der Polynome über einem Körper \mathbb{K} .

Zunächst untersuchen wir die multiplikativ invertierbaren Elemente von R .

Definition 1.1 Es sei R ein Ring mit Eins. Ein Element $x \in R$ heißt **Einheit**, falls ein Element $x' \in R$ existiert, so dass gilt:

$$x \cdot x' = 1 = x' \cdot x.$$

Wir schreiben x^{-1} für x' . Die Menge aller Einheiten von R heißt **Einheitengruppe** und wird mit R^\times bezeichnet.

Hilfssatz 1.2 Die Einheitengruppe R^\times bildet mit der Multiplikation von R als Verknüpfung eine Gruppe mit neutralem Element 1.

Beweis: Die Multiplikation in R ist assoziativ, da R ein Ring ist. Es ist $1 \in R^\times$, da $1 \cdot 1 = 1$. Nach Definition von R^\times existiert für jedes $x \in R$ ein inverses Element $x^{-1} \in R^\times$. Für $x, y \in R^\times$ gilt (wegen der Assoziativität)

$$(xy)(y^{-1}x^{-1}) = 1 = (y^{-1}x^{-1})(xy),$$

also $(xy)^{-1} = y^{-1}x^{-1}$. Somit ist R^\times unter der Multiplikation abgeschlossen. ■

Beispiel 1.3 (Einheitengruppen)

- (a) Für $R = \mathbb{Z}$ ist

$$\mathbb{Z}^\times = \{-1, 1\}.$$

- (b) Es sei $R = \mathbb{Z}/6\mathbb{Z}$. Damit eine Klasse $\bar{n} = n + 6\mathbb{Z}$ eine Einheit in $\mathbb{Z}/6\mathbb{Z}$ ist, muss ein $m \in \mathbb{Z}$ existieren, so dass gilt:

$$\bar{n} \cdot \bar{m} = \bar{1},$$

was äquivalent ist zu

$$n \cdot m = q \cdot 6 + 1$$

für ein geeignetes $q \in \mathbb{Z}$. Von den Zahlen von 1 bis 6 erfüllen dies 1 und 5:

$$\begin{aligned} 1 \cdot 1 &= 0 \cdot 6 + 1, \\ 5 \cdot 5 &= 4 \cdot 6 + 1. \end{aligned}$$

Die übrigen Zahlen $n = 2, 3, 4, 6$ haben jeweils einen Primfaktor p (entweder 2 oder 3) mit 6 gemeinsam. Damit ist $n \cdot m = q \cdot 6 + 1$ äquivalent zu

$$\underbrace{n \cdot m - q \cdot 6}_{\in p\mathbb{Z}} = 1 \notin p\mathbb{Z},$$

ein Widerspruch. Also ist

$$(\mathbb{Z}/6\mathbb{Z})^\times = \{\bar{1}, \bar{5}\}.$$

- (c) Für einen Körper \mathbb{K} sind nach Definition alle Elemente $x \in \mathbb{K} \setminus \{0\}$ invertierbar. Somit ist

$$\mathbb{K}^\times = \mathbb{K} \setminus \{0\}.$$

- (d) Im Polynomring $\mathbb{K}[X]$ über dem Körper \mathbb{K} sind die invertierbaren Elemente genau die konstanten Polynome $\neq 0$, also

$$\mathbb{K}[X]^\times = \mathbb{K}^\times.$$

- (e) Im Ring $\mathbb{K}^{n \times n}$ der $n \times n$ -Matrizen über \mathbb{K} ist die Einheitengruppe (nach Definition) die allgemeine lineare Gruppe,

$$(\mathbb{K}^{n \times n})^\times = \mathbf{GL}_n(\mathbb{K}).$$

Wir überlegen nun, welche Eigenschaften ein Ring besitzen muss, um ähnliche Teilbarkeitseigenschaften wie \mathbb{Z} zu haben. Zunächst verlangen wir dazu, dass der Ring R kommutativ sei, denn sonst würde aus

$$a \cdot b = c \text{ (für ein geeignetes } b)$$

nicht automatisch folgen, dass auch

$$b' \cdot a = c \text{ (für ein geeignetes } b')$$

gilt. Wir müssten im nicht-kommutativen Fall also eine lästige Unterscheidung zwischen „Linksteilern“ und „Rechtsteilern“ machen. Außerdem gilt für alle $a \in \mathbb{Z}$ die Kürzungsregel

$$\text{„aus } a \neq 0 \text{ und } a \cdot b = a \cdot c \text{ folgt } b = c“,$$

die wir auch im allgemeineren Fall behalten wollen. Daher verlangen wir, dass der Ring R nullteilerfrei sei, denn daraus folgt sofort diese Kürzungsregel.

Definition 1.4 Ein kommutativer nullteilerfreier Ring mit Eins heißt **integer** (oder **Integritätsbereich**).

Beispiel 1.5 Die Ringe \mathbb{Z} und $\mathbb{K}[X]$ sind integer. Falls n keine Primzahl ist, besitzen die Ringe $\mathbb{Z}/n\mathbb{Z}$ Nullteiler und sind somit nicht integer. Da sie aber durch Quotientenbildung aus dem integren Ring \mathbb{Z} konstruiert wurden, können viele Eigenschaften von $\mathbb{Z}/n\mathbb{Z}$ aus den Teilbarkeitseigenschaften in \mathbb{Z} abgeleitet werden (wie etwa in Beispiel 1.3 (b) geschehen).

Definition 1.6 Es sei R ein integrer Ring.

- (a) Es seien $x, y \in R$. Wir sagen, x teilt y (geschrieben $x \mid y$), falls ein $q \in R$ existiert mit $q \cdot x = y$.
- (b) Es seien $x_1, \dots, x_k \in R$. Ein Element $g \in R$ heißt **größter gemeinsamer Teiler** von x_1, \dots, x_k (geschrieben $g = \text{ggT}(x_1, \dots, x_k)$), falls gilt
 - g ist ein Teiler der Elemente x_1, \dots, x_k ,
 - ist $d \in R$ ein weiterer Teiler der Elemente x_1, \dots, x_k , so ist d auch ein Teiler von g .
- (c) Gilt $\text{ggT}(x_1, \dots, x_k) = 1$, so heißen x_1, \dots, x_k **teilerfremd** (oder **coprim**).

Bemerkung 1.7 Der ggT ist nur eindeutig bestimmt bis auf Multiplikation mit einer Einheit. Die Schreibweise $\text{ggT}(x, y) = g$ bedeutet also, dass für jedes $u \in R^\times$ das Element $u \cdot g$ ebenfalls ein ggT von x und y ist.

Konvention: Im Fall $R = \mathbb{Z}$ bezeichnen wir mit $\text{ggT}(a, b)$ stets den *positiven* größten gemeinsamen Teiler von a und b ; im Fall $R = \mathbb{K}[X]$ bezeichnen wir mit $\text{ggT}(a, b)$ stets das *normierte* Polynom, das größter gemeinsamer Teiler von a und b ist. Dadurch wird die Schreibweise eindeutig.

Bemerkung 1.8 Für alle $x \in R$ gilt $1 \mid x$ und $x \mid 0$. Andererseits gilt für $x \neq 0$ immer $0 \nmid x$.

Aufgabe 1.9 Die Teilbarkeitsrelation $|$ ist eine *Quasiordnung* auf R , d.h. sie ist reflexiv und transitiv.

Aufgabe 1.10 Aus $x \mid y$ und $x \mid z$ folgt $x \mid y + z$.

Bemerkung 1.11 Es gilt

$$\text{ggT}(x_1, x_2, x_3) = \text{ggT}(\text{ggT}(x_1, x_2), x_3),$$

denn für jeden Teiler d von $\text{ggT}(x_1, x_2)$ und x_3 gilt insbesondere $d \mid x_1$, $d \mid x_2$, und somit $d \mid \text{ggT}(x_1, x_2, x_3)$ nach Definition des größten gemeinsamen Teilers.

Hilfssatz 1.12 (Kürzungsregel) Es sei R ein integrer Ring und $a, x, y \in R$, $a \neq 0$. Wenn $a \cdot x = a \cdot y$ gilt, dann gilt bereits $x = y$.

BEWEIS: Die Bedingung $a \cdot x = a \cdot y$ ist äquivalent zu $a \cdot (x - y) = 0$. Da $a \neq 0$ und R nullteilerfrei ist, muss $x - y = 0$ gelten, also $x = y$. ■

Definition 1.13 Es sei R ein kommutativer Ring mit Eins. Eine Teilmenge $\mathfrak{I} \subset R$ heißt **Ideal** in R , wenn gilt:

- (i) \mathfrak{I} ist eine additive Untergruppe von R , d.h. $0 \in \mathfrak{I}$ und aus $x, y \in \mathfrak{I}$ folgt $x + y \in \mathfrak{I}$ und $-x \in \mathfrak{I}$.
- (ii) Für alle $r \in R$ und $x \in \mathfrak{I}$ gilt $r \cdot x \in \mathfrak{I}$.

Man beachte, dass Ideale in der Regel *keine* Teilringe von R sind, da wir nicht verlangen, dass $1 \in \mathfrak{I}$ gilt (in der Tat ist R selbst das einzige Ideal, dass 1 enthält).

Definition 1.14 Für $x \in R$ nennen wir

$$\langle x \rangle = \{r \cdot x \mid r \in R\}$$

das von x erzeugte **Hauptideal** in R . Allgemeiner schreiben wir

$$\langle x_1, \dots, x_n \rangle = \{r_1 \cdot x_1 + \dots + r_n \cdot x_n \mid r_1, \dots, r_n \in R\}$$

für das von den Elementen $x_1, \dots, x_n \in R$ erzeugte Ideal in R .

Beispiel 1.15 (Ideale)

- (a) In jedem Ring R sind R selbst und $\{0\}$ Ideale, die *trivialen Ideale*.
- (b) Im Ring \mathbb{Z} der ganzen Zahlen ist für jedes $n \in \mathbb{Z}$ die Menge $n\mathbb{Z}$ ein Ideal, denn für $a, b \in \mathbb{Z}$ gilt $an + bn = (a + b)n \in n\mathbb{Z}$ und $a(bn) = (ab)n \in n\mathbb{Z}$. Das Ideal $n\mathbb{Z}$ ist ein Hauptideal mit Erzeuger n .
- (c) Es sei $R = \mathbb{K}[X]$. Für $\alpha \in \mathbb{K}$

$$\langle X - \alpha \rangle = \{h \cdot (X - \alpha) \mid h \in \mathbb{K}[X]\}$$

das Ideal aller Polynome, die α als Nullstelle haben.

- (d) Allgemeiner erzeugt jedes Polynom $f \in \mathbb{K}[X]$ ein Ideal $\langle f \rangle$ in $\mathbb{K}[X]$, das alle Vielfachen von f enthält. Ist $\deg(f) = 0$, also $f \in \mathbb{K} \setminus \{0\}$, so ist $\langle f \rangle = \mathbb{K}[X]$. Andernfalls ist $\langle f \rangle$ ein echtes Ideal in $\mathbb{K}[X]$.

Der von f in $\mathbb{K}[X]$ erzeugte Untervektorraum $[f]$ ist eine echte Teilmenge des Ideals $\langle f \rangle$:

$$[f] = \{\lambda \cdot f \mid \lambda \in \mathbb{K}\} \subsetneq \{h \cdot f \mid h \in \mathbb{K}[X]\} = \langle f \rangle.$$

- (e) Es sei $\Phi : V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen \mathbb{K} -Vektorraumes V . Dann ist

$$\mathfrak{J}_\Phi = \{f \in \mathbb{K}[X] \mid f(\Phi) = 0\}$$

ein Ideal im Polynomring $\mathbb{K}[X]$, nämlich der Kern des Einsetzungshomomorphismus

$$\mathbb{K}[X] \rightarrow \text{End}(V), \quad f \mapsto f(\Phi).$$

Nach dem Satz von Cayley-Hamilton ist das charakteristische Polynom f_Φ von Φ ein Element von \mathfrak{J}_Φ . Aus Satz 1.33 folgt, dass \mathfrak{J}_Φ ein Hauptideal ist. Der normierte Erzeuger von \mathfrak{J}_Φ heißt **Minimalpolynom** von Φ .

Hilfssatz 1.16 Es sei $\Phi : R_1 \rightarrow R_2$ ein Homomorphismus von Ringen. Dann ist $\text{Kern } \Phi = \{x \in R_1 \mid \Phi(x) = 0\}$ ein Ideal in R_1 .

Beweis: Es seien $r \in R_1$ und $x, y \in \text{Kern } \Phi$. Da Φ ein Homomorphismus ist, gilt

$$\begin{aligned} \Phi(x + y) &= \Phi(x) + \Phi(y) = 0 + 0 = 0, \\ \Phi(r \cdot x) &= \Phi(r) \cdot \Phi(x) = \Phi(r) \cdot 0 = 0. \end{aligned}$$

Somit $x + y, r \cdot x \in \text{Kern } \Phi$, d.h. $\text{Kern } \Phi$ ist ein Ideal. ■

Aufgabe 1.17 Ein Ideal \mathfrak{I} enthält genau dann eine Einheit $u \in R^\times$, wenn $\mathfrak{I} = R$ gilt.

Aufgabe 1.18 Ist \mathbb{K} ein Körper, so sind die einzigen Ideale in \mathbb{K} die trivialen.

Teilbarkeit lässt sich auch mit Hilfe der Ideale eines Ringes charakterisieren:

Satz 1.19 Es sei R ein integrer Ring und $x, y \in R$.

- (a) Es gilt $x | y$ genau dann, wenn $\langle y \rangle \subset \langle x \rangle$.
- (b) Es gilt $\langle x \rangle = \langle y \rangle$ genau dann, wenn $x = u \cdot y$ für eine Einheit $u \in R^\times$.
- (c) Falls $\langle x, y \rangle = \langle g \rangle$, so ist $g = \text{ggT}(x, y)$.
- (d) Falls $\langle x, y \rangle = R$, so sind x, y teilerfremd.

BEWEIS:

- (a) Es gilt $y = r \cdot x$ für ein $r \in R$ genau dann, wenn y ein Element von $\langle x \rangle$ ist.
- (b) Gilt $x = u \cdot y$, so gilt auch $y = u^{-1} \cdot x$. Mit Teil (a) folgt nun $\langle x \rangle \subset \langle y \rangle$ und $\langle y \rangle \subset \langle x \rangle$, also $\langle x \rangle = \langle y \rangle$.
Gilt umgekehrt $r \cdot x = y$ und $s \cdot y = x$, so gilt $srx = x$. Kürzen von x liefert $sr = 1$, und somit sind $r, s \in R^\times$. Dann erfüllt $u = s$ die Gleichung $x = u \cdot y$.
- (c) Nach Voraussetzung gibt es $a, b \in R$ mit $g = ax + by$. Da $x, y \in \langle g \rangle$, ist g ein Teiler von x und y . Für jeden gemeinsamen Teiler d von x und y folgt daraus $d | g$. Somit ist $g = \text{ggT}(x, y)$.
- (d) Folgt aus der vorigen Aussage und Aufgabe 1.17. ■

Die entsprechenden Aussagen gelten für eine beliebige Anzahl von Elementen $x_1, \dots, x_k \in R$.

Als Primzahlen bezeichnet man oft diejenigen $p \in \mathbb{N}$, die „nur durch sich selbst und 1 teilbar sind“. Da wir den Ring $\mathbb{Z} \supset \mathbb{N}$ betrachten, müssen wir aber auch -1 noch als Teiler zulassen. Allgemeiner wollen wir für Primelemente alle Einheiten als Teiler zulassen, womit die Bedingung „prim“ bedeuten soll, dass p zu allen Elementen aus $R \setminus R^\times$ teilerfremd ist. Wir geben hier eine noch allgemeinere Definition, die aber (wie wir später sehen werden) in den für uns interessanten Fällen genau die obige Bedingung wiedergibt. Aus praktischen Gründen schließen wir dabei die Einheiten von vornherein aus der Definition aus.

Definition 1.20 Es sei R ein integrer Ring und $p \in R \setminus R^\times$, $p \neq 0$.

- (a) p heißt **irreduzibel**, falls aus einer Zerlegung $p = x \cdot y$ stets folgt, dass eines der beiden Elemente x oder y eine Einheit ist.
- (b) p heißt **prim** (oder **Primelement**), falls aus $p \mid x \cdot y$ mit $x, y \in R$ stets $p \mid x$ oder $p \mid y$ folgt.

Hilfssatz 1.21 Ist $p \in R$ ein Primelement, so ist p auch irreduzibel.

BEWEIS: Es sei $p = x \cdot y$, insbesondere ist p dann ein Teiler von $x \cdot y$. Da p prim ist, gibt es ein $a \in R$, so dass (ohne Einschränkung) gelte: $p \cdot a = x$. Dann gilt

$$p = x \cdot y = p \cdot a \cdot y$$

und wegen der Kürzungsregel gilt dann

$$1 = a \cdot y.$$

Also ist $y \in R^\times$. Somit ist p irreduzibel. ■

Beispiel 1.22 (Primelemente)

- (a) Die Primelemente in \mathbb{Z} sind genau die Elemente $\pm p$, wobei p die Menge der Primzahlen $\{2, 3, 5, 7, 11, \dots\}$ durchläuft.
- (b) In $\mathbb{C}[X]$ sind die Primelemente genau die Polynome

$$p = a_1 X + a_0,$$

wobei $a_1 \in \mathbb{C}^\times$, $a_0 \in \mathbb{C}$. Dies ist eine Folge des Fundamentalsatzes der Algebra und wird in Abschnitt 1.3 wieder aufgegriffen.¹⁾

- (c) In $\mathbb{R}[X]$ sind die irreduziblen Polynome von der Form

$$p = a_1 X + a_0 \quad \text{oder} \quad q = b_2 X^2 + b_1 X + b_0,$$

wobei $a_1, b_2 \in \mathbb{R}^\times$, $a_0, b_1, b_0 \in \mathbb{R}$, so dass das Polynom q keine reelle Nullstelle hat. Beispielsweise ist das Polynom $X^2 + 1$ irreduzibel in $\mathbb{R}[X]$, da es nur die Nullstellen i und $-i$ in \mathbb{C} hat.

¹⁾Bei Polynomen ist es üblich, den Begriff *irreduzibel* anstelle von *prim* zu verwenden. Die beiden Begriffe stimmen in diesem Fall nach Satz 1.33 überein.

- (d) In Polynomringen über anderen Körpern lassen sich die irreduziblen Polynome in der Regel nicht so einfach charakterisieren. Die Polynome $a_1X + a_0$ vom Grad 1 sind immer irreduzibel, aber im Allgemeinen gibt es noch weitere. Es gibt eine Reihe von Kriterien, mit denen man auf Irreduzibilität prüfen kann, siehe Kapitel 2.8 in Bosch [1]. Beispielsweise ist für jede Primzahl $p \in \mathbb{N}$ das Polynom $X^{p-1} + X^{p-2} + \dots + X + 1$ irreduzibel in $\mathbb{Q}[X]$, und das Polynom $X^p - X - 1$ ist irreduzibel in $\mathbb{F}_p[X]$.

Idealtheoretisch lässt sich die Primeigenschaft so formulieren:

Definition 1.23 Ein Ideal \mathfrak{P} in einem Ring R heißt **Primideal**, wenn für alle $x, y \in R$ aus $x \cdot y \in \mathfrak{P}$ stets $x \in \mathfrak{P}$ oder $y \in \mathfrak{P}$ folgt.

Hilfssatz 1.24 Sei R ein integrer Ring und $p \in R$ prim. Dann ist das Hauptideal $\langle p \rangle$ ein Primideal.

BEWEIS: $x \cdot y \in \langle p \rangle$ bedeutet $p \mid x \cdot y$. Also $p \mid x$ oder $p \mid y$, was wiederum bedeutet $x \in \langle p \rangle$ oder $y \in \langle p \rangle$. ■

Um weitere Eigenschaften der ganzen Zahlen \mathbb{Z} (wie z.B. die eindeutige Primfaktorzerlegung oder die Möglichkeit, den ggT zu berechnen) verallgemeinern zu können, müssen wir die Klasse der Ringe, die wir betrachten, noch weiter spezialisieren. Insbesondere wollen wir eine Division mit Rest durchführen können. Diese Ringe sind Gegenstand des nächsten Abschnitts.

1.2 Euklidische Ringe

Definition 1.25 Ein integrer Ring R heißt **euklidischer Ring**, falls es eine Funktion $\delta : R \rightarrow \mathbb{N}_0$ gibt, so dass Folgendes gilt: Für alle $a, b \in R$ gibt es eine Darstellung

$$a = q \cdot b + r \tag{1.1}$$

mit $q, r \in R$, wobei $r = 0$ oder $\delta(r) < \delta(b)$.

In anderen Worten: In euklidischen Ringen ist Teilen mit Rest möglich. Falls die Division nicht aufgeht, ist der Rest r kleiner (bzgl. δ) als der Divisor b .

Beispiel 1.26 (Euklidische Ringe)

- (a) Der Ring \mathbb{Z} ist ein euklidischer Ring mit $\delta(n) = |n|$. Für die ganzzahlige Division schreiben wir $q = a \div b$, mit q, a, b wie in (1.1).

- (b) Der Polynomring $\mathbb{K}[X]$ ist ein euklidischer Ring mit $\delta(f) = \deg(f)$ für $f \neq 0$ und $\delta(0) = 0$. Dies sieht man mit Hilfe der *Polynomdivision*: Es seien $f, g \in \mathbb{K}[X] \setminus \{0\}$ und

$$f = a_m X^m + \dots + a_1 X + a_0, \quad g = b_n X^n + \dots + b_1 X + b_0.$$

mit $a_m, b_n \neq 0$, wobei wir $m \geq n$ annehmen (also $\deg(f) \geq \deg(g)$). Dann setze

$$q_0 = \frac{a_m}{b_n} X^{m-n}$$

und erhalte (nach Ausmultiplizieren und Zusammenfassen nach Potenzen)

$$f_1 = f - q_0 \cdot g = \underbrace{\left(a_m - \frac{a_m}{b_n} b_n \right)}_{=0} X^m + \left(a_{m-1} - \frac{a_m}{b_n} b_{n-1} \right) X^{m-1} + \dots$$

Falls $f_1 = 0$, so erfüllen $r = f_1$ und $q = q_0$ die Bedingung (1.1). Falls $f_1 \neq 0$, so ist $\deg(f_1) \leq m-1 < \deg(f)$. Mit Induktion über $m = \deg(f)$ dürfen wir annehmen, dass es eine Zerlegung $f_1 = q_1 \cdot g + r$ gibt, so dass $\deg(r) < \deg(g)$ erfüllt ist. Nun ist

$$f = (q_1 + q_0) \cdot g + r$$

die gesuchte Zerlegung (1.1) für f mit $\deg(r) < \deg(g)$ und $q = q_1 + q_0$.

Beispiel 1.27 Es seien $f = 2X^4 + X^2 - 3X + 2, g = -X^2 + X - 1 \in \mathbb{R}[X]$. Wir bestimmen mit Hilfe der Polynomdivision wie in Beispiel 1.26 die Elemente $q, r \in \mathbb{R}[X]$, so dass $f = q \cdot g + r$:

- Setze $q_0 = -2X^2$. Dann ist $f_1 = f - q_0 \cdot g = 2X^3 - X^2 - 3X + 2$.
- Setze $q_1 = -2X$. Dann ist $f_2 = f_1 - q_1 \cdot g = X^2 - 5X + 2$.
- Setze $q_2 = -1$. Dann ist $f_3 = f_2 - q_2 \cdot g = -4X + 1$. Nun ist $\deg(f_3) < \deg(g)$ und für $q = q_0 + q_1 + q_2 = -2X^2 - 2X - 1$ gilt

$$\begin{aligned} -4X + 1 &= f_3 = f_2 - q_2 \cdot g \\ &= (f_1 - q_1 \cdot g) - q_2 \cdot g = f_1 - (q_1 + q_2) \cdot g \\ &= (f - q_0 \cdot g) - (q_1 + q_2) \cdot g \\ &= f - (q_0 + q_1 + q_2) \cdot g \\ &= f - q \cdot g. \end{aligned}$$

Der Divisionsrest ist somit $r = f_3 = -4X + 1$.

Algorithmus 1.28 (Euklidischer Algorithmus) Es sei R ein euklidischer Ring und $a_0, a_1 \in R$. Teile solange mit Rest,

$$a_{i-1} = q_i \cdot a_i + a_{i+1} \quad (\text{mit } \delta(a_i) > \delta(a_{i+1})) \text{ für } i = 1, \dots, n, \quad (1.2)$$

bis nach endlich vielen Schritten der Rest $a_{n+1} = 0$ bleibt. Dann ist $a_n = \text{ggT}(a_0, a_1)$.

Beweis: Da die Folge der $\delta(a_i)$ streng monoton fällt und ≥ 0 ist, endet der Algorithmus nach endlich vielen Schritten.

Der Algorithmus hat die Schleifeninvariante $\text{ggT}(a_{i-1}, a_i) = \text{ggT}(a_i, a_{i+1})$ für $i = 1, \dots, n$: Ist $g = \text{ggT}(a_i, a_{i+1})$, so gilt $d \mid g$ für jeden gemeinsamen Teiler $d \in R$ von a_i und a_{i+1} . Nach der Definition von a_{i+1} in (1.2) ist d genau dann ein Teiler von a_i und a_{i+1} , wenn d ein Teiler von a_i und a_{i-1} ist. Also gilt auch $d \mid g$ für jeden gemeinsamen Teiler von a_i und a_{i-1} , d.h. $g = \text{ggT}(a_{i-1}, a_i)$.

Im letzten Schritt des Algorithmus gilt nun $\text{ggT}(a_0, a_1) = \text{ggT}(a_1, a_2) = \dots = \text{ggT}(a_n, a_{n+1}) = \text{ggT}(a_n, 0) = a_n$. ■

Folgerung 1.29 (Lemma von Bézout) Es sei R ein euklidischer Ring und $a, b \in R$. Es gibt $s, t \in R$, so dass

$$\text{ggT}(a, b) = sa + tb. \quad (1.3)$$

Beweis: Zum Beweis verwenden wir eine *erweiterte Variante des euklidischen Algorithmus* (mit Notation wie im Beweis von Algorithmus 1.28 und $a_0 = a$, $a_1 = b$): Der Algorithmus liefert eine Folge von Resten

$$a_{i+1} = a_{i-1} - q_i \cdot a_i, \quad i = 1, \dots, n$$

wobei $a_n = \text{ggT}(a_0, a_1)$ und $a_{n+1} = 0$. Dies lässt sich durch Matrizen ausdrücken:

$$\begin{pmatrix} a_i \\ a_{i+1} \end{pmatrix} = Q_i \cdot \begin{pmatrix} a_{i-1} \\ a_i \end{pmatrix} \quad \text{mit } Q_i = \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix}.$$

In der erweiterten Form des euklidischen Algorithmus berechnet man in jedem Schritt die Matrix $A_i = Q_i \cdot A_{i-1}$, wobei $A_0 = I_2$. Wenn der Algorithmus nach n Schritten endet, gilt

$$\begin{pmatrix} \text{ggT}(a_0, a_1) \\ 0 \end{pmatrix} = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = Q_n \cdots Q_1 \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = A_n \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}.$$

Die gesuchten Elemente s, t sind also die Einträge in der ersten Zeile der Matrix $A_n = \begin{pmatrix} s & t \\ s' & t' \end{pmatrix}$. ■

Bemerkung 1.30 Obwohl der euklidische Algorithmus für $R = \mathbb{Z}$ einer der ältesten und einfachsten Algorithmen überhaupt ist, gestaltet sich die Analyse seines Aufwands keineswegs einfach. Der Aufwand des Algorithmus wird dadurch bestimmt, wie oft die Division mit Rest (1.2) durchgeführt werden muss. Nehmen wir $0 \leq a_1 \leq a_0$ an. Für die Folge der Divisionsreste in (1.2) gilt

$$a_1 > a_2 > \dots > a_n > 0.$$

Für die Quotienten $q_i = a_{i-1} \div a_i$ der ganzzahligen Divisionen in (1.2) gilt demnach

$$q_i \geq 1 \text{ für } 1 \leq i \leq n-1, \quad q_n \geq 2. \quad (*)$$

Für $1 \leq i \leq n-1$ liefert dies

$$\begin{aligned} a_{i-1} &= a_i q_i + a_{i+1} \\ &> a_{i+1} q_i + a_{i+1} \\ &= a_{i+1}(q_i + 1) \end{aligned}$$

und folglich

$$\prod_{i=1}^{n-1} a_{i-1} > \prod_{i=1}^{n-1} a_{i+1}(q_i + 1).$$

Die Faktoren a_2, a_3, \dots, a_{n-2} tauchen auf beiden Seiten auf und können daher rausgekürzt werden:

$$\begin{aligned} a_0 a_1 &> a_{n-1} a_n \prod_{i=1}^{n-1} (q_i + 1) \\ &= a_n^2 q_n \prod_{i=1}^{n-1} (q_i + 1). \end{aligned}$$

Mit (*) und $a_0 \geq a_1$ können wir dies weiter abschätzen zu

$$\begin{aligned} a_0^2 &\geq a_0 a_1 \\ &> a_n^2 \cdot 2 \cdot 2^{n-1} \\ &= 2^n \operatorname{ggT}(a_0, a_1)^2 \\ &\geq 2^n. \end{aligned}$$

Es folgt

$$2 \log(a_0) > n.$$

Der Anzahl der Schleifendurchläufe des Algorithmus liegt also in $\mathcal{O}(\log(m))$ für $m = \max\{|a_0|, |a_1|\}$. Der ungünstigste Fall tritt für $\operatorname{ggT}(f_{n+1}, f_n)$ ein, wobei f_n die

nte Fibonacci-Zahl ist (definiert durch $f_{n+1} = f_n + f_{n-1}$ mit $f_1 = 1, f_0 = 0$). In diesem Fall benötigt der Algorithmus für $n > 1$ genau $n-1$ Schleifendurchläufe. Eine noch genauere Analyse ist in Knuth [8], Abschnitt 4.5.3, zu finden. Für Polynome liegt der Aufwand des Algorithmus in $\mathcal{O}(m^2)$, wobei $m = \max\{\deg(a_0), \deg(a_1)\}$, siehe Lipson [11], Abschnitt VII.2.2.

Bemerkung 1.31 Nach Bemerkung 1.11 kann der (erweiterte) euklidische Algorithmus auch verwendet werden, um den ggT von beliebig vielen Elementen $a_1, \dots, a_k \in R$ zu berechnen und Elemente $s_1, \dots, s_k \in R$ zu bestimmen, so dass gilt

$$\text{ggT}(a_1, \dots, a_k) = s_1 a_1 + \dots + s_k a_k. \quad (1.4)$$

Beispiel 1.32 Wir demonstrieren den erweiterten euklidischen Algorithmus an zwei Beispielen:

(a) Es seien $a = 21, b = 8 \in \mathbb{Z}$.

- $a_0 = 21, a_1 = 8$ liefert $q_1 = 2, a_2 = 5$ und die Matrix $Q_1 = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$.
- $a_1 = 8, a_2 = 5$ liefert $q_2 = 1, a_3 = 3$ und die Matrix $Q_2 = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$.
- $a_2 = 5, a_3 = 3$ liefert $q_3 = 1, a_4 = 2$ und die Matrix $Q_3 = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$.
- $a_3 = 3, a_4 = 2$ liefert $q_4 = 1, a_5 = 1$ und die Matrix $Q_4 = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$.
- $a_4 = 2, a_5 = 1$ liefert $q_5 = 2, a_6 = 0$ und die Matrix $Q_5 = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$. Der Algorithmus endet in diesem Schritt, da $a_6 = 0$.

Somit ist

$$\text{ggT}(21, 8) = a_5 = 1$$

und

$$A_5 = Q_5 Q_4 Q_3 Q_2 Q_1 = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3 & 8 \\ 8 & -21 \end{pmatrix}.$$

Die Elemente in der erste Zeile von A_5 erfüllen

$$1 = (-3) \cdot 21 + 8 \cdot 8.$$

(b) Es seien $a = X^{13} + X^{12} - 2X^9, b = -X^6 - X^5 + X^4 + X^3 + 2X^2 - 2 \in \mathbb{R}[X]$ (die auftretenden Polynomdivisionen in den einzelnen Schritten führen wir nicht aus):

- $a_0 = X^{13} + X^{12} - 2X^9, a_1 = -X^6 - X^5 + X^4 + X^3 + 2X^2 - 2$ liefert $q_1 = -X^7 - X^5 - X^3 - X, a_2 = X^5 + X^4 - 2X$ und die Matrix $Q_1 = \begin{pmatrix} 0 & 1 \\ 1 & X^7 + X^5 + X^3 + X \end{pmatrix}$.

- $a_1 = -X^6 - X^5 + X^4 + X^3 + 2X^2 - 2, a_2 = X^5 + X^4 - 2X$ liefert $q_2 = -X, a_3 = X^4 + X^3 - 2$ und die Matrix $Q_2 = \begin{pmatrix} 0 & 1 \\ 1 & X \end{pmatrix}$.
- $a_2 = X^5 + X^4 - 2X, a_3 = X^4 + X^3 - 2$ liefert $q_3 = X, a_4 = 0$ und die Matrix $Q_3 = \begin{pmatrix} 0 & 1 \\ 1 & -X \end{pmatrix}$. Der Algorithmus endet in diesem Schritt, da $a_4 = 0$.

Somit ist

$$\begin{aligned} \text{ggT}(X^{13} + X^{12} - 2X^9, -X^6 - X^5 + X^4 + X^3 + 2X^2 - 2) \\ = a_3 = X^4 + X^3 - 2 \end{aligned}$$

und

$$\begin{aligned} A_3 = Q_3 Q_2 Q_1 &= \begin{pmatrix} 0 & 1 \\ 1 & -X \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & X \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & X^7 + X^5 + X^3 + X \end{pmatrix} \\ &= \begin{pmatrix} X & X^8 + X^6 + X^4 + X^2 + 1 \\ 1 - X^2 & -X^9 \end{pmatrix}. \end{aligned}$$

Die Elemente in der ersten Zeile von A_3 erfüllen

$$X^4 + X^3 - 2 = X \cdot (X^{13} + X^{12} - 2X^9) + (X^8 + X^6 + X^4 + X^2 + 1) \cdot (-X^6 - X^5 + X^4 + X^3 + 2X^2 - 2).$$

Satz 1.33 Es sei R ein euklidischer Ring.

- Jedes Ideal \mathfrak{J} in R ist ein Hauptideal (d.h. es gibt $x \in R$, so dass $\mathfrak{J} = \langle x \rangle$).
- Ein Element $x \in R$ ist genau dann irreduzibel, wenn es ein Primelement ist.

Beweis:

- Es sei $x \in \mathfrak{J}, x \neq 0$, so dass $\delta(x)$ minimal ist unter allen Elementen aus \mathfrak{J} (das ist möglich, da δ nur Werte aus \mathbb{N}_0 annimmt). Da R euklidisch ist, gibt es für jedes $y \in \mathfrak{J}$ Elemente $q, r \in R$ mit $y = qx + r$, wobei $\delta(r) < \delta(x)$ oder $r = 0$. Nun ist aber $r = y - qx \in \mathfrak{J}$, also $r = 0$ wegen der Minimalität von x . Somit $y \in \langle x \rangle$, und da y beliebig gewählt war, folgt $\mathfrak{J} = \langle x \rangle$.
- Ein Primelement ist irreduzibel nach Hilfssatz 1.21.

Nun sei p irreduzibel und es gelte $p \mid xy$ für gewisse $x, y \in R$. Angenommen, $p \nmid y$. Wir müssen zeigen, dass $p \mid x$ gilt. Es sei $g = \text{ggT}(p, y)$. Da p irreduzibel und $g \mid p$, gilt $p = hg$ mit $g \in R^\times$ oder $h \in R^\times$. Wäre hier $h \in R^\times$, so wäre $h^{-1}p$ ein Teiler von y , und somit auch p ein Teiler von y , im Widerspruch zur Annahme. Also ist $g \in R^\times$. Mit dem erweiterten

euklidischen Algorithmus können wir nun Elemente $s, t \in R$ bestimmen, die

$$g = sp + ty$$

erfüllen. Da g eine Einheit ist, können wir dies mit xg^{-1} multiplizieren und erhalten

$$x = xg^{-1}sp + xg^{-1}ty = xg^{-1}sp + g^{-1}txy.$$

Nach Voraussetzung teilt p beide Summanden auf der rechten Seite, und somit $p \mid x$. Insgesamt folgt, dass $p \mid xy$ stets impliziert, dass p wenigstens eines der Elemente x oder y teilt. Somit ist p prim. ■

Satz 1.34 (Eindeutige Primfaktorzerlegung) *Sei R ein euklidischer Ring. Jedes Element $x \in R \setminus R^\times$, $x \neq 0$, hat eine Zerlegung*

$$x = p_1 \cdots p_k, \quad (1.5)$$

wobei die p_1, \dots, p_k Primelemente in R sind. Diese Zerlegung ist eindeutig bis auf die Reihenfolge und Multiplikation der p_i mit Einheiten aus R^\times .

BEWEIS: Zunächst zeigen wir die Existenz:

Sei $x \in R \setminus R^\times$, $x \neq 0$. Nach Satz 1.33 sind Primelemente in R das selbe wie irreduzible Elemente. Ist x selbst irreduzibel, so ist $x = p_1$. Andernfalls ist x ein Produkt $x = x_1 y_1$ mit $x_1, y_1 \notin R^\times$. Falls x_1, y_1 nicht beide irreduzibel sind, zerlege sie weiter, bis alle Faktoren irreduzibel sind und erhalte so die Primfaktorzerlegung. Dieses Verfahren bricht nach endlich vielen Schritten ab: Haben wir nämlich eine Folge von Elementen $x_i \in R \setminus R^\times$ mit der Eigenschaft $x_i \mid x_{i-1}$ für $i = 1, \dots, n$ und $x_0 = x$, so gilt nach Satz 1.19

$$\langle x_0 \rangle \subset \langle x_1 \rangle \subset \langle x_2 \rangle \subset \dots$$

Dann ist $\mathfrak{I} = \bigcup_{i=0}^{\infty} \langle x_i \rangle$ ein Ideal in R und nach Satz 1.33 sogar ein Hauptideal. Es gibt also ein $z \in R$ mit $\langle z \rangle = \mathfrak{I} = \bigcup_{i=0}^{\infty} \langle x_i \rangle$. Dann muss aber $z \in \langle x_m \rangle$ gelten für ein gewisses $m \geq 0$. Das bedeutet $\langle z \rangle = \langle x_i \rangle$ für alle $i \geq m$. Nach Satz 1.19 (b) unterscheiden sich die x_i für $i \geq m$ also nur um einen Faktor aus R^\times von x_m . Dies ergäbe einen Widerspruch, wenn wir das obige Verfahren beliebig lange fortsetzen könnten.

Nun zeigen wir die Eindeutigkeit:

Es sei $x = q_1 \cdots q_j$ eine weitere Primfaktorzerlegung der Form (1.5), wobei wir ohne Einschränkung $k \geq j$ annehmen. Wir zeigen die Behauptung durch Induktion über k : Falls $k = 1$, so ist x selbst prim, also auch $j = 1$ und $p_1 = q_1 = x$. Ist $k > 1$, so muss der Primfaktor p_k einen der Faktoren q_i teilen; nehmen wir ohne

Einschränkung an es sei q_j . Da q_j selbst prim (also irreduzibel) ist, unterscheiden sich q_j und p_k nur um einen Faktor $u \in R^\times$. Mit der Kürzungsregel (Hilfssatz 1.12) folgt

$$p_1 \cdots (u^{-1} p_{k-1}) = q_1 \cdots q_{j-1}.$$

Die linke Seite hat nun einen Primfaktor weniger als x , also können wir auf sie die Induktionsvoraussetzung anwenden. Das bedeutet $k - 1 = j - 1$ und $p_i = u_i q_{\sigma(i)}$ für $u_i \in R^\times$ und eine geeignete Permutation $\sigma \in S_{k-1}$. Zusammen mit $p_k = u q_k$ folgt die Behauptung. ■

Als Anwendung der eindeutigen Primfaktorzerlegung können wir nun einen der ältesten mathematischen Sätze beweisen:

Satz 1.35 (Euklid) *Es gibt unendlich viele Primzahlen in \mathbb{N} (bzw. in \mathbb{Z}).*

BEWEIS NACH EUKLID: Angenommen, es gäbe nur endlich viele positive Primzahlen p_1, \dots, p_n . Setze $k = 1 + p_1 \cdots p_n$. Für $i = 1, \dots, n$ gilt $k \equiv 1 \pmod{p_i}$, also ist k durch keines der p_i teilbar. Aber k ist auch keine der Einheiten ± 1 von \mathbb{Z} . Also muss die Primfaktorzerlegung von k noch weitere Primzahlen enthalten, die nicht zu p_1, \dots, p_n gehören, im Widerspruch zur Annahme. ■

Aufgabe 1.36 Ähnlich kann man zeigen, dass der Polynomring $\mathbb{K}[X]$ unendlich viele irreduzible Polynome enthält (dies ist bereits klar, falls \mathbb{K} kein endlicher Körper ist, da ja alle Polynome der Form $X + a$ mit $a \in \mathbb{K}$ irreduzibel sind).

Wir betrachten einen weiteren Beweis von Satz 1.35 von Euler, der Methoden aus der Analysis verwendet.

BEWEIS NACH EULER: Angenommen, es gäbe nur endlich viele positive Primzahlen $p_1 < p_2 < \dots < p_n$. Für $1 \leq i \leq n$ gilt mit der Formel für die geometrische Reihe

$$\frac{1}{1 - p_i^{-1}} = \sum_{k=0}^{\infty} p_i^{-k} \in \mathbb{R},$$

wobei die Reihe konvergiert, da $0 < |p_i^{-1}| < 1$. Folglich ist der folgende Ausdruck ein Produkt reeller Zahlen:

$$\begin{aligned} \frac{1}{1 - p_1^{-1}} \cdots \frac{1}{1 - p_n^{-1}} &= \left(\sum_{k_1=0}^{\infty} p_1^{-k_1} \right) \cdots \left(\sum_{k_n=0}^{\infty} p_n^{-k_n} \right) \\ &= \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} (p_1^{-k_1} \cdots p_n^{-k_n}) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \frac{1}{p_1^{k_1} \cdots p_n^{k_n}} \\ &= \sum_{m=1}^{\infty} \frac{1}{m}. \end{aligned}$$

Die letzte Gleichheit gilt, da als Folge von Satz 1.34 die Ausdrücke $p_1^{k_1} \cdots p_n^{k_n}$ sämtliche natürlichen Zahlen $m \in \mathbb{N}$ durchlaufen. Die Reihe $\sum_{m=1}^{\infty} \frac{1}{m}$ ist aber divergent, stellt also keine reelle Zahl dar. Aus diesem Widerspruch folgt, dass es unendlich viele Primzahlen geben muss. ■

1.3 Nullstellen von Polynomen

Im euklidischen Ring $R = \mathbb{K}[X]$ erhalten wir wichtige Zusammenhänge zwischen Teilbarkeit und Nullstellen von Polynomen:

Satz 1.37 Es sei $f \in \mathbb{K}[X]$ und $\alpha, \beta \in \mathbb{K}$. Dann gilt Folgendes:

- (a) Es sei $g \in \mathbb{K}[X]$, $g \neq 0$, mit $g(\alpha) = 0$. Ist $r = f - qg$ der Rest nach Polynomdivision, so gilt

$$r(\alpha) = f(\alpha).$$
- (b) $X - \alpha$ teilt f genau dann, wenn $f(\alpha) = 0$.
- (c) $X - \alpha, X - \beta$ sind teilerfremd genau dann, wenn $\alpha \neq \beta$.
- (d) Ist $n = \deg(f)$, so hat f höchstens n Nullstellen in \mathbb{K} .

BEWEIS:

- (a) Es ist $f(\alpha) = f(\alpha) - 0 = f(\alpha) - q(\alpha)g(\alpha) = r(\alpha)$.
- (b) In (a) mit $g = X - \alpha$ ist r eine Konstante, also $r = f(\alpha)$. Außerdem ist $X - \alpha$ genau dann ein Teiler von f , wenn der Divisionsrest $r = 0$ ist.
- (c) Der Divisonsrest bei Polynomdivision von $X - \alpha$ durch $X - \beta$ ist $\beta - \alpha$.
- (d) Jede Nullstelle λ von f liefert nach (b) einen irreduziblen Teiler $X - \lambda$ von f . Die Nullstellen sind eindeutig bestimmt, da die Primfaktorzerlegung eindeutig ist. Gäbe es $m > n$ Nullstellen, so hätte das Produkt von m Linearfaktoren den Grad $m > \deg(f)$, Widerspruch. ■

Wir erinnern an den Fundamentalsatz der Algebra, dessen Beweis im Rahmen der linearen Algebra nicht möglich ist. Er ist in Bosch [1], Abschnitt 6.3, zu finden.

Satz 1.38 (Fundamentalsatz der Algebra) Jedes Polynom in $\mathbb{C}[X]$ vom Grad ≥ 1 hat eine Nullstelle in \mathbb{C} .

Folgerung 1.39 Jedes Polynom $f \in \mathbb{C}[X]$ vom Grad $n \geq 1$ zerfällt in n Linearfaktoren,

$$f = \alpha \cdot (X - \lambda_1) \cdots (X - \lambda_n),$$

mit $\alpha \in \mathbb{C}^\times$, wobei die $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ nicht zwingend verschieden sein müssen.

BEWEIS: Nach dem Fundamentalsatz der Algebra hat f eine Nullstelle λ_1 , und somit nach Satz 1.37 (b) den Teiler $X - \lambda_1$. Es gibt also ein Polynom h vom Grad $n - 1$ mit $f = (X - \lambda_1) \cdot h$. Mit Induktion kann man nun schließen, dass h in $n - 1$ Linearfaktoren zerfällt, womit die Behauptung gezeigt ist. ■

1.4 Quotientenringe

Man kann sich vorstellen, dass $\mathbb{Z}/n\mathbb{Z}$ aus \mathbb{Z} entsteht, indem man in \mathbb{Z} die zusätzliche Rechenregel „ $n = 0$ “ einführt. Mit dieser neuen Rechenregel sind natürlich auch alle Vielfachen kn identisch 0, und Elemente $a, b \in \mathbb{Z}$, die sich nur um ein Vielfaches von n unterscheiden (d.h. $a = b + kn$ für ein geeignetes k), sind mit der neuen Rechenregel nicht mehr unterscheidbar. Dies drücken wir mit der Schreibweise

$$a \equiv b \pmod{n}$$

aus. Solche Elemente werden in der Restklasse $\bar{a} = a + n\mathbb{Z}$ zusammengefasst. Dann ist $\mathbb{Z}/n\mathbb{Z}$ die Menge all dieser Restklassen, und mit den von \mathbb{Z} induzierten Operationen $+$ und \cdot ist $\mathbb{Z}/n\mathbb{Z}$ ein Ring.

Diese Idee des „Einführens neuer Rechenregeln“ wollen wir nun auf beliebige Ringe verallgemeinern. Dazu überlegen wir, wie sich das obige Beispiel $\mathbb{Z}/n\mathbb{Z}$ mit den in den vorigen Kapiteln eingeführten Begriffen formulieren lässt: Es ist $n\mathbb{Z} = \langle n \rangle$, das von n erzeugte Ideal, und die kanonische Projektion $\pi : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$, $a \mapsto \bar{a}$ ist ein Ringhomomorphismus mit Kern $\pi = n\mathbb{Z}$. Die Bedingung $a \equiv b \pmod{n}$ ist äquivalent zu $a - b \in n\mathbb{Z}$.

Dies lässt sich auf einen beliebigen Ring R übertragen, indem man ein Ideal $\mathfrak{J} \subset R$ auswählt und diejenigen Elemente $a, b \in R$ als äquivalent auffasst, die $a - b \in \mathfrak{J}$ erfüllen. Die Idee hierbei ist, dass \mathfrak{J} genau diejenigen Elemente enthält, die nach dem Einführen neuer Rechenregeln auf 0 gesetzt worden sind. Daher bezeichnet man die Erzeuger von \mathfrak{J} oft als *Relationen* und \mathfrak{J} als das *Relationenideal*.

Satz 1.40 Es sei R ein kommutativer Ring und \mathfrak{J} ein Ideal in R . Dann ist die Menge

$$R/\mathfrak{J} = \{\bar{x} = x + \mathfrak{J} \mid x \in R\} \tag{1.6}$$

ebenfalls ein kommutativer Ring, wenn man $+$ und \cdot über Vertreter $x, y \in R$ definiert:

$$\begin{aligned}\bar{x} + \bar{y} &= \overline{x+y}, \\ \bar{x} \cdot \bar{y} &= \overline{x \cdot y}.\end{aligned}$$

Die kanonische Projektion $\pi : R \rightarrow R/\mathfrak{J}$, $x \mapsto \bar{x}$ ist ein surjektiver Homomorphismus von Ringen mit Kern $\pi = \mathfrak{J}$.

BEWEIS: Die Ringeigenschaften vererben sich ohne weiteres von R auf R/\mathfrak{J} . Es bleibt zu zeigen, dass die Definitionen der Verknüpfungen unabhängig von der Auswahl der Vertreter x, y (d.h. wohldefiniert) sind: Seien $x - x', y - y' \in \mathfrak{J}$. Es gilt

$$\bar{x} + \bar{y} = \overline{x+y} = \overline{(x-x') + (y-y') + x' + y'} = \underbrace{\overline{x-x'}}_{\in \mathfrak{J}} + \underbrace{\overline{y-y'}}_{\in \mathfrak{J}} + \overline{x'} + \overline{y'} = \overline{x'} + \overline{y'}.$$

Somit ist die Addition wohldefiniert. Für die Wohldefiniertheit der Multiplikation ist die Idealeigenschaft von \mathfrak{J} erforderlich:

$$\begin{aligned}\bar{x} \cdot \bar{y} &= \overline{x \cdot y} = \overline{((x-x') + x') \cdot ((y-y') + y')} \\ &= \underbrace{\overline{(x-x') \cdot (y-y')}}_{\in \mathfrak{J}} + \underbrace{\overline{(x-x') \cdot y'}}_{\in \mathfrak{J}} + \underbrace{\overline{x' \cdot (y-y')}}_{\in \mathfrak{J}} + \overline{x' \cdot y'} \\ &= \underbrace{\overline{(x-x') \cdot (y-y')}}_{\in \mathfrak{J}} + \overline{(x-x') \cdot y'} + \overline{x' \cdot (y-y')} + \overline{x' \cdot y'} \\ &= \overline{x' \cdot y'} = \overline{x'} \cdot \overline{y'}.\end{aligned}$$

Die Abbildung π ist nach Konstruktion ein surjektiver Homomorphismus und $\text{Kern } \pi = \{x \in R \mid \pi(x) = \bar{0}\} = \{x \in R \mid x + \mathfrak{J} = \mathfrak{J}\} = \mathfrak{J}$. ■

Definition 1.41 Der Ring R/\mathfrak{J} heißt **Quotientenring** von R über \mathfrak{J} .

