Lineare Algebra I und II für Physiker

Prof. Dr. Klaus Keimel

Wintersemester 2003/2004
Sommersemester 2004
Version 7. April 2006

Dieses Skript basiert auf einem englischen Skript von Prof. Dr. Klaus Keimel, welches von Prof. Dr. Alexander Martin für eine Vorlesung in deutscher Sprache aufbereitet woden ist.

Technische Universität Darmstadt
Fachbereich Mathematik

Inhaltsverzeichnis

Kapitel 1. Einführung in die Lineare Algebra 1
1.1. Vektoren im 2- und 3-dimensionalen Raum 2
1.2. Der Begriff des reellen Vektorraums 7
1.3. Lineare Gleichungssysteme 12
1.4. Matrizen 21

Kapitel 2. Vektorräume 33
2.1. Algebraische Strukturen 33
2.2. Der Körper der komplexen Zahlen 38
2.3. Reelle und komplexe Vektorräume 48
2.4. Lineare Teilräume, affine Teilräume 50
2.5. Linearkombinationen, lineare Unabhängigkeit, Basen 54
2.6. Dimension und Sätze über Basen 61

Kapitel 3. Euklidische und unitäre Vektorräume 65
3.1. Das Skalarprodukt für Vektoren in Ebene und Raum 65
3.2. Der Euklidische Vektorraum $\mathbb{R}^n$ 68
3.3. Der unitäre Vektorraum $\mathbb{C}^n$ 72
3.4. Euklidische und unitäre Vektorräume 75
3.5. Orthogonalprojektionen 78
3.6. Geraden und Ebenenen im $\mathbb{R}^2$ und $\mathbb{R}^3$ 85
3.7. Ein Hauch von Quantenmechanik 88

Kapitel 4. Lineare Abbildungen und Matrizen 94
4.1. Reelle und komplexe Matrizen 94
4.2. Lineare Abbildungen 98
4.3. Lineare Abbildungen $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ und Matrizen 104
4.4. Die Matrix einer linearen Abbildung bzgl. beliebiger Basen 109
4.5. Operieren mit linearen Abbildungen 113
4.6. Basiswechsel 117
4.7. Orthogonale bzw. unitäre Abbildungen und Matrizen 123
4.8. Systeme linearer Gleichungen 128
4.9. Die Spur 132

Kapitel 5. Determinanten 135
5.1. Das Vektorprodukt im Euklidischen $\mathbb{R}^3$ 135
5.2. Die Determinante als „orientierte Fäche“ in $\mathbb{R}^2$ 139
5.3. Die Determinante als „orientiertes Volumen“ in $\mathbb{R}^3$ 141
5.4. Die Determinante einer $n \times n$-Matrix 145
5.5. Die Determinante eines Endomorphismus 155
5.6. Permutationen 156

5.7. Beweis des Hauptsatzes 5.4.2 159

Kapitel 6. Eigenvektoren und Eigenwerte 163
6.1. Eigenwerte und Eigenvektoren von Endomorphismen 163
6.2. Eigenwerte und Eigenvektoren von Matrizen 167
6.3. Ähnlichkeit von Matrizen und Diagonalisierbarkeit 171
6.4. Determinante und Eigenwerte von unitären und orthogonalen Matrizen 175

Kapitel 7. Diagonalisierung: Hauptsätze 179
7.1. Normale Matrizen 179
7.2. Diagonalisierungssätze 182
7.3. Die Adjungierte eines Endomorphismus 187

Kapitel 8. Linearformen, Bilinearformen, Quadratische Formen 190
8.1. Bilinearformen 190
8.2. Quadratische Formen 193
8.3. Hauptachsentransformation 194
8.4. Definitheit von Formen und von Matrizen 204
8.5. Linearformen und Dualraum, Tensoren 209

Kapitel 9. Die Jordansche Normalform 218
9.1. Matrizen in Jordanscher Normalform 218
9.2. Verallgemeinerte Eigenvektoren 220
9.3. Nilptente Endomorphismen bzw. Matrizen 223
9.4. Bestimmung einer Jordanbasis im allgemeinen Fall 227
9.5. Anwendungen 229

Kapitel 10. Gruppen, Symmetrien, Symmetriegruppen 231
10.1. Gruppen 231
10.2. Untergruppen, Homomorphismen, direkte Produkte 232
10.3. Symmetriegruppen 235

KAPITEL 1

Einführung in die Lineare Algebra

In diesem einführenden Kapitel behandeln wir zwei wesentliche Grundlagen der Linearen Algebra: Vektoralgebra in Dimension zwei und drei sowie Gleichungssysteme und Matrizen.

Wir verwenden folgende Notation:

$\mathbb{N}$	die Menge der natürlichen Zahlen	$0, 1, 2, \ldots$
$\mathbb{Z}$	die Menge der ganzen Zahlen	$\ldots, -2, -1, 0, 1, 2, \ldots$
$\mathbb{Q}$	die Menge der rationalen Zahlen	$r = \frac{m}{n}$ mit $m \in \mathbb{Z}$, $0 \neq n \in \mathbb{N}$
$\mathbb{R}$	die Menge der reellen Zahlen	zum Beispiel $\sqrt{2}$, π, e
$\mathbb{C}$	die Menge der komplexen Zahlen	zum Beispiel $2 + 3i$.

Wir gehen davon aus, dass diese Zahlensysteme und ihre Regeln für Addition, Subtraktion, Multiplikation und Division bekannt sind. In der Analysis wird mehr zu den reellen Zahlen gesagt werden. Auf die komplexen Zahlen werden wir allerdings erst später zu sprechen kommen.

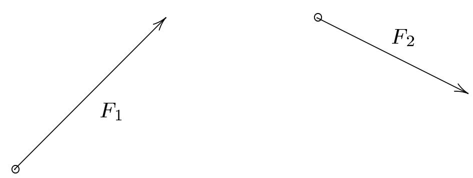
Zur Erinnerung einige griechische Buchstaben, die wir verwenden wollen:

α		alpha
β		beta
γ	Γ	gamma
δ	Δ	delta
ε		epsilon
κ		kappa
λ	Λ	lambda
μ		mü
ν		nü
π	Π	pi
φ	Φ	phi
ψ	Ψ	psi
ρ		rho
σ	Σ	sigma
χ		chi
η		eta
ξ		xi
ζ		zeta
ω	Ω	omega

1.1. Vektoren im 2- und 3-dimensionalen Raum

In diesem Kapitel bauen wir auf unsere Intuition und einige elementare geometrische Grundkenntnisse aus der Schule. In den späteren Kapiteln werden wir nicht auf diese Intuition setzen und die Theorie systematisch aufbauen.

1.1.1 Motivation aus der Physik. Die Vektorrechnung hat einen ihrer Ursprünge in der Physik. Dort gibt es Größen wie Temperatur, elektrische Ladung oder Energie, die in reellen Zahlen gemessen werden. Sie werden skalare Größen genannt. Andere Größen wie Kraft, Geschwindigkeit oder Drehmoment haben nicht nur einen absoluten Wert, sondern auch eine Richtung. Sie werden vektorielle Größen genannt und durch Pfeile dargestellt.



1.1.2 Skalare und Vektoren. Reelle Zahlen werden wir auch als *Skalare* [*scalars*] bezeichnen. Oft verwenden wir griechische Buchstaben

$$\alpha, \beta, \gamma, \ldots, \omega,$$

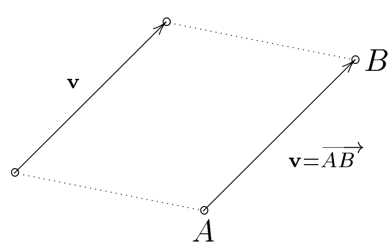
um Skalare zu bezeichnen oder manchmal auch lateinische Buchstaben

$$a, b, c, \ldots, x, y, z.$$

Ein *Vektor* [*vector*] wird durch einen Pfeil dargestellt, genauer gesagt durch eine gerichtete Strecke. Zwei gerichtete Strecken repräsentieren denselben Vektor, falls sie dieselbe Länge und dieselbe Richtung haben, d.h. falls die eine in die andere durch eine Parallelverschiebung überführt werden kann. Vektoren werden wir mit fettgedruckten lateinischen Buchstaben bezeichnen.

$$\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots.$$

(Häufig wird auch die Schreibweise $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}, \ldots$ verwendet. Wir werden späer einfach $u, v, w, \ldots$ schreiben.) Mit $\overrightarrow{AB}$ bezeichnen wir den Vektor, der durch die gerichtete Strecke mit Anfangspunkt A und Endpunkt B dargestellt wird.



Bemerkung. Es macht keinen Unterschied, ob wir uns Vektoren in der Ebene oder im 3-dimensionalen Raum vorstellen. Die Regeln der Vektoralgebra sind unabhängig von der Dimension.

1.1.3 Addition von Vektoren und Multiplikation mit Skalaren.

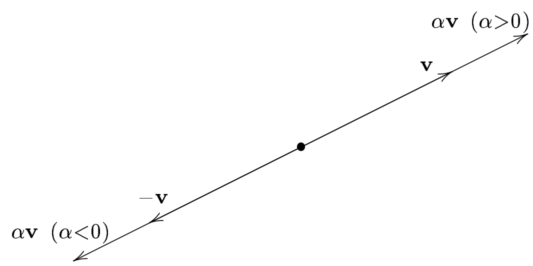
Nullvektor: $\mathbf{0}$ bezeichnet den Nullvektor, d.h. den Vektor mit Länge 0. Dieser Vektor hat keine Richtung.

Skalarmultiplikation: Sei $\mathbf{v}$ ein Vektor. Für einen Skalar α definieren wir das skalare Vielfache $\alpha\mathbf{v}$ wie folgt:

Falls $\alpha > 0$: $\alpha\mathbf{v}$ hat dieselbe Richtung wie $\mathbf{v}$ und seine Länge ist α-mal die Länge von $\mathbf{v}$.

Falls $\alpha = 0$: $0\mathbf{v} := \mathbf{0}$.

Falls $\alpha < 0$: Die Richtung von $\alpha\mathbf{v}$ ist entgegen der von $\mathbf{v}$ und seine Länge ist $|\alpha|$-mal der Länge von $\mathbf{v}$.

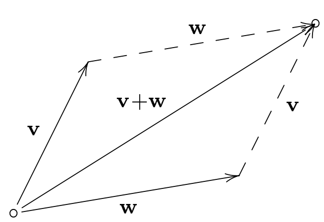


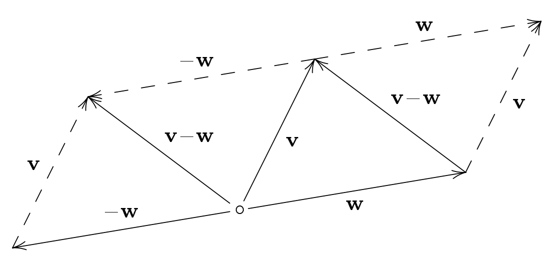
Anstelle von $(-1)\mathbf{v}$ schreiben wir einfach $-\mathbf{v}$ für den in die Gegenrichtung umgeklappten Vektor $\mathbf{v}$.

Addition: Die Summe $\mathbf{v}+\mathbf{w}$ zweier Vektoren $\mathbf{v}$ und $\mathbf{w}$ ist geometrisch wie folgt definiert: Verschiebe $\mathbf{w}$ parallel derart, dass sein Anfangspunkt mit dem Endpunkt von $\mathbf{v}$ übereinstimmt; dann ist $\mathbf{v}+\mathbf{w}$ der Vektor, dessen Anfangspunkt mit dem Anfangspunkt von $\mathbf{v}$ und dessen Endpunkt mit dem Endpunkt von $\mathbf{w}$ übereinstimmt.

Subtraktion: Die Subtraktion eines Vektors ist definiert als die Addition des Gegenvektors:

$$\mathbf{v}-\mathbf{w} := \mathbf{v}+(-\mathbf{w}).$$





1.1.4 Grundregeln der Vektoralgebra.

(V1)	$\mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w}$	für beliebige Vektoren $\mathbf{u},\mathbf{v},\mathbf{w}$
(V2)	$\mathbf{v}+\mathbf{0} = \mathbf{v} = \mathbf{0}+\mathbf{v}$	für beliebige Vektoren $\mathbf{v}$
(V3)	$\mathbf{v}+(-\mathbf{v}) = \mathbf{0} = (-\mathbf{v})+\mathbf{v}$	für beliebige Vektoren $\mathbf{v}$
(V4)	$\mathbf{v}+\mathbf{w} = \mathbf{w}+\mathbf{v}$	für beliebige Vektoren $\mathbf{v},\mathbf{w}$
(V5)	$\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$	für beliebige Skalare λ,μ und Vektoren $\mathbf{v}$
(V6)	$1\mathbf{v} = \mathbf{v}$	für beliebige Vektoren $\mathbf{v}$
(V7)	$\lambda(\mathbf{v}+\mathbf{w}) = \lambda\mathbf{v}+\lambda\mathbf{w}$	für beliebige Skalare λ und Vektoren $\mathbf{v},\mathbf{w}$

(V8) $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ für beliebige Skalare λ, μ und Vektoren $\mathbf{v}$

Die Gültigkeit dieser Beziehungen kann man einfach mit geometrischen Argumenten überprüfen. Mit Bedingung (V1) können wir Klammern weglassen, wenn wir mehrere Vektoren addieren; Bedingung (V5) sagt dasselbe für Skalare, wenn wir einen Vektor nacheinander mit mehreren Skalaren multiplizieren. Regel (V4) impliziert, dass die Anordnung der Summanden beim Addieren mehrerer Vektoren keine Rolle spielt.

1.1.5 Abgeleitete Regeln. Für einen beliebigen Skalar λ und beliebige Vektoren $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ gilt:

(V9) $$\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w} \quad \text{impliziert} \quad \mathbf{u} = \mathbf{v}$$

(V10) $$0\mathbf{v} = \mathbf{0} \quad \text{und} \quad \lambda\mathbf{0} = \mathbf{0}$$

(V11) $$\lambda\mathbf{v} = \mathbf{0} \quad \text{impliziert} \quad \lambda = 0 \text{ oder } \mathbf{v} = \mathbf{0}$$

Diese Regeln sind geometrisch leicht einsichtig. Wir haben diese Regeln nicht zu unseren Grundregeln hinzugenommen, da man sie leicht aus ihnen ableiten kann. Wenn wir das machen, achten wir streng darauf, dass wir bei der Herleitung von Regeln nur Grundregeln oder bereits gezeigte Regeln verwenden. Diese Vorgehensweise hat den Vorteil, dass wir sie auch in komplizierten Situationen anwenden können, wenn uns unsere geometrische Intuition verlässt.

Beweis von (V9). Sei $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$. Addieren wir $-\mathbf{w}$ auf beiden Seiten, so erhalten wir $(\mathbf{u} + \mathbf{w}) + (-\mathbf{w}) = (\mathbf{v} + \mathbf{w}) + (-\mathbf{w})$. Für die linke Seite gilt

$$(\mathbf{u} + \mathbf{w}) + (-\mathbf{w}) = \mathbf{u} + (\mathbf{w} + (-\mathbf{w})) = \mathbf{u} + \mathbf{0} = \mathbf{u}.$$

Eine analoge Rechnung zeigt, dass die rechte Seite gleich $\mathbf{v}$ ist. Folglich gilt $\mathbf{u} = \mathbf{v}$. □

Beweis von (V10). Es gilt $0\mathbf{v} = (1 + (-1))\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ nach (V8), (V6) und (V3). Der Beweis des zweiten Teils von (V10) ist ähnlich und wird in der Übung behandelt. □

Beweis von (V11). Nehmen wir an, dass $\lambda\mathbf{v} = \mathbf{0}$ gilt. Falls $\lambda = 0$, gibt es nichts zu zeigen. Nehmen wir also an, dass $\lambda \neq 0$. Dann betrachten wir das Inverse λ^{-1} und multiplizieren die Gleichung mit diesem Skalar. Wir erhalten: $\lambda^{-1}(\lambda\mathbf{v}) = \lambda^{-1}\mathbf{0}$. Für die linke Seite erhalten wir $\lambda^{-1}(\lambda\mathbf{v}) = (\lambda^{-1}\lambda)\mathbf{v} = 1\mathbf{v} = \mathbf{v}$ und für die rechte Seite $\mathbf{0}$ nach (V10). Also gilt $\mathbf{v} = \mathbf{0}$. □

1.1.6 Übung. Leite die folgenden Regeln für beliebige Vektoren $\mathbf{v}$, $\mathbf{w}$ und beliebige Skalare λ, μ ab:

(V12) $$\lambda(\mathbf{v} - \mathbf{w}) = \lambda\mathbf{v} - \lambda\mathbf{w}$$

(V13) $$(\lambda - \mu)\mathbf{v} = \lambda\mathbf{v} - \mu\mathbf{v}$$

1.1.7 Darstellung geometrischer Objekte. Zunächst wählen wir einen Referenzpunkt O, auch Ursprung genannt. Dieser Punkt sei für den Augenblick fest vorgegeben.

Punkte: Ein Punkt P kann beschrieben werden durch den Vektor $\mathbf{x} = \overrightarrow{OP}$. P ist eindeutig durch $\mathbf{x}$ bestimmt (solange wir von dem oben gewählten festen Referenzpunkt ausgehen).

Geraden: Sei g eine Gerade.

- Wähle einen festen Punkt P_0 auf g, sei $\mathbf{a} = \overrightarrow{OP_0}$ und
- sei $\mathbf{v}$ irgendein Vektor ungleich Null, der parallel zu g ist.

Dann liegt ein Punkt P auf der Geraden g genau dann, wenn der Vektor $\mathbf{x} = \overrightarrow{OP}$ eine Darstellung der Form

$$\mathbf{x} = \mathbf{a} + \lambda \mathbf{v} \quad (\lambda \in \mathbb{R})$$

hat. Diese Darstellung ist gültig für Geraden in der Ebene wie auch für Geraden im 3-Dimensionalen. Sie wird als Parameterdarstellung von g bezeichnet.

Ebenen: Sei E eine Ebene im 3-dimensionalen Raum. Wähle

- einen Punkt P_0 auf E und bestimme den Vektor $\mathbf{a} = \overrightarrow{OP_0}$
- zwei Vektoren $\mathbf{v}$ und $\mathbf{w}$ ungleich Null in der Ebene, die nicht parallel sind (sind etwa P_1 und P_2 Punkte in der Ebene, die mit P_0 ein (nicht entartetes) Dreieck bilden, so kann man $\mathbf{v} = \overrightarrow{P_0P_1}$ und $\mathbf{w} = \overrightarrow{P_0P_2}$) wählen).

Damit liegt ein Punkt P in der Ebene E genau dann, wenn der Vektor $\mathbf{x} = \overrightarrow{OP}$ eine Darstellung der folgenden Form hat:

$$\mathbf{x} = \mathbf{a} + \lambda \mathbf{v} + \mu \mathbf{w} \quad (\lambda, \mu \in \mathbb{R}) .$$

Freie Vektoren und Ortsvektoren. In manchen Büchern werden *freie Vektoren* und *Ortsvektoren* unterschieden. Wir verstehen Vektoren als freie Vektoren in dem Sinne, daß sie im Raume in beliebiger Weise parallel verschoben werden dürfen. (Ein Kraftvektor ist immer derselbe Vektor gleichgültig an welchem Punkt die Kraft angesetzt wird.) Wenn man sich einen festen Bezugspunkt O wählt, dann bezeichnet man den Vektor $\overrightarrow{OP}$ auch als den Ortsvektor des Punktes P, und in diesem Sinne kann man die Punkte durch ihre Ortsvektoren bezüglich des Bezugspunktes O beschreiben. Verändert man den Bezugspunkt, transformiert sich auch der Ortsvektor eines Punktes.

1.1.8 Lineare Unabhängigkeit von zwei Vektoren. Wir haben gesehen, dass geometrische Objekte wie Punkte, Geraden und Ebenen durch Vektoren beschrieben werden können. Wir werden nun den geometrischen Begriff der *Parallelität* in der Sprache der Vektoren ausdrücken:

Definition. Zwei Vektoren $\mathbf{v}$, $\mathbf{w}$ heißen *linear unabhängig* [*linearly independent*], falls keiner der beiden ein skalares Vielfaches des anderen ist, d.h. falls es keinen Skalar λ gibt, so dass $\mathbf{v} = \lambda \mathbf{w}$ oder $\mathbf{w} = \lambda \mathbf{v}$.

Falls es einen Skalar λ gibt, so dass $\mathbf{v} = \lambda \mathbf{w}$ oder $\mathbf{w} = \lambda \mathbf{v}$, dann heißen $\mathbf{v}$, $\mathbf{w}$ *linear abhängig* [*linearly dependent*]. Lineare Abhängigkeit entspricht der Eigenschaft, parallel zu sein.

Beobachtung. *Für zwei Vektoren* $\mathbf{v}$*,* $\mathbf{w}$ *sind die folgenden Eigenschaften äquivalent:*

(1) $\mathbf{v}$*,* $\mathbf{w}$ *sind linear unabhängig.*

(2) *Falls* $\lambda \mathbf{v} + \mu \mathbf{w} = \mathbf{0}$*, dann ist* $\lambda = 0$ *und* $\mu = 0$*.*

(3) *Falls* $\lambda \mathbf{v} + \mu \mathbf{w} = \lambda' \mathbf{v} + \mu' \mathbf{w}$*, dann ist* $\lambda = \lambda'$ *und* $\mu = \mu'$*.*

Beweis. Wir zeigen (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2): Nehmen wir an, dass $\lambda\mathbf{v} + \mu\mathbf{w} = \mathbf{0}$ und $\lambda \neq 0$. Dann können wir die Gleichung mit $\frac{1}{\lambda}$ multiplizieren und erhalten $\mathbf{v} + \frac{\mu}{\lambda}\mathbf{w} = \mathbf{0}$, d.h. $\mathbf{v} = -\frac{\mu}{\lambda}\mathbf{w}$, ein Widerspruch zu (1).

(2) $\Rightarrow$ (3): Nehmen wir an, dass $\lambda\mathbf{v} + \mu\mathbf{w} = \lambda'\mathbf{v} + \mu'\mathbf{w}$. Dann ist $(\lambda - \lambda')\mathbf{v} + (\mu - \mu')\mathbf{w} = \mathbf{0}$. Aus (2) erhalten wir $\lambda - \lambda' = 0$ und $\mu - \mu' = 0$, und damit $\lambda = \lambda'$ sowie $\mu = \mu'$.

(3) $\Rightarrow$ (1): Angenommen $\mathbf{v}$ und $\mathbf{w}$ sind linear abhängig. Dann gibt es einen Skalar λ, so dass $\mathbf{v} = \lambda\mathbf{w}$ oder $\mathbf{w} = \lambda\mathbf{v}$. Im ersten Fall gilt $1\mathbf{v} + 0\mathbf{w} = 0\mathbf{v} + \lambda\mathbf{w}$. Falls (3) gelten würde, würden wir $1 = 0$ folgern, ein Widerspruch! Den zweiten Fall behandelt man analog. □

Die Implikation (3) $\Rightarrow$ (1) im letzten Beweis haben wir durch Widerspruch gezeigt. Sie sollten sich nochmals die logische Form der Argumente klar machen. Vergewissern Sie sich auch, dass wir nur die Regeln (V1)–(V8) und daraus abgeleitete Regeln verwendet haben.

1.1.9 Anwendung. Viele Sätze der Elementargeometrie können mit Hilfe der Vektoralgebra gezeigt werden. Betrachten wir ein Beispiel:

> In einem Parallelogramm ist der Schnittpunkt der beiden Diagonalen gleichzeitig der Mittelpunkt beider Diagonalen.

Man kann einen elementaren Beweis angeben, indem man zeigt, dass die beiden schattierten Dreiecke in der Abbildung kongruent sind. Diese Tatsache kann auch mit Hilfe der Vektoralgebra auf einfache Weise gezeigt werden: Wir wählen O, $\mathbf{v}$, $\mathbf{w}$ wie im Bild:

Nach 1.1.7 können die Punkte der ersten Diagonalen dargestellt werden in der Form

$$\mathbf{x} = \lambda(\mathbf{v} + \mathbf{w}), \quad \lambda \in \mathbb{R},$$

und die Punkte der zweiten Diagonalen in der Form

$$\mathbf{x} = \mathbf{w} + \mu(\mathbf{v} - \mathbf{w}), \quad \mu \in \mathbb{R}.$$

Da der Schnittpunkt auf beiden Diagonalen liegt, kann $\mathbf{x} = \overrightarrow{0S}$ auf zwei Weisen dargestellt werden

$$\mathbf{x} = \lambda(\mathbf{v} + \mathbf{w})$$

$$= \mathbf{w} + \mu(\mathbf{v} - \mathbf{w}).$$

Und damit

$$\begin{aligned} \lambda(\mathbf{v} + \mathbf{w}) &= \mathbf{w} + \mu(\mathbf{v} - \mathbf{w}) \\ \lambda\mathbf{v} + \lambda\mathbf{w} &= \mu\mathbf{v} + (1 - \mu)\mathbf{w}. \end{aligned}$$

Da $\mathbf{v}$ und $\mathbf{w}$ linear unabhängig sind, gilt mit 1.1.8 (3)

$$\lambda = \mu \text{ und } \lambda = 1 - \mu$$

Addition beider Gleichungen liefert $2\lambda = 1$, d.h. $\lambda = \frac{1}{2}$. Zusammen mit der ersten Gleichung erhalten wir $\mu = \frac{1}{2}$ und damit

$$\begin{aligned} \mathbf{x} &= \tfrac{1}{2}(\mathbf{v} + \mathbf{w}) \\ &= \mathbf{w} + \tfrac{1}{2}(\mathbf{v} - \mathbf{w}), \end{aligned}$$

was die Behauptung beweist.

1.1.10 Lineare Unabhängigkeit von drei Vektoren. Nun wollen wir beschreiben, wann drei Vektoren *nicht* in derselben Ebene liegen. Dazu führen wir eine nützliche Definition ein.

Definition. Seien $\mathbf{v}$ und $\mathbf{w}$ zwei Vektoren, dann heißt jeder Vektor der Form $\lambda\mathbf{v} + \mu\mathbf{w}$ $(\lambda, \mu \in \mathbb{R})$ *Linearkombination* [*linear combination*] von $\mathbf{v}$ und $\mathbf{w}$.

Definition. Drei Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}$ heißen *linear unabhängig* [*linearly independent*], falls keiner der drei Vektoren als Linearkombination der beiden anderen dargestellt werden kann. Andernfalls heißen sie *linear abhängig* [*linearly dependent*].

Wie in 1.1.8 hat man folgende Beziehungen (den Beweis führen wir in der Übung):

Beobachtung. *Für $\mathbf{u}, \mathbf{v}, \mathbf{w}$ sind die folgenden Eigenschaften äquivalent:*

(1) $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *sind linear unabhängig.*

(2) $\lambda\mathbf{u} + \mu\mathbf{v} + \nu\mathbf{w} = \mathbf{0}$ *impliziert* $\lambda = \mu = \nu = 0$.

(3) $\lambda\mathbf{u} + \mu\mathbf{v} + \nu\mathbf{w} = \lambda'\mathbf{u} + \mu'\mathbf{v} + \nu'\mathbf{w}$ *impliziert* $\lambda = \lambda'$, $\mu = \mu'$, $\nu = \nu'$.

Frage. Warum haben wir lineare Unabhängigkeit nicht definiert, indem wir nur sagen: $\mathbf{u}$ ist nicht Linearkombination von $\mathbf{v}$ und $\mathbf{w}$?

1.2. Der Begriff des reellen Vektorraums

Im ersten Abschnitt dieses Kapitels haben wir Vektoren im 2- und 3-dimensionalen Raum mit geometrischer Intuition eingeführt. Wir haben erkannt, dass gewisse Regeln grundlegend für die Vektorarithmetik sind. Wir wechseln nun den Blickwinkel. Entgegen dem vorherigen nehmen wir die Grundregeln als Axiome für die Vektorarithmetik und verwenden die geometrische Intuition lediglich als Motivation.

1.2.1 Definition: $\mathbb{R}$-Vektorraum. Ein *reeller Vektorraum* [*real vector space*] (kurz $\mathbb{R}$-*Vektorraum*) besteht aus einer Menge V, in der

- ein spezielles Element $\mathbf{0}$ ausgezeichnet ist,
- wir eine Operation haben, die je zwei Elementen $\mathbf{v}, \mathbf{w} \in V$ ein weiteres Element $\mathbf{v} + \mathbf{w} \in V$ zuordnet (Addition),
- und in der wir eine zweite Operation haben, die jedem $\lambda \in \mathbb{R}$ und jedem $\mathbf{v} \in V$ ein Element $\lambda\mathbf{v} \in V$ so zuordnet (Skalar-Multiplikation),

dass folgende Axiome erfüllt sind:

(V1)	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$
(V2)	$\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$	für alle $\mathbf{v} \in V$
(V3)	$\mathbf{v} + (-1)\mathbf{v} = \mathbf{0} = (-1)\mathbf{v} + \mathbf{v}$	für alle $\mathbf{v} \in V$
(V4)	$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$	für alle $\mathbf{v}, \mathbf{w} \in V$
(V5)	$\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$	für alle $\lambda, \mu \in \mathbb{R}$ und alle $\mathbf{v} \in V$
(V6)	$1\mathbf{v} = \mathbf{v}$	für alle $\mathbf{v} \in V$
(V7)	$\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$	für alle $\lambda \in \mathbb{R}$ und alle $\mathbf{v}, \mathbf{w} \in V$
(V8)	$(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$	für alle $\lambda, \mu \in \mathbb{R}$ und alle $\mathbf{v} \in V$.

Wir schreiben $-\mathbf{v}$ anstelle von $(-1)\mathbf{v}$ und $\mathbf{u} - \mathbf{v}$ anstelle von $\mathbf{u} + (-1)\mathbf{v}$.

Die reellen Zahlen λ, $\mu, \ldots$ heißen *Skalare* [*scalars*] und die Elemente $\mathbf{u}$, $\mathbf{v}, \ldots$ von V heißen *Vektoren* [*vectors*].

1.2.2 Abgeleitete Regeln. Für beliebige Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}$ eines reellen Vektorraums V und beliebige Skalare λ, μ gelten:

(V9)	$\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w} \quad \text{impliziert} \quad \mathbf{u} = \mathbf{v}$
(V10)	$0\mathbf{v} = \mathbf{0} \quad \text{und} \quad \lambda\mathbf{0} = \mathbf{0}$
(V11)	$\lambda\mathbf{v} = \mathbf{0} \quad \text{impliziert} \quad \lambda = 0 \text{ oder } \mathbf{v} = \mathbf{0}$
(V12)	$\lambda(\mathbf{u} - \mathbf{v}) = \lambda\mathbf{u} - \lambda\mathbf{v}$
(V13)	$(\lambda - \mu)\mathbf{v} = \lambda\mathbf{v} - \mu\mathbf{v}$

Die Herleitung dieser Regeln ist wie in 1.1.5 und 1.1.6; denn dort haben wir nur die Axiome verwendet.

Beispiele reeller Vektorräume.

(1) Die Menge aller Vektoren in der Ebene, vgl. Sektion 1.1.

(2) Die Menge aller Vektoren im 3-dimesionalen Raum, vgl. Sektion 1.1.

(3) Der Vektorraum $\mathbb{R}^n$ (siehe 1.2.3).

(4) Funktionenräume (siehe 1.2.4).

1.2.3 Der Vektorraum $\mathbb{R}^n$. Sei n eine natürliche Zahl. $\mathbb{R}^n$ ist die Menge aller n-Tupel [*n-tupels*] (= Liste mit n Einträgen)

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

wobei die Einträge $\alpha_1, \alpha_2, \ldots, \alpha_n$ beliebige reelle Zahlen sind. Wir definieren

$$\mathbf{0} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{v} + \mathbf{w} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}, \quad \lambda \mathbf{v} = \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \vdots \\ \lambda\alpha_n \end{pmatrix}.$$

Durch einfaches Nachrechnen überprüft man, dass die Axiome (V1)–(V8) erfüllt sind.

Spezialfall $n = 0$: Wir setzen $\mathbb{R}^0 := \{\mathbf{0}\}$.

Spezialfall $n = 1$: Wir haben $\mathbb{R}^1 = \mathbb{R}$.

1.2.4 Funktionenräume. Sei I eine Teilmenge der Menge der reellen Zahlen, z.B. ein Intervall $[a, b]$ oder auch $I = \mathbb{R}$. Wir betrachten Funktionen $f : I \to \mathbb{R}$, d.h. Funktionen, welche in jedem x in I definiert sind und deren Werte beliebige reelle Zahlen sein dürfen. Zwei solche Funktionen f und g kann man wie üblich addieren und mit einem Skalar α multiplizieren, indem man $f + g$ und αf wie üblich durch

$$(f + g)(x) = f(x) + g(x) \text{ und } (\alpha f)(x) = \alpha \cdot f(x) \text{ für alle } x \text{ in } I$$

definiert.

Man weist leicht nach, daß die Funktionen $f : I \to \mathbb{R}$ mit dieser Addition und Skalar-Multiplikation einen Vektorraum bilden, den wir mit $\mathcal{F}(I, \mathbb{R})$ bezeichnen wollen.

Dieser Vektorraum ist gewöhnungsbedürftig, denn er ist *unendlichdimensional*. Geometrische Vorstellungen sind nur mit Vorsicht übertragbar. Denn man muß nun Funktionen als *Punkte* in diesem Vektorraum betrachten. Insbesondere muß man sich davon freimachen, eine Funktion sich immer als eine Kurve in der Ebene vorzustellen, d.h. man muß einen abstrakteren Standpunkt einnehmen.

Gewisse Funktionenräume bilden jedoch in der Gestalt der Hilberträume die mathematische Grundlage der Quantenmechanik.

1.2.5 Der Vektorraum der Vektoren in der Ebene und der Vektorraum $\mathbb{R}^2$. Wir setzen nun den Vektorraum $\mathbb{R}^2$ in Beziehung zu den Vektoren in der Ebene, vgl. Sektion 1.1. Dazu führen wir ein Koordinatensystem in der Ebene ein.

Zu diesem Zweck wählen wir in der Ebene einen Bezugspunkt O sowie zwei linear unabhängige Vektoren

$$\mathcal{B} : \ \mathbf{u}_1, \mathbf{u}_2 \,.$$

Einen beliebigen Vektor $\mathbf{v}$ der Ebene können wir nun als Linearkombination

$$\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$$

darstellen. Die Skalare α_1 und α_2 sind eindeutig durch $\mathbf{v}$ bestimmt; denn aus $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2$ folgt $\alpha_1 = \beta_1$ und $\alpha_2 = \beta_2$ gemäß der Beobachtung (3) in 1.1.10.

Wir nennen $\mathcal{B}: \quad \mathbf{u}_1, \mathbf{u}_2$ eine *Vektorraumbasis* und α_1 und α_2 die *Koordinaten von* $\mathbf{v}$ *bezüglich der Basis* $\mathcal{B}$. Wir schreiben

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

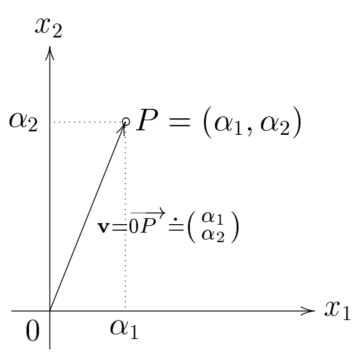
für den Vektor im $\mathbb{R}^2$, der aus den Koordinaten von $\mathbf{v}$ bezüglich der Basis $\mathcal{B}$ besteht. Man beachte, daß die Koordinaten α_1 und α_2 eines Vektors (anders als die Koordinaten eines Punktes) nicht von dem speziell gewählten Bezugspunkt O abhängen; geht man aber zu einer anderen Basis über, so verändern sich die Koordinaten. Wie sich die Koordinaten bei Basiswechsel transformieren, werden wir später eingehend behandeln.

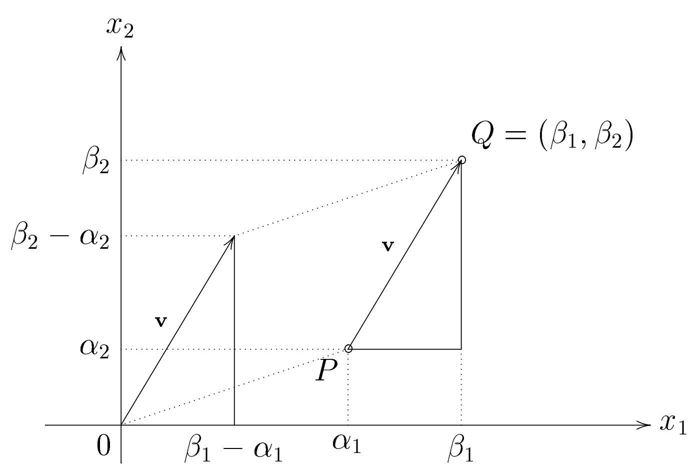
Wir haben durch unsere spezielle Koordinatisierung eine eineindeutige Beziehung zwischen den Vektoren der Ebene und den Elementen des $\mathbb{R}^2$ hergestellt: Jedem Vektor der Ebene haben wir eine eindeutig bestimmes Element des $\mathbb{R}^2$, seinen Koordinatenvektor, zugeordnet. Umgekehrt entspricht jedem Element $\left(\begin{smallmatrix} \alpha-1 \\ \alpha_2 \end{smallmatrix}\right)$ ein eindeutig bestimmter Vektor $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$ der Ebene mit ebendiesen Koordinaten.

In von uns gewählten Koordinatensystem wird ein Punkt P wird durch seine Koordinaten (α_1, α_2) beschrieben. Wir schreiben $P = (\alpha_1, \alpha_2)$.

Sind allgemeiner zwei Punkte $P = (\alpha_1, \alpha_2)$ und $Q = (\beta_1, \beta_2)$ gegeben, dann erhalten wir für $\mathbf{v} = \overrightarrow{PQ}$

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 - \alpha_1 \\ \beta_2 - \alpha_2 \end{pmatrix}.$$





Die Koordinaten des Nullvektors $\mathbf{0}$ sind beide 0, d.h.

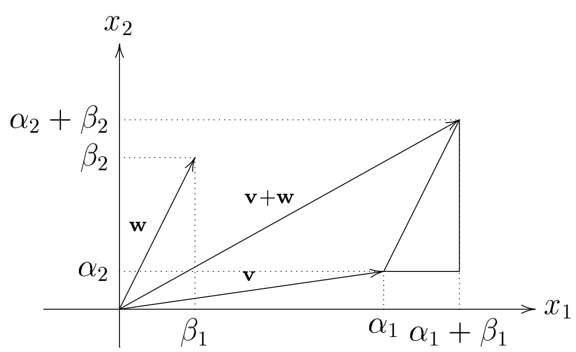
$$[\mathbf{0}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

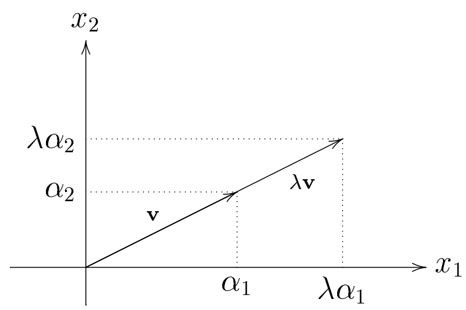
und für $\mathbf{v} \doteq \left(\begin{smallmatrix} \alpha_1 \\ \alpha_2 \end{smallmatrix}\right)$ und $\mathbf{w} \doteq \left(\begin{smallmatrix} \beta_1 \\ \beta_2 \end{smallmatrix}\right)$ haben wir

$$[\mathbf{v} + \mathbf{w}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$$

und

$$[\lambda\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix}\lambda\alpha_1\\ \lambda\alpha_2\end{pmatrix} = \lambda\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \lambda[\mathbf{v}]_{\mathcal{B}}\,.$$





Dies alles können wir auch wie folgt ausdrücken:

- Zum Nullvektor $\mathbf{0}$ in der Ebene korrespondiert der Nullvektor $\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$ im $\mathbb{R}^2$.
- Zu der geometrisch definierten Summe $\mathbf{v}+\mathbf{w}$ zweier Vektoren in der Ebene korrespondiert die algebraisch definierte Summe der zugehörigen Koordinatenvektoren im $\mathbb{R}^2$.
- Zu der geometrisch definierten Skalarmultiplikation $\lambda\mathbf{v}$ korrespondiert die algebraisch definierte Skalarmultiplikation im $\mathbb{R}^2$.

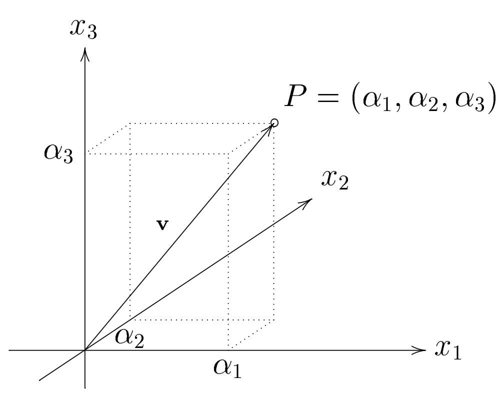
Diese Betrachtungen rechtfertigen, dass wir den Vektorraum der Vektoren in der Ebene, vgl. Sektion 1.1, mit dem Vektorraum $\mathbb{R}^2$ identifizieren.

1.2.6 $\mathbb{R}^3$ und der Vektorraum der Vektoren im 3-dimensionalen Raum.

Wie im ebenen Fall können wir den dreidimensionalen Raum koordinatisieren, indem wir einen Bezugspunkt O und drei linear unabhängige Vektoren

$$\mathcal{B}: \quad \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$$

auswählen. Punkte P können dann durch ihre drei Koordinaten $(\alpha_1, \alpha_2, \alpha_3)$ beschrieben werden.



Ist $\mathbf{v}$ ein Vektor im 3-dimensionalen Raum wie in Sektion 1.1, können wir ihn so verschieben, dass sein Anfangspunkt mit dem Ursprung O des Koordinatensystems übereinstimmt. Die Koordinaten des Endpunkts P werden dann als die Koordinaten von $\mathbf{v}$

bezüglich der Basis $\mathcal{B}$ definiert:

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

Damit korrespondiert zu jedem Vektor $\mathbf{v}$ im 3-dimensionalen Raum genau ein Element $\left(\begin{smallmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{smallmatrix}\right)$ aus $\mathbb{R}^3$ und umgekehrt. Wie im 2-dimensionalen Fall werden wir nun im $\mathbb{R}^3$ arbeiten und argumentieren und werden den „physikalischen" Raum für unsere Intuition verwenden.

1.2.7 Geraden und Ebenen in reellen Vektorräumen.

Basierend auf unserer Intuition in Dimension zwei und drei führen wir nun die Begriffe „Punkt, Gerade, Ebene" in beliebigen reellen Vektorräumen V ein.

Punkte: Ein *Punkt* [*point*] in V ist nichts anderes als ein Element $\mathbf{a} \in V$.

Geraden: Sei $\mathbf{a}$ ein Element in V und $\mathbf{v} \neq \mathbf{0}$ aus V. Die Menge aller $\mathbf{x} \in V$, die in der Form

$$\mathbf{x} = \mathbf{a} + \lambda \mathbf{v} \quad (\lambda \in \mathbb{R})$$

dargestellt werden können, ist die *Gerade* [*line*] durch $\mathbf{a}$ mit Richtung $\mathbf{v}$.

Ebenen: Sei $\mathbf{a}$ ein Punkt (Element) in V und $\mathbf{v}, \mathbf{w}$ zwei linear unabhängige Vektoren in V. Die Menge aller $\mathbf{x} \in V$, die in der Form

$$\mathbf{x} = \mathbf{a} + \lambda \mathbf{v} + \mu \mathbf{w} \quad (\lambda, \mu \in \mathbb{R})$$

dargestellt werden können, heißt *Ebene* [*plane*] in V. Der Punkt $\mathbf{a}$ liegt in der Ebene (setze $\lambda = \mu = 0$) und seine „Richtungen" sind gegeben durch $\mathbf{v}$ und $\mathbf{w}$.

??: (wie würden Sie weitermachen?)

In unserer Definition einer Ebene haben wir den Begriff der linearen Unabhängigkeit von zwei Vektoren verwendet. In einem beliebigen Vektorraum ist dieser Begriff genau auf dieselbe Weise definiert wie in Sektion 1.1. Wenn Sie den Begriff der linearen Unabhängigkeit verstanden haben, dann sollte es für Sie kein Problem sein, fortzufahren und die Äquivalenz der verschiedenen Charakterisierungen von linearer Unabhängigkeit im allgemeinen Fall zu beweisen. Überprüfen Sie Ihr Verständnis an folgender Übung:

1.2.8 Übung. Definieren Sie den Begriff der linearen Unabhängigkeit für zwei Vektoren $\mathbf{v}, \mathbf{w}$ und für drei Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in einem beliebigen reellen Vektorraum und beweisen Sie die Äquivalenz der verschiedenen Charakterisierung linearer Unabhängigkeit wie in 1.1.8 und 1.1.10.

1.3. Lineare Gleichungssysteme

Viele Probleme führen auf lineare Gleichungssysteme. Das Lösen linearer Gleichungssysteme ist eine weitere Motivation für die Beschäftigung mit Linearer Algebra. Wir stellen die folgenden Fragen für ein gegebenes System von Gleichungen:

- Gibt es überhaupt eine Lösung?
- Wenn das System eine Lösung besitzt, ist diese eindeutig?

- Falls es mehr als eine Lösung gibt, wie kann man die Menge aller Lösungen beschreiben?
- Gibt es Methoden, die Lösungen praktisch zu bestimmen?

In dieser Vorlesung betrachten wir primär die prinzipiellen Aspekte dieser Fragen, die praktischen Aspekte werden in Vorlesungen über Numerische Mathematik behandelt.

Lineare Gleichungssysteme würden keine großen Schwierigkeiten darstellen, wenn man nur Systeme mit wenigen Variablen betrachten würde. Jedoch muß man in der Praxis Systeme mit mehreren 100.000 Gleichungen und ebenso vielen Variablen behandeln können. Solche Systeme können mit Techniken gelöst werden, die auf fundierter mathematischer Theorie basieren.

1.3.1 Zur Motivation. Häufig liegt eine Situation vor, in der m skalare Größen $y_1, y_2, \ldots, y_m$ linear von n weiteren skalaren Größen $x_1, x_2, \ldots, x_n$ abhängen. Das heißt, die Beziehungen zwischen den y_i und den x_j kann durch Gleichungen der folgenden Form ausgedrückt werden, wobei die α_{ij} reelle Zahlen bezeichnen:

$$\text{(Z)} \qquad \begin{array}{ccccccccc}
\alpha_{11}x_1 & + & \alpha_{12}x_2 & + \ldots + & \alpha_{1j}x_j & + \ldots + & \alpha_{1n}x_n & = & y_1 \\
\alpha_{21}x_1 & + & \alpha_{22}x_2 & + \ldots + & \alpha_{2j}x_j & + \ldots + & \alpha_{2n}x_n & = & y_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\alpha_{i1}x_1 & + & \alpha_{i2}x_2 & + \ldots + & \alpha_{ij}x_j & + \ldots + & \alpha_{in}x_n & = & y_i \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + \ldots + & \alpha_{mj}x_j & + \ldots + & \alpha_{mn}x_n & = & y_m
\end{array}$$

Beispiel. Eine Firma verwendet m verschiedene Rohstoffe zur Herstellung von n verschiedenen Produkten. Um eine Einheit von Produkt j herzustellen, benötigt die Firma α_{ij} Einheiten des Rohstoffs i. Wir stellen die Zahlen α_{ij} in der folgenden Tabelle dar:

	Produkt 1	Produkt 2	......	Produkt j	......	Produkt n
Rohstoff 1	α_{11}	α_{12}	......	α_{1j}	......	α_{1n}
Rohstoff 2	α_{21}	α_{22}	......	α_{2j}	......	α_{2n}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Rohstoff i	α_{i1}	α_{i2}	......	α_{ij}	......	α_{in}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Rohstoff m	α_{m1}	α_{m2}	......	α_{mj}	......	α_{mn}

Wenn wir $x_1 = a_1$ Einheiten von Produkt 1, $x_2 = a_2$ Einheiten von Produkt 2, ..., und $x_n = a_n$ Einheiten von Produkt n herstellen, erlauben uns obige Gleichungen die notwendigen Mengen $y_1, y_2, \ldots, y_m$ an Rohstoffen $1, 2, \ldots, m$ auszurechnen.

Auf der anderen Seite kann es sein, dass die zur Verfügung stehenden Rohstoffmengen vorgegeben sind, sagen wir $y_1 = \beta_1, y_2 = \beta_2, \ldots, y_m = \beta_m$. Die Frage lautet dann, wieviele Einheiten $x_1, x_2, \ldots, x_n$ der Produkte $1, 2, \ldots, n$ können damit hergestellt werden.

Das heißt, man fragt nach Lösungen des folgenden Gleichungssystems:

$$
\text{(L)} \qquad
\begin{array}{ccccccccc}
\alpha_{11}x_1 & + & \alpha_{12}x_2 & + \ldots + & \alpha_{1j}x_j & + \ldots + & \alpha_{1n}x_n & = & \beta_1 \\
\alpha_{21}x_1 & + & \alpha_{22}x_2 & + \ldots + & \alpha_{2j}x_j & + \ldots + & \alpha_{2n}x_n & = & \beta_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\alpha_{i1}x_1 & + & \alpha_{i2}x_2 & + \ldots + & \alpha_{ij}x_j & + \ldots + & \alpha_{in}x_n & = & \beta_i \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + \ldots + & \alpha_{mj}x_j & + \ldots + & \alpha_{mn}x_n & = & \beta_m
\end{array}
$$

Das ist ein System mit m linearen Gleichungen in n Unbekannten $x_1, x_2, \ldots, x_n$. Wir interessieren uns für die Menge aller

$$
\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,
$$

welche all die obigen Gleichungen gleichzeitig erfüllen.

1.3.2 Der Gauß-Jordan Eliminations-Algorithmus. Hier handelt es sich um eine systematische Vorgehensweise zum Lösen linearer Gleichungssysteme, die auch von großem praktischen Interesse ist. Sie basiert auf der folgenden

Beobachtung. *Die folgenden drei Operationen transformieren ein Gleichungssystem in ein dazu äquivalentes, d.h. in ein System, welches dieselben Lösungen besitzt wie das ursprüngliche Problem:*

(E1) *Addition eines skalaren Vielfachen einer Gleichung zu einer anderen Gleichung.*

(E2) *Vertauschen zweier Gleichungen.*

(E3) *Multiplikation einer Gleichung mit einem Skalar* $\lambda \neq 0$.

Diese drei Operationen heißen elementare (Zeilen-)Umformungen [elementary (row) transformations] *eines Gleichungssystems.*

Im ERSTEN SCHRITT des Gauß-Algorithmus transformiert man das Originalsystem (L) durch elementare Umformungen in ein System (L′) der folgenden Form (tatsächlich braucht man nur die Umformungen vom Typ (E1) und (E2)):

$$
\text{(L}'\text{)} \qquad
\begin{array}{ccccccccc}
\alpha'_{11}x_1 & + & \alpha'_{12}x_2 & + \ldots + & \alpha'_{1j}x_j & + \ldots + & \alpha'_{1n}x_n & = & \beta'_1 \\
 & & \alpha'_{22}x_2 & + \ldots + & \alpha'_{2j}x_j & + \ldots + & \alpha'_{2n}x_n & = & \beta'_2 \\
 & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & \alpha'_{i2}x_2 & + \ldots + & \alpha'_{ij}x_j & + \ldots + & \alpha'_{in}x_n & = & \beta'_i \\
 & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & \alpha'_{m2}x_2 & + \ldots + & \alpha'_{mj}x_j & + \ldots + & \alpha'_{mn}x_n & = & \beta'_m
\end{array}
$$

Dies kann man auf folgende Weise erreichen:

- Falls $\alpha_{11} \neq 0$, lasse die erste Gleichung unverändert und subtrahiere

  $\frac{\alpha_{21}}{\alpha_{11}}$-mal die erste Gleichung von der zweiten
  $\frac{\alpha_{31}}{\alpha_{11}}$-mal die erste Gleichung von der dritten
  $\vdots$
  $\frac{\alpha_{m1}}{\alpha_{11}}$-mal die erste Gleichung von der letzten Gleichung.
- Falls $\alpha_{11} = 0$, suche eine Gleichung, in der der erste Koeffizient $\alpha_{i1} \neq 0$ ist; tausche die erste Gleichung mit der i-ten Gleichung und fahre wie im ersten Schritt fort. (Bemerkung: Aus numerischen Gründen ist es günstig i so zu wählen, dass $|\alpha_{i1}|$ maximal ist, aber das ist im Moment noch nicht wichtig.)
- Falls $\alpha_{i1} = 0$ für alle $i = 1, \ldots, m$, dann ist nichts zu tun und wir fahren mit dem zweiten Schritt fort.

Im ZWEITEN SCHRITT lassen wir die erste Gleichung unberührt und für die verbleibenden $m-1$ Gleichungen verfahren wir wie im ersten Schritt.

Nach höchstens $m-1$ Schritten enden wir mit einem System von *stufenförmiger Gestalt* [*echelon form*]:

$$
\tag{M}
\begin{array}{llllcl}
\boxed{\gamma_{1j_1}}x_{j_1} + \ldots + & \gamma_{1j_2}x_{j_2} + \ldots + & \gamma_{1j_r}x_{j_r} & + \ldots + \gamma_{1n}x_n & = & \delta_1 \\
 & \boxed{\gamma_{2j_2}}x_{j_2} + \ldots + & \gamma_{2j_r}x_{j_r} & + \ldots + \gamma_{2n}x_n & = & \delta_2 \\
 & & \ddots & \qquad \vdots & & \vdots \\
 & & \boxed{\gamma_{rj_r}}x_{j_r} & + \ldots + \gamma_{rn}x_n & = & \delta_r \\
 & & & \qquad 0 & = & \delta_{r+1} \\
 & & & \qquad \vdots & & \vdots \\
 & & & \qquad 0 & = & \delta_m
\end{array}
$$

Die eingerahmten Elemente $\gamma_{1j_1}, \gamma_{2j_2}, \ldots, \gamma_{rj_r}$ sind alle von Null verschieden; man nennt sie *Pivotelemente* [*pivot entries*]. Die zugehörigen Variablen $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ heißen *Pivotvariablen* [*pivot variables*]. Wie wir oben angedeutet haben, kann es vorkommen, dass nach der Transformation in Stufenform die linke Seite einiger Gleichungen Null wird. Die Zahl r an Gleichungen, für die die linke Seite nicht verschwindet, nennt man den *Rang* [*rank*] des linearen Gleichungssystems (L).

Es ist nicht schwierig, das Gleichungssystem zu lösen, wenn wir es in die Stufenform (M) gebracht haben:

1. Fall: Eine der rechten Seiten $\delta_{r+1}, \ldots, \delta_m$ ist ungleich Null. In diesem Fall widersprechen sich die Gleichungen; das Gleichungssystem hat also keine Lösung.

2. Fall: $\delta_{r+1} = \cdots = \delta_m = 0$. In diesem Fall hat das Gleichungssystem mindestens eine Lösung. Man kann alle Lösungen auf die folgende Weise finden: Die Werte der $n-r$ Variablen x_i, die keine Pivotvariablen sind, können beliebig $x_i = \lambda_i$ gewählt werden. Die Werte der Pivotvariablen werden danach von unten beginnend durch Rückwärtseinsetzen $x_{j_r} = --, \ldots, x_{j_2} = --, x_{j_1} = --$ bestimmt.

Beachte, dass die Stufenform $r \leq n$ impliziert. Damit haben wir folgenden Satz gezeigt:

1.3.3 Theorem. *Wenn r der Rang eines Systems von m linearen Gleichungen in n Variablen ist, dann hat es entweder keine Lösung oder die Werte von $n - r$ Variablen können beliebig gewählt werden.*

In den folgenden Fällen hat das System (L) mindestens eine Lösung:

(1) $r = m$, d.h. der Rang ist gleich der Anzahl Gleichungen: In diesem Fall verschwindet die linke Seite von keiner Gleichung.

(2) Das Gleichungssystem (L) ist *homogen* [*homogeneous*], d.h. $\beta_1 = \beta_2 = \cdots = \beta_m = 0$. Ein homogenes System hat mindestens eine Lösung, nämlich $x_1 = x_2 = \cdots = x_n = 0$, die sogenannte *triviale* Lösung.

Schließlich können wir charakterisieren, wann das System (L) eine eindeutige Lösung hat. Dazu beschränken wir uns auf den Fall $n = m$, d.h. den Fall von n Gleichungen in n Variablen:

1.3.4 Theorem. *Ein System mit n linearen Gleichungen in n Variablen hat dann und nur dann eine eindeutige Lösung, falls sein Rang $r = n$ ist.*

1.3.5 Beispiel. Wir betrachten ein System mit 5 Gleichungen in 5 Unbekannten. Wir wollen das System für verschiedene rechte Seiten lösen:

1: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

2: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 1$

3: $\beta_1 = \beta_2 = \beta_3 = 1, \beta_4 = 4, \beta_5 = 2$

	1	2	3	x_1	x_2	x_3	x_4	x_5	β_1	β_2	β_3	β_4	β_5
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = \beta_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-x_1 - x_2 - 2x_3 + x_4 + x_5 = \beta_2$	0	1	1	−1	−1	−2	1	1	0	1	0	0	0
$2x_1 + 4x_3 + 2x_4 + x_5 = \beta_3$	0	1	1	2	0	4	2	1	0	0	1	0	0
$x_1 - 3x_2 + 2x_3 + 6x_4 + 6x_5 = \beta_4$	0	1	4	1	−3	2	6	6	0	0	0	1	0
$-x_1 - 3x_2 - 2x_3 + 3x_4 + 6x_5 = \beta_5$	0	1	2	−1	−3	−2	3	6	0	0	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = \beta_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = \beta_2 + \beta_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$2x_2 - 2x_4 - 5x_5 = \beta_3 - 2\beta_1$	0	−1	−1	0	2	0	−2	−5	−2	0	1	0	0
$-2x_2 + 4x_4 + 3x_5 = \beta_4 - \beta_1$	0	0	3	0	−2	0	4	3	−1	0	0	1	0
$-4x_2 + 5x_4 + 9x_5 = \beta_5 + \beta_1$	0	2	3	0	−4	0	5	9	1	0	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = \beta_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = \beta_2 + \beta_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$x_4 - x_5 = \beta_3 + \beta_2 - \beta_1$	0	1	1	0	0	0	1	−1	−1	1	1	0	0
$x_4 - x_5 = \beta_4 - \beta_2 - 2\beta_1$	0	−2	1	0	0	0	1	−1	−2	−1	0	1	0
$-x_4 + x_5 = \beta_5 - 2\beta_2 - \beta_1$	0	−2	−1	0	0	0	−1	1	−1	−2	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = \beta_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = \beta_2 + \beta_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$x_4 - x_5 = \beta_3 + \beta_2 - \beta_1$	0	1	1	0	0	0	1	−1	−1	1	1	0	0
$0 = \beta_4 - \beta_3 - 2\beta_2 - \beta_1$	0	−3	0	0	0	0	0	0	−1	−2	−1	1	0
$0 = \beta_5 + \beta_3 - \beta_2 - 2\beta_1$	0	−1	0	0	0	0	0	0	−2	−1	1	0	1

Der rechte Teil obiger Tabelle ist nochmals dargestellt in Abbildung 1.1. Sie zeigt ein Schema, in dem man nur die Koeffizienten der x_i und der β_i darstellt, während die Gleichungen transformiert werden.

x_1	x_2	x_3	x_4	x_5	β_1	β_2	β_3	β_4	β_5
1	−1	2	2	3	1				
−1	−1	−2	1	1		1			
2	0	4	2	1			1		
1	−3	2	6	6				1	
−1	−3	−2	3	6					1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	2	0	−2	−5	−2		1		
0	−2	0	4	3	−1			1	
0	−4	0	5	9	1				1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	0	0	1	−5	−2	1	1		
0	0	0	1	3	−1	−1		1	
0	0	0	−1	9	1	−2			1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	0	0	1	−1	−2	1	1		
0	0	0	0	0	−1	−1	−1	1	
0	0	0	0	0	1	−2	1	0	1

ABBILDUNG 1.1. Schema für das Beispiel aus 1.3.5. Beachte, dass die leeren Einträge in diesem Schema Nullen bedeuten.

Die Lösungen für die drei Fälle können nun wie folgt bestimmt werden:

(1) $x_1 = -\frac{3}{2}\lambda - 2\mu$
$x_2 = -\frac{7}{2}\lambda$
$x_3 = \mu$ (beliebig)
$x_4 = \lambda$
$x_5 = \lambda$ (beliebig)

Lösungen:

$$\mathbf{x} = \begin{pmatrix} -\frac{3}{2}\lambda - 2\mu \\ -\frac{7}{2} \\ \mu \\ \lambda \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} -\frac{3}{2} \\ -\frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad (\lambda, \mu \text{ beliebig})$$

(2) Es gibt keine Lösung; die Umformungen führen zu zwei Widersprüchen $0 = -3$ und $0 = -1$.

(3) Die letzten beiden Gleichungen ergeben $0 = 0$. Wir können $x_5 = \lambda$ und $x_3 = \mu$ beliebig wählen und erhalten die Lösungen:

$$\mathbf{x} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (\lambda, \mu \text{ beliebig})$$

Bemerkungen:

- Der Rang des Systems ist $r = 3$.
- Die Pivotvariablen sind x_1, x_2, x_4.
- x_3, x_5 können beliebig gewählt werden.
- Die Lösungen im homogenen Fall (1) bilden eine Ebene im $\mathbb{R}^5$ durch den Punkt $\mathbf{0}$.
- Die Menge der Lösungen im inhomogenen Fall kann leer sein (2) oder eine Ebene im $\mathbb{R}^5$ sein, die parallel zu der Ebene im homogenen Fall (3) ist.
- Im homogenen Fall (1) nennt man die zwei Vektoren

$$\mathbf{v} = \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Basislösungen [*basic solutions*], auch *Fundamentalsystem von Lösungen* [*fundamental system of solutions*] genannt. Alle anderen Lösungen sind „nur“ Linearkombination von $\mathbf{v}$ und $\mathbf{w}$.

1.3.6 Ein paar einfache Fälle.

- $n = 1$, $m = 1$:

$$\alpha x_1 = \beta$$

Wir unterscheiden drei Fälle:

(1) $\alpha \neq 0$: Dann ist $x_1 = \frac{\alpha}{\beta}$ die einzige Lösung.

(2) $\alpha = 0$, $\beta \neq 0$: Es gibt überhaupt keine Lösung.

(3) $\alpha = 0$, $\beta = 0$: $x_1 = \lambda$ ist eine Lösung für jedes $\lambda \in \mathbb{R}$.

- $n = 2$, $m = 1$:

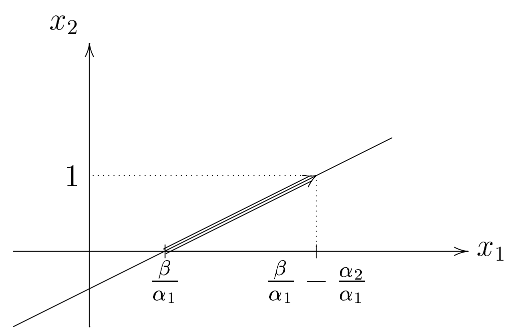
$$\alpha_1 x_1 + \alpha_2 x_2 = \beta$$

Es können vier Fälle auftreten:

(1) $\alpha_1 \neq 0$: Man kann $x_2 = \lambda \in \mathbb{R}$ beliebig wählen und dann $x_1 = -\frac{\alpha_2}{\alpha_1}\lambda + \frac{\beta}{\alpha_1}$ berechnen. Die Lösungen sind demnach

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{\alpha_2}{\alpha_1}\lambda + \frac{\beta}{\alpha_1} \\ \lambda \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\alpha_1} \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{\alpha_2}{\alpha_1} \\ 1 \end{pmatrix}.$$

Dies ist eine Parameterdarstellung der Geraden durch den Punkt $(\frac{\beta}{\alpha_1}, 0)$, deren Richtung festgelegt ist durch den Vektor $\mathbf{v} = \left(\begin{smallmatrix} -\frac{\alpha_2}{\alpha_1} \\ 1 \end{smallmatrix}\right)$.



(2) $\alpha_1 = 0$, $\alpha_2 \neq 0$: Wähle $x_1 = \lambda \in \mathbb{R}$ beliebig und berechne $x_2 = \frac{\beta}{\alpha_2}$

$$\mathbf{x} = \begin{pmatrix} \lambda \\ \frac{\beta}{\alpha_2} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\beta}{\alpha_2} \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Dies beschreibt eine horizontale Gerade durch den Punkt $(0, \frac{\beta}{\alpha_2})$.

(3) $\alpha_1 = 0$, $\alpha_2 = 0$, $\beta \neq 0$: Es gibt keine Lösung.

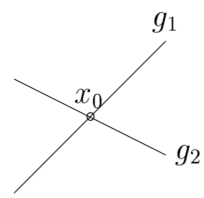
(4) $\alpha_1 = 0$, $\alpha_2 = 0$, $\beta = 0$: Alle Punkte $\left(\begin{smallmatrix} \lambda \\ \mu \end{smallmatrix}\right) \in \mathbb{R}^2$ sind Lösungen.

- $n = 2$, $m = 2$:

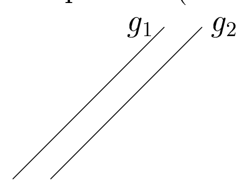
$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 &= \beta_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 &= \beta_2 \end{aligned}$$

Nehmen wir an, dass in jeder Gleichung einer der Koeffizienten $\alpha_{ij} \neq 0$ ist. Dann bilden die Lösungen für Gleichung 1 eine Gerade g_1 in $\mathbb{R}^2$ und für Gleichung 2 eine Gerade g_2. Die Lösungen des Systems sind die Punkte, die in der Schnittmenge der beiden Geraden g_1, g_2 liegen. Es können drei Fälle auftreten:

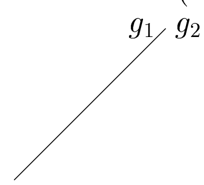
(1) g_1 und g_2 schneiden sich in einem Punkt (eindeutige Lösung):



(2) Die beiden Geraden sind parallel (keine Lösung):



(3) Die beiden Geraden sind identisch (alle Geradenpunkte sind Lösungen):



- $n = 3, m = 1$:

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \beta.$$

Angenommen eines der α_i ist $\neq 0$, zum Beispiel $\alpha_1 \neq 0$. Dann gilt $x_1 = \frac{1}{\alpha_1}(\beta - \alpha_2 x_2 - \alpha_3 x_3)$. Wir können $x_2 = \lambda$, $x_3 = \mu$ beliebig wählen und erhalten $x_1 = \frac{\beta}{\alpha_1} - \frac{\alpha_2}{\alpha_1}\lambda - \frac{\alpha_3}{\alpha_1}\mu$. Damit sind die Lösungen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\alpha_1} - \frac{\alpha_2}{\alpha_1}\lambda - \frac{\alpha_3}{\alpha_1}\mu \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\alpha_1} \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{\alpha_2}{\alpha_1} \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -\frac{\alpha_3}{\alpha_1} \\ 0 \\ 1 \end{pmatrix}.$$

Dies ist eine Parameterdarstellung einer Ebene $\mathbf{x} = \mathbf{a} + \lambda\mathbf{v} + \mu\mathbf{w}$.

- $n = 3, m = 3$:

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 &= \beta_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 &= \beta_2 \\ \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3 &= \beta_3 \end{aligned}$$

Angenommen in jeder Gleichung gibt es einen Koeffizienten α_{ij} ungleich Null. Wie wir aus obigen Fällen wissen, bildet die Lösungsmenge einer jeden Gleichung eine Ebene im $\mathbb{R}^3$.

Die gemeinsamen Lösungen aller drei Gleichungen sind die Punkte, die im Durchschnitt aller drei Ebenen E_1, E_2, E_3 liegen. Es können mehrere Fälle auftreten, siehe Abbildung 1.2.

(a) Die drei Ebenen sind parallel, aber nicht identisch: keine Lösung.

(b) Zwei der Ebenen sind parallel, aber nicht die dritte: Keine Lösung.

(c) Die drei Ebenen sind identisch: die Lösungen sind die Punkte dieser Ebene.

(d) Die drei Ebenen schneiden sich in einer Geraden: die Lösungen sind die Punkte dieser Geraden.

(e) Die drei Ebenen schneiden sich in einem Punkt: Dieser Punkt ist die eindeutige Lösung.

Zur Berechnung dieser Lösungen können wir den Gauß-Jordan Eliminations-Algorithmus verwenden.

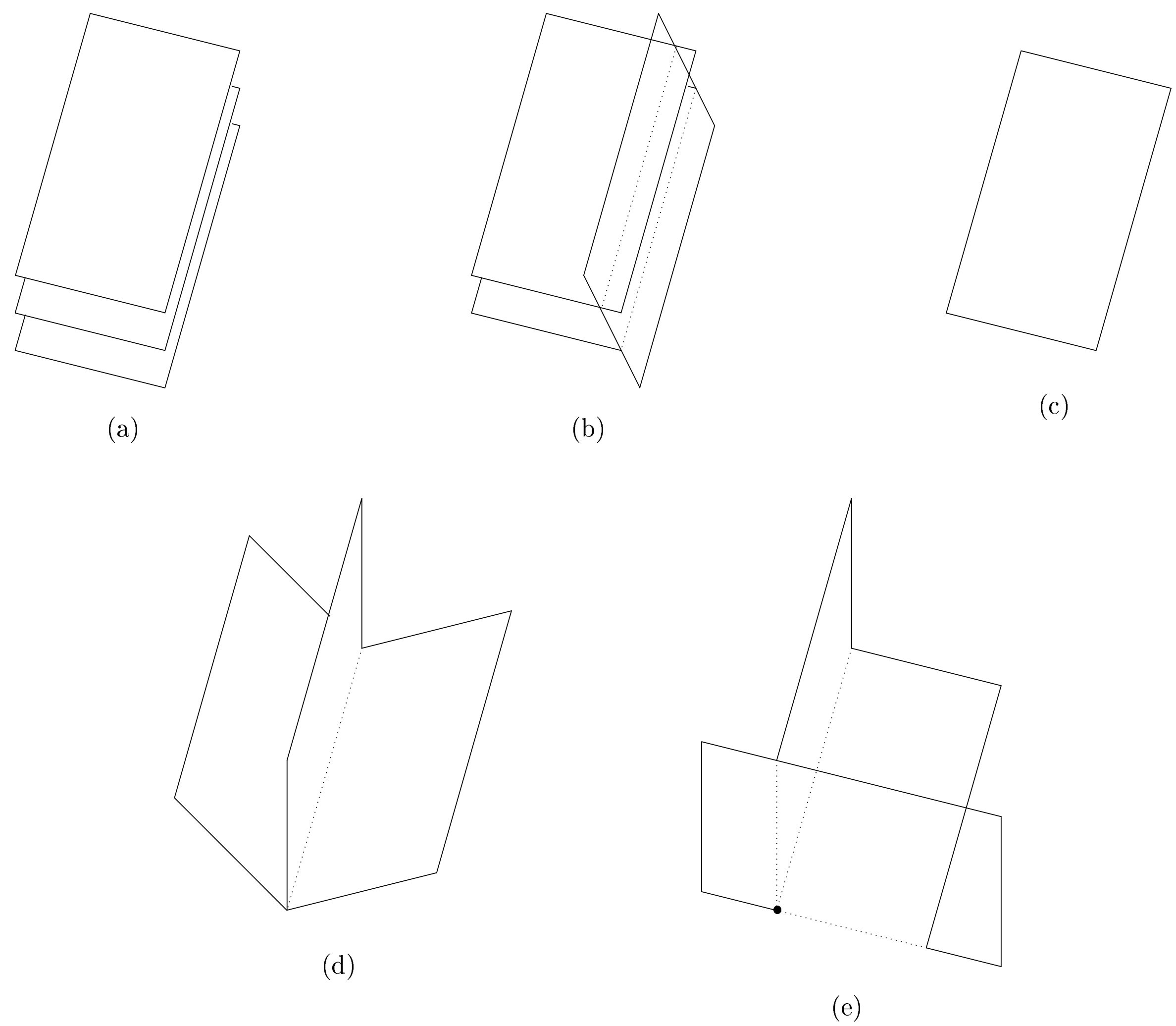


Abbildung 1.2. Mögliche Schnittmengen dreier Ebenen

1.4. Matrizen

In 1.3.5 haben wir gesehen, dass es zur Lösung von Gleichungssystemen nicht notwendig ist, mit den Gleichungen selbst zu arbeiten, sondern die Tabellen der Koeffizienten ausreichen. Diese Tabellen von Koeffizienten nennt man Matrizen:

1.4.1 $\mathbf{m \times n}$-Matrizen mit reellen Einträgen. Eine *reelle $m \times n$-Matrix* [*real $m \times n$-matrix*] ist ein rechteckiges Schema A von reellen Zahlen α_{ij} mit m Zeilen und n

Spalten. Wir schreiben:

$$A = (\alpha_{ij}) = (\alpha_{ij})_{m\times n} = (\alpha_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2j} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mj} & \dots & \alpha_{mn} \end{pmatrix}$$

Die Zahlen α_{ij} heißen *Einträge* [*entries*] der Matrix. Der Index i bezeichnet die *Zeile* [*row*] und der Index j die *Spalte* [*column*], in der Eintrag α_{ij} steht. Wir bezeichnen mit $\mathcal{M}(m \times n, \mathbb{R})$, oder kurz mit $\mathbb{R}^{m\times n}$, die Menge aller $m \times n$-Matrizen A mit reellen Einträgen.

Haben wir ein Gleichungssystem (L) wie in 1.3.1, dann nennt man die $m \times n$-Matrix $A = (\alpha_{ij})$ die *Koeffizientenmatrix* [*matrix of coefficients*] von (L).

Spezialfälle.

- $m \times 1$-Matrizen sind nichts anderes als (Spalten-) Vektoren

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{R}^n.$$

- $1 \times n$-Matrizen sind nichts anderes als (Zeilen-) Vektoren mit n Einträgen:

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}.$$

- Die $m \times n$-Nullmatrix ist die Matrix, deren Einträge alle Null sind:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

- $n \times n$-Matrizen heißen *quadratische* Matrizen:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1i} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2i} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ii} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{ni} & \dots & \alpha_{nn} \end{pmatrix}$$

  In einer quadratischen Matrix bilden die Einträge $\alpha_{11}, \alpha_{22}, \dots, \alpha_{nn}$ die *Hauptdiagonale* [*main diagonal*] der Matrix.