Wir sprechen auch von „ R modulo \mathfrak{J} “ und schreiben für äquivalente $x, y \in R$

$$x \equiv y \pmod{\mathfrak{J}}.$$

Ist \mathfrak{J} ein Hauptideal mit Erzeuger f , so schreiben wir (analog zum Fall $\mathbb{Z}/n\mathbb{Z}$)

$$x \equiv y \pmod{f}$$

oder kürzer

$$x \equiv_f y.$$

Aufgabe 1.42 Ist \mathfrak{P} ein Primideal in R , so ist R/\mathfrak{P} nullteilerfrei.

Aufgabe 1.43 Ist R ein euklidischer Ring und \mathfrak{P} ein Primideal, so ist R/\mathfrak{P} ein Körper.

Satz 1.44 (Homomorphiesatz für Ringe) Es seien R, S kommutative Ringe und $\Phi : R \rightarrow S$ ein Ringhomomorphismus. Dann existiert ein eindeutig bestimmter Ringhomomorphismus $\bar{\Phi} : R/\text{Kern } \Phi \rightarrow S$, so dass gilt: $\bar{\Phi}$ ist injektiv und

$$\bar{\Phi} \circ \pi = \Phi, \quad (1.7)$$

d.h. das folgende Diagramm kommutiert:

$$\begin{array}{ccc} R & \xrightarrow{\Phi} & S \\ \pi \downarrow & \nearrow \bar{\Phi} & \\ R/\text{Kern } \Phi & & \end{array}$$

BEWEIS: Definiere $\bar{\Phi}$ durch

$$\bar{\Phi}(\bar{x}) = \Phi(x)$$

für $x \in R$. Dies ist die Gleichung (1.7).

- $\bar{\Phi}$ ist wohldefiniert, denn ist $x - x' \in \text{Kern } \Phi$, so gilt

$$\bar{\Phi}(\bar{x}) = \Phi(x) = \Phi(x - x' + x') = \underbrace{\Phi(x - x')}_{=0} + \Phi(x') = \Phi(x') = \bar{\Phi}(\bar{x'}).$$

- $\bar{\Phi}$ ist ein Homomorphismus von Ringen: Für $x, y \in R$ gilt

$$\bar{\Phi}(\bar{x} + \bar{y}) = \bar{\Phi}(\bar{x} + \bar{y}) = \Phi(x + y) = \Phi(x) + \Phi(y) = \bar{\Phi}(\bar{x}) + \bar{\Phi}(\bar{y})$$

und analog

$$\bar{\Phi}(\bar{x} \cdot \bar{y}) = \bar{\Phi}(\bar{x}) \cdot \bar{\Phi}(\bar{y}).$$

- $\bar{\Phi}$ ist injektiv: Es sei $\bar{x} \in \text{Kern } \bar{\Phi}$, d.h.

$$0 = \bar{\Phi}(\bar{x}) = \Phi(x).$$

Dann ist $x \in \text{Kern } \Phi$, also $\bar{x} = \text{Kern } \Phi = \bar{0}$. Somit ist $\text{Kern } \bar{\Phi} = \{\bar{0}\}$. ■

Folgerung 1.45 (Isomorphiesatz für Ringe) Ist Φ in Satz 1.44 surjektiv, so ist $\bar{\Phi}$ ein Isomorphismus. Insbesondere gilt

$$R/\text{Kern } \Phi \cong S. \quad (1.8)$$

BEWEIS: $\bar{\Phi}$ ist injektiv, und da auch $\text{Bild } \bar{\Phi} = \text{Bild } \Phi = S$ gilt, ist $\bar{\Phi}$ bijektiv. ■

Der euklidische Algorithmus liefert ein wichtiges Kriterium für Invertierbarkeit in Quotientenringen.

Hilfssatz 1.46 Es sei R ein euklidischer Ring und $a, x \in R$. Es ist $\bar{x} \in R/\langle a \rangle$ genau dann invertierbar, wenn $\text{ggT}(a, x) = 1$ gilt.

BEWEIS: Falls $\text{ggT}(a, x) = 1$ ist, so kann man mit Hilfe des erweiterten euklidischen Algorithmus Elemente $s, t \in R$ finden, die

$$1 = sx + ta$$

erfüllen. Das bedeutet nichts anderes als

$$s \cdot x \equiv 1 \pmod{a}, \quad (*)$$

also $\bar{s} = \bar{x}^{-1}$ in $R/\langle a \rangle$.

Ist umgekehrt \bar{x} invertierbar in $R/\langle a \rangle$ mit Inversem \bar{s} , so folgt aus $(*)$, dass jeder ggT von a und x ein Teiler von 1 ist, also ein Element aus R^\times . Insbesondere ist dann 1 ein ggT von a und x . ■

Mit Hilfe der Quotientenbildung kann man interessante Konstruktionen durchführen, wie die folgenden Beispiele demonstrieren.

Beispiel 1.47 (Einsetzungsabbildung) Es sei $R = \mathbb{R}[X]$ und $\mathfrak{I} = \langle X - \lambda \rangle$. Um das Bild der kanonischen Projektion von $f = a_nX^n + \dots + a_1X + a_0$ zu bestimmen, teilen wir mit Rest,

$$f = (X - \lambda) \cdot h + r,$$

wobei $\deg(r) < \deg(X - \lambda) = 1$, also $r \in \mathbb{R}$. Genauer gilt

$$f(\lambda) = (\lambda - \lambda) \cdot h + r = r.$$

Da $(X - \lambda) \cdot h \in \mathfrak{I}$, ist

$$\bar{f} = \bar{r} = \overline{f(\lambda)}.$$

Das Bild von π enthält somit genau die Restklassen der konstanten Polynome. Wenn wir nun jedes \bar{f} mit seinem konstanten Vertreter $f(\lambda)$ identifizieren, können wir die kanonische Projektion als Einsetzungsabbildung $f \mapsto f(\lambda)$ auffassen. Nach dem Isomorphiesatz gilt dann

$$\mathbb{R}[X]/\mathfrak{I} \cong \mathbb{R}.$$

Beispiel 1.48 (Komplexe Zahlen) Es sei $R = \mathbb{R}[X]$ und $\mathfrak{I} = \langle X^2 + 1 \rangle$. Jedes Polynom $f \in \mathbb{R}[X]$ ist von der Form

$$f = (X^2 + 1) \cdot h + (a_1 X + a_0),$$

so dass

$$f \equiv a_1 X + a_0 \pmod{X^2 + 1}.$$

Insbesondere gilt für das Polynom X :

$$\overline{X} \cdot \overline{X} = \overline{X^2} = \overline{X^2} - (\overline{X^2 + 1}) = \overline{X^2 - X^2 - 1} = -\overline{1},$$

d.h. $\overline{X} = \sqrt{-1}$. Dies motiviert die folgende Abbildung

$$\mathbb{R}[X] \rightarrow \mathbb{C}, \quad f \mapsto a_0 + i a_1.$$

Sie ist ein surjektiver Ringhomomorphismus. Nach dem Isomorphiesatz ist

$$\mathbb{R}[X]/\langle X^2 + 1 \rangle \cong \mathbb{C}.$$

Am Schluss dieses Abschnitts wollen wir eine Konstruktion für endliche Körper angeben. Ein endlicher Körper \mathbb{F} hat notwendigerweise endliche Charakteristik p (prim). Wir kennen bereits die endlichen Körper $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Hilfssatz 1.49 Ist \mathbb{F} ein endlicher Körper mit Charakteristik p , so ist der Körper \mathbb{F}_p ein Unterkörper von \mathbb{F} .

BEWEIS: Die Abbildung $\Phi : \mathbb{Z} \rightarrow \mathbb{F}$, $n \mapsto \text{sgn}(n) \cdot \sum_{i=1}^{|n|} 1$ ist ein Ringhomomorphismus mit Kern $\Phi = p\mathbb{Z}$. Nach dem Isomorphiesatz gibt es einen injektiven Körperhomomorphismus $\overline{\Phi} : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{F}$, d.h. $\mathbb{F}_p \cong \text{Bild } \overline{\Phi} \subset \mathbb{F}$. ■

Folgerung 1.50 Ist \mathbb{F} ein endlicher Körper mit Charakteristik p , so ist \mathbb{F} ein Vektorraum über dem Körper \mathbb{F}_p . Insbesondere ist $|\mathbb{F}| = p^k$, wobei $k = \dim_{\mathbb{F}_p} \mathbb{F}$.

BEWEIS: Nach Hilfssatz 1.49 ist \mathbb{F}_p ein Unterkörper von \mathbb{F} . Also kann man die Elemente von \mathbb{F} mit Skalaren aus \mathbb{F}_p multiplizieren, wodurch \mathbb{F} zu einem \mathbb{F}_p -Vektorraum wird. Da \mathbb{F} endlich ist, muss \mathbb{F} endliche Dimension k über \mathbb{F}_p haben. Dann gibt es p^k Möglichkeiten, Linearkombinationen aus einer gegebenen Basis von \mathbb{F} zu bilden, d.h. \mathbb{F} hat p^k Elemente. ■

Aus Aufgabe 1.36 folgt, dass es in $\mathbb{F}_p[X]$ irreduzible Polynome vom Grad $k > n$ für alle $n \in \mathbb{N}$ geben muss. Das ermöglicht uns die folgende Konstruktion:

Beispiel 1.51 (Endliche Körper) Es seien $R = \mathbb{F}_p[X]$ und $f \in \mathbb{F}_p[X]$ ein irreduzibles Polynom vom Grad k . Schreibe

$$\mathbb{F}_{p^k} = \mathbb{F}_p[X]/\langle f \rangle.$$

Da f irreduzibel ist, besitzt jedes Element in \mathbb{F}_{p^k} ein Inverses, und somit ist \mathbb{F}_{p^k} tatsächlich ein Körper. Schreibe $x = \overline{X}$. Die Elemente $1 = x^0, x, x^2, \dots, x^{k-1}$ erzeugen \mathbb{F}_{p^k} . Sie sind linear unabhängig, denn gilt

$$a_0 + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1} = 0$$

für $a_0, \dots, a_{k-1} \in \mathbb{F}_p$, so bedeutet dies, dass das Polynom $h = a_0 + a_1X + \dots + a_{k-1}X^{k-1}$ im Ideal $\langle f \rangle$ liegt, also ein Vielfaches von f ist. Da $\deg(f) = k > k - 1$ ist, muss h aber das Nullpolynom sein. Die Menge $\{1, x, x^2, \dots, x^{k-1}\}$ ist also eine Basis von \mathbb{F}_{p^k} als \mathbb{F}_p -Vektorraum.

Bemerkung 1.52 Es lässt sich zeigen, dass jeder Körper mit p^k Elementen isomorph ist zu dem in Beispiel 1.51 konstruierten \mathbb{F}_{p^k} . Dies ist der *Hauptsatz der Galois-Theorie für endliche Körper*, siehe Bosch [1], Abschnitt 3.8, Theorem 2.

1.5 Der chinesische Restsatz

Die Lösung eines aufwändigen Problems kann vereinfacht werden, wenn man das Problem *parallelisieren* kann. Das bedeutet, dass man das Problem in mehrere einfacher zu lösende Probleme aufteilt und danach aus deren Teillösungen eine Lösung des ursprünglichen Problems rekonstruieren kann.

Das folgende Beispiel soll diese Idee illustrieren:

Beispiel 1.53 Es soll die Determinante einer Matrix $A \in \mathbb{Z}^{d \times d}$ berechnet werden. Wir nehmen an, die Einträge von A können im Betrag durch eine Konstante c abgeschätzt werden. Mit der Leibniz-Formel für Determinanten schätzt man²⁾

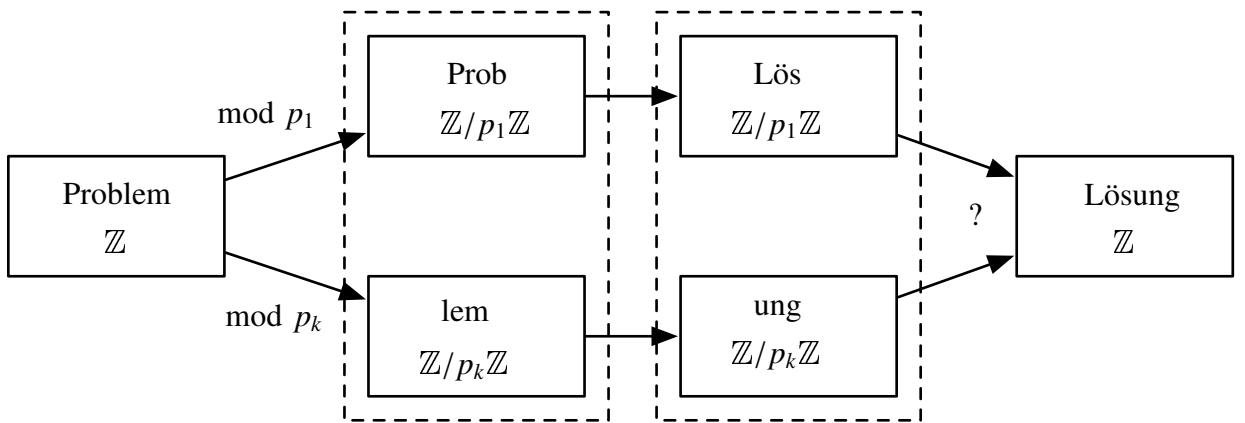
$$|\det(A)| \leq d!c^d.$$

Dank dieser Schranke können wir annehmen, dass wir anstatt in \mathbb{Z} im Ring $\mathbb{Z}/m\mathbb{Z}$ mit einer ungeraden Zahl $m \geq 2d!c^d$ rechnen. Den Faktor 2 benötigen wir, weil die Zahlen $\frac{m-1}{2} + 1, \dots, \frac{m-1}{2}$ in $\mathbb{Z}/m\mathbb{Z}$ die Zahlen $-\frac{m-1}{2}, \dots, -1$ in \mathbb{Z} darstellen sollen. Für großes m wird die direkte Berechnung der Determinante am Computer sehr zeit- und speicherintensiv. Könnte man stattdessen das Problem nach $\mathbb{Z}/p\mathbb{Z}$ mit $p \ll m$ projizieren, so würden die Rechnungen erheblich beschleunigt. Dabei sind zwei Dinge zu beachten:

²⁾Es existieren bessere Abschätzungen.

- (i) Bei der Projektion $\mathbb{Z}/m\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$ müssen die Rechnungen in $\mathbb{Z}/m\mathbb{Z}$ in Rechnungen in $\mathbb{Z}/p\mathbb{Z}$ übergehen, d.h. sie muss ein Homomorphismus sein. Dazu muss p ein Teiler von m sein.
- (ii) Damit dabei keine Information verloren geht, muss man in mehreren solcher $\mathbb{Z}/p_1\mathbb{Z}, \dots, \mathbb{Z}/p_k\mathbb{Z}$ rechnen, so dass man insgesamt m Zahlen codieren kann. Zusammen mit (i) muss also $p_1 \cdots p_k = m$ gelten.

Das Problem wird also von $\mathbb{Z}/m\mathbb{Z}$ nach $\mathbb{Z}/p_1\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k\mathbb{Z}$ übertragen und dort „komponentenweise“ gelöst.



Das Berechnen der Determinanten über $\mathbb{Z}/p_i\mathbb{Z}$ kann parallel erfolgen. Die Frage ist nun, ob und wie man aus diesen Teillösungen die ursprüngliche Determinante $\det(A)$ in $\mathbb{Z}/m\mathbb{Z}$ rekonstruieren kann. Wir werden feststellen, dass dies dann der Fall ist, wenn wir p_1, \dots, p_k so wählen, dass sie paarweise teilerfremd sind (z.B. eine Primfaktorzerlegung von m).

Wir wollen die Rekonstruktion eines Elements in $\mathbb{Z}/m\mathbb{Z}$ aus Elementen in den $\mathbb{Z}/p_i\mathbb{Z}$ an einem einfachen Zahlenbeispiel nachvollziehen.

Beispiel 1.54 Gegeben seien die Elemente $3 \bmod 8 \in \mathbb{Z}/8\mathbb{Z}$ und $7 \bmod 21 \in \mathbb{Z}/21\mathbb{Z}$. Gesucht ist ein $x \in \mathbb{Z}$, so dass gilt

$$\begin{aligned} x &\equiv 3 \bmod 8, \\ x &\equiv 7 \bmod 21. \end{aligned}$$

Es geht also bei dem Problem darum, ein System *simultaner Kongruenzen* zu lösen. Wir machen den folgenden Ansatz:

$$x = 3 + b \cdot 8.$$

Dann ist die erste Kongruenz auf jeden Fall erfüllt. Nun ist b zu bestimmen, so dass auch die zweite Kongruenz gilt:

$$\begin{aligned} 3 + b \cdot 8 &\equiv 7 \pmod{21} \\ \Leftrightarrow b \cdot 8 &\equiv 4 \pmod{21}. \end{aligned}$$

Da 8 und 21 teilerfremd sind, können wir das Inverse von 8 modulo 21 berechnen, wie in Beispiel 1.32 (a) geschehen: $1 = (-3) \cdot 21 + 8 \cdot 8$, also ist 8 das Inverse von 8 modulo 21. Multiplizieren wir beide Seiten der obigen Gleichung mit 8, so erhalten wir

$$b \equiv 4 \cdot 8 \equiv 11 \pmod{21}.$$

Also erfüllt

$$x = 3 + 11 \cdot 8 = 91$$

beide Kongruenzen. Alle Lösungen sind durch die Menge $x + (8 \cdot 21)\mathbb{Z} = 91 + 168\mathbb{Z}$ gegeben. In der Situation von Beispiel 1.53 sind wir an der kleinsten Lösung zwischen 0 und 168 interessiert, die ja ein Element aus $\mathbb{Z}/168\mathbb{Z}$ repräsentieren soll. Dies ist bereits $x = 91$.

Nun betrachten wir die Situation allgemein für euklidische Ringe. Zunächst bemerken wir, dass für Ringe R_1, \dots, R_k auch das Produkt $R_1 \times \dots \times R_k$ ein Ring ist, wenn wir die Verknüpfungen komponentenweise definieren:

$$\begin{aligned} (x_1, \dots, x_k) + (y_1, \dots, y_k) &= (x_1 + y_1, \dots, x_k + y_k), \\ (x_1, \dots, x_k) \cdot (y_1, \dots, y_k) &= (x_1 \cdot y_1, \dots, x_k \cdot y_k). \end{aligned}$$

Satz 1.55 (Chinesischer Restsatz) Es sei R ein euklidischer Ring und $p_1, \dots, p_k \in R$ paarweise teilerfremd. Dann ist für $q = p_1 \cdots p_k$ die Abbildung

$$\Psi : R/\langle q \rangle \rightarrow R/\langle p_1 \rangle \times \dots \times R/\langle p_k \rangle, \quad a \pmod{q} \mapsto (a \pmod{p_1}, \dots, a \pmod{p_k}) \quad (1.9)$$

ein Isomorphismus von Ringen.

BEWEIS: Die Abbildung ist in jeder Komponente wohldefiniert: Für $a \equiv a' \pmod{q}$ gibt es ein $h \in R$, so dass

$$a - a' = hq = h(p_1 \cdots p_k) = (hp_2 \cdots p_k)p_1,$$

also $a \equiv a' \pmod{p_1}$ und entsprechend für p_2, \dots, p_k . Offensichtlich ist Ψ dann ein Homomorphismus von Ringen.

Ψ ist injektiv: Ist $a \pmod{q} \in \text{Kern } \Psi$, so gilt $a \equiv 0 \pmod{p_i}$ für $i = 1, \dots, k$. Also ist a ein Vielfaches von jedem p_i . Da die p_i nach Voraussetzung teilerfremd sind, ist a

auch ein Vielfaches von ihrem Produkt $p_1 \cdots p_k = q$. Also ist $\text{Kern } \Psi = \{0 \bmod q\}$ trivial und Ψ somit injektiv.

Die Surjektivität von Ψ folgt, wenn wir für jedes $(a_1 \bmod p_1, \dots, a_k \bmod p_k) \in R/\langle p_1 \rangle \times \dots \times R/\langle p_k \rangle$ ein Urbild bestimmen können. Dies ist gleichbedeutend damit, das folgende System simultaner Kongruenzen nach $x \in R$ zu lösen:

$$\begin{aligned} x &\equiv a_1 \bmod p_1, \\ &\vdots \\ x &\equiv a_k \bmod p_k. \end{aligned}$$

Dann ist $x \bmod q$ das gesuchte Urbild. Die Lösbarkeit dieses Systems folgt aus dem folgenden Algorithmus 1.56 (oder Algorithmus 1.57). ■

Algorithmus 1.56 (Newton-Interpolation) Es sei R ein euklidischer Ring. Gegeben sei das nach $x \in R$ zu lösende System

$$\begin{aligned} x &\equiv a_1 \bmod p_1, \\ &\vdots \\ x &\equiv a_k \bmod p_k, \end{aligned}$$

mit paarweise teilerfremden p_1, \dots, p_k . Mache den Ansatz

$$x = x_1 + x_2 p_1 + x_3 p_1 p_2 + \dots + x_k p_1 \cdots p_{k-1} \quad (1.10)$$

und bestimmte iterativ die x_i :

- Im ersten Schritt gilt

$$x \equiv x_1 \stackrel{!}{\equiv} a_1 \bmod p_1,$$

wir können also $x_1 = a_1$ wählen.

- Angenommen, x_1, \dots, x_{i-1} seien bereits bestimmt, dann gilt

$$x \equiv x_1 + x_2 p_1 + x_3 p_1 p_2 + \dots + x_i p_1 \cdots p_{i-1} \stackrel{!}{\equiv} a_i \bmod p_i.$$

Umformen ergibt

$$x_i p_1 \cdots p_{i-1} \equiv a_i - x_1 - x_2 p_1 - x_3 p_1 p_2 - \dots - x_{i-1} p_1 \cdots p_{i-2} \bmod p_i.$$

Dann ist alles auf der rechten Seite der Gleichung bereits bekannt und wir können nach x_i auflösen, wenn wir $p_1 \cdots p_{i-1}$ modulo p_i invertieren können. Da die p_1, \dots, p_k teilerfremd sind, ist dies nach Hilfssatz 1.46 möglich, und

das Inverse kann mit dem erweiterten euklidischen Algorithmus berechnet werden. Es bezeichne s_i das Inverse von $p_1 \cdots p_{i-1}$ modulo p_i . Setze

$$x_i = (a_i - x_1 - x_2 p_1 - x_3 p_1 p_2 - \dots - x_{i-1} p_1 \cdots p_{i-2}) \cdot s_i.$$

Nach k dieser Schritte endet der Algorithmus und gibt eine Lösung x aus.

Algorithmus 1.57 (Lagrange-Interpolation) Es sei R ein euklidischer Ring. Gegeben sei das nach $x \in R$ zu lösende System

$$\begin{aligned} x &\equiv a_1 \pmod{p_1}, \\ &\vdots \\ x &\equiv a_k \pmod{p_k}, \end{aligned}$$

mit paarweise teilerfremden p_1, \dots, p_k . Mache den Ansatz

$$x = a_1 q_1 + a_2 q_2 + \dots + a_k q_k, \quad (1.11)$$

wobei

$$q_i = u_i \cdot p_1 \cdots p_{i-1} \cdot p_{i+1} \cdots p_k$$

mit unbekannten $u_i \in R$. Dann ist

$$x \equiv a_i q_i \stackrel{!}{\equiv} a_i \pmod{p_i}.$$

Folglich muss u_i so gewählt werden, dass

$$q_i = u_i \cdot p_1 \cdots p_{i-1} \cdot p_{i+1} \cdots p_k \equiv 1 \pmod{p_i}$$

gilt. Da p_i und $p_1 \cdots p_{i-1} \cdot p_{i+1} \cdots p_k$ nach Voraussetzung teilerfremd sind, kann ein solches u_i mit dem erweiterten euklidischen Algorithmus berechnet werden (Hilfssatz 1.46). Nach dem Berechnen der u_1, \dots, u_k und q_1, \dots, q_k liefert Einsetzen in (1.11) eine Lösung x .

Bemerkung 1.58 Hat man eine Lösung x der Kongruenzgleichungen gefunden, so ist die Menge aller Lösungen gegeben durch

$$x + \langle p_1 \cdots p_k \rangle.$$

Dies folgt aus der Injektivität der Abbildung Ψ im chinesischen Restsatz 1.55.

Bemerkung 1.59 Die Newton-Interpolation hat gegenüber der Lagrange-Interpolation den Vorteil, dass beim Hinzukommen einer weiteren Kongruenz $x \equiv a_{k+1} \pmod{p_{k+1}}$ mit dem bisherigen Ergebnis weitergerechnet werden kann, indem einfach ein weiterer Iterationsschritt angehängt wird (dagegen müssten bei der Lagrange-Interpolation alle q_i neu berechnet werden). Der Vorteil der Lagrange-Interpolation besteht darin, dass die q_i nicht von den a_i abhängen, also für feste p_i nur einmal berechnet werden müssen, um Kongruenzen mit beliebigen a_i zu lösen.

Die Namen *Newton-Interpolation* und *Lagrange-Interpolation* bezeichnen in der numerischen Mathematik Verfahren, um das Interpolationsproblem für Polynome zu lösen: Der Wert eines Polynoms f sei an k Stellen $\alpha_1, \dots, \alpha_k$ bekannt,

$$f(\alpha_i) = \beta_i \quad i = 1, \dots, k.$$

Gesucht ist das Polynom f kleinsten Grades, das diese Gleichungen erfüllt. Dies ist ein lineares Gleichungssystem für die Koeffizienten von f , und man kann mit den üblichen Methoden der linearen Algebra zeigen, dass es für paarweise verschiedene α_i immer lösbar ist.

Der Zusammenhang mit dem chinesischen Restsatz wird klar, wenn man sich an Beispiel 1.47 erinnert. Die Bedingung

$$f(\alpha_i) = \beta_i$$

ist nämlich äquivalent zu

$$f \equiv \beta_i \pmod{X - \alpha_i}.$$

Das Interpolationsproblem kann also durch simultane Kongruenzen ausgedrückt werden:

$$\begin{aligned} f &\equiv \beta_1 \pmod{X - \alpha_1}, \\ &\vdots \\ f &\equiv \beta_k \pmod{X - \alpha_k}, \end{aligned}$$

wobei die $X - \alpha_i$ paarweise teilerfremd sind, wenn die α_i paarweise verschieden sind (Satz 1.37 (c)). Wendet man die Algorithmen 1.56 oder 1.57 in diesem Spezialfall an, so erhält man die ursprünglichen numerischen Verfahren.

Beispiel 1.60 (Polynominterpolation) Gegeben seien die Stellen $-1, 0, 1$ und die Funktionswerte

$$f(-1) = 2, \quad f(0) = -1, \quad f(1) = 1.$$

Gesucht ist ein Polynom $f \in \mathbb{R}[X]$ vom Grad 2, das diese Werte annimmt. Der Ansatz der Newton-Interpolation ist

$$f = f_1 + f_2(X+1) + f_3(X+1)X.$$

Bestimme die f_i :

- Es ist $f \equiv f_1 \equiv 2 \pmod{X+1}$. Setze $f_1 = 2$.
- Es ist $f \equiv 2 + f_2(X+1) \equiv -1 \pmod{X}$. Da $X+1 \equiv 1 \pmod{X}$, muss in diesem Schritt kein Inverses bestimmt werden. Setze $f_2 = -1 - 2 = -3$.
- Es ist $f \equiv 2 - 3(X+1) + f_3(X+1)X \equiv 1 \pmod{X-1}$. Es ist $X+1 \equiv 2 \pmod{X-1}$ und $X \equiv 1 \pmod{X-1}$, d.h. $2 - 3 \cdot 2 + f_3 \cdot 2 \cdot 1 \equiv 1 \pmod{X-1}$. Setze $f_3 = \frac{1}{2}(1 - 2 + 6) = \frac{5}{2}$.

Dann ist

$$f = 2 - 3(X+1) + \frac{5}{2}(X+1)X = \frac{5}{2}X^2 - \frac{1}{2}X - 1.$$

Das selbe Ergebnis leiten wir nun mit Lagrange-Interpolation her. Der Ansatz ist

$$f = 2 \cdot u_1 \cdot X(X-1) + (-1) \cdot u_2 \cdot (X+1)(X-1) + 1 \cdot u_3 \cdot (X+1)X.$$

Um u_1, u_2, u_3 zu bestimmen, beachte dass

$$\begin{aligned} X(X-1) &\equiv 2 \pmod{X+1}, \\ (X+1)(X-1) &\equiv -1 \pmod{X}, \\ (X+1)X &\equiv 2 \pmod{X-1}, \end{aligned}$$

also setze $u_1 = \frac{1}{2}$, $u_2 = -1$, $u_3 = \frac{1}{2}$. Damit ist

$$f = 2 \cdot \frac{1}{2} \cdot X(X-1) + (-1) \cdot (-1) \cdot (X+1)(X-1) + 1 \cdot \frac{1}{2} \cdot (X+1)X = \frac{5}{2}X^2 - \frac{1}{2}X - 1.$$

Schließlich greifen wir das einführende Beispiel 1.53 noch einmal mit konkreten Werten auf:

Beispiel 1.61 (Ganzzahlige Determinante) Es sei

$$A = \begin{pmatrix} -17 & 23 \\ -5 & 22 \end{pmatrix} \in \mathbb{Z}^{2 \times 2}$$

mit $|\det(A)| \leq 2! \cdot 23^2 = 1058$. Für die drei Primzahlen $p_1 = 11$, $p_2 = 13$, $p_3 = 17$ gilt

$$11 \cdot 13 \cdot 17 = 2431 > 2 \cdot 1058 + 1 = 2117.$$

- (i) Wir fassen $\det(A)$ als ein Element von $\mathbb{Z}/2431\mathbb{Z}$ auf, wobei die Zahlen $1215, \dots, 2430$ die negativen Lösungen in \mathbb{Z} repräsentieren.
- (ii) Dann parallelisieren wir die Rechnung, indem wir sie auf $\mathbb{Z}/11\mathbb{Z}$, $\mathbb{Z}/13\mathbb{Z}$ und $\mathbb{Z}/17\mathbb{Z}$ aufteilen. Wir berechnen die Determinanten der Matrizen

$$A_{11} = \begin{pmatrix} \overline{5} & \overline{1} \\ \overline{6} & \overline{0} \end{pmatrix} \in (\mathbb{Z}/11\mathbb{Z})^{2 \times 2}, A_{13} = \begin{pmatrix} \overline{9} & \overline{10} \\ \overline{8} & \overline{9} \end{pmatrix} \in (\mathbb{Z}/13\mathbb{Z})^{2 \times 2}, A_{17} = \begin{pmatrix} \overline{0} & \overline{6} \\ \overline{12} & \overline{5} \end{pmatrix} \in (\mathbb{Z}/17\mathbb{Z})^{2 \times 2} :$$

$$\det(A_{11}) = \overline{5}, \quad \det(A_{13}) = \overline{1}, \quad \det(A_{17}) = \overline{13}.$$

- (iii) Nun rekonstruieren wir $\det(A) \bmod 2431$ aus den Kongruenzen

$$\begin{aligned} \det(A) &\equiv 5 \bmod 11, \\ \det(A) &\equiv 1 \bmod 13, \\ \det(A) &\equiv 13 \bmod 17. \end{aligned}$$

Für die Newton-Interpolation setzt man an:

$$\det(A) = a_1 + a_2 \cdot 11 + a_3 \cdot 11 \cdot 13.$$

- Setze $a_1 = 5$.
- Es gilt $5 + a_2 \cdot 11 \equiv 1 \bmod 13$. Mit dem erweiterten euklidischen Algorithmus findet man $1 = 6 \cdot 11 - 5 \cdot 13$, also ist 6 das Inverse von 11 modulo 13. Damit können wir nach a_2 auflösen: $a_2 \equiv (1 - 5) \cdot 6 \equiv 2 \bmod 13$. Setze $a_2 = 2$.
- Es gilt $5 + 2 \cdot 11 + a_3 \cdot 11 \cdot 13 \equiv 13 \bmod 17$. Das Inverse von $11 \cdot 13 = 143$ modulo 17 ist 5. Also ist $a_3 \equiv (13 - 5 - 22) \cdot 5 \equiv 15 \bmod 17$. Setze $a_3 = 15$.

Dann ist $\det(A) = 5 + 2 \cdot 11 + 15 \cdot 11 \cdot 13 = 2172 \bmod 2431$.

- (iv) Da $2172 > 1214$, müssen wir dieses Ergebnis als negative Zahl in \mathbb{Z} interpretieren. Dann ist

$$\det(A) = 2172 - 2431 = -259.$$

Weitere Anwendungen des chinesischen Restsatzes finden sich bei Lipson [11].

Der Vollständigkeit wegen geben wir noch eine allgemeinere Form des chinesischen Restsatzes an. Diese allgemeine Form liefert jedoch kein Verfahren, um die Umkehrfunktion des Isomorphismus zu berechnen.

Satz 1.62 (Chinesischer Restsatz für beliebige Ringe) *Es seien R ein Ring und $\mathfrak{J}_1, \dots, \mathfrak{J}_k$ paarweise coprime Ideale in R . Weiter sei $\mathfrak{H} = \bigcap_{j=1}^k \mathfrak{J}_j$ (dies ist ebenfalls ein Ideal in R). Dann ist die Abbildung*

$$\Psi : R/\mathfrak{H} \rightarrow R/\mathfrak{J}_1 \times \cdots \times R/\mathfrak{J}_k, \quad x \bmod \mathfrak{H} \mapsto (x \bmod \mathfrak{J}_1, \dots, x \bmod \mathfrak{J}_k) \quad (1.12)$$

ein Isomorphismus von Ringen.

Zum Beweis siehe Bosch [1], Satz 12 in Abschnitt 2.3.

2 Die Jordansche Normalform

2.1 Invariante Untervektorräume

In diesem Abschnitt sei V stets ein \mathbb{K} -Vektorraum.

Definition 2.1 Es sei $\Phi : V \rightarrow V$ ein Endomorphismus. Ein Untervektorraum U von V heißt **Φ -invariant**, falls $\Phi(U) \subset U$ gilt.

Ist U invariant, so ist die Einschränkung $\Phi|_U$ ein Endomorphismus von U .

Definition 2.2 Es sei U ein Φ -invarianter Untervektorraum von V . Ein Untervektorraum W von V , der $V = U \oplus W$ erfüllt, heißt **Komplementärraum** zu U . Ist W ebenfalls Φ -invariant, so nennen wir W ein **invariantes Komplement** zu U .

Beispiel 2.3 (Invariante Unterräume)

- (a) Die Räume $\{0\}$ und V sind für jeden Endomorphismus invariant.
- (b) Jeder Eigenraum (insbesondere der Kern) von Φ ist invariant. Ist E_λ der Eigenraum zu Eigenwert λ , so spannt jeder Eigenvektor $x \in E_\lambda$ seinerseits einen invarianten Untervektorraum $[x]$ von E_λ auf.

Jeder 1-dimensionale invariante Untervektorraum ist ein Eigenraum.

- (c) Es sei e_1, e_2, e_3 die Standardbasis des \mathbb{R}^3 . Es bezeichne Φ_α die Drehung um den Winkel α in der e_2, e_3 -Ebene. Dann ist der Untervektorraum $[e_2, e_3]$ invariant unter Φ_α , da innerhalb dieser Ebene gedreht wird. Die Drehachse ist die e_1 -Achse, d.h. $\Phi_\alpha(e_1) = e_1$. Diese Achse ist somit ein Eigenraum, also ein invariantes Komplement zur Drehebene. Daher lässt sich \mathbb{R}^3 als Summe von zwei invarianten Untervektorräumen schreiben:

$$\mathbb{R}^3 = [e_1] \oplus [e_2, e_3].$$

Diese Invarianz wird durch die Blockform der Abbildungsmatrix von Φ_α wiedergegeben:

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{array} \right).$$

- (d) Die Abbildung

$$\Phi(x) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot x$$

hat den invarianten Unterraum $[e_1]$.

In diesem Beispiel existiert kein invariantes Komplement zu $[e_1]$, denn dieses müsste Dimension $\dim \mathbb{R}^2 - \dim [e_1] = 1$ haben und somit ein Eigenraum sein. Da Φ aber nur den Eigenwert 1 hat, wäre Φ dann diagonalisierbar mit Eigenwert 1, also $\Phi = \text{id}_{\mathbb{R}^2}$, was aber offensichtlich nicht der Fall ist.

Diese Beispiele deuten ein allgemeines Prinzip (bei endlicher Dimension n) an: Sei U ein Φ -invarianter Untervektorraum und $B_U = \{b_1, \dots, b_k\}$ eine Basis von U . Ergänze diese zu einer Basis $B = \{b_1, \dots, b_k, c_{k+1}, \dots, c_n\}$ von V . Die Invarianz bedeutet, dass die Bilder $\Phi(b_i)$ bzgl. der Basis B die folgende Form haben:

$$\Phi(b_i) = \lambda_1 \cdot b_1 + \dots + \lambda_k \cdot b_k + 0 \cdot c_{k+1} + \dots + 0 \cdot c_n$$

Für die Abbildungsmatrix $M_B^B(\Phi)$ von Φ bzgl. der Basis B bedeutet dies, dass sie folgende Blockgestalt hat:

$$M_B^B(\Phi) = \begin{pmatrix} M_{B_U}^{B_U}(\Phi|_U) & M \\ 0 & N \end{pmatrix}$$

für geeignete Matrizen $N \in \mathbb{K}^{(n-k) \times (n-k)}$, $M \in \mathbb{K}^{n \times (n-k)}$. Diese Darstellung lässt sich noch genauer angeben: Gilt

$$\Phi(c_{k+i}) = \mu_{1i} \cdot b_1 + \dots + \mu_{ki} \cdot b_k + \nu_{1i} \cdot c_{k+1} + \dots + \nu_{n-k,i} \cdot c_n, \quad (*)$$

so sind die μ_{ij} gerade die Koeffizienten der Matrix M und die ν_{ij} die Koeffizienten der Matrix N . Da U invariant unter Φ ist, induziert³⁾ Φ einen Endomorphismus

$$\bar{\Phi} : V/U \rightarrow V/U, \quad x + U \mapsto \Phi(x) + U.$$

Eine Basis C von V/U ist durch die Klassen der Basisvektoren c_{k+1}, \dots, c_n gegeben:

$$\bar{c}_1 = c_{k+1} + U, \dots, \bar{c}_{n-k} = c_n + U.$$

Da die Elemente b_1, \dots, b_k in U liegen, werden sie in V/U auf die 0 projiziert. Somit ergibt sich die Darstellung von $\bar{\Phi}$ in der Basis C aus $(*)$ zu

$$\begin{aligned} \bar{\Phi}(\bar{c}_i) &= \underbrace{\mu_{1i} \cdot \bar{b}_1}_{=0} + \dots + \underbrace{\mu_{ki} \cdot \bar{b}_k}_{=0} + \nu_{1i} \cdot \bar{c}_1 + \dots + \nu_{n-k,i} \cdot \bar{c}_{n-k} \\ &= \nu_{1i} \cdot \bar{c}_1 + \dots + \nu_{n-k,i} \cdot \bar{c}_{n-k}, \end{aligned}$$

d.h. die Abbildungsmatrix von $\bar{\Phi}$ bzgl. C ist

$$M_C^C(\bar{\Phi}) = N.$$

³⁾Die Invarianz von U ist für die Wohldefiniertheit von $\bar{\Phi}$ notwendig.

Insgesamt gilt somit

$$M_B^B(\Phi) = \begin{pmatrix} M_{B_U}^{B_U}(\Phi|_U) & M \\ 0 & M_C^C(\bar{\Phi}) \end{pmatrix}. \quad (2.1)$$

Aus dieser Form lesen wir direkt den folgenden Hilfssatz ab:

Hilfssatz 2.4 *Mit $U, \Phi, \bar{\Phi}$ wie oben gilt:*

- (a) $\det(\Phi) = \det(\Phi|_U) \cdot \det(\bar{\Phi})$.
- (b) Für die charakteristischen Polynome gilt $f_\Phi = f_{\Phi|_U} \cdot f_{\bar{\Phi}}$. Insbesondere ist $f_{\Phi|_U}$ ein Teiler von f_Φ .
- (c) $\text{Spec } \Phi = \text{Spec } \Phi|_U \cup \text{Spec } \bar{\Phi}$.

Falls U ein invariantes Komplement W besitzt, so kann man die Basis B_U durch eine Basis B_W zu einer Basis B von V ergänzen. Dann ist die Abbildungsmatrix (2.1) von Φ bzgl. B von der Gestalt

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

wobei A_1 eine Abbildungsmatrix von $\Phi|_U$ ist und A_2 eine Abbildungsmatrix von $\Phi|_W$. Dies motiviert die folgende Notation:

Definition 2.5 Es seien U, W Untervektorräume von V , so dass $U \oplus W = V$ gilt. Für Endomorphismen $\Phi_U : U \rightarrow U$ und $\Phi_W : W \rightarrow W$ definieren wir ihre **direkte Summe**

$$\Phi_U \oplus \Phi_W : V \rightarrow V$$

durch

$$\begin{aligned} (\Phi_U \oplus \Phi_W)(u) &= \Phi_U(u) && \text{für } u \in U, \\ (\Phi_U \oplus \Phi_W)(w) &= \Phi_W(w) && \text{für } w \in W. \end{aligned}$$

Entsprechend definieren wir für Matrizen $A \in \mathbb{K}^{n \times n}$ und $B \in \mathbb{K}^{m \times m}$ die direkte Summe

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathbb{K}^{(n+m) \times (n+m)}.$$

Entsprechend wird die direkte Summe $\Phi_1 \oplus \dots \oplus \Phi_k$ (bzw. $A_1 \oplus \dots \oplus A_k$) von k Endomorphismen (bzw. Matrizen) definiert.

Unsere obigen Überlegungen können wir nun folgendermaßen ausdrücken:

Hilfssatz 2.6 Ist U ein Φ -invarianter Untervektorraum von V mit invariantem Komplement W , so gilt

$$\Phi = \Phi|_U \oplus \Phi|_W.$$

Ist B_U eine Basis von U und B_W eine Basis von W , so wird Φ bzgl. der Basis $B = B_U \cup B_W$ von V durch die Matrix

$$M_B^B(\Phi) = M_{B_U}^{B_U}(\Phi|_U) \oplus M_{B_W}^{B_W}(\Phi|_W)$$

dargestellt.

Satz 2.7 Es seien $\Phi, \Psi \in \text{End}(V)$ und es gelte $\Phi \circ \Psi = \Psi \circ \Phi$. Dann gilt:

- (a) Jeder Eigenraum E_λ von Φ ist invariant unter Ψ .
- (b) Kern Φ ist invariant unter Ψ .
- (c) Bild Φ ist invariant unter Ψ .

BEWEIS:

- (a) Sei $x \in E_\lambda$, $x \neq 0$. Nach Voraussetzung gilt

$$\lambda \cdot \Psi(x) = \Psi(\lambda \cdot x) = \Psi(\Phi(x)) = \Phi(\Psi(x)).$$

Also ist $\Psi(x)$ entweder 0 oder ein Eigenvektor von Φ , d.h. $\Psi(E_\lambda) \subset E_\lambda$.

- (b) $\text{Kern } \Phi = E_0$.
- (c) $\Psi(\text{Bild } \Phi) = \text{Bild}(\Psi \circ \Phi) = \text{Bild}(\Phi \circ \Psi) = \Phi(\text{Bild } \Psi) \subset \text{Bild } \Phi$. ■

Die analogen Aussagen gelten für kommutierende Matrizen.

Aufgabe 2.8 (Lemma von Schur) Es sei V ein \mathbb{C} -Vektorraum, $\dim V < \infty$, und $\Phi \in \text{End}(V)$, so dass $\Phi \circ \Psi = \Psi \circ \Phi$ für alle $\Psi \in \text{End}(V)$ gilt. Dann gibt es ein $\lambda \in \mathbb{C}$, so dass gilt

$$\Phi = \lambda \cdot \text{id}_V.$$

2.2 Nilpotente Endomorphismen

Definition 2.9 Ein Endomorphismus $\Phi : V \rightarrow V$ heißt **nilpotent**, falls $\Phi^k = 0$ für ein $k \in \mathbb{N}$ gilt. Entsprechend heißt eine Matrix $A \in \mathbb{K}^{n \times n}$ nilpotent, falls $A^k = 0$ für ein $k \in \mathbb{N}$. Ist k minimal, d.h. $\Phi^m \neq 0$ für $m < k$, so heißt k die **Stufe** von Φ .

Beispiel 2.10 (Nilpotente Matrizen) Die Matrix

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

ist nilpotent von Stufe 2. Allgemeiner ist jede obere $n \times n$ -Dreiecksmatrix der Form

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

nilpotent von Stufe n : Man rechnet nach, dass gilt

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ 0 & & & & 0 \end{pmatrix}, A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & 1 \\ & & & \ddots & \ddots & 0 \\ & & & & \ddots & 0 \\ 0 & & & & & 0 \end{pmatrix}, \dots, A^{n-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 & & \\ & \ddots & & & & 0 \\ & & \ddots & & & \vdots \\ 0 & & & & & 0 \end{pmatrix}, A^n = 0.$$

Bemerkung 2.11 Die Nilpotenz der Matrizen aus Beispiel 2.10 kommt auch in ihrem charakteristischen Polynom f_A zum Ausdruck: Es ist $f_A = X^n$, und nach dem Satz von Cayley-Hamilton gilt $f_A(A) = A^n = 0$.

Satz 2.12 Es sei V ein n -dimensionaler Vektorraum. $\Phi \in \text{End}(V)$ ist genau dann nilpotent, wenn das charakteristische Polynom $f_\Phi = X^n$ ist.

BEWEIS: Beweis durch Induktion über n : Für $n = 1$ stimmt der Satz offensichtlich. Es sei nun $n > 1$. Da 0 eine Nullstelle von f_Φ ist, gibt es einen Eigenvektor x von Φ zum Eigenwert 0. Wenn wir x zu einer Basis von V ergänzen, können wir Φ bzgl. dieser Basis durch eine nilpotente Matrix

$$A = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \in \mathbb{K}^{n \times n}$$

mit $B \in \mathbb{K}^{(n-1) \times (n-1)}$ darstellen. Daraus lesen wir ab:

$$f_\Phi = X \cdot f_B.$$

Da $A^k = 0$ gilt für ein $k > 0$, gilt auch $B^k = 0$. Multiplikation mit B definiert also einen nilpotenten Endomorphismus von \mathbb{K}^{n-1} . Nach Induktionsvoraussetzung gilt $f_B = X^{n-1}$. Folglich ist $f_\Phi = X \cdot X^{n-1} = X^n$. ■

Folgerung 2.13 Ist Φ nilpotent von Stufe k , so gilt stets $k \leq n$.

BEWEIS: Das charakteristische Polynom von Φ ist X^n . Nach dem Satz von Cayley-Hamilton gilt $\Phi^n = 0$. Somit ist $k \leq n$. ■

Beispiel 2.14 Die beiden 4×4 -Matrizen

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

sind nilpotent. Dabei ist A von Stufe 4 und B von Stufe 2. Beide Matrizen haben das charakteristische Polynom $f_A = f_B = X^4$. Für A gibt es kein Polynom $0 \neq p = p_0 + p_1X + p_2X^2 + p_3X^3$ kleineren Grades, dass $p(A) = 0$ erfüllt, denn

$$p(A) = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 \\ 0 & p_0 & p_1 & p_2 \\ 0 & 0 & p_0 & p_1 \\ 0 & 0 & 0 & p_0 \end{pmatrix} \neq 0.$$

Für B hingegen gilt $B^2 = 0$. Das Polynom $h = X^2$ ist das Polynom minimalen positiven Grades, für das $h(B) = 0$ gilt, denn für $0 \neq p = p_0 + p_1X$ ist

$$p(B) = \begin{pmatrix} p_0 & p_1 & 0 & 0 \\ 0 & p_0 & 0 & 0 \\ 0 & 0 & p_0 & p_1 \\ 0 & 0 & 0 & p_0 \end{pmatrix} \neq 0.$$

Anhand des charakteristischen Polynoms kann man erkennen, ob eine Matrix nilpotent ist. Wie das Beispiel 2.14 zeigt, gibt es jedoch keinen Aufschluss über die Nilpotenzstufe. Dies veranlasst uns zu folgender Definition:

Definition 2.15 Es sei V ein Vektorraum von endlicher Dimension n . Weiter sei $\Phi \in \text{End}(V)$. Wir nennen $h_\Phi \in \mathbb{K}[X]$ das **Minimalpolynom** von Φ , wenn gilt:

- (i) $h_\Phi(\Phi) = 0$.
- (ii) h_Φ ist normiert.
- (iii) h_Φ ist das Polynom minimalen Grades mit den Eigenschaften (i) und (ii).

Entsprechend definieren wir das Minimalpolynom h_A für eine Matrix $A \in \mathbb{K}^{n \times n}$.

Beispiel 2.16 In Beispiel 2.14 haben A und B die jeweiligen Minimalpolynome $h_A = X^4 = f_A$ und $h_B = X^2$.

Bemerkung 2.17 Erinnern wir uns, dass in $\mathbb{K}[X]$ jedes Ideal ein Hauptideal ist (Satz 1.33). Dann ist der normierte Erzeuger h des Ideals

$$\mathfrak{J}_\Phi = \{f \in \mathbb{K}[X] \mid f(\Phi) = 0\} = \langle h \rangle \quad (2.2)$$

gerade das Minimalpolynom von Φ . Insbesondere ist damit gezeigt, dass ein eindeutiges Minimalpolynom für jedes Φ existiert.

Hilfssatz 2.18 Es sei h_Φ das Minimalpolynom von $\Phi \in \text{End}(V)$. Dann gilt:

- (a) Für jede Abbildungsmatrix $A = M_B^B(\Phi)$ von Φ ist $h_A = h_\Phi$.
- (b) h_Φ teilt das charakteristische Polynom f_Φ .