- Die $n \times n$-Matrix mit Einsen in der Hauptdiagonalen und Nullen sonst heißt $n \times n$-*Einheitsmatrix* [*identity matrix*] $E = E_n$. Genauer:

  $$E_n = (\delta_{ij})_{n\times n}, \text{ wobei } \delta_{ij} \text{ das Kronecker Symbol } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \text{ ist.}$$

  $$\text{Zum Beispiel } E_1 = (1), \quad E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Unter den quadratischen Matrizen $A = (\alpha_{ij})$ unterscheiden wir

  Diagonalmatrizen, bei denen $\alpha_{ij} = 0$ für $i \neq j$,
  obere Dreiecksmatrizen, bei denen $\alpha_{ij} = 0$ für $i > j$,
  untere Dreiecksmatrizen, bei denen $\alpha_{ij} = 0$ für $i < j$:

$$\begin{pmatrix} \alpha_{11} & 0 & \cdots & 0 \\ 0 & \alpha_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_{nn} \end{pmatrix}, \quad \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ & \alpha_{22} & \dots & \alpha_{2n} \\ & & \ddots & \vdots \\ 0 & & & \alpha_{nn} \end{pmatrix}, \quad \begin{pmatrix} \alpha_{11} & & & 0 \\ \alpha_{21} & \alpha_{22} & & \\ \vdots & \vdots & \ddots & \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}.$$

1.4.2 Addition von $\mathbf{m \times n}$-Matrizen und Multiplikation mit Skalaren.

Ähnlich wie in unserer Definition von Vektoren definieren wir die Summe zweier $m \times n$-Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{ij})$ und die Skalarmultiplikation $\lambda \in \mathbb{R}$ Eintrag für Eintrag:

$$A + B := (\alpha_{ij} + \beta_{ij}), \qquad \lambda A := (\lambda \alpha_{ij}).$$

Genauer:

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \alpha_{21} & \dots & \alpha_{2n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \beta_{21} & \dots & \beta_{2n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \end{pmatrix} := \begin{pmatrix} \alpha_{11} + \beta_{11} & \dots & \alpha_{1n} + \beta_{1n} \\ \alpha_{21} + \beta_{21} & \dots & \alpha_{2n} + \beta_{2n} \\ \vdots & & \vdots \\ \alpha_{m1} + \beta_{m1} & \dots & \alpha_{mn} + \beta_{mn} \end{pmatrix}$$

$$\lambda \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} := \begin{pmatrix} \lambda\alpha_{11} & \lambda\alpha_{12} & \dots & \lambda\alpha_{1n} \\ \lambda\alpha_{21} & \lambda\alpha_{22} & \dots & \lambda\alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda\alpha_{m1} & \lambda\alpha_{m2} & \dots & \lambda\alpha_{mn} \end{pmatrix}.$$

Die Gültigkeit der Axiome (V1)–(V8) können für die Addition von $m \times n$-Matrizen und für die Skalarmultiplikation leicht überprüft werden. Damit erhalten wir

Proposition. *Mit der Addition und Skalarmultiplikation wie oben ist die Menge $\mathcal{M}(m \times n, \mathbb{R})$ aller $m \times n$-Matrizen ein $\mathbb{R}$-Vektorraum.*

Beispiel.

$$\begin{aligned}
&4\begin{pmatrix}2&1&3\\1&0&-1\end{pmatrix}-3\begin{pmatrix}1&-1&0\\0&-2&1\end{pmatrix}\\
=\;&\begin{pmatrix}8&4&12\\4&0&-4\end{pmatrix}+\begin{pmatrix}-3&3&0\\0&6&-3\end{pmatrix}\\
=\;&\begin{pmatrix}5&7&12\\4&6&-7\end{pmatrix}
\end{aligned}$$

Achtung. Die Addition ist nur für Matrizen derselbe Größe definiert.

1.4.3 Multiplikation von Matrizen. Das Produkt einer $m \times n$-Matrix $A = (\alpha_{ij})$ und einer $n \times p$-Matrix $B = (\beta_{jk})$ ist definiert als die $m \times p$-Matrix $C = (\gamma_{ik})$ mit

$$\gamma_{ik} := \sum_{j=1}^{n} \alpha_{ij}\beta_{jk} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \cdots + \alpha_{in}\beta_{nk},$$

genauer:

$$i \rightarrow \begin{pmatrix}
\alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n}\\
\vdots & \vdots & & \vdots & & \vdots\\
\alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in}\\
\vdots & \vdots & & \vdots & & \vdots\\
\alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mj} & \dots & \alpha_{mn}
\end{pmatrix} \cdot
\begin{array}{c}
\hphantom{\beta_{11}\ \dots\ }k\hphantom{\ \dots\ \beta_{1p}}\\
\begin{pmatrix}
\beta_{11} & \dots & \beta_{1k} & \dots & \beta_{1p}\\
\beta_{21} & \dots & \beta_{2k} & \dots & \beta_{2p}\\
\vdots & & \vdots & & \vdots\\
\beta_{j1} & \dots & \beta_{jk} & \dots & \beta_{jp}\\
\vdots & & \vdots & & \vdots\\
\beta_{n1} & \dots & \beta_{nk} & \dots & \beta_{np}
\end{pmatrix}
\end{array} =$$

$$= \; i \rightarrow
\begin{array}{c}
\hphantom{\gamma_{11}\ \dots\ }k\hphantom{\ \dots\ \gamma_{1p}}\\
\begin{pmatrix}
\gamma_{11} & \dots & \gamma_{1k} & \dots & \gamma_{1p}\\
\vdots & & \vdots & & \vdots\\
\gamma_{i1} & \dots & \gamma_{ik} & \dots & \gamma_{ip}\\
\vdots & & \vdots & & \vdots\\
\gamma_{m1} & \dots & \gamma_{mk} & \dots & \gamma_{mp}
\end{pmatrix}
\end{array}$$

Merke: $\gamma_{ik} = (i\text{-te Zeile von } A) \cdot (k\text{-te Spalte von } B)$

Beispiel.

$$\begin{pmatrix}1&2&1\\-1&0&2\end{pmatrix}\begin{pmatrix}-1&1\\1&1\\2&-1\end{pmatrix}=\begin{pmatrix}3&2\\5&-3\end{pmatrix}$$

$$\begin{pmatrix}-1&1\\1&1\\2&-1\end{pmatrix}\begin{pmatrix}1&2&1\\-1&0&2\end{pmatrix}=\begin{pmatrix}-2&-2&1\\0&2&3\\3&4&0\end{pmatrix}$$

Beachte.

- Das Beispiel zeigt, dass in der Regel AB verschieden von BA ist.
- Das Produkt zweier Matrizen A und B ist nur definiert, wenn die Anzahl Spalten von A mit der Anzahl Zeilen von B übereinstimmt.
- Quadratischer Matrizen derselben Größe können immer multipliziert werden.

Spezialfälle. Multiplikation einer $m \times n$-Matrix A mit einem Vektor $\mathbf{x} \in \mathbb{R}^n$:

$$A\mathbf{x} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mj} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} \quad \text{mit } y_i = \sum_{j=1}^{n} \alpha_{ij} x_j.$$

Damit haben wir eine kurze Schreibweise für Gleichungssystem (Z) in 1.3.1 gefunden, nämlich

$$A\mathbf{x} = \mathbf{y},$$

und wir haben eine sehr komprimierte Schreibweise für das System (L) linearer Gleichungen in 1.3.1:

$$A\mathbf{x} = \mathbf{b} \qquad \text{wobei } \mathbf{b} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Das Problem, das Gleichungssystem (L) zu lösen, kann nun wie folgt ausgedrückt werden: für eine gegebene $m \times n$-Matrix A und einen gegebenen Vektor $\mathbf{b} \in \mathbb{R}^m$, finde alle $\mathbf{x} \in \mathbb{R}^n$ mit $A\mathbf{x} = \mathbf{b}$. Das Gauß-Jordan Eliminations-Verfahren besteht im Wesentlichen darin, die Matrix $(A|\mathbf{b})$, d.h. die Matrix A erweitert um $\mathbf{b}$ als zusätzliche Spalte, durch elementare Zeilenumformungen in eine Matrix in Stufenform zu bringen.

1.4.4 Motivation für das Matrixprodukt. Angenommen $y_1, \dots, y_m$ hängen linear von $x_1, \dots, x_n$ wie in (Z) ab, siehe 1.3.1, d.h. es gibt Skalare α_{ij}, so dass

$$y_i = \sum_{j=1}^{n} \alpha_{ij} x_j$$

für $i = 1, \dots, m$. Verwenden wir die $m \times n$-Matrix $A = (\alpha_{ij})$, $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$, so kann dies geschrieben werden als

$$\mathbf{y} = A\mathbf{x}.$$

Nehmen wir weiter an, dass die $x_1, \dots, x_n$ linear von $t_1, \dots, t_p$ abhängen, d.h. es gibt Skalare β_{jk}, so dass

$$x_j = \sum_{k=1}^{p} \beta_{jk} t_k$$

für $j = 1, \ldots, n$. Verwenden wir die $n \times p$-Matrix $B = (\beta_{jk})$ und den Vektor $\mathbf{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_p \end{pmatrix}$, so können wir dies in der Form

$$\mathbf{x} = B\mathbf{t}$$

schreiben. Die Frage ist: Wie hängen $y_1, \ldots, y_m$ von $t_1, \ldots, t_p$ ab? Verwenden wir obige Gleichungen, erhalten wir

$$\begin{aligned} y_i &= \sum_{j=1}^{n} \alpha_{ij} x_j \\ &= \sum_{j=1}^{n} \alpha_{ij} \Big(\sum_{k=1}^{p} \beta_{jk} t_k \Big) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{p} \alpha_{ij} \beta_{jk} t_k \\ &= \sum_{k=1}^{p} \sum_{j=1}^{n} \alpha_{ij} \beta_{jk} t_k \\ &= \sum_{k=1}^{p} \Big(\sum_{j=1}^{n} \alpha_{ij} \beta_{jk} \Big) t_k . \end{aligned}$$

Setzen wir $\gamma_{ik} := \sum_{j=1}^{n} \alpha_{ij}\beta_{jk}$ und verwenden die $m \times p$-Matrix $C = (\gamma_{ik})$ so sehen wir, dass

$$\mathbf{y} = C\mathbf{t}.$$

Beachte, dass $C = A \cdot B$ gilt, die Produktmatrix wie in 1.4.3 definiert, d.h.

$$\mathbf{y} = (AB)\mathbf{t}.$$

Dies verdeutlicht, dass es Sinn macht, das Produkt zweier Matrizen so zu definieren, wie wir es oben getan haben.

Des Weiteren stellen wir fest, dass $\mathbf{y} = A\mathbf{x} = A(B\mathbf{t})$. Zusammen mit der letzten Gleichung zeigt dies, dass

$$(A \cdot B)\mathbf{t} = A(B\mathbf{t})$$

für beliebige $A \in \mathcal{M}(m \times n, \mathbb{R})$, $B \in \mathcal{M}(n \times p, \mathbb{R})$, $\mathbf{t} \in \mathbb{R}^p$.

1.4.5 Grundregeln für Matrixprodukte. *Seien A, A' $m \times n$-Matrizen, B, B' $n \times p$-Matrizen und C eine $p \times q$-Matrix. Dann gelten die folgenden Regeln:*

(A1) $$(AB)C = A(BC)$$

(A2) $$(A + A')B = AB + A'B$$
$$A(B + B') = AB + AB'$$

(A3) $$E_m A = A = A E_n$$

(A4) $$(\lambda A)B = \lambda(AB) = A(\lambda B)$$

Diese Regeln können durch explizites Ausrechnen bewiesen werden. Die schwierigste Regel ist das Assoziativgesetz (A1). Einen Spezialfall von (A1) haben wir oben gezeigt (für $C = \mathbf{t}$, eine $p \times 1$-Matrix). Aber wenn man genauer nachdenkt, so kann man den allgemeinen Fall aus dem Spezialfall ableiten: berechne einfach die Spalten des Produkts.

Das Produkt AB zweier Matrizen A und B ist nur definiert, wenn die Anzahl Spalten von A mit der Anzahl Zeilen von B übereinstimmt. Wenn wir uns auf quadratische $n \times n$-Matrizen A und B beschränken, dann ist das Produkt AB immer definiert und das Ergebnis wieder eine $n \times n$-Matrix. Damit haben wir für $\mathcal{M}(n \times n, \mathbb{R})$ eine Multiplikation, die mit (A1)–(A4) alle üblichen Regeln außer Kommutativität erfüllt, da im allgemeinen AB ungleich BA ist. Eigenschaft (A3) liest sich wie folgt: Sei $E := E_n$, dann gilt

$$AE = A = EA \quad \text{für alle } A \in \mathcal{M}(n \times n, \mathbb{R}).$$

Das bedeutet, dass die Einheitsmatrix E die Rolle eines neutralen Elements bzgl. der Matrizenmultiplikation spielt, vergleichbar der „1“ bei den reellen Zahlen.

1.4.6. Wir wollen nochmal die Nützlichkeit von Matrizen verdeutlichen. Wir erinnern uns, ein System linearer Gleichungen

$$A\mathbf{x} = \mathbf{b}$$

heißt *homogen* [*homogeneous*], falls $\mathbf{b} = \mathbf{0}$. Andernfalls nennt man es *inhomogen* [*inhomogeneous*].

Proposition. *Sei $\mathbf{x}_s$ eine spezielle Lösung des inhomogenen Systems*

$$\text{(L)} \qquad A\mathbf{x} = \mathbf{b}.$$

Ist $\mathbf{x}_h$ eine Lösung des homogenen Systems

$$\text{(H)} \qquad A\mathbf{x} = \mathbf{0},$$

so ist $\mathbf{x}_h + \mathbf{x}_s$ eine Lösung des inhomogenen Systems (L). *Und umgekehrt, zu jeder Lösung x des inhomogenen Systems* (L) *gibt es eine Lösung $\mathbf{x}_h$ des homogenen Systems* (H)*, so dass $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_s$.*

Beweis. Sei $\mathbf{x}_s$ eine spezielle Lösung von (L).

(1) Ist $\mathbf{x}_h$ eine Lösung von (H), dann gilt

$$A\mathbf{x}_h = \mathbf{0}$$
$$A\mathbf{x}_s = \mathbf{b}.$$

Addieren wir beide Gleichungen, erhalten wir $A\mathbf{x}_h + A\mathbf{x}_s = \mathbf{b}$. Da $A\mathbf{x}_h + A\mathbf{x}_s = A(\mathbf{x}_h + \mathbf{x}_s)$ nach (A2), folgern wir $A(\mathbf{x}_h + \mathbf{x}_s) = \mathbf{b}$, d.h. $\mathbf{x} := \mathbf{x}_h + \mathbf{x}_s$ ist eine Lösung von (L).

(2) Ist $\mathbf{x}$ irgendeine Lösung von (L), dann gilt

$$A\mathbf{x} = \mathbf{b}$$
$$A\mathbf{x}_s = \mathbf{b}.$$

Subtrahieren wir beide Gleichungen, so erhalten wir $A\mathbf{x} - A\mathbf{x}_s = \mathbf{0}$. Wieder mit (A2) gilt $A\mathbf{x} - A\mathbf{x}_s = A(\mathbf{x} - \mathbf{x}_s)$. Und damit $A(\mathbf{x} - \mathbf{x}_s) = \mathbf{0}$, d.h. $\mathbf{x}_h := \mathbf{x} - \mathbf{x}_s$ ist eine Lösung von (H) und folglich $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_s$. □

Merkregel. Die „allgemeine Lösung“ des inhomogene Gleichungssystems ist die Summe einer speziellen Lösung des inhomogenen Systems und der „allgemeinen Lösung“ des homogenen Systems.

1.4.7 Invertierbare Matrizen. Wir betrachten jetzt nur quadratische $n\times n$-Matrizen. Sei $E := E_n$.

Eine Matrix $A \in \mathcal{M}(n \times n, \mathbb{R})$ heißt *invertierbar* [*invertible*], falls es eine Matrix $B \in \mathcal{M}(n \times n, \mathbb{R})$ gibt, so dass

$$AB = E = BA. \tag{I}$$

Andernfalls wird A *singulär* [*singular*] genannt.

Invertierbare Matrizen werden auch *nicht singulär* [*non-singular*] oder *regulär* [*regular*] genannt.

Wir bezeichnen mit $\mathrm{GL}(n, \mathbb{R})$ die Menge aller invertierbaren $n \times n$-Matrizen mit reellen Einträgen.

Beachte. Ist A invertierbar, dann ist die Matrix B mit Eigenschaft (I) eindeutig bestimmt. Denn wenn B' eine weitere Matrix mit $AB' = E = B'A$ ist, dann gilt

$$B = BE = B(AB') = (BA)B' = EB' = B'.$$

Wir nennen B *die Inverse* von A und schreiben $B = A^{-1}$.

Eigenschaften invertierbarer Matrizen.

(1) *E ist invertierbar und es gilt $E = E^{-1}$.*

(2) *Falls A invertierbar ist, dann ist auch A^{-1} invertierbar und es gilt $(A^{-1})^{-1} = A$.*

(3) *Falls A und B invertierbar sind, dann ist auch AB invertierbar und es gilt*
$$(AB)^{-1} = B^{-1}A^{-1}.$$

Beweis.

(1) $E \cdot E = E = E \cdot E$.

(2) $A^{-1} \cdot A = E = A \cdot A^{-1}$.

(3) $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$ und
$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E$. □

Beispiele.

- Die Nullmatrix 0 ist nicht invertierbar.
- Sei $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Dann gilt $AA = E$. Das heißt A ist invertierbar und $A^{-1} = A$.
- Falls eine Zeile (oder eine Spalte) von A nur aus Nullen besteht, dann ist A nicht invertierbar.

  (Denn ist zum Beispiel die erste Spalte von A der Nullvektor $\mathbf{0}$, dann ist die erste Spalte von BA ebenfalls der Nullvektor $\mathbf{0}$ und damit gilt $BA \neq E$ für jede Matrix B.)

- Die 2×2-Matrix $A=\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ ist invertierbar genau dann, wenn $\alpha\delta-\gamma\beta\neq 0$. In diesem Fall ist

$$A^{-1}=(\alpha\delta-\gamma\beta)^{-1}\begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}$$

  (Überprüfen Sie das!)

Für $n\times n$-Matrizen gibt es keine explizite Formel für die Inverse, sofern sie existiert. Von daher fragen wir uns: Gibt es (wenigstens) einen Algorithmus, der herausfindet, ob eine gegebene $n\times n$-Matrix A invertierbar ist, und die Inverse bestimmt, sofern sie existiert? Den Schlüssel zu diesem Problem liefert uns folgendes Lemma:

Lemma. *Seien A und B zwei $n\times n$-Matrizen. Dann ist B die Inverse von A genau dann, wenn*

$$(*)\qquad A\mathbf{x}=\mathbf{y} \iff \mathbf{x}=B\mathbf{y} \quad \textit{für alle } \mathbf{x},\mathbf{y}\in\mathbb{R}^n.$$

Beweis. Nehmen wir zuerst an, dass $B=A^{-1}$. Dann impliziert $A\mathbf{x}=\mathbf{y}$, dass $\mathbf{x}=E\mathbf{x}=BA\mathbf{x}=B\mathbf{y}$ und $\mathbf{x}=B\mathbf{y}$ impliziert, dass $A\mathbf{x}=A(B\mathbf{y})=(AB)\mathbf{y}=E\mathbf{y}=\mathbf{y}$. Also gilt $(*)$.

Nehmen wir nun an, dass $(*)$ gilt. Sei $\mathbf{x}\in\mathbb{R}^n$ beliebig. Definiere $\mathbf{y}:=A\mathbf{x}$. Aus $(*)$ erhalten wir $B\mathbf{y}=\mathbf{x}$. Ersetzen wir in dieser Gleichung $\mathbf{y}$ durch $A\mathbf{x}$, erhalten wir $BA\mathbf{x}=\mathbf{x}=E\mathbf{x}$. Da dies für alle $\mathbf{x}\in\mathbb{R}^n$ gilt, folgern wir $BA=E$. Analog zeigt man $AB=E$. Also ist B die Inverse von A. □

Dies legt uns folgende Vorgehensweise zur Berechnung der Inversen nahe: Sei $A\in\mathcal{M}(n\times n,\mathbb{R})$ gegeben. Betrachte das lineare Gleichungssystem

$$(1)\qquad A\mathbf{x}=\mathbf{y}.$$

Versuche dieses System nach $\mathbf{x}$ aufzulösen und versuche $\mathbf{x}$ durch $\mathbf{y}$ auszudrücken:

$$(2)\qquad \mathbf{x}=B\mathbf{y}.$$

Wenn der Eliminationsprozess so durchgeführt werden kann, dass das zweite System zum ersten äquivalent ist, dann ist A invertierbar und $A^{-1}=B$ nach obigem Lemma.

Wir haben gesehen, dass die elementaren Umformungen (E1)–(E3) ein lineares Gleichungssystem in ein äquivalentes überführen. Wir versuchen nun, eine Erweiterung des Gauß-Jordan Eliminations-Algorithmus anzuwenden, um (1) in (2) zu überführen:

Wenn die Gauß-Jordan Elimination eine Zeile erzeugt, in der alle Koeffizienten der linken Seite Null sind, dann ist A nicht invertierbar. Andernfalls ist A invertierbar und $\mathbf{x}=B\mathbf{y}$ liefert $B=A^{-1}$.

Wir wollen dies an zwei Beispielen illustrieren.

1.4.8 Berechnung der Inversen.

Beispiel 1.

$$A=\begin{pmatrix}1 & 1 & 1\\ 2 & 2 & 4\\ 2 & 1 & 1\end{pmatrix}$$

Nun versuchen wir das System $A\mathbf{x} = \mathbf{y}$ nach $\mathbf{x}$ aufzulösen:

	A			E		
$x_1 + x_2 + x_3 = y_1$	1	1	1	1	0	0
$2x_1 + 2x_2 + 4x_3 = y_2$	2	2	4	0	1	0
$2x_1 + x_2 + x_3 = y_3$	2	1	1	0	0	1
$x_1 + x_2 + x_3 = y_1$	1	1	1	1	0	0
$+2x_3 = -2y_1 + y_2$	0	0	2	-2	1	0
$-x_2 - x_3 = -2y_1 + y_3$	0	-1	-1	-2	0	1
$x_1 + x_2 + x_3 = y_1$	1	1	1	1	0	0
$-x_2 - x_3 = -2y_1 + y_3$	0	-1	-1	-2	0	1
$2x_3 = -2y_1 + y_2$	0	0	2	-2	1	0
$x_1 = -y_1 + y_3$	1	0	0	-1	0	1
$-x_2 - x_3 = -2y_1 + y_3$	0	-1	-1	-2	0	1
$2x_3 = -2y_1 + y_2$	0	0	2	-2	1	0
$x_1 = -y_1 + y_3$	1	0	0	-1	0	1
$x_2 + x_3 = 2y_1 - y_3$	0	1	1	2	0	-1
$x_3 = -y_1 + \frac{1}{2}y_2$	0	0	1	-1	$\frac{1}{2}$	0
$x_1 = -y_1 + y_3$	1	0	0	-1	0	1
$x_2 = 3y_1 - \frac{1}{2}y_2 - y_3$	0	1	0	3	$-\frac{1}{2}$	-1
$x_3 = -y_1 + \frac{1}{2}y_2$	0	0	1	-1	$\frac{1}{2}$	0
	E			A^{-1}		

Dies zeigt, dass A invertierbar ist und

$$A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 3 & -\frac{1}{2} & -1 \\ -1 & \frac{1}{2} & 0 \end{pmatrix}$$

Beachte, dass das System $A\mathbf{x} = \mathbf{b}$ eine eindeutige Lösung hat, wenn A invertierbar ist, nämlich $\mathbf{x} = A^{-1}\mathbf{b}$.

Löse zum Beispiel

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 + 2x_2 + 4x_3 &= -1 \\ 2x_1 + x_2 + x_3 &= 2 \end{aligned}$$

d.h.

$$A\mathbf{x} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 2 & 1 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Die Lösung ist

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} -1 & 0 & 1 \\ 3 & -\frac{1}{2} & -1 \\ -1 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{pmatrix}.$$

Beispiel 2.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 1 & 1 & -1 \end{pmatrix}$$

Das System $A\mathbf{x} = \mathbf{y}$

$$\begin{aligned} x_1 + x_2 + x_3 &= y_1 \\ 2x_1 + 2x_2 + 4x_3 &= \phantom{y_1 +{}} y_2 \\ x_1 + x_2 - x_3 &= \phantom{y_1 + y_2 +{}} y_3 \end{aligned}$$

wird durch elementare Umformungen übergeführt in

$$\begin{aligned} x_1 + x_2 + x_3 &= y_1 \\ 2x_3 &= -2y_1 + y_2 \\ 0 &= -3y_1 + y_2 + y_3 \end{aligned}$$

und dieses System kann nicht übergeführt werden in die Form $\mathbf{x} = B\mathbf{y}$ mit elementaren Umformungen. Also ist A *nicht* invertierbar.

A			E		
1	1	1	1	0	0
2	2	4	0	1	0
1	1	−1	0	0	1
1	1	1	1	0	0
0	0	2	−2	1	0
0	0	2	−1	0	1
1	1	1	1	0	0
0	0	2	−2	1	0
0	0	0	−3	1	1

1.4.9 Die Transponierte einer Matrix. Die *Transponierte* [*transpose*] einer $m \times n$-Matrix

$$A = \begin{array}{rc} & \begin{array}{cccccc} \phantom{\alpha_{11}} & \phantom{\alpha_{12}} & & j & & \phantom{\alpha_{1n}} \end{array} \\ i \rightarrow & \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mj} & \dots & \alpha_{mn} \end{pmatrix} \end{array}$$

ist definiert als die $n \times m$-Matrix

$$A^T := \begin{array}{rc} & \begin{array}{ccccc} \phantom{\alpha_{11}} & & i & & \phantom{\alpha_{m1}} \end{array} \\ j \rightarrow & \begin{pmatrix} \alpha_{11} & \dots & \alpha_{i1} & \dots & \alpha_{m1} \\ \alpha_{12} & \dots & \alpha_{i2} & \dots & \alpha_{m2} \\ \vdots & & \vdots & & \vdots \\ \alpha_{1j} & \dots & \alpha_{ij} & \dots & \alpha_{mj} \\ \vdots & & \vdots & & \vdots \\ \alpha_{1n} & \dots & \alpha_{in} & \dots & \alpha_{mn} \end{pmatrix} \end{array}$$

Die i-te Zeile von A wird die i-te Spalte von A^T, die j-te Spalte von A wird die j-te Zeile A^T.

Regeln für die Transponierte. *Für alle $m \times n$-Matrizen A und A' und jede $n \times p$-Matrix B gilt*

(T1) $$(A^T)^T = A$$

(T2) $$(A + A')^T = A^T + (A')^T$$

(T3) $$(\lambda A)^T = \lambda A^T$$

(T4) $$(AB)^T = B^T A^T$$

Beachte, dass die Transponierte eines Produkts von Matrizen das Produkt der Transponierten Matrizen ist, nur in umgekehrter Reihenfolge (vergleiche dazu auch die Inverse eines Produkts).

Übung. Zeige, dass für jede invertierbare $n \times n$-Matrix A, auch A^T invertierbar ist und bestimme $(A^T)^{-1}$.

KAPITEL 2

Vektorräume

Wir kommen nun zu einer systematischeren Behandlung von Vektorräumen. In diesem Kapitel wollen wir einige grundlegende Begriffe und Eigenschaften darstellen. Einiges hiervon wurde bereits in der Analysis-Vorlesung erläutert. Insbesondere wollen wir voraussetzen, wie man mit Mengen M, N, $\ldots$, direkten Produkten $M \times N$ von Mengen, Schnittmengen $M \cap N$, Vereinigungsmengen $M \cup N$ und Funktionen $f\colon M \to N$ etc. umgeht. Dabei sagen wir häufig *Abbildung* statt *Funktion.*Mathematisch ist genau dasselbe gemeint.

2.1. Algebraische Strukturen

Einige arithmetische Operationen wie die Addition und die Multiplikation auf bestimmten Mengen wie den natürlichen oder den reellen Zahlen sind uns bereits vertraut. Hier wollen wir diese Operationen von einem allgemeineren Standpunkt aus betrachten.

2.1.1 Binäre Verknüpfungen. Eine *(binäre) Verknüpfung* auf einer Menge M ist eine Abbildung

$$\begin{aligned} M \times M &\xrightarrow{*} M \\ (a,b) &\mapsto a * b, \end{aligned}$$

die jedem Paar von Elementen (a,b) aus M ein weiteres Element $a * b$ aus M zuordnet.

Anstelle von $*$ werden oftmals andere Symbole zur Darstellung einer Verknüpfung verwendet, wie zum Beispiel

additive Scheibweise: $a + b$

multiplikative Schreibweise: $a \cdot b$ oder einfach ab

Beispiele von Verknüpfungen.

(1) Die gewöhnliche Addition auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathcal{M}(m \times n, \mathbb{R})$, $\mathbb{R}^n$.

(2) Die gewöhnliche Multiplikation auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathcal{M}(n \times n, \mathbb{R})$.

Folgende sind *keine* Verknüpfungen im obigen Sinne:

(3) Die skalare Multiplikation im $\mathbb{R}^n$:

$$\begin{aligned} \mathbb{R} \times \mathbb{R}^n &\to \mathbb{R}^n \\ (\lambda, \mathbf{v}) &\mapsto \lambda \mathbf{v}. \end{aligned}$$

Wir verbinden hier Objekte verschiedenen Typs: Skalare und Vektoren, was in unserer Definition einer Verknüpfung nicht enthalten ist.

(4) Die Subtraktion $(a, b) \mapsto a - b$ ist keine Verknüpfung auf $\mathbb{N}$, da sie aus dem Bereich der natürlichen Zahlen hinsusführt. Sie ist aber eine Verknüpfung auf $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{R}^n$, $\mathcal{M}(m \times n, \mathbb{R})$.

2.1.2 Assoziative Verknüpfungen – Halbgruppen. Eine Verknüpfung $*$ auf M heißt *assoziativ*, wenn gilt:

(G1) $$a * (b * c) = (a * b) * c \quad \text{für alle } a, b, c \in M.$$

Ein Paar $(M, *)$ bestehend aus einer Menge M zusammen mit einer assoziativen Verknüpfung $*$ auf M wird *Halbgruppe* genannt.

Es gibt beispielsweise fünf Möglichkeiten, ein Produkt aus vier Faktoren zu bilden:

$$\begin{array}{lll} ((a * b) * c) * d & (a * b) * (c * d) & a * (b * (c * d)) \\ (a * (b * c)) * d & & a * ((b * c) * d) \end{array}$$

Bei einer assoziativen Verknüpfung $*$ sind alle diese fünf Produkte gleich (Beweis?). Mittels Induktion kann man zeigen, dass in einem Produkt von n Elementen das Ergebnis unabhängig von der Klammersetzung ist. Bei einer assoziativen Verknüpfung kann man also die Klammern weglassen.

Potenzen. Ist $*$ eine assoziative Verknüpfung auf M, dann können wir die n-te *Potenz* a^n für jedes $a \in M$ und jede natürliche Zahl $n \geq 1$ wie folgt definieren:

$$a^n := \underbrace{a * a * \cdots * a}_{n \text{ mal}}$$

Die Gesetze

$$a^n * a^m = a^{n+m} \qquad\qquad (a^n)^m = a^{n \cdot m}$$

gelten für alle $n, m \geq 1$.

Benutzt man die additive Schreibweise, spricht man von einem Vielfachen (anstelle von einer Potenz), geschrieben als

$$n \cdot a := \underbrace{a + a + \cdots + a}_{n \text{ mal}}$$

Es ergeben sich die folgenden Rechenregeln:

$$n \cdot a + m \cdot a = (n + m) \cdot a \qquad\qquad m \cdot (n \cdot a) = (mn) \cdot a \; .$$

2.1.3 Neutrale Elemente. Ein Element $e \in M$ heißt *neutrales Element* (für die Verknüpfung $*$ auf M), wenn

(G2) $$e * a = a = a * e \quad \text{für alle } a \in M.$$

Es muss kein neutrales Element existieren, gibt es jedoch ein solches, dann ist es auch eindeutig: wenn e und e' neutrale Elemente sind, folgt nämlich $e = e * e' = e'$.

In multiplikativer Schreibweise wird ein neutrales Element oft als 1 dargestellt, in additiver Schreibweise immer als 0.

2.1.4 Invertierbare Elemente. Sei M eine Menge mit einer assoziativen Verknüpfung $*$ und einem Neutralelement e.

Ein Element $a \in M$ heißt *invertierbar*, wenn es ein Element $b \in M$ mit der Eigenschaft

$$a * b = e = b * a.$$

gibt. Ein solches Element b heißt dann *Inverses* von a und wird mit a^{-1} bezeichnet.

Bemerkung. Das Inverse eines invertierbaren Elementes ist eindeutig bestimmt: Sind nämlich b und b' Inverse von a. Dann folgt

$$b \underset{(G2)}{=} b * e = b * (a * b') \underset{(G1)}{=} (b * a) * b' = e * b' \underset{(G2)}{=} b'.$$

Damit ist die Schreibweise a^{-1} für das Inverse gerechtfertigt.

Regeln.

(1) *e ist invertierbar und es gilt $e^{-1} = e$.*

(2) *Ist a invertierbar, dann ist auch a^{-1} invertierbar und es gilt $(a^{-1})^{-1} = a$.*

(3) *Sind a und b invertierbar, so ist auch $a * b$ invertierbar und es gilt $(a * b)^{-1} = b^{-1} * a^{-1}$.*

Beweis.

(1) Da $e * e = e$ aus (G2) folgt, ist e invertierbar und gleich seinem Inversen.

(2) Wenn a invertierbar ist, dann gilt $a * a^{-1} = e = a^{-1} * a$, was zeigt, dass a^{-1} invertierbar ist und dass das Inverse von a^{-1} das Element a ist.

(3) a und b seien invertierbar. Es folgt

$$(a * b) * (b^{-1} * a^{-1}) = a * (b * b^{-1}) * a^{-1} = a * e * a^{-1} = a * a^{-1} = e$$
$$(b^{-1} * a^{-1}) * (a * b) = b^{-1} * (a^{-1} * a) * b = b^{-1} * e * b = b^{-1} * b = e.$$

Also ist auch $a * b$ invertierbar mit der Inversen $b^{-1} * a^{-1}$. □

Bemerkung. In der additiven Schreibweise ist ein Element a invertierbar, falls es ein Element b gibt mit $a + b = 0 = b + a$. Das Inverse von a wird dann mit $-a$ anstelle von a^{-1} bezeichnet.

2.1.5 Gruppen. Unter einer *Gruppe* verstehen wir eine Menge G mit einer assoziativen Verknüpfung $*$ und einem Neutralelement e, in welcher jedes Element invertierbar ist.

Eine Gruppe ist also eine Menge G mit einer Verknüpfung

$$G \times G \xrightarrow{*} G$$
$$(a, b) \mapsto a * b$$

und einem Element $e \in G$ so, daß die folgenden Eigenschaften erfüllt sind:

(G1) $\forall a, b, c \in G : a * (b * c) = (a * b) * c$

(G2) $\forall a \in G : a * e = a = e * a$

(G3) $\forall a \in G : \ \exists b \in G : a * b = e = b * a.$

Beispiele. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}^n, +)$ und $(\mathcal{M}(m \times n, \mathbb{R}), +)$ sind Gruppen.

2.1.6 Kommutativität. Eine Verknüpfung $*$ auf einer Menge M heißt *kommutativ*, falls gilt:

$$a * b = b * a \quad \text{für alle } a, b \in M.$$

Eine Halbgruppe bzw. Gruppe mit einer kommutativen Verknüpfung wird kommutativ genannt. Eine kommutative Gruppe nennt man oft auch eine *abelsche Gruppe*.

Bemerkungen.

- Die Gruppen in den obigen Beispielen sind alle abelsch. Die Multiplikation von Matrizen ist aber nicht kommutativ.
- Die additive Schreibweise wird ausschließlich bei kommutativen Verknüpfungen benutzt.

2.1.7 Theorem. *Es sei $(M, *)$ eine Halbgruppe mit einem neutralen Element e. Wir bezeichnen mit $M^\times$ die Menge der invertierbaren Elemente von M. Dann ist $M^\times$ eine Gruppe mit neutralem Element e bzgl. der Einschränkung der Verknüpfung $*$ auf M^*.*

Beweis. Für $a, b \in M^\times$ liegt das Produkt $a * b$ auch in $M^\times$, wie 2.1.4 (3) zeigt. Also ergibt die Einschränkung von $*$ tatsächlich eine Verknüpfung $M^\times \times M^\times \to M^\times$.

Da die Verknüpfung $*$ auf M assoziativ ist, ist ihre Einschränkung auf $M^\times$ ebenfalls assoziativ, woraus (G1) folgt. Nach 2.1.4 (1) gilt $e \in M^\times$ und natürlich ist e auch ein neutrales Element für $M^\times$, was (G2) zeigt. Außerdem ist jedes $a \in M^\times$ nach Definition invertierbar in M und sein Inverses a^{-1} gehört nach 2.1.4 (2) auch zu $M^\times$. Damit haben wir auch (G3) gezeigt. □

Beispiele.

(1) $(\mathbb{Z}, \cdot)$ ist eine Halbgruppe mit Neutralelement 1. Die einzigen (multiplikativ) invertierbaren Elemente sind $+1$ und -1. Damit ist $\mathbb{Z}^\times = \{+1, -1\}$ und dies ist eine multiplikative Gruppe.

(2) Jede von 0 verschiedene rationale Zahl ist invertierbar bzgl. der Multiplikation. Also is die Menge $\mathbb{Q}^\times$ aller von 0 verschiedenen rationalen Zahlen eine multiplikative Gruppe. Dasselbe gilt, wenn man $\mathbb{Q}$ durch $\mathbb{R}$ ersetzt.

(3) Die Menge $\mathcal{M}(n \times n, \mathbb{R})$ aller $n \times n$-Matrizen mit reellen Einträgen ist eine Halbgruppe bzgl. der Multiplikation von Matrizen, die Einheitsmatrix $E = E_n$ ist das Neutralelement. Aufgrund des obigen Satzes bilden die invertierbaren reellen $n \times n$-Matrizen eine Gruppe bzgl. der Multiplikation von Matrizen. Wir bezeichnen wir diese Gruppe mit $\mathrm{GL}(n, \mathbb{R})$ und nennen sie die *allgemeine lineare Gruppe* über $\mathbb{R}$. Diese Gruppe ist für $n \geq 2$ nicht kommutativ.

In vielen unserer Beispiele hatten wir zwei Verknüpfungen auf derselben Menge: eine Addition und eine Multiplikation.

2.1.8 Ringe. Eine Menge R mit zwei Verknüpfungen $+$ und $\cdot$ wird ein *Ring* genannt, falls

(1) $(R,+)$ eine kommutative Gruppe mit Neutralelement 0 ist,

(2) $(R,\cdot)$ eine Halbgruppe mit Neutralelement 1 ist und

(3) die folgenden *Distributivgesetze* gelten:

$$\forall a,b,c \in R.\ a\cdot(b+c) = a\cdot b + a\cdot c$$
$$\forall a,b,c \in R.\ (b+c)\cdot a = b\cdot a + c\cdot a.$$

Ein Ring wird *kommutativ* genannt, wenn seine Multiplikation ebenfalls kommutativ ist.

Beispiele. $(\mathbb{Z},+,\cdot)$ ist ein kommutativer Ring. $(\mathcal{M}(n\times n,\mathbb{R}),+,\cdot)$ ist ein Ring, der nicht kommutativ ist.

Übung. In einem Ring R gilt $a\cdot 0 = 0 = 0\cdot a$ für alle $a\in R$.

In einem Ring kann allerdings von 0 verschiedene Elemente a,b geben, so daß $a\cdot b=0$ gilt. Beispielsweise gilt in $\mathcal{M}(2\times 2,\mathbb{R})$: $\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)\cdot\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)=\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$.

2.1.9 Körper. Einen kommutativen Ring K mit $0\neq 1$, in dem jedes von 0 verschiedene Element ein multiplikatives Inverses besitzt, nennen wir einen *Körper* [*field*].

Beispiele von Körpern. $\mathbb{Q}$, $\mathbb{R}$ mit der gewöhnlichen Addition und Multiplikation. $\mathbb{Z}$ ist kein Körper.

Bemerkungen.

(1) In einem Körper K bilden die von 0 verschiedenen Elemente eine multiplikative Gruppe $K^\times = K\setminus\{0\}$ (siehe 2.1.7).

(2) Es gibt in einem Körper keine Nullteiler, d.h. aus $ab=0$ folgt $a=0$ oder $b=0$ (Beweis?).

In einem Körper können wir Elemente addieren, subtrahieren, multiplizieren und durch von 0 verschiedene Elemente dividieren, und für diese Operationen gelten die von den reellen Zahlen her vertrauten Rechenregeln.

Zusammenfassung. Wir fassen die definierenden Eigenschaften eines Körpers zusammen:

Ein *Körper* ist eine Menge K mit zwei Verknüpfungen $+$ und $\cdot$ und zwei Elementen $0\neq 1$, so daß die folgenden Gesetze gelten:

(K1)	$\forall a,b,c\in K: a+(b+c)=(a+b)+c$	$a(bc)=(ab)c$
(K2)	$\forall a\in K: a+0=a$	$a\cdot 1 = a$
(K3)	$\forall a\in K:\ \exists b\in K: a+b=0$	$a\neq 0 \implies \exists c\in K:\ ac=1$
(K4)	$\forall a,b\in K: a+b=b+a$	$ab=ba$
(K5)	$\forall a,b,c\in K: a(b+c)=ab+ac$	

2.2. Der Körper der komplexen Zahlen

Komplexe Zahlen spielen sowohl in der Linearen Algebra als auch in der Analysis eine bedeutende Rolle. In dieser Vorlesung sind komplexe Zahlen deshalb von so hoher Bedeutung, da sie einen Körper $\mathbb{C}$ bilden, der den Körper der reellen Zahlen dergestalt erweitert, dass jedes nichtkonstante Polynom mit reellen oder komplexen Koeffizienten eine komplexe Wurzel besitzt. Diese Tatsache nennt man den *Fundamentalsatz der Algebra.*

Im Körper der reellen Zahlen sind alle Quadrate nichtnegativ. Es gibt also insbesondere keine reelle Zahl a mit $a^2 = -1$. Dies ist damit gleichbedeutend, dass die Gleichung

$$x^2 = -1$$

keine Lösung in $\mathbb{R}$ besitzt.

Aus diesem Grund wollen wir eine neue Zahl i, eine *imaginäre Zahl* einführen, für die $i^2 = -1$ gilt. Dabei ist noch nicht klar, dass dies im Einklang mit den üblichen Rechenregeln geschehen kann. Wir werden wie folgt vorgehen: Wir nehmen an, es gibt einen Körper $\mathbb{K}$, der den Körper $\mathbb{R}$ der reellen Zahlen umfaßt und ein Element i mit der Eigenschaft $i^2 = -1$ enthält. Wir werden Folgerungen für das Rechnen in einem solchen Körper herleiten. Dann werden wir erkennen, wie man einen solchen Körper explizit konstruieren kann.

2.2.1 Erster Ansatz zur Konstruktion der komplexen Zahlen. Wir nehmen an, es gibt ein System K von Zahlen, das sowohl den Körper $\mathbb{R}$ der reellen Zahlen als auch ein Element i mit

$$i^2 = -1$$

enthält.

Wir nehmen für dieses Zahlensystem K die Existenz einer Addition und einer Multiplikation an, auf $\mathbb{R}$ mit der üblichen Addition und Multiplikation übereinstimmen und welche auf K die Eigenschaften eines kommutativen Rings erfüllen.

In K können wir i mit einer reellen Zahl β multiplizieren. Danach können wir eine reelle Zahl α zu dem Ergebnis addieren, d.h. wir können Elemente der Form

$$z = \alpha + i\beta \ , \qquad \alpha, \beta \in \mathbb{R}$$

bilden. Elemente in dieser Darstellung nennen wir *komplexe Zahlen* und wir bezeichnen mit $\mathbb{C}$ die Menge all dieser komplexen Zahlen.

Man beachte, daß reelle Zahlen spezielle komplexe Zahlen sind: $\alpha = \alpha + i \cdot 0$. Die komplexen Zahlen der Form $i\beta, \beta \in \mathbb{R}$, heißen *rein imaginär.*

In dem kommutativen Ring K können wir zwei komplexe Zahlen $z_1 = \alpha_1 + i\beta_1$ und $z_2 = \alpha_2 + i\beta_2$ addieren und multiplizieren. Aufgrund der in einem kommutativen Ring geltenden Gesetze und der Gleichung $i^2 = -1$ erhalten wir:

$$z_1 \pm z_2 = (\alpha_1 \pm \alpha_2) + i(\beta_1 \pm \beta_2) \ ,$$

$$z_1 \cdot z_2 = (\alpha_1 + i\beta_1)(\alpha_2 + i\beta_2) = \alpha_1\alpha_2 + i\beta_1\alpha_2 + i\alpha_1\beta_2 + \underbrace{i^2}_{=-1}\beta_1\beta_2 = (\alpha_1\alpha_2 - \beta_1\beta_2) + i(\beta_1\alpha_2 + \alpha_1\beta_2) \ ,$$

d.h. Summen, Differenzen und Produkte zweier komplexer Zahlen sind wieder komplexe Zahlen. Also sind die Addition, Subtraktion und Multiplikation Verknüpfungen auf der Menge $\mathbb{C}$ der komplexen Zahlen und damit ist $\mathbb{C}$ ein kommutativer Ring.

2.2.2 Gleichheit von komplexen Zahlen. Für zwei komplexe Zahlen $z_1 = \alpha_1 + i\beta_1$ und $z_2 = \alpha_2 + i\beta_2$ gilt:

$$z_1 = z_2 \quad \text{genau dann, wenn} \quad \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2 \ .$$

Ist nämlich $\alpha_1 + i\beta_1 = \alpha_2 + i\beta_2$, so folgt $\alpha_1 - \alpha_2 = i(\beta_2 - \beta_1)$. Mittels Quadrieren erhalten wir $(\alpha_1 - \alpha_2)^2 = -(\beta_2 - \beta_1)^2$ wegen $i^2 = -1$. Da auf der linken Seite eine reelle Zahl ≥ 0 und auf der rechten Seite eine reelle Zahl ≤ 0 steht, folgt dass beide Seiten verschwinden, d.h. es gilt $\alpha_1 = \alpha_2$ und $\beta_2 = \beta_1$.

2.2.3 Real- und Imaginärteil. Nach dem vorhergehenden sind für eine komplexe Zahl

$$z = \alpha + i\beta$$

die reellen Zahlen α und β eindeutig bestimmt. Wir nennen α den *Realteil* von z und β den *Imaginärteil* von z; Wir schreiben

$$\Re(z) = \alpha \ ,$$

$$\Im(z) = \beta \ .$$

Nach 2.2.3 sind zwei komplexe Zahlen z_1 und z_2 genau dann gleich, wenn $\Re(z_1) = \Re(z_2)$ und $\Im(z_1) = \Im(z_2)$. Beachte, dass auch der Imaginärteil einer komplexen Zahl eine reelle Zahl β ist.

2.2.4 Die Gaußsche Zahlenebene. Wir können die komplexe Zahl z durch das Paar (α, β) von reellen Zahlen, d.h. durch einen Punkt in $\mathbb{R}^2$ darstellen. Die erste Achse wird für den Realteil benutzt (sie heißt deswegen die *reelle Achse*), die zweite Achse für den Imaginärteil (diese heißt deswegen die *imaginäre Achse*). Wir können also die Ebene mit der Menge der komplexen Zahlen identifizieren. Wir sprechen auch von der *Gaußschen Zahlenebene*.

Die Addition von komplexen Zahlen entspricht der Addition im Vektorraum $\mathbb{R}^2$. Damit haben wir eine geometrische Darstellung der komplexen Zahlen und der Addition der komplexen Zahlen. Die geometrische Interpretation der Multiplikation von komplexen Zahlen geben wir später.

Abbildung: Die Gaußsche Zahlenebene

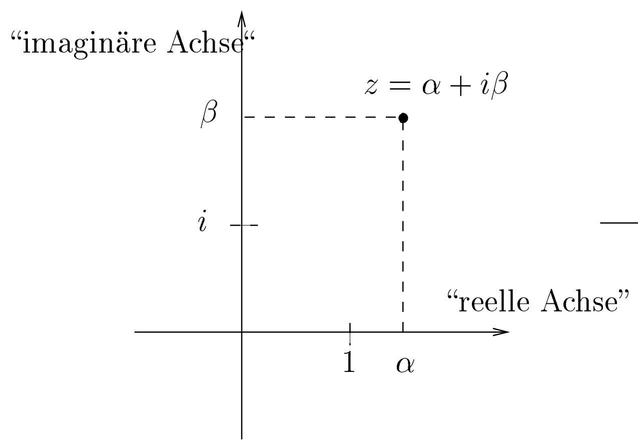
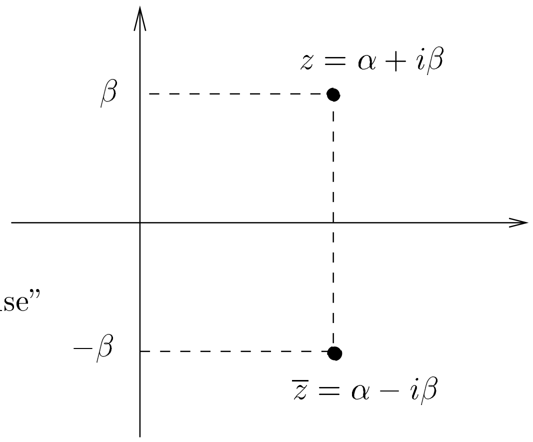


Abbildung: Die komplex konjugierte Zahl



2.2.5 Die komplex konjugierte Zahl. Die *komplex konjugierte Zahl* einer komplexen Zahl

$$z = \alpha + i\beta$$

ist definiert als die komplexe Zahl

$$\overline{z} := \alpha - i\beta \ .$$

Mittels einfacher Rechnung zeigt man die folgenden Regeln für beliebige komplexe Zahlen:

$$\begin{aligned}
\overline{\overline{z}} &= z \\
\overline{z_1 \pm z_2} &= \overline{z_1} \pm \overline{z_2} \\
\overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2} \\
\Re(z) &= \frac{1}{2}(z + \overline{z}) \\
\Im(z) &= \frac{-i}{2}(z - \overline{z})
\end{aligned}$$

2.2.6 Betrag und Argument einer komplexen Zahl. Der *Betrag* einer komplexen Zahl $z = \alpha + i\beta$ ist definiert als die reelle Zahl

$$r = |z| := \sqrt{\alpha^2 + \beta^2} = \sqrt{z \cdot \overline{z}} \ .$$

In der Darstellung von z als Punkt in der Gaußschen Zahlenebene ist der Betrag nichts anderes als der Abstand vom Ursprung. Er hat die folgenden offensichtlichen Eigenschaften:

$$\begin{aligned}
|z| &\geq 0 \ , \quad |z| = 0 \Leftrightarrow z = 0 \\
|z_1 + z_2| &\leq |z_1| + |z_2| \ \text{Dreiecksungleichung}
\end{aligned}$$

Das *Argument* einer komplexen Zahl $z \neq 0$ ist definiert als der Winkel φ, der gegen den Uhrzeigersinn zwischen der positiven rellen Achse und dem Strahl vom Ursprung durch den Punkt z gebildet wird. Das Argument von z ist nur bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt:

$$\arg(z) = \varphi + 2\pi k \quad , k \in \mathbb{Z} \; .$$

Ist $\Re(z) > 0$, d.h. ist $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$, dann ist $\tan\varphi = \frac{\beta}{\alpha}$ und wir können das Argument nach der folgenden Formel berechnen:

$$\varphi = \arctan\frac{\beta}{\alpha} = \arctan\frac{\Im(z)}{\Re(z)}.$$

Ist $\Re(z) < 0$, dann gilt

$$\varphi = \pi + \arctan\frac{\Im(z)}{\Re(z)}.$$

Viele Programmiersprachen verfügen über eine binäre Funktion (oft `atan2()` genannt), die φ aus α und β berechnet.

2.2.7 Polardarstellung von komplexen Zahlen. Jede komplexe Zahl $z = \alpha + i\beta$ kann in der Form

$$z = r(\cos\varphi + i\sin\varphi)$$

mit $r = |z|$ als Betrag und $\varphi = \arg(z)$ als Argument von z dargestellt werden; die Zahlen r und φ heißen Polarkoordinaten von z. Die obige Darstellung heißt *Polardarstellung* einer komplexen Zahl.

Die komplexen Zahlen mit Betrag $|z| = 1$ bilden den *Einheitskreis*, d.h. den Kreis mit Radius 1 um den Ursprung. Dies sind genau die komplexen Zahlen mit der Polardarstellung

$$z = \cos\varphi + i\sin\varphi \; .$$

Zwei komplexe Zahlen in Polardarstellung

$$z_1 = r_1(\cos\varphi + i\sin\varphi) \quad , \quad z_2 = r_2(\cos\psi + i\sin\psi)$$

sind genau dann gleich, wenn

$$r_1 = r_2 \quad \text{und} \quad \varphi = \psi + 2\pi k \text{ für eine Zahl } k$$

gilt.

Abbildung: Polarkoordinaten

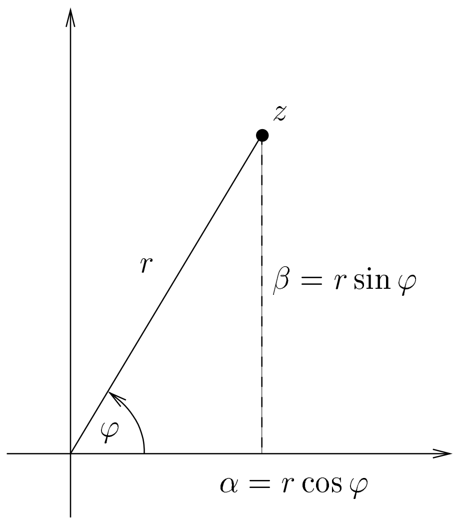
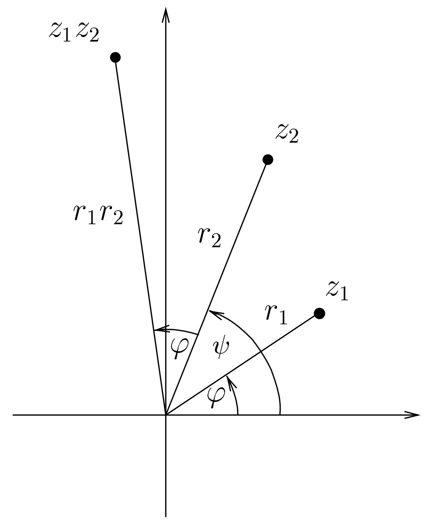


Abbildung: Geometrische Interpretation des Produktes



2.2.8 Geometrische Interpretation des Produktes von komplexen Zahlen. Polarkoordinaten liefern eine geometrische Interpretation des Produktes zweier komplexer Zahlen. Man betrachte dazu zwei komplexe Zahlen in Polardarstellung:

$$z_1 = r_1(\cos\varphi + i\sin\varphi) \quad , \text{ d.h. } r_1 = |z_1|, \quad \varphi = \arg(z_1)\ ,$$

$$z_2 = r_2(\cos\psi + i\sin\psi) \quad , \text{ d.h. } r_2 = |z_2|, \quad \psi = \arg(z_2)\ .$$

Es folgt

$$\begin{aligned} z_1 z_2 &= r_1(\cos\varphi + i\sin\varphi)\cdot r_2(\cos\psi + i\sin\psi) \\ &= r_1 r_2\big(\cos\varphi\cos\psi - \sin\varphi\sin\psi + i(\cos\varphi\sin\psi + \sin\varphi\cos\psi)\big) \\ &= r_1 r_2(\cos(\varphi+\psi) + i\sin(\varphi+\psi))\ , \end{aligned}$$

wobei wir die Additionstheoreme für sinus und cosinus benutzt haben. Wir erhalten wieder eine Darstellung des Produktes in Polardarstellung und folgern:

$$|z_1 z_2| = r_1 r_2 = |z_1|\cdot|z_2| \quad , \quad \arg(z_1 z_2) = \varphi + \psi = \arg(z_1) + \arg(z_2) \quad (\ +2\pi k) \quad ;$$

d.h. der Betrag des Produktes ist das Produkt der Beträge der Faktoren und das Argument des Produktes ist die Summe der Argumente der Faktoren.

2.2.9 Invertierbarkeit von komplexen Zahlen. Für jede komplexe Zahl $z \neq 0$ ist das Produkt $z\overline{z} = |z|^2$ eine echt positive reelle Zahl. Also können wir die Inverse von $|z|^2$ im Körper der reellen Zahlen bilden, weshalb auch $w = \frac{1}{|z|^2}\cdot\overline{z}$ eine komplexe Zahl ist. Es folgt $w \cdot z = \frac{1}{|z|^2}\overline{z}\cdot z = \frac{1}{|z|^2}\cdot|z|^2 = 1$. Also ist z invertierbar und das Inverse ist $z^{-1} = \frac{1}{|z|^2}\cdot\overline{z} = \frac{1}{z\overline{z}}\cdot\overline{z}$. Wir folgern:

Satz *Die komplexen Zahlen bilden einen Körper* $\mathbb{C}$.

Das inverse Element einer komplexen Zahl $z = \alpha + i\beta \neq 0$ ist

$$z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{|z|^2} = \frac{\alpha - i\beta}{\alpha^2 + \beta^2} .$$

Ist eine von 0 verschiedene komplexe Zahl in Polardarstellung $z = r(\cos\varphi + i\sin\varphi)$ gegeben, so ist deren Inverse

$$z^{-1} = \frac{1}{r}(\cos(-\varphi) + \sin(-\varphi)) = \frac{1}{r}(\cos\varphi - i\sin\varphi).$$

Es folgt

$$|z^{-1}| = \frac{1}{|z|} \quad , \quad \arg(z^{-1}) = -\arg(z) \quad (+2\pi k) .$$

2.2.10 Potenzen. Mittels Polardarstellung können Potenzen von komplexen Zahlen leicht berechnet werden. Nutzen wir nämlich die Formel für das Produkt in 2.2.8, so erhalten wir:

$$\begin{aligned} z &= r(\cos\varphi + i\sin\varphi) \\ z^2 &= r^2(\cos 2\varphi + i\sin 2\varphi) \\ z^3 &= r^3(\cos 3\varphi + i\sin 3\varphi) \\ &\vdots \\ z^n &= r^n(\cos n\varphi + i\sin n\varphi) \end{aligned}$$

Die letzte Formel gilt nicht nur für positive, sondern auch für alle negativen ganzen Zahlen, was aus der Formel für die Inverse nach 2.2.9 folgt. Spezialisiert man obige Formel auf komplexe Zahlen mit Betrag $|z| = r = 1$, so ergibt sich die überraschende Formel

$$(\cos\varphi + i\sin\varphi)^n = (\cos n\varphi + i\sin n\varphi) \qquad \text{(Formel von DE MOIVRE)}.$$

2.2.11 Wurzeln. Stellen wir eine komplexe Zahl in Polardarstellung

$$z = r(\cos\varphi + i\sin\varphi)$$

dar, kann man leicht die $n-$ten Wuzeln, d.h. die komplexen Zahlen w, die die Gleichung $w^n = z$ erfüllen, berechnen. Wir schreiben dazu die Unbekannte w in Polardarstellung

$$w = \rho(\cos\psi + i\sin\psi).$$

Aus 2.2.10 folgt

$$w^n = \rho^n(\cos n\psi + i\sin n\psi) .$$

Die Gleichung $w^n = z$ gilt genau dann, wenn

$$\rho^n = r \quad \text{und} \quad n\psi = \varphi + 2\pi k,$$

d.h. genau dann, wenn

$$\rho = \sqrt[n]{r} \quad \text{und} \quad \psi = \frac{\varphi + 2\pi k}{n} \quad , k \in \mathbb{Z} .$$

Wählen wir $k = 0, 1, 2, \ldots, n-1$, so erhalten wir n verschiedene n-te Wurzeln $w_0, w_1, \ldots, w_{n-1}$ von z; alle anderen Zahlen k ergeben dieselben Wurzeln.

Wir haben also gezeigt: *für eine gegebene natürliche Zahl* n *hat jede komplexe Zahl* $z = r(\cos\varphi + i\sin\varphi) \neq 0$ *genau* n *verschiedene komplexe* $n-$*te Wuzeln*

$$\begin{aligned}
w_0 &= \sqrt[n]{r}\ (\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}) \\
w_1 &= \sqrt[n]{r}\ (\cos\frac{\varphi+2\pi}{n} + i\sin\frac{\varphi+2\pi}{n}) \\
w_2 &= \sqrt[n]{r}\ (\cos\frac{\varphi+4\pi}{n} + i\sin\frac{\varphi+4\pi}{n}) \\
\vdots\ & \quad \vdots \\
w_{n-1} &= \sqrt[n]{r}\ (\cos\frac{\varphi+2(n-1)\pi}{n} + i\sin\frac{\varphi+2(n-1)\pi}{n})
\end{aligned}$$

Betrachten wir einige Spezialfälle:

Quadratwurzeln Eine komplexe Zahl $z = r(\cos\varphi + i\sin\varphi) \neq 0$ hat zwei komplexe Quadratwurzeln

$$w_{1/2} = \pm\sqrt{r}\ (\cos\frac{\varphi}{2} + i\sin\frac{\varphi}{2})\ .$$

Einheitswurzeln Betrachten wir den Fall $z = 1$, d.h. $r = 1, \varphi = 0$, dann suchen wir nach den komplexen Lösungen der Gleichung $w^n = 1$. Nach dem Obigen ergeben sich die Lösungen:

$$\begin{aligned}
w_0 &= 1 \\
w_1 &= \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n} \\
w_2 &= \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n} \\
\vdots\ & \quad \vdots \\
w_{n-1} &= \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n}.
\end{aligned}$$

Diese Zahlen heißen die $n-$*ten Einheitswurzeln*. In der komplexen Ebene bilden die $n-$ten Einheitswurzeln die Ecken eines regelmäßigen $n-$Ecks.

Spezialfall $n = 6$: Die sechsten Einheitswurzeln sind:

$$w_0 = 1\ ,\ w_1 = \frac{1}{2}+i\frac{\sqrt{3}}{2}\ ,\ w_2 = -\frac{1}{2}+i\frac{\sqrt{3}}{2}\ ,\ w_3 = -1\ ,\ w_4 = -\frac{1}{2}-i\frac{\sqrt{3}}{2}\ ,\ w_5 = \frac{1}{2}-i\frac{\sqrt{3}}{2}\ .$$

Abbildung: Die 6-ten Einheitswurzeln

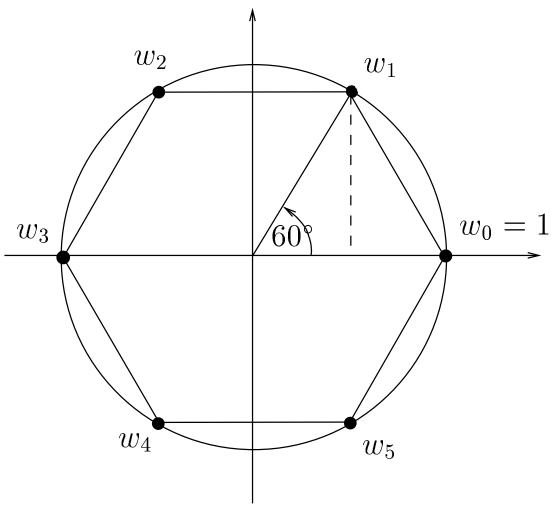
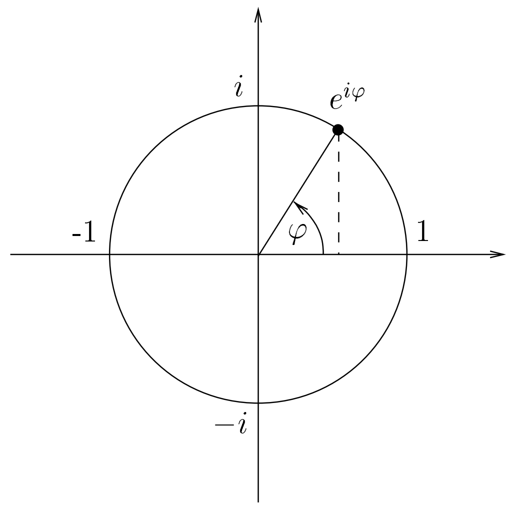


Abbildung: Die komplexe Exponentialfunktion



2.2.12 Die komplexe Exponentialfunktion. Sie kennen sicherlich die *reelle Exponentialfunktion*

$$\exp(x) = e^x,$$

die für alle reellen Zahlen x definiert ist, wobei $e = 2.71\ldots$ die Euler'sche Zahl ist. Die reelle Exponentialfunktion erfüllt die folgende fundamentale Funktionalgleichung:

$$e^{x+y} = e^x \cdot e^y \quad \text{für alle reellen Zahlen} \quad x, y \ .$$

Wir werden nun die Exponentialfunktion $\exp(z)$ für jede komplexe Zahl z so definieren, dass die fundamentale Funktionalgleichung weiterhin gültig bleibt. In der Analysisvorlesung wird dieses Vorgehen später genauer behandelt werden, so dass die Vorgehensweise dann natürlicher erscheinen wird.

Wir definieren zunächst die Exponentialfunktion für rein imaginäre Zahlen $z = i\varphi$ durch

$$\exp(i\varphi) = e^{i\varphi} := \cos\varphi + i\sin\varphi \ .$$

Sie können $e^{i\varphi}$ einfach als eine Kurzschreibweise für $\cos\varphi + i\sin\varphi$ ansehen. Dies scheint natürlich zu sein. Die Darstellung mittels Polarkoordinaten einer komplexen Zahl ergibt sich zu

$$z = r(\cos\varphi + i\sin\varphi) = r \cdot e^{i\varphi} = |z|e^{\arg(z)} \ ,$$

und die $n-$ten Wurzeln von z können wie folgt geschrieben werden:

$$w_k = \sqrt[n]{r} \cdot e^{\frac{\varphi + 2\pi k}{n}} \quad , \quad k = 0, 1, 2, \ldots, n-1 \ .$$

Die $n-$ten Einheitswurzeln sind

$$w_k = e^{\frac{2\pi k}{n}} \quad , \quad k = 0, 1, 2, \ldots, n-1 \ .$$

Für $0 \leq \varphi \leq 2\pi$ ist das Bild der Exponentialfunktion $e^{i\varphi}$ der Einheitskreis in der komplexen Ebene. Da die Funktionen sinus und cosinus periodisch mit der Periode 2π sind, gilt dasselbe auch für $e^{i\varphi}$, d.h.

$$e^{i(\varphi + 2\pi)} = e^{i\varphi} \ .$$

Wir definieren nun für eine beliebige komplexe Zahl $z = \alpha + i\beta$:

$$\exp(z) = e^z := e^\alpha \cdot e^{i\beta} = e^\alpha \cdot (\cos\beta + i\sin\beta)\ .$$

Die komplexe Exponentialfunktion erfüllt also auch die Funktionalgleichung

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}\ .$$

In der Tat folgt für $z_1 = \alpha_1 + i\beta_1$ und $z_2 = \alpha_2 + i\beta_2$ aus der Definition der Exponentialfunktion mittels der Funktionalgleichung der reellen Exponentialfunktion und aus der Formel für das Produkt komplexer Zahlen in 2.2.8:

$$\begin{aligned}
e^{z_1} \cdot e^{z_2} &= e^{\alpha_1}e^{i\beta_1} \cdot e^{\alpha_2}e^{i\beta_2} \\
&= e^{\alpha_1}(\cos\beta_1 + i\sin\beta_1) \cdot e^{\alpha_2}(\cos\beta_2 + i\sin\beta_2) \\
&= e^{\alpha_1+\alpha_2}((\cos(\beta_1+\beta_2) + i\sin(\beta_1+\beta_2)) \\
&= e^{\alpha_1+\alpha_2+i(\beta_1+\beta_2)} \\
&= e^{z_1+z_2}
\end{aligned}$$

2.2.13 Wurzeln von Polynomen. Betrachten wir ein Polynom p mit reellen oder komplexen Koeffizienten a_i:

$$p(z) = a_n z^n + a_{n-1}z_{n-1} + \cdots + a_2 z^2 + a_1 z + a_0 \quad \text{für alle } z \in \mathbb{C}\ .$$

Im Falle $a_n \neq 0$ heißt der Koeffizient a_0 der *Leitkoeffizient* von p.

Eine komplexe Zahl z_0 heißt *Wurzel, Nullstelle* des Polynoms p, wenn sie die Gleichung $p(z_0) = 0$ erfüllt.

Die Bestimmung der Wurzeln von Polynomen stellte ein wichtiges Problem in der Geschichte der Mathematik dar. Die Tatsache, dass das Polynom

$$p(z) = z^2 + 1$$

keine reelle Lösung hat, führte mit zur Entwicklung der komplexen Zahlen. Wir haben in 2.2.11 gesehen, dass für jede komplexe Zahl $a_0 \neq 0$ das Polynom

$$p(z) = z^n - a_0$$

genau n Lösungen hat, die n verschiedenen komplexen Wurzeln von a_0.

Jedes quadratische Polynom

$$p(z) = z^2 + pz + q \quad , \quad p, q \in \mathbb{C}$$

hat zwei komplexe Wurzeln z_1 und z_2 (diese können gleich sein). Es gilt:

$$p(z) = z^2 + pz + \frac{p^2}{4} + (q - \frac{p^2}{4}) = (z + \frac{p}{2})^2 + (q - \frac{p^2}{4})\ .$$

Also ist die Gleichung $p(z) = 0$ äquivalent zu der Gleichung

$$(z + \frac{p}{2})^2 = D\ , \text{ wobei } D = \frac{p^2}{4} - q$$

gilt. Die komplexe Zahl D heißt die *Diskriminante* von p. Wir wissen aus 2.2.11, dass D zwei komplexe Quadratwurzeln $w_{1/2}$ hat. Also gilt

$$z + \frac{p}{2} = w_{1/2}$$

und das gegebene quadratische Polynom $p(z) = z^2 + pz + q$ hat zwei komplexe Wurzeln

$$z_{1/2} = -\frac{p}{2} + w_{1/2} \; .$$

Sind die Koeffizienten p und q reelle Zahlen, dann ist die Diskriminante $D = \frac{p^2}{4} - q$ ebenso eine reelle Zahl. Es gibt drei Fälle:

- $D > 0$: Dann hat D zwei positive Quadratwurzeln $\pm\sqrt{D}$ und das Polynom $p(z) = z^2 + pz + q$ hat zwei reelle Wurzeln $x_{1/2} = -\frac{p}{2} \pm \sqrt{D}$.
- $D = 0$: Das Polynom $p(z) = z^2 + pz + q$ hat genau eine Lösung $x_1 = x_2 = -\frac{p}{2}$.
- $D < 0$: Hier ist $-D > O$ und D hat zwei echt imaginäre Wurzeln $w_{1/2} = \pm i\sqrt{-D}$. Das Polynom $p(z) = z^2 + pz + q$ hat zwei komplex konjugierte Wurzeln $z_{1/2} = -\frac{p}{2} \pm i\sqrt{-D}$.

Es gibt für Polynome von Grad 3 und 4 explizite Formeln zur Bestimmung der komplexen Wurzeln (siehe die Formel von Cardano). Galois hat gezeigt, dass es keine allgemeinen expliziten Formeln zur Bestimmung der komplexen Wurzeln von Polynomen von Grad 5 und höher geben kann.

Dennoch kann man die Existenz von komplexen Wurzeln beweisen:

2.2.14 Fundamentalsatz der Algebra. *Jedes Polynom p mit komplexen Koeffizienten und Grad ≥ 1 hat mindestens eine komplexe Nullstelle.*

Gauss hat diesen Satz als erster bewiesen. Wir können mit unseren Mitteln diesen Satz jetzt noch nicht zeigen. Sie werden einen Beweis in der Vorlesung Funktionentheorie im dritten Semester kennenlernen.

2.2.15 Konstruktion des Körpers der komplexen Zahlen. Wir haben diesen Abschnitt mit der Annahme begonnen, dass es einen kommutativen Ring gibt, der den Körper $\mathbb{R}$ der reellen Zahlen umfasst und ein Element i enthält, für welches $i^2 = -1$ gilt. Die eigenschaften der komplexen Zahlen wurden unter dieser Annahme hergeleitet. Wir werden nun den Körper $\mathbb{C}$ der komplexen Zahlen konstruieren.

Die Darstellung komplexer Zahlen in der komplexen Zahlenebene (siehe 2.2.4) und die Formeln für die Addition und die Multiplikation in 2.2.1 suggerieren die folgende Definition:

Definition Eine *komplexe Zahl* ist ein Paar $z = (\alpha, \beta)$ von reellen Zahlen α, β. Wir bezeichnen mit $\mathbb{C}$ die Menge aller $z = (\alpha, \beta)$, $\alpha, \beta \in \mathbb{R}$. Wir definieren die Summe und das Produkt zweier beliebiger komplexer Zahlen $z_1 = (\alpha_1, \beta_1)$ und $z_2 = (\alpha_2, \beta_2)$ durch:

$$z_1 + z_2 = (\alpha_1 + \alpha_2, \beta_1 + \beta_2) \; ,$$
$$z_1 \cdot z_2 = (\alpha_1\alpha_2 - \beta_1\beta_2, \alpha_1\beta_2 + \beta_1\alpha_2) \; .$$

Aus der Sicht der Addition ist $\mathbb{C}$ völlig identisch mit $\mathbb{R}^2$ bzgl. der üblichen Addition. Also ist $\mathbb{C}$ eine kommutative Gruppe bzgl. der Addition. Man zeigt leicht mittels einfacher Rechnungen, dass die Multiplikation in obiger Form assoziativ und kommutativ ist, dass das Element $(1, 0)$ ein neutrales Element ist und dass die Distributivgesetze gelten. Also ist $\mathbb{C}$ ein kommutativer Ring.