BEWEIS:

- (a) $M_B^B : \text{End}(V) \rightarrow \mathbb{K}^{n \times n}$ ist ein Isomorphismus von Vektorräumen und von Ringen. Somit ist $h_\Phi(A) = h_\Phi(M_B^B(\Phi)) = M_B^B(h_\Phi(\Phi)) = 0$. Außerdem ist h_Φ normiert und es existiert kein Polynom p mit $0 < \deg(p) < \deg(h_\Phi)$, das $p(A) = 0$ erfüllt, denn dann würde auch $p(\Phi) = (M_B^B)^{-1}(p(A)) = 0$ gelten, im Widerspruch dazu, dass h_Φ das Minimalpolynom von Φ ist. Also gilt $h_A = h_\Phi$.
- (b) Es ist $f_\Phi \in \mathfrak{J}_\Phi = \langle h_\Phi \rangle$. ■

Satz 2.19 Es sei V ein n -dimensionaler Vektorraum. $\Phi \in \text{End}(V)$ ist genau dann nilpotent von Stufe k , wenn das Minimalpolynom $h_\Phi = X^k$ ist.

BEWEIS: $\Phi^k = 0$ bedeutet $X^k \in \mathfrak{J}_\Phi$. Dann ist X^k ein Vielfaches von h_Φ und alle Primfaktoren von h_Φ müssen auch Primfaktoren von X^k sein. Da X der einzige Primteiler von X^k ist, muss folglich $h_\Phi = X^m$ für ein gewisses $m > 0$ gelten. Da $\Phi^m \neq 0$ ist für $m < k$ und $\deg(h_\Phi)$ minimal in \mathfrak{J}_Φ ist, gilt $h_\Phi = X^k$.

Ist umgekehrt $h_\Phi = X^k$, so ist $\Phi^k = 0$ und $\Phi^m \neq 0$ für $m < k$ nach Definition des Minimalpolynoms. ■

2.3 Das Programm

Die Herleitung der Jordanschen Normalform ist etwas langwierig und leichter zu überblicken, wenn man das Ziel der Mühen bereits vor Augen hat. Daher wird

in diesem Abschnitt skizziert, welche Schritte zu Bestimmung der Jordanschen Normalform erforderlich sind.

Wir werden sehen, dass jeder Endomorphismus Φ eines n -dimensionalen \mathbb{C} -Vektorraumes V durch seine **Jordansche Normalform**

$$J(\Phi) = \begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_1 & 1 \\ & & & & \lambda_1 \\ \\ \lambda_2 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_2 & 1 \\ & & & & \lambda_2 \\ \\ & & & \ddots & \ddots \\ & & & & \lambda_k & 1 \\ & & & & & \lambda_k & 1 \\ & & & & & & \ddots & \ddots \\ & & & & & & & \lambda_k & 1 \\ & & & & & & & & \lambda_k \end{pmatrix} \in \mathbb{C}^{n \times n}$$

dargestellt werden kann (die λ_i sind dabei nicht notwendigerweise verschieden).

Die in der Matrix $J(\Phi)$ auftauchenden Kästchen bezeichnen wir als **Jordan-Kästchen**

$$J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{n_i \times n_i}$$

der Größe n_i zum Eigenwert λ_i . Damit können wir $J(\Phi)$ kürzer schreiben als

$$J(\Phi) = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \dots \oplus J_{n_k}(\lambda_k).$$

Die Matrix $J(\Phi)$ ist die Summe

$$J(\Phi) = D + N$$

der Diagonalmatrix D mit den Einträgen λ_i und der nilpotenten Matrix $N = J_{n_1}(0) \oplus J_{n_2}(0) \oplus \dots \oplus J_{n_k}(0)$. Die Kenntnis von $J(\Phi)$ ist also gleichbedeutend mit der Kenntnis von D und N .

Wir überlegen, wie eine **Jordan-Basis** B zu bestimmen ist, in der Φ durch die zugehörige Jordansche Normalform $J(\Phi)$ dargestellt wird.

1. Zunächst überlegen wir uns, wie D aus Φ bestimmt werden kann. Offensichtlich haben D und $J(\Phi)$ (also auch Φ) das selbe charakteristische Polynom

$$f_\Phi = \det(X \cdot \text{id}_V - \Phi) = (X - \lambda_1)^{n_1}(X - \lambda_2)^{n_2} \cdots (X - \lambda_k)^{n_k} = f_D,$$

die Diagonaleinträge von D sind also gerade die Eigenwerte von Φ (der Fundamentalsatz der Algebra garantiert uns dabei, dass f_Φ tatsächlich in Linearfaktoren zerfällt). Der Exponent n_i eines Eigenwertes λ_i gibt an, wie oft λ_i auf der Diagonalen von D bzw. $J(\Phi)$ auftaucht.

2. Da die λ_i in $J(\Phi)$ nicht notwendigerweise verschieden sind, hat Φ womöglich weniger als k verschiedene Eigenwerte, etwa $\lambda_1, \dots, \lambda_m$ mit $m \leq k$. Taucht ein Eigenwert λ mit Vielfachheit α auf der Diagonalen von $J(\Phi)$ bzw. D auf, so hat D einen Eigenraum $V_\lambda \subset V$ der Dimension α .
3. Die Jordan-Kästchen zum selben Eigenwert λ können (notfalls durch Umsortieren der Basis) zu einem **Jordan-Block** zusammengefasst werden:

$$\hat{J}_{(m_1, \dots, m_r)}(\lambda) = \begin{pmatrix} & & & & \\ & \boxed{\begin{matrix} \lambda & 1 \\ \lambda & 1 \\ \ddots & \ddots \\ & \lambda & 1 \\ & & \lambda \end{matrix}} & & & \\ & & \ddots & & \\ & & & \boxed{\begin{matrix} \lambda & 1 \\ \lambda & 1 \\ \ddots & \ddots \\ & \lambda & 1 \\ & & \lambda \end{matrix}} & & \\ & & & & \end{pmatrix} \in \mathbb{C}^{(m_1 + \dots + m_r) \times (m_1 + \dots + m_r)},$$

oder kürzer

$$\hat{J}_{(m_1, \dots, m_r)}(\lambda) = J_{m_1}(\lambda) \oplus \dots \oplus J_{m_r}(\lambda) \in \mathbb{C}^{\alpha \times \alpha}$$

mit $m_1 + \dots + m_r = \alpha$. Wie man anhand der Normalform $J(\Phi)$ sieht, ist V_λ invariant unter Φ und die Einschränkung $\Phi|_{V_\lambda}$ wird durch den Jordan-Block $\hat{J}_{(m_1, \dots, m_r)}(\lambda)$ dargestellt in einer geeigneten Basis $B_\lambda \subset B$ von V_λ .

4. Wenn wir für jeden Eigenwert λ_i diese Basis B_{λ_i} von V_{λ_i} kennen, erhalten wir die gesuchte Basis B von V als deren Vereinigung

$$B = B_{\lambda_1} \cup \dots \cup B_{\lambda_m}.$$

5. Um für einen gegebenen Eigenwert λ die Basis B_λ zu bestimmen, betrachten wir den Endomorphismus

$$\Phi - \lambda \cdot \text{id}_V.$$

Die Einschränkung $(\Phi - \lambda \cdot \text{id}_V)|_{V_\lambda}$ wird durch den Jordan-Block

$$\hat{J}_{(m_1, \dots, m_k)}(0) = \hat{J}_{(m_1, \dots, m_k)}(\lambda) - \lambda I_\alpha$$

dargestellt und ist folglich nilpotent von Stufe $\leq \alpha$. Wir erhalten

$$V_\lambda = \text{Kern}(\Phi - \lambda \cdot \text{id}_V)^\alpha.$$

Damit ist die Basis B_λ zwar noch nicht festgelegt, aber das Problem der Bestimmung der Jordanschen Normalform ist reduziert auf den Fall nilpotenter Endomorphismen. Dieses Problem wird in Abschnitt 2.5 untersucht.

Für eine Matrix $A \in \mathbb{C}^{n \times n}$ definieren wir die **Jordansche Normalform** $J(A)$ als die Jordansche Normalform des Endomorphismus $\Phi(x) = Ax$.

2.4 Die Primärzerlegung

In diesem Abschnitt betrachten wir zunächst Vektorräume V endlicher Dimension über einem beliebigen Körper \mathbb{K} .

Satz 2.20 Es sei $\Phi \in \text{End}(V)$. Weiter seien $g, h \in \mathbb{K}[X]$ teilerfremd und $f = gh$, so dass $f(\Phi) = 0$ gilt. Setze $U = \text{Kern } g(\Phi)$ und $W = \text{Kern } h(\Phi)$. Dann gilt

- (a) $V = U \oplus W$.
- (b) $U = \text{Bild } h(\Phi)$ und $W = \text{Bild } g(\Phi)$.
- (c) U und W sind Φ -invariant.

BEWEIS: Da g und h teilerfremd sind, können wir mit dem erweiterten euklidischen Algorithmus Polynome $r, s \in \mathbb{K}[X]$ bestimmen, so dass

$$1 = rg + sh$$

gilt. Setzen wir Φ ein, so bedeutet dies

$$\begin{aligned} \text{id}_V &= r(\Phi) \circ g(\Phi) + s(\Phi) \circ h(\Phi) \\ &= g(\Phi) \circ r(\Phi) + h(\Phi) \circ s(\Phi). \end{aligned} \tag{*}$$

- (a) Nach Voraussetzung ist $0 = f(\Phi) \circ r(\Phi) = (hg)(\Phi) \circ r(\Phi) = h(\Phi) \circ (gr)(\Phi)$ und analog $0 = g(\Phi) \circ (hs)(\Phi)$. Durch Einsetzen von $x \in V$ in (*) sieht man

$$x = \underbrace{g(\Phi)(r(\Phi)(x))}_{\in \text{Kern } h(\Phi)} + \underbrace{h(\Phi)(s(\Phi)(x))}_{\in \text{Kern } g(\Phi)},$$

also $V = U + W$. Ist $x \in U \cap W$, also $g(\Phi)(x) = 0 = h(\Phi)(x)$, so liefert Einsetzen in (*) $x = 0$, also $V = U \oplus W$.

- (b) Ist $x \in U$, so bedeutet (*)

$$x = h(\Phi)(s(\Phi)(x)) \in \text{Bild } h(\Phi).$$

Umgekehrt bedeutet $x = h(\Phi)(y)$, dass gilt

$$g(\Phi)(x) = g(\Phi)(h(\Phi)(y)) = ((gh)(\Phi))(y) = f(\Phi)(y) = 0,$$

d.h. $x \in U$. Also $\text{Bild } h(\Phi) = U$. Analog zeigt man $\text{Bild } g(\Phi) = W$.

- (c) Nach (b) ist $\Phi(U) = \Phi(h(\Phi)(V)) = h(\Phi)(\Phi(V)) \subset h(\Phi)(V) = U$ und analog $\Phi(W) \subset W$. ■

Nach dem Satz über die Primfaktorzerlegung (Satz 1.34) können wir ein normiertes Polynom $f \in \mathbb{K}[X]$ als Produkt von Potenzen irreduzibler normierter Polynome p_i schreiben,

$$f = p_1^{\alpha_1} \cdots p_m^{\alpha_m}.$$

Folgerung 2.21 Es sei $\Phi \in \text{End}(V)$ und $f \in \mathbb{K}[X]$. Weiter sei $f = q_1 \cdots q_m$ eine Faktorisierung mit paarweise teilerfremden $q_i \in \mathbb{K}[X]$. Dann gilt:

$$\text{Kern } f(\Phi) = \text{Kern } q_1(\Phi) \oplus \dots \oplus \text{Kern } q_m(\Phi). \quad (2.3)$$

BEWEIS: Offensichtlich ist $\text{Kern } f(\Phi)$ ein Φ -invarianter Untervektorraum und es gilt $f(\Phi)|_{\text{Kern } f(\Phi)} = 0$. Nach Satz 2.20 (angewandt auf den Vektorraum $\text{Kern } f(\Phi)$) folgt zunächst

$$\text{Kern } f(\Phi) = \text{Kern } q_1(\Phi) \oplus \text{Kern}(q_2 \cdots q_m)(\Phi).$$

Mit Induktion über m folgt nun

$$\text{Kern}(q_2 \cdots q_m)(\Phi) = \text{Kern } q_2(\Phi) \oplus \dots \oplus \text{Kern } q_m(\Phi),$$

und somit die Behauptung. ■

Satz 2.22 Es sei V ein \mathbb{K} -Vektorraum der Dimension n und $\Phi \in \text{End}(V)$. Weiter sei $f_\Phi = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ die Primfaktorzerlegung des charakteristischen Polynoms von Φ . Dann ist

$$V = V_1 \oplus \dots \oplus V_m \quad (2.4)$$

mit $V_i = \text{Kern } p_i^{\alpha_i}(\Phi)$ eine Zerlegung von V in Φ -invariante Unterräume.

BEWEIS: Nach dem Satz von Cayley-Hamilton gilt $V = \text{Kern } f_\Phi(\Phi) = \text{Kern } 0$. Der Satz folgt somit aus Folgerung 2.21, angewandt auf $f = f_\Phi$ und $q_i = p_i^{\alpha_i}$. ■

Die Zerlegung (2.4) von V in Φ -invariante Unterräume entspricht einer Vergrößerung der Primfaktorzerlegung von f_Φ und wird daher **Primärzerlegung** von V bzgl. Φ genannt.

Folgerung 2.23 Sei $\Phi \in \text{End}(V)$. Jeder Primfaktor des charakteristischen Polynoms f_Φ von Φ ist auch ein Primfaktor des Minimalpolynoms h_Φ von Φ .

BEWEIS: Ist $V = V_1 \oplus \dots \oplus V_m$ die Primärzerlegung von V bzgl. Φ , so ist $p_i^{\alpha_i}(\Phi|_{V_i}) = 0$ nach Definition von V_i . Somit ist das Minimalpolynom $h_i = h_{\Phi|_{V_i}}$ von $\Phi|_{V_i}$ ein Teiler von $p_i^{\alpha_i}$, insbesondere ist $h_i = p_i^{k_i}$ für ein $k_i \leq \alpha_i$. Es gilt aber auch $h_\Phi(\Phi|_{V_i}) = 0$ und somit ist h_i ein Teiler von h_Φ . ■

Aus dem Beweis von Folgerung 2.23 ergibt sich, dass $p_i^{k_i}(\Phi|_{V_i}) = 0$ gilt, wobei k_i der Exponent von p_i im Minimalpolynom ist. Zugleich bedeutet das, dass $p_i^r(\Phi|_{V_i}) \neq 0$ ist für $r < k_i$. Dies erlaubt folgenden Schluss:

Folgerung 2.24 Die Aussage von Satz 2.22 bleibt gültig, wenn f_Φ durch das Minimalpolynom h_Φ ersetzt wird. Insbesondere ist $V_i = \text{Kern } p_i^{k_i}(\Phi)$.

Bemerkung 2.25 Sind B_1, \dots, B_m Basen der invarianten Unterräume V_1, \dots, V_m und $B = B_1 \cup \dots \cup B_m$ eine Basis von V , so ist

$$M_B^B(\Phi) = \begin{pmatrix} M_{B_1}^{B_1}(\Phi|_{V_1}) & & & \\ & M_{B_2}^{B_2}(\Phi|_{V_2}) & & \\ & & \ddots & \\ & & & M_{B_m}^{B_m}(\Phi|_{V_m}) \end{pmatrix}.$$

Nun betrachten wir die Primärzerlegung speziell für den Fall $\mathbb{K} = \mathbb{C}$.

Nach Folgerung 1.39 sind die irreduziblen Polynome in $\mathbb{C}[X]$ die Polynome $X - \lambda$ vom Grad 1. Die Primfaktorzerlegung des charakteristischen Polynoms von Φ ist folglich ein Produkt von Linearfaktoren $p_i = X - \lambda_i$,

$$f_\Phi = (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_m)^{\alpha_m},$$

wobei die λ_i die paarweise verschiedenen Eigenwerte von Φ mit algebraischer Vielfachheit α_i sind. Nach Folgerung 2.23 ist das Minimalpolynom ebenfalls ein Produkt dieser Linearfaktoren,

$$h_\Phi = (X - \lambda_1)^{k_1} \cdots (X - \lambda_m)^{k_m}$$

mit $0 < k_i \leq \alpha_i$ für $i = 1, \dots, m$. In der Primärzerlegung (2.4) sind die Φ -invarianten Unterräume V_i somit

$$V_i = \text{Kern } p_i^{k_i}(\Phi) = \text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^{k_i}.$$

Jeder Raum V_i ist also durch einen bestimmten Eigenwert festgelegt. Ist $k_i = 1$, so stimmt V_i sogar mit dem Eigenraum $E_{\lambda_i} = \text{Kern}(X - \lambda_i \cdot \text{id}_V)$ überein. Wir definieren daher:

Definition 2.26 Es sei V ein \mathbb{C} -Vektorraum der Dimension n und $\Phi \in \text{End}(V)$. Ist λ ein Eigenwert von Φ und k der Exponent von $X - \lambda$ im Minimalpolynom von Φ , so heißt

$$V_\lambda = \text{Kern}(\Phi - \lambda \cdot \text{id}_V)^k \quad (2.5)$$

der **verallgemeinerte Eigenraum** (oder **Hauptraum**) von Φ zum Eigenwert λ .

Satz 2.22 lautet im komplexen Fall:

Satz 2.27 Es sei V ein \mathbb{C} -Vektorraum der Dimension n und $\Phi \in \text{End}(V)$. Weiter seien $\lambda_1, \dots, \lambda_m$ die paarweise verschiedenen Eigenwerte von Φ . Dann ist

$$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m} \quad (2.6)$$

wobei V_{λ_i} der verallgemeinerte Eigenraum zum Eigenwert λ_i ist.

Hilfssatz 2.28 Für $i = 1, \dots, m$ gilt:

- (a) $E_{\lambda_i} \subseteq V_{\lambda_i}$.
- (b) $\Phi|_{V_{\lambda_i}} - \lambda_i \cdot \text{id}_{V_{\lambda_i}}$ ist nilpotent von Stufe k_i .
- (c) $\dim V_{\lambda_i} = \alpha_i$.

BEWEIS:

- (a) Klar.
- (b) Das Minimalpolynom von $\Phi|_{V_{\lambda_i}}$ ist $h_i = (X - \lambda_i)^{k_i}$. Die Behauptung folgt, da $h_i(\Phi|_{V_{\lambda_i}}) = 0$ gilt und h_i das Polynom kleinsten Grades mit dieser Eigenschaft ist.

- (c) Da $\Phi|_{V_{\lambda_i}} - \lambda_i \cdot \text{id}_{V_{\lambda_i}}$ nach (b) nilpotent ist, hat es nach Satz 2.12 das charakteristische Polynom $\det(X \cdot \text{id}_{V_{\lambda_i}} - \Phi|_{V_{\lambda_i}} + \lambda_i \cdot \text{id}_{V_{\lambda_i}}) = X^{\dim V_{\lambda_i}}$. Das bedeutet

$$f_i = \det(X \cdot \text{id}_{V_{\lambda_i}} - \Phi|_{V_{\lambda_i}}) = (X - \lambda_i)^{\dim V_{\lambda_i}}$$

ist das charakteristische Polynom von $\Phi|_{V_{\lambda_i}}$. Da außerdem $f_\Phi = f_1 \cdots f_i \cdots f_m$ gilt und der Linearfaktor $X - \lambda_i$ mit Exponent α_i in f_Φ und gleichzeitig in keinem anderen f_j auftaucht, folgt $\dim V_{\lambda_i} = \alpha_i$. ■

Bemerkung 2.29 Sind $B_{\lambda_1}, \dots, B_{\lambda_m}$ beliebige Basen der verallgemeinerten Eigenräume $V_{\lambda_1}, \dots, V_{\lambda_m}$, und A_1, \dots, A_m die Abbildungsmatrizen der jeweiligen $\Phi|_{V_{\lambda_i}}$ bzgl. B_i , so ist

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} = A_1 \oplus A_2 \oplus \dots \oplus A_m \quad (2.7)$$

die Abbildungsmatrix von Φ bzgl. der Basis $B = B_{\lambda_1} \cup \dots \cup B_{\lambda_m}$. Dies ist eine grobe Vorstufe der Jordanschen Normalform. Hierbei entsprechen die Blöcke A_i den Jordan-Blöcken $\hat{J}(\lambda_i)$. Damit tatsächlich $A_i = \hat{J}(\lambda_i)$ gilt, müssen spezielle Basen B_{λ_i} gewählt werden. Dafür untersucht man für jeden Eigenwert λ_i die Einschränkung $\Phi|_{V_{\lambda_i}}$. Da jede Abbildungsmatrix A_i von $\Phi|_{V_{\lambda_i}}$ aber eindeutig einer Abbildungsmatrix $A_i - \lambda_i I_{\alpha_i}$ des nilpotenten Endomorphismus $\Phi|_{V_{\lambda_i}} - \lambda_i \cdot \text{id}_{V_{\lambda_i}}$ entspricht, kann man stattdessen auch diesen untersuchen. Ist nämlich $A_i - \lambda_i I_{\alpha_i} = \hat{J}(0)$ ein Jordan-Block, so ist auch $A_i = \hat{J}(0) + \lambda_i I_{\alpha_i} = \hat{J}(\lambda_i)$ ein Jordan-Block. Das hat den Vorteil, dass man sich auf nilpotente Endomorphismen beschränken kann.

2.5 Die Normalform nilpotenter Endomorphismen

In diesem Abschnitt werden wir zeigen, dass jeder nilpotente Endomorphismus Φ eines endlichdimensionalen \mathbb{K} -Vektorraums V eine Abbildungsmatrix A der Form

$$A = \hat{J}_{(m_1, \dots, m_r)}(0)$$

besitzt. Im Folgenden sei Φ nilpotent von Stufe k (d.h. $\Phi^k = 0$ und $\Phi^{k-1} \neq 0$).

Definition 2.30 Ein Element $x \in V$ hat die **Stufe j** (mit $j \geq 1$), falls gilt:

$$\Phi^j(x) = 0 \quad \text{und} \quad \Phi^{j-1}(x) \neq 0.$$

Da Φ nilpotent ist, ist 0 der einzige Eigenwert von Φ . Offensichtlich sind die Eigenvektoren von Φ zum Eigenwert 0 die Elemente der Stufe 1. Da $\Phi^{k-1} \neq 0$, existieren Elemente der Stufe k (aber keiner höheren Stufe, da $\Phi^k = 0$).

Hilfssatz 2.31 *Es sei $\Phi \in \text{End}(V)$ nilpotent und x ein Element der Stufe j . Dann sind die Elemente*

$$\Phi^{j-1}(x), \dots, \Phi(x), x$$

linear unabhängig.

BEWEIS: Es gelte

$$0 = \lambda_0 x + \lambda_1 \Phi(x) + \dots + \lambda_{j-1} \Phi^{j-1}(x) \quad (*)$$

mit $\lambda_0, \dots, \lambda_{j-1} \in \mathbb{K}$. Wende darauf Φ^{j-1} an:

$$0 = \Phi(0) = \lambda_0 \underbrace{\Phi^{j-1}(x)}_{=0} + \lambda_1 \underbrace{\Phi^j(x)}_{=0} + \dots + \lambda_{j-1} \underbrace{\Phi^{2j-2}(x)}_{\neq 0} = \lambda_0 \underbrace{\Phi^{j-1}(x)}_{\neq 0}.$$

Also ist $\lambda_0 = 0$ und $(*)$ wird zu

$$0 = \lambda_1 \Phi(x) + \dots + \lambda_{j-1} \Phi^{j-1}(x).$$

Wende darauf nun Φ^{j-2} an, um $\lambda_1 = 0$ zu erhalten. Dies setzt man fort, bis man schließlich

$$\lambda_0 = \lambda_1 = \dots = \lambda_{j-1} = 0$$

erhält. Das bedeutet,

$$x, \Phi(x), \dots, \Phi^{j-1}(x)$$

sind linear unabhängig. ■

Daraus folgt sofort:

Folgerung 2.32 *Ist $x \in V$ ein Element der Stufe j , so ist*

$$B_x = \{\Phi^{j-1}(x), \dots, \Phi(x), x\} \quad (2.8)$$

eine Basis des Φ -invarianten Untervektorraums

$$V_x = [\Phi^{j-1}(x), \dots, \Phi(x), x]. \quad (2.9)$$

Die Abbildungsmatrix von $\Phi|_{V_x}$ bzgl. B_x ist

$$J_j(0) = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathbb{K}^{j \times j}. \quad (2.10)$$

Wir nennen x einen **Erzeuger** von V_x .

Beispiel 2.33

- (a) Es sei $\Phi \in \text{End}(\mathbb{R}^4)$ gegeben durch

$$\Phi(x) = \left(\begin{array}{c|cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) \cdot x.$$

Die Matrix ist

$$\hat{J}_{(1,3)}(0) = J_1(0) \oplus J_3(0)$$

und dies ist bereits die Jordansche Normalform von Φ . Der vierte Einheitsvektor e_4 ist ein Erzeuger der Stufe 3, die Einheitsvektoren e_1 und e_2 sind Eigenvektoren, also Erzeuger der Stufe 1. Weiter sind $V_{e_1} = [e_1]$ und

$$V_{e_4} = [e_4, e_3, e_2] \supset V_{e_3} = [e_3, e_2] \supset V_{e_2} = [e_2].$$

- (b) Ist Φ durch

$$\Phi(x) = \left(\begin{array}{c|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) \cdot x$$

gegeben, so gibt es keine Erzeuger der Stufe 3 oder 4, da Φ Nilpotenzstufe 2 hat. Auch hier ist die Matrix bereits die Jordansche Normalform

$$\hat{J}_{(2,2)}(0) = J_2(0) \oplus J_2(0).$$

Erzeuger der Stufe 2 sind z.B. e_2 und e_4 . Hier sind

$$\begin{aligned} V_{e_2} &= [e_2, e_1] \supset V_{e_1} = [e_1], \\ V_{e_4} &= [e_4, e_3] \supset V_{e_3} = [e_3]. \end{aligned}$$

- (c) Ist Φ durch

$$\Phi(x) = \left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) \cdot x$$

gegeben, so hat Φ die Nilpotenzstufe 4 und die Matrix ist bereits die Jordansche Normalform. Ein Erzeuger der Stufe 4 ist e_4 . Es ist

$$\mathbb{R}^4 = V_{e_4} = [e_4, e_3, e_2, e_1] \supset V_{e_3} = [e_3, e_2, e_1] \supset V_{e_2} = [e_2, e_1] \supset V_{e_1} = [e_1].$$

Wir wollen nun zeigen, dass V eine Basis B besitzt, die eine Vereinigung von Basen B_{x_1}, \dots, B_{x_r} vom Typ (2.8) ist. Dann nimmt die Abbildungsmatrix $M_B^B(\Phi)$ die gewünschte Form an:

$$\hat{J}_{(m_1, \dots, m_r)}(0) = \begin{pmatrix} J_{m_1}(0) & & \\ & \ddots & \\ & & J_{m_r}(0) \end{pmatrix}$$

Im Folgenden beschreiben wir die Konstruktion dieser Zerlegung und sehen dabei, wie die Basis B zu bestimmen ist. Wir schreiben abkürzend

$$K^j = \text{Kern } \Phi^j. \quad (2.11)$$

Wegen der Nilpotenz gilt

$$\{0\} = K^0 \subset K^1 \subset K^2 \subset \dots \subset K^{k-1} \subset K^k = V \quad (2.12)$$

und für alle j

$$\Phi(K^j) \subset K^{j-1}. \quad (2.13)$$

Die Elemente der Stufe j sind genau die Elemente der Menge $K^j \setminus K^{j-1}$.

Wähle einen Untervektorraum $W_k \subset V$, so dass gilt

$$V = W_k \oplus K^{k-1}.$$

Dann besteht W_k aus Elementen der Stufe k . Nun können wir auch in K^{k-1} einen Untervektorraum W_{k-1} wählen, so dass gilt

$$K^{k-1} = W_{k-1} \oplus K^{k-2}$$

und W_{k-1} aus Elementen der Stufe $k-1$ besteht. Fahren wir nach diesem Schema fort, so erhalten wir eine Zerlegung von V der Form

$$\begin{aligned} V &= W_k \oplus K^{k-1} \\ &= W_k \oplus W_{k-1} \oplus K^{k-2} \\ &\quad \vdots \\ &= W_k \oplus W_{k-1} \oplus \dots \oplus W_2 \oplus K^1 \\ &= W_k \oplus W_{k-1} \oplus \dots \oplus W_2 \oplus W_1. \end{aligned}$$

Anhand dieser Zerlegung kann man nun die gesuchte Basis von V konstruieren:

Hilfssatz 2.34 Ist W ein Untervektorraum von V mit $W \cap K^j = \{0\}$ für ein $j > 0$, so ist $\Phi|_W$ injektiv.

Beweis: Für $j > 0$ ist $\text{Kern } \Phi = K^1 \subset K^j$. Somit folgt aus der Voraussetzung $\text{Kern } \Phi|_W = \{0\}$, d.h. $\Phi|_W$ ist injektiv. ■

Für $j = 1, \dots, k$ sei

$$d_j = \dim W_j.$$

Es sei $b_1^{[k]}, \dots, b_{d_k}^{[k]}$ eine Basis von W_k . Da $\Phi|_{W_k}$ nach Hilfssatz 2.34 injektiv ist, sind die Elemente

$$\begin{aligned} &\Phi^{k-1}(b_1^{[k]}), \dots, \Phi(b_1^{[k]}), b_1^{[k]}, \\ &\Phi^{k-1}(b_2^{[k]}), \dots, \Phi(b_2^{[k]}), b_2^{[k]}, \\ &\vdots \\ &\Phi^{k-1}(b_{d_k}^{[k]}), \dots, \Phi(b_{d_k}^{[k]}), b_{d_k}^{[k]} \end{aligned}$$

linear unabhängig. Außerdem gilt $\Phi'(b_i^{[k]}) \in W_{k-l}$, insbesondere sind

$$\Phi(b_1^{[k]}), \dots, \Phi(b_{d_k}^{[k]})$$

linear unabhängig in W_{k-1} . Falls diese Elemente noch keine Basis von W_{k-1} bilden, können wir sie durch geeignete Elemente

$$b_1^{[k-1]}, \dots, b_{s_{k-1}}^{[k-1]}$$

zu einer Basis von W_{k-1} ergänzen, wobei $s_{k-1} = d_{k-1} - d_k$. Es folgt aus Hilfssatz 2.34, dass die

$$\Phi^{k-2}(b_i^{[k-1]}), \dots, \Phi(b_i^{[k-1]}), b_i^{[k-1]}, \quad i = 1, \dots, s_{k-1}$$

linear unabhängig sind. So fährt man fort, bis man schließlich eine Basis für V nach dem folgenden Schema konstruiert hat:

W_k	$b_1^{[k]}$	\dots	$b_{s_k}^{[k]}$					
W_{k-1}	$\Phi(b_1^{[k]})$	\dots	$\Phi(b_{s_k}^{[k]})$	$b_1^{[k-1]}$	\dots	$b_{s_{k-1}}^{[k-1]}$		
\vdots	\vdots	\vdots						
W_1	$\Phi^{k-1}(b_1^{[k]})$	\dots	$\Phi^{k-1}(b_{s_k}^{[k]})$	$\Phi^{k-2}(b_1^{[k-1]})$	\dots	$\Phi^{k-2}(b_{s_{k-1}}^{[k-1]})$	\dots	$b_1^{[1]}$
								\dots
								$b_{s_1}^{[1]}$

Hier ist

$$s_k = d_k \quad \text{und} \quad s_i = d_i - s_{i+1} - s_{i+2} - \dots - s_k. \quad (2.14)$$

In einer **Jordan-Basis** B für Φ ordnen wir diese Basisvektoren zunächst von unten nach oben und dann von links nach rechts:

$$\begin{aligned} B &= \{\Phi^{k-1}(b_1^{[k]}), \dots, b_1^{[k]}, \dots, \Phi^{k-1}(b_{s_k}^{[k]}), \dots, b_{s_k}^{[k]}, \Phi^{k-2}(b_1^{[k-1]}), \dots, b_1^{[k-1]}, \dots, b_1^{[1]}, \dots, b_{s_1}^{[1]}\} \\ &= B_{b_1^{[k]}} \cup B_{b_{s_k}^{[k]}} \cup B_{b_1^{[k-1]}} \cup \dots \cup B_{b_1^{[1]}} \cup B_{b_{s_1}^{[1]}}. \end{aligned}$$

Satz 2.35 Die Abbildungsmatrix $M_B^B(\Phi)$ von Φ ist die Jordansche Normalform $J(\Phi)$:

$$\begin{pmatrix} J_k(0) & & & & \\ \ddots & s_k & & & \\ & J_k(0) & & & \\ & & J_{k-1}(0) & & \\ & & & \ddots & \\ & & & & J_{k-1}(0) \\ & & & & & \ddots & \\ & & & & & & 0 \\ & & & & & & & s_1 \\ & & & & & & & & 0 \end{pmatrix}. \quad (2.15)$$

Diese Normalform ist eindeutig bis auf die Reihenfolge der Jordan-Kästchen.

Beweis: Dass $M_B^B(\Phi) = J(\Phi)$ gilt, ergibt sich aus der vorangehenden Diskussion. Die Eindeutigkeit (bis auf die Reihenfolge) folgt daraus, dass die Zahlen s_k, \dots, s_1 gemäß (2.14) eindeutig durch die Dimensionen der Kerne K^i bestimmt sind (da $d_i = \dim W_i = \dim K^j - \dim K^{j-1}$ gilt). Diese Kerne wiederum sind eindeutig durch Φ bestimmt. ■

Bemerkung 2.36 Da $K^1 = W_1$ der Eigenraum E_0 von Φ zum Eigenwert 0 ist, gilt

$$\begin{aligned} \dim E_0 &= s_k + s_{k-1} + \dots + s_1 \\ &= \text{Anzahl der Jordan-Kästchen in } J(\Phi). \end{aligned}$$

Des Weiteren gilt $\Phi^{k-1} \neq 0$, daher muss mindestens ein Jordan-Kästchen $J_k(0)$ der Größe k in $J(\Phi)$ vorkommen. Da $\Phi^k = 0$, darf aber kein größeres Jordan-Kästchen auftauchen. Dies können wir mit dem Minimalpolynom $h_\Phi = X^k$ charakterisieren:

$$\text{Größe des größten Jordan-Kästchens} = k = \deg h_\Phi.$$

Als Folgerung aus Satz 2.35 erhalten wir sofort das entsprechende Resultat für nilpotente Matrizen:

Satz 2.37 Jede nilpotente Matrix $A \in \mathbb{K}^{n \times n}$ ist konjugiert zu einer Matrix der Form (2.15). Die Konjugationsklasse von A ist durch die Zahlen $(s_k, s_{k-1}, \dots, s_1)$ eindeutig festgelegt. Insbesondere gibt es nur endlich viele Konjugationsklassen von nilpotenten Matrizen in $\mathbb{K}^{n \times n}$.

2.6 Die Jordansche Normalform

Wir verbinden nun die Ergebnisse aus den Abschnitten 2.4 und 2.5.

Es sei V ein \mathbb{C} -Vektorraum der Dimension n und $\Phi \in \text{End}(V)$. Weiter seien $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte von Φ , $V_{\lambda_1}, \dots, V_{\lambda_m}$ ihre jeweiligen verallgemeinerten Eigenräume und $\alpha_1, \dots, \alpha_m$ ihre algebraischen Vielfachheiten (insbesondere $\alpha_i = \dim V_{\lambda_i}$).

Mit Ψ_j bezeichnen wir die nach Hilfssatz 2.28 nilpotenten Endomorphismen

$$\Psi_j = \Phi|_{V_{\lambda_j}} - \lambda_j \cdot \text{id}_{V_{\lambda_j}}.$$

Satz 2.38 Es gibt eine Basis B von V , so dass die Abbildungsmatrix $M_B^B(\Phi)$ von Φ von der folgende Form ist:

$$J(\Phi) = \begin{pmatrix} \lambda_1 I_{\alpha_1} + J(\Psi_1) & & & \\ & \ddots & & \\ & & \lambda_m I_{\alpha_m} + J(\Psi_m). \end{pmatrix} \quad (2.16)$$

Diese Normalform ist eindeutig bis auf die Reihenfolge der λ_i und der Jordan-Kästchen in den $J(\Psi_i)$.

BEWEIS: Gemäß Satz 2.22 ist die Primärzerlegung $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m}$ durch die Primfaktorzerlegung des charakteristischen Polynoms eindeutig bestimmt (bis auf die Reihenfolge). Da die Zerlegung Φ -invariant ist, reicht es, die Einschränkungen von Φ auf die V_{λ_i} zu betrachten. Für diese gilt $\Phi|_{V_{\lambda_i}} = \lambda_i \cdot \text{id}_{V_{\lambda_i}} + \Psi_i$. Für den nilpotenten Endomorphismus wählen wir eine Jordan-Basis B_i von V_{λ_i} wie in Satz 2.35. Da $\lambda_i \cdot \text{id}_{V_{\lambda_i}}$ in jeder Basis durch $\lambda_i I_{\alpha_i}$ dargestellt wird, folgt, dass $\Phi|_{V_{\lambda_i}}$ bzgl. der Basis B_i die Abbildungsmatrix

$$J(\Phi|_{V_{\lambda_i}}) = \lambda_i I_{\alpha_i} + J(\Psi_i)$$

hat. Diese wiederum ist nach Satz 2.35 eindeutig (bis auf die Reihenfolge). Für die Basis $B = B_1 \cup \dots \cup B_m$ von V hat Φ dann die Abbildungsmatrix (2.16). ■

Die Basis B aus Satz 2.38 nennen wir eine **Jordan-Basis** von V für Φ .

Bemerkung 2.39 Unter Berücksichtigung von Bemerkung 2.36 gilt

algebraische Vielfachheit $\alpha_i =$ Größe des Jordan-Blocks zum Eigenwert λ_i ,

$\dim E_{\lambda_i} =$ Anzahl der Jordan-Kästchen zum Eigenwert λ_i in $J(\Phi)$,

Exponent k_i von $X - \lambda_i$ in $h_\Phi =$ Größe des größten Jordan-Kästchens zum Eigenwert λ_i .

Wir formulieren Satz 2.38 für komplexe Matrizen:

Folgerung 2.40 Jede Matrix $A \in \mathbb{C}^{n \times n}$ ist konjugiert zu einer Matrix der Form (2.16). Die Konjugationsklasse von A ist durch die Jordansche Normalform von A eindeutig festgelegt.

Folgerung 2.41 Es gibt unendlich viele Konjugationsklassen von Matrizen in $\mathbb{C}^{n \times n}$. Jedoch gibt es nur endlich viele Konjugationsklassen von Matrizen in $\mathbb{C}^{n \times n}$ mit dem selben charakteristischen Polynom.

Eine sehr umfangreiche Untersuchung der Konjugationsklassen von Matrizen über beliebigen Körpern ist in den Abschnitten 11.6 und 11.7 von Brieskorn [2], zu finden.

Bemerkung 2.42 Diagonalmatrizen sind eine Spezialfall der Jordanschen Normalform. Hier haben alle Jordan-Kästchen die Größe 1, und die verallgemeinerten Eigenräume stimmen mit den Eigenräumen überein.

Algorithmus 2.43 Zur praktischen Bestimmung einer Jordan-Basis für Φ geht man wie folgt vor:

- (i) Bestimme das charakteristische Polynom f_Φ und seine Nullstellen $\lambda_1, \dots, \lambda_m$, die Eigenwerte von Φ .
- (ii) Für $i = 1, \dots, m$ führe folgende Schritte aus:
 - (ii.a) Bestimme die kleinste Zahl k_i , so dass

$$\text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^{k_i} = \text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^{k_i+1}.$$

Dann gilt auch $\text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^{k_i} = \text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^n = V_{\lambda_i}$. Die Zahl k_i ist der Exponent von $X - \lambda_i$ im Minimalpolynom h_Φ .

Wenn man die Kerne mit dem Gauß-Algorithmus bestimmt, erhält man hierbei für $s = 1, \dots, k_i$ eine Basis $B_i^{[s]}$ von $\text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^s$, die aus Elementen von Stufe $\leq s$ besteht. Insbesondere ist $B_i^{[k_i]}$ eine Basis von V_{λ_i} (die in der Regel aber noch keine Jordan-Basis ist). Dies ermöglicht den nächsten Schritt.

- (ii.b) Es sei L_i eine zunächst leere Liste von Basisvektoren.

Für $s = k_i, \dots, 1$ führe die folgenden Schritte aus:

Solange ein Element $b \in V_{\lambda_i}$ von Stufe s existiert, dass nicht in der linearen Hülle des bisher in L_i gemerkten Elemente liegt, berechne

$$(\Phi - \lambda_i \cdot \text{id}_V)^j(b), \quad j = s-1, \dots, 0$$

und füge diese Elemente in dieser Reihenfolge zur Liste L_i hinzu. An dernfalls fahre mit Stufe $s - 1$ fort.

Am Ende enthält die Liste L_i eine Jordan-Basis von V_{λ_i} für $\Phi|_{V_{\lambda_i}}$.

- (iii) Fassen wir die Elemente der Listen L_1, \dots, L_m aus Schritt (ii) zu einer Basis B zusammen, so ist B eine Jordan-Basis von Φ . Die Jordansche Normalform erhält man dann als $J(\Phi) = M_B^B(\Phi)$.

Ist man nur an der Bestimmung von $J(\Phi)$ (ohne Basis) interessiert, so kann man Schritt (ii.b) weglassen und in Schritt (ii.a) anhand der Differenzen der Dimensionen der Kerne $\text{Kern}(\Phi - \lambda_i \cdot \text{id}_V)^j$, $j = 1, \dots, k_i$, prüfen, wieviele Jordan-Kästchen einer bestimmten Größe für den Eigenwert λ_i auftauchen.

Wir geben nun zwei Beispiele von Matrizen A und B , die zwar die selben Eigenwerte, das selbe charakteristische Polynom und das selbe Minimalpolynom haben, aber dennoch verschiedene Jordansche Normalformen.

Beispiel 2.44 Es sei

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{C}^{4 \times 4}.$$

- Das charakteristische Polynom von A ist

$$f_A = \det(XI_4 - A) = (X - 2)^4.$$

A hat also nur den Eigenwert 2 mit algebraischer Vielfachheit 4.

- Wir bestimmen mit dem Gauß-Algorithmus $\text{Kern}(A - 2 \cdot I_4)^j$ für $j = 1, 2, \dots$ und stellen fest, dass

$$\text{Kern}(A - 2 \cdot I_4) \neq \mathbb{C}^4, \quad \text{Kern}(A - 2 \cdot I_4)^2 = \mathbb{C}^4$$

gilt. Somit ist das Minimalpolynom $h_A = (X - 2)^2$.

- Wähle einen Vektor $b_1^{[2]} \in \mathbb{C}^4 \setminus \text{Kern}(A - 2 \cdot I_4)$, etwa $b_1^{[2]} = e_1$. Als die ersten beiden Vektoren unserer Jordan-Basis wählen wir

$$(A - 2 \cdot I_4) \cdot b_1^{[2]} = \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad b_1^{[2]} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Da $\dim \text{Kern}(A - 2 \cdot I_4)^2 - \dim \text{Kern}(A - 2 \cdot I_4) = 4 - 3 = 1$ gilt, existieren keine dazu linear unabhängigen Vektoren der Stufe 2.

- Fülle die Jordan-Basis mit den linear unabhängigen Vektoren $b_1^{[1]} = e_3$ und $b_2^{[1]} = e_4$ aus $\text{Kern}(A - 2 \cdot I_4)$ auf. Somit ist unsere Jordan-Basis für A :

$$\left\{ \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

- Die Jordansche Normalform von A ist

$$J(A) = \left(\begin{array}{cc|cc|c} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{array} \right).$$

Beispiel 2.45 Es sei

$$B = \begin{pmatrix} 9 & -7 & 0 & 2 \\ 7 & -5 & 0 & 2 \\ 4 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{C}^{4 \times 4}.$$

- Das charakteristische Polynom von B ist

$$f_B = \det(XI_4 - B) = (X - 2)^4.$$

B hat also nur den Eigenwert 2 mit algebraischer Vielfachheit 4.

- Wir bestimmen mit dem Gauß-Algorithmus $\text{Kern}(B - 2 \cdot I_4)^j$ für $j = 1, 2, \dots$ und stellen fest, dass

$$\text{Kern}(B - 2 \cdot I_4) \neq \mathbb{C}^4, \quad \text{Kern}(B - 2 \cdot I_4)^2 = \mathbb{C}^4$$

gilt. Somit ist das Minimalpolynom $h_B = (X - 2)^2$.

- Wähle einen Vektor $b_1^{[2]} \in \mathbb{C}^4 \setminus \text{Kern}(B - 2 \cdot I_4)$, etwa $b_1^{[2]} = e_1$. Als die ersten beiden Vektoren unserer Jordan-Basis wählen wir

$$(B - 2 \cdot I_4) \cdot b_1^{[2]} = \begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}, \quad b_1^{[2]} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es gilt $\dim \text{Kern}(B - 2 \cdot I_4)^2 - \dim \text{Kern}(B - 2 \cdot I_4) = 4 - 2 = 2$, also können wir noch einen dazu linear unabhängigen Vektoren der Stufe 2 finden.

- Der Vektor $b_2^{[2]} = e_4$ ist ebenfalls in $\mathbb{C}^4 \setminus \text{Kern}(B - 2 \cdot I_4)$ und offensichtlich linear unabhängig zu den beiden Vektoren aus dem vorigen Schritt. Wir ergänzen unsere Jordan-Basis durch die Vektoren

$$(B - 2 \cdot I_4) \cdot b_2^{[2]} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad b_2^{[2]} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

- Somit ist unsere Jordan-Basis für B :

$$\left\{ \begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

- Die Jordansche Normalform von B ist

$$J(B) = \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right).$$

3 Kryptographie

Alice möchte eine geheime Nachricht m an Bob verschicken. Sie kann jedoch nicht sicher sein, dass der Kanal, über den sie m verschickt, abhörsicher ist. Also muss sie eine Methode finden, die Nachricht derart zu verschlüsseln, dass nur Bob sie wieder entschlüsseln und lesen kann. Als **Kryptographie** bezeichnet man die Kunst, solche Methoden zu finden und zu untersuchen. Genauer untersucht man folgende Situation: Gegeben ist eine Menge \mathcal{M} von allgemein verständlichen Nachrichten („Klartext“) die durch eine Verschlüsselungsfunktion

$$\text{enc} : \mathcal{M} \rightarrow \mathcal{C}$$

in einen Code \mathcal{C} abgebildet werden sollen, der nicht allgemein verständlich ist („Chiffrat“). Der Empfänger einer Nachricht soll in der Lage sein, mit Hilfe einer Funktion

$$\text{dec} : \mathcal{C} \rightarrow \mathcal{M}$$

die codierte Nachricht wieder zu entschlüsseln und so den ursprünglichen Klartext zu erhalten. Es soll also

$$\text{dec} \circ \text{enc} = \text{id}_{\mathcal{M}}$$

gelten. Eine Möglichkeit, dies zu erreichen ist es, dass Alice und Bob Verfahren enc und dec verwenden, die sie gegenüber Dritten geheimhalten. Die Probleme dabei sind jedoch offensichtlich: Werden durch Verrat oder Nachlässigkeit die Methoden bekannt, so ist das Verfahren wertlos. Außerdem müsste man bei mehr als zwei Kommunikationsteilnehmern jeweils neue geheime Verfahren ersinnen.

Die Philosophie der **Public Key-Kryptographie** ist daher, öffentlich bekannte Algorithmen zu verwenden, bei denen die Geheimhaltung durch einen zusätzlichen Parameter, den **Schlüssel**, gewährleistet ist. Jede Partei besitzt einen eigenen Schlüssel. Der Schlüssel besteht aus zwei Teilen, dem **öffentlichen Schlüssel** für die Verschlüsselung und dem **geheimen Schlüssel** für die Entschlüsselung.

Die meisten Verfahren der Public Key-Kryptographie nutzen auf raffinierte Weise die algebraischen Eigenschaften der ganzen Zahlen (oder anderer integerer Ringe) aus. In diesem Abschnitt wollen wir Verfahren betrachten, die wir mit unserem bisherigen Kenntnisstand schon verstehen können. Als Vorbereitung studieren wir einige Eigenschaften von endlichen Gruppen und die Einheiten von $\mathbb{Z}/n\mathbb{Z}$.

In diesem Abschnitt gelte die Konvention, dass geheimzuhaltende Daten durch Maschinenschrift m, k, \dots dargestellt werden.

3.1 Der Satz von Lagrange

Definition 3.1 Es sei G eine Gruppe und $g \in G$. Die **Ordnung von G** ist $\text{ord}(G) = |G|$. Die **Ordnung von g** ist $\text{ord}(g) = \text{ord}(\langle g \rangle)$, also die Ordnung der von g erzeugten zyklischen Untergruppe $\langle g \rangle = \{1, g, g^2, g^3, \dots\}$.

Satz 3.2 (Lagrange) Ist G eine endliche Gruppe und H eine Untergruppe von G , so gilt $\text{ord}(H) \mid \text{ord}(G)$. Insbesondere gilt $\text{ord}(g) \mid \text{ord}(G)$ für alle $g \in G$.

BEWEIS: Für jedes $g \in G$ gilt $|H| = |gH|$, da $H \rightarrow gH, h \mapsto gh$ eine Bijektion ist.

Für $g_1, g_2 \in G$ gilt entweder $g_1H = g_2H$ oder $g_1H \cap g_2H = \emptyset$: Falls $a \in g_1H \cap g_2H$ existiert, so bedeutet dies $g_1h_1 = a = g_2h_2$ für gewisse $h_1, h_2 \in H$. Das bedeutet $g_1 = g_2h_2h_1^{-1}$. Wir können nun jedes Element $g_1h \in g_1H$ schreiben als

$$g_1h = g_2(h_2h_1^{-1}h) \in g_2H.$$

Also $g_1H \subset g_2H$ und entsprechend zeigt man $g_2H \subset g_1H$, also $g_1H = g_2H$.

Da jedes Element $g \in G$ in einem gH enthalten ist, folgt, dass G die disjunkte Vereinigung von gewissen endlichen Teilmengen g_1H, \dots, g_kH ist. Daher gilt

$$|G| = \left| \bigcup_{j=1}^k g_jH \right| = \sum_{j=1}^k |g_jH| = k|H|,$$

also ist $|H|$ ein Teiler von $|G|$. ■

Hilfssatz 3.3 Es sei G eine abelsche Gruppe und $g, h \in G$. Weiter gelte $\text{ord}(g) = m$, $\text{ord}(h) = n$ und $\text{ggT}(m, n) = 1$. Dann ist $\text{ord}(gh) = mn$.

BEWEIS: Es sei $r = \text{ord}(gh)$. Da G abelsch ist, gilt

$$(gh)^{mn} = (g^m)^n(h^n)^m = 1 \cdot 1 = 1,$$

d.h. $r \leq mn$. Nach Folgerung 1.29 gibt es $s, t \in \mathbb{Z}$ mit $sm + tn = 1$. Das bedeutet

$$g^r = (g^{sm+tn})^r = (\underbrace{g^m}_{=1})^{sr}(g^n)^{tr} = (g^n)^{tr} = (g^n h^n)^{tr} = ((gh)^r)^{tn} = 1.$$

Also ist m ein Teiler von r , und genauso zeigt man, dass n ein Teiler von r ist. Aus $\text{ggT}(m, n) = 1$ und der eindeutigen Primfaktorzerlegung folgt nun, dass mn ein Teiler von r ist, also $mn \leq r$. Insgesamt gilt also $mn = r$. ■

Hilfssatz 3.4 Es sei G eine endliche abelsche Gruppe und $g \in G$ ein Element von maximaler Ordnung $m = \text{ord}(g)$. Dann ist für jedes $h \in G$ die Ordnung $\text{ord}(h)$ ein Teiler von m .

BEWEIS: Es sei $n = \text{ord}(h)$. Angenommen, es gelte $n \nmid m$. Dann gibt es einen Primfaktor p , der in n einen größeren Exponenten hat als in m , etwa

$$m = p^r a, \quad n = p^s b$$

mit $r < s$ und $p \nmid a, b$. Das bedeutet $\text{ord}(g^{p^r}) = a$ und $\text{ord}(h^b) = p^s$. Da $\text{ggT}(p^s, a) = 1$, folgt mit Hilfssatz 3.3 $\text{ord}(g^{p^r} h^b) = p^s a$. Aber $p^s a > p^r a = m$, d.h. $g^{p^r} h^b$ hat höhere Ordnung als g , im Widerspruch zur Voraussetzung. ■

3.2 Einheiten in $\mathbb{Z}/n\mathbb{Z}$

Satz 3.5 Es gilt

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{\bar{k} \in \mathbb{Z}/n\mathbb{Z} \mid \text{ggT}(k, n) = 1\}. \quad (3.1)$$

BEWEIS: Sind k und n teilerfremd, so können wir mit dem erweiterten euklidischen Algorithmus Elemente $s, t \in \mathbb{Z}$ bestimmten, so dass gilt $sk + tn = 1$. Das bedeutet

$$sk \equiv 1 \pmod{n}.$$

Also ist $\bar{k} = \bar{s}^{-1}$ eine Einheit.

Setzen wir umgekehrt voraus, dass $sk \equiv 1 \pmod{n}$ gilt für ein gewisses $s \in \mathbb{Z}$, so bedeutet dies $sk = 1 + mn$ für ein $m \in \mathbb{Z}$. Ist d ein Teiler von k und n , etwa $ad = k$ und $bd = n$, so gilt $sad = 1 + mbd$. Das ist äquivalent zu

$$1 = (sa - mb)d.$$

Also ist $d \in \mathbb{Z}^\times = \{-1, 1\}$ und k, n sind folglich teilerfremd. ■

Definition 3.6 Die Funktion

$$\varphi : \mathbb{N} \rightarrow \mathbb{N}, \quad n \mapsto |(\mathbb{Z}/n\mathbb{Z})^\times| \quad (3.2)$$

heißt **Eulersche φ -Funktion**.

Nach Satz 3.5 ist $\varphi(n)$ die Anzahl der zu n teilerfremden Zahlen $0 < k < n$.

Satz 3.7 (Euler) Es sei $n \in \mathbb{N}$, $n \geq 2$. Dann gilt für jede zu n teilerfremde Zahl a :

$$a^{\varphi(n)} \equiv 1 \pmod{n}. \quad (3.3)$$

BEWEIS: Gilt $\text{ggT}(a, n) = 1$, so ist \bar{a} eine Einheit in $\mathbb{Z}/n\mathbb{Z}$. Es ist $\text{ord}((\mathbb{Z}/n\mathbb{Z})^\times) = \varphi(n)$ und somit gibt es nach dem Satz von Lagrange ein $k \in \mathbb{N}$ mit $k \cdot \text{ord}(\bar{a}) = \varphi(n)$. Das heißt

$$\bar{a}^{\varphi(n)} = (\bar{a}^{\text{ord}(\bar{a})})^k = 1^k = 1,$$

was äquivalent ist zu (3.3). ■

Ein Spezialfall des Satzes von Euler ist als *Satz von Fermat* bekannt.

Satz 3.8 (Fermat) *Es sei p eine Primzahl. Dann gilt für alle $a \in \mathbb{Z}$, die kein Vielfaches von p sind:*

$$a^{p-1} \equiv 1 \pmod{p}. \quad (3.4)$$

Bemerkung 3.9 Der Satz von Fermat liefert einen Test, ob eine Zahl q keine Primzahl ist. Ist nämlich (3.4) nicht erfüllt für irgendein $0 < a < q$, so ist q nicht prim. Umgekehrt kann man aber nicht folgern, dass q prim ist, wenn (3.4) für alle $0 < a < q$ erfüllt ist. Die Zahlen, die dies erfüllen ohne Primzahlen zu sein, werden **Carmichael-Zahlen** genannt. Die drei kleinsten Carmichael-Zahlen sind 561, 1105 und 1729. In Forster [5] werden sowohl die Eigenschaften der Carmichael-Zahlen als auch einige praktische Primzahltests besprochen.

Der folgende Satz 3.10 ist von fundamentaler Bedeutung für die Kryptographie.

Satz 3.10 *Ist p prim, so ist $(\mathbb{Z}/p\mathbb{Z})^\times$ zyklisch.*

BEWEIS: Da p prim ist, ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper und $(\mathbb{Z}/p\mathbb{Z})^\times$ hat $p - 1$ Elemente.

Es sei g ein Element maximaler Ordnung $m \leq p - 1$ in $(\mathbb{Z}/p\mathbb{Z})^\times$. Dann ist g eine Nullstelle des Polynoms $X^m - 1$. Als Folge von Hilfssatz 3.4 gilt

$$a^m = 1$$

für alle $a \in (\mathbb{Z}/p\mathbb{Z})^\times$, also sind alle $p - 1$ Elemente von $(\mathbb{Z}/p\mathbb{Z})^\times$ ebenfalls Nullstellen von $X^m - 1$. Nach Satz 1.37 muss

$$m = \deg(X^m - 1) \geq p - 1$$

gelten. Folglich gilt $m = p - 1$ und $\langle g \rangle = (\mathbb{Z}/p\mathbb{Z})^\times$. ■

3.3 Schlüsselaustausch nach Diffie und Hellman

Alice und Bob wollen über einen unsicheren Kanal einen gemeinsamen geheimen Schlüssel k erzeugen. Dazu müssen sie „Teilschlüssel“ über den Kanal austauschen, aus denen sie k erzeugen können. Sie können wie folgt vorgehen:

Algorithmus 3.11 (Diffie-Hellman-Protokoll) Es sei g ein Erzeuger von $(\mathbb{Z}/p\mathbb{Z})^\times$, wobei p eine große Primzahl ist.

- (i) Alice wählt eine geheime Zufallszahl a und schickt g^a an Bob.
- (ii) Bob wählt eine geheime Zufallszahl b und schickt g^b an Alice.
- (iii) Alice empfängt g^b und berechnet $k = (g^b)^a$.
- (iv) Bob empfängt g^a und berechnet $k = (g^a)^b$.

Damit haben Alice und Bob den gemeinsamen Schlüssel k , der nicht über den unsicheren Kanal übertragen wurde.

Bemerkung 3.12 Die Sicherheit dieses Verfahrens beruht auf der Schwierigkeit, den diskreten Logarithmus in $(\mathbb{Z}/p\mathbb{Z})^\times$ zu berechnen (dazu ist p möglichst groß zu wählen). Damit könnte ein Angreifer aus den abgehörten Teilschlüsseln g^a und g^b die geheimen Zahlen a und b berechnen und somit k bestimmen. Ein Algorithmus für den diskreten Logarithmus in zyklischen Gruppen ist in §8 von Forster [5] zu finden. Das Protokoll funktioniert im Prinzip mit allen endlichen zyklischen Gruppen, wie z.B. elliptischen Kurven (Forster [5], §20).

3.4 El Gamal-Verschlüsselung

Eine ähnliche Idee wie beim Diffie-Hellman-Protokoll liegt dem Verschlüsselungsverfahren von El Gamal zugrunde. Der Einfachheit wegen nehmen wir an, dass die Menge der Klartexte $\mathcal{M} = (\mathbb{Z}/p\mathbb{Z})^\times$ und die Menge der Chiffre ebenfalls $\mathcal{C} = (\mathbb{Z}/p\mathbb{Z})^\times$ sei.

Algorithmus 3.13 (El Gamal-Verschlüsselung) Alice will eine geheime Nachricht m an Bob schicken. Es sei g ein Erzeuger von $(\mathbb{Z}/p\mathbb{Z})^\times$, wobei p eine große Primzahl ist.

- (i) Bob wählt eine geheime Zufallszahl b (den geheimen Schlüssel) und berechnet $x = g^b$ (den öffentlichen Schlüssel).
- (ii) Alice wählt eine geheime Zufallszahl a , die $\text{ggT}(a, p - 1) = 1$ erfüllt.
- (iii) Alice berechnet $y = g^a$ und $z = x^a m$, und schickt (y, z) an Bob.
- (iv) Bob kann die Nachricht mit Hilfe seines geheimen Schlüssels b durch die folgende Rechnung entschlüsseln:

$$z(y^b)^{-1} = z((g^a)^b)^{-1} = z(g^{ab})^{-1} = x^a m (x^a)^{-1} = m.$$

Auch hier beruht die Sicherheit auf der Schwierigkeit, den diskreten Logarithmus zu berechnen. Die Bedingung $\text{ggT}(a, p - 1) = 1$ ist nicht zwingend erforderlich. Hätten a und $p - 1$ jedoch einen großen gemeinsamen Teiler, so würde die Berechnung des diskreten Logarithmus von g^a erheblich vereinfacht.

3.5 RSA-Verschlüsselung

Das RSA-Verfahren ist benannt nach seinen Erfindern Rivest, Shamir und Adleman.

Aufgabe 3.14 Für Primzahlen $p, q \in \mathbb{N}$ gilt $\varphi(pq) = (p - 1)(q - 1)$.

Algorithmus 3.15 (RSA-Verschlüsselung) Alice will eine Nachricht m an Bob schicken.

- (i) Bob wählt zwei große geheime Primzahlen p, q und veröffentlicht $n = pq$.
- (ii) Bob wählt eine zu $\varphi(n)$ teilerfremde Zahl e (den öffentlichen Schlüssel). Mit dem erweiterten euklidischen Algorithmus bestimmt er eine Zahl d (den geheimen Schlüssel), die $ed \equiv 1 \pmod{\varphi(n)}$ erfüllt.
- (iii) Alice kann mit den öffentlichen Daten eine Nachricht $m \in \mathbb{Z}/n\mathbb{Z}$ wie folgt verschlüsseln:

$$\text{enc} : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}, \quad m \mapsto m^e.$$

Sie schickt diese Nachricht an Bob.

- (iv) Bob kann das Chiffrat m^e mit seinem geheimen Schlüssel entschlüsseln:

$$\text{dec} : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}, \quad x \mapsto x^d.$$

Es ist also $\text{dec}(m^e) = m^{ed} = m$.

Beweis: Zu zeigen ist, dass die Aussage in Schritt (iv) wahr ist.

Es sei $x \in \mathbb{Z}/n\mathbb{Z}$. Nach dem chinesischen Restsatz gibt es einen Isomorphismus

$$\Psi : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}.$$

Es sei $(x_1, x_2) = \Psi(x)$. Ist $x_1 = 0$, so gilt trivialerweise $x_1^{ed} = x_1$. Falls $x_1 \neq 0$, so ist x_1 eine Einheit in $\mathbb{Z}/p\mathbb{Z}$. Da nach Aufgabe 3.14 gilt $\varphi(n) = (p - 1)(q - 1)$, gilt auch $ed \equiv 1 \pmod{p - 1}$ und $ed \equiv 1 \pmod{q - 1}$. Des Weiteren ist $\varphi(p) = p - 1$, und somit folgt aus dem Satz von Euler

$$x_1^{ed} = x_1^{1+k\varphi(p)} = x_1.$$

Entsprechend folgt $x_2^{ed} = x_2$. Das bedeutet

$$(x_1, x_2)^{ed} = (x_1^{ed}, x_2^{ed}) = (x_1, x_2),$$

und somit

$$x^{ed} = \Psi^{-1}((x_1, x_2)^{ed}) = \Psi^{-1}(x_1, x_2) = x.$$

Damit ist $\text{dec} \circ \text{enc} = \text{id}_{\mathbb{Z}/n\mathbb{Z}}$ gezeigt. ■

Man kann sich nun fragen, warum man im Beweis den Umweg über den chinesischen Restsatz nimmt, statt direkt mit dem Satz von Euler zu argumentieren, dass

$$x^{ed} = x^{1+k\varphi(n)} = x$$

gilt. Diese Argumentation ist aber nur für solche x zulässig, die zu n teilerfremd sind. Daher verlagert man das Problem in die Ringe $\mathbb{Z}/p\mathbb{Z}$ und $\mathbb{Z}/q\mathbb{Z}$, denn da p und q prim sind, ist hier sichergestellt, dass für alle $x_i \neq 0$ der Satz von Euler anwendbar ist.

Bemerkung 3.16 Die Sicherheit des RSA-Verfahrens beruht darauf, dass der geheime Schlüssel d schwer zu bestimmen ist. Dazu müsste man $\varphi(n)$ kennen und dann das Inverse von e modulo $\varphi(n)$ bestimmen. Aber um $\varphi(n)$ zu kennen, muss man wiederum die Primfaktorzerlegung pq von n kennen. Diese ist im Allgemeinen aber sehr schwer zu finden.

Teil II

Geometrie in Vektorräumen

Bisher haben wir die algebraischen Eigenschaften von Vektorräumen studiert. Statten wir einen Vektorraum V mit einer zusätzlichen Struktur, einem *Skalarprodukt*, aus, so erhält V eine Geometrie. Genauer gesprochen können wir nun Winkel, Abstände und Längen in V definieren. Wie so oft steht uns hierbei ein elementarer Spezialfall Pate, in diesem Fall das *Standardskalarprodukt* im \mathbb{R}^2 . Durch das Abstrahieren der algebraischen Eigenschaften von seiner konkreten Definition werden wir auf eine abstrakte Definition für Skalarprodukte in beliebigen reellen oder komplexen Vektorräumen geführt.

4 Vektorräume mit Skalarprodukt

4.1 Das Standardskalarprodukt

Das **Standardskalarprodukt** zweier Vektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

ist definiert als

$$\langle x|y \rangle = x^\top \cdot y = x_1 y_1 + \dots + x_n y_n \quad (4.1)$$

Das Ergebnis ist eine reelle Zahl, was den Namen *Skalarprodukt* erklärt.

Wir beschränken uns der Übersichtlichkeit wegen auf den Fall $n = 2$. Zunächst stellen wir einige algebraische Eigenschaften des Skalarprodukts fest:

Bemerkung 4.1 Für $x, y \in \mathbb{R}^2$ gilt

$$\langle x|y \rangle = x_1 y_1 + x_2 y_2 = y_1 x_1 + y_2 x_2 = \langle y|x \rangle. \quad (4.2)$$

Das Skalarprodukt ist also *symmetrisch*.

Bemerkung 4.2 Für alle $x, y, z \in \mathbb{R}^2$ und jede Zahl $\alpha \in \mathbb{R}$ gilt

$$\begin{aligned} \langle x|y + z \rangle &= x_1(y_1 + z_1) + x_2(y_2 + z_2) \\ &= x_1 y_1 + x_1 z_1 + x_2 y_2 + x_2 z_2 \\ &= (x_1 y_1 + x_2 y_2) + (x_1 z_1 + x_2 z_2) \\ &= \langle x|y \rangle + \langle x|z \rangle \end{aligned}$$

und

$$\begin{aligned}\langle x|ay \rangle &= x_1(ay_1) + x_2(ay_2) \\ &= a(x_1y_1) + a(x_2y_2) \\ &= a(x_1y_1 + x_2y_2) \\ &= a\langle x|y \rangle.\end{aligned}$$

Das bedeutet, dass $\langle x|\cdot \rangle$ für festes x eine lineare Funktion ist:

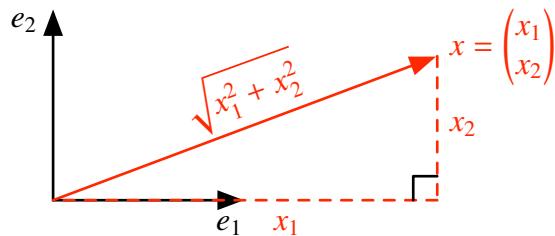
$$\langle x|ay + z \rangle = a\langle x|y \rangle + \langle x|z \rangle. \quad (4.3)$$

Aus der Symmetrie (4.2) folgt, dass auch $\langle \cdot|x \rangle$ für festes x eine lineare Funktion ist. Dies bezeichnen wir als die *Bilinearität* des Skalarprodukts.

Nun wenden wir uns den geometrischen Eigenschaften zu. Setzen wir $x = y$ in (4.1), so erhalten wir

$$\langle x|x \rangle = x_1^2 + x_2^2.$$

Das folgende Bild zeigt, wie wir dies geometrisch deuten können.



In diesem Bild stehen die e_1 -Achse und die e_2 -Achse senkrecht aufeinander. Wir erhalten also ein rechtwinkliges Dreieck, dessen Seiten durch x , die e_1 -Komponente von x und die e_2 -Komponente von x gegeben sind. Die Hypotenuse entspricht hier der Linie vom Ursprung nach x . Der Satz des Pythagoras besagt in dieser Situation

$$x_1^2 + x_2^2 = (\text{Länge des Vektors } x)^2.$$

Das Skalarprodukt erlaubt uns also, die Länge

$$\sqrt{\langle x|x \rangle} = \sqrt{x_1^2 + x_2^2} \quad (4.4)$$

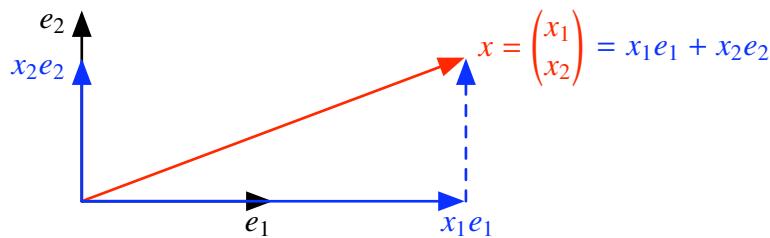
eines Vektors x zu ermitteln. Insbesondere stellen wir folgendes fest:

Bemerkung 4.3 Jeder Vektor $x \neq 0$ hat positive Länge. Der Vektor 0 hat jedoch die Länge 0. Das bedeutet

$$\langle x|x \rangle \geq 0, \text{ und } „= 0“ \text{ nur dann, wenn } x = 0. \quad (4.5)$$

Diese Eigenschaft werden wir später als *positive Definitheit* bezeichnen.

Auch das Skalarprodukt zweier verschiedener Vektoren $x, y \in \mathbb{R}^2$ lässt sich geometrisch deuten. Dazu betrachten wir die Spezialfälle $y = e_1$ und $y = e_2$ mit jeweils beliebigem x . Im folgenden Bild sehen wir, wie x sich aus einem Vielfachen von e_1 und einem Vielfachen von e_2 zusammensetzt.



Dabei ist der Betrag der jeweiligen e_i -Komponente der i te Koeffizient x_i ($i = 1, 2$). Im Fall $y = e_1$ gilt

$$\langle x|y \rangle = \langle x|e_1 \rangle = x_1 \cdot 1 + x_2 \cdot 0 = x_1,$$

und im Fall $y = e_2$ gilt

$$\langle x|y \rangle = \langle x|e_2 \rangle = x_1 \cdot 0 + x_2 \cdot 1 = x_2.$$

Das Skalarprodukt $\langle x|e_i \rangle$ berechnet also den Betrag der senkrechten Projektion von x auf die e_i -Achse. Lax formuliert gibt $\langle x|e_i \rangle$ an, welchen Beitrag der i te Einheitsvektor zum Vektor x leistet. Dies liefert eine Zerlegung von x in orthogonale Komponenten:

$$x = x_1e_1 + x_2e_2 = \langle x|e_1 \rangle e_1 + \langle x|e_2 \rangle e_2.$$

Ist nun y ein beliebiger Vektor der Länge 1, so kann man in einem geeigneten Koordinatensystem annehmen, dass $y = e_1$ gilt. Wir stellen somit fest:

Bemerkung 4.4 Für einen Vektor y der Länge 1 gibt das Skalarprodukt $\langle x|y \rangle$ die Länge der senkrechten Projektion von x auf die y -Achse an. Sind y_1, y_2 zwei senkrecht aufeinanderstehende Vektoren der Länge 1 (z.B. $y_i = e_i$), so lässt sich x mittels des Skalarproduktes in seine orthogonalen Komponenten bzgl. der Basis $\{y_1, y_2\}$ zerlegen:

$$x = \langle x|y_1 \rangle y_1 + \langle x|y_2 \rangle y_2.$$

Diese letzte Eigenschaft erlaubt uns, die Darstellung von x in einer geeigneten Basis allein durch das Berechnen von Skalarprodukten zu ermitteln. Besonders interessant ist dies im Falle unendlicher Dimension, in dem man eine solche Darstellung nicht durch das Lösen linearer Gleichungssysteme erhalten kann.

In den folgenden Abschnitten 4.2 und 4.3 gelangen wir über die Eigenschaften (4.3), (4.2) und (4.5) zu einer abstrakten Definition von Skalarprodukten. Zunächst werden wir dazu den Begriff der Bilinearität von einem sehr allgemeinen Standpunkt aus untersuchen.

4.2 Bilinearformen

In diesem Abschnitt sei \mathbb{K} ein Körper und V ein \mathbb{K} -Vektorraum.

Definition 4.5 Eine Abbildung $s : V \times V \rightarrow \mathbb{K}$ heißt **Bilinearform**, falls für alle $x, y, z \in V$ und alle $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned} s(x+y, z) &= s(x, z) + s(y, z), \\ s(x, y+z) &= s(x, y) + s(x, z), \\ s(\lambda x, y) &= \lambda s(x, y), \\ s(x, \lambda y) &= \lambda s(x, y). \end{aligned}$$

Es bezeichne $\text{Bil}(V)$ die Menge der Bilinearformen auf V .

Bemerkung 4.6 Dass s eine Bilinearform ist, ist äquivalent dazu, dass für jedes $x \in V$ die Funktionen $s(x, \cdot) : V \rightarrow \mathbb{K}$ und $s(\cdot, x) : V \rightarrow \mathbb{K}$ Linearformen sind. $\text{Bil}(V)$ ist mit der üblichen Addition und Skalarmultiplikation von Funktionen ein \mathbb{K} -Vektorraum.

Definition 4.7 Es sei $s : V \times V \rightarrow \mathbb{K}$ eine Bilinearform.

- (a) s heißt **symmetrisch**, falls $s(x, y) = s(y, x)$ gilt für alle $x, y \in V$.
- (b) s heißt **schiefsymmetrisch**, falls $s(x, y) = -s(y, x)$ gilt für alle $x, y \in V$.
- (c) s heißt **alternierend**, falls $s(x, x) = 0$ gilt für alle $x \in V$.

Es bezeichne $\text{Sym}(V)$ die Menge der symmetrischen Bilinearformen auf V .

Aufgabe 4.8 Jede alternierende Bilinearform s ist schiefsymmetrisch. Falls die Charakteristik von \mathbb{K} nicht 2 ist, ist s genau dann alternierend, wenn s schiefsymmetrisch ist.

Beispiel 4.9 (Bilinearformen)

(a) Das Standardskalarprodukt $s = \langle \cdot | \cdot \rangle$ im \mathbb{R}^n ist eine symmetrische Bilinearform (vgl. Abschnitt 4.1).

(b) Auf \mathbb{K}^2 ist durch

$$s(x, y) = \det(x \ y)$$

eine alternierende Bilinearform gegeben.

(c) Auf dem Vektorraum $C([a, b])$ der stetigen reellwertigen Funktionen auf dem Intervall $[a, b]$ ist durch

$$s(f, g) = \int_a^b f(t)g(t) dt$$

ist eine symmetrische Bilinearform gegeben.

(d) Jede Matrix $A \in \mathbb{K}^{n \times n}$ definiert auf \mathbb{K}^n eine Bilinearform durch die Vorschrift

$$s_A(x, y) = x^\top \cdot A \cdot y.$$

In der Tat können alle Bilinearformen von Vektorräumen endlicher Dimension in der Form von Beispiel 4.9 (d) dargestellt werden:

Hilfssatz 4.10 Es sei $\dim V = n$ und $B = \{b_1, \dots, b_n\}$ eine Basis von V . Weiter sei $s : V \times V \rightarrow \mathbb{K}$ eine Bilinearform. Wir setzen $s_{ij} = s(b_i, b_j)$ und definieren die Matrix

$$S_B(s) = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} \in \mathbb{K}^{n \times n}. \quad (4.6)$$

Dann gilt

$$s(x, y) = \Theta_B(x)^\top \cdot S_B(s) \cdot \Theta_B(y). \quad (4.7)$$

BEWEIS: Beide Seiten der Gleichung (4.6) sind bilinear. Es genügt also, die Gleichheit für Basisvektoren $x = b_i$, $y = b_j$ zu zeigen. Es gilt

$$\Theta_B(b_i) = e_i.$$

Also ist

$$\Theta_B(b_i)^\top \cdot S_B(s) \cdot \Theta_B(b_j) = e_i^\top \cdot S_B(s) \cdot e_j = s_{ij} = s(b_i, b_j),$$

wie behauptet. ■

Aufgabe 4.11 Für gegebene Basis B ist die Abbildung

$$S_B : \text{Bil}(V) \rightarrow \mathbb{K}^{n \times n}, \quad s \mapsto S_B(s)$$

bijektiv und sogar ein Isomorphismus von \mathbb{K} -Vektorräumen.

Beispiel 4.12 (Matrizen von Bilinearformen) Es sei B die Standardbasis von \mathbb{K}^n .

(a) Für das Standardskalarprodukt im \mathbb{R}^n gilt $\langle e_i | e_j \rangle = \delta_{ij}$. Also ist

$$S_B(\langle \cdot | \cdot \rangle) = I_n.$$

(b) In Beispiel 4.9 (b) gilt

$$s(e_1, e_1) = 0 = s(e_2, e_2), \quad s(e_1, e_2) = 1, \quad s(e_2, e_1) = -1.$$

Also ist

$$S_B(s) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Hilfssatz 4.13 Es sei $\dim V = n$ und B eine beliebige Basis von V . Weiter sei s eine Bilinearform auf V und $S = S_B(s)$.

(a) s ist genau dann symmetrisch, wenn S eine symmetrische Matrix ist (d.h. $S = S^\top$).

(b) s ist genau dann schiefsymmetrisch, wenn S eine schiefsymmetrische Matrix ist (d.h. $S = -S^\top$). ■

BEWEIS:

(a) s symmetrisch genau dann, wenn $s_{ij} = s_{ji}$ gilt in (4.6). Dies ist äquivalent zu $S = S^\top$.

(b) s schiefsymmetrisch genau dann, wenn $s_{ij} = -s_{ji}$ gilt in (4.6). Dies ist äquivalent zu $S = -S^\top$. ■

Sind für eine Bilinearform s die Matrizen $S_B(s)$ und $S_C(s)$ für zwei Basen B, C von V gegeben, so können wir ähnlich wie bei den Abbildungsmatrizen von Endomorphismen $S_C(s)$ bestimmen, wenn wir $S_B(s)$ kennen.

Satz 4.14 Es sei $\dim V = n$ und es seien B, C Basen von V . Für jede Bilinearform $s : V \times V \rightarrow \mathbb{K}$ gilt

$$S_C(s) = (M_B^C)^\top \cdot S_B(s) \cdot M_B^C. \quad (4.8)$$

Hier bezeichnet $M_B^C = M_B^C(\text{id}_V)$ die Übergangsmatrix für den Basiswechsel von der Basis C zur Basis B .

Beweis: Für alle $x \in V$ gilt $\Theta_C(x) = M_C^B \cdot \Theta_B(x)$.

Also gilt nach (4.7) für alle $x, y \in V$:

$$\begin{aligned}\Theta_B(x)^\top \cdot S_B(s) \cdot \Theta_B(y) &= s(x, y) = \Theta_C(x)^\top \cdot S_C(s) \cdot \Theta_C(y) \\ &= (M_C^B \cdot \Theta_B(x))^\top \cdot S_C(s) \cdot (M_C^B \cdot \Theta_B(y)) \\ &= \Theta_B(x)^\top \cdot (M_C^B)^\top \cdot S_C(s) \cdot M_C^B \cdot \Theta_B(y).\end{aligned}$$

Wegen der Eindeutigkeit von $S_B(s)$ folgt $S_B(s) = (M_C^B)^\top \cdot S_C(s) \cdot M_C^B$. ■

Bemerkung 4.15 Man beachte, dass im Allgemeinen

$$(M_C^B)^\top \neq (M_C^B)^{-1} = M_B^C$$

gilt! Der Basiswechsel für Bilinearformen ist also anders zu berechnen als der Basiswechsel für lineare Abbildungen (in der Tat ist jener leichter zu berechnen, da man M_C^B lediglich transponieren und nicht invertieren muss).

Definition 4.16 Eine Bilinearform $s \in \text{Sym}(V)$ heißt **ausgeartet**, falls $x_0 \in V$, $x_0 \neq 0$, existiert, so dass gilt

$$s(x_0, x) = 0 \quad \text{für alle } x \in V. \tag{4.9}$$

Andernfalls heißt s **nicht ausgeartet**.

Aufgabe 4.17 Es sei $\dim V < \infty$. Es ist $s \in \text{Sym}(V)$ genau dann nicht ausgeartet, wenn für alle $x \in V$ die Abbildungen

$$\begin{aligned}s_{1,x} : V &\rightarrow V^*, \quad x \mapsto s(x, \cdot), \\ s_{2,x} : V &\rightarrow V^*, \quad x \mapsto s(\cdot, x)\end{aligned}$$

Isomorphismen sind.

Beispiel 4.18 In Beispiel 4.9 sind alle Bilinearformen nicht ausgeartet. Ein nicht-triviales Beispiel für eine ausgeartete Bilinearform auf \mathbb{R}^2 ist

$$s(x, y) = x^\top \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot y.$$

Definition 4.19 Es sei $s \in \text{Sym}(V)$. Die Menge

$$\text{Rad } s = \{x_0 \in V \mid s(x_0, x) = 0 \text{ für alle } x \in V\} \tag{4.10}$$

heißt das **Radikal** (oder der **Kern**) von s .

Bemerkung 4.20 Offensichtlich ist $s \in \text{Sym}(V)$ genau dann nicht ausgeartet, wenn $\text{Rad } s = \{0\}$ gilt.

Hilfssatz 4.21 Ist $S_B(s)$ eine Matrix für $s \in \text{Sym}(V)$, so gilt

$$\Theta_B(\text{Rad } s) = \text{Kern } S_B(s).$$

BEWEIS: Ist $x_0 \in \text{Rad } s$, so gilt $e_i^\top \cdot S_B(s) \cdot \Theta_B(x_0) = 0$ für alle Einheitsvektoren e_i . Also ist $S_B(s) \cdot \Theta_B(x_0) = 0$.

Gilt umgekehrt $S_B(s) \cdot v = 0$ für ein $v \in \mathbb{K}^n$, so ist $s(x, \Theta_B^{-1}(v)) = \Theta_B(x)^\top \cdot S_B(s) \cdot v = 0$ für alle $x \in V$. Also ist $\Theta_B^{-1}(v) \in \text{Rad } s$. ■

Folgerung 4.22 $s \in \text{Sym}(V)$ ist genau dann nicht ausgeartet, wenn $S_B(s) \in \mathbf{GL}_n(\mathbb{K})$.

4.3 Euklidische Vektorräume

Um zu einer abstrakten Definition von Skalarprodukten zu gelangen, ergänzen wir den Begriff der Bilinearform nun um die den Eigenschaften (4.2) und (4.5) entsprechenden Begriffe.

Definition 4.23 Es sei V ein \mathbb{R} -Vektorraum. Eine Bilinearform $s \in \text{Sym}(V)$ heißt **positiv definit**, falls für alle $x \in V$, $x \neq 0$, gilt

$$s(x, x) > 0.$$

Ein **Skalarprodukt** $\langle \cdot | \cdot \rangle$ auf V ist eine positiv definite symmetrische Bilinearform.

Definition 4.24 Ein **euklidischer Vektorraum** $(V, \langle \cdot | \cdot \rangle)$ ist ein \mathbb{R} -Vektorraum V zusammen mit einem Skalarprodukt $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{R}$.

Bemerkung 4.25

- (a) Beachte, dass wegen der Bilinearität stets $s(0, 0) = 0$ gilt.
- (b) Wir beschränken uns bei der Definition von Skalarprodukten auf \mathbb{R} -Vektorräume, da die positive Definitheit nur Sinn ergibt, wenn der zugrundeliegende Skalarkörper geordnet ist.⁴⁾
- (c) Neben der Schreibweise $\langle x | y \rangle$ ist auch die Schreibweise $\langle x, y \rangle$ in der Mathematik weit verbreitet. Gelegentlich wird das Standardskalarprodukt im \mathbb{R}^n mit einem Punkt geschrieben: $x \cdot y$.

⁴⁾In Abschnitt 4.5 werden wir ausnutzen, dass \mathbb{R} ein Teilkörper von \mathbb{C} ist, um auch in komplexen Vektorräumen Skalarprodukte zu definieren.

- (d) Die Schreibweise $\langle x|y \rangle$ wurde von Dirac als **Bra-Ket-Notation** (von engl. *bracket*) in der Quantenphysik eingeführt. Dirac interpretiert dabei einen Vektor $y \in V$ als *Ket-Vektor* $|y\rangle$ und bezeichnet die Linearform $x \mapsto \langle x|\cdot\rangle$ aus V^* als *Bra-Vektor* $\langle x|$. Das Anwenden der Linearform $\langle x|$ auf den Vektor $|y\rangle$ liefert dann gerade den Wert des Skalarprodukts $\langle x|y \rangle$.

Beispiel 4.26 (Euklidische Vektorräume)

- (a) Der \mathbb{R}^n mit dem Standardskalarprodukt ist ein euklidischer Vektorraum (dass das Standardskalarprodukt ein Skalarprodukt im Sinne von Definition 4.23 ist, ergibt sich sofort aus den Überlegungen in Abschnitt 4.1).
- (b) Im \mathbb{R}^n ist für jede Matrix $A \in \mathbb{R}^{n \times n}$ durch $s_A(x, y) = x^\top \cdot A \cdot y$ eine Bilinearform gegeben. Wir überlegen, welche Eigenschaften die Matrix A besitzen muss, damit s_A ein Skalarprodukt ist: Zunächst muss A nach Hilfssatz 4.13 eine symmetrische Matrix sein. Positive Definitheit bedeutet

$$s_A(x, x) = x^\top \cdot A \cdot x > 0$$

für alle $x \in \mathbb{R}^n$, $x \neq 0$.

Die Beobachtung im letzten Beispiel veranlasst uns zu folgender Definition:

Definition 4.27 Eine symmetrische Matrix $S \in \mathbb{R}^{n \times n}$ heißt **positiv definit**, falls für alle $x \in \mathbb{R}^n$, $x \neq 0$, gilt:

$$x^\top \cdot S \cdot x > 0. \quad (4.11)$$

Die Bedingung (4.11) direkt nachzuprüfen ist in der Regel sehr schwer. In Abschnitt 5.5 lernen wir einige Kriterien kennen, mit denen man eine symmetrische Matrix leicht auf positive Definitheit prüfen kann.

Hilfssatz 4.28 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum.

- (a) $\langle \cdot | \cdot \rangle$ ist nicht ausgeartet.
- (b) Ist $\dim V < \infty$ mit Basis B , so ist die Matrix $S = S_B(\langle \cdot | \cdot \rangle)$ symmetrisch, positiv definit und invertierbar.
- (c) Alle Eigenwerte der Matrix S aus (b) sind positiv.

BEWEIS:

- (a) Wegen der positiven Definitheit ist $\text{Rad}\langle \cdot | \cdot \rangle = \{0\}$.

- (b) Ergibt sich aus Hilfssatz 4.13, Beispiel 4.26 (b) und Teil (a) zusammen mit Folgerung 4.22.
- (c) Sei λ Eigenwert von S mit Eigenvektor $x \in \mathbb{R}^n$. Dann gilt

$$0 < x^\top \cdot S \cdot x = x^\top \cdot (\lambda \cdot x) = \lambda \cdot \underbrace{x^\top \cdot x}_{>0}.$$

Das bedeutet $\lambda > 0$. ■

Zum Ende dieses Abschnitts betrachten wir noch ein Standardbeispiel eines Skalarprodukts in Funktionsvektorräumen von unendlicher Dimension.

Beispiel 4.29 Der Vektorraum $C([a, b])$ der stetigen Funktionen auf dem Intervall $[a, b]$ ist ein euklidischer Vektorraum mit dem Skalarprodukt

$$\langle f | g \rangle = \int_a^b f(t)g(t)dt. \quad (4.12)$$

Aus Beispiel 4.9 (c) wissen wir bereits, dass dies eine symmetrische Bilinearform ist. Sie ist auch positiv definit: Es ist $\int_a^b f(t)^2 dt \geq 0$ für alle $f \in C([a, b])$, da $f(t)^2 \geq 0$ ist für alle $t \in [a, b]$. Falls $f \neq 0$ ist, so ist $f(t_0)^2 > 0$ an einer Stelle $t_0 \in [a, b]$. Aus der Definition der Stetigkeit folgt, dass $f(t)^2 > 0$ in einer geeigneten Umgebung $[t_0 - \delta, t_0 + \delta]$ von t_0 gilt (mit $\delta > 0$). Dann gilt aber

$$\langle f | f \rangle = \int_a^b f(t)^2 dt \geq \int_{t_0-\delta}^{t_0+\delta} f(t)^2 dt \geq 2\delta \min\{f(t)^2 \mid t \in [t_0 - \delta, t_0 + \delta]\} > 0.$$

Somit ist $\langle \cdot | \cdot \rangle$ in der Tat ein Skalarprodukt.

Wir wollen nun noch heuristisch verstehen, wieso man (4.12) als die Entsprechung des Standardskalarprodukts im \mathbb{R}^n auf Funktionsvektorräumen auffassen kann: Ein Element $x \in \mathbb{R}^n$ ist durch seine Koeffizienten x_1, \dots, x_n festgelegt. Eine andere Sichtweise auf x ist es, x als Funktion $x : \{1, \dots, n\} \rightarrow \mathbb{R}$ aufzufassen, die der Zahl i gerade den Wert des Koeffizienten x_i zuweist: $x(i) = x_i$. Das Standardskalarprodukt zweier Vektoren $x, y \in \mathbb{R}^n$ ist dann

$$\langle x | y \rangle = \sum_{i=1}^n x(i)y(i),$$

die Summe über das Produkt der Funktionswerte von x und y , ausgewertet an allen Elementen von $\{1, \dots, n\}$. Betrachten wir nun anstelle der endlichen Menge $\{1, \dots, n\}$ das Intervall $[a, b]$ und die Funktionen darauf, so entspricht der endlichen Summe $\sum_{i=1}^n$ wegen der Überabzählbarkeit von $[a, b]$ das Integral \int_a^b . Auf diese Weise führt uns das Standardskalarprodukt auf (4.12).

4.4 Normen, Winkel und Orthogonalität

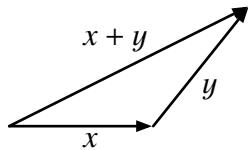
Normen verallgemeinern den elementargeometrischen Längenbegriff:

Definition 4.30 Es sei V ein \mathbb{R} -Vektorraum. Eine Funktion $\|\cdot\| : V \rightarrow \mathbb{R}$ heißt **Norm**, wenn sie folgende Eigenschaften besitzt:

- (i) $\|x\| > 0$ für alle $x \in V \setminus \{0\}$ und $\|0\| = 0$.
- (ii) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ für alle $x \in V$ und $\lambda \in \mathbb{R}$.
- (iii) $\|\cdot\|$ erfüllt die **Dreiecksungleichung** für alle $x, y \in V$:

$$\|x + y\| \leq \|x\| + \|y\|. \quad (4.13)$$

Die Dreiecksungleichung bedeutet, dass die Summe zweier Seitenlängen eines Dreiecks mindestens so groß ist wie die Länge der längsten Seite des Dreiecks.



Die positive Definitheit eines Skalarprodukts erlaubt es, in Analogie zu (4.4) einen Längenbegriff für euklidische Vektorräume einzuführen.

Satz 4.31 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum. Dann ist durch

$$\|x\| = \sqrt{\langle x | x \rangle} \quad (4.14)$$

eine Norm auf V definiert.

BEWEIS: Die Eigenschaft (i) folgt aus der positiven Definitheit des Skalarprodukts. Die Eigenschaft (ii) folgt aus der Bilinearität:

$$\|\lambda \cdot x\| = \sqrt{\langle \lambda x | \lambda x \rangle} = \sqrt{\lambda^2 \langle x | x \rangle} = \sqrt{\lambda^2} \cdot \sqrt{\langle x | x \rangle} = |\lambda| \cdot \|x\|.$$

Für Eigenschaft (iii) verwenden wir die Cauchy-Schwarz-Ungleichung (4.16), die im Anschluss bewiesen wird:

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

Wurzelziehen erhält \leq , und somit folgt die Dreiecksungleichung. ■

Satz 4.32 (Cauchy-Schwarz-Ungleichung) Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum. Für alle $x, y \in V$ gilt:

$$\langle x|y \rangle^2 \leq \langle x|x \rangle \cdot \langle y|y \rangle. \quad (4.15)$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

BEWEIS: Falls $x = 0$ oder $y = 0$ ist nichts zu zeigen. Also nehmen wir $x \neq 0$ und $y \neq 0$ an. Es gilt

$$0 \leq \left\langle \frac{x}{\|x\|} \pm \frac{y}{\|y\|} \middle| \frac{x}{\|x\|} \pm \frac{y}{\|y\|} \right\rangle = \underbrace{\frac{\langle x|x \rangle}{\|x\|^2}}_{=1} + \underbrace{\frac{\langle y|y \rangle}{\|y\|^2}}_{=1} \pm 2 \frac{\langle x|y \rangle}{\|x\| \cdot \|y\|} = 2 \pm 2 \frac{\langle x|y \rangle}{\|x\| \cdot \|y\|}. \quad (*)$$

Daraus folgt

$$\mp \frac{\langle x|y \rangle}{\|x\| \cdot \|y\|} \leq 1 \quad \text{bzw.} \quad \mp \langle x|y \rangle \leq \|x\| \cdot \|y\|.$$

Quadrieren ergibt (4.15).

Sind x, y linear abhängig, so gilt $y = \lambda x$ für ein $\lambda \in \mathbb{R}$. Dann ist

$$\langle x|y \rangle^2 = \lambda^2 \langle x|x \rangle^2 = \lambda^2 \langle x|x \rangle \langle x|x \rangle = \langle x|x \rangle \cdot \langle \lambda x|\lambda x \rangle = \langle x|x \rangle \cdot \langle y|y \rangle.$$

Umgekehrt nehmen wir nun an, es gelte Gleichheit in (4.15), also auch $\langle x|y \rangle = \|x\| \cdot \|y\|$. Einsetzen in (*) ergibt

$$0 \leq \left\langle \frac{x}{\|x\|} - \frac{y}{\|y\|} \middle| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\rangle = 2 - 2 = 0.$$

Aus der positiven Definitheit folgt $0 = \frac{x}{\|x\|} - \frac{y}{\|y\|}$ bzw. $x = \frac{\|x\|}{\|y\|} y$. ■

Die Cauchy-Schwarz-Ungleichung wird auch in der folgenden Form angegeben:

$$\langle x|y \rangle \leq \|x\| \cdot \|y\|. \quad (4.16)$$

Bemerkung 4.33 Es gibt Normen, die nicht von der Form (4.14) sind, also nicht zu einem Skalarprodukt gehören. Ein Beispiel hierfür ist die **Maximumsnorm** auf dem \mathbb{R}^n , gegeben durch

$$\|x\|_{\max} = \max\{x_1, \dots, x_n \mid x_i \text{ Koeffizienten von } x\}.$$

Es lässt sich zeigen, dass eine Norm $\|\cdot\|$ genau dann zu einem Skalarprodukt gehört, wenn alle $x, y \in V$ die **Parallelogrammgleichung** erfüllen:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Siehe dazu Werner [14], Satz V.1.7. In diesem Fall lässt sich das Skalarprodukt durch die Norm ausdrücken. Dies nennen wir **Polarisierung**:

$$\langle x|y \rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2). \quad (4.17)$$

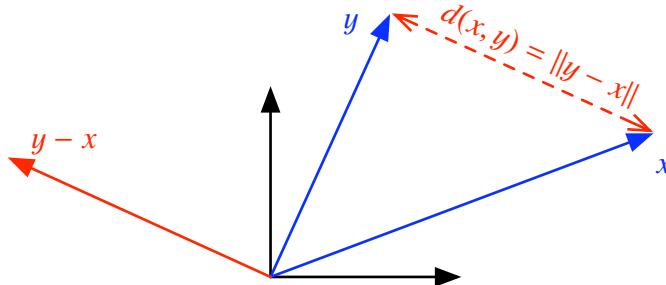
Aufgabe 4.34 Es sei $S \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix mit Einträgen s_{ij} . Dann gilt:

- (a) Es ist $s_{ik}^2 < s_{ii}s_{kk}$ für $i \neq k$.
- (b) Es existiert ein k mit $\max_{i,j} |s_{ij}| = s_{kk}$ (d.h. der betragsgrößte Matrixeintrag liegt auf der Diagonalen).

Können wir Längen bestimmen, so können wir auch Abstände bestimmen. Dazu definieren wir den Abstand zweier Vektoren x, y als die Länge des Verbindungsvektors von x zu y .

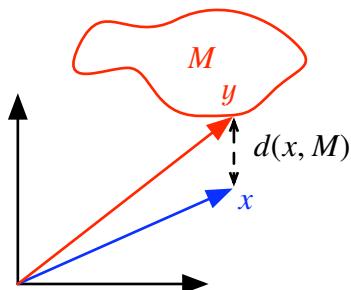
Definition 4.35 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum, M eine Teilmenge von V und $x, y \in V$. Der **Abstand $d(x, y)$ von x zu y** ist

$$d(x, y) = \|y - x\| \quad (4.18)$$



und der **Abstand $d(x, M)$ von x zu M** ist

$$d(x, M) = \inf\{d(x, y) \mid y \in M\}. \quad (4.19)$$



Aufgabe 4.36 Es sei $V = \mathbb{R}^n$ und $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt. Ist $M \subset \mathbb{R}^n$ abgeschlossen, so ist das Infimum in (4.19) ein Minimum. Insbesondere gilt dies, wenn M ein Untervektorraum von V ist.

Aufgabe 4.37 Es seien $x, y, z \in V$. Aus den Eigenschaften der Norm schließt man für den Abstand:

- (i) $d(x, y) > 0$ falls $x \neq y$ und $d(x, y) = 0$ falls $x = y$.
- (ii) $d(x, y) = d(y, x)$.
- (iii) $d(x, z) \leq d(x, y) + d(y, z)$.

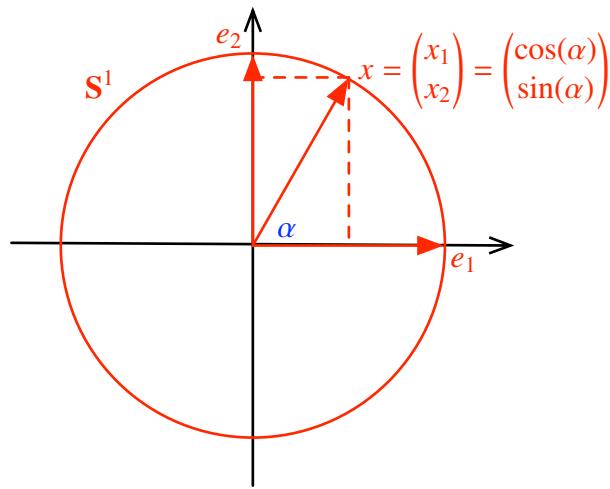
Beispiel 4.38 In einem euklidischen Vektorraum $(V, \langle \cdot | \cdot \rangle)$ nennen wir die Menge aller Vektoren mit Abstand r von einem Punkt $x_0 \in V$ die **Sphäre** vom Radius r mit Mittelpunkt x_0 , geschrieben

$$S_r(x_0) = \{x \in V \mid d(x, x_0) = r\}. \quad (4.20)$$

Die Sphäre $S_1(0)$ heißt **Einheitssphäre** und ihre Elemente **Einheitsvektoren** (oder **normierte Vektoren**). Im Fall $V = \mathbb{R}^n$ mit Standardskalarprodukt $\langle \cdot | \cdot \rangle$ nennen wir $S_r(0)$ die **$n - 1$ -Sphäre**⁵⁾ vom Radius r , geschrieben

$$S_r^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = r\}. \quad (4.21)$$

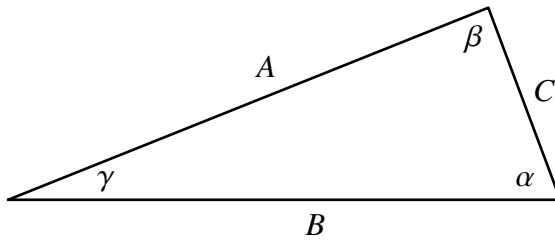
Für die Einheitssphäre im \mathbb{R}^n schreiben wir $S^{n-1} = S_1^{n-1}$.