Wir identifizieren nun eine reelle Zahl α mit der komplexen Zahl $(\alpha, 0)$, so wird z.B. 1 mit $(1, 0)$ identifiziert. Weiterhin setzen wir

$$i := (0, 1) \ .$$

Eine komplexe Zahl kann dann in der Form

$$z = (\alpha, \beta) = (\alpha, 0) + (0, \beta) = (\alpha, 0) + (0, 1) \cdot (\beta, 0) \ ,$$

d.h.

$$z = \alpha + i\beta$$

geschrieben werden. Die Identifizierung der komplexen Zahl $(\alpha, 0)$ mit der reellen Zahl α ist sinnvoll, da $(\alpha_1, 0) + (\alpha_2, 0) = (\alpha_1 + \alpha_2, 0)$ und $(\alpha_1, 0) \cdot (\alpha_2, 0) = (\alpha_1\alpha_2, 0)$ gilt, d.h. komplexe Zahlen der Form $(\alpha, 0)$ werden genau wie die korrespondierenden reellen Zahlen addiert und multipiziert. In diesem Sinn erweitert $\mathbb{C}$ den Körper der reellen Zahlen. $\mathbb{C}$ enthält auch ein Element i mit $i^2 = -1$; es gilt nämlich $i^2 = (0, 1) \cdot (0, 1) = (-1, 0)$, was mit -1 identifiziert wird.

Wir erkennen analog zu 2.2.9, dass jede komplexe Zahl $z \neq 0$ ein Inverses besitzt. Damit ist $\mathbb{C}$ ein Körper.

2.3. Reelle und komplexe Vektorräume

In Kapitel 1 haben wir uns auf reelle Vektorräume beschränkt. Dennoch haben wir aber keine speziellen Eigenschaften der reellen Zahlen benutzt: wir brauchten nur arithmetische Operationen wie die Addition, Subtraktion, Multiplikation und die Division durch von 0 verschiedene Zahlen. Diese Operationen gibt es aber in jedem Körper $\mathbb{K}$. Damit haben wir die Möglichkeit, unseren Begriff des Vektorraumes zu erweitern, indem wir anstelle der reellen Zahlen einen beliebigen Körper für die Menge der Skalare zulassen. Wir beschränken uns neben dem Körper der reellen Zahlen auf den Körper der komplexen Zahlen.

In der Folge sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

2.3.1 Definition. Ein *Vektorraum über* $\mathbb{K}$ (oder kurz ein $\mathbb{K}$-*Vektorraum*) ist eine Menge V mit

- einem Element $\mathbf{0} \in V$
- einer Addition $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w} \colon V \times V \to V$
- einer skalaren Multiplikation $(\lambda, \mathbf{v}) \mapsto \lambda\mathbf{v} \colon \mathbb{K} \times V \to V$

so daß folgende Eigenschaften erfüllt sind:

(V1)	$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V :$	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
(V2)	$\forall \mathbf{v} \in V :$	$\mathbf{v} + \mathbf{0} = \mathbf{v}$
(V3)	$\forall \mathbf{v} \in V :$	$\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}$
(V4)	$\forall \mathbf{v}, \mathbf{w} \in V :$	$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
(V5)	$\forall \mathbf{v} \in V, \forall \lambda, \mu \in \mathbb{K} :$	$\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$
(V6)	$\forall \mathbf{v} \in V :$	$1\mathbf{v} = \mathbf{v}$
(V7)	$\forall \mathbf{v}, \mathbf{w} \in V, \forall \lambda \in \mathbb{K} :$	$\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$

(V8) $\quad \forall \mathbf{v} \in V, \forall \lambda, \mu \in \mathbb{K}: \quad (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}.$

(V1)–(V4) besagen, dass V eine kommutative Gruppe bzgl. der Addition mit $\mathbf{0}$ als neutralem Element ist; das additive Inverse von $\mathbf{v} \in V$ ist $-\mathbf{v} = (-1)\mathbf{v}$.

2.3.2 Folgerungen. In jedem $\mathbb{K}$-Vektorraum V gilt:

(V9) $\quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V: \quad \mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w} \implies \mathbf{u} = \mathbf{v}$

(V10) $\quad \forall \mathbf{v} \in V: \quad 0\mathbf{v} = \mathbf{0}$

$\quad \forall \lambda \in \mathbb{K}: \quad \lambda \cdot \mathbf{0} = \mathbf{0}$

(V11) $\quad \forall \mathbf{v} \in V: \forall \lambda \in \mathbb{K}: \quad \lambda\mathbf{v} = \mathbf{0} \implies \lambda = 0$ oder $\mathbf{v} = \mathbf{0}$.

Der Beweis dieser Eigenschaften kann aus 1.1.5 direkt übertragen werden, da wir dort nur die Grundregeln (V1)–(V8) benutzt haben. (Man prüfe obige Rechenregeln erneut nach.)

2.3.3 Beispiele.

(1) $\mathbb{C}^n$ bezeichne die Menge aller n-Tupel

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

mit $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$. Auf $\mathbb{C}^n$ sind die Addition und die skalare Multiplikation wie folgt definiert

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} \qquad \lambda\mathbf{v} = \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \vdots \\ \lambda\alpha_n \end{pmatrix}.$$

Zusammen mit diesen Verknüpfungen wird $\mathbb{C}^n$ zu einem $\mathbb{C}$-Vektorraum.

(2) Eine $m \times n$-Matrix mit Einträgen aus $\mathbb{C}$ ist ein rechteckiges Schema mit m Zeilen und n Spalten

$$A := (\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} := \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix}$$

mit $\alpha_{ij} \in \mathbb{C}$ für alle i, j. Wir bezeichnen mit $\mathcal{M}(m \times n, \mathbb{C})$ die Menge aller $m \times n$-Matrizen mit Einträgen aus $\mathbb{C}$.

Die Summe $A + B$ zweier $m \times n$-Matrizen A, B und die Multiplikation einer Matrix A mit einem Skalar $\lambda \in \mathbb{C}$ sind auf dieselbe Weise definiert wie für reelle Matrizen. $\mathcal{M}(m \times n, \mathbb{K})$ ist wiederum ein $\mathbb{C}$-Vektorraum, wie man leicht nachweist.

(3) Sei X eine Menge. Wir bezeichnen mit $\mathcal{F}(X,\mathbb{K})$ die Menge aller Funktionen $f\colon X \to \mathbb{K}$. Die konstante Funktion mit Wert $0 \in \mathbb{K}$ sei $\mathbf{0}$; für $f, g \in \mathcal{F}(X,\mathbb{K})$ und $\lambda \in \mathbb{K}$ definieren wir $f+g$ und λf durch

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x) && \text{für alle } x \in X,\\ (\lambda f)(x) &:= \lambda \cdot f(x) && \text{für alle } x \in X. \end{aligned}$$

Man kann wieder zeigen, dass $\mathcal{F}(X,\mathbb{K})$ auf diese Weise zu einem $\mathbb{K}$-Vektorraum wird.

Als Spezialfall kann man für X ein Intervall in $\mathbb{R}$ und $\mathbb{K} = \mathbb{R}$ wählen. Die Menge aller Funktionen $f\colon X \to \mathbb{R}$ mit der gewöhnlichen Addition und der Multiplikation mit Skalaren λ ist ein reeller Vektorraum.

2.4. Lineare Teilräume, affine Teilräume

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

2.4.1 Definition. Eine Teilmenge $U \subseteq V$ wird *linearer Teilraum* [*linear subspace*] oder auch auch *Untervektorraum* genannt, wenn sie die folgenden Eigenschaften erfüllt:

(U1) $$\mathbf{0} \in U,$$

(U2) $$\mathbf{v}, \mathbf{w} \in U \implies \mathbf{v} + \mathbf{w} \in U,$$

(U3) $$\mathbf{v} \in U,\ \lambda \in \mathbb{K} \implies \lambda\mathbf{v} \in U.$$

Bemerkung. (U2) und (U3) können zu einer einzelnen äquivalenten Bedingung zusammengefasst werden:

(U$*$) $$\mathbf{v}, \mathbf{w} \in U,\ \lambda, \mu \in \mathbb{K} \implies \lambda\mathbf{v} + \mu\mathbf{w} \in U.$$

Wiederholte Anwendung von (U$*$) ergibt: Wenn U ein linearer Teilraum von V ist, so gilt:

$$\left.\begin{array}{l} \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in U \\ \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K} \end{array}\right\} \implies \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n \in U.$$

2.4.2 Beispiele.

(1) $\{\mathbf{0}\}$ und V sind lineare Teilräume von V.

(2) Die linearen Teilräume von $\mathbb{R}^2$ sind: $\{\mathbf{0}\}$, Geraden durch $\mathbf{0}$, ganz $\mathbb{R}^2$.

(3) Die linearen Teilräume von $\mathbb{R}^3$ sind: $\{\mathbf{0}\}$, Geraden durch $\mathbf{0}$, Ebenen durch $\mathbf{0}$, ganz $\mathbb{R}^3$.

(4) **Polynome** Ein *reelles Polynom* ist eine Funktion $p\colon \mathbb{R} \to \mathbb{R}$, die durch einen Ausdruck der Form

$$p(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_n x^n \qquad \text{für alle } x \in \mathbb{R}$$

gegeben ist, wobei die Koeffizienten $\lambda_0, \lambda_1, \dots, \lambda_n$ reelle Zahlen sind. Der *Grad* von p ist definiert als

$$\deg(P) := \begin{cases} \max\{i \in \mathbb{N} \mid \lambda_i \neq 0\}, & \\ -\infty, & \text{falls alle Koeffizienten } \lambda_i = 0, \end{cases}$$

Die folgenden Mengen sind lineare Teilräume des reellen Vektorraums $\mathcal{F}(\mathbb{R},\mathbb{R})$ aller Funktionen $f\colon \mathbb{R}\to\mathbb{R}$:

- $\mathcal{P}(\mathbb{R})$: die Menge aller reellen Polynome
- $\mathcal{P}_n(\mathbb{R})$: die Menge aller reellen Polynome vom Höchstgrad n.

Entsprechend heißt eine Funktion $p\colon \mathbb{C}\to\mathbb{C}$ ein Polynom, wenn sie durch einen Ausdruck der Form

$$p(z)=\lambda_0+\lambda_1 z+\lambda_2 z^2+\cdots+\lambda_n z^n \qquad \text{für alle } z\in\mathbb{C}$$

mit $\lambda_0,\lambda_1,\ldots,\lambda_n\in\mathbb{C}$ gegeben ist. Die Menge der komplexen Polynome bildet einen linearen Teilraum des Vektorraums $\mathcal{F}(\mathbb{C},\mathbb{C})$.

(5) **Trigonometrische Polynome** Als *trignometrisches Polynom vom Grad* $\leq n$ bezeichnet man jede reelle Funktion gegeben durch einen Ausdruck der Form

$$T(x)=\frac{\alpha_0}{2}+\sum_{k=1}^{n}\alpha_i\cos kx+\beta_i\sin kx\,.$$

Sowohl die Menge $\mathcal{T}$ aller trigonometrischen Polynome als auch die Menge $\mathcal{T}_n$ aller trigonometrischen Polynome vom Grad $\leq n$ bilden einen Untervektorraum der Vektorraums $\mathcal{F}(\mathbb{R},\mathbb{R})$ aller Funktionen $f\colon\mathbb{R}\to\mathbb{R}$.

Die folgende Proposition rechtfertigt den Namen *Teilvektorraum.*

2.4.3 Proposition. *Ist U ein linearer Teilraum eines $\mathbb{K}$-Vektorraumes V, dann ist auch U ein $\mathbb{K}$-Vektorraum.*

Beweis. Per Definition gilt $\mathbf{0}\in U$ und die Einschränkung der Addition von V auf U ergibt eine Verknüpfung $(\mathbf{v},\mathbf{w})\mapsto\mathbf{v}+\mathbf{w}\colon U\times U\to U$ nach (U2). Auf ähnliche Weise folgt aus (U3), dass die Einschränkung der skalaren Multiplikation auf Elemente $v\in U$ eine skalare Multiplikation $(\lambda,\mathbf{v})\mapsto\lambda\mathbf{v}\colon\mathbb{K}\times U\to U$ ergibt. Selbstverständlich bleiben die Axiome (V1)–(V8) weiterhin gültig, wenn wir uns auf Elemente in U beschränken. Also ist U ein $\mathbb{K}$-Vektorraum. □

Bemerkung. Um nachzuweisen, dass ein gegebenes Objekt U ein Vektorraum ist, kann man die Gültigkeit der Axiome (V1)–(V8) zeigen. Es ist jedoch oftmals erheblich leichter zu prüfen, dass U ein linearer Teilraum eines bereits bekannten Vektorraumes V ist; dann ist nach vorigem Ergebnis auch U ein Vektorraum.

Beispielsweise bilden die Mengen $\mathcal{P}(\mathbb{R})$ und $\mathcal{P}_n(\mathbb{R})$ reeller Polynome reelle Vektorräume, da sie lineare Teilräume des Vektorraumes $\mathcal{F}(\mathbb{R},\mathbb{R})$ aller Funktionen $f\colon\mathbb{R}\to\mathbb{R}$ sind (siehe 2.4.2 (4)).

2.4.4 Proposition. *Sind $U_1,U_2,\ldots,U_n$ lineare Teilräume von V, dann ist auch ihre Schnittmenge $U_1\cap U_2\cap\cdots\cap U_n$ ein linearer Teilraum von V.*

Beweis.

(U1): Da $U_1,U_2,\ldots,U_n$ lineare Teilräume sind, folgt $\mathbf{0}\in U_1$, $\mathbf{0}\in U_2$, …, $\mathbf{0}\in U_n$, was $\mathbf{0}\in U_1\cap U_2\cap\cdots\cap U_n$ impliziert.

(U2): Seien $\mathbf{v}, \mathbf{w} \in U_1 \cap U_2 \cap \cdots \cap U_n$. Dann ist $\mathbf{v}, \mathbf{w} \in U_i$ für jedes $i = 1, \ldots, n$. Da die U_i alle lineare Teilräume sind, folgern wir $\mathbf{v} + \mathbf{w} \in U_i$ für alle $i = 1, \ldots, n$. Deshalb gilt auch $\mathbf{v} + \mathbf{w} \in U_1 \cap U_2 \cap \cdots \cap U_n$.

(U3): Es sei $\mathbf{v} \in U_1 \cap U_2 \cap \cdots \cap U_n$ und $\lambda \in \mathbb{K}$. Dann ist $\mathbf{v} \in U_i$ für jedes $i = 1, \ldots, n$. Da die U_i lineare Teilräume sind, folgt $\lambda\mathbf{v} \in U_i$ für alle $i = 1, \ldots, n$, was $\lambda\mathbf{v} \in U_1 \cap U_2 \cap \cdots \cap U_n$ impliziert. □

Frage. Wenn U_1 und U_2 lineare Teilräume von V sind, ist dann auch $U_1 \cup U_2$ ein linearer Teilraum?

2.4.5 Die Summe linearer Teilräume. Für lineare Teilräume $U_1, U_2, \ldots, U_n$ von V ist die *Summe* definiert als

$$U_1 + U_2 + \cdots + U_n := \{\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_n : \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2, \ldots, \mathbf{u}_n \in U_n\}.$$

Proposition. *Die Summe $U_1 + U_2 + \cdots + U_n$ linearer Teilräume $U_1, U_2, \ldots, U_n$ von V ist ein linearer Teilraum.*

Der Beweis dieser Proposition wird dem Leser als Übung überlassen. □

Beispiel. Betrachte im $\mathbb{R}^3$ zwei Ebenen durch den Ursprung:

$$U_1 = \{\lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2 : \lambda_1, \lambda_2 \in \mathbb{R}\} \qquad U_2 = \{\lambda_1\mathbf{e}_1 + \lambda_3\mathbf{e}_3 : \lambda_1, \lambda_3 \in \mathbb{R}\}.$$

Dann gilt $U_1 \cap U_2 = \{\lambda_1\mathbf{e}_1 : \lambda_1 \in \mathbb{R}\}$ (eine Gerade durch $\mathbf{0}$) und $U_1 + U_2 = \mathbb{R}^3$.

Das Element $\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) \in \mathbb{R}^3$ besitzt mehr als eine Darstellung als Summe $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$ mit $\mathbf{u}_1 \in U_1$ und $\mathbf{u}_2 \in U_2$:

$$\mathbf{v} = \begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\end{pmatrix} = \underbrace{\mathbf{e}_1 + \mathbf{e}_2}_{\in U_1} + \underbrace{\mathbf{e}_3}_{\in U_2} = \underbrace{\mathbf{e}_2}_{\in U_1} + \underbrace{\mathbf{e}_1 + \mathbf{e}_3}_{\in U_2}.$$

2.4.6 Direkte Summen. Es seien $U_1, U_2, \ldots, U_n$ lineare Teilräume von V. Wir sagen, V sei die *direkte Summe* von $U_1, U_2, \ldots, U_n$, und wir schreiben

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_n,$$

wenn jedes Element $v \in V$ eine *eindeutige* Darstellung der Form

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_n \qquad \text{mit } \mathbf{u}_i \in U_i \text{ für } i = 1, 2, \ldots, n$$

besitzt.

Proposition. *Seien U_1 und U_2 lineare Teilräume von V. Dann gilt $V = U_1 \oplus U_2$ genau dann, wenn*

(1) $V = U_1 + U_2$

(2) $U_1 \cap U_2 = \{\mathbf{0}\}$

Beweis. (a) „$\Rightarrow$“: Nehmen wir zunächst $V = U_1 \oplus U_2$ an. Dann gilt insbesondere $V = U_1 + U_2$. Um (2) zu zeigen, sei $\mathbf{v} \in U_1 \cap U_2$. $\mathbf{v}$ hat dann zwei Darstellungen

$$\mathbf{v} = \underbrace{\mathbf{v}}_{\in U_1} + \underbrace{\mathbf{0}}_{\in U_2} = \underbrace{\mathbf{0}}_{\in U_1} + \underbrace{\mathbf{v}}_{\in U_2}$$

als Summe eines Elementes aus U_1 und eines Elementes aus U_2. Aufgrund der Eindeutigkeit einer solchen Darstellung folgern wir $\mathbf{v} = \mathbf{0}$. Also gilt $U_1 \cap U_2 = \{\mathbf{0}\}$.

(b) „$\Leftarrow$“: Es seien nun (1) und (2) erfüllt. Nach (1) hat jedes $\mathbf{v} \in V$ eine Darstellung

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 \quad \text{mit } \mathbf{u}_1 \in U_1,\ \mathbf{u}_2 \in U_2$$

Um die Eindeutigkeit einer solchen Darstellung zu zeigen, nehmen wir weiterhin an, dass gilt:

$$\mathbf{v} = \mathbf{u}_1' + \mathbf{u}_2' \quad \text{mit } \mathbf{u}_1' \in U_1,\ \mathbf{u}_2' \in U_2.$$

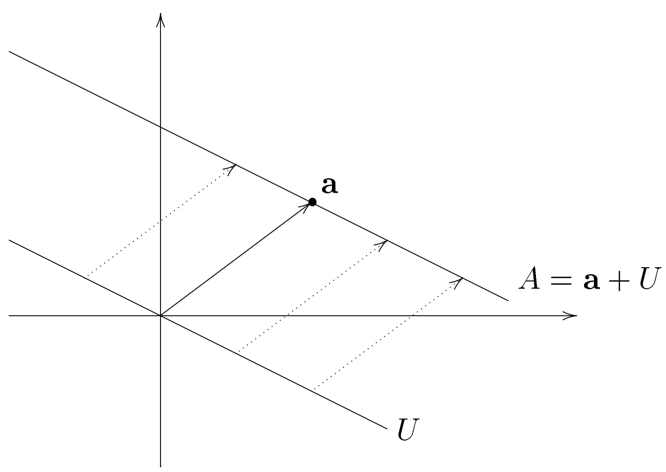
Durch Subtraktion erhalten wir $\mathbf{0} = (\mathbf{u}_1 - \mathbf{u}_1') + (\mathbf{u}_2 - \mathbf{u}_2')$. Daraus folgt $\mathbf{u}_1 - \mathbf{u}_1' = \mathbf{u}_2 - \mathbf{u}_2'$. Da U_1 und U_2 lineare Teilräume sind, gilt weiterhin $\mathbf{u}_1' - \mathbf{u}_1 \in U_1$ und $\mathbf{u}_2' - \mathbf{u}_2 \in U_2$, also

$$\mathbf{u}_1' - \mathbf{u}_1 = \mathbf{u}_2' - \mathbf{u}_2 \in U_1 \cap U_2.$$

Da $U_1 \cap U_2 = \{\mathbf{0}\}$ nach (2) gilt, folgt $\mathbf{u}_1' - \mathbf{u}_1 = \mathbf{0} = \mathbf{u}_2' - \mathbf{u}_2$ und schließlich $\mathbf{u}_1 = \mathbf{u}_1'$ und $\mathbf{u}_2 = \mathbf{u}_2'$. Also ist die Darstellung von $\mathbf{v}$ als Summe eines Elementes $\mathbf{u}_1 \in U_1$ und eines Elementes $\mathbf{u}_2 \in U_2$ eindeutig. □

2.4.7 Affine Teilräume. Eine Teilmenge $A \subseteq V$ wird *affiner Teilraum* von V genannt, wenn es einen linearen Teilraum $U \subseteq V$ und einen Vektor $\mathbf{a} \in V$ gibt, so daß gilt:

$$A = \mathbf{a} + U := \{\mathbf{a} + \mathbf{u} : \mathbf{u} \in U\}.$$



Beispiele.

(1) Die affinen Teilräume des $\mathbb{R}^2$ sind Punkte $\{\mathbf{a}\} = \mathbf{a} + \{\mathbf{0}\}$, beliebige Geraden und $\mathbb{R}^2$.

(2) Die affinen Teilräume des $\mathbb{R}^3$ sind Punkte $\{\mathbf{a}\}$, beliebige Geraden, beliebige Ebenen, $\mathbb{R}^3$.

2.5. Linearkombinationen, lineare Unabhängigkeit, Basen

V bezeichne nun immer einen Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$, zum Beispiel $V = \mathbb{R}^n, \mathbb{C}^n, \mathcal{P}(\mathbb{R}), \mathcal{P}_n(\mathbb{R}), \mathcal{T}, \mathcal{T}_n$.

$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ bezeichne ein System von Elementen aus V.

2.5.1 Linearkombinationen. Jedes Element der Form

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m = \sum_{j=1}^{m} \lambda_j \mathbf{v}_j,$$

mit Skalaren $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{K}$ wird eine *Linearkombination* von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ genannt.

Proposition. *Die Menge*

$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m) := \{\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m \colon \lambda_1, \ldots, \lambda_m \in \mathbb{K}\}$$

aller Linearkombinationen von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ *ist ein linearer Teilraum von* V.

Beweis.

(U1): $\mathbf{0} = 0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_m \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

(U2): Seien $\mathbf{v}, \mathbf{w} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$, d.h. $\mathbf{v}$ und $\mathbf{w}$ sind Linearkombinationen der Form

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m$$
$$\mathbf{w} = \mu_1\mathbf{w}_1 + \mu_2\mathbf{w}_2 + \cdots + \mu_m\mathbf{w}_m.$$

Dann gilt:

$$\mathbf{v} + \mathbf{w} = (\lambda_1 + \mu_1)\mathbf{w}_1 + (\lambda_2 + \mu_2)\mathbf{w}_2 + \cdots + (\lambda_m + \mu_m)\mathbf{w}_m,$$

d.h. $\mathbf{v} + \mathbf{w} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

(U3): Sei $\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$. Für jedes $\lambda \in \mathbb{K}$ gilt

$$\lambda\mathbf{v} = \lambda\lambda_1\mathbf{v}_1 + \lambda\lambda_2\mathbf{v}_2 + \cdots + \lambda\lambda_m\mathbf{v}_m.$$

Also folgt $\lambda\mathbf{v} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$. $\square$

Bemerkung. $\mathbf{v}_1 = 1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_m \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$. Analog folgt $\mathbf{v}_2, \ldots, \mathbf{v}_m \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

Weiterhin ist $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ der kleinste lineare Teilraum von V, der $\mathbf{v}_1$ $\mathbf{v}_2, \ldots, \mathbf{v}_m$ enthält. Das ist klar; denn wenn U ein linearer Teilraum von V ist, der $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ beinhaltet, folgt aus (U2) und (U3), daß U alle Linearkombinationen von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ umfaßt, d.h. U enthält $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

Bemerkung. Sind $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$, so gilt

$$\operatorname{span}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k) \subseteq \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m).$$

Dies ergibt sich direkt aus voriger Bemerkung.

Definition. Wir sagen, U wird von $\mathbf{v}_1$, $\mathbf{v}_2,\ldots,\mathbf{v}_m$ *aufgespannt* oder *erzeugt*, wenn

$$U = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m),$$

d.h. falls jedes Element $\mathbf{v} \in U$ als eine Linearkombination von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ dargestellt werden kann. Wir sagen dann auch, $\mathbf{v}_1$, $\mathbf{v}_2,\ldots,\mathbf{v}_m$ ist ein *Erzeugendensystem* von U.

Beispiele.

(1) $\mathbb{K}^n$ wird aufgespannt von

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \ldots \qquad \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Es gibt aber noch viele weitere erzeugende Mengen. So wird $\mathbb{R}^2$ ebenso von $\mathbf{v}_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ und $\mathbf{v}_2 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ aufgespannt.

(2) Der Vektorraum $\mathcal{P}_n(\mathbb{R})$ der reellen Polynome vom Grad $\leq n$ wird von p_0, $p_1,\ldots$, p_n mit

$$p_j(x) = x^j \quad \text{für alle } x \in \mathbb{R}$$

aufgespannt.

2.5.2 Lineare Unabhängigkeit. $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ werden *linear unabhängig* genannt, wenn keiner der $\mathbf{v}_j$ als Linearkombination der anderen $\mathbf{v}_i$ $(i \neq j)$ dargestellt werden kann. Andernfalls werden sie *linear abhängig* genannt.

Kriterium für lineare Unabhängigkeit. *Die folgenden drei Eigenschaften sind gleichwertig:*

(1) $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ *sind linear unabhängig.*

(2) $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m = \mathbf{0} \implies \lambda_1 = \lambda_2 = \cdots = \lambda_m = 0.$

(3) $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m = \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \cdots + \mu_m\mathbf{v}_m \implies \lambda_i = \mu_i$ *für alle i.*

Beweis.

(1) $\implies$ (3): Wir gehen von der Gleichung

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m = \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \cdots + \mu_m\mathbf{v}_m$$

aus. Sie ist gleichwertig mit der Gleichung

$$(\lambda_1 - \mu_1)\mathbf{v}_1 + \cdots + (\lambda_m - \mu_m)\mathbf{v}_m = 0$$

Wäre (3) nicht erfllt so gäbe es mindestens einen Index i für den $\lambda_i \neq \mu_i$, d.h. $\lambda_i - \mu_i \neq 0$ gilt. Man könnte dann die Gleichung nach $\mathbf{v}_i$ auflösen:

$$\mathbf{v}_i = \frac{1}{\lambda_i - \mu_i} \sum_{j \neq i} (\mu_j - \lambda_j)\mathbf{v}_j$$

und hätte $\mathbf{v}_i$ als Linearkombination der anderen $\mathbf{v}_j$ dargestellt im Widerspruch der in (1) vorausgesetzten linearen Unabhängigkeit.

(3) $\implies$ (2): Setzt man $\mu_i = 0$ für alle i, dann sieht man, daß (2) eine Spezialfall von (3) ist.

(2) $\implies$ (1): Angenommen $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ wären linear abhängig. Dann gäbe es ein j, so daß $\mathbf{v}_j$ eine Linearkombination der $\mathbf{v}_i$ mit $i \neq j$ ist. Nach Umnummerierung kann man $j = 1$ annehmen, d.h., es gilt

$$\mathbf{v}_1 = \mu_2 \mathbf{v}_2 + \dots + \mu_m \mathbf{v}_m$$

für gewisse Skalare μ_i. Es folgt

$$-\mathbf{v}_1 + \mu_2 \mathbf{v}_2 + \dots + \mu_m \mathbf{v}_m = \mathbf{0}$$

Es sind aber nicht alle Koeffizienten 0, denn der erste Koeffizient ist -1 im Widerspruch zu (2). $\square$

Beispiele.

(1) $\mathbf{e}_1 = \binom{1}{0}$ und $\mathbf{e}_2 = \binom{0}{1}$ sind linear unabhängig, ebenso sind $\mathbf{v}_1 = \binom{1}{1}$ und $\mathbf{v}_2 = \binom{1}{-1}$ linear unabhängig in $\mathbb{R}^2$.

(2) $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ sind linear unabhängig in $\mathbb{K}^n$.

(3) $p_0, p_1, \dots, p_n$ sind linear unabhängig in $\mathcal{P}_n(\mathbb{R})$.

(4) Man betrachte eine Matrix in Zeilenstufenform. Dann sind die Zeilen $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r$, die ungleich Null sind, linear unabhängig:

$$\begin{array}{cc} \left(\begin{array}{cccccccccc} 0 & \dots & 0 & \beta_{1j_1} & \dots & \beta_{1j_2} & \dots & \beta_{1j_r} & \dots & \beta_{1n} \\ 0 & \dots & \dots & \dots & 0 & \beta_{2j_2} & \dots & \beta_{2j_r} & \dots & \beta_{2n} \\ \vdots & & & & & & \ddots & \vdots & & \vdots \\ 0 & \dots & \dots & \dots & & \dots & 0 & \beta_{rj_r} & \dots & \beta_{rn} \\ 0 & \dots & \dots & \dots & & \dots & \dots & 0 & \dots & 0 \\ \vdots & & & & & & & \vdots & & \vdots \\ 0 & \dots & \dots & \dots & & \dots & \dots & 0 & \dots & 0 \end{array}\right) & \begin{array}{c} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_r \\ \mathbf{w}_{r+1} \\ \vdots \\ \mathbf{w}_m \end{array} \end{array}$$

2.5.3 Basen. Eine Familie $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ von Vektoren in V wird eine *Basis* von V genannt, wenn

(1) $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m) = V$ ist und

(2) $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ linear unabhängig sind.

In anderen Worten,

> die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ bilden genau dann eine Basis von V, wenn sie linear unabhängig sind und wenn jedes $\mathbf{v} \in V$ als eine Linearkombination von $\mathbf{v}_1, \dots, \mathbf{v}_m$ dargestellt werden kann.

Beispiele.

(1) $\mathbf{e}_1 = \binom{1}{0}$ und $\mathbf{e}_2 = \binom{0}{1}$ ist eine Basis von $\mathbb{R}^2$, ebenso bilden $\mathbf{v}_1 = \binom{1}{1}$ und $\mathbf{v}_1 = \binom{1}{-1}$ eine Basis von $\mathbb{R}^2$.

(2) $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ist eine Basis von $\mathbb{K}^n$.

(3) $p_0, p_1, \dots, p_n$ bilden eine Basis des reellen Vektorraumes $\mathcal{P}_n(\mathbb{R})$.

Unser Ziel ist es nun, einen Algorithmus anzugeben, mit dem wir eine Basis des linearen Teilraums $U = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ finden können, der von einer Familie $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ von Vektoren in $\mathbb{K}^n$ erzeugt wird. Dazu brauchen wir einige Vorbereitungen.

2.5.4 Vorbemerkungen. Es seien $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ und $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ zwei Familien von Vektoren eines $\mathbb{K}$-Vektorraumes V.

(1) *Wenn die Vektoren* $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ *alle als Linearkombinationen von* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ *darstellbar sind, dann gilt*

$$\operatorname{span}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r) \subseteq \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$$

$\operatorname{span}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r)$ ist der kleinste lineare Teilraum, der $\mathbf{w}_1$, $\mathbf{w}_2$, $\ldots$, $\mathbf{w}_r$ umfasst, vgl. zweite Bemerkung in 2.5.1. $\square$

(2) Aus (1) folgern wir sofort: *Wenn*

$$\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$$

und

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \operatorname{span}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r),$$

gilt, dann folgt

$$\operatorname{span}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m),$$

d.h. wenn alle Vektoren $\mathbf{w}_1$, $\mathbf{w}_2$, $\ldots$, $\mathbf{w}_r$ als Linearkombinationen der $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_m$ und umgekehrt darstellbar sind, dann erzeugen die beiden Familien von Vektoren denselben linearen Teilraum.

(3) Aus (2) erkennen wir, dass die folgenden sog. *elementaren Umformungen* einer Familie von Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ den durch diese erzeugten linearen Teilraum nicht ändern, vgl. dazu auch die elementaren (Zeilen-) Umformungen aus Kapitel 1:

(E1): Addition eines skalaren Vielfachen eines Vektors $\mathbf{v}_j$ zu einem anderen dieser Vektoren v_k $(k \neq j)$:

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_k, \ldots, \mathbf{v}_m) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_k + \lambda\mathbf{v}_j, \ldots, \mathbf{v}_m).$$

(E2): Das Vertauschen von Vektoren v_k und v_j:

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k, \ldots, \mathbf{v}_j, \ldots \mathbf{v}_m) = \operatorname{span}(\mathbf{v}_1, \ldots \mathbf{v}_j, \ldots, \mathbf{v}_k, \ldots \mathbf{v}_m).$$

(E3): Multiplikation eines Vektors v_k mit einem Skalar $\lambda \neq 0$:

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k, \ldots, \mathbf{v}_m) = \operatorname{span}(\mathbf{v}_1, \ldots, \lambda\mathbf{v}_k, \ldots, \mathbf{v}_m) \quad (\lambda \neq 0).$$

2.5.5 Algorithmus. Der folgende Algorithmus gestattet es uns, eine Basis des linearen Teilraums $U = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ zu finden, der von den Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{K}^n$ aufgespannt wird:

Bilde die $m \times n$-Matrix A, deren *Zeilen* die Vektoren $\mathbf{v}_1^T, \mathbf{v}_2^T, \ldots, \mathbf{v}_m^T$ sind. Wende den Gauß-Jordan Eliminations-Algorithmus an, um die Matrix A in eine Matrix B in Dreiecksgestalt zu überführen, was mittels elementarer Zeilenumformungen geschieht. Die Zeilen $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ von B, die ungleich Null sind, sind linear unabhängig (vgl. 2.5.2, Beispiel 4) und erzeugen denselben linearen Teilraum wie die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ (vgl. 2.5.4 (3)). Also bilden $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ eine Basis von $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

$$\begin{array}{cccc|c}
\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \mathbf{v}_1 \\
\alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \mathbf{v}_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \mathbf{v}_m
\end{array}$$

$$\Big| \\ \text{Gauß-Jordan Eliminations-Algorithmus} \\ \downarrow$$

$$\begin{array}{ccccccccccc|c}
0 & \dots & 0 & \beta_{1j_1} & \dots & \beta_{1j_2} & \dots & \beta_{1j_r} & \dots & \beta_{1n} & & \mathbf{w}_1 \\
0 & \dots & \dots & \dots & 0 & \beta_{2j_2} & \dots & \beta_{2j_r} & \dots & \beta_{2n} & & \mathbf{w}_2 \\
\vdots & & & & & & \ddots & \vdots & & \vdots & & \vdots \\
0 & \dots & \dots & \dots & \dots & 0 & & \beta_{rj_r} & \dots & \beta_{rn} & & \mathbf{w}_r \\
\hline
0 & \dots & \dots & \dots & \dots & \dots & & 0 & \dots & 0 & & \\
\vdots & & & & & & & \vdots & & \vdots & & \\
0 & \dots & \dots & \dots & \dots & \dots & & 0 & \dots & 0 & &
\end{array}$$

$$\beta_{1j_1} \neq 0, \beta_{2j_2} \neq 0, \dots, \beta_{rj_r} \neq 0$$

ABBILDUNG 2.1. Schema für Algorithmus 2.5.5

2.5.6 Beispiel. Wir betrachten im $\mathbb{R}^4$ den linearen Teilraum U, der von den folgenden Vektoren aufgespannt wird:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ -1 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \qquad \mathbf{v}_4 = \begin{pmatrix} 2 \\ 6 \\ 2 \\ -1 \end{pmatrix} \qquad \mathbf{v}_5 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ -3 \end{pmatrix}.$$

In Abbildung 2.2 führen wir die in 2.5.5 beschriebene Prozedur durch und erhalten die folgende Basis für U:

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -3 \end{pmatrix} \qquad \mathbf{w}_3 = \begin{pmatrix} 0 \\ 0 \\ 4 \\ -9 \end{pmatrix}.$$

Aber wir haben noch mehr als das erreicht: Gleichzeitig zur Transformation der Matrix A transformieren wir die 5×5-Einheitsmatrix E (auf der rechten Seite in der Abbildung) und erhalten eine 5×5-Matrix C. Aus den Zeilen von C können wir ablesen, wie $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ von den $\mathbf{v}_j$ abhängen:

$$\mathbf{w}_1 = \mathbf{v}_1 \qquad \mathbf{w}_2 = -2\mathbf{v}_1 + \mathbf{v}_2 \qquad \mathbf{w}_3 = -5\mathbf{v}_1 + 3\mathbf{v}_2 + \mathbf{v}_3.$$

Aus den letzten zwei Zeilen von C können wir ablesen, wie $\mathbf{v}_4$ und $\mathbf{v}_5$ von $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ abhängen:

$$-\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 + \mathbf{v}_4 = \mathbf{0}, \quad \text{d. h. } \mathbf{v}_4 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3,$$
$$\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 + \mathbf{v}_5 = \mathbf{0}, \quad \text{d. h. } \mathbf{v}_5 = -\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3.$$

Man kann sich überlegen, dass $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ ebenfalls eine Basis von U bilden.

1	2	0	1	$\mathbf{v}_1$						1	0	0	0	0
2	3	1	−1		$\mathbf{v}_2$					0	1	0	0	0
−1	1	1	−1			$\mathbf{v}_3$				0	0	1	0	0
2	6	2	−1				$\mathbf{v}_4$			0	0	0	1	0
0	2	2	−3					$\mathbf{v}_5$		0	0	0	0	1
1	2	0	1	$\mathbf{v}_1$						1	0	0	0	0
0	−1	1	−3	$-2\mathbf{v}_1$	$+\mathbf{v}_2$					−2	1	0	0	0
0	3	1	0	$\mathbf{v}_1$		$+\mathbf{v}_3$				1	0	1	0	0
0	2	2	−3	$-2\mathbf{v}_1$			$+\mathbf{v}_4$			−2	0	0	1	0
0	2	2	−3					$\mathbf{v}_5$		0	0	0	0	1
1	2	0	1	$\mathbf{v}_1$						1	0	0	0	0
0	−1	1	−3	$-2\mathbf{v}_1$	$+\mathbf{v}_2$					−2	1	0	0	0
0	0	4	−9	$-5\mathbf{v}_1$	$+3\mathbf{v}_2$	$+\mathbf{v}_3$				−5	3	1	0	0
0	0	4	−9	$-6\mathbf{v}_1$	$+2\mathbf{v}_2$		$+\mathbf{v}_4$			−6	2	0	1	0
0	0	4	−9	$-4\mathbf{v}_1$	$+2\mathbf{v}_2$			$+\mathbf{v}_5$		−4	2	0	0	1
1	2	0	1	$\mathbf{v}_1$					$\mathbf{w}_1$	1	0	0	0	0
0	−1	1	−3	$-2\mathbf{v}_1$	$+\mathbf{v}_2$				$\mathbf{w}_2$	−2	1	0	0	0
0	0	4	−9	$-5\mathbf{v}_1$	$+3\mathbf{v}_2$	$+\mathbf{v}_3$			$\mathbf{w}_3$	−5	3	1	0	0
0	0	0	0	$-\mathbf{v}_1$	$-\mathbf{v}_2$	$-\mathbf{v}_3$	$+\mathbf{v}_4$			−1	−1	−1	1	0
0	0	0	0	$+\mathbf{v}_1$	$-\mathbf{v}_2$	$-\mathbf{v}_3$		$+\mathbf{v}_5$		1	−1	−1	0	1

ABBILDUNG 2.2. Finden einer Basis des linearen Teilraums in Beispiel 2.5.6

2.5.7 Koordinaten bzgl. einer Basis.

Sei $\mathcal{B} : \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ eine Basis von V.

Behauptung. *Für jedes $\mathbf{v} \in V$ gibt es eindeutig bestimmte Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, für die gilt:*

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n \ .$$

Beweis. Sei $\mathbf{v} \in V$ gegeben. Da V von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ aufgespannt wird, gibt es Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, für die gilt

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n.$$

Die Eindeutigkeit der Skalare λ_i folgt aus der linearen Unabhängikeit der $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ unter Verwendung des Kriteriums (3) aus 2.5.2. $\square$

Definition. Die (eindeutig bestimmten) Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n$ mit $\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots \lambda_n \mathbf{v}_n$ heißen die *Koordinaten* von $\mathbf{v}$ bzgl. der Basis $\mathcal{B}$; wir schreiben auch:

$$[\mathbf{v}]_{\mathcal{B}} := \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} .$$

Auf diese Weise haben wir jedem $\mathbf{v} \in V$ das n-Tupel $[\mathbf{v}]_{\mathcal{B}} \in \mathbb{K}^n$ der Koordinaten von $\mathbf{v}$ bzgl. der Basis $\mathcal{B}$ zugeordnet. Die Abbildung

$$\begin{aligned} V &\to \mathbb{K}^n \\ \mathbf{v} &\mapsto [\mathbf{v}]_{\mathcal{B}} \end{aligned}$$

ist bijektiv. Die inverse Abbildung ist

$$\begin{aligned} \mathbb{K}^n &\to V \\ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} &\mapsto \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n. \end{aligned}$$

Merksatz. Eine Basis in einem Vektorraum festzulegen bedeutet ein Koordinatensystem wählen.

2.5.8 Beispiele.

(1) Sei $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ die Standardbasis $\mathcal{B}$ von $\mathbb{K}^n$. Für

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n$$

gilt

$$\begin{aligned} \mathbf{v} &= \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\ &= \alpha_1 \quad \mathbf{e}_1 \quad + \alpha_2 \quad \mathbf{e}_2 \quad + \ldots + \alpha_n \quad \mathbf{e}_n. \end{aligned}$$

Die Koordinaten von $\mathbf{v}$ bzgl. der Standardbasis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ sind $\alpha_1, \alpha_2, \ldots, \alpha_n$, wie es auch sein sollte.

(2) In einem Vektorraum gibt es i.A. viele verschiedene Basen. In $\mathbb{R}^2$ finden wir beispielsweise

- die Standardbasis $\mathcal{B}$: $\mathbf{e}_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $\mathbf{e}_2 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$
- die Basis $\mathcal{B}'$: $\mathbf{v}_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, $\mathbf{v}_2 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$.

Für $\mathbf{v} = \left(\begin{smallmatrix} 3 \\ -1 \end{smallmatrix}\right)$ sind 3 und -1 die Koordinaten bzgl. der Standardbasis $\mathcal{B}$. Die Koordinaten von $\mathbf{v}$ bzgl. der Basis $\mathcal{B}'$ sind $[\mathbf{v}]_{\mathcal{B}'} = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$, denn

$$\mathbf{v}_1 + 2\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \mathbf{v}.$$

2.6. Dimension und Sätze über Basen

Innerhalb dieses Abschnitts bezeichne V immer einen Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

2.6.1 Dimension. Wenn V eine aus n Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ bestehende Basis besitzt, dann sagen wir V sei ein Vektorraum der *Dimension* n und wir schreiben

$$\dim_{\mathbb{K}} V = n\,.$$

Hat V keine endliche Basis, bezeichnen wir V als einen unendlichdimensionalen Vektorraum und schreiben

$$\dim_{\mathbb{K}} V = \infty.$$

Beispiele.

(1) $\dim\{\mathbf{0}\} = 0$

(2) $\dim_{\mathbb{K}} \mathbb{K}^n = n$

(3) $\dim_{\mathbb{R}} \mathcal{P}_n(\mathbb{R}) = n + 1$

(4) $\dim_{\mathbb{R}} \mathcal{P}(\mathbb{R}) = \infty$, $\dim \mathcal{F}(\mathbb{R}, \mathbb{R}) = \infty$.

Vorsicht. Halten wir für einen Moment inne: Ist unsere Definition der Dimension eines Vektorraumes sinnvoll? Es könnte sein, dass V eine Basis $\mathcal{B}\colon \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{10}$ besitzt, die aus zehn Vektoren besteht, aber auch eine andere Basis $\mathcal{B}'\colon \mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_{11}$, die aus elf Vektoren besteht. Das würde bedeuten, daß V keine wohldefinierte Dimension hätte. Ist die Dimension nun 10 oder 11? Wir beweisen, daß dies nicht auftreten kann, sodaß die Dimension eines Vektorraums wohldefiniert ist:

2.6.2 Theorem. *Je zwei Basen eines Vektorraums haben dieselbe Anzahl von Elementen.*

Der Satz ergibt sich direkt aus der folgenden

2.6.3 Proposition. *Jede Auswahl von $m + 1$ Vektoren $\mathbf{w}_1$, $\mathbf{w}_2$, $\ldots$, $\mathbf{w}_m$, $\mathbf{w}_{m+1}$ aus einem von m Vektoren $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_m$ aufgespannten linearen Teilraum ist linear abhängig.*

Beweis. Wir beweisen diese Proposition mittels Induktion nach m.

$\underline{m = 1}$. Seien $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{span}(\mathbf{v}_1)$. Dann gibt es Skalare λ_1 und λ_2 mit $\mathbf{w}_1 = \lambda_1 \mathbf{v}_1$ und $\mathbf{w}_2 = \lambda_2 \mathbf{v}_1$. Es folgt, daß $\mathbf{w}_1$ und $\mathbf{w}_2$ linear abhängig sind.

Nun ist die Induktionsannahme, dass je m Vektoren, die in einem von $m - 1$ Vektoren aufgespannten Teilraum liegen, linear abhängig sind, und wir beweisen dieselbe Aussage für $m + 1$ anstelle von m.

Also seien $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{w}_{m+1} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$. Wir zeigen, daß $\mathbf{w}_1$, $\mathbf{w}_2$, $\ldots$, $\mathbf{w}_m$, $\mathbf{w}_{m+1}$ linear abhängig sind.

Zunächst stellen wir die $\mathbf{w}_i$ als Linearkombinationen der $\mathbf{v}_j$ dar:

$$\mathbf{w}_1 = \alpha_{11}\mathbf{v}_1 \;\; + \;\; \alpha_{12}\mathbf{v}_2 \; + \; \ldots \; + \; \alpha_{1m}\mathbf{v}_m$$

$$\begin{aligned}
\mathbf{w}_2 &= \alpha_{21}\mathbf{v}_1 + \alpha_{22}\mathbf{v}_2 + \dots + \alpha_{2m}\mathbf{v}_m \\
&\vdots \\
\mathbf{w}_{m+1} &= \alpha_{m+1,1}\mathbf{v}_1 + \alpha_{m+1,2}\mathbf{v}_2 + \cdots + \alpha_{m+1,m}\mathbf{v}_m
\end{aligned}$$

1. Fall: $\alpha_{11} = \alpha_{21} = \cdots = \alpha_{m+1,1} = 0$. Dann sind die m Vektoren $\mathbf{w}_2, \dots, \mathbf{w}_{m+1}$ in dem von den $(m-1)$ Vektoren $\mathbf{v}_2, \dots, \mathbf{v}_m$ aufgespannten Teilraum enthalten. Nach der Induktionsannahme sind die Vektoren $\mathbf{w}_2, \dots, \mathbf{w}_{m+1}$ linear abhängig. Damit sind auch $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{m+1}$ linear abhängig.

2. Fall: Wenigstens einer der Koeffizienten α_{i1} ist von 0 verschieden. Nach Umnummerierung der $\mathbf{w}_i$'s können wir annehmen, daß $\alpha_{11} \neq 0$ ist. Wie im Gauß-Jordan Eliminations-Algorithmus bilden wir nun

$$\begin{aligned}
\mathbf{w}_2' := \mathbf{w}_2 - \tfrac{\alpha_{21}}{\alpha_{11}}\mathbf{w}_1 &= \big(\alpha_{22} - \tfrac{\alpha_{21}}{\alpha_{11}} \cdot \alpha_{12}\big)\mathbf{v}_2 + \dots + \big(\alpha_{2n} - \tfrac{\alpha_{21}}{\alpha_{11}} \cdot \alpha_{1m}\big)\mathbf{v}_m \\
&\vdots \\
\mathbf{w}_{m+1}' := \mathbf{w}_{m+1} - \tfrac{\alpha_{m+1,1}}{\alpha_{11}}\mathbf{w}_1 &= \big(\alpha_{m+1,2} - \tfrac{\alpha_{m+1,1}}{\alpha_{11}} \cdot \alpha_{12}\big)\mathbf{v}_2 + \cdots + \big(\alpha_{m+1,n} - \tfrac{\alpha_{m+1,1}}{\alpha_{11}} \cdot \alpha_{1m}\big)\mathbf{v}_m
\end{aligned}$$

und erkennen, dass die m Vektoren $\mathbf{w}_2', \dots, \mathbf{w}_{m+1}'$ in dem von den $(m-1)$ Vektoren $\mathbf{v}_2$, $\dots, \mathbf{v}_m$ aufgespannten linearen Teilraum liegen. Nach der Induktionsannahme sind die Vektoren $\mathbf{w}_2', \dots, \mathbf{w}_{m+1}'$ linear abhängig. Also gibt es Skalare $\mu_2, \dots, \mu_{m+1}$, die nicht alle Null sind, so daß

$$\mu_2\mathbf{w}_2' + \mu_3\mathbf{w}_3' + \cdots + \mu_{m+1}\mathbf{w}_{m+1}' = 0$$

gilt. Unter Benutzung von $\mathbf{w}_i' = \mathbf{w}_i - \frac{\alpha_{i1}}{\alpha_{11}}\mathbf{w}_1$ erhalten wir

$$\mu_2\big(\mathbf{w}_2 - \tfrac{\alpha_{21}}{\alpha_{11}}\mathbf{w}_1\big) + \cdots + \mu_{m+1}\big(\mathbf{w}_{m+1} - \tfrac{\alpha_{m+1,1}}{\alpha_{11}}\mathbf{w}_1\big) = 0$$

und damit

$$\Big(-\sum_{i=2}^{n+1} \mu_i \tfrac{\alpha_{i1}}{\alpha_{11}}\Big)\mathbf{w}_1 + \mu_2\mathbf{w}_2 + \cdots + \mu_{m+1}\mathbf{w}_{m+1} = 0$$

Da nicht alle $\mu_2, \dots, \mu_{m+1}$ gleich Null sind, sind die Vektoren $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{m+1}$ linear abhängig. Damit ist der Beweis der Proposition beendet. □

Konsequenz. *In einem Vektorraum V der Dimension n besteht jedes linear unabhängige System aus höchstens n Vektoren.*

Das folgende Lemma erscheint zunächst abstrakt, seine Konsequenzen sind jedoch von hoher Bedeutung:

2.6.4 Lemma. *Es seien $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ linear unabhängig in V. Weiterhin seien $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r$ Vektoren in V mit*

$$\mathrm{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r) = V.$$

Dann kann man Indizes $1 \leq i_1 < i_2 < \cdots < i_s \leq r$ auswählen, so dass $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$, $\mathbf{w}_{i_1}, \dots, \mathbf{w}_{i_s}$ eine Basis von V bilden.

Beweis. Wir betrachten alle erzeugenden Familien von V der Form

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_s} \quad \text{mit } 1 \leq i_1 < i_2 < \cdots < i_s \leq r.$$

Nach Voraussetzung gibt es mindestens eine solche Familie, nämlich $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_m$, $\mathbf{w}_1$, $\ldots$, $\mathbf{w}_r$. Unter allen diesen Erzeugendensystemen wählen wir ein minimales, d. h. wenn wir einen der Vektoren $\mathbf{w}_{i_k}$ herausnehmen, erzeugen die verbleibenden Vektoren nicht mehr V. Wir zeigen, dass ein solches minimales Erzeugendensystem in der Tat eine Basis von V ist. Dazu genügt es nachzuweisen, dass $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_s}$ linear unabhängig sind.

Wir nehmen das Gegenteil an. Dann gibt es Skalare $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_s$, die nicht alle 0 sind, mit

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m + \mu_1\mathbf{w}_{i_1} + \cdots + \mu_s\mathbf{w}_{i_s} = 0$$

Die μ_i können nicht alle 0 sein, sonst wäre $\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m = 0$ und daraus würde $\lambda_1 = \cdots = \lambda_m = 0$ folgen, da die $\mathbf{v}_j$ nach Voraussetzung linear unabhängig sind. Also gibt es ein k, so dass $\mu_k \neq 0$ ist. Dann können wir aber $\mathbf{w}_{i_k}$ als Linearkombination der $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_{k-1}}, \mathbf{w}_{i_{k+1}}, \ldots, \mathbf{w}_{i_s}$ ausdrücken. Daraus folgt, dass das neue System immer noch erzeugend für V ist, was ein Widerspruch zur Minimalität des vorigen Systems ist. □

Betrachten wir nun einige Spezialfälle dieses Lemmas. Es sei zunächst $m = 0$; dies bedeutet, dass keine $\mathbf{v}$'s auftauchen. 2.6.4 besagt dann:

2.6.5 Korollar (Basisauswahlsatz). *Besitzt ein Vektorraum V ein endliches Erzeugendensystem von Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_r$, dann kann man Indizes $1 \leq i_1 < i_2 < \cdots < i_n \leq r$ auswählen, so dass $\mathbf{w}_{i_1}, \mathbf{w}_{i_2}, \ldots, \mathbf{w}_{i_n}$ eine Basis von V bilden.*

Folgerungen.

- Wird ein Vektorraum V von endlich vielen Vektoren erzeugt, dann ist V endlichdimensional und es ist $\dim V$ kleiner gleich der Anzahl von Elementen in jedem Erzeugendensystem von V.
- In einem endlichdimensionalen Vektorraum ist jedes minimale Erzeugendensystem eine Basis.

Als weiteren Spezialfall nehmen wir an, es seien linear unabhängige Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in einem Vektorraum V der Dimension n gegeben. Dann hat V eine n-elementige Basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Nach 2.6.4 können wir $n - m$ der Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in solcher Weise auswählen, dass sie zusammen mit $\mathbf{v}_1, \ldots, \mathbf{v}_m$ eine Basis von V bilden. Also erhalten wir

2.6.6 Korollar (Basisergänzungssatz). *Wenn $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in einem Vektorraum V der Dimension n linear unabhängig sind, kann man Vektoren $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_n$ finden, so dass $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n$ eine Basis von V bilden.*

Folgerung. In einem endlichdimensionalen Vektorraum ist jedes maximale linear unabhängige System eine Basis.

2.6.7 Proposition. *Für eine Familie $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ von n Vektoren in einem Vektorraum V der Dimension n sind die folgenden Aussagen äquivalent:*

(1) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *sind linear unabhängig.*

(2) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *erzeugen V.*

(3) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *bilden eine Basis von V.*

Beweis. Per Definition folgen aus Teil 3 die Teile 1 und 2. $1 \Rightarrow 3$ wegen 2.6.6, $2 \Rightarrow 3$ wegen 2.6.5. □

Wir beschäftigen uns nun mit der Dimension linearer Teilräume. Eine erste Beobachtung ergibt sich aus:

2.6.8 Proposition. *Sei V endlichdimensional. Für jeden linearen Teilraum U von V gilt* $\dim U \leq \dim V$. *Gleichheit gilt nur im Falle $U = V$.*

Beweis. Sei $\dim V = n$. Für jede Familie $\mathbf{u}_1, \ldots, \mathbf{u}_m$ linear unabhängiger Vektoren gilt $m \leq n$ nach 2.6.3. Daher können wir ein maximales System linear unabhängiger Vektoren in U betrachten. Aufgrund der Folgerung von 2.6.6 ist dieses eine Basis von U. Also ist U endlichdimensional und $\dim U \leq \dim V$. Ist $\dim U = \dim V$, dann hat U eine Basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ mit n Elementen. Da diese n Vektoren linear unabhängig in V sind, bilden sie auch eine Basis von V, was aus 2.6.7 folgt. Damit ist $U = V$. □

2.6.9 Dimensionssatz. *Es seien U_1 und U_2 lineare Teilräume von V. Dann gilt:*

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

Beweis. Gilt $\dim U_1 = \infty$ oder $\dim U_2 = \infty$, dann gilt auch $\dim(U_1 + U_2) = \infty$ und die Gleichung ist trivialerweise wahr. Nehmen wir also $\dim U_1 = m_1 < \infty$ und $\dim U_2 = m_2 < \infty$ an.

$U_1 \cap U_2$ ist ein linearer Teilraum von U_1 und von U_2. Sei $\dim(U_1 \cap U_2) = m$. Wählen wir eine Basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ von $U_1 \cap U_2$. Nach 2.6.6 gibt es Vektoren $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m_1}$ und $\mathbf{w}_{m+1}, \ldots, \mathbf{w}_{m_2}$, so dass gilt:

- $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m_1}$ ist eine Basis von U_1,
- $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{w}_{m+1}, \ldots, \mathbf{w}_{m_2}$ ist eine Basis von U_2.

Wir behaupten, dass $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m_1}, \mathbf{w}_{m+1}, \ldots, \mathbf{w}_{m_2}$ eine Basis von $U_1 + U_2$ ist. (Übung: Beweise diese Behauptung.)

Wir folgern:

$$\begin{aligned}\dim(U_1 + U_2) &= m + (m_1 - m) + (m_2 - m) \\ &= m_1 + m_2 - m \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)\end{aligned}$$

und daraus ergibt sich der Dimensionssatz. □

KAPITEL 3

Euklidische und unitäre Vektorräume

Um Längen von Vektoren und um Winkel messen zu können, brauchen wir neben der Addition und Skalarmultiplikation eine weitere Struktur auf Vektorräumen: ein *Skalarprodukt* [*scalar product*], auch *inneres Produkt* [*inner product*] genannt – nicht zu verwechseln mit der Skalarmultiplikation. Wir werden uns zunächst auf das Standardskalarprodukt im $\mathbb{R}^n$ und im $\mathbb{C}^n$ beschränken. Gerade die Quantenmechanik erfordert es, auf komplexe Vektorräume einzugehen. Dort wird es auch notwendig sein, auch unendlichdimensionale komplexe Vektorräume – vor allem Funktionenräume – mit einem Skalarprodukt zu betrachten, sogenannte *Hilbert-Räume*. Auf letztere werden wir jedoch nicht sehr intensiv eingehen.

3.1. Das Skalarprodukt für Vektoren in Ebene und Raum

Wir beginnen damit, das Skalarprodukt physikalisch und geometrisch zu motivieren.

3.1.1 Eine physikalische Motivation. Sie kennen das Gesetz

$$\text{Arbeit} = \text{Kraft mal Weg:} \qquad A = F \cdot s\,.$$

Bewegt man einen Wagen auf einer Schiene, so ist die geleistete Arbeit gleich der Länge des durchlaufenen Weges. Dabei wird angenommen, daß die Kraft in Richtung des Schienenwegs wirkt und konstant ist.

Arbeit ist eine skalare Größe A. Dagegen sind Kraft $\mathbf{F}$ und Weg $\mathbf{s}$ eigentlich gerichtete Größen. Wir bezeichnen mit F der Betrag der Kraft $\mathbf{F}$ und mit s die Länge des Weges $\mathbf{s}$. Die obige Formel gilt nicht mehr in dieser Form, wenn die Kraft nicht in Richtung des Weges wirkt. Man muß die Kraft $\mathbf{F}$ in ihre Komponente in Wegrichtung und ihre Komponente senkrecht zur Wegrichtung zerlegen. Die Kraftkomponente senkrecht zur Wegrichtung leistet keine Arbeit, sondern nur die Kraftkomponente in Wegrichtung. Ist γ der Winkel zwischen dem Kraftvektor $\mathbf{F}$ und dem Wegvektor $\mathbf{s}$, so ist die Größe der Kraftkomponente in Wegrichtung $F \cdot \cos\gamma$. Damit gilt für die geleistete Arbeit

$$A = F \cdot s \cdot \cos\gamma\,.$$

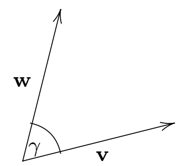
Wir wollen nun ein Produkt für Vektoren im Raum so definieren, daß die Formel

$$A = \mathbf{F} \cdot \mathbf{s}$$

allgemein gilt.

In diesem Unterabschnitt kehren wir in die Ebene und den 3-dimensionalen Raum und deren Geometrie zurück, wie wir sie aus der Schule kennen, vgl. 1.1. Dort haben wir gelernt, wie man die Länge von Strecken und wie man Winkel mißt.

3.1.2 Geometrische Einführung des Skalarprodukts in Ebene und Raum. Für zwei beliebige Vektoren $\mathbf{v}$ und $\mathbf{w}$ bezeichne $\|\mathbf{v}\|$ bzw. $\|\mathbf{w}\|$ deren Länge und γ bezeichne den Winkel zwischen $\mathbf{v}$ und $\mathbf{w}$ mit $0 \leq \gamma \leq \pi$. Wir definieren das Skalarprodukt von $\mathbf{v}$ und $\mathbf{w}$ durch:

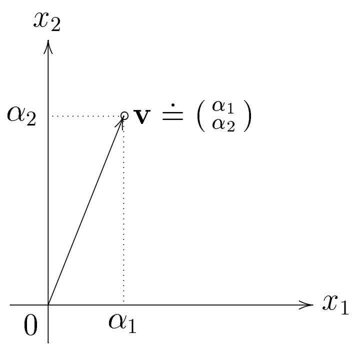


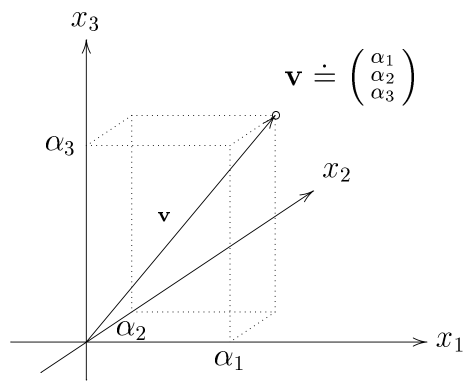
$$\langle \mathbf{v} \mid \mathbf{w} \rangle := \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos\gamma\,.$$

Man nennt dieses Produkt *Skalarprodukt*, da das so definierte Produkt zweier Vektoren kein Vektor sondern ein Skalar ist. Manchmal wird auch die Schreibweise $\mathbf{v} \cdot \mathbf{w}$ statt $\langle \mathbf{v} \mid \mathbf{w} \rangle$ verwendet, und man spricht auch vom *inneren Produkt* statt vom Skalarprodukt.

Wir wollen nun überlegen, wie wir das Skalarprodukt aus den Koordinaten der Vektoren berechnen können.

Dazu statten wir die Ebene und den Raum mit einem *kartesischen* Koordinatensystem aus. Das bedeutet, daß die Achsen zueinander orthogonal sind und daß die Einheitspunkte auf den Achsen den Abstand 1 vom Ursprung haben.





Nun können wir Vektoren durch ihre Koordinaten in diesem Koordinatensystem bezüglich der Basis $\mathcal{B} : \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ beschreiben. Da wir hier nur eine Basis betrachten wollen, identifizieren wir der Einfachheit halber die Vektoren in Ebene und Raum mit ihren Koordinatenvektoren:

$$\mathbf{v} = \begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix} \in \mathbb{R}^2 \quad \text{bzw.} \quad \mathbf{v} = \begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} \in \mathbb{R}^3.$$

3.1.3 Länge von Vektoren in Ebene und Raum. Nach dem Satz von Pythagoras ist die Länge der gerichteten Strecke, die den Vektor $\mathbf{v}$ repräsentiert:

$$\|\mathbf{v}\| = \|\left(\begin{smallmatrix}\alpha_1\\\alpha_2\end{smallmatrix}\right)\| = \sqrt{\alpha_1^2 + \alpha_2^2}$$

beziehungsweise

$$\|\mathbf{v}\| = \left\|\left(\begin{smallmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{smallmatrix}\right)\right\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}\,.$$

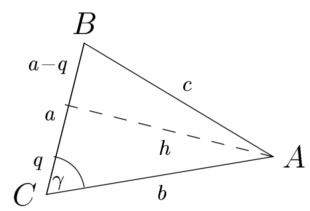
3.1.4 Berechnung des Skalarprodukts aus den Koordinaten. Wir behaupten

$$\text{Für Vektoren } \mathbf{v} = \begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}, \mathbf{w} = \begin{pmatrix}\beta_1\\\beta_2\\\beta_3\end{pmatrix} \text{ gilt } \langle \mathbf{v} \mid \mathbf{w} \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3\ .$$

Beweis. Wir benutzen wir den *Cosinus-Satz* [*cosine law*], den Sie aus der Schule kennen sollten: Für ein Dreieck mit den Bezeichnungen wie in der folgenden Abbildung gilt:

(4) $$a^2 + b^2 = c^2 + 2ab\cos\gamma$$

(Für einen rechten Winkel, d.h. wenn $\gamma = 90°$, dann ist $\cos\gamma = 0$ und (4) liefert den Satz des Pythagoras.)



Der Vollständigkeit halber geben wir einen Beweis des Cosinus-Satzes: Wir fällen das Lot (*Höhe*) vom Punkt A aus auf die Gerade $\overline{BC}$. Sie teilt das Dreieck in zwei rechtwinklige Dreiecke. Von dem in der Skizze unten gelegenen Dreieck erhalten wir:

(1) $$q = b\cos\gamma$$

und nach Pythagoras

$$(a-q)^2 + h^2 = c^2$$

d.h.

(2) $$a^2 - 2aq + q^2 + h^2 = c^2$$

Analog, erhält man vom zweiten rechtwinkligen Dreieck,

(3) $$q^2 + h^2 = b^2.$$

Ersetzen wir $q^2 + h^2$ in (2) durch b^2, so folgt

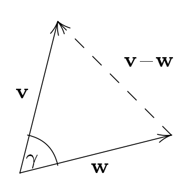
$$a^2 - 2aq + b^2 = c^2$$

Setzen wir hier $q = b\cos\gamma$ aus (1) ein, ergibt sich

$$a^2 - 2ab\cos\gamma + b^2 = c^2,$$

woraus der Cosinus-Satz (4) folgt.

Wir wenden nun den Cosinus-Satz auf die Situation in dem Bild unten an:



mit

$$a := \|\mathbf{v}\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \qquad b := \|\mathbf{w}\| = \sqrt{\beta_1^2 + \beta_2^2 + \beta_3^2}$$
$$c := \|\mathbf{v} - \mathbf{w}\| = \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 + (\alpha_3 - \beta_3)^2}$$

Wir haben nun:

$$2\langle \mathbf{v} \mid \mathbf{w}\rangle = 2\|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos\gamma$$

$$
\begin{aligned}
&= 2ab\cos\gamma = a^2 + b^2 - c^2 \qquad \text{nach dem Cosinus-Satz (4)} \\
&= \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \beta_1^2 + \beta_2^2 + \beta_3^2 - (\alpha_1 - \beta_1)^2 - (\alpha_2 - \beta_2)^2 - (\alpha_3 - \beta_3)^2 \\
&= 2(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3)
\end{aligned}
$$

und daraus ergibt sich die Behauptung $\langle \mathbf{v} \mid \mathbf{w} \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3$. □

3.1.5 Korollar. Sei γ der Winkel zwischen $\mathbf{v}$ und $\mathbf{w}$. Dann gilt:

(1) $\langle \mathbf{v} \mid \mathbf{w} \rangle > 0$ genau dann, wenn γ ein spitzer Winkel [*acute angle*] ist.

(2) $\langle \mathbf{v} \mid \mathbf{w} \rangle = 0$ genau dann, wenn $\mathbf{v}$ orthogonal zu $\mathbf{w}$ ist.

(3) $\langle \mathbf{v} \mid \mathbf{w} \rangle < 0$ genau dann, wenn γ ein stumpfer Winkel [*obtuse angle*] ist.

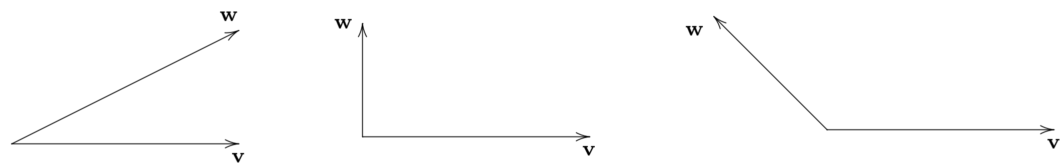


Abbildung 3.1. spitzer Winkel, rechter Winkel, stumpfer Winkel

In der Tat gilt für Winkel γ mit $0 \leq \gamma \leq \pi$

$$
\begin{aligned}
\cos\gamma > 0 &\iff 0 \leq \gamma < \tfrac{\pi}{2} \\
\cos\gamma = 0 &\iff \gamma = \tfrac{\pi}{2} \\
\cos\gamma < 0 &\iff \tfrac{\pi}{2} < \gamma \leq \pi,
\end{aligned}
$$

wie man an folgendem Bild sehen kann.

Wir werden das Skalarprodukt verwenden, um Längen und Winkel zu berechnen:

3.1.6 Korollar. (a) Die Länge eines Vektors $\mathbf{v}$ ist gegeben durch

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}\,.$$

(b) Für Vektoren $\mathbf{v} \neq \mathbf{0}$ und $\mathbf{w} \neq \mathbf{0}$ ist der Winkel γ zwischen $\mathbf{v}$ und $\mathbf{w}$ gegeben durch:

$$\cos\gamma = \frac{\langle \mathbf{v} \mid \mathbf{w} \rangle}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}.$$

3.2. Der Euklidische Vektorraum $\mathbb{R}^n$

Analog zu dem physikalisch und geometrisch motivierten Skalarprodukt für Vektoren in Ebene und Raum definieren wir:

3.2.1 Das Standard-Skalarprodukt im $\mathbb{R}^n$. Für beliebige Vektoren $\mathbf{v}$ und $\mathbf{w}$ im $\mathbb{R}^n$ definieren wir das *Skalarprodukt*, auch *kanonisches Skalarprodukt* oder *Standard-Skalarprodukt* genannt, durch

$$\langle \mathbf{v} \mid \mathbf{w} \rangle = \langle \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \mid \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \rangle := \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n = \sum_{i=1}^{n} \alpha_i\beta_i\ .$$

Oft schreibt man auch $\mathbf{v} \cdot \mathbf{w}$ anstelle von $\langle \mathbf{v} \mid \mathbf{w} \rangle$.

Wenn wir den $\mathbb{R}^n$ nicht nur als reellen Vektorraum betrachten, sondern ihn mit diesem Skalarprodukt ausstatten, dann sprechen wir vom *Euklidischen Vektorraum* $\mathbb{R}^n$.

Das Skalarprodukt hat als Eingabe ein Paar von Vektoren $\mathbf{v}, \mathbf{w}$ und gibt als Ergebnis ein Skalar zurück. Man muß es natürlich von der Skalarmultiplikation unterscheiden, die einen Vektor $\mathbf{v}$ mit einem Skalar λ multipliziert und als Ergebnis einen Vektor $\lambda\mathbf{v}$ liefert.

Man kann das Skalarprodukt auch als Matrixmultiplikation des Zeilenvektors $\mathbf{v}^T$ mit dem Spaltenvektor $\mathbf{w}$ auffassen:

$$\langle \mathbf{v} \mid \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w} = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}.$$

3.2.2 Regeln für das Skalarprodukt. Für beliebige $\mathbf{v}, \mathbf{w}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^n$ und beliebige $\lambda \in \mathbb{R}$ gilt:

$$\begin{array}{lrcl} (\mathrm{B}_1) & \langle \mathbf{v} + \mathbf{v}' \mid \mathbf{w} \rangle & = & \langle \mathbf{v} \mid \mathbf{w} \rangle + \langle \mathbf{v}' \mid \mathbf{w} \rangle \quad , \quad \langle \alpha\mathbf{v} \mid \mathbf{w} \rangle = \alpha \langle \mathbf{v} \mid \mathbf{w} \rangle \\ (\mathrm{B}_2) & \langle \mathbf{v} \mid \mathbf{w} + \mathbf{w}' \rangle & = & \langle \mathbf{v} \mid \mathbf{w} \rangle + \langle \mathbf{v} \mid \mathbf{w}' \rangle \quad , \quad \langle \mathbf{v} \mid \alpha\mathbf{w} \rangle = \alpha \langle \mathbf{v} \mid \mathbf{w} \rangle \\ (\mathrm{S}) & \langle \mathbf{v} \mid \mathbf{w} \rangle & = & \langle \mathbf{w} \mid \mathbf{v} \rangle \\ (\mathrm{P}) & \langle \mathbf{v} \mid \mathbf{v} \rangle & \geq & 0 \quad , \quad \langle \mathbf{v} \mid \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0} \end{array}$$

Die Regeln (B_1) und (B_2) können durch einfaches Nachrechnen verifiziert werden; man kann die Regeln auch als Spezialfälle der Regeln (A2) und (A4) für die Matrixmultiplikation betrachten.

Die Regeln (B_1) und (B_2) werden kurz durch den Begriff der *Bilinearität* des Skalarprodukts gekennzeichnet.

Wegen Eigenschaft (S), die direkt aus der Definition folgt, wird das Skalarprodukt auch als *symmetrisch* beeichnet.

Wegen der Eigenschaft (P) nennt man das Skalarprodukt auch *positiv definit*. Um diese Eigenschaft zu beweisen, betrachten wir

$$\langle \mathbf{v} \mid \mathbf{v} \rangle = \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2.$$

Wir erhalten die Summe von Quadraten reeller Zahlen. Eine solche ist immer ≥ 0. Außerdem ist $\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 = 0$ genau dann, wenn $\alpha_i^2 = 0$ und folglich $\alpha_i = 0$ für jedes $i = 1, \dots, n$ gilt, d.h. genau dann, wenn $\mathbf{v} = 0$.

3.2.3 Länge. Wegen (P) haben wir $\langle \mathbf{v} \mid \mathbf{v} \rangle \geq 0$ für alle $\mathbf{v} \in \mathbb{R}^n$. Damit hat $\langle \mathbf{v} \mid \mathbf{v} \rangle$ eine (eindeutig definierte nicht-negative) Quadratwurzel $\sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle}$. Wir nennen

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} = \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}$$

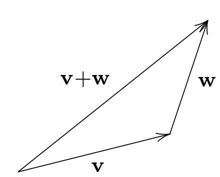
die *Länge* [*length*] oder auch die *Euklidische Norm* des Vektors $\mathbf{v}$.