⁵⁾Dass man hier $n - 1$ statt n wählt liegt daran, dass die Punkte auf S_r^{n-1} bereits durch $n - 1$ Polarkoordinaten eindeutig festgelegt sind. In diesem Sinne ist S_r^{n-1} ein „ $n - 1$ -dimensionales“ Gebilde.

Ein Skalarprodukt auf V ermöglicht es nicht nur, Längen in V zu definieren, sondern auch Winkel. In diesem Sinne enthält ein Skalarprodukt mehr Information als eine Norm.

Betrachten wir zunächst wieder die vertraute Situation im \mathbb{R}^2 . Gegeben sei ein Dreieck mit Seitenlängen A, B, C und entsprechenden Innenwinkeln α, β, γ .



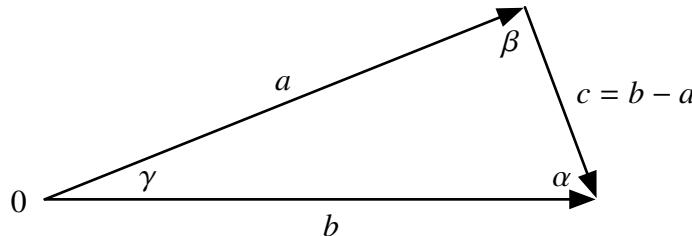
Mit Hilfe des elementargeometrischen **Cosinussatzes**

$$A^2 + B^2 - 2AB \cos(\gamma) = C^2. \quad (4.22)$$

können den Winkel γ in Abhängigkeit von den Seitenlängen angeben:

$$\cos(\gamma) = \frac{A^2 + B^2 - C^2}{2AB}. \quad (4.23)$$

Wir können also Winkel allein durch Längen ausdrücken. Die Längen wiederum können wir durch die Norm $\|\cdot\|$ des Standardskalarprodukts des \mathbb{R}^2 ausdrücken. Dazu nehmen wir an, der Eckpunkt des Dreiecks beim Winkel β sei der Ursprung, und wir wählen Vektoren $a, b \in \mathbb{R}^2$ und $c = b - a$ wie im folgenden Bild:



Insbesondere gilt dann $\|a\| = A$, $\|b\| = B$, $\|c\| = C$, und γ ist der von den Vektoren a, b eingeschlossene Winkel $\measuredangle(a, b)$. Setzen wir dies in (4.23) ein, so ergibt sich wegen $\langle c | c \rangle = \langle b - a | b - a \rangle = \langle b | b \rangle - 2\langle a | b \rangle + \langle a | a \rangle$:

$$\cos(\gamma) = \frac{\langle a | a \rangle + \langle b | b \rangle - \langle b - a | b - a \rangle}{2\|a\| \cdot \|b\|} = \frac{\langle a | b \rangle}{\|a\| \cdot \|b\|}. \quad (4.24)$$

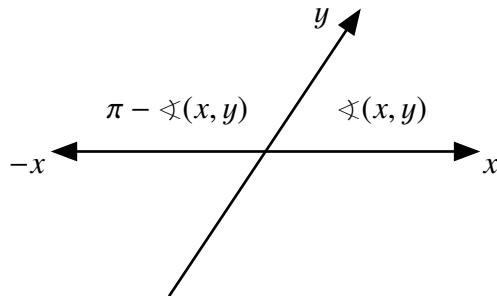
Da wir in jedem abstrakten euklidischen Vektorraum über die Norm einen Längenbegriff haben, können wir (4.24) als definierende Gleichung für den Cosinus des Winkels nehmen (die Cauchy-Schwarz-Ungleichung (4.16) gewährleistet dabei, dass $|\cos(\gamma)| \leq 1$ gilt):

Definition 4.39 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum und $x, y \in V$, $x \neq 0$, $y \neq 0$. Der von x und y eingeschlossene **Winkel** ist die eindeutig bestimmte Zahl $\measuredangle(x, y) \in [0, \pi]$ mit

$$\cos(\measuredangle(x, y)) = \frac{\langle x | y \rangle}{\|x\| \cdot \|y\|}. \quad (4.25)$$

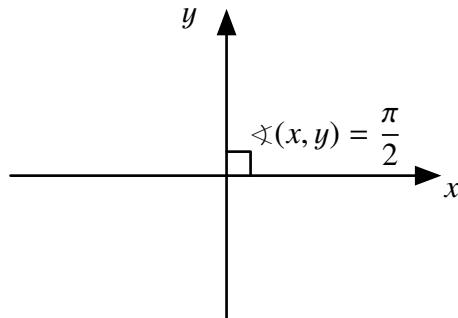
Aufgabe 4.40 Für alle $x, y \in V \setminus \{0\}$ und $\lambda, \mu \in \mathbb{R}^\times$ gilt:

- (a) $\measuredangle(x, y) = \measuredangle(y, x)$.
- (b) $\measuredangle(\lambda x, \mu y) = \begin{cases} \measuredangle(x, y), & \lambda\mu > 0 \\ \pi - \measuredangle(x, y), & \lambda\mu < 0 \end{cases}$.
- (c) $\measuredangle(x, y) = 0$ genau dann, wenn $y = \alpha x$ für ein $\alpha > 0$.
- (d) $\measuredangle(x, y) = \pi$ genau dann, wenn $y = \alpha x$ für ein $\alpha < 0$.



Wie in Abschnitt 4.1 erläutert, lässt sich das Skalarprodukt $\langle x | y \rangle$ (und somit der Winkel) als Maß der Komponente in x -Richtung von y auffassen.

Definition 4.41 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum. Zwei Vektoren $x, y \in V$ heißen **orthogonal** (oder **senkrecht**), falls $\langle x | y \rangle = 0$ gilt, geschrieben $x \perp y$. Zwei Teilmengen $M_1, M_2 \subset V$ heißen **orthogonal**, $M_1 \perp M_2$, falls $x_1 \perp x_2$ gilt für alle $x_1 \in M_1, x_2 \in M_2$.



$x \perp y$ bedeutet zugleich $\cos(\alpha(x, y)) = 0$, was wiederum $\alpha(x, y) = \frac{\pi}{2}$ bedeutet. Dies deckt sich mit unserer Anschauung, dass zwei Vektoren senkrecht aufeinander stehen, wenn sie den Winkel $\frac{\pi}{2}$ ($= 90^\circ$) einschließen.

Beispiel 4.42 (Orthogonale Vektoren)

- (a) In jedem euklidischen Vektorraum gilt $0 \perp x$ für alle x .
- (b) Im \mathbb{R}^3 mit Standardskalarprodukt sind die beiden Vektoren

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

orthogonal. Es gilt nämlich

$$\langle x|y \rangle = x^\top \cdot y = 1 \cdot (-2) + 2 \cdot 1 + 3 \cdot 0 = -2 + 2 = 0.$$

- (c) Im \mathbb{R}^n mit Standardskalarprodukt sind je zwei verschiedene Elemente e_i, e_j der Standardbasis orthogonal.
- (d) Im Vektorraum $C([0, 2\pi])$ mit Skalarprodukt (4.12) sind für $m, n \in \mathbb{N}$ die Funktionen $\cos(mt)$ und $\sin(nt)$ orthogonal, d.h. es gilt

$$\langle \cos(m \cdot) | \sin(n \cdot) \rangle = \int_0^{2\pi} \cos(mt) \sin(nt) dt = 0.$$

In Kapitel 5 werden wir Systeme von orthogonalen Vektoren untersuchen.

Satz 4.43 (Pythagoras) Für Elemente x, y eines euklidischen Vektorraumes gilt $x \perp y$ genau dann, wenn

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2. \quad (4.26)$$

BEWEIS: Es ist

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

Also gilt (4.26) genau dann, wenn $\langle x | y \rangle = 0$, d.h. $x \perp y$. ■

Aus der Motivation für die Definition des Winkels ergibt sich direkt, dass auch der folgende Satz gilt:

Satz 4.44 (Cosinussatz) Für Elemente x, y eines euklidischen Vektorraumes gilt

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \cdot \|y\| \cos(\measuredangle(x, y)). \quad (4.27)$$

4.5 Unitäre Vektorräume

Wir wollen nun den Ansatz, einem Vektorraum V vermöge eines Skalarprodukts mit einer geometrischen Struktur auszustatten, auch auf \mathbb{C} -Vektorräume übertragen. Dafür können wir jedoch nicht die Definition des Skalarprodukts unverändert übernehmen: Ist etwa $V = \mathbb{C}^2$ und betrachten wir das „Standardskalarprodukt“

$$x^\top \cdot y = x_1 y_1 + x_2 y_2$$

für $x, y \in \mathbb{C}^2$, so ist dies zwar symmetrisch und bilinear, aber aufgrund möglicher imaginärer Einträge kann die positive Definitheit verletzt werden. Für

$$x = \begin{pmatrix} 1 \\ i \end{pmatrix} \in \mathbb{C}^2$$

ist beispielsweise

$$x^\top \cdot x = x_1^2 + x_2^2 = 1^2 + i^2 = 0.$$

Durch eine kleine Modifikation kann man aber positive Definitheit erreichen. Dazu definiert man das *Standardskalarprodukt* auf \mathbb{C}^2 durch

$$\langle x | y \rangle = x^\top \cdot \bar{y} = x_1 \bar{y}_1 + x_2 \bar{y}_2.$$

Da $\zeta \bar{\zeta} = |\zeta|^2 \in \mathbb{R}$ für alle komplexen Zahlen $\zeta \in \mathbb{C}$, ist dies positiv definit:

$$\langle x | x \rangle = x^\top \cdot \bar{x} = x_1 \bar{x}_1 + x_2 \bar{x}_2 = |x_1|^2 + |x_2|^2 \geq 0$$

und „ $= 0$ “ nur dann, wenn $x = 0$. Allerdings ist $\langle \cdot | \cdot \rangle$ nur noch im ersten Argument linear, im zweiten Argument gilt

$$\langle x | y + \lambda z \rangle = \langle x | y \rangle + \bar{\lambda} \langle x | z \rangle.$$

Dies bezeichnen wir als *sesquilinear* („ $1\frac{1}{2}$ fach linear“). Auch die Symmetrie ist verlorengegangen. Es gilt stattdessen

$$\langle x|y \rangle = x_1\bar{y}_1 + x_2\bar{y}_2 = \overline{y_1\bar{x}_1 + y_2\bar{x}_2} = \overline{\langle y|x \rangle}.$$

Definition 4.45 Es sei V ein \mathbb{C} -Vektorraum. Eine Abbildung $s : V \times V \rightarrow \mathbb{C}$ heißt **sesquilinear**⁶⁾, falls für alle $x, y, z \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned}s(x+y, z) &= s(x, z) + s(y, z), \\ s(\lambda x, y) &= \lambda s(x, y), \\ s(x, y+z) &= s(x, y) + s(x, z), \\ s(x, \lambda y) &= \bar{\lambda} s(x, y).\end{aligned}$$

Eine Sesquilinearform s heißt **hermitesch**, falls für alle $x, y \in V$ gilt:

$$s(x, y) = \overline{s(y, x)}.$$

Definition 4.46 Es sei V ein \mathbb{C} -Vektorraum. Eine hermitesche Sesquilinearform $s : V \times V \rightarrow \mathbb{C}$ heißt **positiv definit**, falls für alle $x \in V, x \neq 0$, gilt

$$s(x, x) > 0.$$

Ein **unitäres Skalarprodukt** $\langle \cdot | \cdot \rangle$ ist eine positiv definite hermitesche Sesquilinearform.

Definition 4.47 Ein **unitärer Vektorraum** $(V, \langle \cdot | \cdot \rangle)$ ist ein \mathbb{C} -Vektorraum V zusammen mit einem unitären Skalarprodukt $\langle \cdot | \cdot \rangle$.

Beispiel 4.48 (Unitäre Vektorräume) Die Beispiele können als „Komplexifizierung“ der entsprechenden euklidischen Beispiele aufgefasst werden:

- (a) \mathbb{C}^n ist ein unitärer Vektorraum mit dem (unitären) Standardskalarprodukt

$$\langle x|y \rangle = x^\top \cdot \bar{y} = x_1\bar{y}_1 + \dots + x_n\bar{y}_n. \quad (4.28)$$

- (b) Der Vektorraum $C([a, b], \mathbb{C})$ der stetigen \mathbb{C} -wertigen Funktionen auf dem Intervall $[a, b]$ ist unitär mit dem Skalarprodukt

$$\langle f|g \rangle = \int_a^b f(t)\overline{g(t)}dt. \quad (4.29)$$

⁶⁾Sesquilinearformen spielen in der theoretischen Physik eine große Rolle. Dort wird Sesquilinearität überlicherweise jedoch so definiert, dass s im zweiten Argument linear ist und im ersten Argument $s(\lambda x, y) = \bar{\lambda} s(x, y)$ gilt.

Die Resultate über euklidische Vektorräumen lassen sich auf unitäre Vektorräume übertragen, wenn man ihre Beweise so anpasst, dass sie die Sesquilinearität berücksichtigen, und gelegentlich einen Ausdruck z^2 durch $|z|^2$ ersetzt. Wir führen sie daher ohne ausführliche Begründungen auf.

Satz 4.49 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein unitärer Vektorraum der Dimension n . Weiter sei $B = \{b_1, \dots, b_n\}$ eine Basis von V und $S \in \mathbb{C}^{n \times n}$ die Matrix mit Einträgen

$$s_{ij} = \langle b_i | b_j \rangle.$$

Dann gilt für alle $x, y \in V$:

$$\langle x | y \rangle = \Theta_B(x)^\top \cdot S \cdot \overline{\Theta_B(y)}. \quad (4.30)$$

Des Weiteren gilt $S \in \mathbf{GL}_n(\mathbb{C})$ und

$$\overline{S}^\top = S. \quad (4.31)$$

Matrizen mit der Eigenschaft (4.31) heißen **hermitesch**.

Bemerkung 4.50 Jede hermitesche Matrix $S \in \mathbf{GL}_n(\mathbb{C})$ definiert ein unitäres Skalarprodukt auf \mathbb{C}^n durch

$$\langle x | y \rangle = x^\top \cdot S \cdot \overline{y}.$$

Bemerkung 4.51 Eine reelle symmetrische Matrix $S \in \mathbf{GL}_n(\mathbb{R})$ ist insbesondere hermitesch, da $S = S^\top = \overline{S}^\top$ gilt.

Satz 4.52 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein unitärer Vektorraum. Dann ist durch

$$\|x\| = \sqrt{\langle x | x \rangle} \quad (4.32)$$

eine Norm auf V definiert.

Durch die Norm können wir Abstände in V wie in Definition 4.35 definieren.

Die Cauchy-Schwarz-Ungleichung gilt, wenn $\langle x | y \rangle^2$ durch $|\langle x | y \rangle|^2$ ersetzt wird:

Satz 4.53 (Cauchy-Schwarz-Ungleichung) Es sei $(V, \langle \cdot | \cdot \rangle)$ ein unitärer Vektorraum. Für alle $x, y \in V$ gilt:

$$|\langle x | y \rangle|^2 \leq \langle x | x \rangle \cdot \langle y | y \rangle. \quad (4.33)$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Orthogonalität $x \perp y$ wird wieder durch $\langle x|y \rangle = 0$ definiert.

Satz 4.54 (Pythagoras) *Es seien x, y Elemente eines unitären Vektorraumes. Gilt $x \perp y$, so folgt*

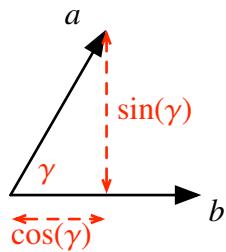
$$\|x\|^2 + \|y\|^2 = \|x + y\|^2. \quad (4.34)$$

Im Folgenden sprechen wir von einem **Vektorraum mit Skalarprodukt**, wenn V entweder ein euklidischer oder unitärer Vektorraum ist.

5 Orthogonalsysteme

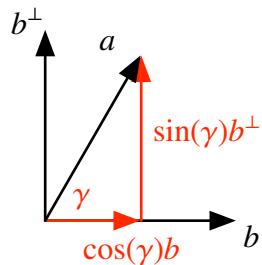
5.1 Mengen orthogonaler Vektoren

Der Cosinus des Winkels γ zwischen zwei *normierten* Vektoren a, b ist gerade $\langle a|b \rangle$. Das Skalarprodukt misst also, wie „verschieden“ die Richtungen von a und b sind. Dies ist konsistent mit der elementargeometrischen Beobachtung, dass die Komponente von a in Richtung b gerade die Länge $\cos(\gamma)$ hat.



Die Komponente senkrecht zu b hat die Länge $\sin(\gamma)$. Wie im Falle des Standard-skalarprodukts aus Abschnitt 4.1 liefert das Skalarprodukt somit eine Zerlegung von a in seine Komponenten parallel zu b und senkrecht zu b . Ist b^\perp ein Einheitsvektor senkrecht zu b , so gilt

$$a = \cos(\gamma)b + \sin(\gamma)b^\perp = \langle a|b \rangle b + \langle a|b^\perp \rangle b^\perp.$$



Wir können b^\perp sogar bestimmen, indem wir von a seine Komponente parallel zu b abziehen, $a - \langle a|b \rangle b$, und dann normieren:

$$b^\perp = \frac{a - \langle a|b \rangle b}{\|a - \langle a|b \rangle b\|}.$$

Wir wollen im Folgenden Systeme solcher paarweise orthogonaler Vektoren untersuchen, insbesondere Basen, und welche geometrischen Informationen wir durch sie gewinnen können.

Hilfssatz 5.1 Es sei V ein Vektorraum mit Skalarprodukt und $M \subset V \setminus \{0\}$ eine Menge, deren Elemente paarweise orthogonal zueinander seien (d.h. $\langle x|y \rangle = 0$ für alle $x, y \in M$). Dann ist M linear unabhängig. Insbesondere ist M endlich, falls $\dim V < \infty$.

BEWEIS: Es seien x_1, \dots, x_m verschiedene Elemente aus M und $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ (oder $\in \mathbb{C}$), so dass gilt

$$\lambda_1 x_1 + \dots + \lambda_m x_m = 0.$$

Nach Voraussetzung gilt für alle $i = 1, \dots, m$:

$$0 = \langle \lambda_1 x_1 + \dots + \lambda_m x_m | x_i \rangle = \lambda_1 \langle x_1 | x_i \rangle + \dots + \lambda_i \langle x_i | x_i \rangle + \dots + \lambda_m \langle x_m | x_i \rangle = \lambda_i \langle x_i | x_i \rangle.$$

Und da $\langle x_i | x_i \rangle > 0$, folgt $\lambda_i = 0$. Somit ist M linear unabhängig. ■

Es stellt sich nun die Frage, ob man in V eine ganze Basis bestehend aus paarweise orthogonalen Vektoren finden kann. Im Falle endlicher Dimension werden wir eine solche Basis im folgenden Abschnitt 5.2 konstruieren.

Definition 5.2 Es sei V ein Vektorraum mit Skalarprodukt. Eine Basis B von V heißt **Orthogonalbasis**, falls $b_1 \perp b_2$ gilt für alle verschiedenen $b_1, b_2 \in B$. Gilt zusätzlich $\|b\| = 1$ für alle $b \in B$, so sprechen wir von einer **Orthonormalbasis**.

Definition 5.3 Es sei V ein Vektorraum mit Skalarprodukt und $M \subset V$ eine Teilmenge. Dann heißt

$$M^\perp = \{x \in V \mid x \perp M\} \tag{5.1}$$

das **orthogonale Komplement** von M in V .

Hilfssatz 5.4 Es sei V ein Vektorraum mit Skalarprodukt und $N \subseteq M \subseteq V$ Teilmengen $\neq \emptyset$. Es gilt:

- (a) $N^\perp \supseteq M^\perp$.
- (b) M^\perp ist ein Untervektorraum von V .
- (c) $[M]^\perp = M^\perp$.
- (d) $[M] \subseteq (M^\perp)^\perp$.
- (e) Ist $\dim V = n$ und U ein Untervektorraum, so gilt $\dim U^\perp = n - \dim U$,

$$V = U \oplus U^\perp \tag{5.2}$$

$$\text{und } (U^\perp)^\perp = U.$$

Beweis:

- (a) Sei $x \in M^\perp$. Da $x \perp y$ für alle $y \in M$, gilt insbesondere $x \perp z$ für alle $z \in N$. Also ist $x \in N^\perp$.
- (b) Es ist $0 \in M^\perp$, also $M \neq \emptyset$. Für $x_1, x_2 \in M^\perp$ und $\lambda \in \mathbb{R}$ (oder $\in \mathbb{C}$) gilt für alle $y \in M$:

$$\langle x_1 + \lambda x_2 | y \rangle = \langle x_1 | y \rangle + \lambda \langle x_2 | y \rangle = 0 + \lambda \cdot 0 = 0.$$

Also ist $x_1 + \lambda x_2 \in M^\perp$.

- (c) Nach Teil (a) gilt $[M]^\perp \subseteq M^\perp$.

Sei nun $x \in M^\perp$ und $y = \lambda_1 y_1 + \dots + \lambda_k y_k$ für $y_1, \dots, y_k \in M$. Dann gilt

$$\langle y | x \rangle = \lambda_1 \langle y_1 | x \rangle + \dots + \lambda_k \langle y_k | x \rangle = 0.$$

Also ist $x \in [M]^\perp$. Folglich gilt auch $M^\perp \subseteq [M]^\perp$.

- (d) Folgt aus der Rechnung in Teil (c).

- (e) Ist $x \in U \cap U^\perp$, so gilt $\langle x | x \rangle = 0$. Wegen der positiven Definitheit ist $x = 0$.

Es sei b_1, \dots, b_k eine Basis von U . Nun bedeutet $y \perp U$, dass y eine Lösung des homogenen linearen Gleichungssystems

$$\langle b_1 | y \rangle = 0, \dots, \langle b_k | y \rangle = 0$$

ist. Dieses lineare Gleichungssystem besteht aus k linear unabhängigen Gleichungen und hat somit einen $n - k$ -dimensionalen Lösungsraum. Nach Definition ist dieser Lösungsraum U^\perp . Es folgt (5.2). Des Weiteren gilt $U \subseteq (U^\perp)^\perp$ und $\dim(U^\perp)^\perp = n - \dim U^\perp = n - (n - k) = k$. Also $U = (U^\perp)^\perp$. ■

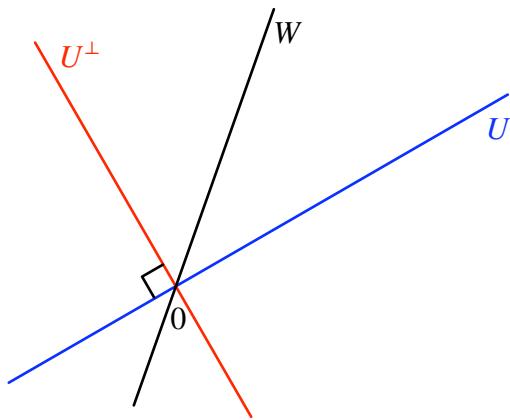
Aufgabe 5.5 Für Teilmengen M_1, M_2 und Untervektorräume U_1, U_2 eines Vektorraums V mit Skalarprodukt gilt:

- (a) $(M_1 \cup M_2)^\perp = M_1^\perp \cap M_2^\perp$.
- (b) $(U_1 \cap U_2)^\perp \supseteq U_1^\perp + U_2^\perp$.
- (c) Im Falle endlicher Dimension gilt Gleichheit in (b).

Bemerkung 5.6 Der Beweis von Hilfssatz 5.4 (e) zeigt insbesondere, dass U^\perp als Lösungsmenge des linearen Gleichungssystems

$$\langle b_1 | y \rangle = 0, \dots, \langle b_k | y \rangle = 0$$

eindeutig bestimmt ist. So kann man zu einem gegebenen Untervektorraum $U \subset V$ beliebig viele Komplementärräume W mit $W \oplus U = V$ finden, aber nur einen einzigen, der senkrecht auf U steht.



Beispiel 5.7 (Orthogonales Komplement) Es sei $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt des \mathbb{R}^4 . Die Vektoren

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^4$$

spannen den Untervektorraum $U = [u_1, u_2]$ im \mathbb{R}^4 auf. Die Elemente y des orthogonalen Komplements U^\perp müssen die Gleichungen

$$\begin{aligned} \langle u_1 | y \rangle &= u_1^\top \cdot y = 0, \\ \langle u_2 | y \rangle &= u_2^\top \cdot y = 0 \end{aligned}$$

erfüllen. Ausgeschrieben ist dies das lineare Gleichungssystem

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

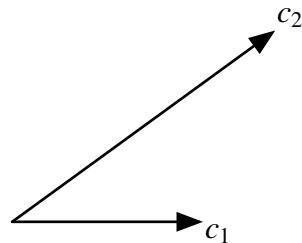
Der Lösungsraum dieses Systems ist

$$U^\perp = \left[\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right].$$

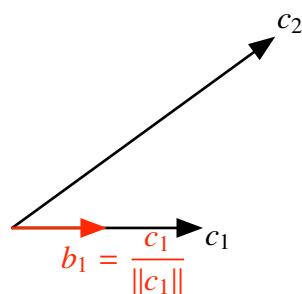
5.2 Das Gram-Schmidt-Verfahren

Ist $C = \{c_1, \dots, c_k\}$ eine Menge von linear unabhängigen Vektoren in einem Vektorraum V mit Skalarprodukt, so können wir aus den c_i eine Menge von paarweise orthogonalen und normierten Vektoren $B = \{b_1, \dots, b_k\}$ konstruieren, die den selben Untervektorraum aufspannen wie c_1, \dots, c_k .

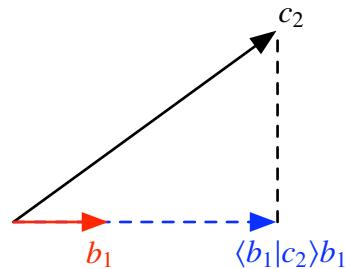
Wir illustrieren das Konstruktionsverfahren zunächst am Beispiel $C = \{c_1, c_2\}$.



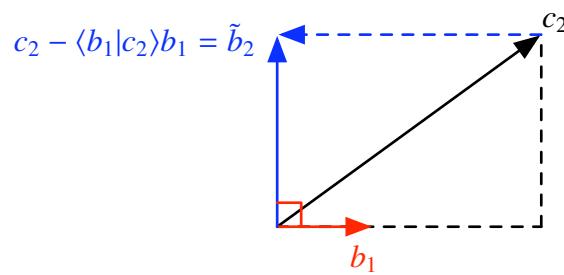
Im ersten Schritt wählt man einen Vektor aus M , etwa c_1 , und normiert ihn. Dann sei $b_1 = \frac{c_1}{\|c_1\|}$ das erste Element der Menge B .



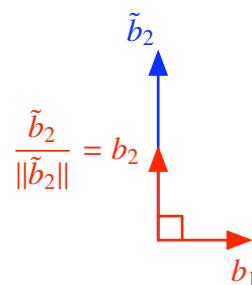
Nun konstruieren wir einen zu b_1 orthogonalen Vektor, der in der linearen Hülle von b_1, c_2 liegt. Dazu erinnern wir uns, dass $\langle b_1 | c_2 \rangle b_1$ die Komponente von c_2 in Richtung des Vektors b_1 ist.



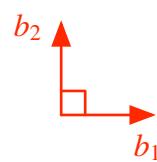
Wenn wir diese Komponente von c_2 abziehen, so erhalten wir einen Vektor $\tilde{b}_2 = c_2 - \langle b_1 | c_2 \rangle b_1$, dessen Komponente in b_1 -Richtung 0 ist. Das bedeutet $\tilde{b}_2 \perp b_1$.



Schließlich normieren wir den Vektor \tilde{b}_2 zu $b_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|}$.



So erhalten wir die gesuchte Menge $B = \{b_1, b_2\}$.



Der folgende Algorithmus erweitert dieses Verfahren auf Mengen mit k Vektoren:

Algorithmus 5.8 (Gram-Schmidt-Orthogonalisierung) Es sei V ein Vektorraum mit Skalarprodukt, $C = \{c_1, \dots, c_k\}$ eine Menge linear unabhängiger Vektoren in V und $U = [C]$ der von den c_i aufgespannte Untervektorraum. Berechne eine Menge paarweise orthogonaler Vektoren $\tilde{B} = \{\tilde{b}_1, \dots, \tilde{b}_k\}$ nach folgender Vorschrift:

(i) Setze

$$\tilde{b}_1 = c_1.$$

(ii) Für $r = 2, \dots, k$ setze

$$\tilde{b}_r = c_r - \sum_{i=1}^{r-1} \frac{\langle c_r | \tilde{b}_i \rangle}{\langle \tilde{b}_i | \tilde{b}_i \rangle} \tilde{b}_i \quad (5.3)$$

Dann gilt $[\tilde{B}] = [C]$, d.h. \tilde{B} ist eine Orthogonalbasis des Untervektorraums U . Normieren der \tilde{b}_i liefert eine Orthonormalbasis $B = \{b_1, \dots, b_k\}$ von U :

(iii) Für $r = 1, \dots, k$ setze

$$b_r = \frac{\tilde{b}_r}{\|\tilde{b}_r\|}. \quad (5.4)$$

BEWEIS: Der Beweis erfolgt durch Induktion über k . Als Induktionsvoraussetzung können wir $[c_1, \dots, c_{k-1}] = [\tilde{b}_1, \dots, \tilde{b}_{k-1}]$ annehmen, und dass $\tilde{b}_i \perp \tilde{b}_j$ gilt für $1 \leq i \neq j \leq k-1$. Nun folgt

$$\tilde{b}_k = c_k - \underbrace{\sum_{i=1}^{k-1} \frac{\langle \tilde{b}_i | c_k \rangle}{\langle \tilde{b}_i | \tilde{b}_i \rangle} \tilde{b}_i}_{\in [c_1, \dots, c_{k-1}]} \in [c_1, \dots, c_k].$$

Lösen wir (5.3) nach c_r auf, so finden wir

$$c_r = \tilde{b}_r + \sum_{i=1}^{r-1} \frac{\langle \tilde{b}_i | c_r \rangle}{\langle \tilde{b}_i | \tilde{b}_i \rangle} \tilde{b}_i \in [\tilde{b}_1, \dots, \tilde{b}_r] \subset [\tilde{b}_1, \dots, \tilde{b}_k].$$

Somit gilt $[\tilde{B}] = [C] = U$. Da C linear unabhängig ist und \tilde{B} ebenfalls eine Erzeugermenge mit k Elementen ist, ist auch \tilde{B} linear unabhängig. Es bleibt zu zeigen, dass $\tilde{b}_r \perp \tilde{b}_k$ für $r < k$ gilt:

$$\langle \tilde{b}_r | \tilde{b}_k \rangle = \langle \tilde{b}_r | c_k \rangle - \sum_{i=1}^{k-1} \frac{\langle \tilde{b}_i | c_k \rangle}{\langle \tilde{b}_i | \tilde{b}_i \rangle} \underbrace{\langle \tilde{b}_i | \tilde{b}_r \rangle}_{=0 \text{ für } i \neq r} = \langle \tilde{b}_r | c_k \rangle - \frac{\langle \tilde{b}_r | c_k \rangle}{\langle \tilde{b}_r | \tilde{b}_r \rangle} \langle \tilde{b}_r | \tilde{b}_r \rangle = 0.$$

Somit ist \tilde{B} eine Orthogonalbasis von U . Es folgt sofort, dass B eine Orthonormalbasis von U ist. ■

Satz 5.9 Es sei V ein Vektorraum mit Skalarprodukt und $\dim V < \infty$. Jede Teilmenge orthogonaler Vektoren von V lässt sich zu einer Orthogonalbasis von V ergänzen. Es existiert eine Orthonormalbasis $\{b_1, \dots, b_n\}$ von V , und jedes $x \in V$ wird in dieser Basis wie folgt dargestellt:

$$x = \langle x|b_1\rangle b_1 + \dots + \langle x|b_n\rangle b_n. \quad (5.5)$$

BEWEIS: Jede Teilmenge orthogonaler Vektoren ist linear unabhängig, lässt sich also zu einer Basis von V ergänzen. Aus dieser Basis kann man mit dem Gram-Schmidt-Verfahren eine Orthonormalbasis $\{b_1, \dots, b_n\}$ konstruieren.

Ist $x = \lambda_1 b_1 + \dots + \lambda_n b_n$, so gilt

$$\langle x|b_1\rangle = \lambda_1 \langle b_1|b_1\rangle + \lambda_2 \langle b_2|b_1\rangle + \dots + \lambda_n \langle b_n|b_1\rangle = \lambda_1,$$

und entsprechend $\lambda_i = \langle x|b_i\rangle$ für $i = 2, \dots, n$. ■

Folgerung 5.10 Es seien C, \tilde{B} Basen von U wie in Algorithmus 5.8. Die Basiswechselmatrix $M_{\tilde{B}}^C$ ist eine obere Dreiecksmatrix mit 1 auf der Diagonale und Einträgen in Zeile i , Spalte j :

$$\frac{\langle \tilde{b}_i|c_j\rangle}{\langle \tilde{b}_i|\tilde{b}_i\rangle} \quad \text{für } i < j$$

und 0 für $i > j$. Die Basiswechselmatrix M_B^C ist eine obere Dreiecksmatrix mit Einträgen $\|\tilde{b}_1\|, \dots, \|\tilde{b}_k\|$ auf der Diagonalen.

BEWEIS: Aus (5.3) folgt für $j = 2, \dots, k$:

$$c_j = \tilde{b}_j + \frac{\langle \tilde{b}_{j-1}|c_j\rangle}{\langle \tilde{b}_{j-1}|\tilde{b}_{j-1}\rangle} \tilde{b}_{j-1} + \dots + \frac{\langle \tilde{b}_1|c_j\rangle}{\langle \tilde{b}_1|\tilde{b}_1\rangle} \tilde{b}_1.$$

Dies liefert die Darstellung der Vektoren aus C in der Basis \tilde{B} .

Die Basiswechselmatrix von \tilde{B} zu B ist die Diagonalmatrix

$$M_B^{\tilde{B}} = \begin{pmatrix} \|\tilde{b}_1\| & & 0 \\ & \ddots & \\ 0 & & \|\tilde{b}_k\| \end{pmatrix}.$$

Daher ist $M_B^C = M_B^{\tilde{B}} \cdot M_{\tilde{B}}^C$ eine obere Dreiecksmatrix mit den Diagonaleinträgen $\|\tilde{b}_1\|, \dots, \|\tilde{b}_k\|$. ■

Beispiel 5.11 (Orthonormalbasis) Es sei $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt des \mathbb{R}^4 .

Die Vektoren

$$c_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, c_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, c_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{R}^4$$

spannen den Untervektorraum $U = [c_1, c_2, c_3]$ im \mathbb{R}^4 auf. Sie bilden jedoch keine Orthogonalbasis von U . Mit dem Gram-Schmidt-Verfahren bestimmen wir nun eine Orthonormalbasis. Wähle

$$\tilde{b}_1 = c_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Dann ist

$$\tilde{b}_2 = c_2 - \frac{\langle \tilde{b}_1 | c_2 \rangle}{\langle \tilde{b}_1 | \tilde{b}_1 \rangle} \tilde{b}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{(-1)}{3} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix},$$

und

$$\begin{aligned} \tilde{b}_3 &= c_3 - \frac{\langle \tilde{b}_1 | c_3 \rangle}{\langle \tilde{b}_1 | \tilde{b}_1 \rangle} \tilde{b}_1 - \frac{\langle \tilde{b}_2 | c_3 \rangle}{\langle \tilde{b}_2 | \tilde{b}_2 \rangle} \tilde{b}_2 \\ &= \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} - \frac{(-1)}{3} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} - \frac{2}{3} \cdot \frac{3}{2} \cdot \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}. \end{aligned}$$

Die Basiswechselmatrix $M_{\tilde{B}}^C$ aus Folgerung 5.10 ist

$$M_{\tilde{B}}^C = \begin{pmatrix} 1 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Durch Normieren von $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3$ erhalten wir eine Orthonormalbasis von U :

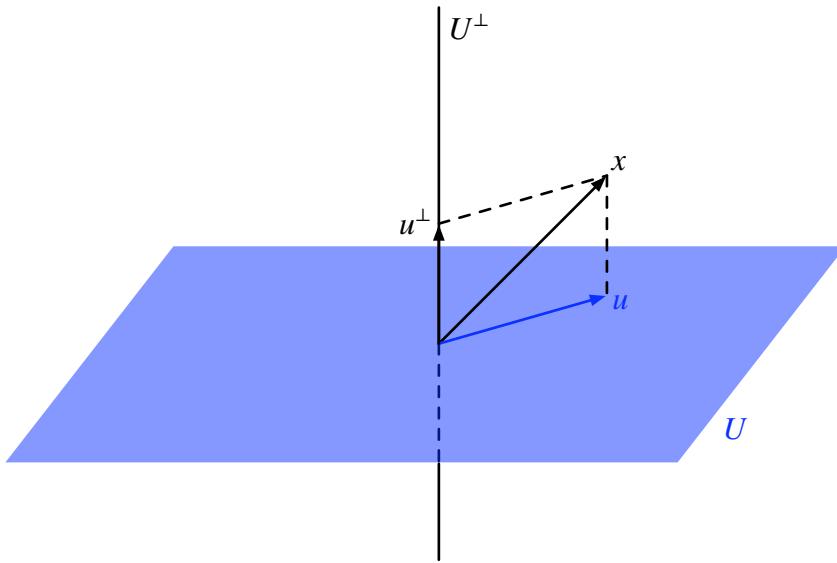
$$b_1 = \frac{1}{\sqrt{3}} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, b_2 = \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, b_3 = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

5.3 Orthogonalprojektionen

Nach Hilfssatz 5.4 (e) ist $V = U \oplus U^\perp$ für einen n -dimensionalen Vektorraum V mit Untervektorraum U . Jedes $x \in V$ lässt sich folglich in eindeutiger Weise zerlegen in

$$x = u + u^\perp \quad (5.6)$$

mit $u \in U$, $u^\perp \in U^\perp$.



Ist $\{b_1, \dots, b_k\}$ eine Orthonormalbasis von U , so können wir diese durch eine Orthonormalbasis $\{b_{k+1}, \dots, b_n\}$ von U^\perp zu einer Basis von V ergänzen. In dieser Basis lässt sich x wie folgt darstellen:

$$x = \underbrace{\langle x|b_1\rangle b_1 + \dots + \langle x|b_k\rangle b_k}_{\in U} + \underbrace{\langle x|b_{k+1}\rangle b_{k+1} + \dots + \langle x|b_n\rangle b_n}_{\in U^\perp}.$$

Wegen der Eindeutigkeit der Zerlegung gilt somit

$$u = \langle x|b_1\rangle b_1 + \dots + \langle x|b_k\rangle b_k, \quad u^\perp = \langle x|b_{k+1}\rangle b_{k+1} + \dots + \langle x|b_n\rangle b_n.$$

Dies veranlasst uns zur folgenden Definition, in der wir nicht verlangen, dass $\dim V < \infty$ ist:

Definition 5.12 Es sei V ein Vektorraum mit Skalarprodukt und U ein k -dimensionaler Untervektorraum mit Orthonormalbasis $\{b_1, \dots, b_k\}$. Die Abbildung

$$\Pi_U : V \rightarrow V, \quad x \mapsto \langle x|b_1\rangle b_1 + \dots + \langle x|b_k\rangle b_k \quad (5.7)$$

wird **Orthogonalprojektion auf U** genannt.

Satz 5.13 Es sei V ein Vektorraum mit Skalarprodukt und U ein k -dimensionaler Untervektorraum von V .

- (a) Π_U ist linear.
- (b) $\Pi_U|_U = \text{id}_U$, Bild $\Pi_U = U$ und $U^\perp \subseteq \text{Kern } \Pi_U$.
- (c) $\Pi_U^2 = \Pi_U$.
- (d) Ist $\dim V < \infty$, so ist $U^\perp = \text{Kern } \Pi_U$ und $\Pi_{U^\perp} = \text{id}_V - \Pi_U$.
- (e) Ist $\dim V < \infty$, so gilt $\|\Pi_U(x)\| \leq \|x\|$ für alle $x \in V$.

Beweis:

- (a) Folgt aus der Linearität der $\langle \cdot | b_i \rangle$, $i = 1, \dots, k$.
- (b) Jedes $u \in U$ lässt sich in der Form $u = \lambda_1 b_1 + \dots + \lambda_k b_k$ schreiben. Also ist

$$\Pi_U(u) = \sum_{i=1}^k \sum_{j=1}^k \lambda_i \underbrace{\langle b_i | b_j \rangle}_{=\delta_{ij}} b_j = \sum_{i=1}^k \lambda_i b_i = u,$$

d.h. $\Pi_U|_U = \text{id}_U$. Da $\Pi_U(x)$ für alle $x \in V$ eine Linearkombination von Elementen aus U ist, folgt Bild $\Pi_U = U$. Für $u^\perp \in U^\perp$ ist

$$\Pi_U(u^\perp) = \langle u^\perp | b_1 \rangle b_1 + \dots + \langle u^\perp | b_k \rangle b_k = 0.$$

- (c) Folgt sofort aus Teil (b).
- (d) Es sei $x \in V$. Gemäß (5.6) ist $x = u + u^\perp$. Dann gilt:

$$\Pi_{U^\perp}(x) = u^\perp = x - u = (\text{id}_V - \Pi_U)(x).$$

Sei nun $x \in \text{Kern } \Pi_U$. Nach Teil (b) gilt

$$0 = \Pi_U(x) = \Pi_U(u) + \Pi_U(u^\perp) = u.$$

Also $x = u^\perp \in U^\perp$.

- (e) Nach dem Satz von Pythagoras gilt für alle $x \in V$:

$$\|x\|^2 = \|u\|^2 + \|u^\perp\|^2 \geq \|u\|^2$$

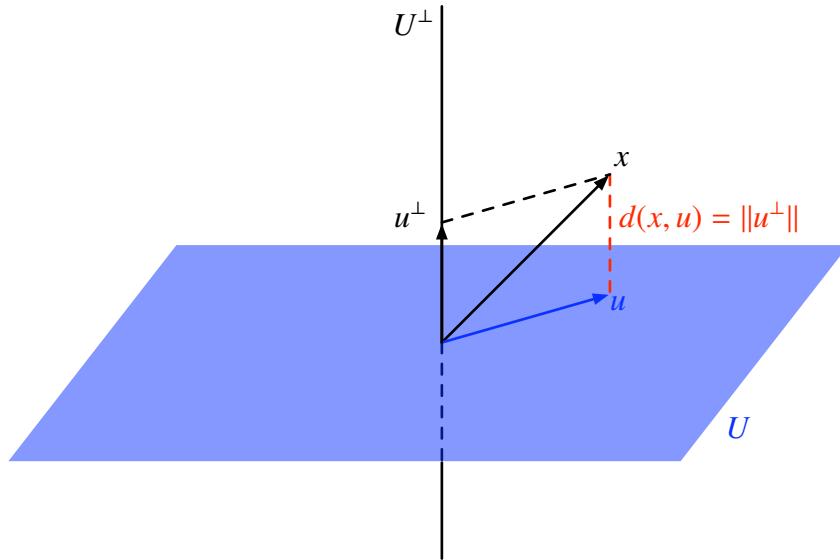
für u, u^\perp wie in (5.6). ■

Aufgabe 5.14 Im \mathbb{R}^n mit Standardskalarprodukt wird die orthogonale Projektion auf einen Untervektorraum U mit Basis b_1, \dots, b_k durch die Matrix

$$P_U = b_1 \cdot b_1^\top + \dots + b_k \cdot b_k^\top$$

dargestellt.

Für den Rest des Abschnitts nehmen wir $\dim V = n$ an. Die orthogonale Projektion erlaubt es uns, den Abstand $d(x, U)$ eines Vektors $x \in V$ zum Untervektorraum U zu berechnen (vgl. Definition 4.35).



Definition 5.15 Es sei V ein Vektorraum mit Skalarprodukt, $\dim V = n$, und U ein Untervektorraum von V . Dann heißt $\Pi_U(x)$ der **Lotfußpunkt** von x auf U , und $x - \Pi_U(x)$ das **Lot** (oder der **Lotvektor**) von x auf U .

Satz 5.16 Ist $\{b_1, \dots, b_k\}$ eine Orthonormalbasis von U , so ist

$$d(x, U)^2 = d(x, \Pi_U(x))^2 = \|x\|^2 - \sum_{i=1}^k |\langle x | b_i \rangle|^2. \quad (5.8)$$

BEWEIS: Zuerst ist zu zeigen, dass $\Pi_U(x)$ unter allen $u \in U$ den Abstand $d(x, u)$ minimiert: Für alle $u \in U$ ist $u - \Pi_U(x) \in U$ und folglich

$$u - \Pi_U(x) \perp x - \Pi_U(x) = \Pi_{U^\perp}(x).$$

Nach dem Satz des Pythagoras ist

$$\begin{aligned} d(x, u)^2 &= \|x - u\|^2 = \|x - \Pi_U(x) + \Pi_U(x) - u\|^2 \\ &= \|x - \Pi_U(x)\|^2 + \|\Pi_U(x) - u\|^2 \\ &= d(x, \Pi_U(x))^2 + \underbrace{\|\Pi_U(x) - u\|^2}_{\geq 0} \end{aligned}$$

Dieser Ausdruck wird minimal, wenn der letzte Summand = 0 ist, also für $u = \Pi_U(x)$. Somit gilt $d(x, U) = d(x, \Pi_U(x))$.

Es bleibt die zweite Gleichheit in (5.8) zu zeigen:

$$d(x, \Pi_U(x))^2 = \|x\|^2 - \|\Pi_U(x)\|^2 = \|x\|^2 - \left\| \sum_{i=1}^k \langle x | b_i \rangle b_i \right\|^2 = \|x\|^2 - \sum_{i=1}^k |\langle x | b_i \rangle|^2,$$

wobei für die letzte Gleichheit wiederholt der Satz des Pythagoras und die Orthogonalität der b_i verwendet wird. Die Betragsstriche um $|\langle x | b_i \rangle|^2$ schreiben wir, um auch den unitären Fall abzudecken. ■

Aufgabe 5.17 Es sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt. Es sei $A \in \mathbb{R}^{n \times k}$ die Matrix mit Spaltenvektoren $a_1, \dots, a_k \in \mathbb{R}^n$ und $U = [a_1, \dots, a_k] \subset \mathbb{R}^n$.

- (a) Es ist $x \in \mathbb{R}^n$ genau dann in U^\perp , wenn $A^\top x = 0$ gilt.
- (b) Für alle $x \in \mathbb{R}^n$ existiert $\xi \in \mathbb{R}^k$, so dass $A^\top A\xi = A^\top x$ und für alle $\zeta \in \mathbb{R}^k$ gilt $\|A\xi - x\| \leq \|A\zeta - x\|$.

Aufgabe 5.18 Für $a_1, \dots, a_k \in V$ heißt

$$g(a_1, \dots, a_k) = \det \begin{pmatrix} \langle a_1 | a_1 \rangle & \cdots & \langle a_1 | a_k \rangle \\ \vdots & \ddots & \vdots \\ \langle a_k | a_1 \rangle & \cdots & \langle a_k | a_k \rangle \end{pmatrix} \quad (5.9)$$

die **Gramsche Determinante** von a_1, \dots, a_k . Ist U ein Untervektorraum mit Basis a_1, \dots, a_k , so gilt für jedes $x \in V$:

$$d(x, U)^2 = \frac{g(a_1, \dots, a_k, x)}{g(a_1, \dots, a_k)}. \quad (5.10)$$

5.4 Vollständige Orthogonalsysteme und Hilbert-Räume

Satz 5.9 über die Existenz einer Orthonormalbasis gilt im Allgemeinen nicht in unendlichdimensionalen Vektorräumen V . Um mit solchen Räumen arbeiten zu

können, ist es dennoch wünschenswert, eine Darstellung ähnlich der in Gleichung (5.5) für alle $x \in V$ zu haben. Eine Vektorraumbasis von V kann überabzählbar viele Elemente haben, und in vielen Fällen wird man sie nicht einmal explizit angeben können. Ist V mit einem Skalarprodukt ausgestattet und erfüllt eine zusätzliche Bedingung (vgl. Definition 5.22), so können wir jedoch einen Ersatz für die unhandlichen Vektorraumbasen finden: Es existiert dann eine Menge B von paarweise orthogonalen Vektoren, so dass jedes $x \in V$ in der Form

$$x = \sum_{b \in B} \langle x | b \rangle b \quad (5.11)$$

dargestellt werden kann. Hierbei tritt die Schwierigkeit auf, dass die Menge B unendlich sein kann, und somit die Summation $\sum_{b \in B}$ eine unendliche Reihe darstellt, deren Konvergenz sicherzustellen ist. Konvergenz ist dabei im Sinne der folgenden Definition zu verstehen:

Definition 5.19 Es sei V ein Vektorraum mit einer Norm $\|\cdot\|$. Eine Folge $(x_n)_{n \in \mathbb{N}}$ in V **konvergiert** in der Norm $\|\cdot\|$ gegen $x \in V$, wenn gilt:

$$\|x_n - x\| \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (5.12)$$

Die Darstellung (5.11) ist auch keine Linearkombination der $b \in B$ (da eine solche nach Definition nur endlich viele Summanden haben dürfte), d.h. B ist keine Vektorraumbasis im üblichen Sinne. Wir definieren daher:

Definition 5.20 Es sei V ein Vektorraum mit Skalarprodukt. Eine Menge $B \subset V$, deren Elemente normiert und paarweise orthogonal sind, heißt **Orthonormalsystem**. Gilt außerdem

$$x = \sum_{b \in B} \langle x | b \rangle b \quad (5.13)$$

für alle $x \in V$ (unabhängig von der Summationsreihenfolge), so nennen wir B ein **vollständiges Orthonormalsystem**.

Das Teilgebiet der Mathematik, das sich mit dem Studium unendlichdimensionaler Funktionsvektorräume befasst, ist die *Funktionalanalysis*. Dort ist es üblich, vollständige Orthonormalsysteme als *Orthonormalbasen* zu bezeichnen, auch dann, wenn sie keine Vektorraumbasen sind.

Definition 5.21 Eine Folge $(x_n)_{n \in \mathbb{N}}$ in einem Vektorraum V mit Norm $\|\cdot\|$ heißt **Cauchy-Folge**, falls für alle $\varepsilon > 0$ ein Index k existiert, so dass für alle $n, m > k$ gilt: $\|x_n - x_m\| < \varepsilon$.

Definition 5.22 Ein Vektorraum V mit Skalarprodukt heißt **Hilbert-Raum**, falls jede Cauchy-Folge in V gegen ein Element aus V konvergiert.

Der folgende Satz erklärt, warum Hilbert-Räume für uns interessant sind:

Satz 5.23 Es sei V ein Hilbert-Raum. Dann existiert ein vollständiges Orthonormalsystem B in V , und für alle $x \in V$ gilt

$$x = \sum_{b \in B} \langle x | b \rangle b, \quad (5.14)$$

wobei diese Reihe unbedingt⁷⁾ konvergiert.

Zum Beweis siehe Satz V.I.8 und V.I.9 in Werner [14]. Es lässt sich zeigen, dass in der Summe (5.14) nur abzählbar viele $\langle x | b \rangle$ verschieden von 0 sind (Werner [14], Lemma V.4.5). In den meisten interessanten Hilbert-Räumen gibt es ein abzählbares vollständiges Orthonormalsystem. Solche Hilbert-Räume heißen **separabel**.

Beispiel 5.24 (Hilbert-Räume)

- (a) Jeder endlichdimensionale \mathbb{R} - oder \mathbb{C} -Vektorraum mit Skalarprodukt ist ein Hilbert-Raum (aus der Analysis ist bekannt, dass Cauchy-Folgen im \mathbb{R}^n konvergieren). Jede Orthonormalbasis ist ein vollständiges Orthonormalsystem.
- (b) Der Vektorraum der quadratsummierbaren komplexen Folgen,

$$\ell^2(\mathbb{C}) = \left\{ (a_n)_{n \in \mathbb{N}} \mid \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}$$

mit dem Skalarprodukt

$$\langle a | b \rangle = \sum_{i=1}^{\infty} a_i \bar{b}_i \quad (5.15)$$

ist ein Hilbert-Raum: Dass die Reihe in (5.15) konvergiert, kann man mit der Cauchy-Schwarz-Ungleichung nachweisen. Dann rechnet man direkt nach, dass es sich um ein unitäres Skalarprodukt handelt. Wir zeigen nun, dass jede Cauchy-Folge $(x_n)_{n \in \mathbb{N}}$ konvergiert: Nun ist $(x_n)_{n \in \mathbb{N}}$ eine Folge von Folgen, d.h. $x_n = (x_{n,i})_{i \in \mathbb{N}}$. Es seien $\varepsilon > 0$ und k so, dass für $m, n > k$ gilt

$$\|x_n - x_m\| = \sqrt{\sum_{i=1}^{\infty} |x_{n,i} - x_{m,i}|^2} < \varepsilon.$$

⁷⁾Das bedeutet, dass die Konvergenz nicht von der Summationsreihenfolge abhängt.

Dann gilt auch für die einzelnen Summanden $|x_{n,i} - x_{m,i}| < \varepsilon$. Also sind für alle i die Folgen $(x_{n,i})_{n \in \mathbb{N}}$ Cauchy-Folgen in \mathbb{C} , die in \mathbb{C} gegen ein Element $x_i \in \mathbb{C}$ konvergieren. Dann ist die Folge $x = (x_i)_{i \in \mathbb{N}}$ ein Element von $\ell^2(\mathbb{C})$ und der Grenzwert von $(x_n)_{n \in \mathbb{N}}$: Für $\varepsilon > 0$ und festes $r \in \mathbb{N}$ gibt es k , so dass für $m, n > k$ gilt $\sum_{i=1}^r |x_{n,i} - x_{m,i}|^2 < \varepsilon$. Für $m \rightarrow \infty$ gilt wegen der Stetigkeit des Betrags $\sum_{i=1}^r |x_{n,i} - x_i|^2 \leq \varepsilon$. Da dies für alle r gilt, folgt

$$\sum_{i=1}^{\infty} |x_{n,i} - x_i|^2 \leq \varepsilon.$$

Also ist die Folge $z = (x_{n,i} - x_i)_{i \in \mathbb{N}}$ ein Element von $\ell^2(\mathbb{C})$. Da $\ell^2(\mathbb{C})$ ein Vektorraum ist, ist auch

$$x = x_n - z \in \ell^2(\mathbb{C}),$$

und da für $n > k$ gilt $\|x_n - x\| \leq \varepsilon$, folgt $x_n \rightarrow x$ für $n \rightarrow \infty$.

Ein vollständiges Orthonormalsystem in $\ell^2(\mathbb{C})$ ist durch die Folgen

$$e_i = (0, \dots, 0, 1, 0, \dots),$$

mit 1 an der i ten Stelle, gegeben.

Der reelle Hilbert-Raum $\ell^2(\mathbb{R})$ ist analog definiert, wobei man in (5.15) auf die komplexe Konjugation verzichten kann.

Bemerkung 5.25 In der Theorie der (separablen) Hilbert-Räume spielen die ℓ^2 -Räume eine ähnliche Rolle wie Räume \mathbb{R}^n und \mathbb{C}^n in der Theorie der endlich-dimensionalen Vektorräumen. Hat man nämlich in V ein abzählbares vollständiges Orthonormalsystem B gewählt, so wird $x \in V$ durch die Reihe

$$x = \sum_{i=1}^{\infty} \langle x | b_i \rangle b_i$$

dargestellt. Dann ist x bereits vollständig durch die Folge der Koeffizienten $x_i = \langle x | b_i \rangle$ bestimmt. Die Zuordnung

$$\Theta_B : V \rightarrow \ell^2(\mathbb{C}), \quad x \mapsto (x_1, x_2, x_3, \dots)$$

ist ein Isomorphismus und nach der Parsevalschen Gleichung (5.18) sogar eine Isometrie (siehe Abschnitt 6.1). Dies entspricht der Koordinatendarstellung in endlicher Dimension durch Spaltenvektoren im \mathbb{C}^n .

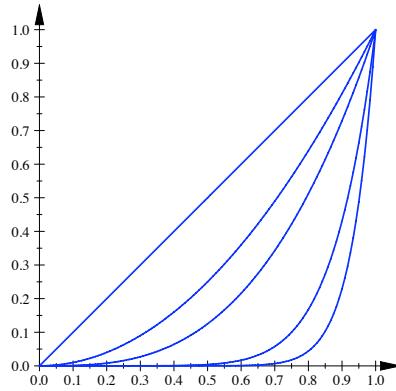
Nicht alle unendlichdimensionalen Vektorräume mit Skalarprodukt sind Hilbert-Räume, wie das folgende Beispiel zeigt.

Beispiel 5.26 Der Vektorraum $C([0, 1])$ mit dem Skalarprodukt

$$\langle f|g \rangle = \int_0^1 f(t)g(t)dt$$

ist *kein* Hilbert-Raum. Betrachte dazu die Folge $(f_n)_{n \in \mathbb{N}}$ der Funktionen

$$f_n(t) = t^n.$$



Dies ist eine Cauchy-Folge. Sie konvergiert punktweise gegen die Funktion

$$f(t) = \begin{cases} 0, & t \in [0, 1) \\ 1, & t = 1 \end{cases}.$$

Da

$$\|f_n - f\|^2 = \int_0^1 (f_n(t) - f(t))^2 dt \rightarrow 0$$

gilt, konvergiert f_n auch in der Norm $\|\cdot\|$ gegen f . Aber f ist nicht stetig und somit kein Element von $C([0, 1])$.

Dennoch gibt es ein vollständiges Orthonormalsystem B in $C([0, 1])$. Dies lässt sich damit erklären, dass $C([0, 1])$ ein Untervektorraum eines größeren Hilbert-Raumes V ist, der auch unstetige Funktionen beinhaltet. In diesem Hilbert-Raum V ist B ein vollständiges Orthonormalsystem, somit auch in $C([0, 1])$. Ein Beispiel für solch ein B ist

$$B = \{1_{[0,1]}\} \cup \{\sqrt{2} \sin(2\pi n \cdot) \mid n \in \mathbb{N}\} \cup \{\sqrt{2} \cos(2\pi n \cdot) \mid n \in \mathbb{N}\}, \quad (5.16)$$

vgl. auch Beispiel 4.42 (d). Die Darstellung (5.14) einer Funktion $f \in C([0, 1])$ (oder $f \in V$)

$$f = \langle f | 1_{[0,1]} \rangle 1_{[0,1]} + \sum_{n=1}^{\infty} 2 \langle f | \sin(2\pi n \cdot) \rangle \sin(2\pi n \cdot) + \sum_{n=1}^{\infty} 2 \langle f | \cos(2\pi n \cdot) \rangle \cos(2\pi n \cdot)$$

ist als **Fourier-Reihe** von f bekannt. Die Skalarprodukte $\langle f|1_{[0,1]}\rangle$, $2\langle f|\sin(2\pi n \cdot)\rangle$ und $2\langle f|\cos(2\pi n \cdot)\rangle$ heißen **Fourier-Koeffizienten** von f . Die Reihe konvergiert in der Norm $\|\cdot\|$. Dies zu beweisen erfordert aufwändige Hilfsmittel aus der Analysis, daher sei auf Körner [9], Theorem 34.1, verwiesen. Es gibt stetige Funktionen, deren Fourier-Reihe nicht auf ganz $[0, 1]$ punktweise gegen f konvergiert. Ein Beispiel für eine Fourier-Reihe einer stetigen Funktion, die an einem einzelnen Punkt sogar divergiert, findet sich in Kapitel 18 von Körner [9]. Ist f jedoch zweimal stetig differenzierbar, so lässt sich zeigen, dass die Fourier-Reihe von f gleichmäßig gegen f konvergiert (Körner [9], Theorem 9.6).

Aufgabe 5.27 Es sei V ein Hilbert-Raum mit einem abzählbaren Orthonormalen System B . Weiter seien $x, y \in V$:

(a) Es gilt die **Besselsche Ungleichung**

$$\sum_{i=1}^{\infty} |\langle x|b_i\rangle|^2 \leq \|x\|^2. \quad (5.17)$$

(b) Ist B vollständig, so gelten die **Parsevalsche Gleichung**

$$\sum_{i=1}^{\infty} |\langle x|b_i\rangle|^2 = \|x\|^2 \quad (5.18)$$

und die **verallgemeinerte Parsevalsche Gleichung**

$$\sum_{i=1}^{\infty} \langle x|b_i\rangle \langle b_i|y\rangle = \langle x|y\rangle. \quad (5.19)$$

Insbesondere gelten (5.17), (5.18) und (5.19) in endlicher Dimension.

5.5 Positiv definite Matrizen

Wir wenden die bisherigen Resultate dieses Kapitels an, um Kriterien für die positive Definitheit von symmetrischen Matrizen herzuleiten.

In diesem Abschnitt sei S stets eine reelle symmetrische $n \times n$ -Matrix mit Einträgen s_{ij} für $1 \leq i, j \leq n$. Es bezeichne S_k die obere linke $k \times k$ -Untermatrix von S mit Einträgen s_{ij} für $1 \leq i, j \leq k$ und $k \leq n$.

Beispiel 5.28 Ist

$$S = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 3 \end{pmatrix} \in \mathbb{R}^{3 \times 3},$$

so haben wir

$$S_1 = (2) \in \mathbb{R}^{1 \times 1}, \quad S_2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \quad S_3 = S \in \mathbb{R}^{3 \times 3}.$$

Hilfssatz 5.29 Ist S positiv definit, so sind auch S_1, \dots, S_{n-1} positiv definit.

BEWEIS: S definiert ein Skalarprodukt auf \mathbb{R}^n und S_k beschreibt die Einschränkung dieses Skalarprodukts auf den Untervektorraum $[e_1, \dots, e_k]$. Daher muss auch S_k positiv definit sein. ■

Satz 5.30 Es sei $S \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Die folgenden Aussagen sind äquivalent:

(i) S ist positiv definit.

(ii) Es existiert eine obere Dreiecksmatrix $R \in \mathbf{GL}_n(\mathbb{R})$ mit positiven Einträgen auf der Diagonalen, so dass gilt:

$$S = R^\top \cdot R. \quad (5.20)$$

(iii) Für $k = 1, \dots, n$ gilt:

$$\det(S_k) > 0 \quad (5.21)$$

BEWEIS: (i) \Rightarrow (ii): Durch die Vorschrift $\langle x|y \rangle = x^\top S y$ ist ein Skalarprodukt auf \mathbb{R}^n definiert. Nach Satz 5.9 existiert eine Orthonormalbasis B von \mathbb{R}^n für dieses Skalarprodukt. In dieser Basis B wird S durch die Einheitsmatrix I_n dargestellt. Bezeichnet C die Standardbasis, so ist nach Folgerung 5.10 die Basiswechselmatrix M_B^C eine obere Dreiecksmatrix mit gewissen Einträgen $\beta_i > 0$ auf der Diagonalen. Es gilt nach Satz 4.14:

$$S = (M_B^C)^\top \cdot I_n \cdot M_B^C.$$

Wählen wir $R = M_B^C$, so gilt (5.20). Es gilt $R \in \mathbf{GL}_n(\mathbb{R})$, da $R^{-1} = M_C^B$.

(ii) \Rightarrow (i): Es gelte $S = R^\top \cdot R$. Dann ist

$$S^\top = (R^\top \cdot R)^\top = R^\top \cdot (R^\top)^\top = R^\top \cdot R = S.$$

Also ist S symmetrisch. Für alle $x \in \mathbb{R}^n$, $x \neq 0$, gilt

$$x^\top \cdot S \cdot x = x^\top \cdot R^\top \cdot R \cdot x = (Rx)^\top \cdot (Rx) > 0,$$

denn der letzte Ausdruck ist gerade das Standardskalarprodukt von Rx mit sich selbst. Folglich ist S positiv definit.

(i) \Rightarrow (iii): Ist S positiv definit, so sind auch die S_k positiv definit (Hilfssatz 5.29). Wegen der bereits bewiesenen Äquivalenz (i) \Leftrightarrow (ii) existieren obere Dreiecksmatrizen $R_k \in \mathbf{GL}_n(\mathbb{R})$ mit positiven Diagonaleinträgen, die $S_k = R_k^\top \cdot R_k$ erfüllen. Daraus folgt

$$\det(S_k) = \det(R_k^\top \cdot R_k) = \det(R_k^\top) \det(R_k) = \det(R_k) \det(R_k) = \det(R_k)^2 > 0.$$

(iii) \Rightarrow (ii): Beweis durch vollständige Induktion über n . Für $n = 1$ ist $S = (s_{11})$ mit $s_{11} > 0$. Setze $R = (r_{11})$ mit $r_{11} = \sqrt{s_{11}}$. Sei nun $n > 1$. Nach Induktionsannahme existiert für S_{n-1} eine obere Dreiecksmatrix $R_{n-1} \in \mathbf{GL}_{n-1}(\mathbb{R})$ mit positiven Diagonaleinträgen, so dass $S_{n-1} = R_{n-1}^\top \cdot R_{n-1}$ gilt. Wir machen den Ansatz

$$R = \begin{pmatrix} R_{n-1} & | & y \\ 0 & | & \beta \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad y \in \mathbb{R}^{n-1}, \beta \in \mathbb{R},$$

wobei y und β so zu bestimmen sind, dass $S = R^\top \cdot R$ gilt. S hat die Gestalt

$$S = \begin{pmatrix} S_{n-1} & | & x \\ x^\top & | & \alpha \end{pmatrix}, \quad x \in \mathbb{R}^n, \alpha \in \mathbb{R}.$$

Dann gilt:

$$R^\top \cdot R = \begin{pmatrix} R_{n-1}^\top & | & 0 \\ y^\top & | & \beta \end{pmatrix} \cdot \begin{pmatrix} R_{n-1} & | & y \\ 0 & | & \beta \end{pmatrix} = \begin{pmatrix} R_{n-1}^\top \cdot R_{n-1} & | & R_{n-1}^\top \cdot y \\ y^\top \cdot R_{n-1} & | & y^\top \cdot y + \beta^2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} S_{n-1} & | & x \\ x^\top & | & \alpha \end{pmatrix}.$$

Es muss somit $x = R_{n-1}^\top y$ gelten, und da R_{n-1} nach Voraussetzung invertierbar ist, gilt $y = (R_{n-1}^\top)^{-1} x$. Des Weiteren muss $y^\top y + \beta^2 = \alpha$ gelten. Setze $\beta = \sqrt{\alpha - y^\top y}$. Für die Determinante von S gilt

$$0 < \det(S) = \det(R_{n-1})^2 \beta^2,$$

also gilt $\beta = \sqrt{\frac{\det(S)}{\det(R_{n-1})^2}}$, und da der Ausdruck unter der Wurzel eine positive reelle Zahl ist, ist auch β eine positive reelle Zahl. Somit ist die Existenz von R mit positiven Diagonaleinträgen gezeigt. ■

Das Kriterium (iii) in Satz 5.30 ist als **Hurwitz-Kriterium** für positive Definitheit bekannt. Die Zerlegung (5.20) aus Satz 5.30,

$$S = R^\top \cdot R,$$

wird **Cholesky-Zerlegung** genannt. Sie ist die Grundlage für ein numerisches Verfahren zur Lösung linearer Gleichungssysteme für positiv definite Matrizen. Der folgende Algorithmus berechnet die Cholesky-Zerlegung für positiv definites S und liefert einen Fehler, falls S nicht positiv definit ist.

Algorithmus 5.31 (Cholesky-Zerlegung) Es sei $S \in \mathbb{R}^{n \times n}$ symmetrisch.

(i) Setze

$$r_{11} = \sqrt{s_{11}}.$$

(ii) Für $i = 2, \dots, n$ setze

$$\begin{aligned} r_{ij} &= \frac{1}{r_{jj}} \cdot \left(s_{ij} - \sum_{k=1}^{j-1} r_{ik} r_{jk} \right), \quad j = 1, \dots, i-1, \\ r_{ii} &= \sqrt{s_{ii} - \sum_{k=1}^{i-1} r_{ik}^2}. \end{aligned}$$

Falls die letzte Wurzel nicht positiv ist, bricht man mit einem Fehler ab. In diesem Fall ist S nicht positiv definit.

Nach n Schritten ist $R^\top = (r_{ij})$ die Matrix aus der Cholesky-Zerlegung von S .

Beweis: Die Formeln ergeben sich aus den Bedingungen $R_{i-1}^\top y = x$ und $\beta = \sqrt{\alpha - y^\top y}$, wobei $\alpha = s_{ii}$ im Beweis von (iii) \Rightarrow (ii), Satz 5.30. ■

Algorithmus 5.32 (Cholesky-Verfahren) Es sei $S \in \mathbb{R}^{n \times n}$ positiv definit und $b \in \mathbb{R}^n$. Gesucht ist eine Lösung x des linearen Gleichungssystems

$$Sx = b.$$

Eine solche Lösung existiert, da S invertierbar ist. Sie kann wie folgt bestimmt werden:

- (i) Bestimme die Cholesky-Zerlegung $S = R^\top \cdot R$.
- (ii) Löse durch Vorwärtseinsetzen $R^\top c = b$ nach c .
- (iii) Löse durch Rückwärtseinsetzen $Rx = c$ nach x .

Dann gilt $Sx = R^\top(Rx) = R^\top c = b$.

Das Cholesky-Verfahren ist schneller als die üblichen numerischen Verfahren zur Lösung linearer Gleichungssysteme, da es die spezielle Struktur der Matrix S ausnutzt. Zwar liegt der Aufwand in beiden Fällen in $O(n^3)$, aber es lässt sich zeigen, dass gegenüber dem Gauß-Algorithmus eine Beschleunigung um den Faktor 2 erzieht wird, siehe dazu Abschnitt 1.3 in Schwarz [13].

Positiv definite Matrizen treten auf natürliche Weise in vielen Problemen auf.

Beispiel 5.33 (Lineare Ausgleichsprobleme) In der Datenanalyse müssen oft unbekannte Parameter durch Messreihen bestimmt werden, die die Parameter in Form linearer Gleichungen festlegen. Bedingt durch Messfehler oder vereinfachte Modellannahmen erhält man in der Regel ein *überbestimmtes* lineares Gleichungssystem, d.h. es liegen mehr Gleichungen vor als Variablen. Das System hat also die Form

$$Ax = b$$

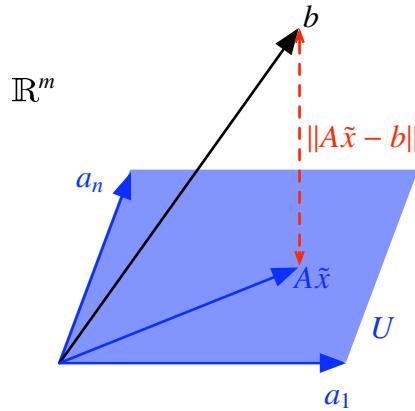
mit $A \in \mathbb{R}^{m \times n}$, wobei $m > n$, $b \in \mathbb{R}^m$ und der Vektor $x \in \mathbb{R}^n$ enthält die zu bestimmenden Parameter. Im Allgemeinen ist ein solches System unlösbar. In diesem Fall möchte man eine Näherungslösung $\tilde{x} \in \mathbb{R}^n$ finden, für die der Fehler

$$\|A\tilde{x} - b\|$$

minimal unter alle $x \in \mathbb{R}^n$ wird. Die geometrische Struktur dieses Problems wird uns auf eine solche optimale Näherungslösung \tilde{x} führen: Wir nehmen an, A habe maximalen Rang n (d.h. überflüssige Gleichungen werden ignoriert). Es seien $a_1, \dots, a_n \in \mathbb{R}^m$ die Spalten von A . Diese erzeugen einen Untervektorraum $U = [a_1, \dots, a_n] \subseteq \mathbb{R}^m$. Falls $b \in U$, so existiert eine Linearkombination $b = \xi_1 a_1 + \dots + \xi_n a_n$. Das System $Ax = b$ ist in diesem Falle exakt lösbar durch

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{R}^n.$$

Falls $b \notin U$, so können wir ein solches x nicht finden. Gesucht ist also das $\tilde{x} \in \mathbb{R}^n$, für dass der Abstand (also der Fehler) $d(A\tilde{x}, b) = \|A\tilde{x} - b\|$ minimal wird.



Dieses $A\tilde{x}$ ist somit die orthogonale Projektion $\Pi_U(b)$ von b auf U . Da $b - A\tilde{x} \perp a_1, \dots, a_n$ ist, muss gelten

$$\langle a_1 | A\tilde{x} \rangle = \langle a_1 | b \rangle, \dots, \langle a_n | A\tilde{x} \rangle = \langle a_n | b \rangle.$$

Das Standardskalarprodukt $\langle a_i | \tilde{x} \rangle$ ist durch $a_i^\top \cdot \tilde{x}$ gegeben. Schreiben wir diese Gleichungen in einen zeilenweise auf, so erhalten wir ein neues lineares Gleichungssystem für \tilde{x} :

$$A^\top A \tilde{x} = A^\top b.$$

Nun ist $S = A^\top A$ symmetrisch, und da A maximalen Rang hat, ist S auch positiv definit. Also ist S invertierbar und es existiert die Näherungslösung

$$\tilde{x} = S^{-1} A^\top b.$$

In der Praxis kann man dieses lineare Gleichungssystem mit dem Verfahren von Cholesky lösen. Dabei können jedoch Probleme mit Rundungsfehlern auftauchen, weswegen man auch andere Verfahren in Betracht ziehen muss (Abschnitte 7.1 und 7.2 in Schwarz [13]).

Beispiel 5.34 (Nichtlineare Optimierung) Es sind die Minimalstellen einer zweifach stetig partiell differenzierbaren Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ gesucht. An einem Punkt $p \in \mathbb{R}^n$ ist die **Hesse-Matrix** von f definiert als

$$H_f(p) = \begin{pmatrix} \partial_{x_1} \partial_{x_1} f(p) & \cdots & \partial_{x_1} \partial_{x_n} f(p) \\ \vdots & \ddots & \vdots \\ \partial_{x_n} \partial_{x_1} f(p) & \cdots & \partial_{x_n} \partial_{x_n} f(p) \end{pmatrix}.$$

Nach dem Satz von Schwarz gilt $\partial_{x_i} \partial_{x_j} f = \partial_{x_j} \partial_{x_i} f$, folglich ist $H_f(p)$ für alle p symmetrisch. Eine lokale Extremstelle liegt vor, falls $\partial_{x_i} f(p) = 0$ für $i = 1, \dots, n$ gilt. Um herauszufinden, ob $f(p)$ ein lokales Maximum oder Minimum von f ist, kann man die Hesse-Matrix auf Definitheit prüfen:

- Ist $H_f(p)$ positiv definit, so handelt es bei $f(p)$ um ein lokales Minimum.
- Ist $H_f(p)$ negativ definit, d.h. ist $-H_f(p)$ positiv definit, so handelt es bei $f(p)$ um ein lokales Maximum.
- Ist $\det(H_f(p)) \neq 0$, aber $H_f(p)$ weder positiv noch negativ definit, so handelt es sich bei $f(p)$ nicht um ein lokales Extremum.
- Ist $\det(H_f(p)) = 0$, so lässt sich keine Aussage treffen.

5.6 Eine Anomalie im \mathbb{R}^3 : Das Vektorprodukt

Der \mathbb{R}^3 sei mit dem Standardskalarprodukt versehen. Es seien zwei linear unabhängige Vektoren

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$$

gegeben, und U der von x, y im \mathbb{R}^3 aufgespannte Untervektorraum. Da $\dim U = 2$, ist das orthogonale Komplement U^\perp eindimensional. Wir wollen x, y in eindeutiger Weise einen zu U orthogonalen Vektor $x \times y$ zuordnen. Für ein solches $x \times y$ muss gelten

$$\langle x \times y | z \rangle = 0 \quad \Leftrightarrow \quad z \in [x, y]. \quad (5.22)$$

Beachte, dass für linear unabhängige x, y ebenfalls gilt

$$\det(x \ y \ z) = 0 \quad \Leftrightarrow \quad z \in [x, y].$$

Dies legt nahe, den Vektor $x \times y$ über die Determinante $\det(x \ y \ z)$ zu definieren. Entwickle diese Determinante nach der dritten Spalte:

$$\begin{aligned} \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} &= z_1 \det \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} - z_2 \det \begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix} + z_3 \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \\ &= z_1 \det \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} + z_2 \det \begin{pmatrix} x_3 & y_3 \\ x_1 & y_1 \end{pmatrix} + z_3 \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \\ &= \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \left| \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \right|. \end{aligned}$$

Der Vektor

$$x \times y = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \quad (5.23)$$

erfüllt somit die Bedingung (5.22). Wir nennen $x \times y$ das **Vektorprodukt** (oder **Kreuzprodukt**) von x und y .

Hilfssatz 5.35 Es seien $x, y, z \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$. Dann gilt:

- (a) $x \times y = -y \times x$, insbesondere $x \times x = 0$.
- (b) $(x + y) \times z = x \times z + y \times z$.
- (c) $(\lambda x) \times y = \lambda(x \times y) = x \times (\lambda y)$.
- (d) $x \times y = 0$ genau dann, wenn x, y linear unabhängig sind.
- (e) $\langle x \times y | z \rangle = \langle x | y \times z \rangle$.
- (f) Es gilt

$$x \times (y \times z) + y \times (z \times x) + z \times (x \times y) = 0. \quad (5.24)$$

BEWEIS: Ergibt sich sofort aus den Rechenregeln für Determinanten und der Definition des Vektorprodukts. ■

Aufgabe 5.36 Das Vektorprodukt ist nicht assoziativ, d.h. im Allgemeinen gilt nicht $x \times (y \times z) = (x \times y) \times z$.

Aufgabe 5.37 Für die Standardbasis des \mathbb{R}^3 gelten folgende Relationen:

$$e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2.$$

Bemerkung 5.38 Die Gleichung (5.24) heißt **Jacobi-Identität**. Ein Vektorraum V mit einem alternierenden Produkt, das die Jacobi-Identität erfüllt, wird **Lie-Algebra** genannt. Somit ist (\mathbb{R}^3, \times) eine Lie-Algebra.

Bemerkung 5.39 Das Vektorprodukt liefert eine Möglichkeit, im \mathbb{R}^3 orthogonale Komplemente zu bestimmen: Ist $x \in \mathbb{R}^3$ gegeben, so wählt man einen beliebigen Vektor v , der zu x linear unabhängig ist. Dann

$$y_1 = x \times v$$

im orthogonalen Komplement von x , und

$$y_2 = x \times y_1$$

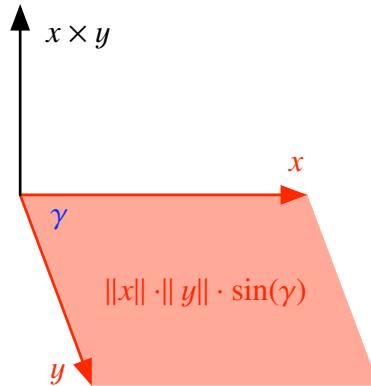
ist im orthogonalen Komplement von x und y_1 . Also ist $\{y_1, y_2\}$ eine Basis des orthogonalen Komplements von x .

Aufgabe 5.40 Für $x, y \in \mathbb{R}^3$ gelten

$$\|x \times y\|^2 = \|x\|^2 \cdot \|y\|^2 - \langle x | y \rangle, \quad (5.25)$$

$$\|x \times y\| = \|x\| \cdot \|y\| \cdot \sin(\gamma(x, y)). \quad (5.26)$$

Die zweite Gleichung gibt den Flächeninhalt des Parallelogramms an, das von x und y aufgespannt wird.



In höheren Dimensionen $n > 3$ gibt es für die lineare Hülle zweier Vektoren x, y keine eindeutige orthogonale Richtung, daher lässt sich das Vektorprodukt nicht direkt verallgemeinern. Für $n - 1$ linear unabhängige Vektoren im \mathbb{R}^n kann man aber eine analoge Definition angeben:

$$x_1 \times \cdots \times x_{n-1} = \begin{pmatrix} (-1)^{1+n} \det(X_{1n}) \\ \vdots \\ (-1)^{i+n} \det(X_{in}) \\ \vdots \\ (-1)^{n+n} \det(X_{nn}) \end{pmatrix},$$

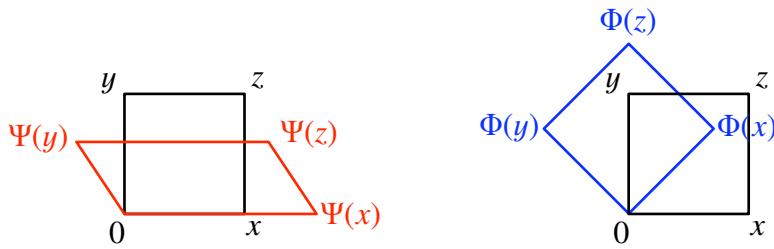
wobei $X = (x_1 \ x_2 \ \cdots \ x_{n-1} \ y) \in \mathbb{R}^{n \times n}$ und X_{in} die Matrix ist, die aus X entsteht durch Streichen der i ten Zeile und n ten Spalte. Dann gilt

$$x_1 \times \cdots \times x_{n-1} \perp [x_1, \dots, x_{n-1}].$$

Hierbei handelt es sich aber nicht mehr um ein Produkt im üblichen Sinne, da diese Formel nur mit $n - 1 > 2$ Vektoren sinnvoll definiert ist.

6 Isometriegruppen

In diesem Kapitel studieren wir eine wichtige Klasse von linearen Abbildungen, die wir anschaulich als die „starren Bewegungen“ im Raum auffassen. Die folgende Zeichnung soll dies illustrieren.



Hier sind $\Phi, \Psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ jeweils lineare Abbildungen. Wir betrachten ihre Wirkung auf einem Quadrat, dessen Ecken durch die Vektoren $0, x, y, z$ gegeben seien. Offensichtlich wird man die Abbildung Ψ nicht als „starr“ auffassen wollen, da das Quadrat hier verzerrt wird (genauer gesagt handelt es sich bei Ψ um die Kombination einer Scherung und einer Skalierung). Die Abbildung Φ hingegen verändert zwar die Lage des Quadrates, aber nicht seine geometrische Gestalt (es handelt sich um eine Drehung um den Ursprung), sie ist in diesem Sinne „starr“.

6.1 Isometrien

Definition 6.1 Es seien $(V_1, \langle \cdot | \cdot \rangle_1)$ und $(V_2, \langle \cdot | \cdot \rangle_2)$ Vektorräume mit Skalarprodukt. Ihre Abstandsfunktionen bezeichnen wir mit d_1 bzw. d_2 . Eine **Isometrie** ist eine bijektive Abbildung $\psi : V_1 \rightarrow V_2$, die für alle $x, y \in V$ die Bedingung

$$d_1(x, y) = d_2(\psi(x), \psi(y)) \quad (6.1)$$

erfüllt.

Beispiel 6.2 Eine Isometrie ist nicht zwingendermaßen eine lineare Abbildung. Beispielsweise erfüllt jede **Translation** $\tau_v : V \rightarrow V$ um einen festen Vektor v , gegeben durch

$$\tau_v : V \rightarrow V, \quad x \mapsto x + v,$$

die Bedingung (6.1):

$$d(\tau_v(x), \tau_v(y)) = \|(y + v) - (x + v)\| = \|y - x\| = d(x, y).$$

Offensichtlich ist die Translation um einen Vektor $v \neq 0$ aber nicht linear.

Lineare Isometrien lassen sich durch das Skalarprodukt charakterisieren:

Satz 6.3 Es seien $(V_1, \langle \cdot | \cdot \rangle_1)$ und $(V_2, \langle \cdot | \cdot \rangle_2)$ Vektorräume mit Skalarprodukt und $\Phi : V_1 \rightarrow V_2$ eine lineare Abbildung. Folgende Aussagen sind äquivalent:

- (i) Φ ist eine Isometrie.
- (ii) $\langle x | y \rangle_1 = \langle \Phi(x) | \Phi(y) \rangle_2$ für alle $x, y \in V_1$.
- (iii) $\|\Phi(x)\|_2 = 1$ für alle $x \in V_1$ mit $\|x\|_1 = 1$.
- (iv) $\|\Phi(x)\|_2 = \|x\|_1$ für alle $x \in V_1$.

BEWEIS: (i) \Rightarrow (iv): Da Φ linear ist, gilt für alle $x \in V_1$:

$$d_2(\Phi(x), \Phi(0)) = \|\Phi(x) - \Phi(0)\|_2 = \|\Phi(x) - 0\|_2 = \|\Phi(x)\|_2.$$

Da Φ eine Isometrie ist, folgt

$$\|x\|_1 = d_1(x, 0) = d_2(\Phi(x), \Phi(0)) = \|\Phi(x)\|_2.$$

(iv) \Rightarrow (iii): Trivial.

(iii) \Rightarrow (iv): Es sei $x \in V_1$. Ist $x = 0$, so gilt $\Phi(0) = 0$ und $\|\Phi(0)\|_2 = 0$. Sei nun $x \neq 0$. Da $\|\frac{x}{\|x\|_1}\|_1 = 1$ ist, können wir folgern:

$$\|\Phi(x)\|_2 = \left\| \frac{\|x\|_1}{\|x\|_1} \cdot \Phi(x) \right\|_2 = \|x\|_1 \cdot \underbrace{\left\| \Phi\left(\frac{x}{\|x\|_1}\right) \right\|_2}_{=1} = \|x\|_1.$$

(iv) \Rightarrow (ii): Seien $x, y \in V_1$. Im euklidischen Fall erlaubt die Polarisierung (4.17) uns folgende Rechnung:

$$\begin{aligned} \langle x | y \rangle_1 &= \frac{1}{2} (\|x + y\|_1^2 - \|x\|_1^2 - \|y\|_1^2) \\ &= \frac{1}{2} (\|\Phi(x + y)\|_2^2 - \|\Phi(x)\|_2^2 - \|\Phi(y)\|_2^2) \\ &= \frac{1}{2} (\|\Phi(x) + \Phi(y)\|_2^2 - \|\Phi(x)\|_2^2 - \|\Phi(y)\|_2^2) \\ &= \langle \Phi(x) | \Phi(y) \rangle_2. \end{aligned}$$

Im unitären Fall gelten ähnliche Gleichungen jeweils für $\operatorname{Re}\langle x | y \rangle$ und $\operatorname{Im}\langle x | y \rangle$, aus denen (i) folgt.

(ii) \Rightarrow (i): Für $x, y \in V_1$ gilt

$$d_1(x, y)^2 = \langle y - x | y - x \rangle_1 = \langle \Phi(y - x) | \Phi(y - x) \rangle_2 = d_2(\Phi(x), \Phi(y))^2,$$

wobei für die letzte Gleichung die Linearität von Φ verwendet wurde. ■

Aufgabe 6.4 $\Phi : V_1 \rightarrow V_2$ ist genau dann eine lineare Isometrie, wenn für jede Orthonormalbasis $B \subset V_1$ auch $\Phi(B) \subset V_2$ eine Orthonormalbasis ist.

6.2 Orthogonale und unitäre Gruppen

Im Rahmen der linearen Algebra interessieren wir uns überwiegend für lineare Isometrien, die V auf sich selbst abbilden.

Definition 6.5 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer oder unitärer Vektorraum. Die Menge der Isometrien von V schreiben wir

$$\text{Iso}(V, \langle \cdot | \cdot \rangle) = \{\psi : V \rightarrow V \mid \psi \text{ ist Isometrie}\}.$$

Als **lineare Isometrien** oder **orthogonale Transformationen** bezeichnen wir die Elemente von

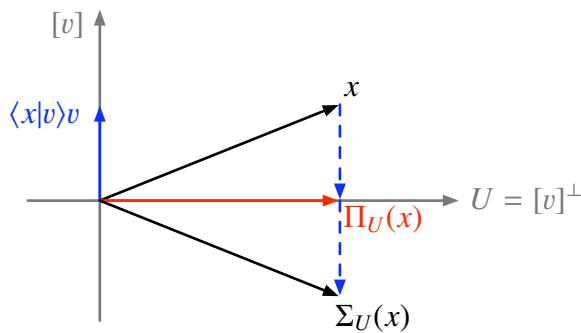
$$\mathbf{O}(V, \langle \cdot | \cdot \rangle) = \mathbf{GL}(V) \cap \text{Iso}(V, \langle \cdot | \cdot \rangle).$$

Im Falle unitärer Vektorräume ist es auch üblich, anstelle von orthogonalen von **unitären Transformationen** zu sprechen und ihre Menge mit $\mathbf{U}(V, \langle \cdot | \cdot \rangle)$ zu bezeichnen.

Beispiel 6.6 (Spiegelungen) Es sei V ein Vektorraum mit Skalarprodukt, $v \in V$, $\|v\| = 1$ und $U = [v]^\perp$. Die **Spiegelung an U** ist definiert als

$$\Sigma_U : V \rightarrow V, \quad x \mapsto x - 2\langle x | v \rangle v. \tag{6.2}$$

Ja nach Bedarf sprechen wir auch von der *Spiegelung entlang der Achse $[v]$* .



Die Spiegelung Σ_U ist eine lineare Isometrie:

$$\begin{aligned}\langle \Sigma_U(x) | \Sigma_U(y) \rangle &= \langle x - 2\langle x|v\rangle v | y - 2\langle y|v\rangle v \rangle \\ &= \langle x|y \rangle - 2\langle x|v\rangle \langle v|y \rangle - 2\langle y|v\rangle \langle x|v \rangle + 4\langle x|v\rangle \langle y|v \rangle \underbrace{\langle v|v \rangle}_{=1} \\ &= \langle x|y \rangle - 4\langle x|v\rangle \langle y|v \rangle + 4\langle x|v\rangle \langle y|v \rangle \\ &= \langle x|y \rangle.\end{aligned}$$

Für einen beliebigen Untervektorraum $W \subset V$ können wir die Spiegelung an W definieren über

$$\Sigma_W(x) = x - 2\langle x|v_1\rangle v_1 - \dots - 2\langle x|v_k\rangle v_k,$$

wobei die Vektoren v_1, \dots, v_k eine Orthonormalbasis von W^\perp bilden.

Aufgabe 6.7 Seien W und v_1, \dots, v_k wie in Beispiel 6.6. Setzen wir $U_i = [v_i]^\perp$, so ist $W = U_1 \cap \dots \cap U_k$ und Σ_W ist die Verknüpfung der Spiegelungen Σ_{U_i} :

$$\Sigma_W = \Sigma_{U_1} \circ \dots \circ \Sigma_{U_k}. \quad (6.3)$$

Hierfür ist es notwendig, dass die v_i paarweise orthogonal zu einander sind.

Aufgabe 6.8 Ist $V = \mathbb{R}^n$ und $U_1 = [e_1]^\perp$, so wird die Spiegelung Σ_{U_1} in der Standardbasis durch die Matrix

$$\begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} \quad (6.4)$$

dargestellt. Für beliebiges $v \in \mathbb{R}^n$ mit $\|v\| = 1$ und $U = [v]^\perp$ wird Σ_U durch die Matrix

$$I_n - 2v \cdot v^\top \quad (6.5)$$

dargestellt.

Aufgabe 6.9 Für jede Spiegelung Σ gilt $\Sigma^2 = \text{id}_V$.

Beispiel 6.10 (Rotationen im \mathbb{R}^2) Im \mathbb{R}^2 mit dem Standardskalarprodukt ist eine **Rotation** (oder **Drehung**) mit Drehwinkel α um den Ursprung durch die Matrix

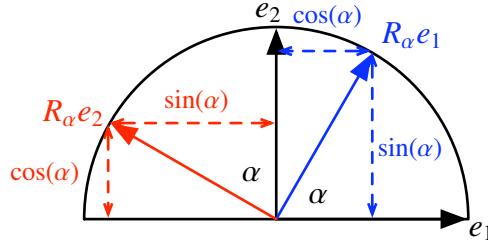
$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (6.6)$$

gegeben. Diese Bezeichnung wird dadurch gerechtfertigt, dass für alle $x \in \mathbb{R}^n$ gilt $\|x\| = \|Ax\|$ und

$$\begin{aligned}\cos(\varphi(x, Ax)) &= \frac{\langle x | Ax \rangle}{\|x\| \cdot \|Ax\|} = \frac{\cos(\alpha)x_1^2 - \sin(\alpha)x_2x_1 + \sin(\alpha)x_1x_2 + \cos(\alpha)x_2^2}{\|x\|^2} \\ &= \cos(\alpha) \cdot \frac{\|x\|^2}{\|x\|^2} = \cos(\alpha).\end{aligned}$$

Besonders klar wird die geometrische Interpretation, wenn man die Bilder der Standardbasis betrachtet:

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}, \quad \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}.$$



Beispiel 6.11 (Rotationen im \mathbb{R}^3) Im \mathbb{R}^3 mit dem Standardskalarprodukt werden die Drehungen um die Achsen $[e_1]$, $[e_2]$ und $[e_3]$ jeweils durch die Matrizen

$$\begin{aligned}R_{1,\alpha} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}, & R_{2,\beta} &= \begin{pmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{pmatrix}, \\ R_{3,\gamma} &= \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

beschrieben. Wie man der Matrixform direkt ansieht, ist die jeweilige **Drehachse** $[e_i]$ ein Eigenraum zum Eigenwert 1 und auf ihrem orthogonalen Komplement, der **Drehebene** $[e_i]^\perp = [e_j, e_k]$, wirkt die Rotation wie eine Rotation (6.6) im \mathbb{R}^2 .

Beispiel 6.12 Ist V ein reeller Hilbert-Raum mit abzählbarem vollständigem Orthonormalsystem B , so ist die Abbildung

$$V \rightarrow \ell^2(\mathbb{R}), \quad x = \sum_{i=1}^{\infty} x_i b_i \mapsto (x_1, x_2, x_3, \dots)$$

eine lineare Isometrie. Dies folgt aus der Parsevalschen Gleichung (5.18).

Hilfssatz 6.13 Es sei V ein Vektorraum mit Skalarprodukt, $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$ und λ ein Eigenwert von Φ . Dann gilt $|\lambda| = 1$. Insbesondere sind im euklidischen Fall 1 und -1 die einzigen möglichen Eigenwerte von Φ .

BEWEIS: Es sei x ein Eigenvektor zu λ . Dann gilt

$$\langle x | x \rangle = \langle \Phi(x) | \Phi(x) \rangle = \langle \lambda x | \lambda x \rangle = |\lambda|^2 \langle x | x \rangle,$$

d.h. $|\lambda|^2 = 1$. ■

Satz 6.14 Die Menge $\mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$ ist eine Gruppe mit der Komposition von Abbildungen als Verknüpfung. $\mathbf{O}(V, \langle \cdot | \cdot \rangle)$ ist eine Untergruppe von $\mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$.

BEWEIS: Offensichtlich ist $\text{id}_V \in \mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$.

Nun seien $\varphi, \psi \in \mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$. Dann ist $\varphi \circ \psi \in \mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$, denn für $x, y \in V$ gilt

$$d(\varphi(\psi(x)), \varphi(\psi(y))) = d(\psi(x), \psi(y)) = d(x, y).$$

Das Inverse ψ^{-1} von ψ existiert, da ψ bijektiv ist. Da ψ eine Isometrie ist, gilt

$$d(x, y) = d(\psi(\psi^{-1}(x)), \psi(\psi^{-1}(y))) = d(\psi^{-1}(x), \psi^{-1}(y)).$$

Somit ist auch $\psi^{-1} \in \mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$.

Die orthogonale Gruppe $\mathbf{O}(V, \langle \cdot | \cdot \rangle)$ ist eine Untergruppe, da sie der Durchschnitt von zwei Gruppen ist. ■

Bemerkung 6.15 Es lässt sich zeigen, dass jede Isometrie $\psi \in \mathbf{Iso}(V, \langle \cdot | \cdot \rangle)$ die Verknüpfung einer Translation τ_v mit einer linearen Isometrie Φ ist, d.h. ψ ist von der Gestalt

$$\psi(x) = \Phi(x) + v$$

für einen festen Vektor $v \in V$ und ein $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$.

Betrachten wir speziell den Fall $V = \mathbb{R}^n$, in dem $\langle \cdot | \cdot \rangle$ durch eine symmetrische Matrix S gegeben ist. Ist Φ eine lineare Isometrie mit Abbildungsmatrix A , so gilt

$$\langle x | y \rangle = x^\top S y = \langle Ax | Ay \rangle = (Ax)^\top S A y = x^\top A^\top S A y$$

für alle $x, y \in \mathbb{R}^n$. Daraus folgt

$$A^\top S A = S.$$

Umgekehrt definiert jede Matrix, die dies erfüllt, eine Isometrie. Aus Abschnitt 5.2 wissen wir, dass wir nach Basiswechsel ohne Einschränkung annehmen können, $\langle \cdot | \cdot \rangle$ sei das Standardskalarprodukt. Dann gilt $S = I_n$, also lautet die Isometriebedingung

$$A^\top \cdot A = I_n.$$

Wir können die zugehörige orthogonale Gruppen somit durch eine Matrixgruppe darstellen.

Definition 6.16 Die **orthogonale Gruppe** ist

$$\mathbf{O}_n = \{A \in \mathbf{GL}_n(\mathbb{R}) \mid A^\top = A^{-1}\}. \quad (6.7)$$

Die **spezielle orthogonale Gruppe** ist

$$\mathbf{SO}_n = \{A \in \mathbf{O}_n \mid \det(A) = 1\} = \mathbf{O}_n \cap \mathbf{SL}_n(\mathbb{R}). \quad (6.8)$$

Ähnlich findet man für \mathbb{C}^n mit dem unitären Standardskalarprodukt heraus, dass die Abbildungsmatrizen A der unitären Transformationen die Bedingung

$$x^\top \bar{y} = x^\top A^\top \bar{A} \bar{y}$$

für alle $x, y \in \mathbb{C}^n$ erfüllen müssen. Das ist äquivalent zu

$$\bar{A}^\top \cdot A = I_n.$$

Definition 6.17 Die **unitäre Gruppe** ist

$$\mathbf{U}_n = \{A \in \mathbf{GL}_n(\mathbb{C}) \mid \bar{A}^\top = A^{-1}\}. \quad (6.9)$$

Die **spezielle unitäre Gruppe** ist

$$\mathbf{SU}_n = \{A \in \mathbf{U}_n \mid \det(A) = 1\} = \mathbf{U}_n \cap \mathbf{SL}_n(\mathbb{C}). \quad (6.10)$$

Die Elemente der orthogonalen bzw. unitären Gruppe werden **orthogonale** bzw. **unitäre** Matrizen genannt.

Bemerkung 6.18 Es sei A eine orthogonale Matrix. Fassen wir A als komplexe Matrix auf, so ist gilt $A = \bar{A}$ und folglich $A^{-1} = A^\top = \bar{A}^\top$. Somit ist A auch unitär.

Hilfssatz 6.19 Für $A \in \mathbf{O}_n$ gilt $\det(A) = \pm 1$. Für $A \in \mathbf{U}_n$ gilt $|\det(A)| = 1$.

BEWEIS: Im euklidischen Fall gilt $1 = \det(I_n) = \det(A^\top \cdot A) = \det(A^\top) \det(A) = \det(A)^2$, also $\det(A) = \pm 1$. Im unitären Fall gilt $1 = \det(I_n) = \det(\bar{A}^\top \cdot A) = \det(\bar{A}) \det(A) = |\det(A)|^2$, also $|\det(A)| = 1$. ■

Die Bedeutung der orthogonalen und unitären Gruppen liegt darin, dass die lineare Isometriegruppe jedes Vektorraumes endlicher Dimension mit Skalarprodukt durch Matrizen aus \mathbf{O}_n bzw. \mathbf{U}_n dargestellt werden kann.

Definition 6.20 Es sei G eine Gruppe. Eine reelle bzw. komplexe **Darstellung** von G ist ein Gruppenhomomorphismus

$$\varrho : G \rightarrow \mathbf{GL}_n(\mathbb{R}) \quad \text{bzw.} \quad \varrho : G \rightarrow \mathbf{GL}_n(\mathbb{C}). \quad (6.11)$$

Satz 6.21 Es sein V ein Vektorraum mit Skalarprodukt und $\dim V = n$.

- (a) Ist V euklidisch, so gilt $\mathbf{O}(V, \langle \cdot | \cdot \rangle) \cong \mathbf{O}_n$.
- (b) Ist V unitär, so gilt $\mathbf{U}(V, \langle \cdot | \cdot \rangle) \cong \mathbf{U}_n$.