Regeln. Für beliebige $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ und ein beliebiges $\lambda \in \mathbb{R}$ gilt:

(1) $\|\mathbf{v}\| \geq 0, \qquad \|\mathbf{v}\| = 0 \iff \mathbf{v} = 0$

(2) $\|\lambda\mathbf{v}\| = |\lambda| \cdot \|\mathbf{v}\|$

(3) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (Dreiecksungleichung)



Beweis. (1) folgt direkt aus der positiven Definitheit (P). Zum Beweis von (2) verwenden wir nur (B_1) und (B_2):

$$\|\lambda\mathbf{v}\| = \sqrt{\langle\lambda\mathbf{v} \mid \lambda\mathbf{v}\rangle} = \sqrt{\lambda\langle\mathbf{v} \mid \lambda\mathbf{v}\rangle} = \sqrt{\lambda^2\langle\mathbf{v} \mid \mathbf{v}\rangle} = \sqrt{\lambda^2}\cdot\sqrt{\langle\mathbf{v} \mid \mathbf{v}\rangle} = |\lambda|\cdot\|\mathbf{v}\|.$$

Schließlich folgt (3) aus der CAUCHY-SCHWARZschen Ungleichung, die wir hier anführen, deren Beweis wir aber in den nächsten Paragraphen zurückstellen, wo wir den allgemeineren komplexen Fall behandeln. Dort wird auch dargestellt, wie die Dreiecksungleichung hergeleitet wird. □

3.2.4 Die CAUCHY-SCHWARZ Ungleichung.

$$|\langle\mathbf{v} \mid \mathbf{w}\rangle| \le \|\mathbf{v}\|\cdot\|\mathbf{w}\| \text{ für alle } \mathbf{v},\mathbf{w} \in \mathbb{R}^n.$$

Wenn man die CAUCHY-SCHWARZsche Ungleichung in Koordinaten hinschreibt, sieht sie kompliziert aus:

$$|\alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n| \le \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}\cdot\sqrt{\beta_1^2 + \beta_2^2 + \cdots + \beta_n^2}$$

oder dazu äquivalent

$$(\alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n)^2 \le (\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2)(\beta_1^2 + \beta_2^2 + \cdots + \beta_n^2).$$

Übungen.

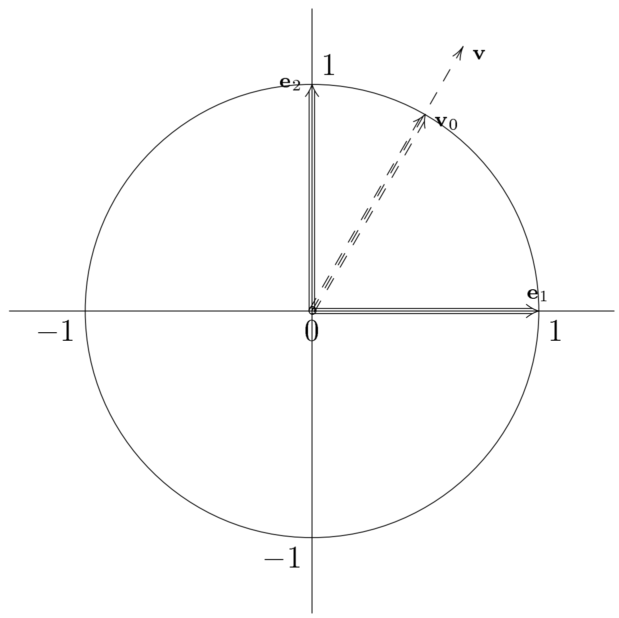
(1) Zeigen Sie, dass die Gleichung $|\langle\mathbf{v} \mid \mathbf{w}\rangle| = \|\mathbf{v}\|\cdot\|\mathbf{w}\|$ genau dann gilt, wenn $\mathbf{v}$ und $\mathbf{w}$ linear abhängig sind.

(2) Unter welchen Bedingungen gilt die Gleichung $\|\mathbf{v}+\mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\|$?

3.2.5 Einheitsvektoren. Jeder Vektor der Länge 1 heißt *Einheitsvektor* [*unit vector*].

Beispiele. Die kanonischen Basisvektoren $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in $\mathbb{R}^n$ sind Einheitsvektoren. Im $\mathbb{R}^2$ sind die Einheitsvektoren von der Form $\binom{\cos\gamma}{\sin\gamma}$ für $\gamma \in \mathbb{R}$.

Normierung. Für einen Vektor $\mathbf{v} \ne \mathbf{0}$, ist der Vektor $\mathbf{v}_0 := \frac{\mathbf{v}}{\|\mathbf{v}\|}$ ein Einheitsvektor („der in dieselbe Richtung zeigt wie $\mathbf{v}$“).



3.2.6 Orthogonalität im Euklidischen Raum $\mathbb{R}^n$. In Übereinstimmung mit der geometrischen Interpretation des Skalarprodukts im 2- und 3-dimensionalen Raum definieren wir:

Definition. Zwei Vektoren $\mathbf{v}$ und $\mathbf{w}$ im Euklidischen Raum $\mathbb{R}^n$ werden *orthogonal* genannt, wir schreiben $\mathbf{v} \perp \mathbf{w}$, falls $\langle \mathbf{v} \mid \mathbf{w} \rangle = 0$.

Nach dieser Definition ist der Nullvektor $\mathbf{0}$ orthogonal zu jedem Vektor.

3.2.7 Winkel zwischen zwei Vektoren im Euklidischen Raum $\mathbb{R}^n$. In Übereinstimmung mit der geometrischen Interpretation des Skalarprodukts definieren wir:

Definition. Seien $\mathbf{v}$ und $\mathbf{w}$ zwei Vektoren im Euklidischen Raum $\mathbb{R}^n$ ungleich dem Nullvektor $\mathbf{0}$. Der *Winkel zwischen* $\mathbf{v}$ *und* $\mathbf{w}$ wird gemessen durch die Zahl γ zwischen 0 und π, die gegeben ist durch

$$\cos \gamma = \frac{\langle \mathbf{v} \mid \mathbf{w} \rangle}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}.$$

Bemerkung. Im Intervall $[0, \pi]$ ist die Funktion $y = \cos \gamma$ streng monoton fallend und der Wertebereich liegt im Intervall $[-1, +1]$. Sie hat eine Umkehrfunktion $\gamma = \arccos y$, die auf dem Intervall $[-1, +1]$ definiert ist und den Wertebereich $[0, \pi]$ hat. Aus der Cauchy-Schwarz Ungleichung folgt:

$$\frac{|\langle \mathbf{v} \mid \mathbf{w} \rangle|}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|} \leq 1,$$

d.h.

$$-1 \leq \frac{\langle \mathbf{v} \mid \mathbf{w} \rangle}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|} \leq 1.$$

Aus diesem Grund macht obige Definition Sinn: Es gibt genau ein $\gamma \in [0, \pi]$, für das die definierte Ungleichung gilt, nämlich

$$\gamma = \arccos(\frac{\langle \mathbf{v} \mid \mathbf{w} \rangle}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}).$$

Konvention. Wenn wir über Länge, Orthogonalität und Winkel im Euklidischen Raum $\mathbb{R}^n$ sprechen, beziehen wir uns immer auf die Definitionen in 3.2, 3.2.6, 3.2.7, und wir verwenden die Ebene und den Raum lediglich zur Anschauung.

3.3. Der unitäre Vektorraum $\mathbb{C}^n$

Wir möchten ein Skalarprodukt auf $\mathbb{C}^n$ definieren, das uns dieselben Begriffe von Länge und Orthogonalität im komplexen Fall zu erklären erlaubt. Wir können allerdings nicht dieselbe Definition wie im reellen Fall benutzen. Die Länge eines Vektors sollte eine nichtnegative reelle Zahl sein. Im Falle einer komplexen Zahl $\alpha = a + ib$, $a, b \in \mathbb{R}$, ist der Betrag nicht durch $\sqrt{\alpha^2}$ wie im reellen Fall gegeben sondern durch

$$|\alpha| = |a + ib| = \sqrt{a^2 + b^2} = \sqrt{(a - ib)(a + ib)} = \sqrt{\overline{\alpha}\alpha}$$

Diese Beobachtung veranlaßt uns zu folgender Definition. Zuvor führen wir jedoch die folgende Schreibweise ein:

$$\text{Für jeden Vektor } \mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{C}^n \text{ schreiben wir } \overline{\mathbf{v}} = \begin{pmatrix} \overline{\alpha_1} \\ \overline{\alpha_2} \\ \vdots \\ \overline{\alpha_n} \end{pmatrix}$$

3.3.1 Das Standardskalarprodukt auf $\mathbb{C}^n$. Wir definieren das *Standard-*(*Hermitesche*) *Skalarprodukt* zweier Vektoren

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{C}^n \text{ durch } \langle \mathbf{v} \mid \mathbf{w} \rangle := \overline{\alpha_1}\beta_1 + \overline{\alpha_2}\beta_2 + \cdots + \overline{\alpha_n}\beta_n = \overline{\mathbf{v}}^T \mathbf{w}$$

Damit ist das Skalarprodukt zweier Vektoren in $\mathbb{C}^n$ eine komplexe Zahl. Der Vektorraum $\mathbb{C}^n$ ausgestattet mit diesem Skalarprodukt wird *unitärer* Vektorraum $\mathbb{C}^n$ genannt. Es sind dies gerade die endlichdimensionalen komplexen *Hilberträume* (bis auf Isomorphie).

Aus der Definition des Skalarproduktes kann man direkt die folgenden Eigenschaften ableiten:

3.3.2 Regeln für das Skalarprodukt. Das Standardskalarprodukt auf $\mathbb{C}^n$ ist (B'$_{1,2}$) semi-bilinear, (H) hermitesch und (P) positiv definit, d.h. für alle $\alpha \in \mathbb{C}$ und alle $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{w}$, $\mathbf{w}_1$, $\mathbf{w}_2 \in \mathbb{C}^n$ gilt:

$$\begin{array}{llcllcl}
(\mathrm{B'}_1) & \langle \mathbf{v}_1 + \mathbf{v}_2 \mid \mathbf{w} \rangle & = & \langle \mathbf{v}_1 \mid \mathbf{w} \rangle + \langle \mathbf{v}_2 \mid \mathbf{w} \rangle & , & \langle \alpha\mathbf{v} \mid \mathbf{w} \rangle & = & \overline{\alpha}\langle \mathbf{v} \mid \mathbf{w} \rangle \\
(\mathrm{B}_2) & \langle \mathbf{v} \mid \mathbf{w}_1 + \mathbf{w}_2 \rangle & = & \langle \mathbf{v} \mid \mathbf{w}_1 \rangle + \langle \mathbf{v} \mid \mathbf{w}_2 \rangle & , & \langle \mathbf{v} \mid \alpha\mathbf{w} \rangle & = & \alpha\langle \mathbf{v} \mid \mathbf{w} \rangle \\
(\mathrm{H}) & \langle \mathbf{v} \mid \mathbf{w} \rangle & = & \overline{\langle \mathbf{w} \mid \mathbf{v} \rangle} & & & & \\
(\mathrm{P}) & \langle \mathbf{v} \mid \mathbf{v} \rangle & & \text{ist eine reelle Zahl} \geq 0 & , & \langle \mathbf{v}, \mathbf{v} \rangle = 0 & \iff & \mathbf{v} = \mathbf{0}
\end{array}$$

Beachte, dass dieses Skalarprodukt wegen der zweiten Gleichung in (B'$_1$) nicht bilinear ist – deshalb die Bezeichnung *semibilinar*. Beschränken wir uns auf reelle Skalare und reelle Vektoren, dann sind diese Eigenschaften genau die Eigenschaften (B$_{1,2}$), (S), (P) aus **??**.

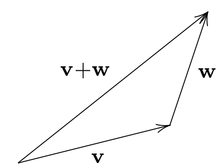
In Analogie zum reellen Fall definieren wir nun Länge und Orthogonalität im $\mathbb{C}^n$:

3.3.3 Länge. Wegen (P) ist $\langle \mathbf{v} \mid \mathbf{v} \rangle$ für alle $\mathbf{v} \in \mathbb{C}^n$ eine nichtnegative reelle Zahl. Diese hat eine (eindeutig definierte nicht-negative) Quadratwurzel $\sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle}$. Wir nennen

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} = \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}$$

die *Länge* [*length*] oder *hermitesche Norm* des Vektors $\mathbf{v}$.

Regeln. Für beliebige $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ und ein beliebiges $\lambda \in \mathbb{C}$ gilt:



(1) $\|\mathbf{v}\| \geq 0, \qquad \|\mathbf{v}\| = 0 \iff \mathbf{v} = 0$

(2) $\|\lambda\mathbf{v}\| = |\lambda| \cdot \|\mathbf{v}\|$

(3) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (Dreiecksungleichung)

(1) folgt direkt aus der positiven Definitheit (P). Zum Beweis von (2) verwenden wir nur (B$_1$) und (B$_2$):

$$\|\lambda\mathbf{v}\| = \sqrt{\langle \lambda\mathbf{v} \mid \lambda\mathbf{v} \rangle} = \sqrt{\overline{\lambda}\langle \mathbf{v} \mid \lambda\mathbf{v} \rangle} = \sqrt{\overline{\lambda}\lambda\langle \mathbf{v} \mid \mathbf{v} \rangle} = \sqrt{\overline{\lambda}\lambda} \cdot \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} = |\lambda| \cdot \|\mathbf{v}\|.$$

Zum Beweis der der Dreiecksungleichung (2) verwenden wir die

3.3.4 Die Cauchy-Schwarzsche Ungleichung. Es gilt für alle Vektoren $\mathbf{v}, \mathbf{w}$:

$$|\langle \mathbf{v} \mid \mathbf{w} \rangle| \leq \|\mathbf{v}\| \cdot \|\mathbf{w}\|$$

oder gleichwertig

$$|\overline{\alpha_1}\beta_1 + \overline{\alpha_2}\beta_2 + \cdots + \overline{\alpha_n}\beta_n| \leq \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2} \cdot \sqrt{|\beta_1|^2 + |\beta_2|^2 + \cdots + |\beta_n|^2}$$

Beweis. Es genügt, das Folgende zu zeigen:

$$|\langle \mathbf{v} \mid \mathbf{w} \rangle|^2 \leq \|\mathbf{v}\|^2\|\mathbf{w}\|^2 \quad \text{d.h.} \quad \overline{\langle \mathbf{v} \mid \mathbf{w} \rangle}\langle \mathbf{v} \mid \mathbf{w} \rangle \leq \langle \mathbf{v} \mid \mathbf{v} \rangle\langle \mathbf{w} \mid \mathbf{w} \rangle$$

(Ziehen wir die Wurzel auf beiden Seiten, ergibt die gewünschte Ungleichung.) Im Fall $\mathbf{w} = \mathbf{0}$ verschwinden beide Seiten und die Ungleichung is sicher richtig. Betrachten wir nun den Fall $\mathbf{w} \neq \mathbf{0}$. In diesem Fall gilt $\langle \mathbf{w} \mid \mathbf{w} \rangle > 0$ nach (P) und damit können wir die komplexe Zahl $\gamma := \frac{\langle \mathbf{w} | \mathbf{v} \rangle}{\langle \mathbf{w} | \mathbf{w} \rangle}$ betrachten. Es folgt

$$\begin{aligned}
0 &\leq \langle \mathbf{v} - \gamma\mathbf{w} \mid \mathbf{v} - \gamma\mathbf{w} \rangle \quad \text{nach (P)} \\
&= \langle \mathbf{v} \mid \mathbf{v} \rangle - \langle \mathbf{v} \mid \gamma\mathbf{w} \rangle - \langle \gamma\mathbf{w} \mid \mathbf{v} \rangle + \langle \gamma\mathbf{w} \mid \gamma\mathbf{w} \rangle \quad \text{nach (B'}_{1,2}) \\
&= \langle \mathbf{v} \mid \mathbf{v} \rangle - \gamma\langle \mathbf{v} \mid \mathbf{w} \rangle - \overline{\gamma}\langle \mathbf{w} \mid \mathbf{v} \rangle + \overline{\gamma}\gamma\langle \mathbf{w} \mid \mathbf{w} \rangle \quad \text{nach (B'}_{1,2}) \\
&= \langle \mathbf{v} \mid \mathbf{v} \rangle - \gamma\langle \mathbf{v} \mid \mathbf{w} \rangle - \overline{\gamma}\langle \mathbf{w} \mid \mathbf{v} \rangle + \overline{\gamma}\langle \mathbf{w} \mid \mathbf{v} \rangle \quad \text{nach der Defintion von } \gamma \\
&= \langle \mathbf{v} \mid \mathbf{v} \rangle - \gamma\langle \mathbf{v} \mid \mathbf{w} \rangle
\end{aligned}$$

Daraus folgern wir $\gamma\langle \mathbf{v} \mid \mathbf{w}\rangle \leq \langle \mathbf{v} \mid \mathbf{v}\rangle$. Ersetzen wir γ durch seinen Wert, so folgt $\frac{\langle \mathbf{w},\mathbf{v}\rangle}{\langle \mathbf{w}|\mathbf{w}\rangle}\langle \mathbf{v} \mid \mathbf{w}\rangle \leq \langle \mathbf{v} \mid \mathbf{v}\rangle$. Daraus folgt direkt $\langle \mathbf{w} \mid \mathbf{v}\rangle\langle \mathbf{v} \mid \mathbf{w}\rangle \leq \langle \mathbf{v},\mathbf{v}\rangle\langle \mathbf{w},\mathbf{w}\rangle$. Mit (H) folgern wir $\overline{\langle \mathbf{v} \mid \mathbf{w}\rangle}\langle \mathbf{v} \mid \mathbf{w}\rangle \leq \langle \mathbf{v},\mathbf{v}\rangle\langle \mathbf{w},\mathbf{w}\rangle$ und damit die gewünschte Ungleichung. □

Beweis der Dreiecksungleichung (3):

$$
\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w} \mid \mathbf{v}+\mathbf{w}\rangle \\
&= \langle \mathbf{v} \mid \mathbf{v}\rangle + \langle \mathbf{v} \mid \mathbf{w}\rangle + \langle \mathbf{w} \mid \mathbf{v}\rangle + \langle \mathbf{w} \mid \mathbf{w}\rangle \qquad \text{nach (B'}_1\text{) und (B}_2\text{)} \\
&\leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\,\|\mathbf{w}\| + \|\mathbf{w}\|^2 \text{ nach der Cauchy-Schwarzschen Ungleichung} \\
&= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2
\end{aligned}
$$

Wurzelziehen auf beiden Seiten ergibt die Dreiecksungleichung. □

Bemerkung. In den Beweisen von 3.3 und **??** haben wir nur die Regeln (B'$_1$),(B$_2$), (H) und (P) für das Skalarprodukt verwendet sowie die Regeln für das Rechnen mit Ungleichungen für reelle Zahlen. Wir haben nicht die konkrete Darstellung des Skalarprodukts benutzt. Dadurch wurden die Beweise einfacher und nicht komplizierter, wie mancher erwarten würde. Das kann man daran sehen, um wieviel komplizierter die Cauchy-Schwarzsche Ungleichung aussieht, wenn man sie in Koordinaten hinschreibt.

3.3.5 Einheitsvektoren. Jeder Vektor der Länge 1 heißt *Einheitsvektor* [*unit vector*].

Normierung. Für einen Vektor $\mathbf{v} \neq \mathbf{0}$, ist der Vektor $\mathbf{v}_0 := \frac{\mathbf{v}}{\|\mathbf{v}\|}$ ein Einheitsvektor („der in dieselbe Richtung zeigt wie $\mathbf{v}$").

3.3.6 Orthogonalität. Zwei Vektoren $\mathbf{v},\mathbf{w} \in \mathbb{C}^n$ werden *orthogonal* genannt und wir schreiben dann $\mathbf{v} \perp \mathbf{w}$, wenn $\langle \mathbf{v} \mid \mathbf{w}\rangle = 0$ gilt.

3.3.7 Satz von Pythagoras. *Aus* $\mathbf{v} \perp \mathbf{w}$ *folgt* $\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$.

Beweis. Sei $\mathbf{v} \perp \mathbf{w}$. Nach Definition gilt $\langle \mathbf{v},\mathbf{w}\rangle = 0$. Aus (H) folgt dann auch $\langle \mathbf{w} \mid \mathbf{v}\rangle = \overline{\langle \mathbf{v} \mid \mathbf{w}\rangle} = 0$. Unter Anwendung von (B'$_{1,2}$) folgern wir

$$\|\mathbf{v}+\mathbf{w}\|^2 = \langle \mathbf{v}+\mathbf{w} \mid \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{v} \mid \mathbf{v}\rangle + \langle \mathbf{v} \mid \mathbf{w}\rangle + \langle \mathbf{w},\mathbf{v}\rangle + \langle \mathbf{w} \mid \mathbf{w}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

□

Übung.

a) Zeige, dass in der Dreiecksungleichung Gleichheit genau dann gilt, wenn $\mathbf{v} = \alpha\mathbf{w}$ für eine reelle Zahl $\alpha > 0$ ist.

b) Zeige, dass die Cauchy-Schwarzsche Ungleichung genau dann mit Gleichheit erfüllt ist, wenn $\mathbf{v}$ und $\mathbf{w}$ linear abhängig sind.

3.4. Euklidische und unitäre Vektorräume

Nun sind wir soweit, eine allgemeine Definition zu geben:

3.4.1 Euklidische und unitäre Vektorräume. Ein *euklidischer* Vektorraum ist ein reeller Vektorraum V ausgestattet mit einer symmetrischen, positiv definiten Bilinearform, d.h. einer Abbildung

$$(\mathbf{v}, \mathbf{w}) \mapsto \langle \mathbf{v} \mid \mathbf{w} \rangle \ : \ V \times V \to \mathbb{R},$$

welche die Eigenschaften ($B_{1,2}$), (S) und (P) aus **??** erfüllt.

Ein *unitärer* Vektorraum ist ein komplexer Vektorraum V mit einer hermiteschen, positiv definiten Semibilinearform, d.h. einer Abbildung

$$(\mathbf{v}, \mathbf{w}) \mapsto \langle \mathbf{v} \mid \mathbf{w} \rangle \ : \ V \times V \to \mathbb{C},$$

welche die Eigenschaften (B'$_{1,2}$), (H) und (P) aus 3.3.2 erfüllt.

Weitere Beispiele von euklidischen und unitären Vektorräumen werden wir später betrachten. Im Augenblick halten wir nur fest, daß jeder lineare Teilraum U von unseren Standardbeispielen $\mathbb{R}^n$ und $\mathbb{C}^n$ wieder ein euklidischer oder unitärer Vektorraum ist: Um U zu einem solchen Raum zu machen, schränken wir einfach das Standardskalarprodukt auf U ein.

Beachte, dass euklidische und unitäre Vektorräume unendlichdimensional sein können. Derartige Räume (insbesondere Funktionenräume) spielen eine bedeutende Rolle in der Quantenmechanik.

Für Elemente $\mathbf{v}, \mathbf{w}$ eines beliebigen euklidischen oder unitären Vektorraumes definieren wir die *Länge* (oder die *Norm*) und *Orthogonalität* genau wie in **??**, **??**, **??** und 3.3.6:

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} \quad , \quad \mathbf{v} \perp \mathbf{w} :\Longleftrightarrow \langle \mathbf{v} \mid \mathbf{w} \rangle = 0\,.$$

Vektoren der Länge 1 heißen *Einheitsvektoren*. Die CAUCHY-SCHWARZsche Ungleichung und die Dreiecksungleichung bleiben gültig, da wir in deren Beweis in 3.3.7 und 3.3.4 nur die Eigenschaften (B'$_{1,2}$), (H) und (P) benutzt haben. Hier zeigt sich erst recht, wie günstig es ist, wenn man sich in Beweisen auf die wesentlichen Eigenschaften des Skalarproduktes stützt und nicht auf die konkrete Darstellung durch Koordinaten. Dadurch werden die hergeleiteten Formeln in beliebigen euklidischen bzw. unitären Vektorräumen allgemeingültig:

$$\begin{aligned} |\langle \mathbf{v} \mid \mathbf{w} \rangle| &\le \|\mathbf{v}\| \cdot \|\mathbf{w}\| && \text{CAUCHY-SCHWARZsche Ungleichung} \\ \|\mathbf{v}\| &\ge 0, \|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0} && \\ \|\lambda\mathbf{v}\| &= |\lambda| \cdot \|\mathbf{v}\| && \\ \|\mathbf{v} + \mathbf{w}\| &\le \|\mathbf{v}\| + \|\mathbf{w}\| && \text{Dreiecksungleichung} \end{aligned}$$

Von nun an sei V ein euklidischer oder unitärer Vektorraum. Man denke der Einfachheit halber an $\mathbb{R}^n$ oder $\mathbb{C}^n$ mit dem Standardskalarprodukt wie in 3.2 und 3.3.

3.4.2 Orthonormalbasen. Ein System $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ von Vektoren in V wird ein *Orthogonalsystem* genannt, wenn die Vektoren alle von 0 verschieden und wenn sie paarweise orthogonal sind, d.h. wenn gilt:

$$\langle \mathbf{v}_i \mid \mathbf{v}_i \rangle \neq 0 \ \text{ für alle } i, \ \text{ aber } \ \langle \mathbf{v}_i \mid \mathbf{v}_j \rangle = 0, \quad \text{falls } \ i \neq j.$$

Sind zusätzlich alle $\mathbf{v}_i$ Einheitsvektoren, sprechen wir von einem *Orthonormalsystem*. Ein Orthonormalsystem wird durch die folgende Bedingung charakterisiert:

$$\langle \mathbf{v}_i \mid \mathbf{v}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

Beobachtung. Jedes Orthogonalsystem $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ und folglich jedes Orthonormalsystem ist linear unabhängig.

Zum Nachweis nehmen wir $\sum_{j=1}^m \alpha_j \mathbf{v}_j = \mathbf{0}$ an. Dann folgt

$$0 = \langle \mathbf{v}_i \mid \sum_{j=1}^m \alpha_j \mathbf{v}_j \rangle = \sum_{j=1}^m \alpha_j \langle \mathbf{v}_i \mid \mathbf{v}_j \rangle = \alpha_i \langle \mathbf{v}_i \mid \mathbf{v}_i \rangle$$

Aus $\langle \mathbf{v}_i \mid \mathbf{v}_i \rangle \neq 0$ folgern wir $\alpha_i = 0$ für jedes i.

Ein Orthonormalsystem $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, welches eine Basis von V ist, heißt *Orthonormalbasis*.

In $\mathbb{R}^n$ und $\mathbb{C}^n$ ist jedes Orthonormalsystem aus n Vektoren eine Orthonormalbasis. Denn in einem n-dimensional Vektorraum bilden je n linear unabhängige Vektoren eine Basis. Die Standardbasis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ ist eine Orthonormalbasis. Im $\mathbb{R}^2$ bilden die Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad , \quad \mathbf{v}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

ebenso eine Orthonormalbasis. Diese ist gegenüber der Standardbasis um den Winkel γ gedreht.

3.4.3 Koordinaten bzgl. einer Orthonormalbasis. Sei $\mathcal{B} : \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ eine Orthonormalbasis von V. Dann hat ein beliebiger Vektor $\mathbf{v} \in V$ eine eindeutige Darstellung als Linearkombination

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n = \sum_{j=1}^m \alpha_j \mathbf{v}_j$$

Die Koeffizienten α_i sind die Koordinaten von $\mathbf{v}$ bezüglich der basis $\mathcal{B}$. Bildet man das Skalarprodukt mit $\mathbf{v}_i$ auf beiden Seiten ergibt

$$\langle \mathbf{v}_i \mid \mathbf{v} \rangle = \langle \mathbf{v}_i \mid \sum_{j=1}^m \alpha_j \mathbf{v}_j \rangle = \sum_{j=1}^m \alpha_j \langle \mathbf{v}_i \mid \mathbf{v}_j \rangle = \sum_{j=1}^m \alpha_j \delta_{ij} = \alpha_i$$

Also ist die i-te Koordinate von $\mathbf{v}$ bzgl. der Orthonormalbasis $\mathcal{B} : \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ gleich $\alpha_i = \langle \mathbf{v}_i \mid \mathbf{v} \rangle$ und die obige Darstellung schreibt sich

$$\mathbf{v} = \langle \mathbf{v}_1 \mid \mathbf{v} \rangle \mathbf{v}_1 + \langle \mathbf{v}_2 \mid \mathbf{v} \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}_n \mid \mathbf{v} \rangle \mathbf{v}_n$$

Beispiel. Im euklidischen $\mathbb{R}^3$ haben die Vektoren

$$\mathbf{v}_1 = \frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix} \quad , \quad \mathbf{v}_2 = \frac{1}{3}\begin{pmatrix}2\\1\\-2\end{pmatrix} \quad , \quad \mathbf{v}_3 = \frac{1}{3}\begin{pmatrix}2\\-2\\1\end{pmatrix}$$

Länge 1 und sie sind paarweise orthogonal (nachprüfen!). Sie bilden also eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^3$. Bezüglich dieser Basis hat der Vektor

$$\mathbf{v} = \begin{pmatrix}-4\\2\\2\end{pmatrix} \in \mathbb{R}^3$$

die Koordinatendarstellung $\mathbf{v} = \langle \mathbf{v}_1 \mid \mathbf{v}\rangle \mathbf{v}_1 + \langle \mathbf{v}_2 \mid \mathbf{v}\rangle \mathbf{v}_2 + \langle \mathbf{v}_3 \mid \mathbf{v}\rangle \mathbf{v}_3 = \frac{4}{3}\mathbf{v}_1 - \frac{10}{3}\mathbf{v}_2 - \frac{10}{3}\mathbf{v}_3$.

3.4.4 Orthogonalräume. Sei V ein euklidischer oder unitärer Vektorraum. Für jede Teilmenge $M \subseteq V$ können wir die Menge $M^\perp$ aller Vektoren betrachten, die auf allen $\mathbf{a} \in M$ senkrecht stehen:

$$M^\perp := \{\mathbf{v} \in V : \mathbf{a} \perp \mathbf{v} \text{ für alle } \mathbf{a} \in M\} = \{\mathbf{v} \in V : \langle \mathbf{a} \mid \mathbf{v}\rangle = 0 \text{ für alle } \mathbf{a} \in M\}$$

Behauptung: $M^\perp$ ist ein linearer Teilraum von V, genannt der *Orthogonalraum* zu M.

Beweis. (U1): $0 \in M^\perp$, da $\langle \mathbf{a} \mid 0\rangle = 0$ für alle $\mathbf{a}$.

(U2): Sei $\mathbf{v}, \mathbf{w} \in M^\perp$. Für alle $\mathbf{a} \in M$ gilt dann $\langle \mathbf{a} \mid \mathbf{v}\rangle = 0$ und $\langle \mathbf{a} \mid \mathbf{w}\rangle = 0$ und folglich $\langle \mathbf{a} \mid \mathbf{v} + \mathbf{w}\rangle \underset{(B2)}{=} \langle \mathbf{a} \mid \mathbf{v}\rangle + \langle \mathbf{a} \mid \mathbf{w}\rangle = 0$, d.h. $\mathbf{v} + \mathbf{w} \in M^\perp$.

(U3): Sei $\mathbf{v} \in M^\perp$ und λ ein Skalar. Dann gilt für jedes $\mathbf{a} \in M$ die Beziehung $\langle \mathbf{a} \mid \mathbf{v}\rangle = 0$ und folglich $\langle \mathbf{a} \mid \lambda\mathbf{v}\rangle \underset{(B2)}{=} \lambda\langle \mathbf{a} \mid \mathbf{v}\rangle = 0$, d.h. $\lambda\mathbf{v} \in M^\perp$ □

Spezialfall. Wähle $V = \mathbb{R}^n$ und ein festes $\mathbf{a} \in \mathbb{R}^n$. Dann gilt $\{a\}^\perp = \{\mathbf{v} \in \mathbb{R}^n : \langle \mathbf{a} \mid \mathbf{v}\rangle = 0\}$. Für

$$\mathbf{a} = \begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{pmatrix} \qquad \text{und} \qquad \mathbf{v} = \begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

kann die Gleichung $\langle \mathbf{a} \mid \mathbf{v}\rangle = 0$ geschrieben werden als

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$

Also besteht $\{\mathbf{a}\}^\perp$ aus den Lösungen $\mathbf{v}$ dieser homogenen linearen Gleichung (also einer Hyperebene von $\mathbb{R}^n$).

Analog dazu besteht für $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ der Orthogonalraum $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}^\perp$ aus den Lösungen des homogenen linearen Gleichungssystems

$$\begin{aligned} \langle \mathbf{a}_1 \mid \mathbf{v}\rangle &= 0 \\ \langle \mathbf{a}_2 \mid \mathbf{v}\rangle &= 0 \\ &\ldots\ldots \\ \langle \mathbf{a}_m \mid \mathbf{v}\rangle &= 0 \end{aligned}$$

Lemma. *Seien* $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_m \in V$ *und sei* U *der von diesen Vektoren aufgespannte lineare Teilraum. Dann gilt* $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}^\perp = U^\perp$.

Beweis. Offensichtlich ist $U^\perp \subseteq \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}^\perp$; denn wenn $\mathbf{v}$ auf allen $\mathbf{a} \in U$ senkrecht steht, dann insbesondere auf den Erzeugern $\mathbf{a}_i$ von U.

Für die umgekehrte Inklusion betrachte ein $\mathbf{v} \in \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}^\perp$. Dann gilt $\langle \mathbf{a}_i \mid \mathbf{v}\rangle = 0$ für $i = 1, 2, \dots, m$. Sei $\mathbf{a}$ ein Element von U. Da der Teilraum U durch $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_m$ aufgespannt wird, kann $\mathbf{a}$ als Linearkombination

$$\mathbf{a} = \alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \cdots + \alpha_m \mathbf{a}_m$$

dargestellt werden. Also folgt

$$\begin{aligned}\langle \mathbf{a} \mid \mathbf{v}\rangle &= \langle \alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \cdots + \alpha_m \mathbf{a}_m \mid \mathbf{v}\rangle \\ &= \overline{\alpha_1}\langle \mathbf{a}_1 \mid \mathbf{v}\rangle + \overline{\alpha_2}\langle \mathbf{a}_2 \mid \mathbf{v}\rangle + \cdots + \overline{\alpha_m}\langle \mathbf{a}_m \mid \mathbf{v}\rangle = 0\end{aligned}$$

Damit ergibt sich $\langle \mathbf{a} \mid \mathbf{v}\rangle = 0$ für jedes $\mathbf{a} \in U$, also $\mathbf{v} \in U^\perp$. □

3.4.5 Bemerkung. Sei V ein euklidscher oder unitärer Vektorraum der Dimension n mit einer Orthonormalbasis $\mathcal{B} : \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Wir berechnen das Skalarprodukt zweier Vektoren $\mathbf{v}$ und $\mathbf{w}$, die durch die Koordinaten

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \quad , \quad \mathbf{w} = \sum_{j=1}^{n} \beta_j \mathbf{v}_j$$

bzgl. der Basis $\mathcal{B}$ dargestellt werden:

$$\begin{aligned}\langle \mathbf{v} \mid \mathbf{w}\rangle &= \langle \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \mid \sum_{j=1}^{n} \beta_j \mathbf{v}_j\rangle = \sum_{i=1}^{n} \overline{\alpha}_i \cdot \langle \mathbf{v}_i, \sum_{j=1}^{n} \beta_j \mathbf{v}_j\rangle \\ &= \sum_{i=1}^{n} \overline{\alpha}_i \cdot \sum_{j=1}^{n} \beta_j \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \sum_{i,j=1}^{n} \overline{\alpha}_i \beta_j \delta_{ij} = \sum_{i=1}^{n} \overline{\alpha}_i \beta_i \qquad .\end{aligned}$$

Wir stellen fest, dass wir denselben Ausdruck wie für das Standardskalarprodukt auf $\mathbb{R}^n$ und $\mathbb{C}^n$ erhalten. Dies kann in grundlegenderer Weise dargestellt werden:

Wir betrachten die Koordinatenabbildung κ von V auf $\mathbb{R}^n$ und entsprechend auf $\mathbb{C}^n$, die jedem Vektor $\mathbf{v} \in V$ seinen Koordinatenvektor bzgl. der Orthonormalbasis $\mathcal{B}$ zuordnet. Diese Abbildung erhält das Skalarprodukt, wie wir oben gesehen haben, d.h. für alle $\mathbf{v}, \mathbf{w} \in V$ gilt $\langle \kappa(\mathbf{v}), \kappa(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$. Abbildungen mit dieser Eigenschaft heißen *Isometrien*, da sie Längen und Winkel erhalten.

Wir folgern daraus, dass jeder euklidische (resp. unitäre) Vektorraum V der Dimension n isometrisch isomorph zu $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) mit dem Standardskalarprodukt ist, vorausgesetzt V hat eine Orthonormalbasis. Wir werden im Folgenden sehen, wie man eine Orthonormalbasis aus jeder beliebigen Basis konstruieren kann.

3.5. Orthogonalprojektionen

Im folgenden sei V ein beliebiger euklidischer oder unitärer Vektorraum. Wir benutzen unsere geometrische Vorstellung aus dem dreidimensionalen, um analoge Konstruktionen in V durchzuführen.

3.5.1 Zerlegung eines Vektors $\mathbf{x}$. Sei $\mathbf{v}$ ein vorgegebener von 0 verschiedener Vektor in V. Wir können das Skalarprodukt verwenden, um jeden beliebigen Vektor $\mathbf{x} \in V$ in seine Komponente $\mathbf{x}_p$ in Richtung von $\mathbf{v}$ und seine Komponente $\mathbf{x}_c$ senkrecht zu $\mathbf{v}$ zu zerlegen:

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_c \text{ mit } \mathbf{x}_p \parallel \mathbf{v} \text{ und } \mathbf{x}_c \perp \mathbf{v}$$

Dazu normieren wir zuerst $\mathbf{v}_0 := \frac{\mathbf{v}}{\|\mathbf{v}\|}$ und setzen

$$\begin{aligned} \mathbf{x}_p &:= \langle \mathbf{v}_0 \mid \mathbf{x} \rangle \cdot \mathbf{v}_0 = \tfrac{\langle \mathbf{v} | \mathbf{x} \rangle}{\langle \mathbf{v} | \mathbf{v} \rangle} \cdot \mathbf{v} \\ \mathbf{x}_c &:= \mathbf{x} - \mathbf{x}_p. \end{aligned}$$

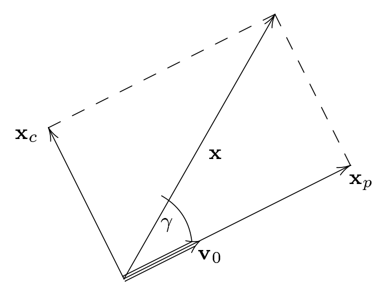


ABBILDUNG 3.2. Zerlegung eines Vektors $\mathbf{x}$

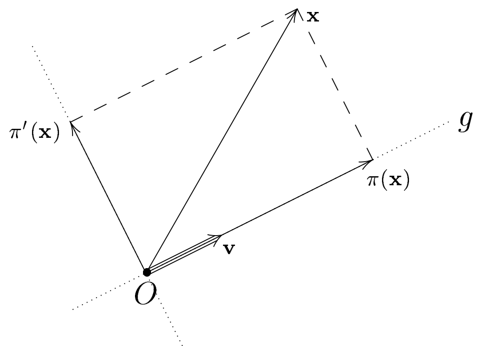
Die Idee zu dieser Vorgehensweise kommt von einem einfachen geometrischen Argument (siehe Abbildung 3.2), indem man die geometrische Interpretation des Skalarprodukts in der Ebene und im Raum verwendet. Dieses geometrische Argument wurde in Formeln übersetzt. Nun verifizieren wir formal, daß die Behauptung allgemein richtig ist, indem wir nur die Grundregeln des Skalarprodukts benutzen:

Per Definiton ist $\mathbf{x}_p$ parallel zu $\mathbf{v}$. Das heißt, wir müssen nur $\mathbf{x}_c \perp \mathbf{v}$ überprüfen. Dafür genügt es zu zeigen $\langle \mathbf{v} \mid \mathbf{x}_c \rangle = 0$:

$$\begin{aligned} \langle \mathbf{v} \mid \mathbf{x}_c \rangle &= \langle \mathbf{v} \mid \mathbf{x} - \mathbf{x}_p \rangle \\ &= \langle \mathbf{v} \mid \mathbf{x} \rangle - \langle \mathbf{v} \mid \mathbf{x}_p \rangle && \text{nach } (\mathrm{B}_2) \\ &= \langle \mathbf{v} \mid \mathbf{x} \rangle - \langle \mathbf{v} \mid \tfrac{\langle \mathbf{v} | \mathbf{x} \rangle}{\langle \mathbf{v} | \mathbf{v} \rangle} \cdot \mathbf{v} \rangle \\ &= \langle \mathbf{v} \mid \mathbf{x} \rangle - \tfrac{\langle \mathbf{v} | \mathbf{x} \rangle}{\langle \mathbf{v} | \mathbf{v} \rangle} \cdot \langle \mathbf{v} \mid \mathbf{v} \rangle && \text{nach } (\mathrm{B}_2) \\ &= \langle \mathbf{v} \mid \mathbf{x} \rangle - \langle \mathbf{v} \mid \mathbf{x} \rangle = 0. \end{aligned}$$

3.5.2 Orthogonale Projektion auf eine Gerade durch den Ursprung. Sei g eine Gerade durch den Ursprung in V mit dem Richtungsvektor $\mathbf{v} \neq 0$. Die Abbildung $\pi_g \colon V \to V$, die jedem Vektor $\mathbf{x} \in V$ seine Komponente $\pi_g(\mathbf{x}) = \mathbf{x}_p$ in Richtung von g zuordnet, wird die *Orthogonalprojektion* auf g genannt; denn $\mathbf{x}_c = \mathbf{x} - \pi_g(\mathbf{x})$ ist orthogonal zu g. Nach 3.5.1 wird $\pi_g(\mathbf{x})$ gegeben durch

$$\pi_g(\mathbf{x}) = \mathbf{x}_p = \frac{\langle \mathbf{v} \mid \mathbf{x} \rangle}{\langle \mathbf{v} \mid \mathbf{v} \rangle} \cdot \mathbf{v}.$$



Normieren wir $\mathbf{v}$, dann vereinfacht sich die Formel zu

$$\pi_g(\mathbf{x}) = \langle \mathbf{v}_0 \mid \mathbf{x} \rangle \cdot \mathbf{v}_0.$$

Nun betrachten wir gemäß 3.4.4 den Orthogonalraum $g^\perp = \mathbf{v}^\perp$ zu g, welcher ein linearer Teilraum von V ist. Für jeden Vektor $\mathbf{x} \in V$ ist die zu g senkrechte Komponente $\mathbf{x}_c$ ein Element des Orthogonalraums $g^\perp$. Wir können also eine Abbildung $\pi'_g \colon V \to V$ dadurch definieren, daß wir jedem $\mathbf{x} \in V$ seine Komponente $\pi'_g(\mathbf{x}) = \mathbf{x}_c$ senkrecht zu g zuordnen. Da $\mathbf{x}_c = \mathbf{x} - \mathbf{x}_p = \mathbf{x} - \pi_g(\mathbf{x})$, wird π'_g gegeben durch die Formel

$$\pi'_g(\mathbf{x}) = \mathbf{x} - \frac{\langle \mathbf{v} \mid \mathbf{x} \rangle}{\langle \mathbf{v} \mid \mathbf{v} \rangle} \cdot \mathbf{v} = \mathbf{x} - \langle \mathbf{v}_0 \mid \mathbf{x} \rangle \cdot \mathbf{v}_0.$$

Wir nennen π'_g die *Orthogonalprojektion* auf the Orthogonalraum $g^\perp$; denn $\mathbf{x} - \pi'_g(\mathbf{x}) = \pi_g(\mathbf{x})$ liegt in g is also orthogonal zu $g^\perp$.

3.5.3 Orthogonale Projektion auf einen linearen Teilraum U. Die Geraden durch 0 sind die eindimensionalen linearen Teilräume. Wir verallgemeinern nun das Vorgehen in 3.5.2 auf endlichdimensionale lineare Teilräume U eines euklidischen oder unitären Vektorraums V. Zu diesem Zweck betrachten wir eine Orthonormalbasis

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$$

von U. (Im Anschluß werden wir ein Verfahren kennenlernen, eine Orthonormalbasis zu bestimmen.) Wir definieren Abbildungen $\pi_U \colon V \to V$ und $\pi'_U \colon V \to V$, in dem wir für alle $\mathbf{x} \in V$ setzen

$$\pi_U(\mathbf{x}) = \langle \mathbf{v}_1 | \mathbf{x} \rangle \cdot \mathbf{v}_1 + \langle \mathbf{v}_2 | \mathbf{x} \rangle \cdot \mathbf{v}_2 + \cdots + \langle \mathbf{v}_m | \mathbf{x} \rangle \cdot \mathbf{v}_m = \sum_{i=1}^m \langle \mathbf{v}_i | \mathbf{x} \rangle \cdot \mathbf{v}_i$$

$$\pi'_U(\mathbf{x}) = \mathbf{x} - \pi_U(\mathbf{x})$$

Diese haben die folgenden Eigenschaften:

(0) Linearität: $\pi_U(\mathbf{x}+\mathbf{y}) = \pi_U(\mathbf{x}) + \pi_U(\mathbf{y}), \quad \pi_U(\lambda\mathbf{x}) = \lambda\pi_U(\mathbf{x})$
für beliebige $\mathbf{x}, \mathbf{y} \in V$ und beliebige Skalare λ. Dasselbe gilt für π'_U.

Unter Verwendung der Definition von π_U und des Axioms (B_2) sehen wir $\pi_U(\mathbf{x}+\mathbf{y}) = \sum_{i=1}^m \langle \mathbf{v}_i | \mathbf{x}+\mathbf{y} \rangle \mathbf{v}_i = \sum_{i=1}^m \big(\langle \mathbf{v}_i | \mathbf{x} \rangle \mathbf{v}_i + \langle \mathbf{v}_i | \mathbf{y} \rangle \mathbf{v}_i \big) = \sum_{i=1}^m \langle \mathbf{v}_i | \mathbf{x} \rangle \mathbf{v}_i + \sum_{i=1}^m \langle \mathbf{v}_i | \mathbf{y} \rangle \mathbf{v}_i = \pi_U(\mathbf{x}) + \pi_U(\mathbf{y})$ und $\pi_U(\lambda\mathbf{x}) = \sum_{i=1}^m \langle \mathbf{v}_i | \lambda\mathbf{x} \rangle \mathbf{v}_i = \sum_{i=1}^m \lambda \langle \mathbf{v}_i | \mathbf{x} \rangle \mathbf{v}_i = \lambda \sum_{i=1}^m \langle \mathbf{v}_i | \mathbf{x} \rangle \mathbf{v}_i = \pi_U(\mathbf{x})$.

(1) $\pi_U(\mathbf{x}) \in U$ für alle $\mathbf{x} \in V$.

In der Tat ist $\pi_U(\mathbf{x})$ eine Linearkombination der Basisvektoren $\mathbf{v}_1, \dots, \mathbf{v}_m$ von U.

(2) $\pi_U(\mathbf{u}) = \mathbf{u}$ für alle $\mathbf{u} \in U$.

Ist nämlich $\mathbf{u} = \sum_{j=1}^{m} \alpha_j \mathbf{v}_j$, so folgt

$$\pi_U(\mathbf{u}) = \sum_{i=1}^{m} \langle \mathbf{v}_i \mid \sum_{j=1}^{m} \alpha_j \mathbf{v}_j \rangle \mathbf{v}_i = \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_j \underbrace{\langle \mathbf{v}_i \mid \mathbf{v}_j \rangle}_{\delta_{ij}} \mathbf{v}_i = \sum_{i=1}^{m} \alpha_i \mathbf{v}_i = \mathbf{u}$$

(3) $\pi_U(\pi_U(\mathbf{x})) = \pi_U(\mathbf{x})$ für alle $\mathbf{x} \in V$, d.h. $\pi_U \circ \pi_U = \pi_U$.

In der Tat ist $\pi_U(\mathbf{x}) \in U$ nach (1) und damit $\pi_U(\pi_U(\mathbf{x})) = \pi_U(\mathbf{x})$ nach (2).

(4) Das Bild der Abbildung π_U ist U und $\{\mathbf{x} \in V : \pi_U(\mathbf{x}) = 0\} = U^\perp$.

Die Behauptung für das Bild folgt aus (1) und (2). Ferner gilt

$$\begin{aligned} \pi_U(\mathbf{x}) = 0 &\iff \langle \mathbf{v}_i \mid \mathbf{x} \rangle = 0 \qquad \text{für alle } i = 1, \dots, m \\ &\iff \mathbf{x} \in \{\mathbf{v}_1, \dots, \mathbf{v}_m\}^\perp \underset{\text{(Lemma)}}{=} U^\perp \end{aligned}$$

(5) $\pi'_U(\mathbf{x}) \in U^\perp$ für alle $\mathbf{x} \in V$; insbesondere gilt $\pi_U(\mathbf{x}) \perp \pi'_U(\mathbf{x})$.

Denn

$$\pi_U(\pi'_U(\mathbf{x})) = \pi_U(\mathbf{x} - \pi_U(\mathbf{x}))) \underset{0}{=} \pi_U(\mathbf{x}) - \pi_U(\pi_U(\mathbf{x})) \underset{(2)}{=} \pi_U(\mathbf{x}) - \pi_U(\mathbf{x}) = 0$$

und damit $\pi'_U(\mathbf{x}) \underset{(4)}{\in} U^\perp$.

(6) $\|\mathbf{x} - \pi_U(x)\| \leq \|\mathbf{x} - \mathbf{u}\|$ für alle $\mathbf{u} \in U$. Das bedeutet, daß $\pi_U(\mathbf{x})$ der Punkt von U mit kleinstem Abstand von $\mathbf{x}$ ist. In anderen Worten ist „$\pi_U(\mathbf{x})$ die beste Approximation von $\mathbf{x}$ durch ein Element aus U“. Beweis: Übung.

Die oben definierte Abbildung π_U wird *Orthogonalprojektion von V auf U* und $\pi'_U : V \to V$ ist die Orthogonalprojektion auf $U^\perp$ ist. Wir haben $\mathbf{x} = \pi_U(\mathbf{x}) + \pi'_U(\mathbf{x})$ und k2onnen dies auffassen als Zerlegung von x in eine Komponente aus U und eine Komponente senkrecht zu U.

3.5.4 Proposition. *Für jeden endlichdimensionalen linearen Teilraum U von V gilt*

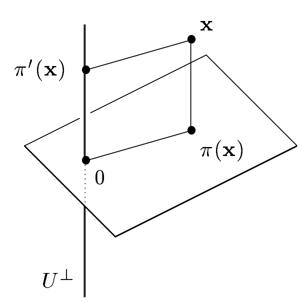
$$V = U \oplus U^\perp$$

Beweis.

(i) $U \cap U^\perp = \{0\}$, denn aus $\mathbf{x} \in U \cap U^\perp$ folgt $\langle \mathbf{x} | \mathbf{x} \rangle = 0$ und folglich $\mathbf{x} = 0$.

(ii) $U + U^\perp = V$, denn für jedes $\mathbf{x} \in V$ gilt $\mathbf{x} = \pi_U(\mathbf{x}) + \pi'_U(\mathbf{x})) \in U + U^\perp$.

□



3.5.5 Folgerung.

$$\dim V = \dim U + \dim U^{\perp}$$

3.5.6 Das Gram-Schmidtsche Orthogonalisierungsverfahren. Wir betrachten ein System $\mathbf{b}_1$, $\mathbf{b}_2$, ..., $\mathbf{b}_m$ von m linear unabhängigen Vektoren in einem euklidischen oder unitären Vektorraum V. Diese spannen einen m-dimensionalen linearen Teilraum U von V auf. Der folgende Algorithmus liefert eine Orthonormalbasis $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ von U. Beginnen wir insbesondere mit einer Basis von V, so erhalten wir eine Orthonormalbasis von V:

$$\begin{aligned}
\mathbf{u}_1 &:= \mathbf{b}_1 & \mathbf{v}_1 &:= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}\\
\mathbf{u}_2 &:= \mathbf{b}_2 - \langle \mathbf{v}_1|\mathbf{b}_2\rangle \mathbf{v}_1 & \mathbf{v}_2 &:= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}\\
\mathbf{u}_3 &:= \mathbf{b}_3 - \langle \mathbf{v}_1|\mathbf{b}_3\rangle \mathbf{v}_1 - \langle \mathbf{v}_2|\mathbf{b}_3\rangle \mathbf{v}_2 & \mathbf{v}_3 &:= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}\\
&\vdots & &\vdots\\
\mathbf{u}_m &:= \mathbf{b}_m - \sum_{i=1}^{m-1} \langle \mathbf{v}_i|\mathbf{b}_m\rangle \mathbf{v}_i & \mathbf{v}_m &:= \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|}
\end{aligned}$$

Diese Verfahren kann man so beschreiben:
Im ersten Schritt normiert man $\mathbf{b}_1$ und erhält $\mathbf{v}_1$.
Im zweiten Schritt bestimmt man die Komponente $\mathbf{u}_2$ von $\mathbf{b}_2$ orthogonal zu $\mathbf{v}_1$; durch Normierung von $\mathbf{u}_2$ erhält man $\mathbf{v}_2$.
Im dritten Schritt bestimmt man die Komponente $\mathbf{u}_3$ von $\mathbf{b}_2$ senkrecht zu $\mathbf{v}_1$ und $\mathbf{v}_2$ gemäß 3.5.3; durch Normierung von $\mathbf{u}_3$ erhält man $\mathbf{v}_3$.
usw.
Das so gewonnene Orthonormalsystem $\mathbf{v} - 1, \ldots, \mathbf{v}_m$ besteht aus Linearkombinationen der Vektoren $\mathbf{b}_1, \ldots, \mathbf{b}_m$, liegt also in dem von $\mathbf{b}_1, \ldots, \mathbf{b}_m$ aufgespannten linearen Teilraum U. Da ein Orthonormalsystem linear unabhängig ist und dieses aus m Vektoren besteht, ist es eine Basis von U.

Für praktische Zwecke schreibt man das Schema oftmals in der folgenden Weise:

$$\begin{aligned}
\mathbf{u}_1 &:= \mathbf{b}_1 & \mathbf{v}_1 &:= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}\\
\mathbf{u}_2 &:= \mathbf{b}_2 - \frac{\langle \mathbf{u}_1|\mathbf{b}_2\rangle}{\langle \mathbf{u}_1|\mathbf{u}_1\rangle} \mathbf{u}_1 & \mathbf{v}_2 &:= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}\\
\mathbf{u}_3 &:= \mathbf{b}_3 - \frac{\langle \mathbf{u}_1|\mathbf{b}_3\rangle}{\langle \mathbf{u}_1|\mathbf{u}_1\rangle} \mathbf{u}_1 - \frac{\langle \mathbf{u}_2|\mathbf{b}_3\rangle}{\langle \mathbf{u}_2|\mathbf{u}_2\rangle} \mathbf{u}_2 & \mathbf{v}_3 &:= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}\\
&\vdots & &\vdots\\
\mathbf{u}_m &:= \mathbf{b}_m - \sum_{i=1}^{m-1} \frac{\langle \mathbf{u}_i|\mathbf{b}_m\rangle}{\langle \mathbf{u}_i|\mathbf{u}_i\rangle} \mathbf{u}_i & \mathbf{v}_m &:= \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|}
\end{aligned}$$

Da jeder endlichdimensionale Vektorraum eine Basis hat, können wir das Gram-Schmidtsche Orthogonalisierungsverfahren auf diese anwenden und erhalten die folgende

3.5.7 Proposition. *Jeder endlichdimensionalen euklidische oder unitäre Vektorraum und insbesondere jeder lineare Teilraum von $\mathbb{R}^n$ und von $\mathbb{C}^n$ hat eine Orthonormalbasis.*

3.5.8 Beispiel. Sei U der lineare Teilraum von $\mathbb{R}^4$, der durch die homogene lineare Gleichung

$$x_1 + x_2 + x_3 + x_4 = 0$$

gegeben ist. Man zeigt leicht, dass die folgenden Vektoren Lösungen dieser Gleichung sind und dass sie linear unabhängig sind. Sie bilden also eine Basis von U

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad , \quad \mathbf{b}_2 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ -3 \end{pmatrix} \quad , \quad \mathbf{b}_3 = \begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

Wir benutzen nun das GRAM-SCHMIDTsche Orthogonalisierungsverfahren, um eine orthonormale Basis von U zu finden:

$$\mathbf{u}_1 = \mathbf{b}_1 \qquad \|\mathbf{u}_1\| = 2 \qquad \mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$\mathbf{u}_2 = \mathbf{b}_2 - \langle \mathbf{v}_1, \mathbf{b}_2 \rangle \mathbf{v}_1 = \mathbf{b}_2 - 2\mathbf{v}_1 = \begin{pmatrix} 2 \\ 2 \\ -2 \\ -2 \end{pmatrix} \qquad \|\mathbf{u}_2\| = 4 \qquad \mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{u}_3 &= \mathbf{b}_3 - \langle \mathbf{v}_1, \mathbf{b}_3 \rangle \mathbf{v}_1 - \langle \mathbf{v}_2, \mathbf{b}_3 \rangle \mathbf{v}_2 \\ &= \mathbf{b}_3 - 2\mathbf{v}_1 - 2\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \qquad \|\mathbf{u}_3\| = 2 \qquad \mathbf{v}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \end{aligned}$$

3.5.9 Die Q-R-Zerlegung einer Matrix B. Sei B eine $n \times m$-Matrix mit reell- oder komplexwertigen Einträgen, deren Spalten

$$\mathbf{b}_1,\ \mathbf{b}_2,\ \ldots,\mathbf{b}_m$$

als linear unabhängig vorausgesetzt werden. Dann hat B eine Zerlegung

$$B = QR$$

wobei Q eine $n \times m$-Matrix ist, deren Spalten $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_m$ ein Orthonormalsystem bilden, und R ist eine obere $m \times m$-Dreiecksmatrix.

Zur Berechnung dieser Zerlegung wende das GRAM-SCHMIDTsche Orthogonalisierungsverfahren auf $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{b}_m$ an. Man erhält ein Orthonormalsystem $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_m$.

Diese Vektoren bilden die Spalten von Q. Für R wähle

$$R = \begin{pmatrix} \|\mathbf{u}_1\| & \langle \mathbf{v}_1|\mathbf{b}_2\rangle & \dots & \langle \mathbf{v}_1|\mathbf{b}_m\rangle \\ 0 & \|\mathbf{u}_2\| & \dots & \langle \mathbf{v}_2|\mathbf{b}_m\rangle \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & & \vdots \\ 0 & \dots & 0 & \|\mathbf{u}_m\| \end{pmatrix}$$

Mit Gram-Schmidt erhalten wir:

$$\begin{aligned} \mathbf{b}_1 &= \|\mathbf{u}_1\| \cdot \mathbf{v}_1 \\ \mathbf{b}_2 &= \langle \mathbf{v}_1|\mathbf{b}_2\rangle \mathbf{v}_1 + \|\mathbf{u}_2\|\mathbf{v}_2 \\ &\vdots \\ \mathbf{b}_m &= \langle \mathbf{v}_1|\mathbf{b}_m\rangle \mathbf{v}_1 + \langle \mathbf{v}_2|\mathbf{b}_m\rangle \mathbf{v}_2 + \dots + \|\mathbf{u}_m\|\mathbf{v}_m \end{aligned}$$

Beachte, dass aus diesen Gleichungen

$$\|\mathbf{u}_i\| = \langle \mathbf{v}_i|\mathbf{b}_i\rangle$$

folgt und daß obige Gleichungen auch in der Form $B = Q \cdot R$ geschrieben werden können.

Beispiel. Wir betrachten die Matrix B, wobei die Spalten b_1, b_2, b_3 die Vektoren aus Beispiel 3.5.8 sind:

$$B = \begin{pmatrix} 1 & 3 & 3 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -3 & -1 \end{pmatrix}$$

Wir berechnen nun

$\langle \mathbf{v}_i \| \mathbf{b}_j\rangle$	$\mathbf{b}_1$	$\mathbf{b}_2$	$\mathbf{b}_3$
$\mathbf{v}_1$	2	2	2
$\mathbf{v}_2$	0	4	2
$\mathbf{v}_3$	0	0	2

Aus den Ergebnissen des Gram-Schmidtschen Algorithmus in 3.5.8 erhalten wir

$$\underbrace{\begin{pmatrix} 1 & 3 & 3 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -3 & -1 \end{pmatrix}}_{B} = \underbrace{\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix}}_{Q} \cdot \underbrace{\begin{pmatrix} 2 & 2 & 2 \\ 0 & 4 & 2 \\ 0 & 0 & 2 \end{pmatrix}}_{R}$$

3.6. Geraden und Ebenenen im $\mathbb{R}^2$ und $\mathbb{R}^3$

3.6.1 Hyperebenen durch den Ursprung im $\mathbb{R}^n$. Wie in 3.5.2 sei $\mathbf{n}$ ein Vektor im $\mathbb{R}^n$ ungleich dem Nullvektor. Sei H die Menge aller $\mathbf{x} \in \mathbb{R}^n$, die orthogonal zu $\mathbf{n}$ sind, d.h. H ist die Menge aller $\mathbf{x} \in \mathbb{R}^n$, die die Gleichung

$$(*) \qquad \langle \mathbf{n} | \mathbf{x} \rangle = 0$$

erfüllen. Dann ist H ein linearer Teilraum der Dimension $n-1$. Deshalb nennt man H eine *Hyperebene* [*hyperplane*]. Der Vektor $\mathbf{n}$ wird ein *Normalenvektor* [*normal vector*] der Hyperebene H genannt und der normierte Vektor $\mathbf{n}_0 = \frac{\mathbf{n}}{\|\mathbf{n}\|}$ ein *Einheitsnormalenvektor* [*unit normal vector*]. Ist

$$\mathbf{n} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

so kann Gleichung $(*)$ geschrieben werden als

$$H \colon \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0.$$

Das heißt, H ist die Menge aller Vektoren $\mathbf{x}$, welche Lösungen dieser homogenen linearen Gleichung sind.

Spezialfall $\mathbf{n = 2}$:

$$H \colon \alpha_1 x_1 + \alpha_2 x_2 = 0.$$

Das ist die Gleichung einer Geraden im $\mathbb{R}^2$ durch den Ursprung. Das heißt, H ist die Gerade durch den Ursprung, die senkrecht zu dem Vektor $\mathbf{n} = \binom{\alpha_1}{\alpha_2}$ ist. Der Vektor $\mathbf{n}$ heißt Normalenvektor der Geraden. Beachte, dass die Koordinaten dieses Normalenvektors nichts anderes als die Koeffizienten in der Gleichung der Geraden H sind.

Spezialfall $\mathbf{n = 3}$:

$$H \colon \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0.$$

Dies ist die Gleichung einer Ebene im $\mathbb{R}^3$, die durch den Ursprung geht. Das heißt, H ist die Ebene, die durch den Ursprung geht und senkrecht zum Vektor $\mathbf{n} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ ist. Der Vektor $\mathbf{n}$ heißt Normalenvektor der Ebene H. Beachte, dass die Koordinaten dieses Normalenvektors nichts anderes als die Koeffizienten der Gleichung der Ebene H sind.

Wir behalten die eben festgelegten Bezeichnungen in der Folge bei.

3.6.2 Orthogonalprojektion auf eine Huperebene durch den Ursprung. Nun sei $\pi \colon \mathbb{R}^n \to \mathbb{R}^n$ die orthogonale Projektion auf die Normale zu der Hyperebene H. Die Normale hat den Richtungsvektor $\mathbf{n}$ und den Normaleneinheitsvektor $\mathbf{n}_0 = \frac{\mathbf{n}}{\|\mathbf{n}\|}$. Wie in 3.5.2 wird π gegeben durch:

$$\pi(\mathbf{x}) = \tfrac{\langle \mathbf{n} | \mathbf{x} \rangle}{\langle \mathbf{n} | \mathbf{n} \rangle} \cdot \mathbf{n} = \langle \mathbf{n}_0 | \mathbf{x} \rangle \mathbf{n}_0 \quad \text{für alle } \mathbf{x} \in \mathbb{R}^n.$$

Die orthogonale Projektion π'auf die Hyperebene H wird gegeben durch

$$\pi'(\mathbf{x}) = x - \tfrac{\langle \mathbf{n} | \mathbf{x} \rangle}{\langle \mathbf{n} | \mathbf{n} \rangle} \cdot \mathbf{n} = x - \langle \mathbf{n}_0 | \mathbf{x} \rangle \mathbf{n}_0 \quad \text{für alle } \mathbf{x} \in \mathbb{R}^n.$$

Beispiel. Betrachten wir im Euklidischen Raum $\mathbb{R}^3$ die Ebene, die durch folgende Gleichung gegeben ist:

$$E\colon x_1 - x_2 + 2x_3 = 0.$$

Ein Normalenvektor für die Ebene E ist durch die Koeffizienten der die Ebene definierenden Gleichung gegeben:

$$\mathbf{n} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

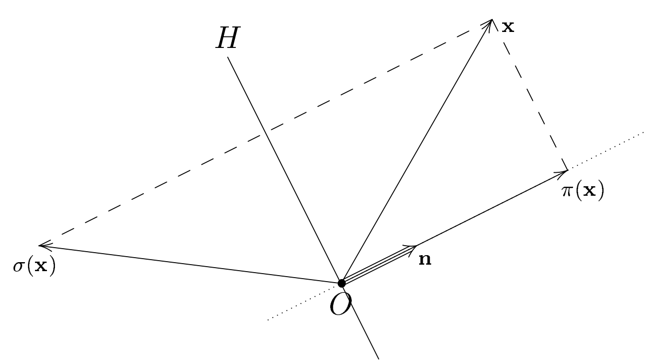
Wir berechnen $\langle \mathbf{n}|\mathbf{n}\rangle = 1^2 + (-1)^2 + 2^2 = 6$. Für die Orthogonalprojektion auf die Ebene E gilt nach obigem

$$\pi'(\mathbf{x}) = \mathbf{x} - \tfrac{1}{6} \cdot \langle \left(\begin{smallmatrix} 1 \\ -1 \\ 2 \end{smallmatrix}\right) |\mathbf{x}\rangle \cdot \left(\begin{smallmatrix} 1 \\ -1 \\ 2 \end{smallmatrix}\right).$$

3.6.3 Spiegelung an einer Hyperebene durch den Ursprung. Sei H eine Hyperebene durch den Ursprung im $\mathbb{R}^n$ mit Normalenvektor $\mathbf{n} \neq \mathbf{0}$. Wir betrachten die Spiegelung σ des $\mathbb{R}^n$ an der Hyperebene H.

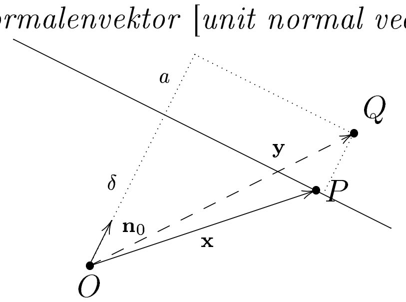
Für jeden Vektor $\mathbf{x}$ ist das Spiegelbild gegeben durch

$$\sigma(\mathbf{x}) = \mathbf{x} - 2\pi(\mathbf{x}) = \mathbf{x} - 2\tfrac{\langle \mathbf{n}|\mathbf{x}\rangle}{\langle \mathbf{n}|\mathbf{n}\rangle}\mathbf{n} = \mathbf{x} - 2\langle \mathbf{n}_0|\mathbf{x}\rangle \mathbf{n}_0.$$



Die obige Formel ist unabhängig von der Dimension n. Sie kann für die Spiegelung des $\mathbb{R}^2$ an einer Geraden durch den Ursprung, für die Spiegelung des $\mathbb{R}^3$ an einer Ebene durch den Ursprung, usw. verwendet werden. Dies zeigt die Nützlichkeit der Vektorrechnung.

3.6.4 Beliebige Geraden im Euklidischen $\mathbb{R}^2$. Sei g eine beliebige Gerade im Euklidischen Raum $\mathbb{R}^2$. Wir bezeichnen mit δ ihren Abstand vom Ursprung und mit $\mathbf{n}_0$ den Einheitsvektor, der senkrecht auf g steht und vom Ursprung aus Richtung g zeigt, dem sogenannten *Einheitsnormalenvektor* [*unit normal vector*] von g.



Ein Punkt P liegt auf der Geraden g genau dann, wenn der Vektor $\mathbf{x} = \overrightarrow{OP}$ folgende Gleichung erfüllt:

$$\langle \mathbf{x} | \mathbf{n}_0 \rangle = \delta.$$

Falls $\mathbf{n}_0 = \binom{\alpha_1}{\alpha_2}$ und $\mathbf{x} = \binom{x_1}{x_2}$, dann hat die Gleichung die Form

$$\alpha_1 x_1 + \alpha_2 x_2 = \delta.$$

Diese Gleichung nennt man die Hessesche *Normalform* der Geraden g.

Sei a der Abstand eines Punktes Q von der Geraden g. Für den Vektor $\mathbf{y} = \overrightarrow{OQ}$ gilt

$$\langle \mathbf{n}_0, \mathbf{y} \rangle = \delta \pm a,$$

wobei das Vorzeichen von a entweder $+$ oder $-$ ist in Abhängigkeit davon, ob Q auf der gegenüberliegenden oder derselben Seite der Geraden wie der Ursprung liegt. Damit ist

$$a = |\langle \mathbf{n}_0, \mathbf{y} \rangle - \delta|.$$

Nun wollen wir mit einer Gleichung der Form

$$\beta_1 x_1 + \beta_2 x_2 = \gamma$$

beginnen, wobei mindestens einer der Koeffizienten β_1, β_2 ungleich Null sei. Wir nehmen an, dass $\gamma \geq 0$ (andernfalls multiplizieren wir die Gleichung mit -1). Sei weiter $\mathbf{n} = \binom{\beta_1}{\beta_2}$ und $\mathbf{n}_0 = \frac{\mathbf{n}}{\|\mathbf{n}\|}$. Multiplizieren wir obige Gleichung mit $\frac{1}{\|\mathbf{n}\|}$, so erhalten wir

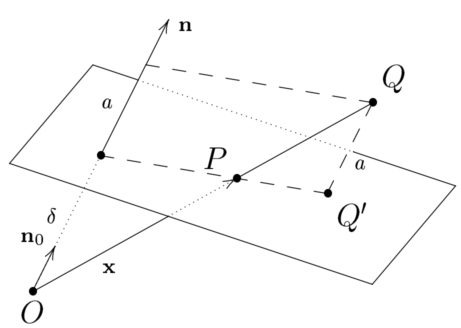
$$\tfrac{\beta_1}{\|\mathbf{n}\|} x_1 + \tfrac{\beta_2}{\|\mathbf{n}\|} x_2 = \tfrac{\gamma}{\|\mathbf{n}\|},$$

welche auch als

$$\langle \mathbf{n}_0 | \mathbf{x} \rangle = \frac{\gamma}{\|\mathbf{n}\|}.$$

geschrieben werden kann. Dies ist die Hessesche Normalform einer Geraden g. Der Vektor $\mathbf{n}_0$ ist der Einheitsnormalenvektor für g, $\mathbf{n}$ ist ein Normalenvektor und $\delta = \frac{\gamma}{\|\mathbf{n}\|}$ ist der Abstand von g zum Ursprung.

3.6.5 Beliebige Ebenen im Euklidischen Raum $\mathbb{R}^3$. Sei E eine beliebige Ebene im Euklidischen Raum $\mathbb{R}^3$. Wir bezeichnen mit δ ihren Abstand zum Ursprung und mit $\mathbf{n}_0$ den Einheitsvektor, der senkrecht auf der Ebene E steht und vom Ursprung O in Richtung E zeigt, den sogenannten *Einheitsnormalenvektor* [*unit normal vector*] der Ebene.



Ein Punkt P liegt in E genau dann, wenn der Vektor $\mathbf{x} = \overrightarrow{OP}$ die folgende Gleichung erfüllt:

$$\langle \mathbf{x} | \mathbf{n}_0 \rangle = \delta.$$

Mit

$$\mathbf{n}_0 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \text{ und } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

liest sich die Gleichung wie folgt:

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \delta.$$

Diese Gleichung heißt HESSESCHE *Normalform* der Ebene E.

Sei a der Abstand eines Punktes Q von der Ebene E. Für den Vektor $\mathbf{y} = \overrightarrow{OQ}$ gilt

$$\langle \mathbf{n}_0 | \mathbf{y} \rangle = \delta \pm a,$$

wobei das Vorzeichen von a entweder $+$ oder $-$ ist in Abhängigkeit davon, ob Q auf der anderen oder derselben Seite der Ebene wie der Ursprung liegt. Damit

$$a = |\langle \mathbf{n}_0 | \mathbf{y} \rangle - \delta|.$$

Umgekehrt, betrachte eine Gleichung der Form

$$\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = \gamma,$$

wobei mindestens einer der Koeffizienten $\beta_1, \beta_2, \beta_3$ von Null verschieden sei. Wir können annehmen, dass $\gamma \geq 0$ (andernfalls multiplizieren wir die Gleichung mit -1). Sei weiterhin

$$\mathbf{n} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \text{ und } \mathbf{n}_0 = \frac{\mathbf{n}}{\|\mathbf{n}\|}.$$

Multiplizieren wir obige Gleichung mit $\frac{1}{\|\mathbf{n}\|}$, so erhalten wir

$$\tfrac{\beta_1}{\|\mathbf{n}\|} x_1 + \tfrac{\beta_2}{\|\mathbf{n}\|} x_2 + \tfrac{\beta_3}{\|\mathbf{n}\|} x_3 = \tfrac{\gamma}{\|\mathbf{n}\|},$$

das zu

$$\langle \mathbf{n}_0 | \mathbf{x} \rangle = \frac{\gamma}{\|\mathbf{n}\|}.$$

umformuliert werden kann. Dies ist die HESSESCHE Normalform einer Ebene E. Der Vektor $\mathbf{n}_0$ ist der Einheitsnormalenvektor von E, $\mathbf{n}$ ein Normalenvektor, und $\delta = \frac{\gamma}{\|\mathbf{n}\|}$ ist der Abstand von E zum Ursprung.

Man beachte wieder, daß wir dieselben Formeln für den zwei- und den dreidimensionalen Fall haben.

Übung. Geben Sie eine Formel für die orthogonale Projektion des $\mathbb{R}^3$ auf eine Ebene E an, die nicht durch den Ursprung geht.