BEWEIS: Es sei V euklidisch und B eine Orthonormalbasis von V . Das Skalarprodukt $\langle \cdot | \cdot \rangle$ wird bzgl. B durch die Einheitsmatrix I_n dargestellt. Durch B ist eine isomorphe Darstellung

$$\varrho : \mathbf{O}(V, \langle \cdot | \cdot \rangle) \rightarrow \mathbf{O}_n, \quad \Phi \mapsto M_B^B(\Phi)$$

definiert: Da

$$M_B^B(\Phi)^\top \cdot I_n \cdot M_B^B(\Phi) = I_n$$

gilt, ist $M_B^B(\Phi) \in \mathbf{O}_n$. Umgekehrt definiert jede orthogonale Matrix A eine Isometrie $(M_B^B)^{-1}(A)$. Es gilt also Bild $\varrho = \mathbf{O}_n$, und ϱ ist ein Isomorphismus, da M_B^B bijektiv ist.

Entsprechend zeigt man die Aussage im unitären Fall. ■

Folgerung 6.22 Für $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$ gilt $\det(\Phi) = \pm 1$. Für $\Phi \in \mathbf{U}(V, \langle \cdot | \cdot \rangle)$ gilt $|\det(\Phi)| = 1$.

Beispiel 6.23 Die Gruppe \mathbf{O}_2 enthält die Spiegelungen und Drehungen der Form

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad S_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

wie wir aus den Beispielen 6.6 und 6.10 bereits wissen. Es ist

$$\det(S_i) = -1 \quad \text{und} \quad \det(R_\alpha) = \cos(\alpha)^2 + \sin(\alpha)^2 = 1.$$

Das bedeutet $R_\alpha \in \mathbf{SO}_2$. Wir bestimmen nun die allgemeine Form der Elemente von \mathbf{O}_2 und \mathbf{SO}_2 : Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{O}_2.$$

Die Bedingung $A^\top \cdot A = I_2$ lautet ausgeschrieben

$$A^\top \cdot A = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Koeffizientenvergleich liefert die Gleichungen

$$\begin{aligned} a^2 + c^2 &= 1, \\ b^2 + d^2 &= 1. \end{aligned}$$

Diese Gleichungen besagen, dass es Winkel $\alpha, \beta \in [0, 2\pi)$ gibt, so dass gilt

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

Insbesondere ist wegen Hilfssatz 6.19

$$\begin{aligned} \pm 1 &= \det(A) = \cos(\alpha) \sin(\beta) - \cos(\beta) \sin(\alpha) \\ &= \sin(\beta - \alpha), \end{aligned}$$

wie man mit Hilfe der Additionstheoreme (Anhang B.2) sieht.

Unterscheiden wir nun die Fälle $\det(A) = 1$ und $\det(A) = -1$:

- Im Fall $\det(A) = 1$ gilt $\sin(\beta - \alpha) = 1$, also $\beta = \alpha + \frac{\pi}{2}$ und somit

$$\sin(\beta) = \cos(\alpha), \quad \cos(\beta) = -\sin(\alpha).$$

Die Matrix $A \in \mathbf{SO}_2$ hat also die Gestalt einer Rotationsmatrix:

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

- Im Fall $\det(A) = -1$ gilt $\sin(\beta - \alpha) = -1$, also $\beta = \alpha + \frac{3\pi}{2}$ und somit

$$\sin(\beta) = -\cos(\alpha), \quad \cos(\beta) = \sin(\alpha).$$

Die Matrix A ist von der Form

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} = R_\alpha \cdot S_1.$$

Jedes Element von \mathbf{O}_2 ist also entweder eine Rotation oder die Verkettung einer Rotation und einer Spiegelung (in Satz 6.40 werden wir sehen, dass sich jede Rotation wiederum als Produkt von Spiegelungen schreiben lässt). Des Weiteren prüft man mit Hilfe der Additionstheoreme leicht nach, dass für zwei Rotationen $R_\alpha, R_\beta \in \mathbf{SO}_2$ gilt:

$$R_\alpha R_\beta = R_{\alpha+\beta} = R_\beta R_\alpha.$$

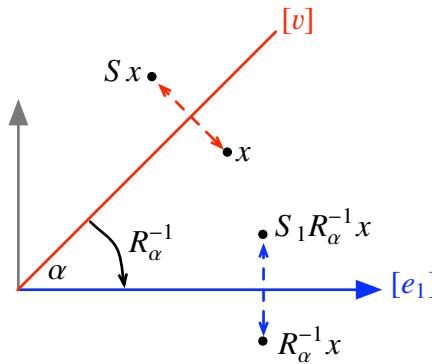
Die Gruppe \mathbf{SO}_2 ist also abelsch. Im Allgemeinen gilt aber

$$R_\alpha S_1 \neq S_1 R_\alpha.$$

Also ist \mathbf{O}_2 nicht abelsch. Setzen wir

$$S = R_\alpha S_1 R_\alpha^{-1} = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix},$$

wobei man die zweite Gleichheit erneut mit den Additionstheoremen nachrechnet, so können wir S als die Spiegelung an der Geraden auffassen, die den Winkel α mit der e_1 -Achse einschließt. Man kann sich dies so vorstellen, dass R_α^{-1} die Spiegelungsachse $[v]$ auf die e_1 -Achse rotiert, dann durch S_1 an der e_1 -Achse gespiegelt wird und schließlich die gespiegelten Punkte durch R_α in die Lage rotiert werden, in der sie gespiegelt an der $[v]$ -Achse liegen.



Das charakteristische Polynom von S ist

$$f_S = \det(XI_2 - S) = X^2 - \cos(2\alpha)^2 - \sin(2\alpha)^2 = X^2 - 1.$$

Es gibt also stets einen Eigenraum E_1 zum Eigenwert 1 (die Spiegelungsachse) und einen Eigenraum E_{-1} zum Eigenwert -1 senkrecht dazu. Man rechnet direkt nach, dass gilt

$$E_1 = \left[\begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \right], \quad E_{-1} = \left[\begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \right].$$

Die Elemente aus $\mathbf{O}_2 \setminus \mathbf{SO}_2$ sind also gerade die Spiegelungen im \mathbb{R}^2 .

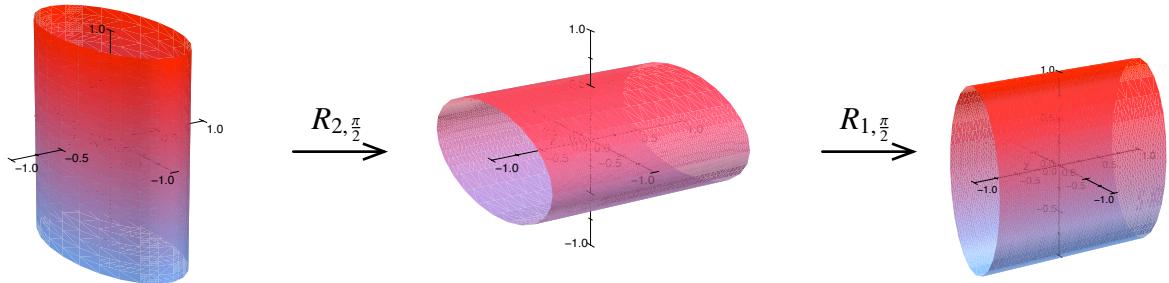
Beispiel 6.24 Es seien $R_{1,\alpha}, R_{2,\beta}, R_{3,\gamma}$ die Rotationsmatrizen aus Beispiel 6.11. Dann gilt für $i = 1, 2, 3$

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

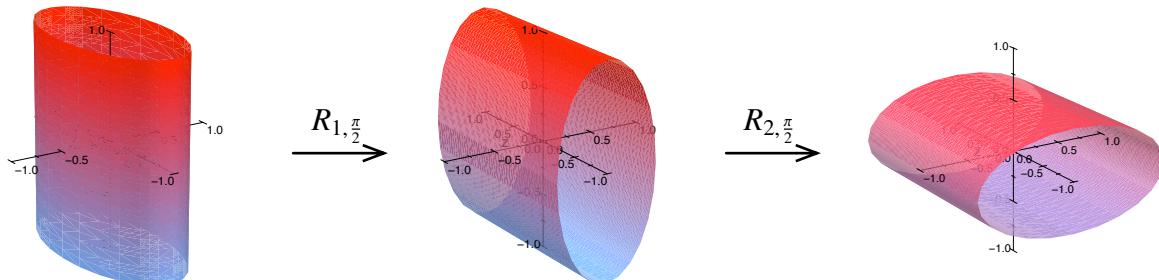
also $R_{i,\alpha} \in \mathbf{SO}_3$. Sei nun speziell $\alpha = \beta = \gamma = \frac{\pi}{2}$ (Drehungen um $\frac{\pi}{2}$ in ihrer jeweiligen Drehebene). Es ist

$$R_{1,\frac{\pi}{2}} R_{2,\frac{\pi}{2}} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} = R_{2,\frac{\pi}{2}} R_{1,\frac{\pi}{2}}.$$

Die Gruppe \mathbf{SO}_3 ist also nicht abelsch. Dieses Phänomen lässt sich geometrisch veranschaulichen: Die Fläche in der folgenden Abbildung wird zunächst mit $R_{2,\frac{\pi}{2}}$ um die e_2 -Achse gedreht, anschließend mit $R_{1,\frac{\pi}{2}}$ um die e_1 -Achse. Dies entspricht der verketteten Transformation durch $R_{1,\frac{\pi}{2}} R_{2,\frac{\pi}{2}}$.



Führen wir die Drehungen in umgekehrter Reihenfolge aus, so wird die Fläche zunächst durch $R_{1,\frac{\pi}{2}}$ um die e_1 -Achse gedreht und anschließend durch $R_{2,\frac{\pi}{2}}$ um die e_2 -Achse. Dies entspricht der verketteten Transformation durch $R_{2,\frac{\pi}{2}} R_{1,\frac{\pi}{2}}$.



Offenbar erhalten wir zwei verschiedene Transformationen. In Dimensionen ≥ 3 ist es daher wichtig, bei der Anwendung geometrischer Transformationen auf die Reihenfolge zu achten.

Folgerung 6.25 \mathbf{SO}_n ist abelsch für $n = 2$ und nicht abelsch für $n > 2$.

Beispiel 6.26 \mathbf{U}_1 ist die Gruppe

$$\mathbf{U}_1 = \{z \in \mathbb{C} \mid z\bar{z} = |z|^2 = 1\}$$

der komplexen Zahlen vom Betrag 1. Diese Zahlen haben die Form $z = e^{i\alpha}$ (vgl. (A.2)). Die Abbildung

$$\Phi : \mathbf{U}_1 \rightarrow \mathbf{SO}_2, \quad e^{i\alpha} \mapsto R_\alpha$$

ist ein Isomorphismus von Gruppen.

Beispiel 6.27 Die spezielle unitäre Gruppe \mathbf{SU}_2 ist

$$\mathbf{SU}_2 = \left\{ \begin{pmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{pmatrix} \mid z_1, z_2 \in \mathbb{C}, |z_1|^2 + |z_2|^2 = 1 \right\}.$$

Diese Gestalt leitet man leicht aus den Bedingungen $A^\top \bar{A} = I_2$ und $\det(A) = 1$ her. Eine Matrix $A \in \mathbf{U}_2$ hat dann die Gestalt

$$A = e^{i\alpha} A_0,$$

wobei $A_0 \in \mathbf{SU}_2$ und $\alpha \in [0, 2\pi)$.

Hilfssatz 6.28 Für eine Matrix $A \in \mathbb{R}^{n \times n}$ (bzw. $A \in \mathbb{C}^{n \times n}$) sind äquivalent:

- (i) A ist orthogonal (bzw. unitär).
- (ii) Die Spalten von A bilden eine Orthonormalbasis von \mathbb{R}^n (bzw. \mathbb{C}^n) mit dem Standardskalarprodukt.
- (iii) Die Zeilen von A bilden eine Orthonormalbasis von \mathbb{R}^n (bzw. \mathbb{C}^n) mit dem Standardskalarprodukt.

Beweis: Die Einträge in $A^\top \cdot \bar{A}$ sind gerade die Skalarprodukte der Spalten von A . Es gilt also $A^\top \cdot \bar{A} = I_n$ genau dann, wenn die Spalten eine Orthonormalbasis von \mathbb{R}^n (bzw. \mathbb{C}^n) bilden. Entsprechendes gilt für die Zeilen. ■

Aufgabe 6.29 \mathbf{SO}_n ist ein Normalteiler in \mathbf{O}_n und \mathbf{SU}_n ist ein Normalteiler in \mathbf{U}_n . Des Weiteren ist $\mathbf{O}_n/\mathbf{SO}_n \cong (\{\pm 1\}, \cdot)$ und $\mathbf{U}_n/\mathbf{SU}_n \cong \mathbf{U}_1$.

Satz 6.30 Die Gruppen \mathbf{O}_n und \mathbf{SO}_n sind kompakte Teilmengen des Vektorraums $\mathbb{R}^{n \times n}$, versehen mit der Norm $\|A\| = \sqrt{\text{tr}(A^\top A)}$.

Beweis: Wir zeigen, dass \mathbf{O}_n und \mathbf{SO}_n beschränkt und abgeschlossen sind.

Da $A \in \mathbf{O}_n$ genau dann, wenn $A^\top A = I_n$ gilt, ist \mathbf{O}_n in der Sphäre mit Radius $\sqrt{n} = \sqrt{\text{tr}(I_n)}$ in $\mathbb{R}^{n \times n}$ enthalten. Somit sind \mathbf{O}_n und \mathbf{SO}_n beschränkt.

Die Abgeschlossenheit von \mathbf{O}_n und \mathbf{SO}_n kann man wie folgt sehen: Die Abbildung $f : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, A \mapsto A^\top A$, ist stetig, da sie durch Addition und Multiplikation der Matrixkoeffizienten gegeben ist. Es gilt per Definition

$$f^{-1}(\{I_n\}) = \mathbf{O}_n.$$

Als Urbild der abgeschlossenen Menge $\{I_n\}$ unter der stetigen Abbildung f ist \mathbf{O}_n selbst abgeschlossen. Die Gruppe \mathbf{SO}_n ist $\mathbf{O}_n \cap \mathbf{SL}_n(\mathbb{R})$. Die Gruppe $\mathbf{SL}_n(\mathbb{R})$ ist wiederum eine abgeschlossene Teilmenge des $\mathbb{R}^{n \times n}$, da $\mathbf{SL}_n(\mathbb{R}) = \det^{-1}(\{1\})$ ist, und die Determinante eine stetige Funktion ist. Als Durchschnitt von abgeschlossenen Mengen ist \mathbf{SO}_n selbst abgeschlossen. ■

6.3 Die Normalform für lineare Isometrien

Die Jordansche Normalform klassifiziert die Konjugationsklassen von Endomorphismen komplexer Vektorräume endlicher Dimension. Im Falle eines unitären Endomorphismus $\Phi \in \mathbf{U}(V, \langle \cdot | \cdot \rangle)$ werden wir sehen, dass die Jordansche Normalform eine besonders einfache Gestalt hat: Φ ist stets diagonalisierbar, und die Konjugationsklasse von Φ ist durch die Eigenwerte und ihre Vielfachheit bereits eindeutig bestimmt. Wir können auch eine Normalform für orthogonale Endomorphismen ableiten, selbst in den Fällen, in denen das charakteristische Polynom nicht in reelle Linearfaktoren zerfällt.

Hilfssatz 6.31 *Es sei V ein Vektorraum mit Skalarprodukt und Φ eine lineare Isometrie. Weiter sei $\lambda \in \mathbb{R}$ (oder $\in \mathbb{C}$) ein Eigenwert von Φ und $x \neq 0$ ein Eigenvektor zum Eigenwert λ . Dann gilt:*

- (a) *Das orthogonale Komplement $[x]^\perp$ ist ein Φ -invarianter Untervektorraum.*
- (b) *Ist $\mu \in \text{Spec } \Phi$ und $\mu \neq \lambda$, so gilt $x \perp E_\mu$.*

Beweis:

- (a) Es sei $y \in [x]^\perp$. Da Φ invertierbar ist, ist $\lambda \neq 0$. Es gilt

$$\langle x | \Phi(y) \rangle = \frac{1}{\lambda} \langle \lambda x | \Phi(y) \rangle = \frac{1}{\lambda} \langle \Phi(x) | \Phi(y) \rangle = \frac{1}{\lambda} \langle x | y \rangle = 0.$$

Also ist $\Phi(y) \in [x]^\perp$.

(b) Es sei $y \neq 0$ ein Eigenvektor zum Eigenwert μ . Dann gilt

$$\langle x|y \rangle = \langle \Phi(x)|\Phi(y) \rangle = \langle \lambda x|\mu y \rangle = \lambda \bar{\mu} \langle x|y \rangle.$$

Falls $\langle x|y \rangle \neq 0$, so folgt $\lambda^{-1} = \bar{\mu}$. Außerdem gilt $|\lambda| = 1$ nach Hilfssatz 6.13, das bedeutet $\lambda^{-1} = \bar{\lambda}$. Also gilt $\lambda = \overline{\lambda^{-1}} = \bar{\mu} = \mu$, im Widerspruch zur Voraussetzung. Folglich gilt $x \perp y$. ■

Satz 6.32 (Unitäre Normalform) *Es sei $(V, \langle \cdot | \cdot \rangle)$ ein unitärer Vektorraum der Dimension n .*

- (a) *Für $\Phi \in \mathbf{U}(V, \langle \cdot | \cdot \rangle)$ existiert eine Orthonormalbasis B von V , die aus Eigenvektoren von Φ besteht.*
- (b) *Für $A \in \mathbf{U}_n$ existiert $S \in \mathbf{U}_n$, so dass gilt*

$$\bar{S}^\top \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad (6.12)$$

mit $\lambda_i \in \mathbb{C}$, $|\lambda_i| = 1$.

BEWEIS:

- (a) Beweis durch Induktion über n : Für $n = 1$ ist nichts zu zeigen. Es sei nun $n > 1$. Das charakteristische Polynom $f_\Phi \in \mathbb{C}[X]$ hat eine Nullstelle $\lambda_1 \in \mathbb{C}$. Sei $x \neq 0$ ein Eigenvektor von Φ zum Eigenwert λ_1 . Nach Hilfssatz 6.31 ist $U = [x]^\perp$ ein Φ -invarianter Untervektorraum. Also ist $\Phi|_U$ eine unitäre Abbildung von U . Da $\dim U = n - 1$, können wir mit der Induktionsvoraussetzung annehmen, dass es eine Orthonormalbasis $\{b_2, \dots, b_n\}$ von U gibt, die aus Eigenvektoren von Φ besteht. Diese ergänzen wir durch $b_1 = \frac{x}{\|x\|}$ zu einer Orthonormalbasis B von V aus Eigenvektoren von Φ .
- (b) Nach Teil (a) existiert eine Orthonormalbasis $B = \{b_1, \dots, b_n\}$ aus Eigenvektoren von A . Es sei $S \in \mathbb{C}^{n \times n}$ die Matrix mit Spalten b_1, \dots, b_n . Nach Hilfssatz 6.28 ist $S \in \mathbf{U}_n$. Insbesondere ist $S^{-1} = \bar{S}^\top$ und S ist die Basiswechselmatrix von der Standardbasis zur Basis B . Folglich gilt

$$\bar{S}^\top \cdot A \cdot S = S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei λ_i der Eigenwert des Eigenvektors b_i ist. Da A unitär ist, gilt $|\lambda_i| = 1$ für alle i (Hilfssatz 6.13). ■

Um eine Basis B wie in Satz 6.32 zu bestimmen, ermittelt man zunächst die Basen der Eigenräume von Φ bzw. A und wendet dann auf jede dieser Basen das Gram-Schmidt-Verfahren an.

Hilfssatz 6.33 *Es sei $f \in \mathbb{R}[X] \subset \mathbb{C}[X]$ ein normiertes Polynom.*

(a) *Ist $\zeta \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle von f , so ist auch $\bar{\zeta}$ eine Nullstelle von f .*

(b) *Die Primfaktorzerlegung von f in $\mathbb{R}[X]$ hat die Form*

$$f = (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_k)^{\alpha_k} \cdot q_1^{\beta_1} \cdots q_m^{\beta_m}, \quad (6.13)$$

wobei $\lambda_i \in \mathbb{R}$ und die $q_j = X^2 + a_j X + b_j \in \mathbb{R}[X]$ quadratische Polynome ohne reelle Nullstellen sind.

(c) *Die Primfaktorzerlegung von f in $\mathbb{C}[X]$ hat die Form*

$$f = (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_k)^{\alpha_k} (X - \zeta_1)^{\beta_1} (X - \bar{\zeta}_1)^{\beta_1} \cdots (X - \zeta_m)^{\beta_m} (X - \bar{\zeta}_m)^{\beta_m}, \quad (6.14)$$

wobei $\zeta_j, \bar{\zeta}_j \in \mathbb{C} \setminus \mathbb{R}$ die Nullstellen des Polynoms q_j aus Teil (b) sind.

BEWEIS:

(a) Da $f \in \mathbb{R}[X]$, gilt $\bar{f} = f$. Die komplexe Konjugation ist ein Körperautomorphismus von \mathbb{C} . Ist also $f(\zeta) = 0$, so gilt auch

$$0 = \bar{0} = \overline{f(\zeta)} = f(\bar{\zeta}).$$

(b) Es sei p ein über $\mathbb{R}[X]$ irreduzibler Faktor vom Grad > 1 von f . Insbesondere besitzt p keinen reellen Linearfaktor. Da p keine reellen Nullstellen hat, besitzt es nach Teil (a) ein Nullstellenpaar $\zeta, \bar{\zeta} \in \mathbb{C} \setminus \mathbb{R}$. Als komplexes Polynom hat p somit den Teiler

$$q = (X - \zeta)(X - \bar{\zeta}) = X^2 - 2\operatorname{Re}(\zeta)X + |\zeta|^2.$$

Letzteres ist aber wieder ein reelles Polynom, und somit ist der Quotient $r = \frac{p}{q}$ ein reelles Polynom. Da p irreduzibel ist und $p = rq$ gilt, muss r ein konstanter Faktor aus \mathbb{R}^\times sein.

(c) Folgt direkt aus (b), wenn man jedes q_j in seine komplexen Linearfaktoren zerlegt. ■

Für die Bestimmung einer Normalform für orthogonale Transformationen eines euklidischen Vektorraumes V nutzen wir aus, dass jede orthogonale Matrix auch unitär ist.

Aufgabe 6.34 Für $z \in \mathbb{C}^n$ gilt: $z + \bar{z} \in \mathbb{R}^n$ und $\frac{1}{i}(z - \bar{z}) \in \mathbb{R}^n$.

Aufgabe 6.35 Fassen wir den \mathbb{R}^n als Teilmenge des \mathbb{C}^n auf, so ist das euklidische Standardskalarprodukt die Einschränkung des unitären Standardskalarprodukts auf Argumente aus \mathbb{R}^n .

Hilfssatz 6.36 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum der Dimension n und $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$. Besitzt das charakteristische Polynom f_Φ von Φ eine Nullstelle $\zeta \in \mathbb{C} \setminus \mathbb{R}$, so gilt:

- (a) $\zeta = \cos(\alpha) + i \sin(\alpha)$ für ein gewisses $\alpha \in [0, 2\pi)$.
- (b) Es existiert ein Φ -invarianter Untervektorraum U von V mit $\dim U = 2$.
- (c) Es existiert eine Orthonormalbasis B_U von U , bzgl. der die Einschränkung $\Phi|_U$ die folgende Abbildungsmatrix hat:

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \in \mathbf{SO}_2. \quad (6.15)$$

BEWEIS:

- (a) Nach Hilfssatz 6.13 ist $|\zeta| = 1$, also ist ζ von der geforderten Form.
- (b) Betrachte \mathbb{R}^n und \mathbb{C}^n mit dem jeweiligen Standardskalarprodukt. Es sei $A \in \mathbf{O}_n$ eine Abbildungsmatrix von Φ . Wir können A als unitäre Matrix auffassen. Nach Hilfssatz 6.33 treten die Nullstellen von $f_\Phi = f_A$ aus $\mathbb{C} \setminus \mathbb{R}$ in Paaren $\zeta, \bar{\zeta} \in \mathbb{C} \setminus \mathbb{R}$ auf. Ist $z \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert ζ , so gilt

$$A\bar{z} = \overline{Az} = \overline{\zeta z} = \bar{\zeta}z.$$

Folglich ist $\bar{z} \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\bar{\zeta}$. Setze

$$x = \frac{1}{i\sqrt{2}}(z - \bar{z}) \in \mathbb{R}^n, \quad y = \frac{1}{\sqrt{2}}(z + \bar{z}) \in \mathbb{R}^n.$$

Diese Vektoren sind linear unabhängig, da z und \bar{z} es sind. Nun gilt mit $\zeta = \cos(\alpha) + i \sin(\alpha)$:

$$\begin{aligned} Ax &= \frac{1}{i\sqrt{2}}(\zeta z - \bar{\zeta}z) = \cos(\alpha)x + \sin(\alpha)y, \\ Ay &= \frac{1}{\sqrt{2}}(\zeta z + \bar{\zeta}z) = -\sin(\alpha)x + \cos(\alpha)y. \end{aligned}$$

Somit ist $U' = [x, y] \subset \mathbb{R}^n$ invariant unter A und $\dim U' = 2$. Dem Untervektorraum $U' \subset \mathbb{R}^n$ entspricht ein Φ -invarianter Untervektorraum $U \subset V$.

- (c) Aus der letzten Rechnung im Beweis von (b) lesen wir ab, dass Φ bzgl. der Basis $B_U = \{x, y\}$ durch R_α dargestellt wird. Wenn wir z in (b) mit $\|z\| = 1$ wählen, so folgt aus $z \perp \bar{z}$ (Hilfssatz 6.31) und dem Satz von Pythagoras:

$$\|x\|^2 = \frac{1}{2}\|z - \bar{z}\|^2 = \frac{1}{2}(\|z\|^2 + \|\bar{z}\|^2) = \frac{1}{2}(1 + 1) = 1.$$

Entsprechend $\|y\| = 1$. Außerdem ist

$$\langle x|y \rangle = \frac{i}{2}\langle z - \bar{z}|z + \bar{z} \rangle = \frac{i}{2}(\langle z|z \rangle - 2\langle z|\bar{z} \rangle - \langle \bar{z}|\bar{z} \rangle) = \frac{i}{2}(1 - 0 - 1) = 0.$$

Also ist $\{x, y\}$ eine Orthonormalbasis. ■

Satz 6.37 (Euklidische Normalform) Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum der Dimension n .

- (a) Für $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$ existiert eine Orthonormalbasis B von V , so dass die Abbildungsmatrix von Φ bzgl. B die folgende Gestalt hat:

$$\tilde{A} = \begin{pmatrix} I_r & & & 0 \\ & -I_s & & \\ & & R_{\alpha_1} & \\ 0 & & & \ddots & \\ & & & & R_{\alpha_k} \end{pmatrix} \quad (6.16)$$

wobei $\alpha_1, \dots, \alpha_k \in (0, \pi)$, $r + s + 2k = n$, und die $R_{\alpha_i} \in \mathbf{SO}_2$ Rotationsmatrizen wie in (6.6) sind. Die Parameter $r, s, \alpha_1, \dots, \alpha_k$ sind durch Φ eindeutig bestimmt.

- (b) Für $A \in \mathbf{O}_n$ existiert ein $S \in \mathbf{O}_n$, so dass

$$S^\top \cdot A \cdot S = \tilde{A}$$

gilt für eine Matrix $\tilde{A} \in \mathbf{O}_n$ der Form (6.16).

BEWEIS: Beweis durch Induktion über n . Für $n = 1$ ist nichts zu zeigen. Es sei nun $n > 1$. Hat das charakteristische Polynom f_Φ keine reelle Nullstelle, so folgt aus Hilfssatz 6.33, dass $n = 2k$ ist für ein $k \in \mathbb{N}$. Für ein Eigenwertpaar $\zeta, \bar{\zeta} \in \mathbb{C} \setminus \mathbb{R}$ von Φ existiert nach Hilfssatz 6.36 ein Untervektorraum U mit Orthonormalbasis $B_U = \{x, y\}$, auf dem $\Phi|_U$ durch R_α dargestellt wird. Dabei ist α durch $\zeta = \cos(\alpha) + i \sin(\alpha)$ bestimmt. Die Fälle $\alpha = 0, \pi$ sind dabei ausgeschlossen, da hier reelle Eigenwerte existieren würden. Sollte $2\pi > \alpha > \pi$ gelten, so gelangt man durch den Basiswechsel

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

in U zu einem R_α mit $\alpha \in (0, \pi)$. Das orthogonale Komplement $W = U^\perp$ hat Dimension $n-2$ und ist invariant unter Φ : Da U invariant ist, gilt nämlich $\langle \Phi(w)|u \rangle = \langle w|\Phi^{-1}(u) \rangle = \langle w|u' \rangle = 0$, für $w \in W$, $u \in U$ mit $\Phi(u) = u' \in U$. Nach Induktionsvoraussetzung existiert eine Orthonormalbasis B_W von W , so dass $\Phi|_W$ durch eine Matrix der Form (6.16) dargestellt wird. Setze $B = B_U \cup B_W$, dann hat \tilde{A} die geforderte Gestalt (in diesem Fall mit $r = s = 0$).

Besitzt das charakteristische Polynom f_Φ eine reelle Nullstelle $\lambda = \pm 1$, so folgt die Behauptung analog zum unitären Fall (Satz 6.32). ■

Algorithmus 6.38 Zur praktischen Bestimmung der euklidischen Normalform \tilde{A} von $A \in \mathbf{O}_n$ geht man wie folgt vor:

- (i) Bestimme die reellen und komplexen Eigenwerte von A .
- (ii) Falls reelle Eigenwerte ± 1 existieren, bestimme mit dem Gram-Schmidt-Verfahren Orthonormalbasen B_1 und B_{-1} ihrer jeweiligen Eigenräume.
- (iii) Für jedes komplexe konjugierte Eigenwertpaar $\zeta, \bar{\zeta} \in \mathbb{C} \setminus \mathbb{R}$:
 - (iii.a) Bestimme eine Orthonormalbasis z_1, \dots, z_j des Eigenraumes in \mathbb{C}^n zum Eigenwert ζ .
 - (iii.b) Für $i = 1, \dots, j$: Setze

$$\begin{aligned} x_i &= \frac{1}{i\sqrt{2}}(z_i - \bar{z}_i), \\ y_i &= \frac{1}{\sqrt{2}}(z_i + \bar{z}_i). \end{aligned}$$

Nach Hilfssatz 6.36 ist $\{x_i, y_i\}$ Orthonormalbasis eines Untervektorraums $U_i \subset \mathbb{R}^n$, auf dem A durch eine Drehmatrix R_{α_i} dargestellt wird. Falls $2\pi > \alpha_i > \pi$, vertausche die Reihenfolge von x_i und y_i in der Basis.

- (iii.c) Vereinige die x_i, y_i zu einer Basis

$$B_\zeta = \{x_1, y_1, \dots, x_j, y_j\}.$$

- (iv) Die gesuchte Orthonormalbasis ist

$$B = B_1 \cup B_{-1} \cup B_{\zeta_1} \cup \dots \cup B_{\zeta_k}.$$

Die Matrix $S \in \mathbf{O}_n$ erhalten wir, indem wir die Elemente von B spaltenweise in eine Matrix schreiben. Dann gilt

$$S^\top \cdot A \cdot S = \tilde{A}.$$

Beispiel 6.39 Es sei

$$A = \frac{1}{8} \begin{pmatrix} 3\sqrt{2}-2 & 2\sqrt{6} & \sqrt{6}+2\sqrt{3} \\ -2\sqrt{6} & 4\sqrt{2} & -2\sqrt{2} \\ \sqrt{6}+2\sqrt{3} & 2\sqrt{2} & \sqrt{2}-6 \end{pmatrix} \in \mathbf{O}_3.$$

Das charakteristische Polynom von A ist

$$f_A = \det(XI_3 - A) = (X + 1) \cdot (X^2 - \sqrt{2}X + 1).$$

Seine Nullstellen sind

$$\lambda = -1, \quad \zeta = \frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}, \quad \bar{\zeta} = \frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}.$$

Der Eigenraum E_{-1} von A ist

$$E_{-1} = \text{Kern}(A + I_3) = \left[\frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ \sqrt{3} \end{pmatrix} \right].$$

Der Vektor

$$b_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ \sqrt{3} \end{pmatrix}$$

ist normiert und wird als erstes Element einer Orthonormalbasis B von \mathbb{R}^3 gewählt. Um eine geeignete Orthonormalbasis $\{b_2, b_3\}$ von E_{-1}^\perp zu bestimmen, bestimmen wir zunächst den Eigenraum $E_\zeta^\mathbb{C}$ in \mathbb{C}^3 :

$$E_\zeta^\mathbb{C} = \text{Kern}(A - \zeta I_3) = \left[\begin{pmatrix} -\frac{\sqrt{3}i}{2} \\ 1 \\ -\frac{i}{2} \end{pmatrix} \right]$$

Dann bilden

$$z = \frac{1}{\sqrt{2}} \begin{pmatrix} -\frac{\sqrt{3}i}{2} \\ 1 \\ -\frac{i}{2} \end{pmatrix}, \quad \bar{z} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{\sqrt{3}i}{2} \\ 1 \\ \frac{i}{2} \end{pmatrix}$$

eine Orthonormalbasis von $E_\zeta^\mathbb{C} \oplus E_{\bar{\zeta}}^\mathbb{C}$. Wie im Beweis von Hilfssatz 6.36 setze

$$x = \frac{1}{i\sqrt{2}}(z - \bar{z}) = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}, \quad y = \frac{1}{\sqrt{2}}(z + \bar{z}) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Dann bilden $b_2 = x, b_3 = y$ eine Orthonormalbasis von E_{-1}^\perp . Schreiben wir b_1, b_2, b_3 in die Spalten einer Matrix, so erhalten wir eine orthogonale Basiswechselmatrix

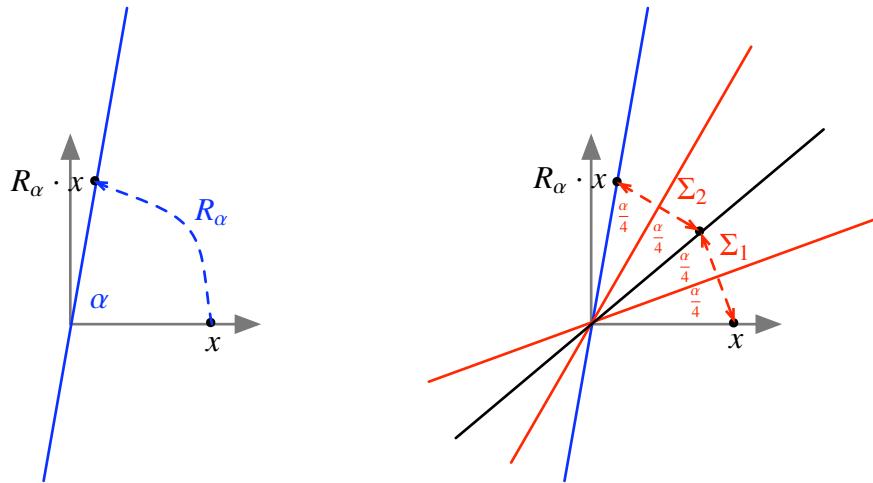
$$S = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \end{pmatrix} \in \mathbf{O}_3.$$

Nun ist

$$\tilde{A} = S^T \cdot A \cdot S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ 0 & \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{pmatrix}$$

die euklidische Normalform von A .

Als Anwendung der Normalform für lineare Isometrien wollen wir zeigen, dass jede lineare Isometrie als Produkt von Spiegelungen der Form (6.2) darstellbar ist. Dazu betrachten wir zunächst den zweidimensionalen Fall: Eine Rotation R_α im \mathbb{R}^2 um den Winkel α kann als Verknüpfung zweier Spiegelungen $\Sigma_2 \circ \Sigma_1$ dargestellt werden.



Die Zeichnung veranschaulicht die Lage der beiden Spiegelungsachsen. Sie schließen den Winkel $\frac{\alpha}{2}$ ein. Die erste Spiegelungsachse ist die um $\frac{\alpha}{4}$ gedrehte e_1 -Achse, somit wird Σ_1 durch die Matrix

$$\begin{pmatrix} \cos(\frac{\alpha}{4}) & -\sin(\frac{\alpha}{4}) \\ \sin(\frac{\alpha}{4}) & \cos(\frac{\alpha}{4}) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{\alpha}{4}) & -\sin(\frac{\alpha}{4}) \\ \sin(\frac{\alpha}{4}) & \cos(\frac{\alpha}{4}) \end{pmatrix}^{-1} = \begin{pmatrix} \cos(\frac{\alpha}{2}) & \sin(\frac{\alpha}{2}) \\ \sin(\frac{\alpha}{2}) & -\cos(\frac{\alpha}{2}) \end{pmatrix}$$

dargestellt. Die zweite Spiegelungsachse ist die um $\frac{3\alpha}{4}$ gedrehte e_1 -Achse, d.h. Σ_2 wird durch

$$\begin{pmatrix} \cos(\frac{3\alpha}{4}) & -\sin(\frac{3\alpha}{4}) \\ \sin(\frac{3\alpha}{4}) & \cos(\frac{3\alpha}{4}) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{3\alpha}{4}) & -\sin(\frac{3\alpha}{4}) \\ \sin(\frac{3\alpha}{4}) & \cos(\frac{3\alpha}{4}) \end{pmatrix}^{-1} = \begin{pmatrix} \cos(\frac{3\alpha}{2}) & \sin(\frac{3\alpha}{2}) \\ \sin(\frac{3\alpha}{2}) & -\cos(\frac{3\alpha}{2}) \end{pmatrix}$$

dargestellt. In der Tat ergibt die Verkettung $\Sigma_2 \circ \Sigma_1$ nun

$$\begin{pmatrix} \cos(\frac{3\alpha}{2}) & \sin(\frac{3\alpha}{2}) \\ \sin(\frac{3\alpha}{2}) & -\cos(\frac{3\alpha}{2}) \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{\alpha}{2}) & \sin(\frac{\alpha}{2}) \\ \sin(\frac{\alpha}{2}) & -\cos(\frac{\alpha}{2}) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} = R_\alpha.$$

Satz 6.40 (Cartan-Dieudonné) *Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum der Dimension n . Die Gruppe $\mathbf{O}(V, \langle \cdot | \cdot \rangle)$ wird von den Spiegelungen der Form (6.2) erzeugt. Insbesondere lässt sich jede lineare Isometrie Φ als Produkt von höchstens n Spiegelungen darstellen:*

$$\Phi = \Sigma_k \circ \cdots \circ \Sigma_1, \quad (6.17)$$

wobei $k \leq n$ und $\Sigma_i = \Sigma_{U_i}$ für geeignete $n - 1$ -dimensionale Untervektorräume U_i .

BEWEIS: Stellt man Φ in der Form (6.16) dar, so sieht man, dass jede Rotationsmatrix R_{α_i} wie oben beschrieben also Produkt von zwei Spiegelungen geschrieben werden kann. Da die Anzahl dieser Blöcke R_{α_i} immer $k \leq \frac{n}{2}$ beträgt, lässt sich Φ als Produkt von $2k \leq n$ Spiegelungen darstellen. ■

6.4 Euler-Winkel

In technischen Anwendungen bedient man sich gerne einer sehr einfachen Darstellung der Elemente von \mathbf{SO}_3 :

Satz 6.41 *Es sei $A \in \mathbf{SO}_3$. Dann gibt es Winkel α, β, γ , so dass gilt*

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Winkel α, β, γ heißen **Euler-Winkel** von A .

BEWEIS: Die Standardbasis von \mathbb{R}^3 sei $B = \{e_1, e_2, e_3\}$. Weiter sei

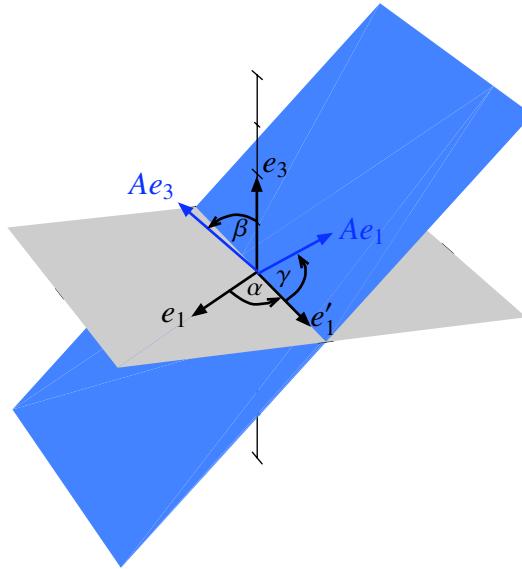
$$e'_1 \in [e_1, e_2] \cap [Ae_1, Ae_2] \quad \text{mit} \quad \|e'_1\| = 1.$$

Setze

$$\alpha = \varphi(e_1, e'_1), \quad \beta = \varphi(e_3, Ae_3), \quad \gamma = \varphi(e'_1, Ae_1).$$

Weiter bezeichne

- Ψ_1 die Drehung um $[e_3]$ um den Winkel α ,
- Ψ_2 die Drehung um $[e'_1]$ um den Winkel β ,
- Ψ_3 die Drehung um $[Ae_3]$ um den Winkel γ .



Also gilt:

$$\begin{aligned}\Psi_1(e_3) &= e_3, & \Psi_1(e_1) &= e'_1, \\ \Psi_2(e'_1) &= e'_1, & \Psi_2(e_3) &= Ae_3, \\ \Psi_3(Ae_3) &= Ae_3, & \Psi_3(e'_1) &= Ae_1.\end{aligned}$$

Es sei nun $\Psi = \Psi_1 \circ \Psi_2 \circ \Psi_3$. Dann gilt:

$$\begin{aligned}\Psi(e_1) &= \Psi_3(\Psi_2(e'_1)) = \Psi_3(e'_1) = Ae_1, \\ \Psi(e_3) &= \Psi_3(\Psi_2(e_3)) = \Psi_3(Ae_3) = Ae_3,\end{aligned}$$

und somit ist A die Abbildungsmatrix der Drehung Ψ bzgl. der Standardbasis, $A = M_B^B(\Psi)$. Wir führen zwei weitere Basen des \mathbb{R}^3 ein:

$$B' = \Psi_1(B), \quad B'' = \Psi_2(B') = \Psi_2(\Psi_1(B)).$$

Die Übergangsmatrizen für die entsprechenden Basiswechsel werden mit $M_{B'}^{B'}$, $M_{B'}^{B''}$, $M_B^{B''}$ bezeichnet. Die Abbildungsmatrizen bzgl. dieser Basen haben folgende

Gestalt:

$$\begin{aligned}
 R_{3,\alpha} &= M_B^B(\Psi_1) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} = M_B^{B'}, \\
 R_{1,\beta} &= M_{B'}^{B'}(\Psi_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{pmatrix} = (M_B^{B'})^{-1} M_B^B(\Psi_2) M_B^{B'} = R_{\alpha}^{-1} M_B^B(\Psi_2) R_{\alpha} = M_{B'}^{B''}, \\
 R_{3,\gamma} &= M_{B''}^{B''}(\Psi_3) = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix} = (M_B^{B''})^{-1} M_B^B(\Psi_3) M_B^{B''} = (M_{B'}^{B''})^{-1} (M_B^{B'})^{-1} M_B^B(\Psi_3) M_B^{B'} M_{B'}^{B''}.
 \end{aligned}$$

Daraus folgt

$$\begin{aligned}
 A &= M_B^B(\Psi) = M_B^B(\Psi_3) M_B^B(\Psi_2) M_B^B(\Psi_1) \\
 &= R_{3,\alpha} R_{1,\beta} R_{3,\gamma} R_{1,\beta}^{-1} R_{3,\alpha}^{-1} R_{3,\alpha} R_{1,\beta} R_{3,\alpha}^{-1} R_{3,\alpha} \\
 &= R_{3,\alpha} R_{1,\beta} R_{3,\gamma},
 \end{aligned}$$

und dies ist gerade die gewünschte Darstellung. ■

Aufgabe 6.42 Es gilt

$$R_{3,\alpha} \cdot R_{1,\beta} \cdot R_{3,\gamma} = R_{3,\alpha+\pi} \cdot R_{1,-\beta} \cdot R_{3,\gamma+\pi}. \quad (6.18)$$

Durch Koeffizientenvergleich kann man die Euler-Winkel einer gegebenen Matrix $A \in \mathbf{SO}_3$ ablesen, wie das folgende Beispiel zeigt:

Beispiel 6.43 Es sei

$$A = \begin{pmatrix} \frac{\sqrt{3}}{4} & -\frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \end{pmatrix} \in \mathbf{SO}_3.$$

Wir machen den Ansatz

$$A = R_{3,\alpha} R_{1,\beta} R_{3,\gamma} = \begin{pmatrix} * & * & \sin(\alpha) \sin(\beta) \\ * & * & -\cos(\alpha) \sin(\beta) \\ \sin(\beta) \sin(\gamma) & \cos(\gamma) \sin(\beta) & \cos(\beta) \end{pmatrix}.$$

Die letzte Zeile und Spalte von A enthalten genug Information, um die Euler-Winkel zu bestimmen. Mit a_{ij} bezeichnen wir die Einträge von A . Der Eintrag a_{33} liefert

$$\cos(\beta) = a_{33} = 0,$$

und dies bedeutet

$$\beta = \frac{\pi}{2} \quad \text{oder} \quad \beta = -\frac{\pi}{2}.$$

Nach (6.18) führen beide Möglichkeiten auf zulässige Euler-Winkel. Wir entscheiden uns für

$$\beta = \frac{\pi}{2}.$$

Die Einträge a_{31}, a_{32} teilen wir durch $\sin(\beta) = 1$ und erhalten

$$\frac{a_{31}}{\sin(\beta)} = \sin(\gamma) = \frac{\sqrt{3}}{2}, \quad \frac{a_{32}}{\sin(\beta)} = \cos(\gamma) = \frac{1}{2}.$$

Folglich ist

$$\gamma = \frac{\pi}{3}.$$

Genauso bestimmen wir α :

$$\frac{a_{13}}{\sin(\beta)} = \sin(\alpha) = \frac{1}{2}, \quad \frac{a_{23}}{\sin(\beta)} = -\cos(\alpha) = -\frac{\sqrt{3}}{2}.$$

Folglich ist

$$\alpha = \frac{\pi}{6}.$$

Es ist also

$$\begin{pmatrix} \frac{\sqrt{3}}{4} & -\frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi}{6}) & -\sin(\frac{\pi}{6}) & 0 \\ \sin(\frac{\pi}{6}) & \cos(\frac{\pi}{6}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ 0 & \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) & 0 \\ \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bemerkung 6.44 In technischen Anwendungen ist auch die ähnliche Darstellung

$$A = R_{1,\alpha}R_{2,\beta}R_{3,\gamma}$$

für $A \in \mathbf{SO}_3$ gebräuchlich. Die Winkel α, β, γ werden **Roll-Nick-Gier-Winkel** genannt, in Anlehnung an die drei Freiheitsgrade zur Steuerung eines Flugzeugs. Beachte, dass hier um drei verschiedene Achsen rotiert wird.

7 Selbstadjungierte Endomorphismen

7.1 Die adjugierte Abbildung

Definition 7.1 Es seien V_1, V_2 Vektorräume mit Skalarprodukt und $\Phi : V_1 \rightarrow V_2$ eine lineare Abbildung. Eine lineare Abbildung $\Phi^* : V_2 \rightarrow V_1$ heißt die zu Φ **adjugierte Abbildung**, falls für alle $x_1 \in V_1, x_2 \in V_2$ gilt:

$$\langle \Phi(x_1) | x_2 \rangle_2 = \langle x_1 | \Phi^*(x_2) \rangle_1. \quad (7.1)$$

Hilfssatz 7.2 Falls die adjugierte Abbildung Φ^* existiert, so ist sie eindeutig.

BEWEIS: Es seien $\Phi^*, \Phi' : V_2 \rightarrow V_1$ lineare Abbildungen, die (7.1) erfüllen. Für alle $x_1 \in V_1, x_2 \in V_2$ gilt also

$$\langle x_1 | \Phi'(x_2) \rangle_1 = \langle \Phi(x_1) | x_2 \rangle_2 = \langle x_1 | \Phi^*(x_2) \rangle_1,$$

was gleichbedeutend ist mit

$$0 = \langle x_1 | \Phi^*(x_2) - \Phi'(x_2) \rangle_1.$$

Wegen der positiven Definitheit des Skalarproduktes muss

$$\Phi^*(x_2) = \Phi'(x_2)$$

gelten für alle $x_2 \in V_2$. ■

Beispiel 7.3 Sind $V_1 = \mathbb{C}^n$ und $V_2 = \mathbb{C}^m$ mit den jeweiligen unitären Standardskalarprodukten, und ist $A \in \mathbb{C}^{m \times n}$ die Abbildungsmatrix von Φ bzgl. der jeweiligen Standardbasen, so gilt

$$\langle \Phi(x_1) | x_2 \rangle_2 = \langle Ax_1 | x_2 \rangle_2 = x_1^\top \cdot A^\top \cdot \bar{x}_2 = x_1^\top \cdot \overline{A^\top \cdot x_2} = \langle x_1 | \overline{A^\top} x_2 \rangle_1.$$

Die Abbildung $\Phi^* : V_2 \rightarrow V_1, x \mapsto \overline{A^\top} x$, erfüllt also die Eigenschaften der zu Φ adjungierten Abbildung.

Haben V_1, V_2 endliche Dimension, so lassen sich die jeweiligen Skalarprodukte durch Wahl von Orthonormalbasen stets durch die Standardskalarprodukte in \mathbb{C}^n bzw. \mathbb{C}^m darstellen. Beispiel 7.3 zeigt dann, dass die zu Φ adjungierte Abbildung in diesem Fall existiert und durch die Matrix $\overline{A^\top}$ dargestellt wird. Es gilt also:

Satz 7.4 Es sei $\Phi : V_1 \rightarrow V_2$ eine lineare Abbildung von Vektorräumen mit Skalarprodukt. Sind V_1, V_2 endlichdimensional, so existiert eine eindeutige zu Φ adjugierte Abbildung $\Phi^* : V_2 \rightarrow V_1$.

Bemerkung 7.5 Die Matrix $\overline{A^\top}$ wird auch die zu A **adjungierte Matrix** genannt. Gebräuchliche Schreibweisen sind $A^* = \overline{A^\top} = A^\dagger$. Ist $A \in \mathbb{R}^{m \times n}$, so ist natürlich $A^* = \overline{A^\top} = \overline{A}^\top = A^\top$.