3.7. Ein Hauch von Quantenmechanik

3.7.1 Das Stern-Gerlach Experiment. Das *Stern-Gerlach Experiment* funktioniert wie folgt: ein horizontaler Strahl von Partikeln (Silberatomen) wird durch ein Paar speziell präparierter Magneten gesendet. Das führt zu einer Aufteilung des Strahls in zwei Strahlen, einer wird nach oben abgelenkt (um etwa 30°) und der andere wird nach unten abgelenkt (ebenfalls um etwa $-30°$). Der Physiker sagt uns, die nach oben gehenden Partikel haben spin $1/2$ und die nach unten gehenden haben spin $-1/2$. Das magnetische Feld kann also die Partikel mit spin $1/2$ von den Partikeln mit spin $-1/2$ trennen.

Sendet man den Strahl mit spin 1/2 durch eine weitere Anordnung gleich orientierter Magneten, dann findet keine weitere Trennung mehr statt: der eintretende Strahl wird nach oben abgelenkt (um etwa 30°). Dasselbe gilt für den Strahl mit spin $-1/2$.

Senden wir nun den Strahl mit spin 1/2 durch ein Paar von Magneten, die um 90° gedreht worden sind, so findet wiederum eine Aufspaltung in zwei Strahlen statt: einer wird nach links (um etwa 30°), der andere nach rechts (um etwa $-30°$) abgelenkt. Folgendes ist überraschend: wählen wir den der vier Strahlen, der durch die ersten Magneten nach oben abgelenkt wurde und durch das zweite Magnetpaar geschickt wurde, oder den anderen Strahl und senden ihn durch ein weiteres Magnetpaar, das wieder wie die ersten Magneten orientiert ist, so erhalten wir wieder eine Aufspaltung in zwei Strahlen, der eine aufwärts abgelenkt, der andere nach unten. Es scheint so, als ob die dazwischenliegende Ablenkung, oder Polarisierung wie es die Physiker bezeichnen, den spin nochmals verschoben hätte.

Wir zeigen nun, wie dieses ungewöhnliche Verhalten durch den unitären Vektorraum $\mathbb{C}^2$ modelliert werden kann: Der *Zustand* eines Partikels ist ein eindimensionaler Teilraum von $\mathbb{C}^2$, der durch einen Einheitsvektor $s \in \mathbb{C}^2$ dargestellt werden kann.

Die Ablenkung durch die anfängliche Menge von Magneten kann als sog. *Polarisierung* bzgl. der Zerlegung $\mathbb{C}^2 = U_1 \oplus U_2$ mit $U_1 = \operatorname{span}(\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right))$ und $U_2 = \operatorname{span}(\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))$ verstanden werden. Das heißt, dass nach der Polarisierung der Zustand eines Partikels entweder in U_1 oder in U_2 sein wird, was mit Wahrscheinlichkeit $||\pi_i(s)||^2$ geschieht, wobei $\pi_i : \mathbb{C}^2 \to U_i$ die Projektion auf U_i ist. Nach der Polarisierung ist der Zustand des Partikels gegeben durch $s' := \pi_i(s)/||\pi_i(s)||$ wobei i der Index der Teilräume ist, zu denen die Polarisierung geführt hat. Man erinnere sich, dass $i = 1$ oder $i = 2$ im allgemeinen beide möglich sind; ein Partikel i tritt mit der Wahrscheinlichkeit $||\pi_i(s)||^2$ auf. Der neue Zustand ist also entweder in U_1 oder in U_2, so dass eine weitere Polarisierung nichts Neues mehr bringt. Ist nämlich $s' \in U_1$, dann folgt $||\pi_1(s')||^2 = 1$ und $||\pi_2(s')||^2 = 0$, so dass wir mit Wahrscheinlichkeit 1 wieder in U_1 landen werden. Insbesondere ist der neue Zustand nach der zweiten Polarisierung wieder s' bzgl. $\pi_1 \circ \pi_1 = \pi_1$.

Wir bemerken, dass – auch wenn der Zustand völlig gegeben ist – das Ergebnis der Polarisierung nicht exakt bestimmt ist; nur dessen Wahrscheinlichkeitsverteilung kann aus dem Ausgangszustand bestimmt werden. Wir bemerken auch, dass das Polarisieren den Zustand durch Projektion auf einen Teilraum verändert.

Wie kann die Ablenkung durch die um 90° gedrehten Magneten erklärt werden? Diese werden als eine Polarisierung bzgl. der Zerlegung $\mathbb{C}^2 = V_1 \oplus V_2$ angesehen, wobei $V_1 = \operatorname{span}(\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right))$ und $V_2 = \operatorname{span}(\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right))$ gilt.

Betrachten wir, was mit $s \in U_1$ geschieht, d.h. $s = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ bzgl. dieser Polarisierung. Dies entspricht einer Sendung des nach oben abgelenkten Strahls durch die gedrehte Anordnung von Magneten, die wie betrachtet zu einer Aufsplittung in zwei Strahlen führt. Die Projektion von s auf V_1 ist gleich $\frac{1}{2}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$, so dass die Wahrscheinlichkeit hierfür $(1/\sqrt{2})^2 = 1/2$ beträgt. Analogerweise ist die Wahrscheinlichkeit, dass s auf V_2 projiziert wird gleich $1/2$, was die Aufsplittung in zwei Strahlen erklärt. Nach dieser Polarisierung ist der Zustand entweder $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right))$ oder $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right))$. Quantenmechanische Experimente können durch unitäre Vektorräume modelliert werden; Zustände werden mit eindimensionalen linearen Teilräumen identifiziert, die durch Einheitsvektoren dargestellt werden (zwei Einheitsvektoren werden als gleich betrachtet, wenn einer aus dem anderen durch Multiplikation mit einer Zahl λ vom Betrag 1 gewonnen werden kann.)

Ein Experiment mit m möglichen Ergebnissen wird durch eine Zerlegung $V = U_1 \oplus \cdots \oplus U_m$ beschrieben, wobei die U_i paarweise orthogonal sind. Die Wahrscheinlichkeit von Endzustand i ist das Quadrat der Länge der Projektion des Zustandes s auf U_i. Es folgt: der Zustand *ist* die normierte Projektion auf den gewählten U_i. Es gibt keine Möglichkeit, Information über s zu erhalten ohne es zu modifizieren, es sei denn, s wäre bereits in einem der U_i gewesen, und in diesem Fall ist das Ergebnis vorhersagbar (U_i erscheint mit Wahrscheinlichkeit 1) und s bleibt gleich.

Experimente mit unendlich vielen möglichen Ergebnissen, z.B. die Messung von Position oder Impulse eines Partikels können nicht mit endlich dimensionalen Vektorräumen beschrieben werden. Eine besondere Art von unendlichdimensionalen Vektorräumen wird benötigt: ein *Hilbertraum*. Solange wir noch nicht den notwendigen Hintergrund besitzen, um Hilberträume zur Modellierung allgemeiner quantenmechanischer Experimente zu benutzen, führen wir diese als Quelle einiger schöner Beispiele ein.

3.7.2 Der Hilbertraum ℓ_2. Sei ℓ_2 die Menge aller Folgen

$$\mathbf{v} = (\alpha_i)_{i=0,1,2,\dots} = (\alpha_0, \alpha_1, \alpha_2, \dots)$$

von komplexen Zahlen mit

$$\sum_{i=0}^{\infty} |\alpha_i|^2 < +\infty$$

Wir definieren für zwei Folgen $\mathbf{v} = (\alpha_i)_i$ und $\mathbf{w} = (\beta_i)_i$ in ℓ_2

$$\mathbf{v} + \mathbf{w} = (\alpha_i + \beta_i)_i \qquad \text{und} \qquad \lambda\mathbf{v} = (\lambda\alpha_i)_i \qquad \text{für } \lambda \in \mathbb{C}.$$

Wenn Ihnen unendliche Summen komplexer Zahlen noch nicht vertraut sind, dann beschränken sie sich einfach auf Folgen reeller Zahlen.

Übung. Zeige $\sum_{i=0}^{\infty} |\alpha_i + \beta_i|^2 < +\infty$, $\sum_{i=0}^{\infty} |\lambda\alpha_i|^2 < \infty$ und dass $\sum_{i=0}^{\infty} \alpha_i\beta_i$ absolut konvergent ist.

Sind also $\mathbf{v}, \mathbf{w} \in \ell_2$, so folgt $\mathbf{v} + \mathbf{w} \in \ell_2$ und $\lambda v \in \ell_2$ und ℓ_2 wird zu einem komplexen Vektorraum.

Wir definieren weiterhin für alle $\mathbf{v}, \mathbf{w} \in \ell_2$ ein Skalarprodukt durch

$$\langle \mathbf{v} | \mathbf{w} \rangle := \sum_{i=0}^{\infty} \bar{\alpha}_i \beta_i$$

(Da die unendliche Reihe absolut konvergent ist, handelt es sich hier um eine wohldefinierte komplexe Zahl.)

Man zeigt leicht, dass die Eigenschaften (B_1'), (B_2), (H), (P) eines Skalarproduktes erfüllt sind.

Damit wird ℓ_2 zu einem unitären Vektorraum. Wir geben ein unendliches Orthonormalsystem für ℓ_2 an:

$$\begin{aligned}
e_0 &= (1,0,0,0,\dots) \\
e_1 &= (0,1,0,0,\dots) \\
e_2 &= (0,0,1,0,\dots) \\
&\vdots
\end{aligned}$$

Die hermitesche Norm von $\mathbf{v} = (\alpha_i)_i \in \ell_2$ ist $\|\mathbf{v}\| = \sqrt{\sum_i |\alpha_i|^2}$.

3.7.3 Der Vektorraum $\mathcal{C}([0,1])$ mit der L_2-Norm. Sei $\mathcal{C}([0,1])$ die Menge aller stetigen Funktionen $f : [0,1] \to \mathbb{R}$. Für die übliche Addition und Multiplikation mit Konstanten

$$(f+g)(x) := f(x) + g(x) \qquad (\lambda f)(x) := \lambda \cdot f(x) \quad (\lambda \in \mathbb{R})$$

wird $\mathcal{C}([0,1])$ zu einem reellen Vektorraum. Wir definieren für $f, g \in \mathcal{C}([0,1])$ ein Skalarprodukt

$$\langle f|g\rangle := \int_0^1 f(x)g(x)dx$$

Übung. Die Eigenschaften ($\mathrm{B}_{1,2}$),(S), (P) eines euklidischen Skalarproduktes sind in der Tat erfüllt.

Ausgestattet mit diesem Skalarprodukt wird $\mathcal{C}([0,1])$ ein euklidischer Vektorraum. Die euklidische Norm einer Funktion f ist

$$\|f\|_2 \;=\; \sqrt{\int_0^1 f(x)^2dx} \;=\; \left(\int_0^1 f(x)^2dx\right)^{1/2}$$

Wie in jedem euklidischen Vektorraum gilt die CAUCHY-SCHWARZsche Ungleichung 3.4.1:

$$|\langle f|g\rangle| \le \|f\|_2 \cdot \|g\|_2$$

d.h.
$$\left|\int_0^1 f(x)g(x)dx\right| \;\le\; \left(\int_0^1 f(x)^2dx\right)^{1/2}\left(\int_0^1 g(x)^2dx\right)^{1/2}$$

Ebenso gilt die Dreiecksungleichung 3.4.1:

$$\|f+g\|_2 \le \|f\|_2 + \|g\|_2$$

d.h.
$$\left(\int_0^1 (f(x)+g(x))^2dx\right)^{1/2} \;\le\; \left(\int_0^1 f(x)^2dx\right)^{1/2} + \left(\int_0^1 g(x)^2dx\right)^{1/2}$$

Beispiele von orthogonalen Funktionen erhält man aus den folgenden *Orthogonalitätsrelationen* für sinus und cosinus:

$$\int_0^1 \sin(2\pi mx)\sin(2\pi nx)\,dx = \begin{cases} 0 & \text{für } n \neq m \\ \frac{1}{2} & \text{für } n = m > 0 \end{cases}$$

$$\int_0^1 \sin(2\pi mx)\cos(2\pi nx)\,dx = \quad 0 \quad \text{für alle } n, m$$

$$\int_0^1 \cos(2\pi mx)\cos(2\pi nx)\,dx = \begin{cases} 0 & \text{für } n \neq m \\ \frac{1}{2} & \text{für } n = m > 0 \\ 1 & \text{für } n = m = 0 \end{cases}$$

Betrachten wir die Funktionen

$$v_0, v_1, v_2, \ldots \qquad\qquad w_1, w_2, \ldots$$

die für $x \in [0,1]$ durch

$$v_k(x) := \cos(2\pi kx) \qquad w_k(x) := \sin(2\pi kx) \qquad \text{für } k = 0, 1, 2, \ldots$$

definiert sind, dann bilden nach den obigen Orthogonalitätsrelationen diese Funktionen ein Orthogonalsystem in $\mathcal{C}([0,1])$. Durch Normierung erhalten wir ein Orthonormalsystem von Funktionen.

Übung (Analysis). Beweisen Sie die Orthogonalitätsrelationen.

Wir wählen nun endlich viele von diesen Funktionen, z.B.

$$v_0, v_1, \ldots, v_n \qquad w_1, \ldots, w_n$$

Diese bilden eine Orthogonalbasis eines linearen Teilraums $\mathcal{T}_n$ von $\mathcal{C}([0,1])$. $\mathcal{T}_n$ besteht aus allen Linearkombinationen der Form

$$T = \alpha_0 v_0 + \sum_{k=1}^{n} (\alpha_k v_k + \beta_k w_k)$$

d.h.
$$T(x) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \cos(2\pi kx) + \beta_k \sin(2\pi kx)\ .$$

Solche Funktionen heißen *trigonometrische Polynome* vom Grad $\leq n$.

Wir bezeichnen mit π_n die Orthogonalprojektion von $\mathcal{C}([0,1])$ auf den linearen Teilraum $\mathcal{T}_n$.

Für jede stetige Funktion $f \in \mathcal{C}([0,1])$ ist dessen Projektion $\pi_n(f)$ ein trigonometrisches Polynom vom Grad $\leq n$; nach **??** (6) ist $\pi_n(f)$ die beste Approximation von f durch ein trigonometrisches Polynom vom Grad $\leq n$ bzgl. der L_2-Norm. Nach der allgemeinen Formel **??** ist $\pi_n(f)$ gegeben durch

$$\pi_n(f) = \langle \widetilde{v}_0 | f \rangle \widetilde{v}_0 + \sum_{k=1}^{n} \left(\langle \widetilde{v}_k | f \rangle \widetilde{v}_k + \langle \widetilde{w}_k | f \rangle \widetilde{w}_k \right)$$

wobei $\widetilde{v}_0$, $\widetilde{v}_1$, $\widetilde{v}_2$, $\ldots$, $\widetilde{w}_1$, $\widetilde{w}_2$, $\ldots$ die normierten $v_0, v_1, v_2, \ldots, w_1, w_2, \ldots$ sind.

Dies bedeutet explizit:

$$\pi_n(f)(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n} \alpha_k \cos(2\pi kx) + \beta_k \sin(2\pi kx)$$

mit

$$\alpha_k = 2\int_0^1 f(x)\cos(2\pi kx)dx \ , \quad k = 0, 1, 2, \ldots$$

$$\beta_k = 2\int_0^1 f(x)\sin(2\pi kx)dx \ , \quad k = 1, 2, 3, \ldots$$

Die Koeffizienten α_k and β_k heißen *Fourier-Koeffizienten* von f.

Bemerkung Die unendliche Reihe

$$\frac{\alpha_0}{2} + \sum_{k=1}^{\infty} \alpha_k \cos(2\pi kx) + \beta_k \sin(2\pi kx)$$

heißt *Fourierreihe* der Funktion f. Diese konvergiert im Sinne der L_2-Norm gegen f. Leider ist es nicht richtig, dass die Fourierreihe von f für jedes x gegen $f(x)$ konvergiert. Ist allerdings die Funktion f zweimal stetig differenzierbar und $f(0) = f(1)$, dann konvergiert die Fourierreihe von f gleichmäßig nach f. Das Teilgebiet der Mathematik, in welchem derartige Fragen untersucht werden, heißt *harmonische Analysis*.

KAPITEL 4

Lineare Abbildungen und Matrizen

$\mathbb{K}$ stehe in diesem Kapitel immer einen der Körper $\mathbb{R}$ oder $\mathbb{C}$. Ferner seien V und W beliebige Vektorräume über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ (so z.B. $V = \mathbb{K}^n, W = \mathbb{K}^m$). Ein Skalarprodukt spielt zunächst keine Rolle. Weiterhin seien ab jetzt zur Vereinfachung die Vektoren, d.h. die Elemente der Vektorräume mit gewöhnlichen Buchstaben $u, v, w, \ldots$ bezeichnet (nicht mehr mit fettgedruckten Buchstaben wie bisher $\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$).

4.1. Reelle und komplexe Matrizen

Alles, was wir in Kapitel 1 über reelle Matrizen gesagt haben, bleibt gültig für Matrizen mit komplexen Einträgen. Die Beweise bleiben dieselben, da wir dort nur die in beliebigen Körpern geltenden Gesetze benutzt haben. Wir wiederholen:

4.1.1 $m \times n$**-Matrizen mit Einträgen aus** $\mathbb{K}$**.** Eine $m \times n$-*Matrix mit Einträgen aus* $\mathbb{K}$ ist ein rechteckiges Schema aus m Zeilen und n Spalten

$$A = (\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1j} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2j} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \ldots & \alpha_{ij} & \ldots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} & \ldots & \alpha_{mn} \end{pmatrix}$$

mit $\alpha_{ij} \in \mathbb{K}$ für $i = 1, \ldots, m$, $j = 1, \ldots, n$.

$\mathcal{M}(m \times n, \mathbb{K})$ bezeichne die Menge aller $m \times n$-Matrizen mit Einträgen aus $\mathbb{K}$. Oft schreibt man $\mathbb{K}^{m \times n}$ anstelle von $\mathcal{M}(m \times n, \mathbb{K})$.

Spezialfälle.

- $m \times 1$-Matrix = Spaltenvektor $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{K}^m$
- $1 \times n$-Matrix = Zeilenvektor $\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$
- Nullmatrix $0 := \begin{pmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$
- $n \times n$-Matrix = quadratische Matrix

Spezielle quadratische Matrizen.

- Die *Einheitsmatrix*
$$E = E_n == (\delta_{ij}) \quad \text{mit } \delta_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{falls } i \neq j \end{cases}$$
- *Diagonalmatrizen*: $\alpha_{ij} = 0$, falls $i \neq j$:
- *obere Dreiecksmatrizen*: $\alpha_{ij} = 0$, falls $i > j$:

4.1.2 $\mathcal{M}(m \times n, \mathbb{K})$ als Vektorraum. Für zwei $m \times n$-Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{ij})$ mit Einträgen aus $\mathbb{K}$ und für alle $\lambda \in \mathbb{K}$ definieren wir:

$$\begin{aligned} A + B = (\alpha_{ij}) + (\beta_{ij}) &:= (\alpha_{ij} + \beta_{ij}) \\ \lambda A = \lambda(\alpha_{ij}) &:= (\lambda\alpha_{ij}). \end{aligned}$$

Man zeigt leicht, dass die Vektorraumaxiome (V1)–(V8) für die Addition und die skalare Multiplikation von $m \times n$-Matrizen gelten. So wird $\mathcal{M}(m \times n, \mathbb{K})$ ein $\mathbb{K}$-Vektorraum (vergleiche 1.4.2). Man beachte, dass die Addition von Matrizen nur definiert ist, wenn die Matrizen dieselbe „Größe“ haben.

Um die Dimension von $\mathcal{M}(m \times n, \mathbb{K})$ als $\mathbb{K}$-Vektorraum zu bestimmen, betrachten wir die Matrizen

$$E_{ij} := \begin{array}{r c} & \begin{array}{ccccc} & & j & & \\ & & \downarrow & & \end{array} \\ \begin{array}{r} \\ \\ i \rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array}$$

mit einer 1 als Eintrag in der i-ten Zeile und j-ten Spalte und 0 sonst. Diese speziellen Matrizen E_{ij} mit $i = 1, \dots, m$ und $j = 1, \dots, n$ bilden eine Basis des Vektorraumes $\mathcal{M}(m \times n, \mathbb{K})$. Da die Anzahl dieser Matrizen $m \cdot n$ ist, gilt

$$\dim \mathcal{M}(m \times n, \mathbb{K}) = m \cdot n.$$

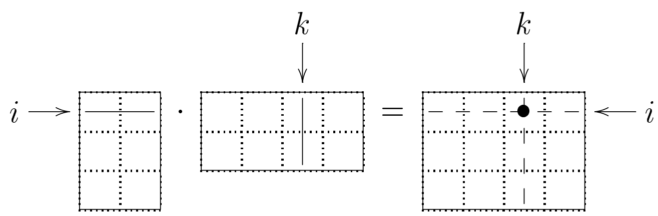
4.1.3 Multiplikation von Matrizen. Das Produkt $A \cdot B$ zweier Matrizen ist definiert, wenn die Anzahl n der Spalten von A gleich der Anzahl der Zeilen von B (siehe 1.4.3) ist. Es sei also $A = (\alpha_{ij})$ eine $m \times n$-Matrix und $B = (\beta_{ik})$ eine $n \times p$-Matrix mit Einträgen aus $\mathbb{K}$.

Dann ist das Produkt $A \cdot B$ definiert als die $m \times p$-Matrix $C = (\gamma_{ik})$ mit

$$\gamma_{ik} := \sum_{j=1}^{n} \alpha_{ij}\beta_{jk} = \alpha_{i1}\beta_{1k} + \dots + \alpha_{in}\beta_{nk}$$

oder weniger formal:

$$\gamma_{ik} = (i\text{-te Zeile von } A) \cdot (k\text{-te Spalte von } B)$$



Regeln für das Matrixprodukt. *Es seien A, A' zwei $m \times n$-Matrizen, B, B' zwei $n \times p$-Matrizen, C eine $p \times q$-Matrix mit Einträgen aus $\mathbb{K}$ und $\lambda \in \mathbb{K}$. Dann gilt:*

(A1) $$A(BC) = (AB)C$$

(A2) $$E_m A = A = A E_n$$

(A3) $$(A + A')B = AB + A'B \quad , \quad A(B + B') = AB + AB'$$

(A4) $$(\lambda A)B = \lambda(AB) = A(\lambda B)$$

Dies sind dieselben Regeln wie für das Produkt von reellen Matrizen. Sie werden genauso bewiesen, wie wir es für den reellen Fall getan haben (siehe 1.4.4 und 1.4.5); tatsächlich wurden dort nur die arithmetischen Gesetze benutzt, die in jedem Körper $\mathbb{K}$ gelten.

Beschränken wir uns auf quadratische Matrizen, so sehen wir, daß zwei $n \times n$-Matrizen A und B immer multipliziert werden können und dass das Produkt $A \cdot B$ wieder eine $n \times n$-Matrix ist. Also ist die Multiplikation eine (innere) Verknüpfung auf der Menge $\mathcal{M}(n \times n, \mathbb{K})$ aller $n \times n$-Matrizen mit Einträgen aus $\mathbb{K}$. Spezialisieren wir die obigen Regeln auf quadratische Matrizen, erhalten wir (vergleiche 4.5.4):

4.1.4 $\mathcal{M}(n \times n, \mathbb{K})$ als $\mathbb{K}$-Algebra. Für alle $A, A', B, B', C \in \mathcal{M}(n \times n, \mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt:

(A1) $$A(BC) = (AB)C$$

(A2) $$E_n A = A = A E_n$$

(A3) $$(A + A')B = AB + A'B \quad , \quad A(B + B') = AB + AB'$$

(A4) $$(\lambda A)B = \lambda(AB) = A(\lambda B)$$

Also ist die Matrixmultiplikation assoziativ (A1), hat ein neutrales Element E_n (A2) und die Distributivgesetze gelten (A3). Also ist $\mathcal{M}(n \times n, \mathbb{K})$ nicht nur ein $\mathbb{K}$-Vektorraum (der Dimension n^2), sondern auch ein Ring. Wegen (A4) ist $\mathcal{M}(n \times n, \mathbb{K})$ sogar eine $\mathbb{K}$-Algebra (im Sinne von 4.5.4).

4.1.5 Die Gruppe $\mathrm{GL}(n, \mathbb{K})$ der invertierbaren Matrizen. Bzgl. der Multiplikation bilden die $n \times n$-Matrizen eine Halbgruppe mit dem neutralen Element $E = E_n$. Wie im allgemeinen Fall (2.1.4) heißt eine $n \times n$-Matrix A *invertierbar*, wenn es eine Matrix $B \in \mathcal{M}(n \times n, \mathbb{K})$ gibt, so dass

$$A \cdot B = E = B \cdot A$$

gilt. Die Inverse B ist, falls sie existiert, eindeutig bestimmt und man schreibt $B = A^{-1}$.

Wir bezeichnen mit $\mathrm{GL}(n, \mathbb{K})$ die Menge aller invertierbaren Matrizen $A \in \mathcal{M}(n \times n, \mathbb{K})$. Nach 2.1.7 ist $\mathrm{GL}(n, \mathbb{K})$ eine Gruppe bzgl. der Matrixmultiplikation. Speziell erhalten wir (siehe 2.1.4):

(1) E ist invertierbar mit $E^{-1} = E$.

(2) Ist A invertierbar, dann ist auch A^{-1} invertierbar mit $(A^{-1})^{-1} = A$.

(3) Sind A und B invertierbar, so ist das Produkt AB auch invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Für praktische Zwecke können wir den (erweiterten) Gauß-Jordan Eliminations-Algorithmus benutzen, um zu entscheiden, ob eine gegebene Matrix A invertierbar ist und um gleichzeitig die Inverse A^{-1} auszurechnen (falls sie existiert), was in derselben Weise wie bei reellen Matrizen in 1.4.8 geschieht.

4.1.6 Transponierte Matrix (vergleiche 1.4.9). Wie für reelle Matrizen ist die *Transponierte* einer $m \times n$-Matrix

$$A = \begin{array}{c} \\ i \rightarrow \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} \phantom{\alpha_{11}} & \phantom{\alpha_{12}} & & j\,\downarrow & & \phantom{\alpha_{1n}} \end{array} \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1j} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mj} & \dots & \alpha_{mn} \end{pmatrix} \end{array}$$

definiert als die $n \times m$-Matrix

$$A^T = \begin{array}{c} \\ \\ \\ j \rightarrow \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} \phantom{\alpha_{11}} & & i\,\downarrow & & \phantom{\alpha_{m1}} \end{array} \\ \begin{pmatrix} \alpha_{11} & \dots & \alpha_{i1} & \dots & \alpha_{m1} \\ \alpha_{12} & \dots & \alpha_{i2} & \dots & \alpha_{m2} \\ \vdots & & \vdots & & \vdots \\ \alpha_{1j} & \dots & \alpha_{ij} & \dots & \alpha_{mj} \\ \vdots & & \vdots & & \vdots \\ \alpha_{1n} & \dots & \alpha_{in} & \dots & \alpha_{mn} \end{pmatrix} \end{array}.$$

Die i-te Zeile von A wird zur i-ten Spalte von A^T, die j-te Spalte von A wird zur j-ten Zeile von A^T.

Regeln für das Transponieren. *Für alle $m \times n$-Matrizen A und A' und jede $n \times p$-Matrix B gilt:*

(T1) $$(A^T)^T = A$$

(T2) $$(A + A')^T = A^T + (A')^T$$

(T3) $$(\lambda A)^T = \lambda A^T$$

(T4) $$(AB)^T = B^T A^T$$

Übung. Für eine invertierbare $n \times n$-Matrix A zeige man, daß A^T ebenfalls invertierbar ist und daß gilt

$$(A^T)^{-1} = (A^{-1})^T.$$

4.2. Lineare Abbildungen

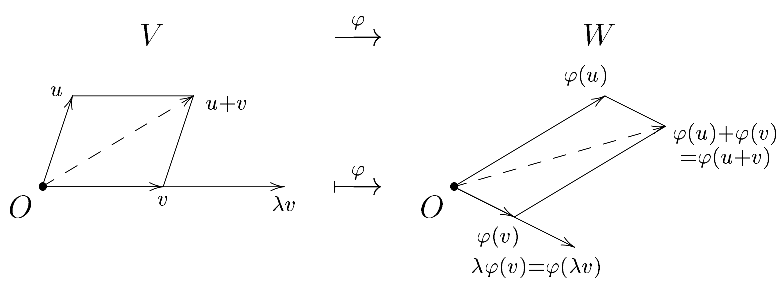
4.2.1 Definition. Eine Abbildung $\varphi\colon V \to W$ wird *lineare Abbildung* (oder auch *linearer Operator*) genannt, wenn sie die folgenden Eigenschaften erfüllt:

(L1) $\quad \varphi(u+v) = \varphi(u) + \varphi(v) \quad$ für alle $u, v \in V$,

(L2) $\quad \varphi(\lambda v) = \lambda\varphi(v) \quad$ für alle $v \in V,\ \lambda \in \mathbb{K}$.

Die Eigenschaften (L1) und (L2) können zu einer einzigen Aussage zusammengefasst werden:

(L*) $\quad \varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v) \quad$ für alle $u, v \in V,\ \lambda, \mu \in \mathbb{K}$.



Eine lineare Abbildung $\varphi\colon V \to W$ nennt man manchmal auch einen *Homomorphismus*. Ist eine lineare Abbildung bijektiv, d.h. injektiv und surjektiv, wird sie auch *Isomorphismus* genannt.

Eine lineare Abbildung $\varphi\colon V \to V$ eines Vektorraumes V in sich wird *Endomorphismus* von V genannt. Einen bijektiven Endomorphismus nennen wir *Automorphismus*.

Alle eben definierten Begriffe und Notationen werden in der folgenden Tabelle zusammengefasst.

Homomorphismus:		lineare Abbildung $\varphi\colon V \to W$
Isomorphismus:	bijektive	lineare Abbildung $\varphi\colon V \to W$
Endomorphismus:		lineare Abbildung $\varphi\colon V \to V$
Automorphismus:	bijektive	lineare Abbildung $\varphi\colon V \to V$

Tabelle 1. Spezielle lineare Abbildungen

$\mathrm{Hom}(V,W) :=$	Menge aller linearen Abbildungen $\varphi\colon V \to W$
$\mathrm{End}(V) :=$	Menge aller Endomorphismen $\varphi\colon V \to V$
$\mathrm{Aut}(V) :=$	$\mathrm{GL}(V)$ = Menge aller Automorphismen $\varphi\colon V \to V$

Tabelle 2. Bezeichnungen für Mengen von linearen Abbildungen

4.2.2 Beispiel. Aus geometrischen Gründen kann man sehen, daß die folgenden Abbildungen linear sind:

Im Euklidischen $\mathbb{R}^2$: Drehung um einen Winkel γ um den Ursprung.
Im Euklidischen $\mathbb{R}^2$: Spiegelung an einer Geraden durch den Ursprung.
Im Euklidischen $\mathbb{R}^3$: Drehung um eine Achse durch den Urprung um einen Winkel γ.
Im Euklidischen $\mathbb{R}^3$: Spiegelung an einer Ebene durch den Ursprung.

In 3.5.3 (0) hatten wir nachgerechnet, daß in einem beliebigen Euklidischen oder unitären Vektorraum die Orthogonalprojektion π_U auf einen endlichdimensionalen linearen Teilraum linear ist.

4.2.3 Eigenschaften von linearen Abbildungen $\varphi\colon V \to W$.

(1) $\varphi(0) = 0$.

(2) $\varphi(-v) = -\varphi(v)$.

(3) $\varphi(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda_1 \varphi(v_1) + \lambda_2 \varphi(v_2) + \cdots + \lambda_n \varphi(v_n)$.

(4) *Ist U ein linearer Teilraum von V, dann ist sein Bild $\varphi(U)$ ein linearer Teilraum von W.*

(5) *Ist U ein linearer Teilraum von W, dann ist dessen Urbild $\varphi^{-1}(U) := \{v \in V \mid \varphi(v) \in U\}$ ein linearer Teilraum von V.*

(6) *Ist A ein affiner Teilraum von V, dann ist das Bild $\varphi(A)$ ein affiner Teilraum von W.*

(7) *Ist B ein affiner Teilraum von W, dann ist dessen Urbild $\varphi^{-1}(B)$ ein affiner Teilraum von V oder die leere Menge.*

Beweis. (1) Wähle $\lambda = 0$ in (L2).

(2) Wähle $\lambda = -1$ in (L2).

(3) Wende (L1) und (L2) mehrfach an.

(4) Übung.

(5) (U1): $\varphi(0) = 0 \in U$, da U ein linearer Teilraum ist und (U1) erfüllt. Also gilt $0 \in \varphi^{-1}(U)$.

(U2): Seien $u, v \in \varphi^{-1}(U)$; dann ist $\varphi(u) \in U$ und $\varphi(v) \in U$. Da U ein linearer Teilraum ist, erfüllt U (U1) und es folgt $\varphi(u) + \varphi(v) \in U$. Es gilt aber auch $\varphi(u) + \varphi(v) = \varphi(u + v)$ nach (L1). Damit folgt $u + v \in \varphi^{-1}(U)$.

(U3): Sei $v \in \varphi^{-1}(U)$. Dann ist $\varphi(v) \in U$. Da U ein linearer Teilraum ist, gilt (U3) und damit $\lambda\varphi(v) \in U$ für alle $\lambda \in \mathbb{K}$. Da $\lambda\varphi(v) = \varphi(\lambda v)$ nach (L2) ist, folgern wir $\lambda v \in \varphi^{-1}(U)$.

(6) Es sei A ein affiner Teilraum von V. Dann gibt es einen linearen Teilraum U und ein Element $a \in V$, so dass $A = a + U$ ist. Es folgt $\varphi(A) = \varphi(a + U) = \varphi(a) + \varphi(U)$ nach (L1). Wegen (4) ist $\varphi(U)$ ein linearer Teilraum von W. Also ist $\varphi(A)$ ein affiner Teilraum von W.

(7) Übung. Man bedenke dabei, dass das Urbild von B leer ist, wenn B die Menge $\varphi(V)$ nicht trifft. Die leere Menge ist aber nach unserer Definition kein affiner Raum. □

4.2.4 Bild und Kern. Für jede lineare Abbildung $\varphi\colon V \to W$ definieren wir das *Bild* als die Menge

$$\operatorname{im}\varphi := \varphi(V) = \{\varphi(v) \mid v \in V\} \subseteq W$$

und den *Kern* als die Menge

$$\ker\varphi := \varphi^{-1}(0) = \{v \in V \mid \varphi(v) = 0\} \subseteq V.$$

Nach 4.2.3 (4) und 4.2.3 (5) sind $\operatorname{im}\varphi$ und $\ker\varphi$ lineare Teilräume von W bzw. V.

Beispiele. (a) Für die orthogonale Projektion $\pi\colon \mathbb{R}^2 \to \mathbb{R}^2$ auf eine Gerade g (durch 0) gilt $\operatorname{im}\pi = g$ und $\ker\pi = g^\perp$, die Gerade durch 0 senkrecht zu g.

(b) Für eine Drehung φ des $\mathbb{R}^2$ um den Winkel γ um 0 ist $\operatorname{im}\varphi = \mathbb{R}^2$ und $\ker\varphi = \{0\}$.

4.2.5 Kriterium für die Injektivität einer linearen Abbildung. *Eine lineare Abbildung $\varphi\colon V \to W$ ist genau dann injektiv, wenn* $\ker\varphi = \{0\}$.

Beweis. $\Rightarrow$: Da $\varphi(0) = 0$ ist, gilt immer $0 \in \ker\varphi$. Es sei nun $v \in \ker\varphi$ beliebig. Nach Definition folgt $\varphi(v) = 0$, also gilt $\varphi(v) = \varphi(0)$. Ist φ injektiv, folgt $v = 0$, d.h. $\ker\varphi = \{0\}$.

$\Leftarrow$: Wir nehmen nun $\ker\varphi = \{0\}$ an. Um die Injektivität von φ zu zeigen, seien u, v Elemente in V mit $\varphi(u) = \varphi(v)$. Es folgt $\varphi(u) - \varphi(v) = 0$. Aufgrund der Linearität von φ schließen wir $\varphi(u - v) = 0$, d.h. $u - v \in \ker\varphi = \{0\}$. Also ist $u - v = 0$, d.h. $u = v$. $\square$

4.2.6 Dimensionssatz. *Für jede lineare Abbildung $\varphi\colon V \to W$ gilt*

$$\dim(\operatorname{im}\varphi) + \dim(\ker\varphi) = \dim V.$$

Beweis. Wir führen den Beweis nur im Fall $\dim V = n < \infty$. Dabei werden wir wesentlich mehr als nur den Dimensionssatz zeigen.

Wir wissen aus 4.2.4, dass $\ker\varphi$ ein linearer Teilraum von V ist. Sei $k := \dim(\ker\varphi)$. Dann hat $\ker\varphi$ eine aus k Vektoren $v_1, \ldots, v_k$ bestehende Basis. Nach dem Basisergänzungssatz (2.6.6) können wir Vektoren $v_{k+1}, \ldots, v_n$ finden, so dass

$$\mathcal{B} : v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$$

eine Basis von V ist. Betrachten wir nun die Bildvektoren

$$w_1 := \varphi(v_{k+1}), \ldots, w_{n-k} := \varphi(v_n).$$

Behauptung. $w_1, \ldots, w_{n-k}$ *ist eine Basis von* $\operatorname{im}(\varphi)$.

Dies beendet dann den Beweis, da daraus $\dim(\operatorname{im}\varphi) = n - k = \dim V - \dim(\ker\varphi)$ folgt.

Beweis der Behauptung. (a) $w_1, \ldots, w_{n-k}$ erzeugen $\operatorname{im}\varphi$: Um dies einzusehen, sei $w \in \operatorname{im}\varphi$. Dann gibt es $v \in V$, so dass $w = \varphi(v)$ ist. Der Vektor v kann als Linearkombination der oben betrachteten Basis $\mathcal{B}$ dargestellt werden:

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \cdots + \lambda_n v_n$$

Also folgt

$$w = \varphi(v)$$

$$= \lambda_1\varphi(v_1) + \cdots + \lambda_m\varphi(v_k) + \lambda_{k+1}\varphi(v_{k+1}) + \cdots + \lambda_n\varphi(v_n)$$
$$= 0 + \cdots + 0 + \lambda_{k+1}w_1 + \cdots + \lambda_n w_{n-k},$$

da $v_1, \ldots, v_k \in \ker\varphi$ ist.

(b) $w_1, \ldots, w_{n-k}$ sind linear unabhängig: Wir nehmen $\mu_1 w_1 + \cdots + \mu_{n-k}w_{n-k} = 0$ an. Dann ist $\mu_1\varphi(v_{k+1}) + \cdots + \mu_{n-k}\varphi(v_n) = 0$. Damit gilt $\varphi(\mu_1 v_{k+1} + \cdots + \mu_{n-k}v_n) = 0$ aufgrund der Linearität von φ. Also folgt $\mu_1 v_{k+1} + \cdots + \mu_{n-k}v_n \in \ker\varphi$. Die Vektoren in $\ker\varphi$ haben aber eine eindeutige Darstellung als Linearkombination der Vektoren $v_1, \ldots, v_k$. Daraus ergibt sich $\mu_1 = \cdots = \mu_{n-k} = 0$. □

Man kann den Dimensionssatz an den beiden Beispielen in 4.2.4 illustrieren.

4.2.7 Der Rang einer linearen Abbildung $\varphi\colon V \to W$. Der *Rang* einer linearen Abbildung φ ist definiert als die Dimension des Bildes von φ:

$$\operatorname{rank}\varphi := \dim(\operatorname{im}\varphi) = \dim(\varphi(V)).$$

Der Dimensionssatz besagt $\operatorname{rank}\varphi = \dim V - \dim(\ker\varphi)$.

4.2.8. Wir betrachten nun einen n-dimensionalen Vektorraumes V mit einer Basis $v_1, \ldots, v_n$ und eine lineare Abbildung $\varphi : V \to W$. Wir untersuchen Eigenschaften der Abbildung φ an Hand der Bilder $\varphi(v_1), \ldots, \varphi(v_n)$ der Basisvektoren:

Eigenschaften.

(1) Ist $\psi\colon V \to W$ eine weitere lineare Abbildung mit $\psi(v_j) = \varphi(v_j)$ für $j = 1, \ldots, n$, dann folgt $\psi = \varphi$; d.h. durch die Bilder der Basisvektoren ist eine lineare Abbildung eindeutig bestimmt.

(2) Die Bilder $\varphi(v_1), \ldots, \varphi(v_n)$ der Basisvektoren erzeugen $\operatorname{im}\varphi$.

(3) Der Rang von φ ist gleich der maximalen Anzahl von linear unabhängigen Vektoren unter den $\varphi(v_1), \ldots, \varphi(v_n)$.

(4) φ ist surjektiv $\iff \operatorname{rank}\varphi = \dim W \iff \varphi(v_1), \ldots, \varphi(v_n)$ erzeugen W.

(5) φ ist injektiv $\iff \varphi(v_1), \ldots, \varphi(v_n)$ sind linear unabhängig.

(6) φ ist bijektiv $\iff \varphi(v_1), \ldots, \varphi(v_n)$ ist eine Basis von W.

Beweis. (1) Sei $\psi\colon V \to W$ linear mit $\psi(v_j) = \varphi(v_j$ für $j = 1, \ldots, n$. Sei $v \in V$ beliebig. Dann hat v eine Darstellung $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Wegen der Linearität von ψ und φ ergibt sich $\psi(v) = \alpha_1\psi(v_1) + \cdots + \alpha_n\psi(v_n) = \alpha_1\varphi(v_1) + \cdots + \alpha_n\varphi(v_n) = \varphi(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \varphi(v)$.

(2) Sei $w \in \operatorname{im}\varphi = \varphi(V)$. Dann gibt es ein $v \in V$, so dass $w = \varphi(v)$ ist. Da $v_1, \ldots, v_n$ eine Basis von V ist, kann der Vektor v als eine Linearkombination

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

dargestellt werden. Da φ linear ist, folgern wir

$$w = \varphi(v) = \lambda_1\varphi(v_1) + \cdots + \lambda_n\varphi(v_n).$$

(3) In der Tat ist $\operatorname{rank}\varphi = \dim(\operatorname{im}\varphi)$ die Zahl der Elemente einer Basis von $\operatorname{im}\varphi$. Wegen der Folgerung aus 2.6.6 ist jedes maximale linear unabhängige Teilsystem einer erzeugenden Menge eine Basis von $\operatorname{im}\varphi$.

(4)
$$\begin{aligned} \operatorname{rank}\varphi = \dim W &\iff \dim(\operatorname{im}\varphi) = \dim W \\ &\iff \operatorname{im}\varphi = W \\ &\iff \varphi \text{ ist surjektiv.} \end{aligned}$$

(5)
$$\begin{aligned} \varphi \text{ ist injektiv} &\iff \ker\varphi = \{0\} && \text{nach 4.2.5} \\ &\iff \dim(\ker\varphi) = 0 \\ &\iff \operatorname{rank}\varphi = \dim V - \underbrace{\dim(\ker\varphi)}_{=0} = n && \text{nach 4.2.7} \\ &\iff \varphi(v_1), \dots, \varphi(v_n) \text{ sind lin. unabh.} && \text{nach (3)} \end{aligned}$$

(6) Dies folgt direkt aus (4) und (5). □

4.2.9 Theorem. *Für jede lineare Abbildung $\varphi\colon V \to V$ eines endlichdimensionalen Vektorraumes V in sich sind die folgenden Eigenschaften äquivalent:*

(1) *φ ist injektiv.*

(2) *φ ist surjektiv.*

(3) *φ ist bijektiv.*

Beweis. Es sei $v_1, \dots, v_n$ eine Basis von V. Der Satz folgt direkt aus 4.2.8 (4), (5), (6); denn nach 2.6.7 sind für die n Vektoren $\varphi(v_1), \dots, \varphi(v_n)$ in dem n-dimensionalen Vektorraum V die folgenden Aussagen äquivalent:

(c) $\varphi(v_1), \dots, \varphi(v_n)$ sind linear unabhängig;
(d) $\varphi(v_1), \dots, \varphi(v_n)$ erzeugen V;
(e) $\varphi(v_1), \dots, \varphi(v_n)$ bilden eine Basis von V.

□

Auf unendlichdimensionalen Vektorräumen V gibt es surjektive lineare Abbildungen $\varphi\colon V \to V$, die nicht injejektiv sind. Es sei zum Beispiel $V = \mathcal{P}(\mathbb{R})$ der Vektorraum aller reellen Polynome und $\varphi\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ die Abbildung

$$p(x) \mapsto x \cdot p(x).$$

Diese Abbildung ist linear und injektiv, jedoch nicht surjektiv. Dagegen ist die Ableitung $p(x) \mapsto p'(x)$ eine lineare Abbildung auf $\mathcal{P}(\mathbb{R})$ die surjektiv, aber nicht injektiv ist.

4.2.10 Ergänzung. *Seien V, W Vektorräume. Sei $v_1, \dots, v_n$ eine Basis von V. Weiterhin seien $w_1, \dots, w_n$ beliebig vorgegebene Vektoren in W. Dann gibt es genau eine lineare Abbildung $\varphi\colon V \to W$, so dass gilt:*

$$\varphi(v_1) = w_1,\ \varphi(v_2) = w_2,\ \dots, \varphi(v_n) = w_n.$$

Beweis. Jedes $v \in V$ besitzt eine eindeutige Darstellung als Linearkombination

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\,.$$

Also können wir φ definieren durch

$$\varphi(v) := \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n\,.$$

Es gilt dann $\varphi(v_j) = w_j$ für $j = 1, \ldots, n$ (da $v_j = 0 \cdot v_1 + \cdots + 0 \cdot v_{j-1} + 1 \cdot v_j + 0 \cdot v_{j+1} + \cdots + 0 \cdot v_n$ gilt, woraus $\varphi(v_j) = w_j$ folgt). Die Linearität von φ rechnet man leicht aus der Definition nach. Also gibt es eine lineare Abbildung $\varphi\colon V \to W$ mit der gewünschten Eigenschaft. Nach 4.2.8 (1) ist eine lineare abbildung durch ihre Werte auf einer Basis eindeutig bestimmt.

□

4.2.11 $\mathbb{K}^n$ als Prototyp eines n-dimensionalen $\mathbb{K}$-Vektorraumes. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Man wähle eine Basis $\mathcal{B}\colon v_1, v_2, \ldots, v_n$ in V. Wir betrachten in $\mathbb{K}^n$ die Standardbasis $e_1, e_2, \ldots, e_n$. Nach 4.2.9 gibt es eine eindeutig bestimmte lineare Abbildung $\iota_{\mathcal{B}}\colon \mathbb{K}^n \to V$ mit $\iota_{\mathcal{B}}(e_1) = v_1, \ldots, \iota_{\mathcal{B}}(e_n) = v_n$. Nach 4.2.8 (6) ist diese Abbildung bijektiv, d.h. ein Isomorphismus. Speziell gilt damit folgende

Proposition. *Jeder n-dimensionale $\mathbb{K}$-Vektorraum V ist isomorph zu $\mathbb{K}^n$.*

Bemerkung. Zwei Vektorräume heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Zwei isomorphe Vektorräume haben (im Sinn der Definition von Vektorräumen) völlig gleiche Eigenschaften.

Stellen wir die Abbildung $\iota_{\mathcal{B}}$ genauer dar. Wegen

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n$$

folgt aus der Linearität

$$\iota_{\mathcal{B}} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n,$$

d.h. $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ wird auf den Vektor $v \in V$ abgebildet, dessen Koordinaten bzgl. der Basis $\mathcal{B}$ $\lambda_1, \lambda_2, \ldots, \lambda_n$ sind. Daraus folgt, dass die Inverse der Abbildung $\iota_{\mathcal{B}} : K^n \to V$ diejenige Abbildung $\kappa_{\mathcal{B}}\colon V \to \mathbb{K}^n$ ist, die zu jedem Vektor $v \in V$ den Vektor $[v]_{\mathcal{B}}$ mit Koordinaten bzgl. der Basis $\mathcal{B}$ (siehe 4.4.8) zuordnet.

4.2.12 Beispiel. Der Vektorraum $\mathcal{P}_n(\mathbb{R})$ der reellen Polynome von Grad $\leq n$ ist isomorph zu $\mathbb{R}^{n+1}$.

Um einen Isomorphismus zu finden, können wir die Basis $\mathcal{B} : p_0, p_1, \ldots, p_n$ mit $p_j(x) = x^j$ in $\mathcal{P}_n(\mathbb{R})$ benutzen. Der Isomorphismus $\iota_{\mathcal{B}}$ ist dann gegeben durch:

$$\iota_{\mathcal{B}}\colon \mathbb{R}^{n+1} \to \mathcal{P}_n(\mathbb{R})$$

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$$

4.3. Lineare Abbildungen $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ und Matrizen

In diesem Abschnitt stellen wir Beziehungen zwischen linearen Abbildungen φ und Matrizen her. Wir betrachten lineare Abbildungen $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ und Matrizen $A \in \mathcal{M}(m \times n, \mathbb{K})$. Wir beschränken uns zunächst immer auf die Standardbasis

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \ldots, \qquad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

des $\mathbb{K}^n$ und analog auf diejenige des $\mathbb{K}^m$.

4.3.1 Von Matrizen zu linearen Abbildungen. Es sei $A = (\alpha_{ij})$ eine $m \times n$-Matrix mit Einträgen $\alpha_{ij} \in \mathbb{K}$. Jeder Vektor $x \in \mathbb{K}^n$ kann als eine $n \times 1$-Matrix betrachtet werden. Also ist das Matrixprodukt Ax wohldefiniert und Ax ist eine $m \times 1$-Matrix, d.h. ein Vektor $Ax \in \mathbb{K}^m$:

$$Ax = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_1 + & \ldots & +\alpha_{1n}x_n \\ \alpha_{21}x_1 + & \ldots & +\alpha_{2n}x_n \\ & \vdots & \\ \alpha_{m1}x_1 + & \ldots & +\alpha_{mn}x_n \end{pmatrix}$$

Wir können also eine Abbildung $\varphi_A\colon \mathbb{K}^n \to \mathbb{K}^m$ durch $\varphi_A(x) = Ax$ für alle $x \in \mathbb{K}^n$ definieren.

Behauptung. *Die Abbildung $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ ist linear.*

Beweis. (L1) $\varphi_A(x+y) = A \cdot (x+y) = Ax + Ay = \varphi_A(x) + \varphi_A(y)$ nach (A3).

(L2) $\varphi_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda\varphi_A(x)$ nach (A4). □

Wir bemerken:

$$\varphi_A(e_j) = Ae_j = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{nj} \end{pmatrix} = j\text{-te Spalte von } A.$$

In anderen Worten: Die Bilder der Standardbasisvektoren $e_1, e_2, \ldots, e_n$ sind die Spalten $w_1 = Ae_1, w_2 = Ae_2, \ldots, w_n = Ae_n$ der Matrix A:

$$A = \begin{array}{c} \begin{array}{cccc} \overbrace{\hphantom{\alpha_{11}}}^{w_1} & \overbrace{\hphantom{\alpha_{12}}}^{w_2} & \cdots & \overbrace{\hphantom{\alpha_{1n}}}^{w_n} \end{array} \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \end{array}.$$

4.3.2 Die Matrix einer linearen Abbildung. Sei $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung. Wir bilden die $m \times n$-Matrix, deren Spalten die Bilder $\varphi(e_1), \ldots, \varphi(e_n)$ der Vektoren der Standardbasis $e_1, \ldots, e_n$ in $\mathbb{K}^n$ sind:

$$[\varphi] := A := \begin{pmatrix} | & | & & | \\ | & | & & | \\ \varphi(e_1) & \varphi(e_2) & \ldots & \varphi(e_n) \\ | & | & & | \\ | & | & & | \end{pmatrix}.$$

Diese Matrix $[\varphi] = A$ heißt *die Matrix von φ bzgl. der Standardbasen* in $\mathbb{K}^n$ bzw. $\mathbb{K}^m$.

Es gilt $\varphi(x) = Ax$ für alle $x \in \mathbb{K}^n$, d.h.

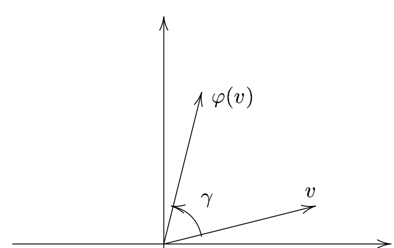
$$\varphi = \varphi_A.$$

In der Tat gilt $Ae_j = (j\text{-te Spalte von } A) = \varphi(e_j)$ für alle $j = 1, \ldots, n$, also auch $\varphi_A(e_j) = \varphi(e_j)$ für alle $j = 1 \ldots, n$. Stimmen zwei lineare Abbildungen auf einer Basis überein, dann haben sie auch gleiche Werte für alle x nach 4.2.8 (1).

Wir können also jede lineare Abbildung $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ bezüglich der Standardbasen durch eine eindeutig bestimmte $m \times n$-Matrix $[\varphi] = A$ darstellen.

4.3.3 Beispiel: Rotationen in der Ebene.

Sei φ die Drehung des Euklidischen $\mathbb{R}^2$ um den Ursprung 0 um den Winkel γ. Wir wollen diese lineare Abbildung durch ihre Matrix bezüglich der kanonischen Basis darstellen:

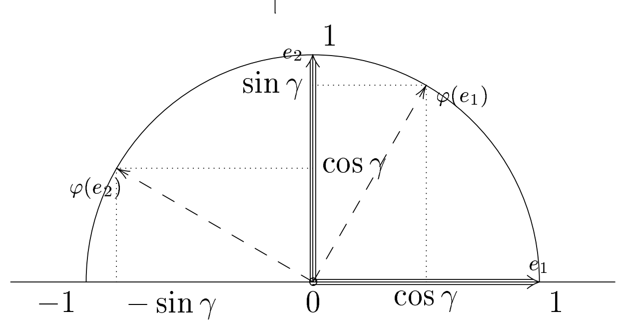


Wir bestimmen

$$\varphi(e_1) = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \varphi(e_2) = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}.$$

Damit kann φ durch folgende *Drehmatrix* dargestellt werden:

$$D = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$$



Im Spezialfall $\gamma := \frac{\pi}{3} = 60°$ gilt $\cos\frac{\pi}{3} = \frac{1}{2}$, $\sin\frac{\pi}{3} = \frac{1}{2}\sqrt{3}$. Deshalb ist die Drehmatrix um 60°

$$D = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$$

Das Bild von $a = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ nach der Rotation um 60° ist

$$\varphi(a) = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-\sqrt{3}\\1+\sqrt{3}\end{pmatrix}.$$

4.3.4 Das Bild von φ. Sei $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung und A die Matrix von φ (bezüglich der Standardbasen). Wir wollen Eigenschaften von φ durch Eigenschaften seiner Matrix A beschreiben.

Aus 4.2.8 (2) wissen wir, dass $\operatorname{im}\varphi$ von den Bildern $Ae_1, \ldots, Ae_1$ der Standardbasisvektoren erzeugt wird. Aus voriger Beobachtung können wir folgern:

Proposition. *Das Bild von φ ist der lineare Teilraum von $\mathbb{K}^m$, der von den Spaltenvektoren $w_1, w_2, \ldots w_n$ von A aufgespannt wird.*

Wir erinnern uns, dass der Rang einer linearen Abbildung φ als die Dimension von $\operatorname{im}\varphi$ definiert ist. Also ist der Rang von φ die Dimension des linearen Teilraumes, der von den Spaltenvektoren $w_1, w_2, \ldots, w_n$ der Mtrix A von φ aufgespannt wird.

4.3.5 Der Rang einer Matrix $A \in \mathcal{M}(m \times n, \mathbb{K})$. Wir definieren den *Rang einer Matrix* A als die Dimension des linearen Teilraums von $\mathbb{K}^m$, der von den Spaltenvektoren $w_1, w_2, \ldots, w_n$ von A aufgespannt wird, d.h.

$$\operatorname{rank} A := \dim(\operatorname{span}(w_1, w_2, \ldots, w_n)).$$

Aufgrund der vorigen Bemerkung gilt für die Matrix A einer linearen Abbildung φ

$$\operatorname{rank} A = \operatorname{rank}\varphi$$

Eine praktische Methode zur Bestimmung einer Basis des Bildes einer linearen Abbildung φ sowie ihres Rangs r besteht darin, die Matrix A von φ (bezüglich der Standardbasen) zu bestimmen und den Gauß-Jordan Eliminations-Algorithmus auf die Spaltenvektoren $w_1, \ldots, w_n$ von A (siehe 2.5.5) anzuwenden.

Beispiel. Sei

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & -4 \end{pmatrix}$$

Wir verwenden den Gauß-Jordan Eliminations-Algorithmus, siehe Abbildung 4.1. Die Rechnung ergibt: $\operatorname{rank} A = 2$. Die Vektoren $\left(\begin{smallmatrix}1\\2\\-1\end{smallmatrix}\right)$ und $\left(\begin{smallmatrix}0\\1\\2\end{smallmatrix}\right)$ bilden eine Basis des linearen Teilraums von $\mathbb{R}^3$, der von den Spalten von A aufgespannt wird. Also ist $\operatorname{rank}\varphi_A = 2$; das Bild von φ_A hat die beiden obigen Vektoren als Basis.

1	2	−1	w_1
0	1	2	w_2
2	3	−4	w_3
1	2	−1	
0	1	2	
0	−1	−2	
1	2	−1	
0	1	2	
0	0	0	

Abbildung 4.1. Der Gauß-Jordan Eliminations-Algorithmus für die Matrix in 4.3.5

4.3.6 Spaltenrang und Zeilenrang. Sei $A \in \mathcal{M}(m \times n, \mathbb{K})$. In 4.3.5 wurde der Rang von A als die Dimension des linearen Teilraums von $\mathbb{K}^m$ definiert, der von den Spalten $w_1, \ldots, w_n$ von A aufgespannt wird. Dieser Rang wird auch der *Spaltenrang* von A genannt.

Wir können auch die Zeilen $z_1, \ldots, z_m \in \mathbb{K}^n$ von A betrachten. Die Dimension des linearen Teilraums von $\mathbb{K}^n$, der von den Zeilenvektoren $z_1, \ldots, z_m$ aufgespannt wird, heißt auch der *Zeilenrang* von A.

Der Zeilenrang kann bestimmt werden, indem wir den Gauß-Jordan Eliminations-Algorithmus auf die Zeilenvektoren von A anwenden.

Beispiel. Sei

$$A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 0 & 1 & 3 \\ -1 & 2 & 0 & -4 \end{pmatrix} \in \mathcal{M}(3 \times 4, \mathbb{R})$$

Wir berechnen den Zeilenrang im linken Schema und den Spaltenrang im rechten Schema von Abbildung 4.2.

Wir erkennen in diesem Beispiel, dass der Spaltenrang gleich dem Zeilenrang ist, obwohl die Zeilen Vektoren des $\mathbb{R}^4$, die Spalten aber Vektoren des $\mathbb{R}^3$ sind. Dies gilt auch allgemein:

Theorem. *Für jede Matrix A ist der Spaltenrang gleich dem Zeilenrang.*

Dieser Satz ist keineswegs offensichtlich. Wir verschieben den Beweis auf später. Aufgrund dieses Satzes müssen wir nicht Zeilen- und Spaltenrang unterscheiden und können einfach vom Rang der Matrix A sprechen.

4.3.7 Charakterisierung des Rangs einer Matrix A. *Der Rang der Matrix A ist sowohl gleich der maximalen Anzahl linear unabhängiger Spalten von A als auch gleich der maximalen Anzahl linear unabhängiger Zeilen von A.*

Dies folgt aus 2.6.4 und dem vorigen Satz. Vergleiche auch 4.2.8 (3). □

$$\begin{array}{rrrrl} 1 & 2 & 1 & -1 & z_1 \\ 2 & 0 & 1 & 3 & z_2 \\ -1 & 2 & 0 & -4 & z_3 \\ \hline 1 & 2 & 1 & -1 & \\ 0 & -4 & -1 & 5 & \\ 0 & 4 & 1 & -5 & \\ \hline 1 & 2 & 1 & -1 & \\ 0 & -4 & -1 & 5 & \\ 0 & 0 & 0 & 0 & \\ \hline \end{array} \qquad \begin{array}{rrrl} 1 & 2 & -1 & w_1 \\ 2 & 0 & 2 & w_2 \\ 1 & 1 & 0 & w_3 \\ -1 & 3 & -4 & w_4 \\ \hline 1 & 2 & -1 & \\ 0 & -4 & 4 & \\ 0 & -1 & 1 & \\ 0 & 5 & -5 & \\ \hline 1 & 2 & -1 & \\ 0 & -4 & 4 & \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \\ \hline \end{array}$$

Zeilenrang von $A = 2$ Spaltenrang von $A = 2$

Abbildung 4.2. Der Spalten- und Zeilenrang der Matrix A in 4.3.6

4.3.8 Injektivität, Surjektivität, Bijektivität. Ist $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ linear und $A \in \mathcal{M}(m \times n, \mathbb{K})$ die Matrix von φ (bezüglich der Standardbasen), dann gilt:

(1) φ ist surjektiv $\iff$ $\operatorname{rank} A = m$
$\iff$ die Zeilen von A sind linear unabhängig.

(Die erste Äquivalenz folgt aus 4.2.8 (4) und aus der Tatsache, dass $\operatorname{rank} A = \operatorname{rank} \varphi_A$ ist. Die zweite Äquivalenz folgt aus dem Satz in 4.3.6, der besagt, dass der Spaltenrang gleich dem Zeilenrang ist. Da es nur m Zeilen in A gibt, sind diese genau dann linear unabhängig, wenn $\operatorname{rank} A = m$ gilt.)

(2) φ ist injektiv $\iff$ $\operatorname{rank} A = n$
$\iff$ die Spalten von A sind linear unabhängig.

(Die zweite Äquivalenz folgt aus 4.2.8 (5), die erste folgt aus der zweiten.)

(3) φ ist bijektiv, d.h. φ ist ein Isomorphismus genau dann, wenn $\operatorname{rank} A = n = m$.

(4) $n = m$: Ist φ eine lineare Abbildung von $\mathbb{K}^n$ in sich, dann ist ihre Matrix A eine $n \times n$-Matrix, und es gilt:

φ ist bijektiv $\iff$ $\operatorname{rank} A = n$
$\iff$ die Spalten von A sind linear unabhängig
$\iff$ die Zeilen von A sind linear unabhängig

4.3.9 Satz (Kriterien für Invertierbarkeit). *Für eine $n \times n$-Matrix A gilt:*

A ist invertierbar $\iff$ $\operatorname{rank} A = n$
$\iff$ *die Spalten von A sind linear unabhängig*
$\iff$ *die Zeilen von A sind linear unabhängig*
$\iff$ *es gibt eine $n \times n$-Matrix B mit $AB = E$*
$\iff$ *es gibt eine $n \times n$-Matrix B mit $BA = E$*

In den letzten beiden Fällen ist jeweils $B = A^{-1}$.

Beweis. Aus der Invertierbarkeit von A folgen natürlich die beiden letzten Aussagen auf der rechten Seite. Die Äquivalenz der ersten vier Aussagen folgt aus 4.3.8 (4). Aus der Invertierbarkeit von A folgen natürlich die beiden letzten Aussagen auf der rechten Seite. Nehmen wir nun an, für zwei $n \times n$-Matrizen A und B gilt $AB = E$. Für alle $v \in \mathbb{K}^n$ gilt dann $\varphi_A(\varphi_B(v)) = A(Bv) = (AB)v = Ev = v$. Wir erkennen daraus, daß φ_A surjektiv sein muß; denn jedes $v \in \mathbb{K}^n$ kommt im Bild von φ_A vor. Nach 4.3.8 (1) ist dies äquivalent dazu, daß die Spalten der Matrix A linear unabhängig sind. Aus der Äquivalenz der vorhergehenden Eigenschaften folgt nun, daß A invertierbar ist. Multiplikation der Gleichung $AB = E$ mit A^{-1} von links ergibt $B = A^{-1}$, und damit ist B ebenfalls invertierbar. Den Fall $BA = E$ kann man also auf den vorhergehenden Fall zurückführen. □

Bemerkung. Gilt $ab = e$ in einer Halbgruppe mit Neutralelement e, so braucht im allgemeinen a nicht invertierbar zu sein. Für die Invertierbarkeit von a hatten wir die Existenz eines beidseitigen Inversen gefordert. Für $n \times n$-Matrizen ist die Existenz eines einseitigen Inversen schon ausreichend für Invertierbarkeit.

4.3.10 Der Kern von $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$. Sei A die Matrix von φ (bezüglich der Standardbasen). Dann ist $\ker\varphi = \{x \in \mathbb{K}^n \mid \varphi(x) = 0\} = \{x \in \mathbb{K}^n \mid Ax = 0\}$. Also ist der Kern von φ die Menge aller Lösungen x des homogenen linearen Gleichungssystems

$$\text{(H)} \qquad Ax = 0.$$

Nach 4.2.4 ist $\ker\varphi$ ein linearer Teilraum von $\mathbb{K}^n$. Also ist die Menge aller Lösungen von (H) ein linearer Teilraum von $\mathbb{K}^n$ der Dimension $\dim(\ker\varphi) = n - \operatorname{rank}\varphi = n - \operatorname{rank} A$ nach 4.2.7. Sei $k := n - \operatorname{rank} A$. Dann besteht jede Basis von $\ker\varphi$ aus k Vektoren $v_1, \ldots, v_k$ von $\ker\varphi$. Diese werden auch ein *Fundamentalsystem* von Lösungen von (H) genannt. Die allgemeine Lösung von (H) ist dann durch die sämtlichen Linearkombinationen

$$x_H = \lambda_1 v_1 + \ldots + \lambda_k v_k$$

mit $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ gegeben. Damit ist das Problem der Lösung eines homogenen Systems linearer Gleichungen auf das Auffinden von k linear unabhängigen Lösungen $v_1, \ldots, v_k$ mit $k = n - \operatorname{rank} A$ zurückgeführt.

4.4. Die Matrix einer linearen Abbildung bzgl. beliebiger Basen

In diesem Abschnitt sei
V ein $\mathbb{K}$-Vektorraum der Dimension n mit einer Basis $\mathcal{B}\colon v_1, v_2, \ldots, v_n$ und
W ein $\mathbb{K}$-Vektorraum der Dimension m mit einer Basis $\mathcal{C}\colon w_1, w_2, \ldots, w_m$.

4.4.1 Wiederholung: Koordinaten. Wir wissen aus 2.5.8, dass jedes $v \in V$ eine Darstellung als Linearkombination

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \quad \text{mit } \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$$

besitzt. Die Skalare $\lambda_1, \lambda_2, \dots, \lambda_n$ sind eindeutig bestimmt und werden die *Koordinaten* von v bzgl. der Basis $\mathcal{B}$ genannt. Wir schreiben

$$[v]_{\mathcal{B}} := \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$$

für den Vektor der Koordinaten von v bzgl. der Basis $\mathcal{B}$. Die Abbildung

$$\begin{aligned} \kappa_{\mathcal{B}} \colon V &\to \mathbb{K}^n \\ v &\mapsto [v]_{\mathcal{B}}, \end{aligned}$$

die jedem $v \in V$ den Koordinatenvektor $[v]_{\mathcal{B}}$ zuordnet, ist bijektiv und linear und damit ein Isomorphismus zwischen Vektorräumen, wie wir bereits in 4.2.11 gesehen haben. Für die Koordinaten der Basisvektoren v_j gilt:

$$[v_j]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_j.$$

4.4.2 Die Matrix einer linearen Abbildung $\varphi \colon V \to W$. Wir betrachten die Bilder der Basisvektoren $v_1, \dots, v_n$ und bestimmen die Koordinaten von $\varphi(v_j)$ bzgl. der Basis $\mathcal{C} \colon w_1, \dots, w_m$ von W:

$$\varphi(v_j) = \alpha_{1j} w_1 + \dots + \alpha_{mj} w_m.$$

Wir bilden nun eine $m \times n$-Matrix, indem wir die Koordinaten von $\varphi(v_j)$ bzgl. der Basis $\mathcal{C}$ in die j-te Spalte schreiben:

$$A = [\varphi]^{\mathcal{B}}_{\mathcal{C}} := \begin{array}{c} \begin{array}{ccc} [\varphi(v_1)]_{\mathcal{C}} & \qquad\qquad & [\varphi(v_n)]_{\mathcal{C}} \end{array} \\ \begin{pmatrix} \alpha_{11} & \cdots\cdots & \alpha_{1n} \\ \alpha_{21} & \cdots\cdots & \alpha_{2n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots\cdots & \alpha_{mn} \end{pmatrix} \end{array}$$

Wir nennen $A = [\varphi]^{\mathcal{B}}_{\mathcal{C}}$ die *Matrix von φ bzgl. der Basen $\mathcal{B}$* und $\mathcal{C}$.

Grobe Merkregel: In den Spalten der Matrix A von φ stehen die Koordinaten der Bilder der Basisvektoren.

Behauptung. *$[\varphi]^{\mathcal{B}}_{\mathcal{C}}[v]_{\mathcal{B}} = [\varphi(v)]_{\mathcal{C}}$ für alle $v \in V$.*

Die Matrix A stellt die lineare Abbildung φ in dem Sinn dar, daß man die Koordinaten von $\varphi(v)$ erhält, indem man die Matrix A auf den Koordinatenvektor von v anwendet. In anderen Worten: Ist $A = [\varphi]^{\mathcal{B}}_{\mathcal{C}}$ die Matrix von φ bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$, sind

$[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die Koordinaten des Vektors $v \in V$, und $[\varphi(v)]_{\mathcal{C}} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ die Koordinaten des Bildvektors $\varphi(v)$ dann gilt::

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Das folgende Diagramm faßt diese Aussagen zusammen:

$$\begin{tikzcd}
v \arrow[r, mapsto] \arrow[d, mapsto] & \varphi(v) \arrow[d, mapsto] \\
{[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}} \arrow[r, mapsto] & A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}
\end{tikzcd}$$

$$\begin{tikzcd}
V \arrow[r, "\varphi"] \arrow[d, "\kappa_{\mathcal{B}}"'] & W \arrow[d, "\kappa_{\mathcal{C}}"] \\
\mathbb{K}^n \arrow[r, "{A=[\varphi]^{\mathcal{B}}_{\mathcal{C}}}"'] & \mathbb{K}^m
\end{tikzcd}$$

Beweis der Behauptung. Da die Abbildungen $v \to A[v]_{\mathcal{B}}$ und $v \to [\varphi(v)]_{\mathcal{C}}$ beide linear sind, genügt es nach 4.2.8(1) zu zeigen, daß sie für die Basisvektoren $v_1, \ldots, v_n$ übereinstimmen. Da $[v_j]_{\mathcal{B}} = e_j$ ist, gilt $A[v_j]_{\mathcal{B}} = Ae_j = (j\text{-te Spalte von } A) = [\varphi(v_j)]_{\mathcal{C}}$, woraus die Behauptung folgt. □

4.4.3 Spezialfall: Die Matrix eines Endomorphismus. Für einen Endomorphismus $\varphi\colon V \to V$ wählt man üblicherweise nur eine Basis $\mathcal{B} = \mathcal{C}$ in V.

Die Matrix $A = [\varphi]^{\mathcal{B}}_{\mathcal{B}}$ wird *die Matrix von φ bzgl. der Basis $\mathcal{B}$* genannt. Dies ist eine quadratische $n \times n$-Matrix. Nach der Vorgehensweise in 4.4.2 wird diese auf folgende Art gebildet: Man stellt die Bilder $\varphi(v_j)$ der Basisvektoren $v_1, \ldots, v_n$ durch ihre Koordinaten bzgl. derselben Basis dar:

$$\varphi(v_j) = \alpha_{1j} v_1 + \ldots + \alpha_{nj} v_n.$$

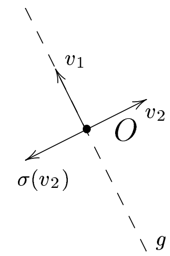
Die Koeffizienten $\alpha_{1j}, \ldots, \alpha_{nj}$ bilden dabei die j-te Spalte von A. Wir schreiben anstelle von $[\varphi]^{\mathcal{B}}_{\mathcal{B}}$ einfach $[\varphi]_{\mathcal{B}}$.

Beispiel. Die Matrix der identischen Abbildung $\mathrm{id}_V\colon V \to V$ bezüglich einer beliebigen Basis ist die Einheitsmatix:

$$[\mathrm{id}_V]_{\mathcal{B}} = E$$

4.4.4 Beispiele.

(a) Sei $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ die (orthogonale) Spiegelung an einer Geraden g durch 0. Um σ durch eine Matrix darzustellen, wählen wir eine Basis v_1, v_2 wie folgt: Sei v ein Vektor, der in die Richtung der Geraden g zeigt und v_2 irgendein Vektor, der senkrecht auf g steht. Es folgt

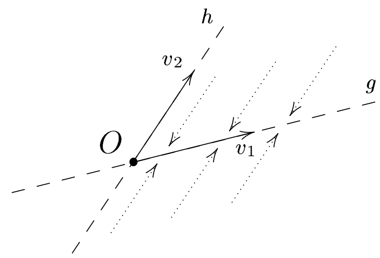


$$\sigma(v_1) = v_1 = 1 \cdot v_1 + 0v_2$$
$$\sigma(v_2) = -v_2 = 0 \cdot v_1 + (-1)v_2$$

Also hat die Matrix von σ bzgl. der Basis v_1, v_2 die Gestalt

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

(b) Seien g und h zwei Geraden im $\mathbb{R}^2$, die sich in O schneiden. Sei $\pi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Parallelprojektion auf die Gerade g in Richtung h. Um π durch eine Matrix darzustellen, wählen wir eine geeignete Basis $\mathcal{B}$: v_1, v_2. Sei v_1 ein Vektor, der die Richtung von g hat, v_2 ein Vektor in Richtung von h. Dann gilt



$$\pi(v_1) = v_1 = 1 \cdot v_1 + 0 \cdot v_2$$
$$\pi(v_2) = 0 = 0 \cdot v_1 + 0 \cdot v_2$$

Die Matrix von π bzgl. der Basis $\mathcal{B}$ ist

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

Wir stellen fest: die Matrix einer linearen Abbildung kann einfach werden, wenn wir eine Basis wählen, die der speziellen Situation angepaßt ist.

(c) Seien V und W endlichdimensionale Vektorräume und $\varphi\colon V \to W$ eine lineare Abbildung. Wir behaupten: Es gibt eine Basis $\mathcal{B}$ in V und eine Basis $\mathcal{C}$ in W, so dass die Matrix von φ bzgl. $\mathcal{B}$ und $\mathcal{C}$ die folgende Darstellung besitzt:

$$A = [\varphi]^{\mathcal{B}}_{\mathcal{C}} = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots \cdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & \vdots & & \vdots \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \cdots \cdots & 0 \\ \vdots & & & & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots \cdots & 0 \end{array}\right)$$

Die „Größe“ der Einheitsmatrix in der linken oberen Ecke ist $r = \operatorname{rank} \varphi$.

Beweisskizze. Wie im Beweis der Dimensionsformel 4.2.6 wählen wir eine Basis $v_{r+1}, \dots, v_n$ von $\ker \varphi$. Nach 2.6.6 (Basisergänzungssatz) gibt es Vektoren $v_1, \dots, v_r$, so dass $v_1, \dots, v_r, v_{r+1}, \dots, v_n$ eine Basis $\mathcal{B}$ von V ist. Wie im Beweis von 4.2.6 erkennen wir, dass $w_1 := \varphi(v_1), \dots w_r := \varphi(v_r)$ eine Basis von

$\operatorname{im}\varphi \subseteq W$ ist. Nach dem Basisergänzungssatz gibt es Vektoren $w_{r+1}, \dots, w_m$, so dass $w_1, \dots, w_r, w_{r+1}, \dots, w_m$ eine Basis von W ist. Da $\varphi(v_1) = w_1, \dots, \varphi(v_r) = w_r$, $\varphi(v_{r+1}) = 0, \dots, \varphi(v_n) = 0$ gilt, hat die Matrix von φ bzgl. $\mathcal{B}$ und $\mathcal{C}$ obige Form.

Jeder linearen Abbildung $\varphi\colon V \to W$ haben wir ihre Matrix $[\varphi]_{\mathcal{C}}^{\mathcal{B}}$ bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$ zugeordnet: Dies führt auf eine Abbildung

$$\begin{aligned} \operatorname{Hom}(V,W) &\to \mathcal{M}(m \times n, \mathbb{K}) \\ \varphi &\mapsto [\varphi]_{\mathcal{C}}^{\mathcal{B}}. \end{aligned}$$

Sind φ und ψ zwei verschiedene lineare Abbildungen von V nach W, dann können sie nach 4.2.8(1) nicht auf allen Basisvektoren v_j übereinstimmen. Wenn z.B. $\varphi(v_j) \neq \psi(v_j)$, dann unterscheiden sich die Matrizen von φ und ψ in der j-ten Spalte. Die beiden Matrizen sind also verschieden. Umgekehrt ist jede $m \times n$-Matrix $A = (\alpha_{ij})$ die Matrix einer linearen Abbildung, und zwar derjenigen linearen Abbildung mit $\varphi(v_j) = \alpha_{1j}w_1 + \dots + \alpha_{mj}w_m$ für alle $j = 1, \dots, n$. Damit haben wir die folgende Proposition bewiesen:

4.4.5 Proposition. *Die Abbildung* $\varphi \mapsto [\varphi]_{\mathcal{C}}^{\mathcal{B}}\colon \operatorname{Hom}(V,W) \to \mathcal{M}(m \times n, \mathbb{K})$ *ist bijektiv.*

4.5. Operieren mit linearen Abbildungen

In diesem Abschnitt seien U, V, W zunächst beliebige Vektorräume über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Wir erinnern uns daran, dass $\operatorname{Hom}(V,W)$ die Menge aller linearen Abbildungen von V nach W ist.

4.5.1 $\operatorname{Hom}(V,W)$ **als Vektorraum.** Für $\varphi, \psi \in \operatorname{Hom}(V,W)$ und $\lambda \in \mathbb{K}$ definieren wir Abbildungen $\varphi + \psi$ und $\lambda\varphi$ von V nach W durch

$$\begin{aligned} (\varphi + \psi)(v) &:= \varphi(v) + \psi(v) \\ (\lambda\varphi)(v) &:= \lambda \cdot \varphi(v) \end{aligned}$$

für alle $v \in V$.

Behauptung. $\varphi + \psi$ *und* $\lambda\varphi$ *sind linear und damit Elemente von* $\operatorname{Hom}(V,W)$.

Beweis. Die Abbildung $\varphi + \psi$ ist linear wegen

$$\begin{aligned} (\varphi + \psi)(u + v) &= \varphi(u + v) + \psi(u + v) && \text{nach Definition von } \varphi + \psi \\ &= \varphi(u) + \varphi(v) + \psi(u) + \psi(v) && \text{nach (L1)} \\ &= \varphi(u) + \psi(u) + \varphi(v) + \psi(v) \\ &= (\varphi + \psi)(u) + (\varphi + \psi)(v) && \text{nach Definition von } \varphi + \psi \end{aligned}$$

und

$$\begin{aligned} (\varphi + \psi)(\mu v) &= \varphi(\mu v) + \psi(\mu v) && \text{nach Definition von } \varphi + \psi \\ &= \mu \cdot \varphi(v) + \mu \cdot \psi(v) && \text{nach (L2)} \\ &= \mu \cdot \big(\varphi(v) + \psi(v)\big) \end{aligned}$$

$$= \mu \cdot (\varphi + \psi)(v) \qquad \text{nach Definition von } \varphi + \psi$$

Auf ähnliche Weise zeigt man die Linearität von $\lambda\varphi$. □

Damit haben wir eine Addition und eine skalare Multiplikation auf $\mathrm{Hom}(V, W)$ konstruiert. Man zeigt leicht, dass die Vektorraumaxiome (V1)–(V8) erfüllt sind (Übung). Also gilt folgendes

Theorem. *Mit der oben definierten Addition und skalaren Multiplikation ist die Menge* $\mathrm{Hom}(V, W)$ *aller linearen Abbildungen* $\varphi\colon V \to W$ *ein* $\mathbb{K}$*-Vektorraum.*

4.5.2 Die Hintereinanderausführung von linearen Abbildungen.

(1) *Sind* $\psi\colon U \to V$ *und* $\varphi\colon V \to W$ *linear, dann ist* $\varphi \circ \psi\colon U \to W$ *ebenfalls linear.*

(2) *Die identische Abbildung* $\mathrm{id}_V\colon V \to V$ *ist linear.*

(3) *Ist* $\varphi\colon V \to W$ *ein Isomorphismus, dann ist die inverse Abbildung* $\varphi^{-1}\colon W \to V$ *ebenfalls ein Isomorphismus.*

Beweis. 1. Es gilt:

$$\begin{aligned}
(\varphi \circ \psi)(u_1 + u_2) &= \varphi\big(\psi(u_1 + u_2)\big) \\
&= \varphi\big(\psi(u_1) + \psi(u_2)\big) && \text{wegen der Linearität von } \psi \\
&= \varphi\big(\psi(u_1)\big) + \varphi\big(\psi(u_2)\big) && \text{wegen der Linearität von } \varphi \\
&= (\varphi \circ \psi)(u_1) + (\varphi \circ \psi)(u_2) \\
(\varphi \circ \psi)(\lambda u) &= \varphi\big(\psi(\lambda u)\big) \\
&= \varphi\big(\lambda \cdot \psi(u)\big) && \text{wegen der Linearität von } \psi \\
&= \lambda \cdot \varphi\big(\psi(u)\big) && \text{wegen der Linearität von } \varphi \\
&= \lambda \cdot (\varphi \circ \psi)(u)
\end{aligned}$$

2. Die Linearität der identischen Abbildung ist offensichtlich.

3. Wir haben nur zu zeigen, dass φ^{-1} linear ist. Dazu sei $w_1, w_2 \in W$. Seien $v_1 = \varphi^{-1}(w_1)$ und $v_2 = \varphi^{-1}(w_2)$. Dann gilt $\varphi(v_1) = w_1$ und $\varphi(v_2) = w_2$. Also folgt

$$\begin{aligned}
\varphi^{-1}(w_1 + w_2) &= \varphi^{-1}\big(\varphi(v_1) + \varphi(v_2)\big) \\
&= \varphi^{-1}\big(\varphi(v_1 + v_2)\big) && \text{wegen der Linearität von } \varphi \\
&= v_1 + v_2 \\
&= \varphi^{-1}(w_1) + \varphi^{-1}(w_2) \\
\varphi^{-1}(\lambda w_1) &= \varphi^{-1}\big(\lambda\varphi(v_1)\big) \\
&= \varphi^{-1}\big(\varphi(\lambda v_1)\big) && \text{wegen der Linearität von } \varphi \\
&= \lambda v_1 \\
&= \lambda\varphi^{-1}(w_1).
\end{aligned}$$

□

4.5.3 Regeln für die Hintereinanderausführung. *Für beliebige lineare Abbildungen $\psi, \psi' \colon U \to V$ und $\varphi, \varphi' \colon V \to W$ und alle $\lambda \in \mathbb{K}$ gilt:*

(A1) *Die Hintereinanderausführung von linearen Abbildungen ist assoziativ.*

(A2) $$\mathrm{id}_W \circ \varphi = \varphi = \varphi \circ \mathrm{id}_V$$

(A3) $$\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi' \quad , \quad (\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi$$

(A4) $$(\lambda\varphi) \circ \psi = \lambda \cdot (\varphi \circ \psi) = \varphi \circ (\lambda\psi)$$

Beweis. Die Regeln (A1) und (A2) gelten für beliebige Abbildungen, damit auch für lineare Abbildungen.

(A3): Für alle $u \in U$ gilt

$$\begin{aligned}
\big(\varphi \circ (\psi + \psi')\big)(u) &= \varphi\big((\psi + \psi')(u)\big) && \text{nach Definition von } \circ \\
&= \varphi\big(\psi(u) + \psi'(u)\big) && \text{nach Definition von } + \\
&= \varphi\big(\psi(u)\big) + \varphi\big(\psi'(u)\big) && \text{wegen der Linearität von } \varphi \\
&= (\varphi \circ \psi)(u) + (\varphi \circ \psi')(u) \\
&= (\varphi \circ \psi + \varphi \circ \psi')(u),
\end{aligned}$$

also $\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi'$.

Auf ähnliche Weise zeigt man den zweiten Teil von (A3) und (A4). □

4.5.4 $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ als $\mathbb{K}$-Algebra. Mit $V = W$ in 4.5.1 erkennen wir, dass $\mathrm{End}(V) = \mathrm{Hom}(V, V)$, also die Menge aller linearen Abbildungen $\varphi \colon V \to V$, ein $\mathbb{K}$-Vektorraum mit der hierfür definierten Addition und skalaren Multiplikation ist. Da die Hintereinanderausführung von linearen Abbildungen nach 4.5.2 (1) linear ist, haben wir eine weitere Verknüpfung $\circ$ auf $\mathrm{End}(V)$:

$$\begin{aligned}
\mathrm{End}(V) \times \mathrm{End}(V) &\to \mathrm{End}(V) \\
(\varphi, \psi) &\mapsto \varphi \circ \psi
\end{aligned}$$

Diese Verknüpfung kann als eine Multiplikation angesehen werden. Nach 4.5.3 hat diese nämlich die folgenden Eigenschaften für alle $\varphi, \varphi', \psi, \psi', \eta \in \mathrm{End}(V)$:

(A1) $$\varphi \circ (\psi \circ \eta) = (\varphi \circ \psi) \circ \eta$$

(A2) $$\mathrm{id}_V \circ \varphi = \varphi = \varphi \circ \mathrm{id}_V$$

(A3) $$\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi' \quad , \quad (\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi$$

(A4) $$(\lambda \cdot \varphi) \circ \psi = \lambda \cdot (\varphi \circ \psi) = \varphi \circ (\lambda\psi)$$

Da $\mathrm{id}_V \in \mathrm{End}(V)$ nach 4.5.2 (2) gilt, erkennen wir wegen (A1) und (A2), dass $\circ$ eine assoziative Verknüpfung auf $\mathrm{End}(V)$ mit id_V als neutralem Element ist. Nach (A3) ist $\circ$ distributiv bzgl. der Addition. Also ist $\mathrm{End}(V)$ ein Ring bzgl. $+$ und $\circ$. Gilt für die Addition (A4), dann nennen wir $\mathrm{End}(V)$ eine $\mathbb{K}$-Algebra in folgendem Sinn:

Definition. Eine $\mathbb{K}$-*Algebra* ist ein $\mathbb{K}$-Vektorraum zusammen mit einer zusätzlichen Operation $*$, die assoziativ ist, ein neutrales Element besitzt, distributiv bzgl. $+$ ist und die Eigenschaft (A4) erfüllt.

Man beachte die Analogie der obigen Regeln mit den Regeln für das Produkt von Matrizen (siehe 4.1.4).

4.5.5 Die Gruppe $\mathrm{Aut}(V) = \mathrm{GL}(V)$**.** Wir erinnern uns, dass $\mathrm{Aut}(V)$ die Menge aller Automorphismen von V, d.h. die Menge aller bijektiven Homomorphismen $\varphi\colon V \to V$ bezeichnet. ($\mathrm{GL}(V)$ ist eine andere Schreibweise für diese Menge, die sich aus dem Französischen „groupe linéaire“ herleitet.)

Für $\varphi \in \mathrm{Aut}(V)$ ist die Umkehrabbildung φ^{-1} auch linear nach 4.5.2 (3) und damit ein Automorphismus. Also ist $\mathrm{Aut}(V)$ die Menge der invertierbaren Elemente von $\mathrm{End}(V)$ bzgl. der Operation $\circ$. Damit folgt aus Satz 2.1.7, dass $\mathrm{Aut}(V)$ eine Gruppe bzgl. $\circ$ mit id_V als neutralem Element ist.

Nun nehmen wir an, die Vektorräume U, V und W seien endlichdimensional und
$\mathcal{A} : u_1, \dots, u_p$ sei eine Basis von U,
$\mathcal{B} : v_1, \dots, v_n$ sei eine Basis von V und
$\mathcal{C} : w_1, \dots, w_m$ sei eine Basis von W.

Jeder linearen Abbildung $\varphi\colon V \to W$ hatten wir ihre Matrix $[\varphi]^{\mathcal{B}}_{\mathcal{C}}$ bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$ zugeordnet: Dies führt nach **??** zu einer Bijektion

$$\begin{aligned} \mathrm{Hom}(V, W) &\to \mathcal{M}(m \times n, \mathbb{K}) \\ \varphi &\mapsto [\varphi]^{\mathcal{B}}_{\mathcal{C}}. \end{aligned}$$

4.5.6 Proposition. *Für alle* $\varphi, \psi \in \mathrm{Hom}(V, W)$ *und* $\lambda \in \mathbb{K}$ *gilt:*
Ist A die Matrix von φ und B die Matrix von ψ, dann ist $A + B$ die Matrix von $\varphi + \psi$ und λA die Matrix von $\lambda\varphi$, jeweils bezüglich den Basen $\mathcal{B}$ und $\mathcal{C}$; anders gesagt:

$$[\varphi + \psi]^{\mathcal{B}}_{\mathcal{C}} = [\varphi]^{\mathcal{B}}_{\mathcal{C}} + [\psi]^{\mathcal{B}}_{\mathcal{C}} \qquad\qquad [\lambda\varphi]^{\mathcal{B}}_{\mathcal{C}} = \lambda[\varphi]^{\mathcal{B}}_{\mathcal{C}}$$

Die Vektorräume $\mathrm{Hom}(V, W)$ *und* $\mathcal{M}(m \times n, \mathbb{K})$ *sind isomorph und*

$$\dim\big(\mathrm{Hom}(V, W)\big) = m \cdot n = \dim V \cdot \dim W.$$

Beweis. Mit den Bezeichnungen von oben gilt für alle $v \in V$:
$(A+B)[v]_{\mathcal{B}} = A[v]_{\mathcal{B}} + B[v]_{\mathcal{B}} = [\varphi(v)]_{\mathcal{C}} + [\psi(v)]_{\mathcal{C}} == [\varphi(v) + \psi(v)]_{\mathcal{C}} = [(\varphi+\psi)(v)]_{\mathcal{C}}$. Also ist $A + B$ die Matrix von $\varphi + \psi$.
$(\lambda A)[v]_{\mathcal{B}} = \lambda(A[v]_{\mathcal{B}}) = \lambda[\varphi(v)]_{\mathcal{C}} = [\lambda\varphi(v)]_{\mathcal{C}}$. Also ist λA die Matrix von $\lambda\varphi$. Damit ist die erste Eigenschaft bewiesen. Sie besagt, daß die bijektive Abbildung, die jeder linearen Abbildung $\varphi\colon V \to W$ ihre Matrix bezüglich der Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W zuordnet, linear ist, also ein Vektorraum-Isomorphismus. Damit ist auch die zweite Behauptung gezeigt. □

4.5.7 Proposition. *Betrachte lineare Abbildungen $\psi\colon U \to V$ und $\varphi\colon V \to W$. Dann gilt: Ist A die Matrix von φ bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$*
und ist B die Matrix von ψ bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$,
dann ist AB die Matrix von $\varphi \circ \psi$ bezüglich der Basen $\mathcal{A}$ und $\mathcal{C}$; kurz:

$$[\varphi \circ \psi]^{\mathcal{A}}_{\mathcal{C}} = [\varphi]^{\mathcal{B}}_{\mathcal{C}}[\psi]^{\mathcal{A}}_{\mathcal{B}}\,.$$

Diese Aussage ist in folgendem Diagramm zusammengefasst:

$$\begin{tikzcd} U_{\mathcal{A}} \arrow[r, "\psi"] \arrow[d, "\kappa_{\mathcal{A}}"'] & V_{\mathcal{B}} \arrow[r, "\varphi"] \arrow[d, "\kappa_{\mathcal{B}}"'] & W_{\mathcal{C}} \arrow[d, "\kappa_{\mathcal{C}}"] \\ \mathbb{K}^p \arrow[r, "B"'] & \mathbb{K}^n \arrow[r, "A"'] & \mathbb{K}^m \end{tikzcd}$$

Beweis. Für alle $u \in U$ gilt nach Voraussetzung: $AB[u]_{\mathcal{A}} = A[\psi(u)]_{\mathcal{B}} = [\varphi(\psi(u))]_{\mathcal{C}} = [(\varphi \circ \psi)(u)]_{\mathcal{C}}$. Also ist AB die Matrix von $\varphi \circ \psi$ bezüglich der Basen $\mathcal{A}$ und $\mathcal{C}$. □

Beispiel. Sei $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die lineare Abbildung, die man durch eine Drehung (um den Ursprung) um den Winkel γ und anschließender Spiegelung an der Geraden $y = -x$ erhält. Dann gilt

$$\varphi = \sigma \circ \delta,$$

wobei δ die Drehung und σ die Spiegelung ist. Man erhält die Matrix von φ (bzgl. der Standardbasis e_1, e_2) durch Multiplikation der Matrix von σ mit der Matrix von δ:

$$[\varphi] = [\sigma][\delta] = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} = \begin{pmatrix} -\sin\gamma & -\cos\gamma \\ -\cos\gamma & \sin\gamma \end{pmatrix}.$$

4.5.8 Proposition. *Sei $\varphi\colon V \to W$ eine lineare Abbildung und $A = [\varphi]^{\mathcal{B}}_{\mathcal{C}}$ ihre Matrix bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$. Genau dann ist φ bijektiv (und damit ein Isomorphismus), wenn $\dim V = \dim W$ und wenn die Matrix A invertierbar ist. In diesem Fall ist A^{-1} die Matrix der Umkehrabbildung $\varphi^{-1}\colon W \to V$ bezüglich der Basen $\mathcal{C}$ und $\mathcal{B}$:*

$$[\varphi^{-1}]^{\mathcal{C}}_{\mathcal{B}} = \left([\varphi]^{\mathcal{B}}_{\mathcal{C}}\right)^{-1}.$$

Beweis. Nehmen wir zunächst an, die lineare Abbildung $\varphi\colon V \to W$ sei bijektiv. Dann haben V und W dieselbe Dimension. Die Umkehrabbildung $\varphi^{-1}\colon W \to V$ ist ebenfalls linear und es gilt $\varphi \circ \varphi^{-1} = \mathrm{id}_W$ und $\varphi^{-1} \circ \varphi = \mathrm{id}_V$. Sei $A = [\varphi]^{\mathcal{B}}_{\mathcal{C}}$ die Matrix von φ bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$ und $B = [\varphi^{-1}]^{\mathcal{C}}_{\mathcal{B}}$ die Matrix von φ^{-1} bzgl. der Basen $\mathcal{C}$ und $\mathcal{B}$. Da die Matrix der identischen Abbildung die $n \times n$-Einheitsmatrix ist, folgen nach 4.5.7 die Gleichungen $AB = E$ und $BA = E$, d.h. A ist invertierbar und $B = A^{-1}$. □

Beispiel. Die Abbildung φ im Beispiel am Ende von 4.5.7 ist bijektiv. Die Matrix der Umkehrabbildung φ^{-1} ist

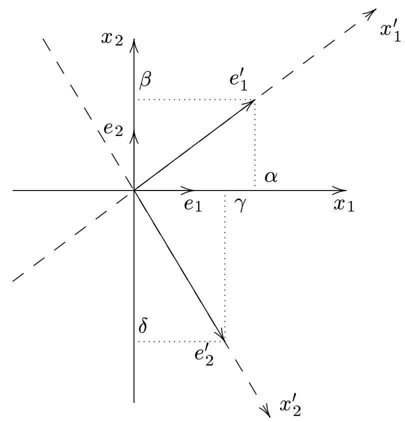
$$[\varphi^{-1}] = [\varphi]^{-1} = [\sigma \circ \delta]^{-1} = ([\sigma][\delta])^{-1} = [\delta]^{-1}[\sigma]^{-1} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

4.6. Basiswechsel

Wir betrachten in diesem Abschnitt zwei verschiedene Basen eines Vektorraumes und möchten untersuchen, wie die Koordinaten eines Vektors bzgl. verschiedener Basen miteinander in Beziehung stehen und wie sich die Matrizen einer linearen Abbildung bzgl. verschiedener Basen ausdrücken.

4.6.1 Koordinatentransformation in $\mathbb{R}^2$.

Betrachte im $\mathbb{R}^2$ die Standardbasis $\mathcal{B}$: e_1, e_2 und eine andere Basis $\mathcal{B}'$: e_1', e_2'.
Für ein beliebiges $v \in \mathbb{R}^2$ seien x_1, x_2 die Koordinaten bzgl. der Standardbasis $\mathcal{B}$:



(1) $$v = x_1 e_1 + x_2 e_2 \qquad [v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Analog seien x_1', x_2' die Koordinaten von v bzgl. der neuen Basis $\mathcal{B}'$:

(2) $$v = x_1' e_1' + x_2' e_2' \qquad [v]_{\mathcal{B}'} = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

Frage: Wie hängen die neuen Koordinaten x_1', x_2' von den Standardkoordinaten x_1, x_2 ab?

Dazu stellen wir die neuen Basisvektoren e_1', e_2' durch ihre Koordinaten bzgl. der Standardbasis dar.

(3) $$e_1' = \alpha e_1 + \beta e_2 \qquad [e_1']_{\mathcal{B}} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$e_2' = \gamma e_1 + \delta e_2 \qquad [e_2']_{\mathcal{B}} = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$$

Ersetzen wir diese Ausdrücke für e_1', e_2' in (2), so erhalten wir:

$$\begin{aligned} v &= x_1' e_1' + x_2' e_2' \\ &= x_1'(\alpha e_1 + \beta e_2) + x_2'(\gamma e_1 + \delta e_2) \\ &= (x_1'\alpha + x_2'\gamma) e_1 + (x_1'\beta + x_2'\delta) e_2. \end{aligned}$$

Wir vergleichen die Koeffizienten von e_1 und e_2 mit denen in (1):

(4) $$\begin{aligned} x_1 &= \alpha x_1' + \gamma x_2' \\ x_2 &= \beta x_1' + \delta x_2' \end{aligned}$$

Sei

$$S := \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

die Matrix der Koeffizienten. Dann kann dieses System linearer Gleichungen in folgender Form geschrieben werden:

(5) $$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} \qquad \text{oder} \qquad [v]_{\mathcal{B}} = S[v]_{\mathcal{B}'}.$$

Die Matrix S heißt *Transformationsmatrix*. Beachte, dass die Spalten von S die Koordinaten der neuen Basisvektoren e_1', e_2' bzgl. der Standardbasis $\mathcal{B}$ sind. Daraus folgt, dass die Spalten von S linear unabhängig sind. Also ist S nach 4.3.9 invertierbar. Multiplikation mit S^{-1} auf beiden Seiten ergibt folgende Formel:

(6) $$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

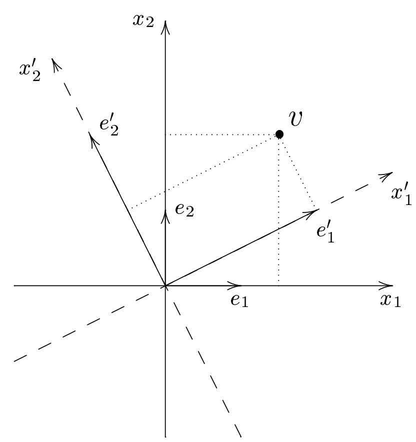
Beispiel. Betrachte die Standardbasis $\mathcal{B}$: $e_1 = \binom{1}{0}$, $e_2 = \binom{0}{1}$ und die neue Basis sei $\mathcal{B}'$: $e'_1 = \binom{2}{1}$, $e'_2 = \binom{-1}{2}$.

Dann ist die Transformationsmatrix:

$$S = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Die Inverse ist (Probe!):

$$S^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$



Es folgt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$
$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

d.h.

$$\begin{aligned} x_1 &= 2x'_1 - x'_2 & \qquad x'_1 &= \tfrac{1}{5}(2x_1 + x_2) \\ x_2 &= x'_1 + 2x'_2 & \qquad x'_2 &= \tfrac{1}{5}(-x_1 + 2x_2). \end{aligned}$$

4.6.2 Koordinatentransformation im allgemeinen Fall.

Seien

$$\begin{aligned} \mathcal{B}&: v_1, \ldots, v_n \\ \mathcal{B}'&: v'_1, \ldots, v'_n \end{aligned}$$

zwei Basen in einem n-dimensionalen Vektorraum V über einem Körper $\mathbb{K}$. Für ein beliebiges $v \in V$ seien $x_1, \ldots, x_n$ die Koordinaten bzgl. der Basis $\mathcal{B}$ und $x'_1, \ldots, x'_n$ die Koordinaten bzgl. $\mathcal{B}'$, d.h.

$$[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad [v]_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Die *Transformationsmatrix* $S := (\tau_{ij})$ wird auf folgende Art gebildet: Berechne die Koordinaten der neuen Basisvektoren v'_j bzgl. der alten Basis $\mathcal{B}$:

$$v'_j = \tau_{ij} v_1 + \ldots + \tau_{nj} v_n.$$

Diese Koordinaten schreibe in die j-te Spalte von S:

$$S = \begin{array}{c} \begin{array}{ccccc} [v'_1]_{\mathcal{B}} & & [v'_j]_{\mathcal{B}} & & [v'_n]_{\mathcal{B}} \end{array} \\ \begin{pmatrix} \tau_{11} & \cdots & \tau_{1j} & \cdots & \tau_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \tau_{n1} & \cdots & \tau_{nj} & \cdots & \tau_{nn} \end{pmatrix} \end{array}.$$

Da die $v_1',\dots,v_n'$ linear unabhängig sind, sind auch die Spalten von S linear unabhängig. Also ist S nach 4.3.9 invertierbar. Wir behaupten:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}, \qquad \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Kurz gesagt:

$$[v]_{\mathcal{B}} = S[v]_{\mathcal{B}'}\,, \qquad [v]_{\mathcal{B}'} = S^{-1}[v]_{\mathcal{B}}\,.$$

Dies kann genau wie in 4.6.1 gezeigt werden. Ein eleganter Beweis arbeitet wie folgt: Wir erkennen, dass S die Matrix der identischen Abbildung bzgl. der Basen $\mathcal{B}'$ und $\mathcal{B}$ ist, d.h. $S = [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$; es folgt $S[v]_{\mathcal{B}'} = [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$ wegen 4.4.2.

$$\begin{tikzcd}
V_{\mathcal{B}'} \arrow[r, "\mathrm{id}_V"] \arrow[d, "\kappa_{\mathcal{B}'}"'] & V_{\mathcal{B}} \arrow[d, "\kappa_{\mathcal{B}}"] \\
\mathbb{K}^n \arrow[r, "S"'] & \mathbb{K}^n
\end{tikzcd}$$

4.6.3 Transformationsformel für Matrizen. Wir betrachten

- eine lineare Abbildung $\varphi\colon V \to W$ zwischen $\mathbb{K}$-Vektorräumen V und W,
- Zwei Basen $\mathcal{B}\colon v_1,\dots,v_n$ und $\mathcal{B}'\colon v_1',\dots,v_n'$ in V und S die Transformationsmatrix wie in 4.6.2,
- Zwei Basen $\mathcal{C}\colon w_1,\dots,w_m$ und $\mathcal{C}'\colon w_1',\dots,w_m'$ in W und R die dazugehörende Transformationsmatrix.
- die Matrix $A := [\varphi]_{\mathcal{C}}^{\mathcal{B}}$ von φ bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$,
- die Matrix $A' := [\varphi]_{\mathcal{C}'}^{\mathcal{B}'}$ von φ bzgl. der Basen $\mathcal{B}'$ und $\mathcal{C}'$.

Behauptung. *$A' = R^{-1}AS$ und $A = RA'S^{-1}$.*

Beweis. Für ein beliebiges $v \in V$ seien $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$ die Koordinaten von v bzgl. $\mathcal{B}$ und $\mathcal{B}'$. Weiterhin seien $\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ und $\begin{pmatrix} y_1' \\ \vdots \\ y_m' \end{pmatrix}$ die Koordinaten von $\varphi(v)$ bzgl. $\mathcal{C}$ und $\mathcal{C}'$. Aus 4.6.2 folgt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = R \begin{pmatrix} y_1' \\ \vdots \\ y_m' \end{pmatrix}.$$

Andererseits gilt

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \qquad \text{und} \qquad A' \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} y_1' \\ \vdots \\ y_m' \end{pmatrix}.$$

Die erste Gleichung führt auf

$$AS\begin{pmatrix}x_1'\\ \vdots\\ x_n'\end{pmatrix}=R\begin{pmatrix}y_1'\\ \vdots\\ y_m'\end{pmatrix}\qquad\text{d.h.}\qquad R^{-1}AS\begin{pmatrix}x_1'\\ \vdots\\ x_n'\end{pmatrix}=\begin{pmatrix}y_1'\\ \vdots\\ y_m'\end{pmatrix}.$$

Da dies für alle $\begin{pmatrix}x_1'\\ \vdots\\ x_n'\end{pmatrix}$ gilt, folgern wir $A'=R^{-1}AS$. □

Alternativer Beweis.

$$A'=[\varphi]^{\mathcal{B}'}_{\mathcal{C}'}=[\mathrm{id}_W\circ\varphi\circ\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{C}'}=[\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'}\cdot[\varphi]^{\mathcal{B}}_{\mathcal{C}}\cdot[\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{C}'}=R^{-1}\cdot A\cdot S.$$

$$\begin{tikzcd}
V_{\mathcal{B}'} \arrow[r, "\mathrm{id}_V"'] \arrow[d] \arrow[rrr, bend left, "\varphi"] & V_{\mathcal{B}} \arrow[r, "\varphi"'] \arrow[d] & W_{\mathcal{C}} \arrow[r, "\mathrm{id}_W"'] \arrow[d] & W_{\mathcal{C}'} \arrow[d] \\
\mathbb{K}^n \arrow[r, "S"] \arrow[rrr, bend right, "A'"'] & \mathbb{K}^n \arrow[r, "A"] & \mathbb{K}^m \arrow[r, "R^{-1}"] & \mathbb{K}^m
\end{tikzcd}$$

4.6.4 Transformation der Matrix eines Endomorphismus. *Sei $\varphi\colon V\to V$ ein Endomorphismus von V. Seien $\mathcal{B}$ und $\mathcal{B}'$ zwei Basen von V. Weiter sei A die Matrix von φ bzgl. $\mathcal{B}$ und A' die Matrix von φ bzgl. $\mathcal{B}'$ wie in 4.4.3. Dann gilt*

$$A'=S^{-1}AS\qquad\textit{und}\qquad A=SA'S^{-1}.$$

Dabei ist S die Transformationsmatrix nach 4.6.2.

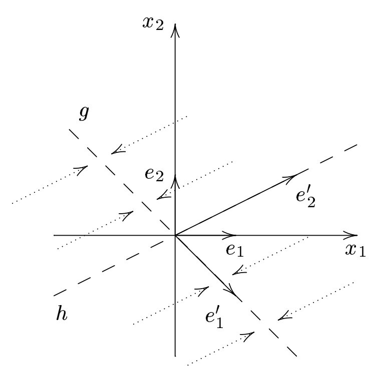
Beweis. Setze $V=W$, $\mathcal{B}=\mathcal{C}$ und $\mathcal{B}'=\mathcal{C}'$ in dem Vorhergehenden. Es folgt $R=S$ und $A'=S^{-1}AS$ nach 4.6.3. □

4.6.5 Beispiele. Wir betrachten zwei Beispiele:

(a) Betrachte in $\mathbb{R}^2$ die beiden Geraden g und h mit den Gleichungen:

$$g\colon x_1+x_2=0$$
$$h\colon x_1-2x_2=0$$

Sei $\varphi:\mathbb{R}^2\to\mathbb{R}^2$ die Projektion auf g, jedoch nicht die orthogonale Projektion, sondern die Projektion parallel zur Geraden h.
Aufgabe. Finde die Matrix A von φ bzgl. der Standardbasis $\mathcal{B}: e_1=\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right),\ e_2=\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$.



Wir wissen aus Beispiel 4.4.4 (b), dass bei Wahl einer geeigneten Basis $\mathcal{B}'\colon e_1', e_2'$, wobei e_1' ein Vektor in Richtung von g und e_2' ein Vektor in Richtung von h ist, die Matrix von φ bzgl. $\mathcal{B}'$ folgende Gestalt hat:

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

Wir wählen in diesem Fall $e'_1 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ und $e'_2 = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$. Mit der Transformationsformel in 4.6.4 ergibt sich $A = SA'S^{-1}$, wobei S die Transformationsmatrix ist. Hier gilt

$$S = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix},$$

und man berechnet

$$S^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}.$$

Daraus folgt:

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \end{aligned}$$

(b) Betrachte im $\mathbb{R}^2$ die zwei Basen $\mathcal{B}\colon e_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right),\ e_2 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ und $\mathcal{B}'\colon e'_1 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right),\ e'_2 = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$ wie im vorigen Beispiel. Wir betrachten den Endomorphismus $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$, der durch die Matrix

$$A = \frac{1}{3}\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix}$$

gegeben ist und zwar bzgl. der Standardbasis $\mathcal{B}$.

Aufgabe. Finde die Matrix A' von φ bzgl. der Basis $\mathcal{B}'$.

Nach 4.6.4 gilt $A' = S^{-1}AS$ mit S und S^{-1} wie in 4.6.5(a). Es folgt:

$$\begin{aligned} A' &= \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 12 \\ -3 & 6 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 9 & 0 \\ 0 & 18 \end{pmatrix}, \end{aligned}$$

also

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

Da A' eine Diagonalmatrix ist, können wir erkennen, wie die Abbildung φ auf $\mathbb{R}^2$ wirkt:

$$\varphi(e'_1) = e'_1,\ \varphi(e'_2) = 2e'_2.$$

Die Basis $\mathcal{B}'$ ist günstig zur Beschreibung der Wirkung von φ. Sie ist darum viel besser geeignet als die Basis $\mathcal{B}$. Wir werden noch lernen, wie wir eine Basis $\mathcal{B}'$ finden können, bzgl. der ein gegebener Endomorphismus φ eine einfache Darstellung besitzt.

4.7. Orthogonale bzw. unitäre Abbildungen und Matrizen

In diesem Abschnitt sei V ein euklidischer oder unitärer Vektorraum. Wir werden hier alle Überlegungen parallel für den reellen (euklidischen) und komplexen (unitären) Fall durchführen. Natürlich werden $\mathbb{R}^n$ und $\mathbb{C}^n$ mit dem Standardskalarprodukt (siehe die Abschnitte 3.2 und 3.3) die Hauptbeispiele sein.

Wir wollen lineare Abbildungen $\varphi\colon V \to V$ betrachten, welche Längen und Winkel erhalten.

4.7.1 Orthogonale [unitäre] Operatoren. Eine Abbildung $\varphi\colon V \to V$ wird *orthogonal* [im komplexen Fall *unitär*] genannt, wenn φ linear und bijektiv ist und wenn φ das Skalarprodukt invariant läßt, d.h.:

$$\text{(I)} \qquad \langle \varphi(\mathbf{v}) \mid \varphi(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \quad \text{für alle} \quad \mathbf{v}, \mathbf{w} \in V.$$

Übung. (a) Zeigen Sie, daß aus der Invarianzeigenschaft (I) die Linearität von φ gefolgert werden kann. (Man kann also auf die Forderung der Linearität in der obigen Definition verzichten, da sie aus (I) folgt.)

(b) Zeigen Sie, daß aus der Invarianzeigenschaft (I) die Injektivität der Abbildung φ folgt. (Ist V endlichdimensional, dann kann man auch auf die Forderung der Bijektivität in der obigen Definition verzichten, da ein injektiver Endomorphismus eines endlichdimensionalen Vektorraums nach 4.2.9 sogar bijektiv ist.)

(c) Zeigen Sie: Für unendlichdimensionale euklidische Vektorräume gibt es lineare Abbildungen, welche die Invarianzeigenschaft (I) besitzen aber nicht bijektiv sind. (Im Raum ℓ_2 (siehe 3.7.2) betrachte den *Shiftoperator* $(\alpha_0, \alpha_1, \alpha_2, \dots) \mapsto (0, \alpha_0, \alpha_1, \alpha_2, \dots)$.)

Beispiele für orthogonale Abbildungen. :

- in $\mathbb{R}^2$: Drehungen um den Ursprung und Spiegelungen an Geraden durch den Ursprung;
- in $\mathbb{R}^3$: Drehungen um eine Achse durch den Ursprung, Spiegelungen an Ebenen durch den Ursprung und Spiegelungen an Geraden durch den Ursprung.

Achtung: Orthogonale Projektionen sind keine orthogonalen Abbildungen im obigen Sinn, da sie weder die Länge von Vektoren noch Orthogonalität erhalten.

Bemerkung. Orthogonale [unitäre] Abbildungen φ erhalten Länge und Orthogonalität, d.h.

$$\|\varphi(\mathbf{v})\| = \|\mathbf{v}\| \quad , \quad \mathbf{v} \perp \mathbf{w} \Rightarrow \varphi(\mathbf{v}) \perp \varphi(\mathbf{w}).$$

In der Tat gilt $\|\varphi(\mathbf{v})\| = \sqrt{\langle \varphi(\mathbf{v}) | \varphi(\mathbf{v}) \rangle} = \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle} = \|\mathbf{v}\|$, und aus $\langle \mathbf{v} | \mathbf{w} \rangle = 0$ folgt $\langle \varphi(\mathbf{v}) | \varphi(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle = 0$.

Allgemeiner sind orthogonale Abbildungen winkelerhaltend.

Um zu zeigen, daß die orthogonalen bzw. unitären Operatoren eine Gruppe bilden, erweist es sich als günstig, den Begriff einer Untergruppe einzuführen:

4.7.2 Untergruppen. Sei G eine Gruppe mit der Operation $*$ und dem Neutralelement e. Eine Teilmenge H von G wird *Untergruppe* [*subgroup*] genannt, falls die folgenden Eigenschaften erfüllt sind:

(U1) $e \in H$,

(U2) $a, b \in H \implies a * b \in H$,

(U3) $a \in H \implies a^{-1} \in H$.

Aus der Definition folgt unmittelbar, daß eine Untergruppe wiederum eine Gruppe ist bezüglich der auf H eingeschränkten Operation $*$.

Die Untergruppenaxiome sind analog zu denen eines linearen Teilraums in einem Vektorraum. In der Tat, ist jeder lineare Teilraum eines Vektorraums eine additive Untergruppe. Allerdings gibt es in Vektorräumen additive Untergruppen, die keine linearen Teilräume sind. Beispielsweise bilden die ganzen Zahlen eine additive Untergruppe $\mathbb{Z}$ von $\mathbb{R}$. Allgemeiner: Die Gitterpunkte mit ganzzahligen Koordinaten bilden eine additive Untergruppe $\mathbb{Z}^n$ des $\mathbb{R}^n$.

4.7.3 Die orthogonale und die unitäre Gruppe. Wir bezeichnen mit $\mathrm{O}(V)$ die Menge aller orthogonalen Operatoren auf einem euklidischen Vektorraum und mit $\mathrm{U}(V)$ die Menge aller unitären Operatoren auf einem unitären Vektorraum V. Wir behaupten, daß $\mathrm{O}(V)$ und $\mathrm{U}(V)$ Gruppen sind (mit der Hintereinanderausführung von Abbildungen als Operation). Dazu zeigen wir, daß diese Mengen Untergruppen der Gruppe $\mathrm{Aut}(V)$ aller Automorphismen von V sind. Wir führen dies im euklidischen Fall durch (die Beweise und Aussagen sind im unitären Fall gleich):

(U1) $\mathrm{id}_V \in \mathrm{O}(V)$, da die identische Abbildung offensichtlich orthogonal ist.

(U2) $\varphi, \psi \in \mathrm{O}(V) \Rightarrow \varphi \circ \psi \in \mathrm{O}(V)$

Wenn φ und ψ das Skalarprodukt erhalten, folgt

$$\langle (\varphi \circ \psi)(\mathbf{v}) \mid (\varphi \circ \psi)(\mathbf{w}) \rangle = \langle \varphi(\psi(\mathbf{v})) \mid \varphi(\psi(\mathbf{w})) \rangle = \langle \psi(\mathbf{v}) \mid \psi(\mathbf{w}) \rangle = \langle \mathbf{v} \mid \mathbf{w} \rangle$$

(U3) $\varphi \in \mathrm{O}(V) \Rightarrow \varphi^{-1} \in \mathrm{O}(V)$

Eine orthogonale Abbildung φ ist per Definition bijektiv; sie hat also eine Umkehrabbildung φ^{-1} und es gilt $v = \varphi(\varphi^{-1}(v))$ für alle $v \in V$. Also folgt

$$\langle v|w \rangle = \langle \varphi(\varphi^{-1}(v)) \mid \varphi(\varphi^{-1}(w)) \rangle = \langle \varphi^{-1}(v) \mid \varphi^{-1}(w) \rangle$$

für alle v, w. Folglich ist φ^{-1} eine orthogonale Abbildung.

4.7.4 Proposition. *Der Vektorraum V sei endlichdimensional. Eine lineare Abbildung $\varphi \colon V \to V$ ist genau dann orthogonal [unitär], wenn das Bild einer Orthonormalbasis wieder eine Orthonormalbasis von V ist.*

Beweis. Sei $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ eine Orthonormalbasis von V. Dann gilt $\langle \mathbf{v}_i | \mathbf{v}_j \rangle = \delta_{ij}$ für alle i, j.

Ist φ orthogonal [unitär], so folgt per definitionem $\langle \varphi(\mathbf{v}_i) \mid \varphi(\mathbf{v}_j) \rangle = \langle \mathbf{v}_i | \mathbf{v}_j \rangle = \delta_{ij}$; damit ist $\varphi(\mathbf{v}_1), \varphi(\mathbf{v}_2), \ldots, \varphi(\mathbf{v}_n)$ ebenfalls eine Orthonormalbasis.

Ist umgekehrt $\varphi(\mathbf{v}_1), \varphi(\mathbf{v}_2), \ldots, \varphi(\mathbf{v}_n)$ eine Orthonormalbasis von V, dann gilt $\langle\varphi(\mathbf{v}_i) \mid \varphi(\mathbf{v}_j)\rangle = \delta_{ij}$. Wir zeigen, dass die Abbildung φ orthogonal [unitär] ist: Für beliebiges $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$ und $\mathbf{w} = \sum_{j=1}^n \beta_j \mathbf{v}_j$ haben wir dann (siehe 3.4.3):

$$\begin{aligned}\langle\varphi(\mathbf{v}) \mid \varphi(\mathbf{w})\rangle &= \langle\varphi(\sum_{i=1}^n \alpha_i \mathbf{v}_i) \mid \varphi(\sum_{j=1}^n \beta_j \mathbf{v}_j)\rangle = \langle\sum_{i=1}^n \alpha_i \varphi(\mathbf{v}_i) \mid \sum_{j=1}^n \beta_j \varphi(\mathbf{v}_j)\rangle \\ &= \sum_{i,j} \overline{\alpha_i}\beta_j \langle\varphi(\mathbf{v}_i) \mid \varphi(\mathbf{v}_j)\rangle = \sum_{i,j} \overline{\alpha_i}\beta_j \delta_{ij} = \sum_{i=1}^n \overline{\alpha_i}\beta_i = \langle\mathbf{v}|\mathbf{w}\rangle.\end{aligned}$$

□

4.7.5 Orthogonale und unitäre Matrizen. Sei Q eine $n \times n$-Matrix mit reellen [komplexen] Einträgen. Wir geben nun zwei äquivalente Definitionen:

Erste Definition. Wir nennen Q eine *orthogonale [unitäre]* Matrix, wenn die durch Q definierte lineare Abbildung

$$\begin{array}{rcl} \varphi_Q : \mathbb{R}^n & \to & \mathbb{R}^n \\ v & \mapsto & Qv \end{array} \qquad \left[\begin{array}{rcl} \varphi_Q : \mathbb{C}^n & \to & \mathbb{C}^n \\ v & \mapsto & Qv \end{array} \right]$$

orthogonal [unitär] ist, d.h. wenn gilt:

$$\langle Qv|Qw\rangle = \langle v|w\rangle \quad \text{für alle } v, w$$

Betrachte die Standardbasis $e_1, e_2, \ldots, e_n$ in $\mathbb{R}^n [\mathbb{C}^n]$, die eine Orthonormalbasis ist. Wir wissen aus 4.7.4, dass Q genau dann orthogonal [unitär] nach der ersten Definition ist, wenn die Vektoren Qe_1, Qe_2, $\ldots$, Qe_n ebenfalls eine Orthonormalbasis bilden. Da Qe_1, Qe_2, $\ldots$, Qe_n gerade die Spalten von Q sind, ergibt sich die folgende äquivalente Definition:

Zweite Definition. Die Matrix Q ist genau dann *orthogonal [unitär]*, wenn ihre Spalten $s_1, s_2, \ldots, s_n$ ein Orthonormalsystem von Vektoren in $\mathbb{R}^n [\mathbb{C}^n]$ bilden, d.h. wenn

$$\langle s_i|s_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

Beachte, dass die orthogonalen Matrizen genau diejenigen unitären Matrizen sind, deren Einträge alle reell sind.

Beispiel. $Q = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & 2 & 1 \\ 2 & -1 & -2 \end{pmatrix}$ ist eine orthogonale Matrix.

4.7.6 Proposition. *Ein Endomorphismus φ eines endlichdimensionalen euklidischen [unitären] Vektorraumes V ist genau dann orthogonal [unitär], wenn seine Matrix Q bzgl. einer beliebigen Orthonormalbasis $\mathcal{B} : v_1, v_2, \ldots, v_n$ eine orthogonale [unitäre] Matrix ist.*

Dies folgt aus 4.7.4 und aus der Tatsache, dass die Spalten der Matrix Q von φ bzgl. der Basis $\mathcal{B}$ die Koordinatenvektoren $[\varphi(v_1)]_{\mathcal{B}}, \ldots, [\varphi(v_n)]_{\mathcal{B}}$ sind.

4.7.7 Charakterisierung von orthogonalen [unitären] Matrizen. Für eine beliebige komplexe Matrix $M = (\alpha_{ij})$ benutzen wir folgenden Bezeichnungen:

$$\overline{M} := (\overline{\alpha_{ij}}) \quad , \quad M^* := \overline{M}^T = \overline{M^T} \qquad \text{(die Transponierte von } \overline{M}\text{)}$$

Man nennt M^* die zu M adjungierte Matrix. Für reele Matrizen ist die Adjungierte gleich der Transponierten.

Beispiel.

$$M = \begin{pmatrix} 1+i & 1 \\ i & 2-i \end{pmatrix} \quad , \quad \overline{M} = \begin{pmatrix} 1-i & 1 \\ -i & 2+i \end{pmatrix} \quad , \quad M^* = \begin{pmatrix} 1-i & -i \\ 1 & 2+i \end{pmatrix}$$

Satz. Für eine $n \times n$-Matrix Q mit reellen [komplexen] Einträgen sind die folgenden Eigenschaften äquivalent:

(i) Q ist orthogonal [unitär], d.h. die Spalten von Q bilden ein Orthonormalsystem;

(ii) Die Zeilen von Q bilden ein Orthonormalsystem;

(iii) $Q^T Q = E_n$ $[Q^* Q = E_n]$;

(iv) $QQ^T = E_n$ $[QQ^* = E_n]$;

(v) Q ist invertierbar und es gilt $Q^{-1} = Q^T$ $[Q^{-1} = Q^*]$.

Beweis.

(i) $\Leftrightarrow$ (iii): Wir betrachten $Q^T Q = (\gamma_{ij})$ mit:

$$\begin{aligned} \gamma_{ij} &= (i-\text{te Zeile von } Q^T) \cdot (j-\text{te Spalte von } Q) \\ &= (i-\text{te Spalte von } Q) \cdot (j-\text{te Spalte von } Q) \end{aligned}$$

Also gilt $Q^T Q = E_n = (\delta_{ij})$, genau dann wenn

$$(i-\text{te Spalte von } Q) \cdot (j-\text{te Spalte von } Q) = \delta_{ij}$$

d.h. genau dann, wenn die Spalten von Q ein Orthonormalsystem bilden.

(ii) $\Leftrightarrow$ (iv): wird analog gezeigt.

(iii) $\Rightarrow$ (v): Gilt $Q^T Q = E_n$, so ist Q nach 4.3.9 invertierbar und $Q^{-1} = Q^T$. Ebenso folgt aus $Q^* Q = E_n$, daß Q invertierbar ist und $Q^{-1} = Q^*$.

(iv) $\Rightarrow$ (v): folgt ganz analog.

(v) $\Rightarrow$ (iii) und (iv): Ist Q invertierbar und $Q^{-1} = Q^T$ $[\,Q^{-1} = Q^*]$, dann folgt $E_n = QQ^T = Q^T Q$ $[E_n = QQ^* = Q^* Q]$.

□

Beispiel. Für die orthogonale Matrix im Beispiel aus 4.7.5 gilt:

$$Q^{-1} = Q^T = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ -2 & 2 & -1 \\ 2 & 1 & -2 \end{pmatrix}$$

4.7.8 Die orthogonalen und unitären Matrixgruppen.

Die Eigenschaften in 4.7.3 ergeben:

(U1) E_n ist eine orthogonale [unitäre] $n \times n$-Matrix.

(U2) Sind P und Q beide orthogonale [unitäre] $n \times n$-Matrizen, dann ist PQ ebenfalls orthogonal [unitär].

(U3) Ist Q eine orthogonale [unitäre] $n \times n$-Matrix, dann ist Q^{-1} auch orthogonal [unitär].

Also sind

$$\begin{array}{lcll} \mathrm{O}(n) & := & \text{Menge aller orthogonalen} & n \times n\text{-Matrizen } Q \\ \mathrm{U}(n) & := & \text{Menge aller unitären} & n \times n\text{-Matrizen } Q \end{array}$$

Untergruppen der Gruppen $\mathrm{GL}(n, \mathbb{R})$ bzw. $\mathrm{GL}(n, \mathbb{C})$ aller invertierbaren reellen bzw. komplexen $n \times n$-Matrizen. Insbesondere sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ Gruppen bzgl. der Matrixmultiplikation.

4.7.9 Orthogonale 2×2-Matrizen Q.

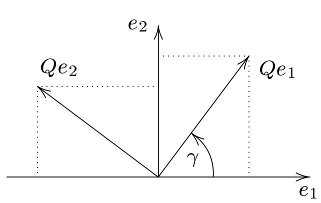
Es sei Q eine orthogonale 2×2-Matrix. Dann gilt:

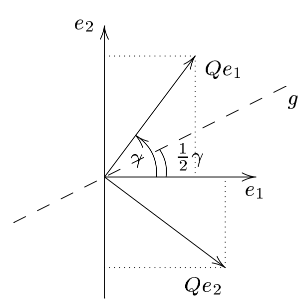
- entweder beschreibt Q eine Drehung um den Ursprung um einen Winkel γ $(0 \leq \gamma \leq 2\pi)$:

$$Q = D(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$$

- oder Q beschreibt eine Spiegelung an einer Geraden durch den Ursprung, die mit der x_1-Achse den Winkel $\frac{\gamma}{2}$ (entgegen dem Uhrzeigersinn) bildet:

$$Q = \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & -\cos\gamma \end{pmatrix}$$





Damit sind die einzigen orthogonalen Operatoren auf $\mathbb{R}^2$ Drehungen um den Ursprung O und Spiegelungen an Geraden durch O. Die ersteren erhalten die Orientierung, die zweiten kehren die Orientierung um. Wir werden spärter sehen, daß man die beiden Fälle mit Hilfe der Determinante unterscheiden kann.

4.8. Systeme linearer Gleichungen

Wie in Abschnitt 1.3 betrachten wir ein System von m linearen Gleichungen in n Unbekannten $x_1, \ldots, x_n$:

$$\text{(L)} \qquad \begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + \ldots + & \alpha_{1j}x_j & + \ldots + & \alpha_{1n}x_n & = & \beta_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + \ldots + & \alpha_{2j}x_j & + \ldots + & \alpha_{2n}x_n & = & \beta_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{i1}x_1 & + & \alpha_{i2}x_2 & + \ldots + & \alpha_{ij}x_j & + \ldots + & \alpha_{in}x_n & = & \beta_i \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + \ldots + & \alpha_{mj}x_j & + \ldots + & \alpha_{mn}x_n & = & \beta_m. \end{array}$$

Die Koeffizienten α_{ij} und die rechten Seiten β_i sollen dabei Elemente des Körpers $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ sein.

Betrachten wir die Matrix der Koeffizienten und die Vektoren

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \in \mathcal{M}(m \times n, \mathbb{K}), \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix},$$

so kann dieses System in der Form

$$\text{(L)} \qquad Ax = b$$

geschrieben werden. Ist $b \neq 0$, so heißt das System *inhomogen*; das System

$$\text{(H)} \qquad Ax = 0$$

heißt das zu (L) gehörende *homogene* System.

Unser Problem besteht darin, die Menge aller Lösungen von (L) zu beschreiben, d.h. die Menge aller $x \in \mathbb{K}^n$, die (L) erfüllen, zu suchen.

4.8.1 Die Lösungen des homogenen Systems (H)**.** Wie wir in 4.3.10 gesehen haben, ist die Menge aller Lösungen von (H) ein linearer Teilraum U (mit Dimension $k = n - r$) in $\mathbb{K}^n$, wobei $r := \operatorname{rank} A$ ist. (Man erinnere sich, dass der Rang von A die maximale Zahl linear unabhängiger Spalten (Zeilen) von A ist.)

Sei $v_1, \ldots, v_k$ eine Basis des linearen Teilraums U der Lösungen von (H). Wir sagen auch, dass $v_1, \ldots, v_k$ ein Fundamentalsystem von Lösungen von (H) ist. Die „allgemeine“ Lösung von (H) ist dann gegeben durch

$$x_H = \lambda_1 v_1 + \cdots \lambda_k v_k$$

mit beliebigen $\lambda_i \in \mathbb{K}$.

4.8.2 Die Existenz von Lösungen für (L)**.** Das homogene System (H) hat immer eine Lösung, nämlich $x = 0$. Diese heißt die *triviale* Lösung. Ein inhomogenes System (L) braucht keine Lösung zu haben. Wir wissen aus 4.3.4, dass

$$\text{(L)} \qquad Ax = b$$

genau dann eine Lösung hat, wenn b in dem linearen Teilraum von $\mathbb{K}^n$ liegt, der von den Spalten von $w_1, \ldots, w_n$ von A aufgespannt wird, d.h. es gibt genau dann mindestens eine Lösung, wenn
$$\operatorname{rank} A \quad = \quad \operatorname{rank}(A \mid b)$$
ist, wobei $(A \mid b)$ die erweiterte $m \times (n+1)$-Matrix ist, die aus A durch Hinzufügen von b als zusätzlicher Spalte gewonnen wird:
$$(A|b) = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} & \beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} & \beta_m \end{array} \right).$$

4.8.3 Die Lösungen des inhomogenen Systems (L)**.** Wir nehmen an, (L) habe wenigstens eine Lösung $x_S \in \mathbb{K}^n$ (eine *spezielle* Lösung). Dann ist die Menge aller Lösungen ein affiner Teilraum von $\mathbb{K}^n$, nämlich $x_S + U$, wobei U die Menge der Lösungen des zugeordneten homogenen Systems (H) ist. Die Dimension dieses affinen Teilraums ist $k = n - r$ mit $r = \operatorname{rank} A$.

(Dies stimmt, denn es gilt: $x \in x_S + U \iff x = x_S + x_H \iff Ax = A(x_S + x_H) = Ax_S + Ax_H = b + 0 = b$.)

Also kann die „allgemeine“ Lösung von (L) in der Form
$$\boxed{x = x_S + x_H = x_s + \lambda_1 v_1 + \ldots + \lambda_k v_k \ , \ \lambda_i \in \mathbb{K}}$$
geschrieben werden, wobei x_S eine spezielle Lösung von (L) und $v_1, \ldots, v_k$ ein Fundamentalsystem der Lösungen von (H) ist.

4.8.4 Eindeutigkeit von Lösungen. Wir beschränken uns auf den Fall $n = m$. Dann ist A eine quadratische Matrix.

Behauptung. *Für eine $n \times n$-Matrix A und ein $b \in \mathbb{K}^n$ sind die folgenden Aussagen äquivalent:*

(1) *Das System $Ax = b$ hat eine eindeutige Lösung x_0.*

(2) *A ist invertierbar.*

(3) $\operatorname{rank} A = n$ *(d.h. die Spalten von A sind linear unabhängig.)*

Die Äquivalenz von (2) und (3) folgt aus 4.3.9.

(1) $\Rightarrow$ (2): Hat $Ax = b$ eine eindeutige Lösung x_0, so ist die Menge aller Lösungen von $Ax = b$ nulldimensional; aus 4.8.3 folgt damit $0 = k = n - r$, d.h. $n = r = \operatorname{rank} A$.

(2) $\Rightarrow$ (1): Ist A invertierbar, folgt $Ax = b \iff x = A^{-1}b$, d.h. $A^{-1}b$ ist die einzige Lösung von $Ax = b$.

Korollar. *Für eine $n \times n$-Matrix A sind die folgenden Aussagen äquivalent:*

(1) *Das homogene System $Ax = 0$ hat eine Lösung $x \neq 0$ (eine solche Lösung heißt nichttrivial).*

(2) *A ist nicht invertierbar.*

(3) $\operatorname{rank} A < n$ *(d.h. die Spalten von A sind linear abhängig.)*

Eigenschaft (1) bedeutet, dass $Ax = 0$ mehr als eine Lösung hat, da ein homogenes System immer die Lösung $x = 0$ hat. Also folgt das Korollar aus der vorigen Behauptung durch Negation der dort angegebenen Eigenschaften.

Ergänzung. Auf den vorigen zwei Seiten haben wir zusammengefasst, wie die Menge der Lösungen eines Systems linearer Gleichungen im Prinzip beschrieben werden kann. Wir wollen nun eine Methode angeben, wie wir explizit diese Lösungsmenge beschreiben können:

Gegeben sei ein System linearer Gleichungen $(L) \quad Ax = b$. Wir stellen uns folgende Fragen:

- Gibt es überhaupt eine Lösung? Wenn ja, finde eine spezielle Lösung x_S.
- Finde ein Fundamentalsystem $v_1, \ldots, v_k$ von Lösungen für das dazugehörige homogene System $(H) \quad Ax = 0$ mit $k = n - \operatorname{rank}(A)$.

Wie in 1.3 kann dies mit dem Gauß-Jordan Eliminations-Algorithmus getan werden:

Bilde die erweiterte Matrix

$$(A|b) = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} & \beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} & \beta_m \end{array}\right).$$

Transformiere diese Matrix mittels elementarer Zeilentransformationen in Stufenform:

$$\left(\begin{array}{cccccccc|c} \gamma_{1j_1} & * & * & * & \ldots & \ldots & \ldots & * & \delta_1 \\ & & \gamma_{2j_2} & * & \ldots & \ldots & \ldots & * & \delta_2 \\ & & & \ddots & & & & \vdots & \vdots \\ & & & & \gamma_{rj_r} & * & \ldots & * & \delta_r \\ & \bigcirc & & & & & & & \delta_{r+1} \\ & & & & & & & & \vdots \\ & & & & & & & & \delta_m \end{array}\right)$$

wobei $j_1 < j_2 < \ldots < j_r$ und $\gamma_{ij_i} \neq 0$ für $i = 1, \ldots, r$.

Erster Fall: Einer der Skalare $\delta_{r+1}, \ldots, \delta_m$ ist $\neq 0$. Dann hat das System $(L) \quad Ax = b$ keine Lösung.

Zweiter Fall: $\delta_{r+1} = \ldots = \delta_m = 0$. Dann hat das System $(L) \quad Ax = b$ Lösungen. Finde eine spezielle Lösung x_S, z.B. durch Wahl aller Nicht-Pivotvariablen $x_i = 0$, $i \neq j_1, \ldots, j_r$.

Betrachte dann das zugehörige homogene System $(H)\, Ax = 0$. Um ein Fundamentalsystem $v_1, \ldots, v_k$ zu finden, betrachte das transformierte System in Dreiecksgestalt mit $\delta_1 = \ldots = \delta_r = 0$ und löse dieses System in den folgenden k Fällen:

Wähle als Nicht-Pivotvariablen $x_i, i \neq j_1, \ldots, j_r$ nacheinander die Werte

$$\begin{array}{cccccc} 1, & 0, & 0, & \dots, & 0 \\ 0, & 1, & 0, & \dots, & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0, & \dots & \dots & 0, & 1 \end{array}$$

4.8.5 Beispiel ($\mathbb{K} = \mathbb{R}$)**.** Betrachte das System:

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & - & x_3 & + & 2x_4 & + & 3x_5 & = & 1 \\ 2x_1 & + & 4x_2 & - & 2x_3 & + & 5x_4 & + & 5x_5 & = & 3 \\ -x_1 & - & 2x_2 & - & x_3 & + & x_4 & + & x_5 & = & 1 \\ 2x_1 & + & 4x_2 & & & + & 2x_4 & + & x_5 & = & 1 \\ 3x_1 & + & 6x_2 & - & 3x_3 & + & 7x_4 & + & 8x_5 & = & 4 \end{array}$$

1	2	−1	2	3	1
2	4	−2	5	5	3
−1	−2	−1	1	1	1
2	4	0	2	1	1
3	6	−3	7	8	4
1	2	−1	2	3	1
0	0	0	1	−1	1
0	0	−2	3	4	2
0	0	2	−2	−5	−1
0	0	0	1	−1	1
1	2	−1	2	3	1
0	0	−2	3	4	2
0	0	0	1	−1	1
0	0	2	−2	−5	−1
0	0	0	1	−1	1
1	2	−1	2	3	1
0	0	−2	3	4	2
0	0	0	1	−1	1
0	0	0	1	−1	1
0	0	0	1	−1	1
1	2	−1	2	3	1
0	0	−2	3	4	2
0	0	0	1	−1	1
0	0	0	0	0	0
0	0	0	0	0	0

Gauß-Jordan Elimination ergibt: $r = 3$, $k = 2$.

(1) Das System hat Lösungen.

(2) Die Pivotvariablen sind x_1, x_3, x_4. Für eine spezielle Lösung x_S sei $x_2 = x_5 = 0$ im inhomogenen System in Dreiecksgestalt. Dies führt auf

$$x_S = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}.$$

(3) Es sei für ein Fundamentalsystem v_1, v_2 von Lösungen des homogenen Systems:

$$x_2 = 1, x_5 = 0 : \qquad v_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$x_2 = 0, x_5 = 1 : \qquad v_2 = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{7}{2} \\ 1 \\ 1 \end{pmatrix}$$

Die „allgemeine“ Lösung:

$$x = x_s + \lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{7}{2} \\ 1 \\ 1 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \text{ beliebig.}$$

4.8.6 Die QR-Zerlegung einer invertierbaren Matrix B. Sei B eine invertierbare $n \times n$-Matrix mit reellen [komplexen] Koeffizienten. Dann sind die n Spalten von B linear unabhängig und wir können die QR-Zerlegung von B wie in 3.5.9 konstruieren und erhalten

$$B = QR$$

Dabei ist Q eine orthogonale [unitäre] Matrix und R eine obere Dreicksmatrix mit positiven reellen Einträgen auf der Diagonalen.

Übung Zeige, dass die Zerlegung einer invertierbaren Matrix X in ein Produkt aus einer orthogonalen [unitären] Matrix Q und einer oberen Dreiecksmatrix R mit positiven reellen Einträgen eindeutig ist.

Anwendung auf das Lösen von linearen Gleichungssystemen

$$Bx = b$$

im Falle, dass B eine invertierbare $n \times n$-Matrix mit reellen [komplexen] Einträgen ist. Ersetzt man B durch ihre QR -Zerlegung, so erhält man

$$QRx = b$$

und wegen $Q^{-1} = Q^T$ $[Q^{-1} = Q^*]$ können wir das System in der Form

$$Rx = Q^T b \quad [Rx = Q^* b]$$

schreiben. Das letzte Gleichungssystem kann leicht gelöst werden, da R obere Dreiecksgestalt besitzt.

4.9. Die Spur

Die Defintion der Spur einer Matrix erscheint einfach und eine geometrische Motivation für diese ist schwer zu finden. Dennoch ist dieser Begriff von Nutzen, da die Spur unter Basiswechsel invariant ist.

4.9.1 Die Spur einer quadratischen Matrix. Sei $A = (\alpha_{ij}) \in \mathcal{M}(n \times n, \mathbb{K})$. Die Summe $\sum_{i=1}^n \alpha_{ii}$ der Elemente auf der Diagonalen von A heißt die *Spur* von A. Formaler:

$$\operatorname{tr} A = \operatorname{tr} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} := \alpha_{11} + \alpha_{22} + \ldots + \alpha_{nn} \in \mathbb{K}.$$

Beispiele.

(1)
$$\operatorname{tr} E_n = n$$
(2)
$$\operatorname{tr}\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 + 2\cos\gamma.$$

4.9.2 Regeln für die Spur. Für beliebige $n \times n$-Matrizen A, B und $\lambda \in \mathbb{K}$ gilt:

(a)
$$\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$$
(b)
$$\operatorname{tr}(\lambda A) = \lambda \operatorname{tr} A$$
(c)
$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Die beiden ersten Eigenschaften sind offensichtlich. Sie bedeuten, dass die Spur linear ist, wenn man sie als eine Abbildung vom Vektorraum $\mathcal{M}(n \times n, \mathbb{K})$ aller $n \times n$-Matrizen in den eindimensionalen Vektorraum $\mathbb{K}$ auffasst. Zum Beweis von (c) sei $A = (\alpha_{ij})$ und $B = (\beta_{ij})$. Dann gilt:

$$\begin{aligned}
\operatorname{tr}(AB) &= \sum_{i=1}^n \gamma_{ii} && \text{mit } \gamma_{ii} = \sum_{j=1}^n \alpha_{ij}\beta_{ji} \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}\beta_{ji} \\
\operatorname{tr}(BA) &= \sum_{i=1}^n \gamma'_{ii} && \text{mit } \gamma'_{ii} = \sum_{j=1}^n \beta_{ij}\alpha_{ji} \\
&= \sum_{i=1}^n \sum_{j=1}^n \beta_{ij}\alpha_{ji} = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ji}\beta_{ij} = \sum_{j=1}^n \sum_{i=1}^n \alpha_{ji}\beta_{ij} \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}\beta_{ji} && \text{(durch Wechsel der Rollen von } i \text{ und } j) \\
&= \operatorname{tr}(AB).
\end{aligned}$$

Bemerkung. Im Allgemeinen gilt $\operatorname{tr}(AB) \neq \operatorname{tr} A \cdot \operatorname{tr} B$.

Es sei z.B. $A = B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Dann ist $\operatorname{tr} A = \operatorname{tr} B = 0$ und $\operatorname{tr} A \cdot \operatorname{tr} B = 0$, aber $\operatorname{tr}(A \cdot B) = \operatorname{tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2$.

4.9.3 Proposition. Für jede invertierbare $n \times n$-Matrix S und jede $n \times n$-Matrix A gilt

$$\boxed{\operatorname{tr}(S^{-1}AS) = \operatorname{tr} A}$$

Beweis.

$$\operatorname{tr}(S^{-1}AS) = \operatorname{tr}((S^{-1}A)S) \underset{(c)}{=} \operatorname{tr}(S(S^{-1}A)) = \operatorname{tr}((S^{-1}S)A) = \operatorname{tr} A.$$

□

4.9.4 Die Spur eines Endomorphismus. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\varphi : V \to V$ eine lineare Abbildung.

Betrachte die Matrix A von φ bzgl. einer Basis $\mathcal{B}$ von V und die Matrix A' von φ bzgl. einer anderen Basis $\mathcal{B}'$ von V.

Nach 4.6.4 gibt es eine invertierbare Matrix S (die Transformationsmatrix), so dass $A' = S^{-1}AS$ gilt. Aus 4.9.3 folgern wir

$$\operatorname{tr} A' = \operatorname{tr} A,$$

d.h. die Spur einer Matrix ist invariant unter Koordinatentransformationen. Die Spur der Matrix von φ hängt nicht von der Wahl der Basis von V ab.

Demnach können wir die *Spur von* φ als die Spur der Matrix von φ bzgl. irgendeiner Basis $\mathcal{B}$ von V definieren.

Beispiele.

(a) $\operatorname{tr}(\mathrm{id}_V) = n$ $(= \dim V)$

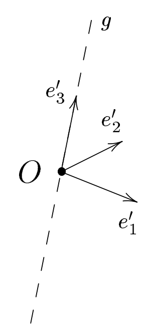
In der Tat ist E_n die Matrix von id_V bzgl. einer Basis von V und es gilt $\operatorname{tr} E_n = n$.

(b) Im euklidischen Raum $\mathbb{R}^3$ sei δ eine Drehung um eine Achse g durch den Ursprung um einen Winkel γ.

Wir behaupten

$$\operatorname{tr} \delta = 1 + 2\cos\gamma.$$

Zum Beweis wähle eine Basis $\mathcal{B}'$ von $\mathbb{R}^3$ in folgender Weise:
Wähle e_1', e_2' senkrecht zueinander und senkrecht zur Geraden g, wähle e_3' in Richtung von g. Weiterhin normiere e_1', e_2', e_3' auf Länge 1.



Die Matrix von δ bzgl. dieser Basis $\mathcal{B}'$ ist

$$A' = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Also gilt $\operatorname{tr}\delta = \operatorname{tr} A' = 1 + 2\cos\gamma$.

Ist die Matrix $A = (\alpha_{ij})$ von δ bzgl. der Standardbasis $\mathcal{B}$ gegeben und kennen wir den Winkel γ der Drehung nicht, können wir dennoch $\cos\gamma$ bestimmen:

$\operatorname{tr}\delta = 1 + 2\cos\gamma\,,$

also ist $\cos\gamma = \frac{1}{2}(\operatorname{tr}\delta - 1) = \frac{1}{2}(\operatorname{tr} A - 1) = \frac{1}{2}(\alpha_{11} + \alpha_{22} + \alpha_{33} - 1)$.

Dadurch ist der Drehwinkel γ noch nicht eindeutig bestimmt, denn $\cos\gamma = \cos -\gamma$.

KAPITEL 5

Determinanten

Der Begriff der Determinante wird motiviert als „orientierte Fläche“ im $\mathbb{R}^2$ und als „orientiertes Volumen“ im $\mathbb{R}^3$. Die Determinante beinhaltet im Reellen zwei Arten Information, nämlich VOLUMEN und ORIENTIERUNG.

Vorab betrachten wir im ersten Abschnitt dieses Kapitels eine Besonderheit des Euklidischen Raums $\mathbb{R}^3$: Im 3-dimensionalen Raum kann man ein Produkt von Vektoren definieren, das auf andere Dimensionen nicht verallgemeinert werden kann.

5.1. Das Vektorprodukt im Euklidischen $\mathbb{R}^3$

In diesem Abschnitt betrachten wir den Eulklidischen $\mathbb{R}^3$ mit dem üblichen Skalarprodukt $\langle \mathbf{v} \mid \mathbf{w} \rangle$.

5.1.1 Die Definition des Vektorprodukts. Für zwei Vektoren

$$\mathbf{v} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{und} \quad \mathbf{w} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

im $\mathbb{R}^3$ definieren wir das *Vektorprodukt* [*vector product*] (manchmal auch *Kreuzprodukt* [*cross product*] oder *äußeres Produkt* [*outer product*] genannt) durch

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} := \begin{pmatrix} \alpha_2\beta_3 - \alpha_3\beta_2 \\ -\alpha_1\beta_3 + \alpha_3\beta_1 \\ \alpha_1\beta_2 - \alpha_2\beta_1 \end{pmatrix}.$$

Anders als das Skalarprodukt ist das Vektorprodukt zweier Vektoren wieder ein Vektor.