Beispiel 7.6 Es sei $\Phi \in \mathbf{O}(V, \langle \cdot | \cdot \rangle)$ eine lineare Isometrie. Dann gilt

$$\langle \Phi(x) | y \rangle = \langle \Phi(x) | \Phi(\Phi^{-1}(y)) \rangle = \langle x | \Phi^{-1}(y) \rangle$$

für alle $x, y \in V$. Also gilt für lineare Isometrien:

$$\Phi^* = \Phi^{-1}. \quad (7.2)$$

Dies ist konsistent mit der Bedingung $A^* = A^\top = A^{-1}$ für orthogonale Matrizen bzw. $A^* = \overline{A^\top} = A^{-1}$ für unitäre Matrizen.

Die Schreibweise für die adjungierte Abbildung Φ^* erinnert nicht zufällig an die Schreibweise für die duale Abbildung $\Phi^* : V^* \rightarrow V$. Im Falle endlicher Dimension kann man über das Skalarprodukt den Raum V mit seinem Dualraum V^* identifizieren:

Satz 7.7 (Fréchet-Riesz) *Es sei V ein Vektorraum mit Skalarprodukt und $\dim V < \infty$. Für jede Linearform $\varphi \in V^*$ existiert ein eindeutiges $x_\varphi \in V$, so dass für alle $y \in V$ gilt:*

$$\varphi(y) = \langle y | x_\varphi \rangle. \quad (7.3)$$

BEWEIS: Es sei $B = \{b_1, \dots, b_n\}$ eine Orthonormalbasis von V . Schreibe $\lambda_k = \varphi(b_k)$ für $k = 1, \dots, n$. Nach Satz 5.9 ist $y = \langle y | b_1 \rangle b_1 + \dots + \langle y | b_n \rangle b_n$. Dann ist

$$\varphi(y) = \langle y | b_1 \rangle \lambda_1 + \dots + \langle y | b_n \rangle \lambda_n = \langle y | \bar{\lambda}_1 b_1 + \dots + \bar{\lambda}_n b_n \rangle.$$

Das Element

$$x_\varphi = \bar{\lambda}_1 b_1 + \dots + \bar{\lambda}_n b_n$$

erfüllt also (7.3) für alle $y \in V$. Die Eindeutigkeit von x_φ folgt aus der positiven Definitheit. ■

Folgerung 7.8 *Ist der Vektorraum V in Satz 7.7 euklidisch, so ist die Zuordnung $\iota : V \rightarrow V^*$, $x \mapsto \langle \cdot | x \rangle$, ein Isomorphismus von Vektorräumen.*

BEWEIS: Im euklidischen Fall ist die Abbildung ι linear, und für $\varphi = \langle \cdot | x \rangle$ ist $x_\varphi = x$ in (7.3). Es ist $\text{Kern } \iota = \{0\}$. Also ist ι ein Isomorphismus. ■

Für unitäres V ist die Abbildung ι eine semilineare Abbildung, d.h. ι ist zwar additiv, aber für $\lambda \in \mathbb{C}$ gilt $\iota(\lambda x) = \langle \cdot | \lambda x \rangle = \bar{\lambda} \langle \cdot | x \rangle = \bar{\lambda} \iota(x)$.

Es sei $\Phi : V \rightarrow V$ linear. Unter der Identifizierung $V \cong V^*$ aus Folgerung 7.8 entspricht die adjungierte Abbildung Φ^* der dualen Abbildung Φ^* :

$$\iota(\Phi^*(x))(y) = \langle y|\Phi^*(x)\rangle = \langle\Phi(y)|x\rangle = \iota(x)(\Phi(y)) = \Phi^*(\iota(x))(y),$$

oder kurz

$$\iota \circ \Phi^* = \Phi^* \circ \iota.$$

Der Satz von Fréchet-Riesz gilt in unendlicher Dimension, falls V ein Hilbert-Raum ist und man anstelle des ganzen Dualraums V^* nur die stetigen Linearformen $V^* \cap C(V, \mathbb{C})$ betrachtet (Werner [14], Theorem V.3.6).

Im Allgemeinen muss im Falle unendlicher Dimension die adjungierte Abbildung aber gar nicht existieren:

Beispiel 7.9 Betrachte die euklidischen Vektorräume $V_1 = \mathbb{R}[X] \cap C([0, 1])$ und $V_2 = C([0, 1])$, jeweils mit dem Skalarprodukt

$$\langle f|g \rangle = \int_0^1 f(t)g(t)dt.$$

Es sei $\Phi : V_1 \rightarrow V_2$ die identische Einbettung, $\Phi(f) = f$. Angenommen, die adjungierte Abbildung $\Phi^* : V_2 \rightarrow V_1$ existiert. Dann gilt für alle $f \in V_1, g \in V_2$:

$$\langle f|g \rangle = \langle \Phi(f)|g \rangle = \langle f|\Phi^*(g) \rangle.$$

Das bedeutet $\langle f|g - \Phi^*(g) \rangle = 0$ für alle $f \in V_1$, also

$$g - \Phi^*(g) \perp V_1$$

für alle $g \in V_2$. Aber $V_1^\perp = \{0\}$, also gilt $g = \Phi^*(g) \in V_1$. Dies ist ein Widerspruch, da $g \in V_2 \setminus V_1$ existieren.

Wir studieren nun die Eigenschaften der adjungierten Abbildungen.

Hilfssatz 7.10 Es seien V_1, V_2, V_3 Vektorräume mit Skalarprodukt und es seien $\Phi : V_1 \rightarrow V_2, \Psi : V_2 \rightarrow V_3$ lineare Abbildungen. Wir nehmen an, dass die adjungierten Abbildungen $\Phi^* : V_2 \rightarrow V_1, \Psi^* : V_3 \rightarrow V_2$ existieren. Dann gilt:

- (a) $(\Phi^*)^* = \Phi$.
- (b) $(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*$.
- (c) $\text{Kern } \Phi = (\text{Bild } \Phi^*)^\perp$ und $\text{Kern } \Phi^* = (\text{Bild } \Phi)^\perp$.

Beweis:

(a) Folgt direkt aus der Definition.

(b) Es ist

$$\langle \Psi(\Phi(x))|y\rangle_3 = \langle \Phi(x)|\Psi^*(y)\rangle_2 = \langle x|\Phi^*(\Psi^*(y))\rangle_1$$

für alle $x \in V_1, y \in V_3$.

(c) $x \in \text{Kern } \Phi$ bedeutet $\Phi(x) = 0$. Wegen der positiven Definitheit ist das äquivalent zu

$$0 = \langle \Phi(x)|y\rangle_2 = \langle x|\Phi^*(y)\rangle_1$$

für alle $y \in V_2$, bzw. zu $x \perp \Phi^*(V_2) = \text{Bild } \Phi^*$. ■

Aufgabe 7.11 Ist Φ^* surjektiv, so ist Φ injektiv.

Hilfssatz 7.12 Es sei V ein Vektorraum mit Skalarprodukt und $\dim V = n$. Weiter sei $\Phi \in \text{End}(V)$. Dann ist $\lambda \in \mathbb{C}$ genau dann ein Eigenwert von Φ , wenn $\bar{\lambda}$ ein Eigenwert von Φ^* ist.

Beweis: Es seien A und $A^* = \overline{A^\top}$ die Abbildungsmatrizen von Φ und Φ^* bzgl. einer Orthonormalbasis von V . Ist $\lambda \in \mathbb{C}$ eine Nullstelle des charakteristischen Polynoms f_A , so gilt

$$\begin{aligned} 0 = \bar{0} &= \overline{f_A(\lambda)} = \overline{\det(\lambda I_n - A)} = \overline{\det((\lambda I_n - A)^\top)} = \overline{\det(\lambda I_n - A^\top)} \\ &= \det(\bar{\lambda} I_n - \overline{A^\top}) = f_{A^*}(\bar{\lambda}). \end{aligned}$$

Also ist $\bar{\lambda}$ ein Eigenwert von A^* und somit von Φ^* . ■

7.2 Selbstadjungierte Endomorphismen

Definition 7.13 Es sei V ein Vektorraum mit Skalarprodukt. $\Phi \in \text{End}(V)$ heißt **selbstadjungiert**, falls $\Phi = \Phi^*$ gilt.

Hilfssatz 7.14 Es sei V ein Vektorraum mit Skalarprodukt und B eine Orthonormalbasis von V . Ein Endomorphismus Φ von V ist genau dann selbstadjungiert, wenn die Abbildungsmatrix A von Φ bzgl. B symmetrisch ist (im euklidischen Fall) bzw. hermitesch ist (im unitären Fall).

Beweis: Die Abbildungsmatrix von Φ^* ist A^* . Also gilt $A = A^*$. Im euklidischen Fall bedeutet das, dass $A = A^\top$ symmetrisch ist, im unitären Fall, dass $A = \overline{A^\top}$ hermitesch ist. ■

Beispiel 7.15 (Selbstadjungierte Abbildungen)

- (a) Orthogonalprojektionen $\Pi_U : V \rightarrow V$ auf einen Untervektorraum U (siehe (5.7)) sind selbstadjungiert. Es gilt nämlich für alle $x, y \in V$:

$$\begin{aligned}\langle \Pi_U(x)|y \rangle &= \langle \Pi_U(x)|\Pi_U(y) + (y - \Pi_U(y)) \rangle = \langle \Pi_U(x)|\Pi_U(y) \rangle + \underbrace{\langle \Pi_U(x)|\Pi_{U^\perp}(y) \rangle}_{=0} \\ &= \langle \Pi_U(x)|\Pi_U(y) \rangle + \underbrace{\langle \Pi_{U^\perp}(x)|\Pi_U(y) \rangle}_{=0} = \langle \Pi_U(x) + (x - \Pi_U(x))|\Pi_U(y) \rangle \\ &= \langle x|\Pi_U(y) \rangle.\end{aligned}$$

- (b) Spiegelungen Σ sind selbstadjungiert, denn einerseits gilt $\Sigma = \Sigma^{-1}$ für jede Spiegelung, andererseits gilt $\Sigma^{-1} = \Sigma^*$, da Spiegelungen Isometrien sind.
- (c) Jede symmetrische Matrix $S \in \mathbb{R}^{n \times n}$ definiert einen selbstadjungierten Endomorphismus $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Sx$.

Aufgabe 7.16 Es sei $\Phi \in \text{End}(V)$. Dann sind $\Phi \circ \Phi^*$ und $\Phi^* \circ \Phi$ selbstadjungiert.

Satz 7.17 Es sei V ein Vektorraum mit Skalarprodukt und $\dim V < \infty$. Ist Φ ein selbstadjungierter Endomorphismus von V , so sind alle Eigenwerte von Φ reell. Insbesondere zerfällt das charakteristische Polynom f_Φ in reelle Linearfaktoren.

Beweis: Es sei $x \neq 0$ ein Eigenvektor von Φ zum Eigenwert λ . Da Φ selbstadjungiert ist, gilt

$$\lambda \langle x|x \rangle = \langle \lambda x|x \rangle = \langle \Phi(x)|x \rangle = \langle x|\Phi^*(x) \rangle = \langle x|\Phi(x) \rangle = \langle x|\lambda x \rangle = \bar{\lambda} \langle x|x \rangle.$$

Also gilt $\lambda = \bar{\lambda} \in \mathbb{R}$. Im unitären Fall zerfällt f_Φ über $\mathbb{C}[X]$ in Linearfaktoren, und nach dem gerade gezeigten sind alle Linearfaktoren reell. Somit zerfällt f_Φ über $\mathbb{R}[X]$ in Linearfaktoren.

Im euklidischen Fall verwendet man, dass $f_\Phi = f_A$ gilt für eine geeignete symmetrische Abbildungsmatrix A von Φ . Da A insbesondere hermitesch ist, folgt aus dem unitären Fall, dass f_A auch hier in reelle Linearfaktoren zerfällt. ■

7.3 Der Spektralsatz

Der letzte Satz 7.17 impliziert, dass selbstadjungierte Endomorphismen endlich-dimensionalen Vektorräumen mit Skalarprodukt stets Eigenvektoren besitzen. Sie sind sogar diagonalisierbar.

Satz 7.18 (Spektralsatz) Es sei V ein Vektorraum mit Skalarprodukt der Dimension n . Ist Φ ein selbstadjungierter Endomorphismus von V , so existiert eine Orthonormalbasis B von V aus Eigenvektoren von Φ .

BEWEIS: Beweis durch Induktion über $n = \dim V$: Für $n = 1$ ist nichts zu zeigen. Es sei nun $n > 1$. Nach Satz 7.17 existiert (auch im euklidischen Fall) ein Eigenvektor $x \neq 0$ zum Eigenwert λ von Φ . Setze $b_1 = \frac{x}{\|x\|}$. Dann ist $W = [b_1]^\perp$ ein Φ -invarianter Untervektorraum: Ist $w \in W$, so gilt

$$\langle \Phi(w)|b_1 \rangle = \langle w|\Phi(b_1) \rangle = \langle w|\lambda b_1 \rangle = \lambda \underbrace{\langle w|b_1 \rangle}_{=0} = 0.$$

Es folgt $\Phi(W) \subset W$. Also ist $\Phi|_W$ eine selbstadjungierte Abbildung von W und $\dim W = n - 1$. Nach Induktionsvoraussetzung existiert somit eine Orthonormalbasis $\{b_2, \dots, b_n\}$ von W aus Eigenvektoren von $\Phi|_W$. Dann ist $B = \{b_1, b_2, \dots, b_n\}$ eine Orthonormalbasis von V aus Eigenvektoren von Φ . ■

Folgerung 7.19 Es sei $A \in \mathbb{C}^{n \times n}$ (bzw. $A \in \mathbb{R}^{n \times n}$) eine hermitesche (bzw. symmetrische) Matrix. Dann existiert $T \in \mathbf{U}_n$ (bzw. $T \in \mathbf{O}_n$), so dass gilt

$$T^* \cdot A \cdot T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{mit } \lambda_1, \dots, \lambda_n \in \mathbb{R}. \quad (7.4)$$

BEWEIS: A definiert einen selbstadjungierten Endomorphismus Φ von \mathbb{C}^n (bzw. \mathbb{R}^n). Nach dem Spektralsatz existiert für Φ eine Orthonormalbasis B aus Eigenvektoren. Wähle T als die Basiswechselmatrix für den Übergang von der B zur Standardbasis. ■

Wir erhalten ein weiteres Kriterium für positive Definitheit:

Folgerung 7.20 Eine symmetrische Matrix $S \in \mathbb{R}^{n \times n}$ ist genau dann positiv definit, wenn alle Eigenwerte von S positiv sind.

BEWEIS: Eine Richtung des Beweises ist Hilfssatz 4.28 (c).

Sind umgekehrt alle Eigenwerte $\lambda_i > 0$, so folgt aus Folgerung 7.19, dass S zu einer Diagonalmatrix \tilde{S} mit positiven Diagonaleinträgen $\lambda_i > 0$ durch Konjugation mit einer Matrix $T \in \mathbf{O}_n$ ähnlich ist. Die Unterdeterminanten $\det(\tilde{S}_k) = \lambda_1 \cdots \lambda_k$ sind positiv für $k = 1, \dots, n$. Nach dem Hurwitz-Kriterium (Satz 5.30 (iii)) ist \tilde{S} positiv definit. Nun ist

$$x^\top \cdot S \cdot x = x^\top \cdot T \cdot \tilde{S} \cdot T^\top \cdot x = (T^\top x)^\top \cdot \tilde{S} \cdot (T^\top x) > 0$$

für alle $x \in \mathbb{R}^n$ und somit ist S positiv definit. ■

Folgerung 7.21 Ist $S \in \mathbb{R}^{n \times n}$ eine positiv definite Matrix, so hat S eine Wurzel, d.h. es gibt eine positiv definite Matrix \sqrt{S} , die

$$S = \sqrt{S} \cdot \sqrt{S}$$

erfüllt.

BEWEIS: Es sei $T \in \mathbf{O}_n$ die Matrix aus (7.4). Setze

$$\sqrt{S} = T \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot T^\top,$$

wobei die $\lambda_i > 0$ die Eigenwerte von S sind. Dann gilt wegen $T^\top = T^{-1}$:

$$\begin{aligned} \sqrt{S} \cdot \sqrt{S} &= T \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot T^\top \cdot T \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot T^\top \\ &= T \cdot \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \cdot T^\top = T \cdot T^\top \cdot S \cdot T \cdot T^\top \\ &= S. \end{aligned}$$

Somit ist \sqrt{S} die Wurzel von S . ■

Aufgabe 7.22 (Spektraldarstellung) Es sei V ein Vektorraum mit Skalarprodukt der Dimension n . Ist Φ ein selbstadjungierter Endomorphismus von V mit verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$, so lässt sich Φ wie folgt darstellen:

$$\Phi = \lambda_1 \Pi_{E_{\lambda_1}} + \dots + \lambda_k \Pi_{E_{\lambda_k}}. \quad (7.5)$$

7.4 Normale Endomorphismen

Die Beweise des Spektralsatzes und des Satzes 6.32 über die unitäre Normalform sind nahezu identisch. In der Tat gehören selbstadjungierte und unitäre Endomorphismen zu einer größeren Klasse von Endomorphismen, die viele Eigenschaften gemeinsame Eigenschaften besitzen.

Definition 7.23 Es sei V ein Vektorraum mit Skalarprodukt. Ein Endomorphismus Φ von V heißt **normal**, falls gilt:

$$\Phi \circ \Phi^* = \Phi^* \circ \Phi. \quad (7.6)$$

Beispiel 7.24 Die wichtigsten Klassen normaler Endomorphismen sind die folgenden:

- Die *unitären* und *orthogonalen* mit $\Phi^* = \Phi^{-1}$.
- Die *selbstadjungierten* mit $\Phi^* = \Phi$.
- Die *schieferhermiteschen* und *schiefsymmetrischen* mit $\Phi^* = -\Phi$.

Entsprechend gelten die folgenden Resultate für alle diese Klassen.

Hilfssatz 7.25 Es sei V ein Vektorraum mit Skalarprodukt. $\Phi \in \text{End}(V)$ ist genau dann normal, wenn

$$\langle \Phi(x) | \Phi(y) \rangle = \langle \Phi^*(x) | \Phi^*(y) \rangle \quad (7.7)$$

gilt für alle $x, y \in V$.

BEWEIS: Ist Φ normal, so gilt

$$\langle \Phi(x) | \Phi(y) \rangle = \langle x | \Phi^*(\Phi(y)) \rangle = \langle x | \Phi(\Phi^*(y)) \rangle = \langle \Phi^*(x) | \Phi^*(y) \rangle.$$

Umgekehrt folgt aus dieser Gleichheit

$$\begin{aligned} 0 &= \langle \Phi(x) | \Phi(y) \rangle - \langle \Phi^*(x) | \Phi^*(y) \rangle = \langle (\Phi^* \circ \Phi)(x) | y \rangle - \langle (\Phi \circ \Phi^*)(x) | y \rangle \\ &= \langle (\Phi^* \circ \Phi - \Phi \circ \Phi^*)(x) | y \rangle \end{aligned}$$

und da x, y beliebig sind, folgt $\Phi^* \circ \Phi - \Phi \circ \Phi^* = 0$. ■

Hilfssatz 7.26 Es sei V ein Vektorraum mit Skalarprodukt und $\Phi \in \text{End}(V)$ ein normaler Endomorphismus. Dann gilt:

- (a) $\text{Kern } \Phi = \text{Kern } \Phi^*$.
- (b) $E_\lambda(\Phi) = E_{\bar{\lambda}}(\Phi^*)$
- (c) Für alle $\lambda, \mu \in \mathbb{C}$, $\lambda \neq \mu$, ist $E_\lambda(\Phi) \perp E_\mu(\Phi)$.

BEWEIS:

- (a) Aus Hilfssatz 7.25 folgt $\|\Phi(x)\| = \|\Phi^*(x)\|$ für alle $x \in V$, insbesondere für $x \in \text{Kern } \Phi$.
- (b) $E_\lambda(\Phi) = \text{Kern}(\Phi - \lambda \text{id}_V) = \text{Kern}(\Phi - \lambda \text{id}_V)^* = \text{Kern}(\Phi^* - \bar{\lambda} \text{id}_V) = E_{\bar{\lambda}}(\Phi^*)$.

(c) Es sei $x \in E_\lambda(\Phi)$ und $y \in E_\mu(\Phi) = E_{\bar{\mu}}(\Phi^*)$. Dann:

$$\lambda \langle x|y \rangle = \langle \lambda x|y \rangle = \langle \Phi(x)|y \rangle = \langle x|\Phi^*(y) \rangle = \langle x|\bar{\mu}y \rangle = \mu \langle x|y \rangle.$$

Da $\lambda \neq \mu$, gilt $\langle x|y \rangle = 0$. ■

Der folgende Satz ist eine verallgemeinerte Form des Spektralsatzes:

Satz 7.27 Es sei V ein Vektorraum mit Skalarprodukt und $\dim V < \infty$. Ist Φ ein normaler Endomorphismus von V und zerfällt das charakteristische Polynom f_Φ in Linearfaktoren, so existiert eine Orthonormalbasis aus Eigenvektoren von Φ .

Der Beweis ist im wesentlichen identisch mit dem Beweis von Satz 7.18. Die Bedingung an f_Φ ist im unitären Fall immer erfüllt. Eine allgemeinere Normalform⁸⁾ für normale Endomorphismen auch im euklidischen Fall liefert der folgende Satz:

Satz 7.28 Es sei $(V, \langle \cdot | \cdot \rangle)$ ein euklidischer Vektorraum $\dim V = n$. Für einen normalen Endomorphismus Φ existiert eine Orthonormalbasis B von V , so dass die Abbildungsmatrix von Φ bzgl. B die folgende Gestalt hat:

$$\tilde{A} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_s & \\ & & & P_1 \\ 0 & & & & \ddots \\ & & & & & P_r \end{pmatrix} \quad (7.8)$$

wobei $\lambda_1, \dots, \lambda_s$ die reellen Eigenwerte von Φ sind, $s + 2r = n$, und die $P_i \in \mathbb{R}^{2 \times 2}$ sind Matrizen der Gestalt

$$P_i = \begin{pmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{pmatrix},$$

wobei die $\alpha_i + i\beta_i \in \mathbb{C}$ die komplexen Eigenwerte von Φ sind.

Der Beweis ist im wesentlichen identisch mit dem Beweis von Satz 6.37. Der Unterschied besteht darin, dass die Eigenwerte hier nicht den Betrag 1 haben müssen.

Aufgabe 7.29 Es sei $X \in \mathbb{R}^{n \times n}$ eine schiefsymmetrische Matrix (d.h. $X^\top = -X$). Alle Eigenwerte von X sind rein imaginär, und es existiert eine Matrix $S \in \mathbf{O}_n$, so

⁸⁾Wir wollen davon absehen, dies als *normale Normalform* zu bezeichnen.

dass gilt

$$S^\top \cdot X \cdot S = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 0 & -\beta_1 \\ & & & \beta_1 & 0 \\ & & & & \ddots \\ & & & & & 0 & -\beta_r \\ & & & & & \beta_r & 0 \end{pmatrix}.$$

Teil III

Anhang

A Geometrie der komplexen Zahlen

Der Körper \mathbb{C} der **komplexen Zahlen** besteht aus den Zahlen

$$z = x + iy$$

mit $x, y \in \mathbb{R}$ und $i = \sqrt{-1}$.

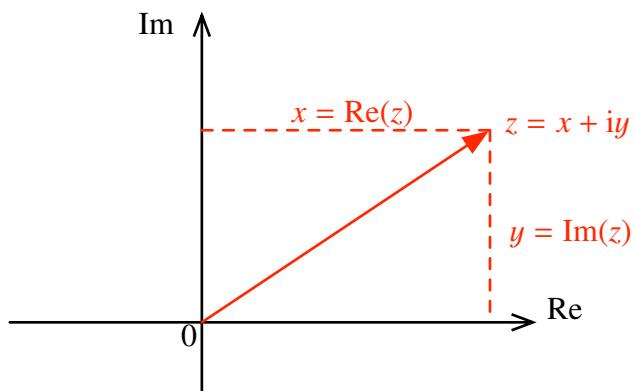
Aufgabe A.1 Schränkt man die Multiplikation in \mathbb{C} auf die reellen Zahlen $\mathbb{R} \subseteq \mathbb{C}$ ein, so wird \mathbb{C} zu einem \mathbb{R} -Vektorraum der Dimension 2.

A.1 Die Gaußsche Zahlenebene

Über die Zuordnung

$$x + iy \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

wird \mathbb{C} mit der Ebene \mathbb{R}^2 identifiziert, die in diesem Kontext als **Gaußsche Zahlenebene** bezeichnet wird.

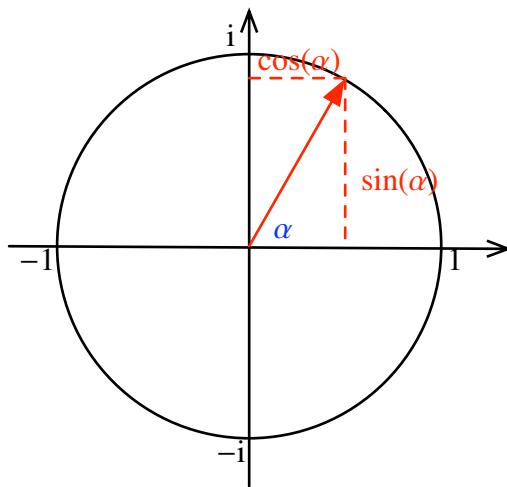


Die x -Achse stellt die reellen Zahlen dar, die y -Achse die rein imaginären Zahlen. Die Zahl 1 entspricht dem ersten Einheitsvektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, die Zahl i dem zweiten Einheitsvektor $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Der Betrag $|z| = \sqrt{\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2}$ einer komplexen Zahl entspricht in dieser Darstellung der euklidischen Länge $\sqrt{x^2 + y^2}$ des Vektors $\begin{pmatrix} x \\ y \end{pmatrix}$.

A.2 Komplexe Multiplikation

Es sei nun $u \in \mathbb{C}$ eine komplexe Zahl mit $|u| = 1$. Dann liegt u auf dem Einheitskreis in der Gaußschen Zahleebene. Der **Polarwinkel** α ist der Winkel zwischen der reellen Achse und der Geraden, die u mit den Nullpunkt verbindet.



Wie die Zeichnung verdeutlicht, hat die Zahl u Real- und Imaginärteil

$$\operatorname{Re}(u) = \cos(\alpha), \quad \operatorname{Im}(u) = \sin(\alpha).$$

Das bedeutet

$$u = \cos(\alpha) + i \sin(\alpha). \quad (\text{A.1})$$

Setzt man für Sinus und Cosinus ihre jeweilige Potenzreihendarstellung ein, so erhält man formal

$$u = e^{i\alpha}. \quad (\text{A.2})$$

In der Tat gelten die selben Rechenregeln wie für die reelle Exponentialfunktion,

$$e^0 = 1, \quad e^{i\alpha} e^{i\beta} = e^{i(\alpha+\beta)}.$$

Wir wollen nun die Multiplikation mit $u = e^{i\alpha}$ geometrisch im \mathbb{R}^2 deuten: Für beliebiges $z = x + iy \in \mathbb{C}$ gilt

$$\begin{aligned} e^{i\alpha} \cdot z &= (\cos(\alpha) + i \sin(\alpha)) \cdot (x + iy) \\ &= (\cos(\alpha)x + i \sin(\alpha)x + i \cos(\alpha)y - \sin(\alpha)y) \\ &= (\cos(\alpha)x - \sin(\alpha)y) + i(\sin(\alpha)x + \cos(\alpha)y). \end{aligned}$$

In den Koordinaten des \mathbb{R}^2 entspricht dies der Abbildung

$$\begin{pmatrix} \cos(\alpha)x - \sin(\alpha)y \\ \sin(\alpha)x + \cos(\alpha)y \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

Die Multiplikation mit $e^{i\alpha}$ entspricht als einer Rotation um den Winkel α gegen den Uhrzeigersinn, die durch die Rotationsmatrix

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

beschrieben wird.

Aufgabe A.2 Die Menge

$$\mathbf{S}^1 = \{u \in \mathbb{C} \mid |u| = 1\}$$

bildet mit der komplexen Multiplikation eine Gruppe, die isomorph ist zur speziellen orthogonalen Gruppe \mathbf{SO}_2 .

Nun sei $w \in \mathbb{C}$ beliebig. Da $w = |w| \cdot \frac{w}{|w|}$ gilt, können wir w schreiben als Produkt einer reellen Zahl $r = |w|$ und einer komplexen Zahl $u = \frac{w}{|w|}$ auf dem Einheitskreis. Wegen (A.2) können wir schreiben:

$$w = r \cdot e^{i\alpha}. \quad (\text{A.3})$$

Dies ist die Darstellung von w in **Polarcoordinaten** (r, α) . Die Multiplikation von $z = x + iy \in \mathbb{C}$ mit $r \in \mathbb{R}$ liefert

$$r \cdot z = rx + i ry,$$

also $|r \cdot z| = r \cdot |z|$. Dies ist entspricht der Streckung von $\begin{pmatrix} x \\ y \end{pmatrix}$ um den Faktor r ,

$$\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} rx \\ ry \end{pmatrix}.$$

Die Multiplikation mit einer beliebigen Zahl $w = r \cdot e^{i\alpha} \in \mathbb{C}$ ist also die Kombination einer Drehung um den Winkel α und einer Streckung um den Faktor $r = |w|$:

$$\begin{pmatrix} r \cos(\alpha)x - r \sin(\alpha)y \\ r \sin(\alpha)x + r \cos(\alpha)y \end{pmatrix} = \underbrace{\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}}_{\text{Streckung}} \cdot \underbrace{\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}}_{\text{Drehung}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

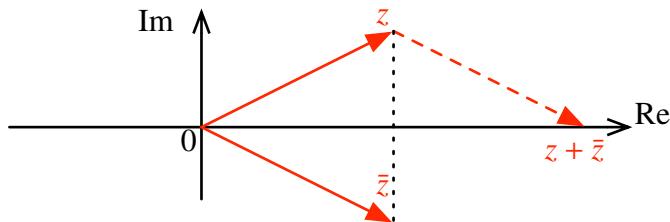
A.3 Komplexe Konjugation

Die **komplexe Konjugation** ist die Abbildung

$$\mathbb{C} \rightarrow \mathbb{C}, \quad z = x + iy \mapsto \bar{z} = x - iy.$$

Dies entspricht der Spiegelung an der reellen Achse in der Gaußschen Zahlen-ebene:

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



Wie man der Zeichnung entnimmt, ist $z + \bar{z} \in \mathbb{R}$ und $\frac{1}{2}(z + \bar{z}) = \operatorname{Re}(z)$ ist die Projektion von z auf die reelle Achse. Ähnlich sieht man, dass $\frac{1}{2}(z - \bar{z}) = i \cdot \operatorname{Im}(z)$ die Projektion von z auf die imaginäre Achse ist.

Der Betrag $|z|$ kann durch die komplexe Konjugation ausgedrückt werden:

$$z \cdot \bar{z} = (x + iy) \cdot (x - iy) = x^2 + y^2 = |z|^2.$$

Dies ergibt eine elegante Darstellung für das Inverse von z :

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

A.4 Einheitswurzeln

Es sei $n \in \mathbb{N}$ und

$$\Omega_n = \{\omega \in \mathbb{C} \mid \omega^n = 1\}. \tag{A.4}$$

Die Elemente von Ω_n heißen **nte Einheitswurzeln**. Sie sind die Nullstellen des Polynoms $X^n - 1$.

Aufgabe A.3 Ω_n ist eine endliche Untergruppe von S^1 . Insbesondere ist $|\omega| = 1$ für alle $\omega \in \Omega_n$. Wie im Beweis von Satz 3.10 zeigt man, dass Ω_n zyklisch ist. Es folgt $|\Omega_n| = n$.

Ein Erzeuger von Ω_n heißt **primitive nte Einheitswurzel**.

Beispiel A.4 (Einheitswurzeln)

- (a) $\Omega_2 = \{1, -1\}$. Eine primitive Einheitswurzel ist -1 .
- (b) $\Omega_4 = \{1, i, -1, -i\}$. Die primitiven Einheitswurzeln sind i und $-i$.

Jedes Element $\omega \in \Omega_n$ hat nach (A.2) die Form $\omega = e^{i\alpha}$. Da

$$1 = \omega^n = e^{ian}$$

gilt, muss wegen (A.1) für ein geeignetes $k \in \mathbb{N}$ gelten:

$$\alpha = 2\pi \frac{k}{n}.$$

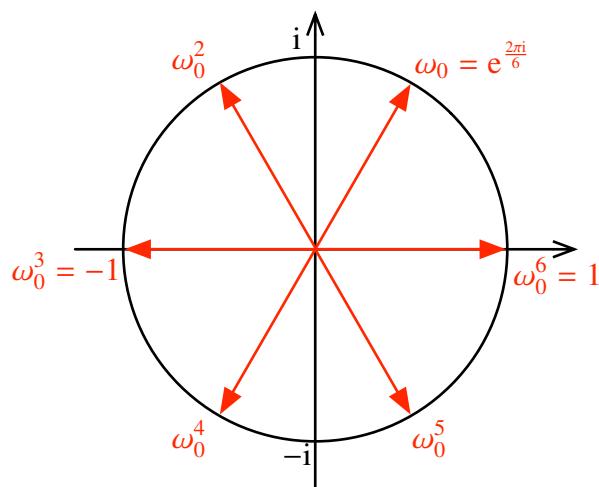
Insbesondere ist

$$\omega_0 = e^{\frac{2\pi i}{n}}$$

für alle n eine primitive n te Einheitswurzel und

$$\Omega_n = \langle \omega_0 \rangle = \{1, e^{\frac{2\pi i}{n}}, e^{\frac{4\pi i}{n}}, \dots, e^{\frac{2\pi i(n-1)}{n}}\}.$$

Die Zeichnung zeigt Ω_6 .



Man sieht hier auch die Gruppe $\Omega_3 \subset \Omega_6$, erzeugt von ω_0^2 .

Aufgabe A.5 Ist ω_0 eine primitive $2m$ te Einheitswurzel, so ist ω_0^2 eine primitive m te Einheitswurzel.

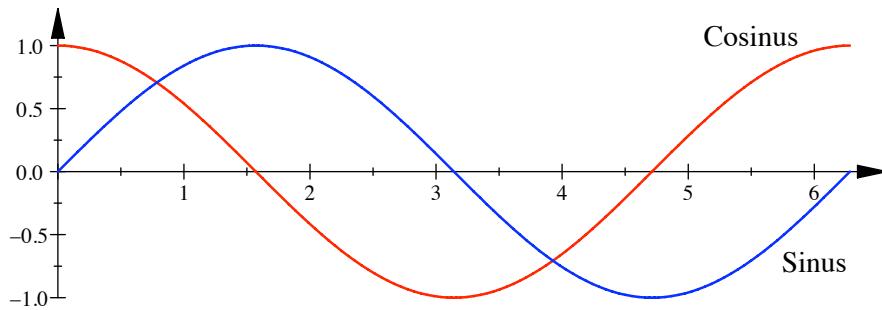
Aufgabe A.6 Mit der Formel für die endliche geometrische Reihe folgt für eine primitive n te Einheitswurzel ω_0 und $0 < k < n$:

$$0 = 1 + \omega_0^k + \omega_0^{2k} + \dots + \omega_0^{(n-1)k}.$$

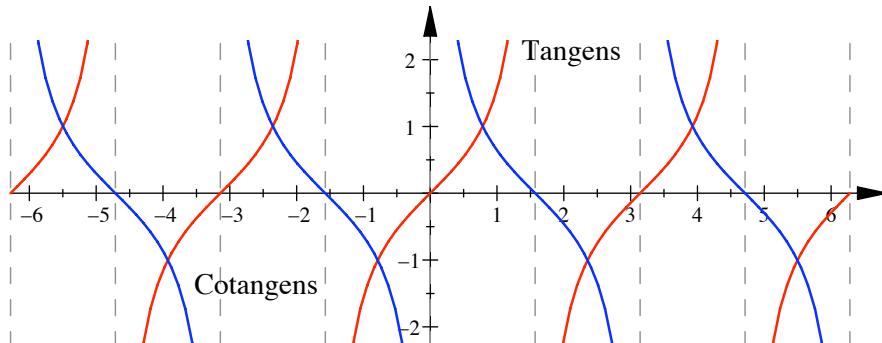
B Trigonometrische Funktionen

B.1 Funktionsgraphen

Die Graphen von Sinus und Cosinus auf dem Intervall $[0, 2\pi]$:



Die Graphen von Tangens und Cotangens auf dem Intervall $[-2\pi, 2\pi]$:



B.2 Rechenregeln

Satz des Pythagoras:

$$\sin(x)^2 + \cos(x)^2 = 1$$

Tangens und Cotangens:

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}$$

Additionstheoreme:

$$\sin(x \pm y) = \sin(x)\cos(y) \pm \cos(x)\sin(y)$$

$$\cos(x \pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y)$$

$$\tan(x \pm y) = \frac{\tan(x) \pm \tan(y)}{1 \mp \tan(x)\tan(y)}$$

$$\cot(x \pm y) = \frac{-1 \pm \cot(x)\cot(y)}{\cot(x) \pm \cot(y)}$$

$$\sin(2x) = 2\sin(x)\cos(x)$$

$$\cos(2x) = \cos(x)^2 - \sin(x)^2$$

$$= 2\cos(x)^2 - 1$$

$$= 1 - 2\sin(x)^2$$

$$\tan(2x) = \frac{2\tan(x)}{1 - \tan(x)^2}$$

$$\cot(2x) = \frac{\cot(x)^2 - 1}{2\cot(x)}$$

$$\sin(x) + \sin(y) = 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$$

$$\sin(x) - \sin(y) = 2\cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$$

$$\cos(x) + \cos(y) = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$$

$$\cos(x) - \cos(y) = -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$$

$$\sin(x)\sin(y) = \frac{\sin(x+y) - \sin(x-y)}{2}$$

$$\sin(x)\cos(y) = \frac{\sin(x+y) + \sin(x-y)}{2}$$

$$\cos(x)\cos(y) = \frac{\cos(x+y) + \cos(x-y)}{2}$$

Periodizität:

$$\sin(-x) = -\sin(x)$$

$$\sin\left(x \pm \frac{\pi}{2}\right) = \pm \cos(x)$$

$$\sin(x \pm \pi) = -\sin(x)$$

$$\cos(-x) = \cos(x)$$

$$\cos\left(x \pm \frac{\pi}{2}\right) = \mp \sin(x)$$

$$\cos(x \pm \pi) = -\cos(x)$$

Komplex:

$$\begin{aligned} e^{ix} &= \cos(x) + i \sin(x) \\ \sin(x) &= \frac{e^{ix} - e^{-ix}}{2i} \\ \cos(x) &= \frac{e^{ix} + e^{-ix}}{2} \end{aligned}$$

B.3 Wertetabelle

α	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin(\alpha)$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
$\cos(\alpha)$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0	-1	0	1
$\tan(\alpha)$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$		0		0

Literatur

- [1] S. BOSCH
Algebra (7. Aufl.)
Springer, 2009
- [2] E. BRIESKORN
Lineare Algebra I & II
Vieweg, 1983 & 1985
- [3] G. FISCHER
Analytische Geometrie (7. Aufl.)
Vieweg, 2001
- [4] G. FISCHER
Lineare Algebra (17. Aufl.)
Vieweg, 2010
- [5] O. FORSTER
Algorithmische Zahlentheorie
Vieweg, 1996
- [6] C. GASQUET, P. WITOMSKI
Fourier Analysis and Applications
Springer, 1999
- [7] B.C. HALL
Lie Groups, Lie Algebras, and Representations
Springer, 2003
- [8] D.E. KNUTH
The Art of Computer Programming 2: Seminumerical Algorithms (3. Aufl.)
Addison-Wesley Longman, 1997
- [9] T.W. KÖRNER
Fourier Analysis
Cambridge University Press, 1988
- [10] E. LEUZINGER
Lineare Algebra
Vorlesungsskript KIT, 2011
- [11] J.D. LIPSON
Elements of Algebra and Algebraic Computing
Benjamin Cummings, 1981

- [12] H.P. REHM, W. TRINKS
Lineare Algebra und analytische Geometrie
Vorlesungsskript Universität Karlsruhe (TH), 2000
- [13] H.R. SCHWARZ
Numerische Mathematik
Teubner, 1997
- [14] D. WERNER
Funktionalanalysis (7. Aufl.)
Springer, 2011
- [15] H. WEYL
Symmetry
Princeton University Press, 1952
- [16] H. ZIESCHANG
Lineare Algebra und Geometrie
Teubner, 1997

Index

- $a \div b$ (ganzzahlige Division), 9
- $A \oplus B$ (direkte Summe von Matrizen), 35
- A^*, A^\dagger (adjungierte Matrix), 138
- $\text{Aut}(V)$ (Automorphismen), vii
- $\text{Bil}(V)$ (Bilinearformen), 68
- $C(M)$ (auf M stetige Funktionen), 69
- $d(x, y)$ (Abstand), 77
- $\deg(f)$ (Grad von f), vii
- δ_{ij} (Kronecker-Symbol), vii
- E_λ (Eigenraum zum Eigenwert λ), vii
- $\text{End}(V)$ (Endomorphismen), vii
- \mathbb{F}_{p^n} (endlicher Körper), 23
- $\text{GL}_n(\mathbb{K})$ (invertierbare Matrizen), vii
- $\text{Hom}(V, W)$ (Homomorphismen), vii
- I_n (Einheitsmatrix), vii
- $\text{Iso}(V, \langle \cdot, \cdot \rangle)$ (Isometrien), 115
- $J_n(\lambda)$ (Jordan-Kästchen), 40
- $\hat{J}_{(n_1, \dots, n_k)}(\lambda)$ (Jordan-Block), 41
- ℓ^2 (Hilbert-Raum), 101
- $M_B^C(\Phi)$ (Abbildungsmatrix von Φ), vii
- Ω_n (nte Einheitswurzeln), 150
- O_n (orthogonale Gruppe), 119
- Φ^* (duale Abbildung zu Φ), vii
- Φ^* (adjungierte Abbildung zu Φ), 137
- $\Phi \oplus \Psi$ (direkte Summe), 35
- Π_U (Orthogonalprojektion), 96
- R^\times (Einheitengruppe), 2
- R_α (Rotationsmatrix), 116
- Rad s (Radikal einer Bilinearform), 71
- $R^{m \times n}$ (Matrizen mit Einträgen aus R), vii
- SO_n (spezielle orthogonale Gruppe), 119
- Σ_U (Spiegelung an U), 115
- SU_n (spezielle unitäre Gruppe), 119
- $S_B(s)$ (Matrix einer Bilinearform s), 69
- Spec Φ (Spektrum von Φ), vii
- $\text{Sym}(V)$ (symmetrische Bilinearformen), 68
- $\Theta_B(x)$ (Koordinatendarstellung), vii
- U^\perp (orthogonales Komplement von U), 88
- U_n (unitäre Gruppe), 119
- V^* (Dualraum zu V), vii
- $x \perp y$ (orthogonal), 80
- $x \times y$ (Vektorprodukt), 110
- $\langle x|y \rangle$ (Skalarprodukt), 72
- $\measuredangle(x, y)$ (Winkel), 80
- $\langle x \rangle$ (von x erzeugtes Ideal), 5
- $x | y$ (x teilt y), 4
- Abstand, 77
- adjungierte Abbildung, 137
- adjungierte Matrix, 138
- Algorithmus
 - Diffie-Hellman-Protokoll, 61
 - El Gamal, 61
 - erweiterter euklidischer, 11
 - euklidisch, 11
 - euklidische Normalform, 130
 - Jordan-Basis, 53
 - Lagrange-Interpolation, 27
 - Newton-Interpolation, 26
 - RSA, 62
- Alice, 57
- alternierende Bilinearform, 68
- ausgeartete Bilinearform, 71
- Bézout, Lemma von, 11
- Besselsche Ungleichung, 104
- bilinear, 66, 68
- Bilinearform, 68
 - alternierend, 68
 - ausgeartet, 71
 - Kern, 71
 - Radikal, 71
 - schiefsymmetrisch, 68
 - symmetrisch, 68
- Bob, 57
- Bra-Ket-Notation, 73
- Carmichael-Zahl, 60
- Cartan-Dieudonné, Satz von, 133
- Cauchy-Folge, 100
- Cauchy-Schwarz-Ungleichung, 76, 84
- Chiffrat, 57
- chinesischer Restsatz, 25
- Cholesky-Verfahren, 107
- Cholesky-Zerlegung, 106
- coprim (siehe teilerfremd), 4
- Cosinussatz, 79, 82
- Darstellung, 120
- Diffie-Hellman-Protokoll, 61
- direkte Summe
 - von Endomorphismen, 35
- Drehachse, 117

- Drehebene, 117
Drehung, 116
Dreiecksungleichung, 75
- Eigenraum
 verallgemeinert, 45
eindeutige Primfaktorzerlegung, 15
- Einheit, 2
Einheitengruppe, 2
Einheitssphäre, 78
Einheitsvektor, 78
Einheitswurzel, 150
 primitiv, 151
El Gamal-Verschlüsselung, 61
endlicher Körper, 23
Erzeuger, 48
Euklid, Primzahlsatz von, 16
euklidische Normalform, 129
euklidischer Algorithmus, 11
 erweitert, 11
euklidischer Ring, 9
euklidischer Vektorraum, 72
Euler-Winkel, 133
Eulersche φ -Funktion, 59
- Fibonacci-Zahl, 13
Fourier-Koeffizienten, 104
Fourier-Reihe, 104
Fréchet-Riesz, Satz von, 138
Fundamentalsatz der Algebra, 17
Funktionalanalysis, 100
- Gaußsche Zahlenebene, 147
geheimer Schlüssel, 57
größter gemeinsamer Teiler, 4
Gram-Schmidt-Verfahren, 93
Gramsche Determinante, 99
Gruppe
 Einheiten-, 2
 Isometrie, 115
 orthogonal, 115, 119
 speziell orthogonal, 119
 speziell unitär, 119
 unitär, 115, 119
- Hauptideal, 5
Hauptraum, 45
hermitesche Form, 83
hermitesche Matrix, 84
Hesse-Matrix, 109
- Hilbert-Raum, 101
 separabel, 101
Hurwitz-Kriterium, 106
- Ideal, 5
 Haupt-, 5
integrer Ring, 4
invariante Unterraum, 33
invariantes Komplement, 33
irreduzibel, 8
Isometrie, 113
 linear, 115
Isometriegruppe, 115
- Jacobi-Identität, 111
Jordan-Basis, 41, 50, 52
Jordan-Block, 41
Jordan-Kästchen, 40
Jordansche Normalform, 1
 einer Matrix, 42
 eines Endomorphismus, 40
- Kürzungsregel, 5
Klartext, 57
Komplementärraum, 33
 invariant, 33
komplexe Konjugation, 150
komplexe Zahl, 147
Kongruenzen, 24
Konjugation, 150
Konvergenz, 100
Kreuzprodukt (siehe Vektorprodukt), 110
Kryptographie, 57
- Lagrange, Satz von, 58
Lagrange-Interpolation, 27
Lemma von Bézout, 11
Lemma von Schur, 36
Lie-Algebra, 111
lineare Isometrie, 115
Lot, 98
Lotfußpunkt, 98
Lotvektor, 98
- Maximumsnorm, 76
Minimalpolynom, 6, 38
- Newton-Interpolation, 26
nilpotent
 Endomorphismus, 36

- Matrix, 36
- Norm, 75
 - Maximums-, 76
- normaler Endomorphismus, 143
- Normalform, 40
 - euklidische, 129
 - Jordansche, 1
 - unitäre, 126
- normierter Vektor, 78
- öffentlicher Schlüssel, 57
- Ordnung, 58
- orthogonal, 80
- Orthogonalbasis, 88
- orthogonale Gruppe, 115, 119
- orthogonale Matrix, 119
- orthogonales Komplement, 88
- Orthogonalisierungsverfahren, 93
- Orthogonalprojektion, 96
- Orthonormalbasis, 88
 - im Sinne der Funktionalanalysis, 100
- Orthonormalsystem, 100
 - vollständig, 100
- Parallelisierung, 23
- Parallelogrammgleichung, 76
- Parsevalsche Gleichung, 104
 - verallgemeinert, 104
- Polarisierung, 76
- Polarkoordinaten, 149
- Polarwinkel, 148
- Polynomdivision, 10
- positiv definit, 67
 - Bilinearform, 72
 - Hurwitz-Kriterium, 106
 - Matrix, 73
 - Sesquilinearform, 83
- prim, 8
- Primärzerlegung, 44
- Primfaktorzerlegung, 15
- Primideal, 9
- Primzahlsatz von Euklid, 16
- Public Key, 57
- Pythagoras, Satz von, 66, 81, 85
- Quasiordnung, 5
- Quotientenring, 19
- Radikal, 71
- Relationen, 18
- Ring
 - euklidisch, 9
 - integer, 4
- Roll-Nick-Gier-Winkel, 136
- Rotation, 116
- RSA-Verschlüsselung, 62
- Satz von Cartan-Dieudonné, 133
- Satz von Fréchet-Riesz, 138
- Satz von Lagrange, 58
- Satz von Pythagoras, 66, 81, 85
- schiefsymmetrische Bilinearform, 68
- schiefsymmetrische Matrix, 70
- Schlüssel, 57
 - geheim, 57
 - öffentlich, 57
- Schlüsselaustausch, 60
- Schur
 - Lemma von, 36
- selbstadjungiert, 140
- senkrecht (siehe orthogonal), 80
- separabler Hilbert-Raum, 101
- Sesquilinearform, 83
- simultane Kongruenzen, 24
- Skalarprodukt, 65, 72
 - unitär, 83
- Spektraldarstellung, 143
- Spektralsatz, 142
- spezielle orthogonale Gruppe, 119
- spezielle unitäre Gruppe, 119
- Sphäre, 78
- Spiegelung, 115
- Standardskalarprodukt, 65
 - unitär, 83
- Stufe, Nilpotenz, 36
- symmetrische Bilinearform, 68
- symmetrische Matrix, 70
- Teiler, 4
 - größter gemeinsamer, 4
- teilerfremd, 4
- Translation, 113
- unitäre Gruppe, 115, 119
- unitäre Matrix, 119
- unitäre Normalform, 126
- unitärer Vektorraum, 83
- Vektorprodukt, 110
- Vektorraum mit Skalarprodukt, 85

verallgemeinerter Eigenraum, 45

Verschlüsselung

El Gamal, 61

RSA, 62

vollständiges Orthonormalsystem, 100

Winkel, 80