Beispiele.

$$\begin{aligned} \mathbf{e}_1 \times \mathbf{e}_2 &= \mathbf{e}_3 \\ \mathbf{e}_2 \times \mathbf{e}_3 &= \mathbf{e}_1 \\ \mathbf{e}_3 \times \mathbf{e}_1 &= \mathbf{e}_2 \\ \mathbf{e}_1 \times \mathbf{e}_3 &= -\mathbf{e}_2 \\ \mathbf{e}_3 \times \mathbf{e}_2 &= -\mathbf{e}_1 \\ \mathbf{e}_2 \times \mathbf{e}_2 &= \mathbf{0} \\ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} &= \begin{pmatrix} 5 \\ 0 \\ 5 \end{pmatrix}. \end{aligned}$$

Beachte, dass für das Vektorprodukt das Assoziativgesetz nicht gilt; denn

$$(\mathbf{e}_1 \times \mathbf{e}_2) \times \mathbf{e}_2 = \mathbf{e}_3 \times \mathbf{e}_2 = -\mathbf{e}_1 \text{ aber } \mathbf{e}_1 \times (\mathbf{e}_2 \times \mathbf{e}_2) = \mathbf{e}_1 \times \mathbf{0} = \mathbf{0}.$$

5.1.2 Regeln für das Vektorprodukt. Für beliebige Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathbb{R}^3$ und einen beliebigen Skalar $\lambda \in \mathbb{R}$ gilt:

(P1) $(\mathbf{u}+\mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$ (Distributivität)

$(\lambda\mathbf{u}) \times \mathbf{w} = \lambda(\mathbf{u} \times \mathbf{w})$

(P2) $\mathbf{u} \times (\mathbf{v}+\mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ (Distributivität)

$\mathbf{u} \times \lambda\mathbf{w} = \lambda(\mathbf{u} \times \mathbf{w})$

(P3) $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$ (Antikommutativität)

(P4) $\mathbf{v} \times \mathbf{v} = \mathbf{0}$

(P5) $\langle \mathbf{v} \mid \mathbf{v} \times \mathbf{w} \rangle = 0$

$\langle \mathbf{w} \mid \mathbf{v} \times \mathbf{w} \rangle = 0$

(P6) $\|\mathbf{v} \times \mathbf{w}\|^2 = \|\mathbf{v}\|^2\|\mathbf{w}\|^2 - \langle \mathbf{v} \mid \mathbf{w} \rangle^2$ (Lagrange Identität)

(P7) $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = \langle \mathbf{u} \mid \mathbf{w} \rangle \cdot \mathbf{v} - \langle \mathbf{v} \mid \mathbf{w} \rangle \cdot \mathbf{u}$ (Grassmann Identität)

(P8) $\langle \mathbf{u} \times \mathbf{v} \mid \mathbf{w} \times \mathbf{x} \rangle = \langle \mathbf{u} \mid \mathbf{w} \rangle \cdot \langle \mathbf{v} \mid \mathbf{x} \rangle - \langle \mathbf{v} \mid \mathbf{w} \rangle \cdot \langle \mathbf{u} \mid \mathbf{x} \rangle$ (Laplace Identität).

Eigenschaft (P1) kann auch auf folgende Art ausgedrückt werden: für jedes feste $\mathbf{w} \in \mathbb{R}^3$ ist die Abbildung

$$\mathbf{x} \mapsto \mathbf{x} \times \mathbf{w} \colon \mathbb{R}^3 \to \mathbb{R}^3$$

linear. Ähnlich ist (P2) äquivalent dazu, daß für jedes feste $\mathbf{u} \in \mathbb{R}^3$ die Abbildung

$$\mathbf{x} \mapsto \mathbf{u} \times \mathbf{x} \colon \mathbb{R}^3 \to \mathbb{R}^3$$

linear ist.

Man kann die obigen Identitäten durch direktes Nachrechnen verifizieren, in dem man die Definitionen des Vektor- und Skalarprodukts verwendet. Dies ist recht einfach für die Eigenschaften (P1)–(P4). Zeigen wir den ersten Teil von (P5):

$$\begin{aligned}
\langle \mathbf{v} \mid \mathbf{v} \times \mathbf{w} \rangle &= \left\langle \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \,\middle|\, \begin{pmatrix} \alpha_2\beta_3 - \alpha_3\beta_2 \\ -\alpha_1\beta_3 + \alpha_3\beta_1 \\ \alpha_1\beta_2 - \alpha_2\beta_1 \end{pmatrix} \right\rangle \\
&= \alpha_1(\alpha_2\beta_3 - \alpha_3\beta_2) + \alpha_2(-\alpha_1\beta_3 + \alpha_3\beta_1) + \alpha_3(\alpha_1\beta_2 - \alpha_2\beta_1) \\
&= 0.
\end{aligned}$$

Analog zeigt man den zweiten Teil von (P5). Als nächstes zeigen wir die Lagrange Identität (P6):

$$\begin{aligned}
\|\mathbf{v} \times \mathbf{w}\|^2 &= (\alpha_2\beta_3 - \alpha_3\beta_2)^2 + (-\alpha_1\beta_3 + \alpha_3\beta_1)^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)^2 \\
\|\mathbf{v}\|^2\|\mathbf{w}\|^2 - \langle \mathbf{v} \mid \mathbf{w} \rangle^2 &= (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_1^2 + \beta_2^2 + \beta_3^2) - (\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3)^2.
\end{aligned}$$

Wenn man beide Ausdrücke ausrechnet, stellt man fest, dass sie identisch sind. Ähnlich zeigt man (P7) und (P8). □

Es gibt elegantere Argumente, um (P5) und (P6) zu zeigen. Diese basieren auf der geometrischen Interpretation des Vektorprodukts, die wir nun darstellen möchten:

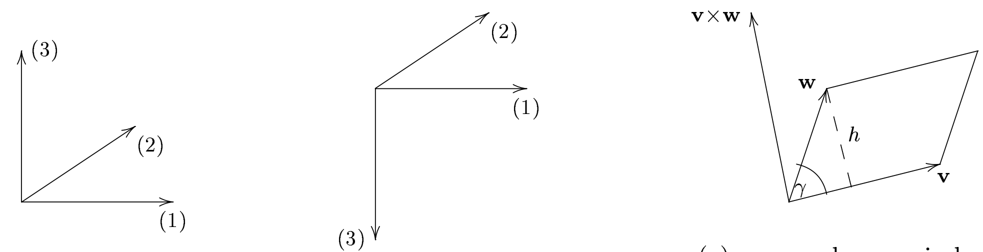
5.1.3 Geometrische Interpretation des Vektorprodukts. Sei γ der Winkel zwischen den Vektoren $\mathbf{v}$ und $\mathbf{w}$ mit $0 \leq \gamma \leq \pi$. Dann ist $\mathbf{v} \times \mathbf{w}$ durch folgende Eigenschaften vollständig festgelegt:

(1) $\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \sin\gamma$

Das heißt $\|\mathbf{v} \times \mathbf{w}\|$ ist die Fläche F des Parallelogramms, das durch $\mathbf{v}$ und $\mathbf{w}$ festgelegt ist.

(2) $\mathbf{v} \times \mathbf{w}$ ist orthogonal zu $\mathbf{v}$ und $\mathbf{w}$.

(3) Die Vektoren $\mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w}$ (genommen in dieser Reihenfolge) formen ein Rechtssystem; vgl. Abbildung 5.1.



(a) Rechtssystem

(b) Linkssystem

(c) $\mathbf{v}$, $\mathbf{w}$ und $\mathbf{v}\times\mathbf{w}$ sind rechtsdrehend

Abbildung 5.1

Beweis. (1)

$$
\begin{aligned}
\|\mathbf{v} \times \mathbf{w}\|^2 &= \|\mathbf{v}\|^2 \cdot \|\mathbf{w}\|^2 - \langle \mathbf{v} \mid \mathbf{w}\rangle^2 && \text{nach Lagrange Identität} \\
&= \|\mathbf{v}\|^2 \cdot \|\mathbf{w}\|^2 - (\|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos\gamma)^2 && \text{nach 3.1.1} \\
&= \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 (1 - \cos^2\gamma) \\
&= \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \sin^2\gamma \\
&= (\|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \sin\gamma)^2
\end{aligned}
$$

Daraus folgt $\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \sin\gamma$. Auf der anderen Seite ist die Fläche des Parallelogramms

$$
\begin{aligned}
F &= (\text{Grundseite}) \cdot (\text{Höhe}) \\
&= \|\mathbf{v}\| \cdot h = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \sin\gamma.
\end{aligned}
$$

(2) ist nichts anderes als Eigenschaft (P5) in 5.1.2.

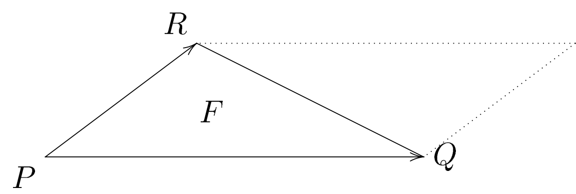
(3) Eigenschaft 3 spiegelt unsere Intuition des Raums, der uns umgibt, wieder. In der aktuellen Form, ist die Aussage nur richtig, wenn wir das kartesische Koordinatensystem so gewählt haben, dass die Basisvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (in dieser Reihenfolge) ein Rechtssystem bilden. □

5.1.4 Folgerung. $\mathbf{v} \times \mathbf{w} = \mathbf{0}$ *genau dann, wenn* $\mathbf{v}$ *und* $\mathbf{w}$ *linear abhängig sind.*

Beweis. Offensichtlich ist $\mathbf{v} \times \mathbf{w} = \mathbf{0}$ genau dann, wenn $\|\mathbf{v} \times \mathbf{w}\| = 0$. Nach 10.3.6(1) ist dies der Fall genau dann, wenn die Fläche des Parallelogramms, das durch $\mathbf{v}$ und $\mathbf{w}$ bestimmt ist, gleich 0 ist. Das wiederum ist der Fall genau dann, wenn das Parallelogramm ‘zusammenfällt’, d.h., wenn $\mathbf{v}$ und $\mathbf{w}$ parallel sind. □

5.1.5 Anwendungen des Vektorprodukts. (1) Berechnung der Flächen von Parallelogrammen und Dreiecken im $\mathbb{R}^3$: Zum Beispiel ist die Fläche eines Dreiecks mit den Ecken P, Q, R gegeben durch

$$F = \tfrac{1}{2} \cdot \|\overrightarrow{PQ} \times \overrightarrow{PR}\|.$$



(2) Berechnung der Flächen von Parallelogrammen und Dreiecken im $\mathbb{R}^2$: Um die Fläche F eines Parallelogramms zu berechnen, das durch die Vektoren $\mathbf{v} = \left(\begin{smallmatrix}\alpha_1\\ \alpha_2\end{smallmatrix}\right)$ und $\mathbf{w} = \left(\begin{smallmatrix}\beta_1\\ \beta_2\end{smallmatrix}\right)$ im $\mathbb{R}^2$ bestimmt wird, betten wir sie in den $\mathbb{R}^3$ ein, indem wir die dritte Koordinate als Null annehmen, und betrachten

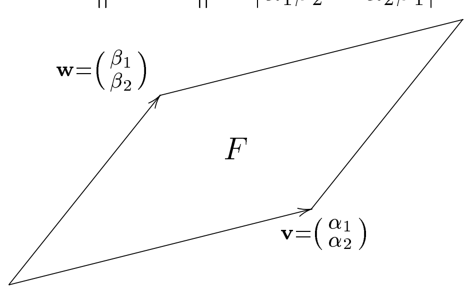
$$\widetilde{\mathbf{v}} = \left(\begin{smallmatrix}\alpha_1\\ \alpha_2\\ 0\end{smallmatrix}\right) \quad \text{and} \quad \widetilde{\mathbf{w}} = \left(\begin{smallmatrix}\beta_1\\ \beta_2\\ 0\end{smallmatrix}\right)$$

anstelle von $\mathbf{v}$ und $\mathbf{w}$. Dann gilt

$$\widetilde{\mathbf{v}} \times \widetilde{\mathbf{w}} = \begin{pmatrix} 0 \\ 0 \\ \alpha_1\beta_2 - \alpha_2\beta_1 \end{pmatrix}$$

Es folgt erstens

$$F = \|\widetilde{\mathbf{v}} \times \widetilde{\mathbf{w}}\| = |\alpha_1\beta_2 - \alpha_2\beta_1|.$$



und zweitens (da $\mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w}$ ein Rechtssystem bilden)

$\alpha_1\beta_2 - \alpha_2\beta_1 > 0$ genau dann, wenn $\mathbf{v}, \mathbf{w}$ positiv orientiert sind,

$\alpha_1\beta_2 - \alpha_2\beta_1 < 0$ genau dann, wenn $\mathbf{v}, \mathbf{w}$ negativ orientiert sind,

$\alpha_1\beta_2 - \alpha_2\beta_1 = 0$ genau dann, wenn $\mathbf{v}, \mathbf{w}$ linear abhängig sind.

Dabei sagen wir, $\mathbf{v}, \mathbf{w}$ seien positiv orientiert, falls gilt: Dreht man $\mathbf{v}$ auf kürzestem Wege zum Vektor $\mathbf{w}$ hin, so ist dies entgegen dem Uhrzeigersinn.

(3) Berechnung eines Normalenvektors $\mathbf{n}$ für eine Ebene E, die durch die drei Punkte P, Q, R gegeben ist:

$$\mathbf{n} = \overrightarrow{PQ} \times \overrightarrow{PR}.$$

Die Ebene E kann nun durch die Gleichung

$$\langle \mathbf{n} \mid \mathbf{x} \rangle = \langle \mathbf{n} \mid \overrightarrow{OP} \rangle$$

beschrieben werden.

5.2. Die Determinante als „orientierte Fäche“ in $\mathbb{R}^2$

In diesem Abschnitt arbeiten wir in der euklidischen Ebene $\mathbb{R}^2$ versehen mit dem Standardskalarprodukt wie in Kapitel 1. Wir nehmen an, die Standardbasis von $\mathbb{R}^2$ sei positiv orientiert.

5.2.1 Die Determinante eines Paares von Vektoren in $\mathbb{R}^2$. Für jedes Paar $u = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$ und $v = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$ von Vektoren in $\mathbb{R}^2$ definieren wir

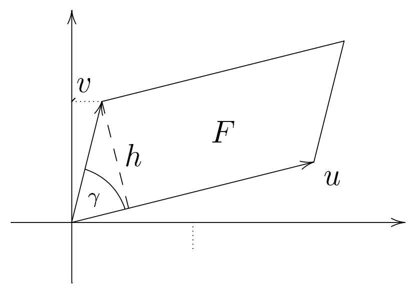
$$\det(u, v) := \alpha_1\beta_2 - \alpha_2\beta_1 \,.$$

5.2.2 Interpretation als orientierte Fläche. Aus 5.1.5(2) erhalten wir

(1)

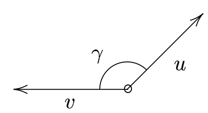
$$F = |\det(u, v)|$$

ist der Flächeninhalt F des von u und v aufgespannten Parallelogramms.

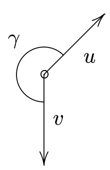


(2)

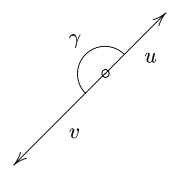
$$\det(u, v) > 0 \iff u, v \text{ bilden eine positiv orientierte Basis von } \mathbb{R}^2$$



$$\det(u, v) < 0 \iff u,\ v \text{ bilden eine negativ orientierte Basis von } \mathbb{R}^2$$



$$\det(u, v) = 0 \iff u, v \text{ sind linear abhängig.}$$



5.2.3 Regeln.

(D1)
$$\begin{aligned} \det(u + u',\, v) &= \det(u, v) + \det(u',\, v) \quad , \quad \det(\lambda u,\, v) = \lambda \det(u, v) \\ \det(u,\, v + v') &= \det(u, v) + \det(u,\, v') \quad , \quad \det(u,\, \lambda v) = \lambda \det(u, v) \end{aligned}$$

In anderen Worten: Für festes v hängt $\det(u, v)$ linear von u ab, für festes u hängt $\det(u, v)$ linear von v ab.

(D2) $\det(u, u) = 0$ für alle $u \in \mathbb{R}^2$

(D3) $\det(e_1, e_2) = 1$

Übung. Wie kann man die Eigenschaft (D1) geometrisch interpretieren?

5.2.4 Die Determinante einer 2×2-Matrix.

Eine reelle 2×2-Matrix $A = \begin{pmatrix} \overbrace{\alpha_{11}}^{s_1} & \overbrace{\alpha_{12}}^{s_2} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ besteht aus zwei Spaltenvektoren s_1, s_2.
Wir definieren die Determinante von A als die Determinante der beiden Spaltenvektoren:

$$\det\, A := \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} := \det(s_1, s_2) := \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}$$

Man beachte, dass die Determinante einer Matrix immer ein Skalar ist. Wir haben also eine ‘Determinantenabbildung’

$$\begin{aligned} \det : \mathcal{M}(2 \times 2, \mathbb{R}) &\longrightarrow \mathbb{R} \\ A &\longmapsto \det A \end{aligned}$$

Diese Abbildung hat nach 5.4.3 die folgenden Eigenschaften:

5.2.5 Regeln für die Determinante einer Matrix.

(D1) $\begin{vmatrix} \alpha_{11} + \alpha'_{11} & \alpha_{12} \\ \alpha_{21} + \alpha'_{21} & \alpha_{22} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} + \begin{vmatrix} \alpha'_{11} & \alpha_{12} \\ \alpha'_{21} & \alpha_{22} \end{vmatrix}$, $\begin{vmatrix} \lambda\alpha_{11} & \alpha_{12} \\ \lambda\alpha_{21} & \alpha_{22} \end{vmatrix} = \lambda \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}$ $\begin{vmatrix} \alpha_{11} & \alpha_{12} + \alpha'_{12} \\ \alpha_{21} & \alpha_{22} + \alpha'_{22} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha'_{12} \\ \alpha_{21} & \alpha'_{22} \end{vmatrix}$, $\begin{vmatrix} \alpha_{11} & \lambda\alpha_{12} \\ \alpha_{21} & \lambda\alpha_{22} \end{vmatrix} = \lambda \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}$

D.h., wenn wir eine der Spalten von A festhalten, hängt $\det A$ linear von der anderen Spalte ab.

(D2) Sind die beiden Spalten von A gleich, ist $\det A = 0$.

(D3) $\det E_2 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$

Da eine 2×2-Matrix A nach Satz 5.2.3(b) genau dann invertierbar ist, wenn ihre Spalten linear unabhängig sind, erhalten wir:

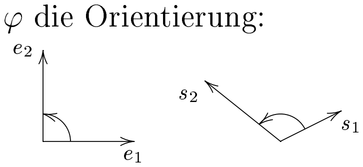
(D2’) A ist genau dann invertierbar, wenn $\det A \neq 0$.

Bemerkung: Für eine komplexe 2×2-Matrix A kann man die Determinante $\det A$ genau wie in 5.2.4 definieren. Die Regeln 5.2.5 bleiben gültig. Nur die geometrische Interpretation der Determinante in 5.2.2 ist dann nicht mehr sinnvoll.

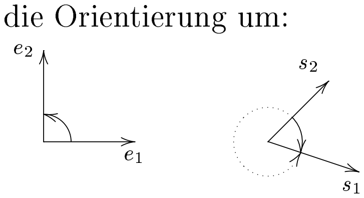
5.2.6 Geometrische Bedeutung der Determinante für Endomorphismen des $\mathbb{R}^2$.

Sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ eine lineare Abbildung und A die Matrix von φ bzgl. der Standardbasis. Die Spalten von A sind $s_1 := \varphi(e_1)$ und $s_2 := \varphi(e_2)$.

Ist $\det A > 0$, dann erhält φ die Orientierung:



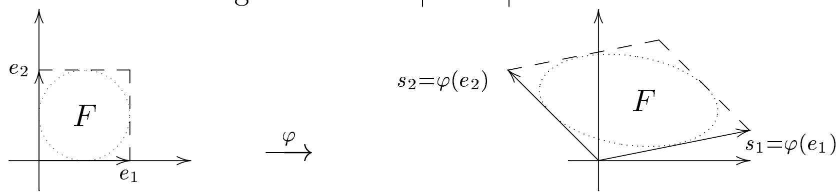
Ist $\det A < 0$, dann kehrt φ die Orientierung um:



$$A = \begin{pmatrix} \overbrace{\alpha_{11}}^{s_1} & \overbrace{\alpha_{12}}^{s_2} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

Beispiele von orientierungserhaltenden Abbildungen : Drehungen.
Beispiele von orientierungsumkehrenden Abbildungen : Spiegelungen.

Weiterhin ist das Bild des Einheitsquadrates (aufgespannt von e_1 und e_2) das von den Spalten $s_1 = \varphi(e_1)$ und $s_2 = \varphi(e_2)$ von A aufgespannte Parallelogramm. Also ist dder Flächeninhalt dieses Parallelogramms $F = |\det A|$ nach 5.2.2:



Das Bild des in das Einheitsquadrat einbeschriebenen Kreises ist eine in das Parallelogramm einbeschriebene Ellipse. Ist F der Flächeninhalt des Kreises, beträgt der Flächeninhalt F' der Ellipse

$$F' = \det A \cdot F\,.$$

In der Vorlesung Analysis III werden Sie lernen, wie man den Flächeninhalt F von weitaus allgemeineren Teilmengen $M \subseteq \mathbb{R}^2$ berechnet. Betrachten wir das Bild von M unter einer linearen Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, ist dann die Fläche F' des Bildes $\varphi(M)$ von M

$$F' = \det A \cdot F.$$

Das bedeutet, der Flächeinhalt des Bildes $\varphi(M)$ einer Teilmenge M von $\mathbb{R}^2$ unter der durch die Matrix A gegebenen linearen Abbildung ist das $\det A$-fache des Flächeninhalts von M.

5.3. Die Determinante als „orientiertes Volumen" in $\mathbb{R}^3$

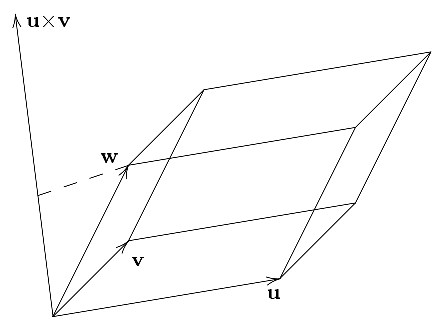
In diesem Abschnitt wollen wir ganz analog zu dem vorigen vorgehen. Im Euklidischen Raum $\mathbb{R}^3$ haben wir das Vektorprodukt (siehe 5.1) eingeführt. Unter Verwendung von Determinanten von 2×2-Matrizen, können wir das Vektorprodukt in der folgenden Weise schreiben:

$$u \times v = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \alpha_2\beta_3 - \alpha_3\beta_2 \\ -(\alpha_1\beta_3 - \alpha_3\beta_1) \\ \alpha_1\beta_2 - \alpha_2\beta_1 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{vmatrix} \\ -\begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_3 & \beta_3 \end{vmatrix} \\ \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \end{pmatrix}$$

5.3.1 Die Determinante dreier Vektoren im $\mathbb{R}^3$. Die Determinante dreier Vektoren $u, v, w \in \mathbb{R}^3$ ist definiert als

$$\det(u, v, w) := \langle u \times v \mid w \rangle.$$

Auch die Determinante dreier Vektoren des $\mathbb{R}^3$ ist eine reelle Zahl. Manchmal wird dies auch das *Spatprodukt* (oder auch das *gemischte Produkt*) der drei Vektoren genannt. Der Name *Spatprodukt* kommt von der geometrischen Interpretation: Ein *Spat/Parallelflach* [*parallelepipedon*] ist das 3-dimensionale Analogon eines Parallelogramms. Die Abbildung zeigt den Spat, der durch drei Vektoren bestimmt ist.



Unter Benutzung der soeben für das Vektorprodukt eingeführten Schreibweise erhalten wir

$$\begin{aligned}
\det(u,v,w) &= \begin{vmatrix} \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{vmatrix} \cdot \gamma_1 - \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_3 & \beta_3 \end{vmatrix} \cdot \gamma_2 + \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \cdot \gamma_3 \\
&= \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 + \alpha_1\beta_2\gamma_3 - \alpha_3\beta_2\gamma_1 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3
\end{aligned}$$

Diesen Ausdruck kann man sich über die „Regel von Sarrus“ merken:

$$\left|\begin{array}{ccc|cc} \alpha_1 & \beta_1 & \gamma_1 & \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 & \gamma_2 & \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 & \gamma_3 & \alpha_3 & \beta_3 \end{array}\right.$$

5.3.2 Interpretation der Determinante als ‘orientiertes Volumen’.

(1) Der Betrag $|\det u, v, w|$ ist das Volumen V des von den Vektoren u, v, w aufgespannten Spats.

(2) $\det(u,v,w) > 0 \iff u, v, w$ bilden ein Rechtssystem.
$\det(u,v,w) < 0 \iff u, v, w$ bilden ein Linkssystem.

(3) $\det(u,v,w) = 0 \iff u, v, w$ sind linear abhängig.

Beweis. Sei γ der Winkel zwischen $u \times v$ und w mit $0 \le \gamma \le \pi$. Mit 3.1.1 gilt

$$\det(u,v,w) = \langle u \times v \mid w \rangle = \|u \times v\| \cdot \|w\| \cdot \cos\gamma.$$

(1) Das Volumen des Spats ist

$$V = (\text{Fläche } F \text{ des Basisparallelogramms}) \cdot (\text{Höhe } h).$$

Nach 10.3.6 ist $F = \|u \times v\|$ und, da $u \times v$ senkrecht auf dem von u und v aufgespannten Parallelogramm steht, ist $h = \|w\| \cdot |\cos\gamma|$. Damit gilt $V = F \cdot h = \|u \times v\| \cdot \|w\| \cdot |\cos\gamma| = |\langle u \times v \mid w \rangle| = \det(u,v,w)$.

(2) $\det(u,v,w) = \langle u \times v \mid w \rangle > 0$ genau dann, wenn $u \times v \neq 0$, $w \neq 0$ und $\cos\gamma > 0$. Letzteres ist der Fall genau dann, wenn $0 \le \gamma < \frac{\pi}{2}$. Da $u, v, u \times v$ ein Rechtssystem bilden, folgt dasselbe für u, v, w.

Im Falle $\det(u,v,w) = \langle u \times v \mid w \rangle < 0$ hat man $\cos\gamma < 0$, d.h. $\frac{\pi}{2} < \gamma \le \pi$. Das heißt, u, v, w ist ein Linkssystem.

(3) $\det(u,v,w) = \langle u \times v \mid w \rangle = 0$ genau dann, wenn das Volumen V des Spats Null ist, d.h. genau dann, wenn der Spat zusammenfällt. Das bedeutet, daß u, v, w in einer Ebene liegen, was wiederum bedeutet, daß sie linear abhängig sind.

5.3.3 Regeln.

(D1)

$$\begin{array}{rclcll} \det(u+u',v,w) &=& \det(u,v,w) &+& \det(u',v,w), & \det(\lambda u,v,w) = \lambda\det(u,v,w)\\ \det(u,v+v',w) &=& \det(u,v,w) &+& \det(u,v',w), & \det(u,\lambda v,w) = \lambda\det(u,v,w)\\ \det(u,v,w+w') &=& \det(u,v,w) &+& \det(u,v,w'), & \det(u,v,\lambda w) = \lambda\det(u,v,w)\end{array}$$

Halten wir also zwei der Vektoren u, v, w fest, dann hängt $\det(u, v, w)$ linear vom dritten Vektor ab.

(D2) Sind zwei der Vektoren u, v, w gleich, gilt $\det(u, v, w) = 0$.

(D3) $\det(e_1, e_2, e_3) = 1$.

Beweis.

(D1)

$$\begin{array}{rcl} \det(u+u',v,w) &=& \langle (u+u')\times v \mid w\rangle \underset{\textbf{??}}{=} \langle u\times v + u'\times v \mid w\rangle = \\ &=& \langle u\times v \mid w\rangle + \langle u', \times v \mid w\rangle \\ &=& \det(u,v,w) + \det(u',v,w)\end{array}$$

Die anderen Eigenschaften aus 5.3.3 (D1) werden auf ähnliche Weise aus den Eigenschaften des Vekorproduktes und des Skalarproduktes hergeleitet.

(D2) Sind zwei der Vektoren u, v, w gleich, dann sind u, v, w linear abhängig. Also folgt (D2) aus Teil **??** (b) des vorigen Satzes.

(D3) $\det(e_1, e_2, e_3) = \langle e_1 \times e_2 \mid e_3\rangle = \langle e_3 \mid e_3\rangle = 1$.

□

5.3.4 Die Determinante einer 3×3-Matrix.

Betrachte die Matrix

$$A = \begin{pmatrix} \overbrace{\alpha_{11}}^{s_1} & \overbrace{\alpha_{12}}^{s_2} & \overbrace{\alpha_{13}}^{s_3} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{pmatrix}$$

Wir definieren $\det A := \det(s_1, s_2, s_3)$ als die Determinante der Spaltenvektoren von A.

Genauer gilt:

$$\det A := \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{vmatrix} := \begin{array}{rl} & \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} \\ - & \alpha_{13}\alpha_{22}\alpha_{31} - \alpha_{12}\alpha_{21}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{33}\end{array}$$

Wir können auch die in **??** gegebenen Formeln oder die Regel von Sarrus benutzen.

Ordnen wir der 3×3-Matrix A ihre Determinante zu, so erhalten wir eine Abbildung

$$\begin{array}{rcl} \det : \mathcal{M}(3\times 3, \mathbb{R}) &\to& \mathbb{R} \\ A &\mapsto& \det A \end{array}$$

Aus 5.3.3 folgen

5.3.5 Regeln für die Determinante einer 3×3-Matrix A.

(D1) $\det A$ hängt linear von jeder Spalte von A ab.

(D2) Sind zwei Spalten von A gleich, ist $\det A = 0$.

(D3) $\det E_3 = 1$.

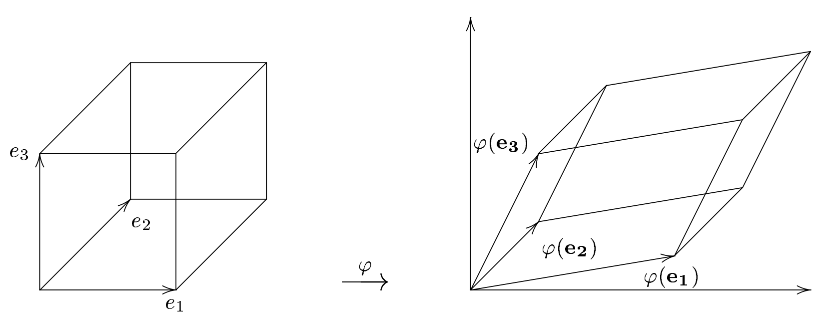
Da eine 3×3-Matrix A genau dann invertierbar ist, wenn ihre Spalten linear unabhängig sind (siehe Satz **??** (b)), erhalten wir:

(D2') A ist genau dann invertierbar, wenn $\det A \neq 0$.

Bemerkung: Wir können die Determinante einer 3×3-Matrix A mit komplexen Einträgen mit derselben Formel wie in 5.3.4 berechnen. Alle Eigenschaften aus 5.3.5 bleiben gültig. Wir möchten den Begriff der Determinante auf beliebige $n \times n$-Matrizen A mit reellen oder komplexen Einträgen in solcher Weise ausdehnen, dass alle diese Eigenschaften weiterhin gültig bleiben.

5.3.6 Geometrische Bedeutung der Determinante eines Endomorphismus $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$.

Für eine gegebene lineare Abbildung $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ betrachte deren Matrix A bzgl. der Standardbasis. Die Spalten von A sind $s_1 := \varphi(e_1)$, $s_2 := \varphi(e_2)$, $s_3 := \varphi(e_3)$.



Der Einheitswürfel (mit Volumen 1) wird auf den Spat abgebildet, der von $s_1 = \varphi(e_1)$, $s_2 = \varphi(e_2)$, $s_3 = \varphi(e_3)$, d.h. von den Spalten von A aufgespannt wird. Das Volumen des Spats ist $|\det(s_1, s_2, s_3)| = |\det A|$ nach **??** (a).

Allgemeiner gilt: Ist M eine Teilmenge von $\mathbb{R}^3$ mit Volumen V, dann hat deren Bild unter der linearen Abbildung φ das Volumen

$$V' = |\det A| \cdot V,$$

wobei A die Matrix von φ bzgl. der Standardbasis ist.

Aus **??** (b) erhalten wir:

$$\det A > 0 \iff \varphi \text{ ist orientierungserhaltend}$$
$$\det A < 0 \iff \varphi \text{ ist orientierungsumkehrend}$$

5.4. Die Determinante einer $n \times n$-Matrix

Ab jetzt betrachten wir wieder simultan die Fälle $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

Wir betrachten $n \times n$-Matrizen A mit Einträgen $\alpha_{ij} \in \mathbb{K}$. Wir bezeichnen die Spaltenvektoren von A mit $s_1, s_2, \ldots, s_n$ und schreiben

$$A = (s_1, s_2, \ldots, s_n).$$

$$A = \begin{array}{c} \begin{matrix} \overbrace{\phantom{\alpha_{11}}}^{s_1} & \overbrace{\phantom{\alpha_{12}}}^{s_2} & \ldots & \overbrace{\phantom{\alpha_{11}}}^{s_n} \end{matrix} \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{11} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix} \end{array}$$

Es ist nicht offensichtlich, wie man den Begriff der Determinante von 2×2- und 3×3-Matrizen auf beliebiges n verallgemeinern soll. Wir machen einen axiomatischen Ansatz und verlangen, daß die Determinante von $n \times n$-Matrien wenigstens die grundlegenden Eigenschaften (D1), (D2) und (D3) aus 5.2.5 und 5.3.5 besitzen soll:

5.4.1 Determinantenabbildungen. Eine Funktion $\Delta : \mathcal{M}(n \times n, \mathbb{K}) \to \mathbb{K}$ wird *Determinantenabbildung* genannt, wenn sie die folgenden Eigenschaften erfüllt:

(D1) $\Delta(A)$ hängt linear von jeder Spalte von A ab, wenn die anderen Spalten fest bleiben; in anderen Worten, für jedes j gilt

$$\begin{aligned} \Delta(\ldots, s_j + s'_j, \ldots) &= \Delta(\ldots, s_j, \ldots) + \Delta(\ldots, s'_j, \ldots) \\ \Delta(\ldots, \lambda \cdot s_j, \ldots) &= \lambda \cdot \Delta(\ldots, s_j, \ldots) \end{aligned}$$

(D2) Sind zwei Spalten von A gleich, ist $\Delta(A) = 0$.

(D3) $\Delta(E_n) = 1$

Es ist nicht klar, daß Determinantenabbildungen für $n > 3$ existieren. Wir werden zeigen, daß für jedes n genau eine Determinantenabbildung existiert. Dies bedeutet insbesondere, daß die für $2 \times 2-$ und $3 \times 3-$Matrizen in 5.2.4 und 5.3.4 definierten Determinanten nur auf diese Weise definiert werden konnten, wenn man die Gültigkeit der Eigenschaften (D1), (D2) und (D3) erreichen will:

5.4.2 Theorem. [Hauptsatz der Determinantentheorie] *Es gibt eine und nur eine Determinantenabbildung* $\Delta : \mathcal{M}(n \times n, \mathbb{K}) \to \mathbb{K}$. *Überdies hat diese die Eigenschaft*

(D4) $$\Delta(A^T) = \Delta(A) \quad \textit{für alle } A \in \mathcal{M}(n \times n, \mathbb{K})$$

Wir werden die Existenz einer Determinantenabbildung später in Abschnitt 5.7 beweisen. Wir gehen nun von der Existenz aus und ziehen einige Folgerungen aus den Eigenschaften (D1) - (D4). Anstelle von $\Delta(A)$ schreiben wir:

$$\det(A) := \begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix} := \Delta(A)$$

5.4.3 Eigenschaften der Determinante.

(D1) $\det(A)$ hängt linear von den Spalten von A ab; in anderen Worten:

$$\begin{vmatrix} \alpha_{11}+\alpha'_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21}+\alpha'_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1}+\alpha'_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \begin{vmatrix} \alpha'_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha'_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha'_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

$$\begin{vmatrix} \lambda\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \lambda\alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda\alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \lambda \cdot \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

und analog für die anderen Spalten.

(D2) Sind zwei Spalten von A gleich, ist $\det A = 0$.

(D3) $\det(E_n) = 1$.

(D4) $\det(A^T) = \det(A)$, wobei A^T die Transponierte von A ist.

Da die Zeilen von A die Spalten von A^T sind, folgt aus (D4), dass (D1), (D2) ebenso wie die folgenden Eigenschaften wahr bleiben, wenn wir das Wort „Spalte“ durch „Zeile“ ersetzen.

(D5) Hat A eine nur aus Nullen bestehende Spalte, ist $\det A = 0$.

(D6) Die Addition eines skalaren Vielfachen von einer Spalte zu einer anderen Spalte verändert nicht den Wert der Determinante.

(D7) Das Vertauschen zweier Spalten ändert nur das Vorzeichen der Determinante.

(D8) Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente:

$$\begin{vmatrix} \alpha_{11} & * & \dots & * \\ 0 & \alpha_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \alpha_{nn} \end{vmatrix} = \alpha_{11} \cdot \alpha_{22} \cdot \ldots \cdot \alpha_{nn}$$

(D9) $\det A \neq 0 \Longleftrightarrow A$ ist invertierbar.

(D10) $\det(A \cdot B) = \det A \cdot \det B$ (Produktsatz)

(D11) Ist A invertierbar, dann gilt $\det(A^{-1}) = \dfrac{1}{\det A}$.

(D12) $\det \left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array} \right) = \det A \cdot \det B$

Wir erkennen, dass (D1) - (D3) nur Wiederholungen der Eigenschaften sind, die eine Determinantenabbildung nach Definition 5.4.1 hat. Wir zeigen, dass die Eigenschaften (D5) - (D8) aus (D1) - (D4) abgeleitet werden können. Da diese Eigenschaften es erlauben, die Determiannte einer jeden Matrix zu berechnen, muß die Determinantenform eindeutig sein, sofern sie existiert. Zum Beweis von (D9) - (D12) können wir also die Eindeutigkeitsaussage benutzen.

(D5) $\det(\dots,0,\dots) = \det(\dots,0\cdot 0,\dots) \underset{(D1)}{=} 0\cdot\det(\dots,0,\dots) = 0.$

(D6) Sei $i<j$:

$$\begin{aligned}
&\det(\dots,s_i,\dots,s_j+\lambda s_i,\dots) = \\
\underset{(D1)}{=}\ &\det(\dots,s_i,\dots,s_j,\dots) + \det(\dots,s_i,\dots,\lambda s_i,\dots) \\
\underset{(D1)}{=}\ &\det(\dots,s_i,\dots,s_j,\dots) + \lambda\cdot\underbrace{\det(\dots,s_i,\dots,s_i,\dots)}_{=\,0\text{ nach (D2)}}
\end{aligned}$$

(D7) Sei $i<j$:

$$\begin{aligned}
&\det(\dots,s_i,\dots,s_j,\dots) + \det(\dots,s_j,\dots,s_i,\dots) \\
\underset{(D6)}{=}\ &\det(\dots,s_i+s_j,\dots,s_j,\dots) + \det(\dots,s_j+s_i,\dots,s_i,\dots) \\
\underset{(D1)}{=}\ &\det(\dots,s_i+s_j,\dots,s_j+s_i,\dots) \underset{(D2)}{=} 0
\end{aligned}$$

also gilt $\det(\dots,s_i,\dots,s_j,\dots) = -\det(\dots,s_j,\dots,s_i,\dots)$

(D8) Durch Subtraktion geeigneter skalarer Vielfacher der ersten Spalte von den anderen Spalten erhält man Nullen in der ersten Zeile ohne den Wert der Determinante zu ändern:

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{vmatrix} \underset{(D1)}{=} \alpha_{11}\cdot \begin{vmatrix} 1 & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{vmatrix} \underset{(D6)}{=} \alpha_{11}\cdot \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{vmatrix}$$

Nun wiederholt man diese Schritte mit $\alpha_{22},\dots$ etc.

$$= \dots = \alpha_{11}\cdot\alpha_{22}\cdot\ldots\cdot\alpha_{nn}\cdot \begin{vmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{vmatrix} \underset{(D3)}{=} \alpha_{11}\alpha_{22}\dots\alpha_{nn}$$

(D9) Sei r der Rang von A. Mit den elementaren Zeilentransformationen

(E1) Addition eines skalaren Vielfachen einer Zeile zu einer anderen;

(E2) Vertauschen zweier Zeilen:

können wir A in eine Matrix in Dreiecksgestalt (Algorithmus von Gauß-Jordan) umformen:

$$A' = \begin{pmatrix} \alpha'_{11} & \alpha'_{12} & * & \dots & \dots & \dots & * \\ 0 & \alpha'_{22} & * & \dots & \dots & \dots & * \\ & & \ddots & & & \vdots & \\ & & & \alpha'_{rr} & * & \dots & * \\ & 0 & & & & \dots & \\ & & & & & \dots & \end{pmatrix}$$

Anwendung von (D6) und (D7) auf Zeilen anstatt auf Spalten zeigt, daß die elmentaren Zeilenumformungen (E1) und (E2) die Determinante unverändert lassen bis auf evtl. das Vorzeichen, d.h.,

$$\det A = \pm \det A' .$$

- Erster Fall: $r < n$. Mindestens eine Zeile von A' hat nur Nullen (nach (D5)). Es folgt $\det A' = 0$ und damit $\det A = 0$.
- Zweiter Fall: $r = n$. Dann ist A' von der Form

$$A' = \begin{pmatrix} \alpha'_{11} & * & \dots & \dots & * \\ 0 & \alpha'_{22} & * & \dots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \dots & \dots\ 0 & & \alpha'_{nn} \end{pmatrix}, \text{ wobei } \alpha'_{ii} \neq 0 \text{ für alle } i .$$

Damit ist $\det A = \pm \det A' \underset{(D8)}{=} \pm \alpha'_{11} \cdot \alpha'_{22} \cdot \ldots \cdot \alpha'_{nn} \neq 0$.

Da genau dann $r = n$ gilt, wenn A invertierbar ist, haben wir (D9) gezeigt.

(D10) Erster Fall: A ist nicht invertierbar. Dann ist $A \cdot B$ nicht invertierbar. Aus (D9) folgt dann $\det A = 0$ und $\det(AB) = 0$. Also gilt $\det(AB) = 0 = \det A \cdot \det B$.

Zweiter Fall: A ist invertierbar. Dann folgt $\det A \neq 0$ aus (D9). Also können wir für jedes $B \in \mathcal{M}(n \times n, \mathbb{K})$ definieren

$$\Delta(B) := \frac{\det(AB)}{\det A}$$

Behauptung: $\Delta : \mathcal{M}(n \times n, \mathbb{K}) \to \mathbb{K}$ ist eine Determinantenabbildung.

Beweis.

(D3) $\Delta(E_n) = \frac{\det(AE_n)}{\det A} = \frac{\det A}{\det A} = 1$.

(D2) Es seien zwei Spalten von B gleich. Dann sind die Spalten von B linear abhängig und damit B nicht invertierbar. Daraus folgt, dass AB nicht invertierbar ist. Aus (D 9) folgern wir $\det(AB) = 0$. Es ergibt sich

$$\Delta(B) = \frac{\det(AB)}{\det A} = 0$$

(D1) Seien $t_1, \ldots, t_n$ die Spalten von B, d.h.

$$B = (t_1, t_2, \ldots, t_n).$$

Man erinnere sich, dass $AB = (At_1, At_2, \dots, At_n)$ gilt.

Wir zeigen, dass $\Delta(B)$ linear von der ersten Spalte von B abhängt, z.B.:

$$\begin{aligned}
\Delta(t_1 + t_1', t_2, \dots, t_n) &= \frac{1}{\det A} \det(A(t_1 + t_1', t_2, \dots, t_n)) \\
&= \frac{1}{\det A} \det(A(t_1 + t_1'), At_2, \dots, At_n) \\
&= \frac{1}{\det A} \det(At_1 + At_1', At_2, \dots, At_n) \\
&\underset{(\mathrm{D1})}{=} \frac{1}{\det A} \big(\det(At_1, At_2, \dots, At_n) \\
&\qquad + \det(At_1', At_2, \dots, At_n)\big) \\
&= \Delta(t_1, t_2, \dots, t_n) + \Delta(t_1', t_2, \dots, t_n) \\
\Delta(\lambda t_1, t_2, \dots, t_n) &= \frac{1}{\det A} \det(A(\lambda t_1, t_2, \dots, t_n)) \\
&= \frac{1}{\det A} \det(A(\lambda t_1), At_2, \dots, At_n) \\
&= \frac{1}{\det A} \det(\lambda At_1, At_2, \dots, At_n) \\
&= \frac{\lambda}{\det A} \det(At_1, At_2, \dots, At_n) \\
&= \lambda \frac{1}{\det A} \det(A(t_1, t_2, \dots, t_n)) \\
&= \lambda \Delta(t_1, \dots, t_n).
\end{aligned}$$

Damit ist unsere Behauptung bewiesen. Aufgrund des Hauptsatzes 5.4.2 ist die Determinantenabbildung eindeutig und damit $\Delta(B) = \det B$ für alle $B \in \mathcal{M}(n \times n, \mathbb{K})$, d.h.

$$\frac{\det(AB)}{\det A} = \det B$$

woraus $\det(AB) = \det A \cdot \det B$ folgt.

□

(D11) Sei A invertierbar. Dann gilt $\det A \cdot \det(A^{-1}) \underset{(\mathrm{D10})}{=} \det(AA^{-1}) = \det E \underset{(\mathrm{D3})}{=} 1$.

Es folgt

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

(D12) Wir beweisen zunächst den Spezialfall:

$$\det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & B \end{array} \right) = \det B$$

Beachte, dass $\det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & B \end{array}\right) = \det \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right)$ gilt: Subtrahiere geeignete skalare Vielfache der ersten Spalte von den anderen Spalten, was nach (D6) die Determinante nicht ändert. Nun definieren wir

$$\Delta(B) := \det \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right) \qquad \text{für alle } B \in \mathcal{M}(n{-}1 \times n{-}1, \mathbb{K}) \,.$$

Man zeigt leicht, dass $\Delta(B)$ eine Determinantenabbildung auf $\mathcal{M}(n{-}1 \times n{-}1, \mathbb{K})$ ist. Da eine solche Abbildung eindeutig ist, folgern wir $\Delta(B) = \det B$, d.h.

$$\det \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right) = \det B \,.$$

Nach mehrfacher Wiederholung dieses Vorgehens erkennen wir:

$$\det \left(\begin{array}{c|c} E & 0 \\ \hline 0 & B \end{array}\right) = \det B \qquad \text{und analog} \qquad \det \left(\begin{array}{c|c} A & 0 \\ \hline 0 & E \end{array}\right) = \det A$$

Nun folgt der allgemeine Fall aus der Tatsache, dass

$$\left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right) = \left(\begin{array}{c|c} E & 0 \\ \hline 0 & B \end{array}\right) \cdot \left(\begin{array}{c|c} A & * \\ \hline 0 & E \end{array}\right)$$

nach den Regeln der blockweisen Matrixmultiplikation gilt. Damit folgt nach (D10)

$$\det \left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right) = \det \left(\begin{array}{c|c} E & 0 \\ \hline 0 & B \end{array}\right) \cdot \det \left(\begin{array}{c|c} A & * \\ \hline 0 & E \end{array}\right) = \det B \cdot \det A = \det A \cdot \det B.$$

Bemerkungen.

(i) Im Allgemeinen ist $\det(A + B) \neq \det A + \det B$. Beispielsweise gilt:

$$1 = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det \Bigg(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{B} \Bigg), \text{ aber}$$

$$0 = \det \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A} + \det \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{B}$$

(ii) $\det(\lambda A) = \lambda^n \det A$ für jede $n \times n$-Matrix A.

(iii) Die Regel von Sarrus kann nicht auf 4×4-Matrizen angewendet werden.

5.4.4 Praktische Methode zur Determinantenberechnung. Die Benutzung von elementaren Zeilen- oder Spaltentransformationen erlaubt uns die Umformung der Matrix A in Dreiecksgestalt (Gauß-Jordan-Algorithmus). Aus 5.4.2 (D6) und 5.4.2 (D7) wissen wir, dass dies höchstens das Vorzeichen der Determinante ändert. Aus 5.4.2 (D5) und 5.4.2 (D8) ergibt sich, wie die Determinante einer Matrix in Dreiecksgestalt berechnet wird.

Beispiel.

$$\begin{vmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 2 & 3 \\ 2 & 1 & 0 & 3 \\ 3 & 1 & 1 & 1 \end{vmatrix} \underset{(D7)}{=} - \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 2 & 1 & 0 & 3 \\ 3 & 1 & 1 & 1 \end{vmatrix} \underset{(D6)}{=} \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & -4 & -3 \\ 0 & 1 & -5 & -8 \end{vmatrix} \underset{(D6)}{=} \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -6 & -6 \\ 0 & 0 & -7 & -11 \end{vmatrix}$$

$$\underset{(D1)}{=} (-6) \cdot \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -7 & -11 \end{vmatrix} \underset{(D6)}{=} (-6) \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -4 \end{vmatrix}$$

$$\underset{(D8)}{=} (-6) \cdot 1 \cdot 1 \cdot 1 \cdot (-4) = 24$$

5.4.5 Entwicklung nach einer Zeile oder Spalte.

Sei $A = (\alpha_{ij})$ eine $n \times n$-Matrix. Wir bezeichnen mit M_{ij} die $(n-1) \times (n-1)$-Matrix, die wir aus A erhalten, indem wir die $i-$te Zeile und die $j-$te Spalte streichen.

$$M_{ij} := \begin{array}{cc} & \begin{array}{ccccc} \phantom{\alpha_{11}} & & j & & \phantom{\alpha_{1n}} \\ & & \downarrow & & \end{array} \\ i \rightarrow & \begin{pmatrix} \alpha_{11} & \dots & \alpha_{ij} & \dots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \dots & \alpha_{ij} & \dots & \alpha_{in} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{nj} & \dots & \alpha_{nn} \end{pmatrix} \end{array}$$

Unter Benutzung der Eigenschaften der Determinante erhalten wir folgende explizite Formel

$$\boxed{\det A = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det(M_{ij})} \qquad \text{(Entwicklung nach der } j\text{-ten Spalte)}$$

und analog

$$\boxed{\det A = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det(M_{ij})} \qquad \text{(Entwicklung nach der } i\text{-ten Zeile)}$$

Für $j = 1$ ergibt sich die Entwicklung nach der ersten Spalte zu:

$$\det A = \alpha_{11} \det(M_{11}) - \alpha_{21} \det(M_{21}) + \ldots + (-1)^n \alpha_{n1} \det(M_{n1})$$

Man kann sich leicht das Vorzeichen $(-1)^{i+j}$ mittels des folgenden Schachbrettmusters merken:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

Hier der Beweis der Entwicklungsformel für den Fall $j = 1$:

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

$$\underset{(D1)}{=} \alpha_{11} \cdot \begin{vmatrix} 1 & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \alpha_{21} \cdot \begin{vmatrix} 0 & \alpha_{12} & \dots & \alpha_{1n} \\ 1 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \dots + \alpha_{n1} \cdot \begin{vmatrix} 0 & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \\ 1 & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

$$\underset{(D12,\ D7)}{=} \alpha_{11} \det(M_{11}) - \alpha_{21} \det(M_{21}) + \dots + (-1)^{n+1} \alpha_{n1} \det(M_{n1})$$

Beispiel. Beinhaltet eine Zeile oder Spalte mehrere Nullen, dann ist die Entwicklung nach dieser Zeile oder Spalte hilfreich für die Berechnung der Determinante:

$$\begin{vmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{vmatrix} = +1 \cdot \begin{vmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{vmatrix}$$

$$= - \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} - \underbrace{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{vmatrix}}_{0} - \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

5.4.6 Die Adjunkte. Für eine gegebene $n \times n$-Matrix $A = (\alpha_{ij})$ betrachten wir die $(n-1) \times (n-1)$-Matrizen M_{ij} wie in 5.4.5 definiert und definieren Skalare

$$A_{ij} := (-1)^{i+j} \det(M_{ij}) , \qquad (i = 1, \dots, n;\ j = 1, \dots, n) .$$

Diese Skalare heißen *Kofaktoren* von A. Sie können als Einträge einer neuen Matrix

$$\operatorname{adj} A := (A_{ij})_{ij} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix} \in \mathcal{M}(n \times n, \mathbb{K})$$

aufgefasst werden. Diese Matrix wird *Adjunkte* von A genannt.

Satz.

$$\boxed{(\operatorname{adj} A)^T \cdot A = \det A \cdot E}$$

Beweis.

$$\begin{aligned} & (j\text{-te Zeile von } (\operatorname{adj} A)^T) \cdot (j\text{-te Spalte von } A) \\ = \ & (j\text{-te Spalte von } \operatorname{adj} A) \cdot (j\text{-te Spalte von } A) \\ = \ & \sum_{i=1}^{n} A_{ij} \cdot \alpha_{ij} = \sum_{i=1}^{n} (-1)^{i+j} \det(M_{ij}) \alpha_{ij} \underset{5.4.5}{=} \det A \end{aligned}$$

Damit steht auf der Diagonale der Produktmatrix $(\operatorname{adj} A)^T A$ immer die Zahl $\det A$. Alle anderen nicht auf der Diagonale stehenden Einträge sind Nullen (was den Satz beweist), da für $j \neq k$ gilt:

$$\begin{aligned}
& (j\text{-te Zeile von } (\operatorname{adj} A)^T) \cdot (k\text{-te Spalte von } A) \\
= \; & (j\text{-te Spalte von } \operatorname{adj} A) \cdot (k\text{-te Spalte von } A) \\
= \; & \sum_{i=1}^{n} A_{ij}\alpha_{ik} = \sum_{i=1}^{n}(-1)^{i+j}\det(M_{ij})\alpha_{ik} = \sum_{i=1}^{n}(-1)^{i+j}\alpha_{ik}\det(M_{ij}) \\
= \; & 0
\end{aligned}$$

(Im letzten Schritt beachte man die aus A durch Ersetzung der j-ten Spalte von A durch die k-te Spalte gewonnene Matrix; in diesem Fall stimmen die j-te und die k-te Spalte überein. Also ist die Determinante Null.) □

5.4.7 Eine Formel für die Inverse. Für eine invertierbare $n \times n$-Matrix A gilt:

$$\boxed{A^{-1} = \frac{1}{\det A}(\operatorname{adj} A)^T}$$

Denn ist A invertierbar, dann ist $\det A \neq 0$ nach (D9) und aus dem Satz in 5.4.6 folgern wir

$$\frac{1}{\det A}(\operatorname{adj} A)^T \cdot A = E\,.$$

Multiplikation mit A^{-1} von rechts ergibt das gewünschte Resultat.

Damit haben wir eine explizite Formel für die Inverse erhalten. Man beachte aber, dass diese Formel für praktische Zwecke nutzlos ist, da wir n^2 Determinanten der $(n-1) \times (n-1)$-Matrizen M_{ij} und die Determinante von $\det A$ berechnen müssen. Der Gauß-Jordan-Algorithmus zur Berechnung der Inversen (1.4.8) arbeitet viel schneller.

Beispiele.

(a)
$$A = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \qquad \operatorname{adj} A = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$$

Für eine invertierbare Matrix A gilt

$$A^{-1} = \frac{1}{\det A}(\operatorname{adj} A)^T = \frac{1}{\det A}\begin{pmatrix} \delta & -\gamma \\ -\beta & \alpha \end{pmatrix}$$

(b)
$$A = \begin{pmatrix} \lambda & 0 & 1 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}, \text{also } \operatorname{adj} A = \begin{pmatrix} \lambda^2 & -\lambda & 1 \\ 1 & \lambda^2 & -\lambda \\ -\lambda & 1 & \lambda^2 \end{pmatrix}$$

Eine kurze Rechnung ergibt $\det A = \lambda^3 + 1$. Für alle λ mit $\lambda^3 \neq -1$ gilt $\det A \neq 0$. In diesem Fall ist A invertierbar und es gilt:

$$A^{-1} = \frac{1}{\lambda^2 + 1}\begin{pmatrix} \lambda^3 & 1 & -\lambda \\ -\lambda & \lambda^2 & 1 \\ 1 & -\lambda & \lambda^2 \end{pmatrix}$$

5.4.8 Die CRAMERsche Regel. Sei A eine *invertierbare* $n \times n$-Matrix und $b \in \mathbb{K}^n$. Betrachte das folgende inhomogene System von n linearen Gleichungen in n Unbekannten

$$\text{(L)} \qquad Ax = b.$$

Wir wissen aufgrund der Invertierbarkeit von A, dass dieses System eine eindeutige Lösung

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

hat, nämlich $x = A^{-1}b$. Aus 5.4.7 wissen wir $x = \frac{1}{\det A}(\operatorname{adj} A)^T b$, also

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

Daraus folgt

$$\begin{aligned} x_i &= \frac{1}{\det A}(A_{1i}\beta_1 + A_{2i}\beta_2 + \ldots + A_{ni}\beta_n) \\ &= \frac{1}{\det A}\left((-1)^{i+1}\beta_1 \det M_{1i} + \ldots + (-1)^{i+n}\beta_n \det M_{ni}\right) \\ &= \frac{\det B_i}{\det A} \qquad \text{für } i = 1, \ldots, n \end{aligned}$$

wobei B_i die aus A durch Ersetzung der i-ten Spalte durch die rechte Seite b gewonnene Matrix ist:

$$B_i := \begin{array}{c} \begin{matrix} \phantom{\alpha_{11}} & & i & & \phantom{\alpha_{1n}} \\ & & \downarrow & & \end{matrix} \\ \begin{pmatrix} \alpha_{11} & \dots & \beta_1 & \dots & \alpha_{1n} \\ \alpha_{21} & \dots & \beta_2 & \dots & \alpha_{2n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \dots & \beta_n & \dots & \alpha_{nn} \end{pmatrix} \end{array}$$

Der letzte Schritt folgt aus der Entwicklung der Determinante von B_i nach der i-ten Spalte:

$$\det B_i = \sum_{j=1}^{n} (-1)^{i+j} \beta_j \det(M_{ji})$$

Die Cramersche Regel ergibt eine explizite Formel zur Lösung von (L), falls A invertierbar ist. Man bedenke aber, dass diese Formel für praktische Zwecke nutzlos ist, da man $n+1$ Determinanten berechnen muss. Für praktische Zwecke sollte man den Gauß-Jordan-Algorithmus verwenden.

Beispiel. Das System $\left\{\begin{array}{rcr} \lambda x_1 & + & x_3 = 1 \\ x_1 + \lambda x_2 & & = 1 \\ x_2 & + & \lambda x_3 = 1 \end{array}\right\}$ hat die Lösung

$$x_1 = \frac{\begin{vmatrix} 1 & 0 & 1 \\ 1 & \lambda & 0 \\ 1 & 1 & \lambda \end{vmatrix}}{\lambda^3+1}, \qquad x_2 = \frac{\begin{vmatrix} \lambda & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & \lambda \end{vmatrix}}{\lambda^3+1}, \qquad x_3 = \frac{\begin{vmatrix} \lambda & 0 & 1 \\ 1 & \lambda & 1 \\ 0 & 1 & 1 \end{vmatrix}}{\lambda^3+1}$$

Es folgt

$$x_1 = \frac{\lambda^2-\lambda+1}{\lambda^3+1}, \qquad x_2 = \frac{\lambda^2-\lambda+1}{\lambda^3+1}, \qquad x_3 = \frac{\lambda^2-\lambda+1}{\lambda^3+1}$$

$$x_1 = \frac{1}{\lambda+1}, \qquad x_2 = \frac{1}{\lambda+1}, \qquad x_3 = \frac{1}{\lambda+1}$$

(vergleiche mit Beispiel (b) in 5.4.7)

5.5. Die Determinante eines Endomorphismus

Wir betrachten nun einen endlichdimensionalen Vektorraum V über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und einen Endomorphismus

$$\varphi : V \to V.$$

Sei A die Matrix von φ bzgl. einer Basis $\mathcal{B}$ von V und sei A' die Matrix von φ bzgl. einer anderen Basis $\mathcal{B}'$ von V.

Wir wissen aus 4.6.4, dass es eine invertierbare Matrix S, die Transformationsmatrix, gibt, so daß gilt:

$$A' = S^{-1}AS.$$

Aus den Eigenschaften (D10) und (D11) in 5.4.3 folgt

$$\begin{aligned} \det(A') &= \det(S^{-1}AS) = \det(S^{-1}) \cdot \det A \cdot \det S = (\det S)^{-1} \cdot \det A \cdot \det S \\ &= \det(A) \end{aligned}$$

Also hängt die Determinante der Matrix eines Endomorphismus φ nicht von der zur Aufstellung der Matrix gewählten Basis ab. Wir können analog wie bei der Spur (siehe 4.9) die Determinante eines Endomorphismus definieren:

Definition. *Für einen Endomorphismus $\varphi : V \to V$ eines endlichdimensionalen Vektorraumes definieren wir dessen Determinante als*

$$\det \varphi := \det A,$$

also als die Determinante der Matrix A von φ bzgl. irgendeiner Basis $\mathcal{B}$ von V.

Die Eigenschaften (D3), (D9), (D10), (D11) in 5.4.3 ziehen folgende Aussagen nach sich:

Eigenschaften.

(i) $\det(id_V) = 1$.

(ii) $\det \varphi \neq 0 \Longleftrightarrow \varphi$ *ist ein Automorphismus. In diesem Fall gilt* $\det(\varphi^{-1}) = \frac{1}{\det \varphi}$.

(iii) $\det(\varphi \circ \psi) = \det \varphi \cdot \det \psi$.

Beispiele.

(a) Sei δ eine Drehung im $\mathbb{R}^3$. Dann gibt es eine Basis in $\mathbb{R}^3$, so dass die Matrix von δ bzgl. dieser Basis folgende Gestalt annimmt:

$$\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Also ist $\det \delta = \cos^2\gamma + \sin^2\gamma = 1$.

(b) Sei σ die Spiegelung im $\mathbb{R}^3$ an einer Ebene. Bzgl. einer geeigneten Basis hat die Matrix von σ die Form

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

Also ist $\det \sigma = -1$.

5.6. Permutationen

Eine *Permutation* der n-elementigen Menge $\{1, 2, \dots, n\}$ ist eine bijektive Abbildung

$$\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$$

Wir schreiben $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$. Wir bezeichnen mit $\mathcal{S}_n$ die Menge aller Permutationen σ der n-elementigen Menge $\{1, 2, \dots, n\}$. Es gibt $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ Permutationen einer n-elementigen Menge.

Bzgl. der Hintereinanderausführung von Abbildungen ist $(\mathcal{S}_n, \circ)$ eine Gruppe.

$$\mathcal{S}_3 = \left\{ \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}_{id}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}}_{\sigma_1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}}_{\sigma_2}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}}_{\tau_{12}}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}}_{\tau_{13}}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}}_{\tau_{23}} \right\}$$

Es gilt z.B. $\sigma_1^2 = \sigma_1 \circ \sigma_1 = \sigma_2$, $\sigma_1^{-1} = \sigma_2$, $\sigma_1^3 = \sigma_1 \circ \sigma_1 \circ \sigma_1 = id$.

5.6.1 Das Signum einer Permutationen. Betrachten wir eine Permutation $\sigma \in \mathcal{S}_n$.

Jedes Paar (i, j) mit $i < j$ aber $\sigma(i) > \sigma(j)$ heißt eine *Inversion* von σ. Wir bezeichnen mit $I(\sigma)$ die Anzahl der Inversionen von σ.

Ist die Zahl $I(\sigma)$ der Inversionen gerade, so heißt σ eine *gerade* Permutation; ist $I(\sigma)$ ungerade, heißt σ eine *ungerade* Permutation. Wir definieren *das Signum von* σ durch

$$\boxed{\operatorname{sign}(\sigma) := (-1)^{I(\sigma)}}.$$

Damit ergibt sich:

$$\operatorname{sign}(\sigma)=\begin{cases}+1 & \text{wenn } \sigma \text{ eine gerade Permutation ist,}\\ -1 & \text{wenn } \sigma \text{ eine ungerade Permutation ist.}\end{cases}$$

$\mathcal{A}_n := \{\sigma \in \mathcal{S}_n : \operatorname{sign}(\sigma) = +1\}$ sei die Menge der geraden Permutationen und
$\mathcal{B}_n := \{\sigma \in \mathcal{S}_n : \operatorname{sign}(\sigma) = -1\}$ die Menge der ungeraden Permutationen.

Beispiele.

(a) In $\mathcal{S}_3$ gilt:
- id, σ_1, σ_2 sind gerade,
- τ_{12}, τ_{13}, τ_{23} sind ungerade.

(b) Die Identität hat keine Inversion, d.h. $I(id) = 0$, weshalb $\operatorname{sign}(id) = +1$ ist.

(c) Betrachte die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$$

Inversionen sind unterhalb der Zeichnung markiert, also ist $I(\sigma) = 5$ und damit

$\operatorname{sign}(\sigma) = (-1)^5 = -1.$

5.6.2 Satz. Für beliebige Permutationen $\sigma, \tau \in \mathcal{S}_n$ gilt

$$\boxed{\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau).}$$

Beweis. Wir betrachten die folgende Funktion in n Variablen:

$$f(x_1, x_2, \dots, x_n) := \prod_{i<j} (x_i - x_j)$$

Für $n = 3$ gilt zum Beispiel $f(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$. Wir behaupten:

Lemma. $f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = \operatorname{sign}(\sigma) \cdot f(x_1, x_2, \dots, x_n)$

Beweis des Lemmas:

$$\begin{aligned} f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) &= \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)}) \\ &= \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} (x_{\sigma(i)} - x_{\sigma(j)}) \cdot \prod_{\substack{i<j\\ \sigma(i)>\sigma(j)}} (x_{\sigma(i)} - x_{\sigma(j)}) \\ &= \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} (x_{\sigma(i)} - x_{\sigma(j)}) \cdot (-1)^{I(\sigma)} \prod_{\substack{i<j\\ \sigma(i)>\sigma(j)}} (x_{\sigma(j)} - x_{\sigma(i)}) \\ &= \operatorname{sign}(\sigma) \cdot \prod_{i<j} (x_i - x_j) = \operatorname{sign}(\sigma)\, f(x_1, x_2, \dots, x_n). \end{aligned}$$

Wir fahren mit dem Beweis des Satzes fort:

$$\underline{\operatorname{sign}(\sigma \circ \tau)} \cdot f(x_1, x_2, \dots, x_n)$$

$$
\begin{aligned}
&= f(x_{\sigma(\tau(1))},\, x_{\sigma(\tau(2))}, \ldots, x_{\sigma(\tau(n))}) && \text{nach dem Lemma}\\
&= \mathrm{sign}(\sigma)\cdot f(x_{\tau(1)},\, x_{\tau(2)}, \ldots, x_{\tau(n)}) && \text{wieder nach dem Lemma}\\
&= \underline{\mathrm{sign}(\sigma)\cdot\mathrm{sign}(\tau)}\cdot f(x_1, x_2, \ldots, x_n) && \text{wieder nach dem Lemma}
\end{aligned}
$$

Daraus folgt: $\mathrm{sign}(\sigma\circ\tau) = \mathrm{sign}(\sigma)\cdot\mathrm{sign}(\tau)$ □

5.6.3 Folgerungen.

(a) $\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma)$, d.h. ist σ gerade, dann ist auch σ^{-1} gerade, ist σ ungerade, dann ist auch σ^{-1} ungerade.

(b) Das Produkt zweier gerader Permutationen ist gerade, ebenso ist das Produkt zweier ungerader Permutationen ungerade. Das Produkt einer ungeraden und einer geraden Permutation (in beliebiger Reihenfolge) ist ungerade.

Beweis. (a) Aus Satz 6.5.2 folgt

$$\mathrm{sign}(\sigma)\cdot\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma\circ\sigma^{-1}) = \mathrm{sign}(id) = 1$$

Ist also $\mathrm{sign}(\sigma) = +1$, dann ist auch $\mathrm{sign}(\sigma^{-1}) = +1$. Ist $\mathrm{sign}(\sigma) = -1$, dann ist auch $\mathrm{sign}(\sigma^{-1}) = -1$.

(b) Dies folgt direkt aus 5.6.2. □

5.6.4 Transpositionen. Eine Permutation, die zwei Elemente $i < j$ miteinander vertauscht und alle anderen Elemente festhält, heißt *Transposition*. Wir bezeichnen eine solche Transposition auch mit τ_{ij}. Formal:

$$\tau_{ij}(k) = \begin{cases} j & \text{wenn } k = i\\ i & \text{wenn } k = j\\ k & \text{wenn } k \neq i, j\end{cases}$$

$$\tau_{ij} = \begin{pmatrix} 1 & 2 & \ldots & i-1 & i & i+1 & \ldots & j-1 & j & j+1 & \ldots & n\\ 1 & 2 & \ldots & i-1 & j & i+1 & \ldots & j-1 & i & j+1 & \ldots & n\end{pmatrix}$$

Jede Transposition ist ungerade, denn die Anzahl der Inversionen von τ_{ij} ist $I(\tau_{ij}) = (j-i) + (j-i-1) = 2(j-i) - 1$, eine ungerade Zahl.

5.6.5 Gerade und ungerade Permutationen. $\mathcal{S}_n$ unterteilt sich in zwei Klassen von Permutationen: gerade und ungerade

$$\mathcal{S}_n = \mathcal{A}_n \cup \mathcal{B}_n \quad , \quad \mathcal{A}_n \cap \mathcal{B}_n = \emptyset\,.$$

Nun sei τ eine Transposition, z.B. $\tau = \tau_{12}$. Aus 5.6.4 folgt, dass τ ungerade ist.

Für jede gerade Permutation σ ist das Produkt $\sigma\circ\tau$ nach 5.6.3 (b) ungerade. Also haben wir eine Funktion

$$\begin{array}{lcl} \mathcal{A}_n & \to & \mathcal{B}_n\\ \sigma & \mapsto & \sigma\circ\tau \end{array}$$

Diese Funktion ist surjektiv, denn für jedes $\sigma' \in \mathcal{B}_n$ ist die Permutation $\sigma := \sigma'\circ\tau$ gerade und $\sigma' = \sigma'\circ id = \sigma'\circ\tau\circ\tau = \sigma\circ\tau$. Die Funktion ist auch injektiv, da man aus $\sigma_1\circ\tau = \sigma_2\circ\tau$ folgern kann $\sigma_1 = \sigma_1\circ id = \sigma_1\circ\tau\circ\tau = \sigma_2\circ\tau\circ\tau = \sigma_2\circ id = \sigma_2$.

Damit ist die obige Funktion eine Bijektion. Also haben $\mathcal{A}_n$ und $\mathcal{B}_n$ dieselbe Anzahl an Elementen. Da es insgesamt $n!$ Permutationen gibt, gibt es $\frac{n!}{2}$ gerade und $\frac{n!}{2}$ ungerade Permutationen.

5.7. Beweis des Hauptsatzes 5.4.2

5.7.1 Explizite Formel für die Determinante. Um die Existenz einer Determinantenabbildung zu zeigen, definieren wir für jede $n \times n$-Matrix $A = (\alpha_{ij})$:

$$\Delta(A) := \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \alpha_{\sigma(1),1} \cdot \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n}$$

Behauptung 1. $\Delta : \mathcal{M}(n \times n, \mathbb{K}) \to \mathbb{K}$ *ist eine Determinantenabbildung.*

Beweis. Wir zeigen, dass (D1), (D2), (D3) erfüllt sind.

(D1): Wir zeigen, dass $\Delta(A)$ linear von der ersten Spalte von A abhängt (der Beweis verläuft analog für die j-te Spalte):

$$\Delta \begin{pmatrix} \alpha_{11} + \alpha'_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} + \alpha'_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} + \alpha'_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)(\alpha_{\sigma(1),1} + \alpha'_{\sigma(1),1}) \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n}$$

$$= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \alpha_{\sigma(1),1} \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n} + \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \alpha'_{\sigma(1),1} \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n}$$

$$= \Delta \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} + \Delta \begin{pmatrix} \alpha'_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha'_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha'_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} \qquad \text{und}$$

$$\Delta \begin{pmatrix} \lambda\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \lambda\alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda\alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)(\lambda \alpha_{\sigma(1),1}) \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n}$$

$$= \lambda \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \alpha_{\sigma(1),1} \alpha_{\sigma(2),2} \dots \alpha_{\sigma(n),n} = \lambda \cdot \Delta \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}$$

(D2): Wir nehmen an, A sei eine Matrix, bei der die erste Spalte gleich der zweiten Spalte ist, d.h. $\alpha_{i1} = \alpha_{i2}$ für $i = 1, \dots, n$. Wir möchten zeigen, dass $\Delta(A) = 0$ gilt (falls die j-te und die k-te Spalte gleich sind, ist der Beweis analog: ersetze einfach τ_{12} durch τ_{jk}):

Bekanntermaßen bezeichnen A_n und B_n die Mengen der geraden und ungeraden Permutationen. Es folgt:

$$\begin{aligned}\Delta(A) &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma)\cdot\alpha_{\sigma(1),1}\cdot\alpha_{\sigma(2),2}\cdot\ldots\cdot\alpha_{\sigma(n),n}\\ &= \sum_{\sigma\in A_n} \underbrace{\operatorname{sign}(\sigma)}_{+1}\cdot\alpha_{\sigma(1),1}\cdot\alpha_{\sigma(2),2}\cdot\ldots\cdot\alpha_{\sigma(n),n}\\ &+ \sum_{\sigma_1\in B_n} \underbrace{\operatorname{sign}(\sigma_1)}_{-1}\cdot\alpha_{\sigma_1(1),1}\cdot\alpha_{\sigma_1(2),2}\cdot\ldots\cdot\alpha_{\sigma_1(n),n}\end{aligned}$$

Es sei nun $\tau = \tau_{12}$ und wir wissen aus 5.6.5, dass es für jedes $\sigma_1 \in B_n$ ein eindeutiges $\sigma \in A_n$ gibt, so dass $\sigma_1 = \sigma \circ \tau$ ist. Dann gilt:

$$\begin{array}{lll} \sigma_1(1) & = \sigma(\tau(1)) & = \sigma(2) \\ \sigma_1(2) & = \sigma(\tau(2)) & = \sigma(1) \\ \sigma_1(i) & = \sigma(\tau(i)) & = \sigma(i) \qquad \text{für } i > 2 \end{array}$$

Dementsprechend können wir folgern:

$$\begin{aligned}\Delta(A) &= \sum_{\sigma\in A_n} \alpha_{\sigma(1),1}\cdot\alpha_{\sigma(2),2}\cdot\ldots\cdot\alpha_{\sigma(n),n}\\ &+ \sum_{\sigma\in A_n} -\alpha_{\sigma(2),1}\cdot\alpha_{\sigma(1),2}\cdot\ldots\cdot\alpha_{\sigma(n),n}\\ &= 0\end{aligned}$$

da nach Voraussetzung $\alpha_{\sigma(1),1} = \alpha_{\sigma(1),2}$ und $\alpha_{\sigma(2),2} = \alpha_{\sigma(2),1}$.

(D3): Sei $A = E_n$. Dann gilt

$$\begin{aligned}&\alpha_{11} = \alpha_{22} = \ldots = \alpha_{nn} = 1 \quad \text{und}\\ &\alpha_{ij} = 0 \qquad \text{falls } i \neq j\end{aligned}$$

Für jede Permutation σ mit $\sigma \neq id$ gibt es ein j, so dass $\sigma(j) \neq j$ und damit $\alpha_{\sigma(j),j} = 0$ gilt. Also gilt für jedes Produkt $\alpha_{\sigma(1),1}\cdot\alpha_{\sigma(2),2}\cdot\ldots\cdot\alpha_{\sigma(n),n} = 0$ für $\sigma \neq id$. Damit gilt

$$\Delta(E_n) = \operatorname{sign}(id)\alpha_{11}\alpha_{22}\ldots\alpha_{nn} = 1\,.$$

□

Dies beendet den Beweis von Behauptung 1.

Behauptung 2. $\Delta(A^T) = \Delta(A)\,.$

Beweis. Da Transponieren einer Matrix die Rollen von Zeilen und Spalten vertauscht, erhält man

$$\Delta(A^T) = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma)\cdot\alpha_{1,\sigma(1)}\cdot\alpha_{2,\sigma(2)}\cdot\ldots\cdot\alpha_{n,\sigma(n)}$$

Wir bemerken, dass $\alpha_{1,\sigma(1)}\cdot\alpha_{2,\sigma(2)}\cdot\ldots\cdot\alpha_{n,\sigma(n)} = \alpha_{\sigma^{-1}(1),1}\cdot\alpha_{\sigma^{-1}(2),2}\cdot\ldots\cdot\alpha_{\sigma^{-1}(n),n}$ gilt.

In der Tat stehen auf beiden Seiten dieselben Faktoren nur in anderer Reihenfolge. Daraus ergibt sich:

$$\Delta(A^T) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \cdot \alpha_{\sigma^{-1}(1),1} \cdot \alpha_{\sigma^{-1}(2),2} \cdot \cdots \cdot \alpha_{\sigma^{-1}(n),n}$$

Nach 5.6.3 ist $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$. Demnach ist

$$\Delta(A^T) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma^{-1}) \cdot \alpha_{\sigma^{-1}(1),1} \cdot \alpha_{\sigma^{-1}(2),2} \cdot \ldots \cdot \alpha_{\sigma^{-1}(n),n}$$

Läuft σ über alle Permutationen, so gilt auch dasselbe für σ^{-1}. Daraus folgt

$$\begin{aligned}\Delta(A^T) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \cdot \alpha_{\sigma(1),1} \cdot \alpha_{\sigma(2),2} \cdot \ldots \cdot \alpha_{\sigma(n),n} \\ &= \Delta(A),\end{aligned}$$

was den Beweis von Behauptung 2 beendet. □

Wir stehen damit auch am Ende des Beweises der Existenz einer Determinantenabbildung. Die Eindeutigkeit folgt aus (D6), (D7), (D8). Diese Eigenschaften sind Folgerungen von (D1), (D2), (D3), (und (D4)), nach denen wir die Determinante explizit berechnen können. Damit haben wir eine explizite Formel für die Determinante einer $n \times n$-Matrix $A = (\alpha_{ij})$:

(Leibniz)
$$\boxed{\det A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \alpha_{\sigma(1),1} \cdot \alpha_{\sigma(2),2} \cdot \ldots \cdot \alpha_{\sigma(n),n}}$$

Es könnte sinnvoll erscheinen, mittels dieser Formel die Determinante zu berechnen. Man bedenke aber, dass dies für praktische Zwecke nur für kleine n möglich ist. Es gibt nämlich $n!$ Permutationen auf n Elementen. Wir müssen also $n!$ Produkte von n Faktoren bestimmen.

$$\begin{array}{ll} n = 2 & n! = 2 \\ n = 3 & n! = 6 \\ n = 4 & n! = 24 \\ n = 5 & n! = 120 \\ \vdots & \vdots \end{array}$$

Bereits für $n = 4$ ist die Formel von Leibniz zu komplex um sie in der Praxis zu verwenden.

Um die Formel von Leibniz besser zu verstehen, betrachten wir die Fälle $n = 2$ und $n = 3$. Wir erhalten dann die folgenden bekannten Ausdrücke:

- Für $n = 2$ ist die Permutationsgruppe $\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} = \{\underset{+}{id}, \underset{-}{\tau_{12}}\}$ (beachte die Vorzeichen) und daraus folgt

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{22}.$$

- Für $n = 3$ gilt $\mathcal{S}_3 = \{\underset{+}{id}, \underset{+}{\sigma_1}, \underset{+}{\sigma_2}, \underset{-}{\tau_{12}}, \underset{-}{\tau_{13}}, \underset{-}{\tau_{23}}\}$ (beachte wieder die Vorzeichen). Dies führt auf

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \begin{matrix} \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{21}\alpha_{32}\alpha_{13} + \alpha_{31}\alpha_{12}\alpha_{23} \\ -\alpha_{21}\alpha_{12}\alpha_{33} - \alpha_{31}\alpha_{22}\alpha_{13} - \alpha_{11}\alpha_{32}\alpha_{23} \end{matrix}$$

  Dieses Ergebnis erhalten wir auch über die Regel von Sarrus.

Allgemeiner betrachten wir das Produkt

$$\alpha_{\sigma(1),1} \cdot \alpha_{\sigma(2),2} \cdot \ldots \cdot \alpha_{\sigma(n),n}$$

Für $\sigma = \mathrm{id}$ ergibt sich das Produkt $\alpha_{11} \cdot \alpha_{22} \cdot \ldots \cdot \alpha_{nn}$. Wir wählen im Allgemeinen

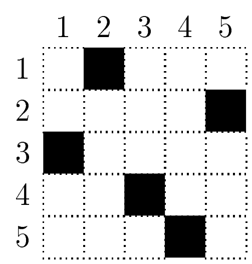
- den ersten Faktor in der ersten Spalte und in der Zeile mit der Nummer $\sigma(1)$
- den zweiten Faktor in der zweiten Spalte und in der Zeile mit der Nummer $\sigma(2)$

etc.

Insgesamt wählen wir in jeder Zeile und in jeder Spalte genau ein Element und dieses auf alle möglichen Arten, da σ eine beliebige Permutation ist.

Beispiel.

Für $n = 5$ veranschaulichen wir $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 5 & 2 \end{pmatrix}$ als



Beobachtung. *Für $\sigma \neq id$ enthält das Produkt*

$$\alpha_{\sigma(1),1} \cdot \alpha_{\sigma(2),2} \cdot \ldots \cdot \alpha_{\sigma(n),n}$$

höchstens $n - 2$ Faktoren auf der Hauptdiagonale.

Ist nämlich $\sigma \neq id$, so gibt es ein j mit $\sigma(j) \neq j$. Dann muss es aber auch ein $k \neq j$ mit $\sigma(k) = j$ geben.

KAPITEL 6

Eigenvektoren und Eigenwerte

6.1. Eigenwerte und Eigenvektoren von Endomorphismen

In diesem Abschnitt sei V stets ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und φ ein Endomorphismus von V, d.h. eine lineare Abbildung $\varphi : V \to V$.

6.1.1 Eigenvektoren und Eigenwerte. Eine Zahl $\lambda \in \mathbb{K}$ wird *Eigenwert* von φ genannt, wenn es einen Vektor $0 \neq v \in V$ gibt, so dass

$$\varphi(v) = \lambda v$$

gilt; jeder Vektor $v \neq 0$, der diese Gleichung erfüllt, wird *Eigenvektor* von φ zum Eigenwert λ genannt.

Wir können die Gleichung $\varphi(v) = \lambda v$ folgendermaßen umformen:

$$\begin{aligned} \varphi(v) = \lambda v &\iff \varphi(v) - \lambda v = 0 \\ &\iff \varphi(v) - \lambda \operatorname{id}(v) = 0 \\ &\iff (\varphi - \lambda \operatorname{id})(v) = 0 \\ &\iff v \in \ker(\varphi - \lambda \operatorname{id}) \end{aligned}$$

Dabei ist $\operatorname{id} = \operatorname{id}_V$ die identische Abbildung auf V. Daraus ergibt sich die folgende Charakterisierung von Eigenwerten und Eigenvektoren:

6.1.2 Proposition.

$$\begin{aligned} &\lambda \textit{ ist ein Eigenwert von } \varphi \\ \iff &\textit{Die Gleichung } (\varphi - \lambda \operatorname{id})(v) = 0 \textit{ hat eine Lösung } v \neq 0 \\ \iff &\ker(\varphi - \lambda \operatorname{id}) \neq \{0\} \end{aligned}$$

$$\begin{aligned} &v \textit{ ist ein Eigenvektor von } \varphi \textit{ zum Eigenwert } \lambda \\ \iff &v \textit{ ist eine nichtverschwindende Lösung von } (\varphi - \lambda \operatorname{id})v = 0 \\ \iff &0 \neq v \in \ker(\varphi - \lambda \operatorname{id}) \end{aligned}$$

6.1.3 Eigenräume. Für jeden Eigenwert λ von φ betrachten wir

$$U_\lambda := \{v \in V : \varphi(v) = \lambda v\} = \ker(\varphi - \lambda \operatorname{id})$$

Da der Kern einer linearen Abbildung immer ein linearer Teilraum ist (siehe 5.1.3), ist insbesondere U_λ ein linearer Teilraum von V.

U_λ besteht aus dem Nullvektor und aus allen Eigenvektoren von φ, die zum Eigenwert λ gehören. Aus diesem Grund heiße U_λ der *Eigenraum* von φ zum Eigenwert λ.

Die Dimension des Eigenraums

$$t := \dim U_\lambda \geq 1$$

wird die *geometrische Vielfachheit* des Eigenwertes λ genannt. Um eine Basis von U_λ zu finden, genügt es, t linear unabhängige Lösungen der Gleichung $(\varphi - \lambda\,\mathrm{id})(v) = 0$ zu berechnen, sofern U_λ endlichdimensional ist.

Wenden wir uns jetzt dem Fall eines Vektorraum V endlicher Dimension n zu. Dann gilt:

$$\begin{aligned}
\ker(\varphi - \lambda\,\mathrm{id}) = \{0\} &\iff \varphi - \lambda\,\mathrm{id} \text{ ist injektiv} && \text{(nach 4.2.5)}\\
&\iff \varphi - \lambda\,\mathrm{id} \text{ ist bijektiv} && \text{(nach 4.2.7)}\\
&\iff \det(\varphi - \lambda\,\mathrm{id}) \neq 0 && \text{(nach 5.4.3 (D9))}
\end{aligned}$$

Also kann man Proposition 6.1.2 folgendermaßen präzisieren:

6.1.4 Satz. Sei φ ein Endomorphismus eines endlichdimensionalen Vektrorraums. Dann istλ genau dann ein Eigenwert von φ, wenn $\det(\varphi - \lambda\,\mathrm{id}) = 0$.

6.1.5 Motivation. Um φ durch eine Matrix A darstellen zu können, wählen wir eine Basis

$$\mathcal{B} : v_1, v_2, \ldots, v_n$$

von V und berechnen die Bilder der Basisvektoren

$$\varphi(v_j) = \sum_{i=1}^{n} \alpha_{ij} v_i \qquad (j = 1, \ldots, n)\,.$$

Dann wird φ (bzgl. der Basis $\mathcal{B}$) durch die $(n\times n)$-Matrix

$$A = [\varphi]_{\mathcal{B}} = (\alpha_{ij})$$

dargestellt. Man möchte für die Matrix A eine möglichst einfache Gestalt finden, d.h. wir möchten eine Basis $\mathcal{B}$ auswählen, so dass die Form der Matrix A einfach wird (dann kann man leichter Berechnungen mit A durchführen).

Die einfachsten Matrizen sind Diagonalmatrizen

$$D = \begin{pmatrix} \lambda_1 & & & \Large 0 \\ & \lambda_2 & & \\ & & \ddots & \\ \Large 0 & & & \lambda_n \end{pmatrix}$$

Die Matrix von φ bzgl. der Basis $\mathcal{B}$ hat genau dann Diagonalgestalt D, wenn

$$\varphi(v_j) = \lambda_j v_j \qquad \text{für } j = 1, \ldots, n\,,$$

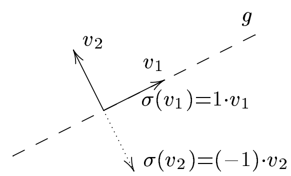
d.h. wenn die Basisvektoren v_j Eigenvektoren von φ sind. Die Skalare λ_j auf der Diagonalen sind dann Eigenwerte von φ

Dies führt zu folgender Frage: *Gibt es in V eine Basis $\mathcal{B} : v_1, \ldots, v_n$ von Eigenvektoren von φ und wie können wir die Eigenvektoren und Eigenwerte finden?*

6.1.6 Beispiele. In einigen Fällen können wir eine Basis von Eigenvektoren aus geometrischen Gründen angeben: Sei $V = \mathbb{R}^2$ und g eine Gerade durch 0.

a) Sei σ die (orthogonale) Spiegelung an der Geraden g. Wir wählen einen Vektor v_1 in Richtung von g und einen Vektor v_2 orthogonal zu g. Dann folgt

$$\sigma(v_1) = 1 \cdot v_1 \quad , \quad \sigma(v_2) = (-1) \cdot v_2 \, ,$$



d.h. v_1 ist ein Eigenvektor zu Eigenwert $\lambda_1 = 1$ und v_2 ein Eigenvektor zu Eigenwert $\lambda_2 = -1$. Die Matrix von σ bzgl. der Basis v_1, v_2 ist eine Diagonalmatrix mit den Eigenwerten auf der Hauptdiagonalen:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Der Eigenraum zum Eigenwert $\lambda_1 = 1$ ist die Gerade g, der Eigenraum zum Eigenwert $\lambda_2 = -1$ ist die Gerade durch 0 senkrecht zu g. Beide Eigenwerte haben die geometrische Vielfachheit 1.

b) Sei π die Orthogonalprojektion auf g. Für die obige Basis v_1, v_2 erhalten wir

$$\sigma(v_1) = 1 \cdot v_1 \quad , \quad \sigma(v_2) = 0 = 0 \cdot v_2$$

In diesem Fall ist v_1 ein Eigenvektor zum Eigenwert 1 und v_2 ein Eigenvektor zu Eigenwert 0. Die Matrix von D bzgl. v_1, v_2 lautet

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

Wiederum ist der Eigenraum zum Eigenwert $\lambda_1 = 1$ die Gerade g und der Eigenraum zum Eigenwert $\lambda_2 = 0$ die Gerade durch 0 senkrecht zu g. Beide Eigenwerte haben die geometrische Vielfachheit 1.

6.1.7 Satz. Seien $\lambda_1, \lambda_2, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte von φ. Die Summe der zugehörigen Eigenräume ist eine direkte Summe:

$$U_{\lambda_1} + U_{\lambda_2} + \ldots + U_{\lambda_r} = U_{\lambda_1} \oplus U_{\lambda_2} \oplus \ldots \oplus U_{\lambda_r} \, .$$

Beweis. Mittels Induktion nach r: Für $r = 1$ ist nichts zu zeigen. Wir nehmen nun an, unsere Behauptung sei richtig für $r-1$ Eigenwerte $\lambda_1, \ldots, \lambda_{r-1}$ und wir zeigen, dass diese dann auch richtig für r Eigenwerte ist: Es seien

$$\begin{aligned} u &= u_1 + \ldots + u_{r-1} + u_r \\ &= u_1' + \ldots + u_{r-1}' + u_r' \end{aligned} \tag{1}$$

zwei Darstellungen eines Vektors u durch Vektoren $u_i, u_i' \in U_{\lambda_i}$. Wir müssen zeigen, dass $u_i = u_i'$ für alle $i = 1, \ldots, r$ gilt.

Aus (1) erhalten wir durch Subtraktion

$$0 = (u_1 - u_1') + \ldots + (u_{r-1} - u_{r-1}') + (u_r - u_r') \, . \tag{2}$$

Anwendung von φ ergibt:

$$0 = \varphi(0) = \varphi(u_1 - u_1') + \ldots + \varphi(u_{r-1} - u_{r-1}') + \varphi(u_r - u_r') \, . \tag{3}$$

Da U_{λ_i} ein linearer Teilraum ist und da $u_i, u_i' \in U_{\lambda_i}$, folgt

$$u_i - u_i' \in U_{\lambda_i} \, , \text{ und damit } \varphi(u_i - u_i') = \lambda_i(u_i - u_i') \, .$$

Es gilt also

$$0 = \lambda_1(u_1 - u_1') + \ldots + \lambda_{r-1}(u_{r-1} - u_{r-1}') + \lambda_r(u_r - u_r') \,. \tag{4}$$

Wir multiplizieren Gleichung (2) mit λ_r und subtrahieren diese von (4); wir erhalten

$$0 = (\lambda_1 - \lambda_r)(u_1 - u_1') + \ldots + (\lambda_{r-1} - \lambda_r)(u_{r-1} - u_{r-1}') \,.$$

Aus der Induktionsannahme folgt nun $(\lambda_i - \lambda_r)(u_i - u_i') = 0$ für $i = 1, \ldots, r-1$. Wegen $\lambda_i \neq \lambda_r$ für $i = 1, \ldots, r-1$ ergibt sich $u_i - u_i' = 0$, d.h. $u_i = u_i'$ für $i = 1, \ldots, r-1$. Dann folgt aus Gleichung (1) auch $u_r = u_r'$. □

6.1.8 Folgerung.

Sind	$\lambda_1, \ldots, \lambda_r$	paarweise verschiedene Eigenwerte von φ
und	$v_1, \ldots, v_r$	zugehörige Eigenvektoren,
dann sind	$v_1, \ldots, v_r$	linear unabhängig.

Folgerung: Ein Endomorphismus φ eines n-dimensionalen Vektorraumes V kann höchstens n verschiedene Eigenwerte besitzen. (Hat φ nämlich r verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_r$, dann können wir zugehörige Eigenvektoren $v_1, \ldots, v_r$ wählen. Diese sind nach 6.1.7 linear unabhängig. Es folgt $r \leq n$.)

6.1.9 Folgerung.

Seien	$\lambda_1, \ldots, \lambda_r$	verschiedene Eigenwerte von φ
	$t_1, \ldots, t_r$	deren geometrische Vielfachheiten.

Wähle	t_1	linear	unabhängige	Eigenvektoren	zugehörig	zu	λ_1
–	t_2	–	–	–	–	–	λ_2
⋮	⋮	⋮	⋮	⋮	⋮		⋮
–	t_r	–	–	–	–	–	λ_r

Dann sind alle diese Vektoren linear unabhängig und bilden eine Basis des Teilraums $U_{\lambda_1} \oplus \cdots \oplus U_{\lambda_r}$. Gilt $t_1 + t_2 + \cdots + t_r = n$, dann ist dieser Teilraum gleich V. Also hat V eine Basis von Eigenvektoren von φ und die Matrix von φ bzgl. dieser Basis ist eine Diagonalmatrix:

$$\begin{array}{r} t_1\left\{\vphantom{\begin{matrix}a\\b\\c\end{matrix}}\right. \\ \\ \\ t_r\left\{\vphantom{\begin{matrix}a\\b\\c\end{matrix}}\right. \end{array}
\begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_1\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}\lambda_r & & 0\\ & \ddots & \\ 0 & & \lambda_r\end{matrix}} \end{pmatrix}$$

6.2. Eigenwerte und Eigenvektoren von Matrizen

Sei A eine beliebige $n \times n$-Matrix mit Einträgen aus $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

Da A einen Endomorphismus $v \mapsto Av : \mathbb{K}^n \to \mathbb{K}^n$ definiert, können wir alle unsere Begriffe und Ergebnisse des vorigen Abschnittes von Endomorphismen auf Matrizen übertragen:

6.2.1 Eigenwerte und Eigenvektoren. Eine Zahl $\lambda \in \mathbb{K}$ heißt *Eigenwert* von A, wenn es einen Vektor $0 \neq v \in \mathbb{K}^n$ gibt mit

$$Av = \lambda v\,.$$

Jeder Vektor $v \neq 0$ in $\mathbb{K}^n$, der diese Gleichung erfüllt, heißt *Eigenvektor* von A zu dem Eigenwert λ.

Wir können die Gleichung $Av = \lambda v$ auf die folgende Art umformen:

$$\begin{aligned} Av = \lambda v &\iff Av - \lambda v = 0 \\ &\iff Av - \lambda E v = 0 \\ &\iff (A - \lambda E)v = 0 \end{aligned}$$

wobei $E = E_n$ die $n \times n$-Einheitsmatrix bezeichne. Wir folgern daraus

6.2.2 Proposition. *λ ist genau dann ein Eigenwert von A, wenn das homogene System linearer Gleichungen*

$$\text{(H)} \qquad (A - \lambda E)v = 0$$

eine nichttriviale Lösung $v \neq 0$ besitzt. Die nichttrivialen Lösungen von (H) sind die Eigenvektoren von A zum Eigenwert λ.

Wir wissen aus 4.8.4, dass ein homogenes System (H) linearer Gleichungen genau dann eine nichttriviale Lösung hat, wenn die Matrix $A - \lambda E$ nicht invertierbar ist, was nach 5.4.3(D9) genau dann gilt, wenn $\det(A - \lambda E) = 0$ ist. Daraus ergibt sich der folgende

Folgerung. *λ ist Eigenwert von A* $\iff$ $\boxed{\det(A - \lambda E) = 0}$.

Die letzte Formel wird auch *charakteristische Gleichung* von A genannt.

6.2.3 Eigenräume. Für jeden Eigenwert λ von A ist die Menge

$$U_\lambda := \{v \in \mathbb{K}^n : Av = \lambda v\} = \text{Menge der Lösungen von } (H)$$

ein linearer Teilraum von $\mathbb{K}^n$. Er besteht aus den Eigenvektoren von A zum Eigenwert λ und dem Nullvektor und wird *Eigenraum* von A zum Eigenwert λ. Die Dimension $t := \dim U_\lambda$ wird die *geometrische Vielfachheit* des Eigenwertes λ genannt.

Eine Basis von U_λ besteht aus t linear unabhängigen Lösungen von (H), d.h., es handelt sich um ein Fundamentalsystem von Lösungen von (H).

Bemerkung: Die geometrische Vielfachheit t ist $t = n - \operatorname{rank}(A - \lambda E)$.

6.2.4 Beispiel. (für $\mathbb{K} = \mathbb{R}$):

Sei $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Dann ist $A - \lambda E = \begin{pmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{pmatrix}$.

Eigenwerte: $\det(A - \lambda E) = \begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 9 = 0$ und damit

$$(1-\lambda)^2 = 9,\ 1-\lambda = \pm 3,\ \lambda_1 = 4, \lambda_2 = -2$$

Eigenvektoren: Wir müssen beide Eigenwerte berücksichtigen.

- $\lambda_1 = 4$:

$$(A - \lambda_1 E)v = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} v = 0 \qquad \Longrightarrow v = \gamma \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ für beliebiges } \gamma \in \mathbb{R}$$

- $\lambda_2 = -2$:

$$(A - \lambda_2 E)v = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} v = 0 \qquad \Longrightarrow v = \gamma \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ wieder für beliebiges } \gamma \in \mathbb{R}$$

Also sind $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ Eigenvektoren von A. Diese bilden eine Basis $\mathcal{B}'$ von $\mathbb{R}^2$.

Bemerkung: Die lineare Abbildung $\varphi_A : \mathbb{R}^2 \to \mathbb{R}^2$, die durch die Matrix A bzgl. der Standardbasis $\mathcal{B}$ gegeben ist, besitzt die folgende Matrix A' bzgl. der Basis $\mathcal{B}' : v_1, v_2$

$$A' = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

Tatsächlich ist $Av_1 = 4v_1$, $Av_2 = -2v_2$.

Die Matrizen A und A' sind über die folgende Formel miteinander verbunden:

$$A' = S^{-1}AS$$

Dabei ist S die Transformationsmatrix $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ (siehe 4.6.4). Die Spalten der Transformationsmatrix S sind die Koordinaten der neuen Basisvektoren v_1, v_2 (bzgl. der Standardbasis.)

6.2.5 Beispiel. Man betrachte $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Aus $A - \lambda E = \begin{pmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{pmatrix}$ berechnen wir die

Eigenwerte: $\det(A - \lambda E) = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 1 = 0$ und damit $(1-\lambda)^2 = -1$.

Diese Gleichung hat in $\mathbb{R}$ keine Lösung. Also hat A keine reellen Eigenwerte und keine reellen Eigenvektoren. Deshalb suchen wir nach komplexen Lösungen und erweitern den Körper $\mathbb{R}$ der reellen Zahlen zum Körper $\mathbb{C}$ der komplexen Zahlen:

Die charakteristische Gleichung ergibt nun $1-\lambda = \pm i$. A hat zwei konjugiert komplexe Eigenwerte $\lambda_{1/2} = 1 \pm i$.

Eigenvektoren:

- zum Eigenwert $\lambda_1 = 1+i$:
$$(A-\lambda_1 E)v = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} v = 0 \qquad \Longrightarrow \; v = \gamma \begin{pmatrix} -i \\ 1 \end{pmatrix} \text{ für } \gamma \in \mathbb{C}$$
- zum Eigenwert $\lambda_2 = 1-i$: Analog erhalten wir $v = \gamma \begin{pmatrix} i \\ 1 \end{pmatrix}$.

Die Vektoren $v_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ bilden eine Basis von $\mathbb{C}^2$ aus Eigenvektoren von A.

Die Eigenschaften 6.1.7, 6.1.8, 6.1.9 können auf Matrizen übertragen werden:

6.2.6 Satz. Seien $\lambda_1, \lambda_2, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte einer $n\times n$-Matrix A und seien $t_1, t_2, \ldots, t_r$ ihre geometrischen Vielfachheiten.

a) Die Summe der Eigenräume ist eine direkte Summe:
$$U_{\lambda_1} + U_{\lambda_2} + \cdots + U_{\lambda_r} = U_{\lambda_1} \oplus U_{\lambda_2} \oplus \cdots \oplus U_{\lambda_r}\,.$$

b) Sind $v_1, v_2, \ldots, v_r$ Eigenvektoren zu den Eigenwerten $\lambda_1, \lambda_2, \ldots, \lambda_r$, so sind diese Eigenvektoren linear unabhängig.

Da es höchstens n linear unabhängige Vektoren in $\mathbb{K}^n$ gibt, folgern wir $r \le n$. Also hat eine $n\times n$-Matrix höchstens n verschiedene Eigenwerte.

Hat A tatsächlich n verschiedene Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_n$, dann erhalten wir durch Wahl eines Eigenvektors v_1, v_2, $\ldots$, v_n für jeden der Eigenwerte eine Basis von $\mathbb{K}^n$.

c) Wählen wir
$$\begin{array}{lll} t_1 & \text{linear unabhängige Eigenvektoren von} & \lambda_1 \\ t_2 & \text{linear unabhängige Eigenvektoren von} & \lambda_2 \\ \vdots & \vdots & \vdots \\ t_r & \text{linear unabhängige Eigenvektoren von} & \lambda_r, \end{array}$$
so bilden diese Vektoren eine Basis $\mathcal{B}$ von $U_{\lambda_1} \oplus \cdots \oplus U_{\lambda_r}$. Gilt $t_1 + t_2 + \cdots + t_r = n$, so folgt $V = U_{\lambda_1} \oplus U_{\lambda_2} \oplus \cdots \oplus U_{\lambda_r}$ und $\mathcal{B}$ ist eine Basis von V aus Eigenvektoren von A.

6.2.7 Das charakteristische Polynom. Wir betrachten für eine $n\times n$-Matrix A die Determinante $\det(A-\lambda E)$ als Funktion in Abhängigkeit von λ:
$$P(\lambda) := \det(A-\lambda E) = \begin{vmatrix} \alpha_{11}-\lambda & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn}-\lambda \end{vmatrix}$$

Beispiel. $n = 2$:

$$P(\lambda) = \begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda \end{vmatrix} = (\alpha_{11} - \lambda)(\alpha_{22} - \lambda) - \alpha_{12}\alpha_{21}$$

$$\begin{aligned} P(\lambda) &= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \\ &= \lambda^2 - \mathrm{tr}(A)\lambda + \det(A) \end{aligned}$$

Allgemein gilt:

Proposition. *$P(\lambda)$ ist ein Polynom vom Grad n mit Koeffizienten aus $\mathbb{K}$; etwas präziser:*

$$P(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \underbrace{(\alpha_{11} + \cdots + \alpha_{nn})}_{(A)} \lambda^{n-1} + \ldots \cdots + \det(A)$$

$P(\lambda)$ heißt *charakteristisches Polynom* der Matrix A.

Nach 6.2 sind die Eigenwerte von A genau die *Nullstellen (= Wurzeln)* des charakteristischen Polynoms $P(\lambda)$. Da ein Polynom vom Grad n höchstens n verschiedene Nullstellen hat, erhalten wir damit einen neuen Beweis für die Aussage in 6.2.6 (b), dass eine $n \times n$-Matrix höchstens n verschiedene Eigenwerte hat.

Beweis der Proposition. Berechnen wir $\det(A - \lambda E)$ mittels der Formel von Leibniz nach 6.6.1, so erhalten wir

$$\begin{aligned} P(\lambda) \; &= (\alpha_{11} - \lambda)(\alpha_{22} - \lambda) \ldots (\alpha_{nn} - \lambda) + \text{Produkte mit höchstens } n-2 \\ &\qquad \text{Linearfaktoren } (\alpha_{ii} - \lambda) \\ &= (-1)^n \lambda^n + (-1)^{n-1}(\alpha_{11} + \cdots + \alpha_{nn})\lambda^{n-1} + \text{Terme niedrigeren Grades} \end{aligned}$$

Für das konstante Glied gilt $a_0 = P(0) = \det(A)$.

6.2.8 Vielfachheit von Nullstellen von Polynomen. Man erinnere sich an die folgenden Tatsachen für en Polynom $P(\lambda)$ vom Grad n mit Koeffizienten aus $\mathbb{K}$. Eine Zahl λ_0 wird *Nullstelle* oder auch *Wurzel* von $P(\lambda)$ genannt, wenn $P(\lambda_0) = 0$ ist.

Ist λ_0 eine Nullstelle von $P(\lambda)$, so kann man den Linearfaktor $\lambda - \lambda_0$ abspalten und erhält

$$P(\lambda) = (\lambda - \lambda_0)P_1(\lambda)$$

wobei $P_1(\lambda)$ ein Polynom vom Grad $n-1$ ist. Ist $P_1(\lambda_0) \neq 0$, so heißt λ_0 eine *einfache Nullstelle* von $P(\lambda)$.

Wenn aber wieder $P_1(\lambda_0) = 0$ gilt, kann man nochmals den Linearfaktor $\lambda - \lambda_0$ abspalten:

$$\begin{aligned} P(\lambda) = (\lambda - \lambda_0)P_1(\lambda) &= (\lambda - \lambda_0)(\lambda - \lambda_0)P_2(\lambda) \\ &= (\lambda - \lambda_0)^2 P_2(\lambda) \end{aligned}$$

und so weiter. Wir nennen λ_0 eine *Nullstelle der Vielfachheit* s_0, wenn P eine Zerlegung

$$P(\lambda) = (\lambda - \lambda_0)^{s_0} \cdot Q(\lambda)$$

besitzt, so dass $Q(\lambda_0) \neq 0$ gilt.

6.2.9 Algebraische Vielfachheit von Eigenwerten. Wir sagen, der Eigenwert λ_0 der Matrix A habe die *algebraische Vilefachheit* s_0 wenn λ_0 eine Nullstelle des charakteristische Polynoms $P(\lambda)$ von A der Vielfachheit s_0 ist.

Übung. Für die geometrische Vielfachheit t_0 und die algebraischen Vielfachheit s_0 eines jeden Eigenwerts λ_0 einer Matrix A gilt $s_0 \leq t_O$. .

6.2.10 Beispiel. ($\mathbb{K} = \mathbb{R}$): Betrachte $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Eigenwerte:

$$P(\lambda) = \det(A - \lambda E) = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = (\lambda - 1)^2$$

Also ist $\lambda_0 = 1$ der einzige Eigenwert, und dieser hat die algebraische Vielfachheit $s_0 = 2$.

Eigenvektoren:

$$(A - 1 \cdot E)v = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} v = 0 \qquad \Longrightarrow \; v = \gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ für beliebiges } \gamma$$

Also ist die geometrische Vielfachheit des einzigen Eigenwertes $\lambda_0 = 1$ gleich $t_0 = 1$. In diesem Beispiel ist $t_0 < s_0$ und es gibt keine Basis von Eigenvektoren.

6.3. Ähnlichkeit von Matrizen und Diagonalisierbarkeit

A, A', S, T bezeichnen in diesem Abschnitt immer $n \times n$-Matrizen mit Koeffizienten aus $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

6.3.1 Ähnlichkeit von Matrizen. Zwei $n \times n$-Matrizen A und A' werden *ähnlich* genannt und wir schreiben dann $A \approx A'$, wenn es eine invertierbare $n \times n$-Matrix S gibt mit

$$A' = S^{-1} A S\,.$$

Eigenschaften. Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation, d.h., sie erfüllt die folgenden Eigenschaften:

(i) $A \approx A \qquad$ für alle $A \in \mathcal{M}(n \times n, \mathbb{K})$ (Reflexivität)

(ii) $A \approx A' \Rightarrow A' \approx A$ (Symmetrie)

(iii) $A \approx A'\,,\; A' \approx A'' \quad \Rightarrow A \approx A''$ (Transitivität)

Beweis.

(i) Wähle $S = E_n$ (Einheitsmatrix); dann gilt $A = E_n^{-1} A E_n$, d.h. $A \approx A$.

(ii) Gilt $A \approx A'$, so gibt es eine invertierbare Matrix S, so dass $A' = S^{-1} A' S$ gilt. Durch Multiplikation mit S von der linken Seite und durch S^{-1} von der rechten Seite erhalten wir $SA'S^{-1} = A$. Wählen wir $T = S^{-1}$, so erhalten wir $A = T^{-1} A' T$ und damit $A' \approx A$.

(iii) Es gelte $A \approx A'$ und $A' \approx A''$. Dann gibt es invertierbare Matrizen S und T, so dass $A' = S^{-1}AS$, $A'' = T^{-1}A'T$ gilt. Durch Ersetzung von A' in der zweiten Gleichung durch den in der ersten Gleichung gegebenen Wert erhalten wir $A'' = T^{-1}S^{-1}AST = (ST)^{-1}A(ST)$. Also gilt $A \approx A''$.

Geometrische Bedeutung der Ähnlichkeit. Zwei Matrizen A und A' sind genau dann ähnlich, wenn sie dieselbe lineare Abbildung $\varphi : V \to V$ auf einem n-dimensionalen Vektorraum V bzgl. zweier verschiedener Basen $\mathcal{B}$ und $\mathcal{B}'$ darstellen.

Beweis. Sind A und A' Matrizen einer linearen Abbildung $\varphi : V \to V$ bzgl. zweier Basen $\mathcal{B}$ und $\mathcal{B}'$ in V, so gibt es nach 4.6.4 eine invertierbare Matrix S (die Transformationsmatrix), so daß $A' = S^{-1}AS$, d.h. $A \approx A'$.

Umgekehrt sei $A \approx A'$. Sei $\varphi\colon \mathbb{K}^n \to \mathbb{K}^n$ die lineare Abbildung mit der Matrix A bzgl. der Standardbasis $\mathcal{B}$. Nach Annahme gilt $A' = S^{-1}AS$ für eine invertierbare Matrix S. Die Zeilen $v_1, \ldots, v_n$ von S bilden eine Basis $\mathcal{B}'$ von $\mathbb{K}^n$. Die Matrix von φ bzgl. der Basis $\mathcal{B}'$ ist $A' = S^{-1}AS$ aufgrund der Transformationsformel 4.6.4. Also stellen A und A' dieselbe lineare Abbildung bzgl. der Basen $\mathcal{B}$ und $\mathcal{B}'$ dar.

Übung. Man beweise, daß ähnliche Matrizen dieselbe Determinante, dieselbe Spur und das gleiche charakteristische Polynom haben.

Wir kommen nun zu dem Problem zurück, welches zu Beginn des Kapitels als Motivation zur Untersuchung von Eigenwerten diente: Für eine gegebene Matrix A finde eine dazu ähnliche Matrix möglichst einfacher Gestalt.

6.3.2 Diagonalisierbarkeit. Eine Matrix A wird *diagonalisierbar* genannt, wenn sie einer Diagonalmatrix A' ähnlich ist, d.h. wenn es eine invertierbare Matrix S so gibt, daß

$$A' := S^{-1}AS \quad \text{eine Diagonalmatrix ist.}$$

Kriterium zur Diagonalisierbarkeit. A ist genau dann diagonalisierbar, wenn es in $\mathbb{K}^n$ eine Basis $v_1, v_2, \ldots, v_n$ aus Eigenvektoren zu den Eigenwerten $\lambda_2, \lambda_2, \ldots, \lambda_n$ von A gibt. In diesem Fall gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & & \Large 0 \\ & \lambda_2 & & \\ & & \ddots & \\ \Large 0 & & & \lambda_n \end{pmatrix} \quad (= A'),$$

wobei S die Transformationsmatrix mit den Eigenvektoren $v_1, \ldots, v_n$ als Spalten ist. (Die λ_i müssen nicht paarweise verschieden sein.)

6.3.3 Beispiele ($\mathbb{K} = \mathbb{R}, n = 2$)**.**

a) Die Matrix $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ aus Beispiel 6.2.4 ist diagonalisierbar:

Eigenwerte: $\lambda_1 = 4 \quad \lambda_2 = -2$

Basis von Eigenvektoren $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$$S^{-1}AS = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{mit} \quad S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

b) Die Matrix $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ aus Beispiel 6.2.5 hat keine reellen Eigenwerte. Sie kann also über dem Körper der reellen Zahlen nicht diagonalisiert werden. A hat aber komplexe

$$\begin{array}{lll} \text{Eigenwerte} & \lambda_1 = 1 + i & \lambda_2 = 1 - i \quad \text{und} \\ \text{Eigenvektoren} & v_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}, & v_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}. \end{array}$$

Also ist A über dem Körper der komplexen Zahlen diagonalisierbar:

$$S^{-1}AS = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \text{ mit } S = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}.$$

Dieses Beispiel zeigt, daß es reelle Matrizen gibt, die über den komplexen, nicht aber über den reellen Zahlen diagonalisierbar sind.

c) Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ aus Beispiel 6.2.10 ist weder über dem Körper der reellen noch über dem Körper der komplexen Zahlen diagonalisierbar. A hat nämlich einen einzigen Eigenwert $\lambda = 1$ der algebraischen Vielfachheit 2. Aber der Eigenraum bzgl. dieses Eigenwertes ist nur eindimensional. Weder in $\mathbb{R}^2$ noch $\mathbb{C}^2$ gibt es eine Basis von Eigenvektoren.

6.3.4 Bedingungen für Diagonalisierbarkeit über $\mathbb{C}$. Sei A eine $n \times n$-Matrix mit reellen oder komplexen Einträgen und

$$P(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$$

deren charakteristisches Polynom. Nach dem Fundamentalsatz der Algebra zerfällt $P(\lambda)$ in Linearfaktoren. Seien $\lambda_1, \lambda_2, \ldots, \lambda_r$ die verschiedenen komplexen Wurzeln von $P(\lambda)$, d.h. die Eigenwerte von A in $\mathbb{C}$ und $s_1, s_2, \ldots, s_r$ die dazugehörigen algebraischen Vielfachheiten, dann hat $P(\lambda)$ eine Zerlegung

$$P(\lambda) = (\lambda - \lambda_1)^{s_1} \ldots (\lambda - \lambda_r)^{s_r}.$$

Man beachte:

$$n = \text{ grad } (P(\lambda)) = s_1 + \cdots + s_r$$

- **Erster Fall:** $r = n$, d.h. A hat n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$. Für jeden Eigenwert berechnen wir je einen Eigenvektor $v_1, \dots, v_n$. Diese sind linear unabhängig und bilden deshalb eine Basis von $\mathbb{C}^n$. Also ist A nach 6.3.2 diagonalisierbar.

  Proposition 1. *Hat eine $n\times n$-Matrix n verschiedene Eigenwerte, dann ist sie diagonalisierbar.*

- **Zweiter Fall:** Sei $r < n$ und wir betrachten die Eigenräume $U_{\lambda_1}, \dots, U_{\lambda_r}$ und deren Dimensionen, d.h. die geometrischen Vielfachheiten $t_1, \dots, t_r$. Wir wählen eine Basis $\mathcal{B}_j$ in jedem der Eigenräume U_{λ_j}. Nach 6.2.6 (c) bilden die Basen $\mathcal{B}_1, \dots, \mathcal{B}_r$ zusammen ein linear unabhängiges System $\mathcal{B}$ aus $t_1+\dots+t_r$ Vektoren in $\mathbb{K}^n$. Da in $\mathbb{K}^n$ jeder Eigenvektor von A zu einem U_{λ_j} gehört, können wir feststellen:

  $$\begin{aligned} \mathbb{K}^n \text{ hat eine Basis von Eigenvektoren von } A &\iff t_1 + \dots + t_r = n \\ &\iff t_1 + t_2 + \dots + t_r = s_1 + \dots + s_r \end{aligned}$$

  Wegen $t_j \le s_j$ für alle j gilt dies genau dann, wenn $t_j = s_j$ für alle $j = 1, \dots, r$. Wir folgern:

  Proposition 2. *Eine $n\times n$-Matrix A ist genau dann diagonalisierbar über $\mathbb{C}$, falls für jeden Eigenwert λ_j die geometrische Vielfachheit t_j gleich der algebraischen Vielfachheit s_j ist.*

6.3.5 Spezialfall: $\mathbb{K} = \mathbb{R}$. Eine reelle $n\times n$-Matrix A kann als spezielle komplexe Matrix aufgefasst werden und man kann sich zunächst fragen, ob diese über den komplexen Zahlen diagonalisierbar ist. In diesem Fall und falls alle Eigenwerte reell sind, ist die Matrix bereits über den reellen Zahlen diagonalisierbar.

6.3.6 Anwendungen der Diagonalisierung. Wir beachten zunächst, dass für beliebige $n\times n$-Matrizen A, B und für jede invertierbare $n\times n$-Matrix S die folgenden Beziehungen gelten:

(i) $$S^{-1}(A + B)S = S^{-1}AS + S^{-1}BS$$

(ii) $$S^{-1}(\lambda A)S = \lambda \cdot S^{-1}AS$$

(iii) $$S^{-1}(AB)S = (S^{-1}AS) \cdot (S^{-1}BS)$$

(Dies bedeutet, dass die Abbildung $\Psi_S : \mathcal{M}(n\times n, \mathbb{K}) \to \mathcal{M}(n\times n, \mathbb{K})$, die die Matrix A auf die Matrix $S^{-1}AS$ abbildet, ein Automorphismus ist mit der Inversen $A \longmapsto SAS^{-1}$.)

Im Spezialfall $A = B$ in (iii) erhalten wir $S^{-1}A^2S = (S^{-1}AS)^2$; mittels Induktion folgt:

(iv) $$S^{-1}A^nS = (S^{-1}AS)^n \ .$$

Setzen wir $A' = S^{-1}AS$, so gilt

(v) $$A^n = SA'^nS^{-1}$$

Beispiel. Wie in 6.2.4 sei $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Dann $S^{-1}AS = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ mit

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad S^{-1} = -\frac{1}{2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Nun können wir Folgendes berechnen:

$$\begin{aligned} A^n &= S(S^{-1}AS)^n S^{-1} = S \cdot 2^n \cdot \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}^n S^{-1} &&= 2^n \cdot S \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix} S^{-1} \\ &= 2^n \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix} \cdot \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} &&= 2^{n-1} \begin{pmatrix} 2^n & (-1)^n \\ 2^n & -(-1)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= 2^{n-1} \begin{pmatrix} 2^n + (-1)^n & 2^n - (-1)^n \\ 2^n - (-1)^n & 2^n + (-1)^n \end{pmatrix}. \end{aligned}$$

6.4. Determinante und Eigenwerte von unitären und orthogonalen Matrizen

Wir kommen nun auf die in Abschnitt 4.7.5 eingeführten orthogonalen bzw. unitären Matrizen zurück. Was wir über diese sagen, gilt genauso für orthogonale und unitäre Abbildungen.

6.4.1 Eigenschaften von orthogonalen [unitären] Matrizen Q.

(i) $|\det(Q)| = 1$ (für orthogonale Matrizen bedeutet dies $\det(Q) = \pm 1$)

(ii) Für jeden Eigenwert λ_i von Q gilt $|\lambda_i| = 1$.
(Die einzig möglichen reellen Eigenwerte sind also $\lambda_i = \pm 1$, es können aber auch komplexe Eigenwerte auftreten.)

(iii) Sind v_1 und v_2 Eigenvektoren zweier verschiedener Eigenwerte λ_1 und λ_2, dann sind v_1 und v_2 orthogonal zueinander.

(iv) Ist U ein Q-invarianter linearer Teilraum von $\mathbb{R}^n$ $[\mathbb{C}^n]$, dann ist $U^\perp$ ebenfalls Q-invariant.

(Bemerkung: Ein linearer Teilraum U von $\mathbb{R}^n$ $[\mathbb{C}^n]$ heißt Q-*invariant*, wenn $v \in U \Rightarrow Qv \in U$ gilt.)

Bemerkung. Die entsprechenden Eigenschaften sind für orthogonale [unitäre] Transformationen φ eines euklidischen [unitären] Vektorraums gültig.

Beweis.

(i) $\det(Q^*) = \det(\overline{Q^T}) = \det(\overline{Q}) = \overline{\det(Q)}$. Aus $Q^*Q = E_n$ (nach 6.4.1 (iv)) folgt

$$1 = \det(E_n) = \det(Q^*Q) = \det(Q^*) \cdot \det(Q) = \overline{\det(Q)} \cdot \det(Q) = |\det(Q)|^2$$

und damit $|\det(Q)| = 1$.

(ii) Sei $\lambda \in \mathbb{C}$ ein Eigenwert von Q. Dann gibt es einen Eigenvektor $0 \neq v \in \mathbb{C}^n$ zum Eigenwert λ. Da Q orthogonal [unitär] ist, erhalten wir

$$\langle v|v\rangle = \langle Qv|Qv\rangle = \langle \lambda v|\lambda v\rangle = \overline{\lambda}\lambda\langle v|v\rangle$$

Aus $\langle v|v\rangle \neq 0$ ergibt sich $\overline{\lambda}\lambda = 1$, d.h. $|\lambda|^2 = 1$, $|\lambda| = 1$.

(iii) Seien $\lambda_1 \neq \lambda_2$ zwei verschiedene Eigenwerte (in $\mathbb{C}$) von Q und seien v_1, v_2 zugehörige Eigenvektoren. Dann gilt

$$\langle v_1|v_2\rangle = \langle Qv_1|Qv_2\rangle = \langle \lambda_1 v_1|\lambda_2 v_2\rangle = \overline{\lambda}_1\lambda_2\langle v_1|v_2\rangle\ .$$

Wir folgern $\langle v_1|v_2\rangle = 0$ (d.h. $v_1 \perp v_2$) oder $\overline{\lambda}_1\lambda_2 = 1$. Im zweiten Fall ist $\lambda_1 = \overline{\lambda}_2^{-1}$; aus $|\lambda_2| = 1$ folgt $\overline{\lambda}_2^{-1} = \lambda_2$ und damit ergibt sich $\lambda_2 = \lambda_1$ im Widerspruch zur Annahme.

(iv) Wir müssen zeigen, dass $v \in U^\perp$ die Gültigkeit von $Qv \in U^\perp$ nach sich zieht. Dazu sei $v \in U^\perp$. Es folgt $\langle v|u\rangle = 0$ für alle $u \in U$. Da Q orthogonal [unitär] ist, schließen wir $\langle Qv|Qu\rangle = 0$ für alle $u \in U$. Da Q orthogonal [unitär] ist und da U nach Annahme Q-invariant ist, induziert Q eine orthogonale [unitäre] Abbildung von U auf sich selbst. Q induziert also insbesondere eine bijektive Abbildung von U auf sich. Wir folgern $\langle Qv|w\rangle = 0$ für alle $w \in U$, also $Qu \in U^\perp$. □

Sei

$\mathrm{O}(n)$	:=	Menge aller orthogonalen	$n \times n$-Matrizen Q
$\mathrm{U}(n)$	:=	Menge aller unitären	$n \times n$-Matrizen Q
$\mathrm{SO}(n)$	:=	Menge aller orthogonalen	$n \times n$-Matrizen Q mit $\det Q = +1$
$\mathrm{SU}(n)$	:=	Menge aller unitären	$n \times n$-Matrizen Q mit $\det Q = +1$

Schon in 4.7.7 haben wir gesehen, daß $\mathrm{O}(n)$ und $\mathrm{U}(n)$ Untergruppen der Gruppen $\mathrm{GL}(n, \mathbb{R})$ und $\mathrm{GL}(n, \mathbb{C})$ aller invertierbaren reellen und komplexen $n \times n$-Matrizen sind. Insbesondere sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ somit Gruppen bzgl. der Matrixmultiplikation.

Übung. Zeige, dass $\mathrm{SO}(n)$ und $\mathrm{SU}(n)$ Untergruppen von $\mathrm{O}(n)$ und $\mathrm{U}(n)$ sind. Sie heißen *spezielle orthogonale* und die *spezielle unitäre Gruppe*.

6.4.2 Orthogonale 2×2-Matrizen Q. Es sei Q eine orthogonale 2×2-Matrix. Dann ist entweder

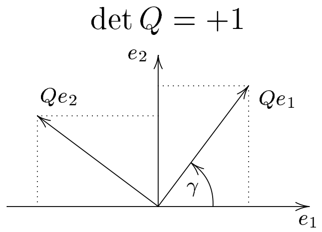
- $\det Q = +1$: dann beschreibt Q eine Drehung um den Ursprung um einen Winkel γ $(0 \leq \gamma \leq 2\pi)$:

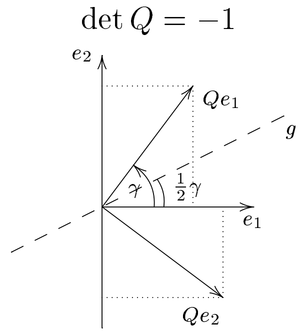
$$Q = D(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$$

- oder $\det Q = -1$: Dann kann Q als eine Spiegelung R an der x_1-Achse mit anschließender Drehung aufgefasst werden:

$$Q = \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & -\cos\gamma \end{pmatrix} = D(\gamma) \cdot R = D(\gamma) \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dies ist in der Tat eine Spiegelung an der Geraden g, die den Winkel $\frac{\gamma}{2}$ mit der x_1-Achse einschließt.





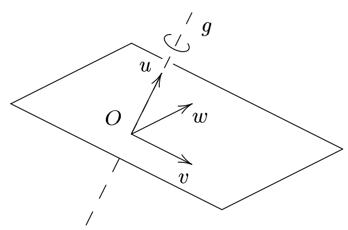
Alle orthogonalen 2×2-Matrizen sind von dieser Gestalt.

Damit sind die einzigen orthogonalen Transformationen des $\mathbb{R}^2$ die Drehungen um den Ursprung und die Spiegelungen an Geraden durch O. Das Vorzeichen der Determinante gibt an, welcher der beiden Fälle vorliegt.

6.4.3 Orthogonale 3×3-Matrizen. Sei $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ eine orthogonale Transformation und sei Q die Matrix von φ bzgl. der Standardbasis e_1, e_2, e_3.

Wir zeigen Folgendes:

1. Ist $\det Q = +1$, dann ist φ eine Drehung um eine Achse g durch den Ursprung O. Die Achse g wird durch einen Eigenvektor u zum Eigenwert 1 gegeben.
   Für den Winkel γ der Drehung gilt (siehe 4.9.4)
   $$\cos\gamma = \frac{1}{2}(\mathrm{tr}(Q) - 1)$$



2. Ist $\det Q = -1$, dann ist φ eine Spiegelung an einer Ebene U gefolgt von einer Drehung um die Gerade g, die orthogonal zu U ist. Also ist φ eine Spiegelung oder eine Drehspiegelung.

   Die Achse g wird erzeugt durch einen Eigenvektor u zum Eigenwert $\lambda = -1$.

Wählt man eine Orthogonalbasis u, v, w, wie in der Zeichnung angedeutet ist, so lautet die Matrix Q' von φ bzgl. dieser neuen Basis

1. $\det Q = +1$: $\qquad Q' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$
2. $\det Q = -1$: $\qquad Q' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$

Im zweiten Fall liegt eine Spiegelung vor, wenn $\gamma = 0$ ist. Es handelt sich um eine Spiegelung gefolgt von einer Drehung, wenn $\gamma \neq 0$ gilt.

Beweis. Wir bemerken zunächst, daß nach 6.4.1 entweder $\det Q = 1$ oder -1.

Da Q eine reelle 3×3-Matrix ist, hat das charakteristische Polynom von Q reelle Koeffizienten und ist vom Grad 3, also von ungeradem Grad. Jedes reelle Polynom ungeraden Grades besitzt mindestens eine reelle Nullstelle λ_1 (dies ist ein Satz aus der Analysis). Aus 6.4.1 (ii) folgt $\lambda_1 = \pm 1$.

Erster Fall: $\det Q = 1$. Wir behaupten: Q hat den Eigenwert 1. Seien $\lambda_1, \lambda_2, \lambda_3$ die (evtl. komplexen) Eigenwerte von Q, wobei $\lambda_1 = \pm 1$ nach dem oben Gesagten. Ist λ_2 echt komplex, dann auch λ_3 und es gilt $\lambda_3 = \overline{\lambda_2}$. In diesem Fall gilt $1 = \det Q = \lambda_1 \lambda_2 \lambda_3 = \lambda_1 |\lambda_2|^2 = \lambda_1$, also $\lambda_1 = 1$. Ist also $\lambda_1 = -1$, dann sind λ_2, λ_3 beide reell vom Betrag 1 und so muß einer von ihnen 1 sein.

Wir wählen nun einen normierten Eigenvektor u zum Eigenwert 1.

Sei $U = \{u\}^\perp$ der zu u orthogonale zweidimensionale lineare Teilraum. Nach 6.4.1 (iv) ist U invariant unter φ. Also wirkt φ als orthogonale Abbildung auf diesem zweidimensionalen Raum. Aus 6.4.2 wissen wir, daß φ auf diesem zweidimensionalen Raum U als Drehung oder Spiegelung wirkt. Also ist φ eine Drehung um die Achse g (denn würde φ auf U als Spiegelung wirken, dann wäre $\det Q = -1$).

Zweiter Fall: $\det Q = -1$. Mittels einer ähnlichen Überlegung wie im vorigen Fall sieht man, daß -1 ein Eigenwert sein muß. Wir wählen dann einen normierten Eigenvektor u zum Eigenwert -1. Wieder ist $U := \{u\}^\perp$ invariant unter φ und wir fahren wie im ersten Fall fort. Auf U wirkt φ als identische Abbildung (in diesem Fall ist φ die Spiegelungan der Ebene U) oder als Drehung (und in diesem Fall ist φ eine Drehspiegelung, genauer eine Spiegelung an der Ebene U gefolgt von einer Drehung um die Achse g).

$\square$

KAPITEL 7

Diagonalisierung: Hauptsätze

In diesem Kapitel werden wir die komplexen Matrizen A beschreiben, die mittels einer unitären Matrix Q diagonalisiert werden können. Wir wollen also die komplexen Matrizen A charakterisieren, welche die Eigenschaft besitzen, daß es eine unitäre Matrix Q gibt, so daß

$$A' := Q^{-1}AQ \quad \text{eine Diagonalmatrix ist.}$$

Da nach 4.7.7 für unitäre Matrizen $Q^{-1} = Q^*$ gilt, ist dies äquivalent zu der Aussage, daß

$$A' := Q^*AQ \quad \text{eine Diagonalmatrix ist.}$$

In diesem Kapitel seien $A, B, \ldots$ immer $n \times n$-Matrizen mit komplexen oder reellen Einträgen und $\mathbb{R}^n$ und $\mathbb{C}^n$ seien immer mit dem Standardskalarprodukt versehen.

7.1. Normale Matrizen

Wir definieren für Matrizen A mit komplexen Einträgen den Begriff der adjungierten Matrix A^*. Man beachte aber, dass dieser Begriff einer Adjungierten nichts mit dem Begriff der Adjunkten gemein hat, wie er in 5.4.6 definiert wurde.

7.1.1 Die adjungierte Matrix. Für eine $n \times m$-Matrix mit komplexen Einträgen

$$A = (\alpha_{ij})$$

setzen wir wie schon früher

$$\overline{A} = (\overline{\alpha_{ij}})$$

$$A^* = \overline{A}^T = \overline{A^T}$$

und wir nennen A^* die *Adjungierte von A*. Für die Adjungierte gelten folgende Regeln, wobei wir zum Beweis die entsprechenden Regeln für die Transponierte verwenden:

$$(A+B)^* = A^* + B^* \qquad (\lambda A)^* = \overline{\lambda} A^* \qquad (\text{für } \lambda \in \mathbb{C})$$

$$(AB)^* = B^*A^* \qquad A^{**} = A$$

Die beiden ersten und die vierte Regel sind offensichtlich; zum Nachweis der dritten:

$$(AB)^* = (\overline{AB})^T = (\overline{A}\,\overline{B})^T = \overline{B}^T\overline{A}^T = B^*A^*$$

Sind alle Einträge α_{ij} von A reell, so gilt $\overline{A} = A$ und $A^* = A^T$.

7.1.2 Spezialfälle. Für eine komplexe $n \times n$-Matrix A gilt:

$$A \text{ ist } \textit{reell} \iff A = \overline{A}$$

$$A \text{ ist } \textit{rein imaginär} \iff A = -\overline{A}$$

Wir definieren weiterhin:

$$A \text{ ist } \textit{symmetrisch} \iff A = A^T$$

$$\begin{aligned} A \text{ ist schiefsymmetrisch} &\iff A = -A^T \\ A \text{ ist } \textit{hermitesch } (= \textit{selbstadjungiert}) &\iff A = A^* \\ A \text{ ist } \textit{schiefhermitesch} &\iff A = -A^* \end{aligned}$$

Bemerkung. Man beachte, dass die rellen (schief-)hermiteschen Matrizen genau die (schief-) symmetrischen sind. Einige Beispiele:

symmetrisch	schiefsymmetrisch	hermitesch	schiefhermitesch
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 2 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -i & 1-i \\ i & 2 & 2+i \\ 1+i & 2-i & -2 \end{pmatrix}$	$\begin{pmatrix} i & i & 1-i \\ i & 2i & 2+i \\ -1-i & -2+i & -2i \end{pmatrix}$

Jede komplexe $n \times n$-Matrix A kann dargestellt werden als

a) Summe einer reellen und einer rein imaginären Matrix:

$$A = \frac{1}{2}(A + \overline{A}) + \frac{1}{2}(A - \overline{A})$$

b) Summe einer symmetrischen und einer schiefsymmetrischen Matrix:

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

c) Summe einer hermiteschen und einer schiefhermiteschen Matrix:

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*)$$

Bemerkung. Für einen Spaltenvektor $v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{C}^n$ ist der Adjungierte ein Zeilenvektor $v^* = (\overline{\alpha_1} \quad \dots \quad \overline{\alpha_n})$ und das übliche hermitesche Skalarprodukt läßt sich schreiben:

$$\langle v|w\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\beta_i = (\overline{\alpha_1} \quad \dots \quad \overline{\alpha_n}) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = v^* w$$

Sei A eine $n \times n$-Matrix mit komplexen Einträgen.

7.1.3 Proposition. $\boxed{\langle v|Aw\rangle = \langle A^*v|w\rangle}$ *und* $\boxed{\langle Av|w\rangle = \langle v|A^*w\rangle}$ *für alle* $v, w \in \mathbb{C}^n$.

Beweis. $\langle v|Aw\rangle = v^*Aw = v^*A^{**}w = (A^*v)^*w = \langle A^*v|w\rangle$. Wegen $A^{**} = A$ folgt aus dem eben Gezeigten $\langle Av|w\rangle = \langle A^{**}v|w\rangle = \langle v|A^*w\rangle$. $\square$

Ist A eine *reelle* Matrix, so gilt $A^* = A^T$ und folglich:

$$\langle v|Aw\rangle = \langle A^Tv|w\rangle \quad , \quad \langle Av|w\rangle = \langle v|A^Tw\rangle.$$

Kommen wir nun zum zentralen Begriff dieses Kapitels:

7.1.4 Normale Matrizen. Eine komplexe $n \times n$-Matrix A wird *normal* genannt, wenn sie mit ihrer Adjungierten kommutiert, d.h. falls gilt:

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

Beispiele:

a) $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ ist nicht normal, da $A^TA = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = AA^T$ gilt.

b) Diagonalmatrizen sind normal.

c) Unitäre und (reelle) orthogonale Matrizen Q sind normal. (Ist nämlich Q unitär, so gilt $Q^*Q = E = QQ^*$ nach 4.7.7.)

d) Reelle symmetrische und reelle schiefsymmetrische Matrizen sind normal.

(Ist nämlich A reell und $A = \pm A^T$, so folgt $A^* = \overline{A}^T = A^T = \pm A$ und damit $A^*A = \pm AA = \pm A^2 = A \cdot (\pm A) = AA^*$.)

e) Komplexe hermitesche und schiefhermitesche Matrizen sind normal.

(Ist nämlich $A = \pm A^*$, dann gilt $A^*A = \pm AA = \pm A^2 = A \cdot (\pm A) = AA^*$.)

Diese Beispiele zeigen, dass es eine große Anzahl normaler Matrizen gibt.

7.1.5 Proposition. *Ist A normal, dann ist*

(a) *$A - \lambda E$ normal für jedes $\lambda \in \mathbb{C}$ und*

(b) *$A' = Q^*AQ$ ist normal für jede unitäre Matrix Q.*

Beweis.

$$(A - \lambda E)(A - \lambda E)^* = (A - \lambda E)(A^* - \overline{\lambda}E) = AA^* - \lambda A^* - \overline{\lambda}A + \lambda\overline{\lambda}E$$

$$(A - \lambda E)^*(A - \lambda E) = (A^* - \overline{\lambda}E)(A - \lambda E) = A^*A - \overline{\lambda}A - \lambda A^* + \overline{\lambda}\lambda E$$

und die rechten Seiten sind gleich, da $AA^* = A^*A$ wegen der Voraussetzung der Normalität. Analog

$$(Q^*AQ)(Q^*AQ)^* = (Q^*A\underbrace{Q)(Q^*}_{E}A^*Q^{**}) = Q^*AA^*Q$$

$$(Q^*AQ)^*(Q^*AQ) = (Q^*A^*Q^{**})(Q^*AQ) = Q^*A^*\underbrace{QQ^*}_{E}AQ = Q^*A^*AQ$$

und die rechten Seiten sind gleich, da $AA^* = A^*A$ wegen der Voraussetzung der Normalität. $\square$

7.1.6 Proposition. *Sei A eine normale Matrix. Ist $\lambda \in \mathbb{C}$ ein Eigenwert von A und $v \in \mathbb{C}^n$ ein Eigenvektor von A bzgl. λ, dann ist $\overline{\lambda}$ ein Eigenwert von A^* und v ist ein Eigenvektor von A^* bzgl. $\overline{\lambda}$.*

Beweis. Da A normal ist, gilt:

$$\langle (A - \lambda E)v \mid (A - \lambda E)v \rangle \underset{7.1.3}{=} \langle (A - \lambda E)^*(A - \lambda E)v \mid v \rangle$$

$$\underset{7.1.5}{=}\langle (A-\lambda E)(A-\lambda E)^* v \mid v\rangle \underset{7.1.3}{=} \langle (A-\lambda E)^* v \mid (A-\lambda E)^* v\rangle \quad = \langle (A^*-\overline{\lambda}E)v \mid (A^*-\overline{\lambda}E)v\rangle$$

v ist aber genau dann ein Eigenvektor von A zum Eigenwert λ, wenn

$$Av = \lambda v \iff (A-\lambda E)v = 0 \iff \langle (A-\lambda E)v \mid (A-\lambda E)v\rangle = 0$$

Ebenso ist v genau dann ein Eigenvektor von A^* zum Eigenwert $\overline{\lambda}$, wenn

$$A^* v = \overline{\lambda} v \iff (A^*-\overline{\lambda}E)v = 0 \iff \langle (A^*-\overline{\lambda}E)v \mid (A^*-\overline{\lambda}E)v\rangle = 0$$

Da nach der ersten Rechnung die beiden Ausdrücke für normale Matrizen A gleich sind, erhalten wir das gewüschte Ergebnis. □

7.1.7 Proposition. *Ist A eine normale Matrix, dann sind die Eigenvektoren von A zu verschiedenen Eigenwerten orthogonal.*

Beweis. Wir nehmen an, v und w seien Eigenvektoren zu verschiedenen Eigenwerten λ und μ, d.h. es gilt $Av = \lambda v$ und $Aw = \mu w$. Es folgt

$$\mu \cdot \langle v|w\rangle \underset{(\mathrm{B}_2)}{=} \langle v|\mu w\rangle = \langle v|Aw\rangle \underset{(7.1.3)}{=} \langle A^* v|w\rangle \underset{(7.1.6)}{=} \langle \overline{\lambda} v|w\rangle \underset{(\mathrm{B'}_1)}{=} \lambda \cdot \langle v|w\rangle$$

Wegen $\mu \neq \lambda$ ergibt sich $\langle v|w\rangle = 0$. Also sind v und w orthogonal. □

7.1.8 Proposition. *Ist A eine normale Matrix und $v \in \mathbb{C}^n$ ein Eigenvektor von A, dann ist der $(n-1)$-dimensionale lineare Teilraum $v^\perp$ invariant unter A.*

Beweis. Sei $\lambda \in \mathbb{C}$ der Eigenwert, für den v ein Eigenvektor ist. Sei $w \in v^\perp$, d.h. $\langle v|w\rangle = 0$. Wir wollen $\langle v|Aw\rangle = 0$ zeigen:

$$\langle v|Aw\rangle \underset{7.1.3}{=} \langle A^* v|w\rangle \underset{7.1.6}{=} \langle \overline{\lambda} v|w\rangle \underset{(\mathrm{B'}_1)}{=} \lambda \cdot \langle v|w\rangle = 0.$$

□

7.2. Diagonalisierungssätze

Nun können wir unsere wichtigsten Sätze angeben.

7.2.1 Hauptssatz über normale Matrizen. *Für eine $n \times n$-Matrix A mit komplexen Einträgen sind die folgenden Eigenschaften äquivalent:*

(i) *A ist normal.*

(ii) *Es gibt in $\mathbb{C}^n$ eine Orthonormalbasis von Eigenvektoren von A.*

(iii) *Es gibt eine unitäre Matrix Q, so daß*

$$A' = Q^* A Q$$

eine Diagonalmatrix ist.

Beweis. (ii) $\Rightarrow$ (iii). Angenommen, $\mathcal{B}' : v_1, v_2, \ldots, v_n$ sei eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren von A. Sei Q die Matrix, deren Spalten die Vektoren $v_1, \ldots, v_n$ sind. Dann ist

$$A' = Q^{-1}AQ$$

eine Diagonalmatrix (die Einträge auf der Diagonalen von A' sind die zugehörigen Eigenwerte, siehe 6.3.2). Da Q unitär ist, folgt $Q^{-1} = Q^*$ aus 4.7.7.

(iii) $\Rightarrow$ (i). Eine Diagonalmatrix A' ist immer normal. Gilt $A' = Q^*AQ$ für eine unitäre Matrix Q, so folgt $QA'Q^* = QQ^*AQQ^* = A$. Wir folgern aus 7.1.5 (b), dass A ebenfalls normal ist.

(i) $\Rightarrow$ (ii). Dies beweisen wir durch Induktion nach n.

Für $n = 1$ ist die Behauptung offensichtlich. Wir nehmen an, wir haben die Implikation für alle normalen $(n-1) \times (n-1)$-Matrizen bewiesen, und betrachten eine normale $n \times n$-Matrix A. Eine komplexe $n \times n$-Matrix hat immer einen Eigenwert $\lambda_1 \in \mathbb{C}$. Sei v_1 ein Eigenvektor von A zu diesem Eigenwert λ_1. Wir nehmen $\|v_1\| = 1$ an. Nach 7.1.8 ist der $(n-1)$-dimensionale lineare Teilraum $v_1^\perp$ invariant unter A. Nach der Induktionsannahme hat $v_1^\perp$ eine Orthonormalbasis von Eigenvektoren $v_2, \ldots, v_n$ von A. Damit ist $v_1, v_2, \ldots, v_n$ eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren von A. □

Bemerkung. Der letzte Beweis enthält eine Lücke: Da der $(n-1)$-dimensionale Teilraum $v_1^\perp$ nicht gleich $\mathbb{C}^{n-1}$ ist, können wir die Induktionsannahme nicht direkt anwenden. Wir können aber wie folgt schließen:

Sei $w_2, \ldots, w_n$ eine Orthonormalbasis des $(n-1)$-dimensionalen linearen Teilraums $v_1^\perp$. Sei Q die Transformationsmatrix zwischen der Standardbasis $\mathcal{B}$ von $\mathbb{C}^n$ und der Basis

$$v_1, w_2, \ldots, w_n \,.$$

Da diese Basis orthonormal ist, ist Q eine unitäre Matrix, d.h. es gilt $Q^{-1} = Q^*$. Transformieren wir die gegebene Matrix A in die neue Basis $v_1, w_2, \ldots, w_n$, so erhält diese die Form

$$A_1 = Q^*AQ = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{array}\right)$$

Nach 7.1.5 (b) ist A_2 ebenfalls normal. Deshalb ist A_2 eine normale $(n-1) \times (n-1)$-Matrix und wir können die Induktionsannahme auf A_2 anwenden.

Wir wenden nun weiterhin vorigen Satz auf die speziellen Klassen normaler Matrizen an, die in den Beispielen von 7.1.4 betrachtet wurden. Wir beginnen mit hermiteschen (selbstadjungierten) Matrizen:

7.2.2 Lemma. *Alle (komplexen) Eigenwerte einer hermiteschen Matrix sind reell. (Insbesondere sind alle komplexen Eigenwerte einer reellen symmetrischen Matrix reell.)*

Beweis. Sei $A = A^*$ und $\lambda \in \mathbb{C}$ ein Eigenwert von A. Dann gibt es einen nichtverschwindenden Vektor v mit

$$Av = \lambda v \,.$$

Aus 7.1.6 folgt $A^*v = \overline{\lambda}v$ und aus $A^* = A$ folgern wir $Av = \overline{\lambda}v$ und damit $\lambda v = \overline{\lambda}v$, was $\lambda = \overline{\lambda}$ impliziert. Also ist λ reell. □

7.2.3 Hauptsatz über hermitesche Matrizen. *Sei A eine komplexe $n \times n$-Matrix. Dann ist A genau dann hermitesch, wenn es eine unitäre Matrix Q gibt, so dass*

$$A' := Q^*AQ$$

eine reelle Diagonalmatrix ist.

Beweis. (i) Sei A hermitesch. Da dann A normal ist, gibt es nach 7.2.1 eine unitäre Matrix Q, so dass

$$A' := Q^*AQ$$

eine Diagonalmatrix ist. Die Einträge λ_i auf der Diagonalen von A' sind die Eigenwerte von A. Wegen 7.2.2 sind alle λ_i reell.

(ii) Nehmen wir an, es gibt eine unitäre Matrix Q derart, daß $A' = Q^*AQ$ eine reelle Diagonalmatrix ist. Es folgt dann $A = QA'Q^*$ somit

$$A^* = (QA'Q^*)^* = Q^{**}A'^*Q^* = QA'Q^* = A\,,$$

d.h. A ist hermitesch. □

7.2.4 Hauptssatz über reelle symmetrische Matrizen. *Für eine $n \times n$-Matrix A mit* reellen *Einträgen sind die folgenden Eigenschaften äquivalent:*

(i) *A ist symmetrisch.*

(ii) *Es gibt in $\mathbb{R}^n$ eine Orthonormalbasis von Eigenvektoren von A.*

(iii) *Es gibt eine orthogonale Matrix Q, so dass*

$$A' := Q^TAQ$$

eine (reelle) Diagonalmatrix ist.

Beweis. (i) $\Rightarrow$ (ii). Die Vorgehensweise geschieht genau wie im Beweis von 7.2.1. Da nach 7.2.2 alle Eigenwerte von A reell sind und da A eine reelle Matrix ist, kann man reelle Eigenvektoren für alle Eigenwerte finden und der Beweis in 7.2.1 kann vollständig im Reellen durchgeführt werden.

(ii) $\Rightarrow$ (iii): Genau wie in 7.2.1.

(iii) $\Rightarrow$ (i): Angenommen, Q ist eine orthogonale Matrix und $A' = Q^TAQ$ ist eine Diagonalmatrix, dann gilt $A = QA'Q^T$ und damit $A^T = (QA'Q^T)^T = Q^{tt}A'^TQ^T = QA'Q^T = A$, d.h. A ist symmetrisch. □

Bemerkung. Eine komplexe symmetrische Matrix muss nicht normal sein; wähle z.B.

$$A = \begin{pmatrix} i & 1 \\ 1 & 0 \end{pmatrix}, \quad A^* = \begin{pmatrix} -i & 1 \\ 1 & 0 \end{pmatrix}; \qquad AA^* = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix} = A^*A.$$

Ist eine komplexe symmetrische Matrix normal, so müssen deren Eigenwerte nicht reell sein. Betrachte z.B. die Matrix $A = \begin{pmatrix} 1 & i \\ i & 0 \end{pmatrix}$, welche normal und symmetrisch ist; deren Eigenwerte sind $\lambda_{1/2} = \frac{1 \pm i\sqrt{3}}{2}$.

7.2.5 Beispiel. Betrachte die Matrix

$$A = \frac{1}{9}\begin{pmatrix} 2 & 2 & 10 \\ 2 & 11 & -8 \\ 10 & -8 & 5 \end{pmatrix}$$

A ist reell und symmetrisch. Also hat diese Matrix im $\mathbb{R}^3$ eine Orthonormalbasis aus Eigenvektoren v_1, v_2, v_3. Wir wollen eine solche Orthonormalbasis aus Eigenvektoren explizit berechnen:

Charakteristisches Polynom:

$$\det(A - \lambda E) = \begin{vmatrix} \frac{2}{9} - \lambda & 2/9 & 10/9 \\ 2/9 & \frac{11}{9} - \lambda & -8/9 \\ 10/9 & -8/9 & \frac{5}{9} - \lambda \end{vmatrix} = -\lambda^3 + 2\lambda^2 + \lambda - 2$$

Eigenwerte: $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$.

Eigenvektoren: $w_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$, $w_2 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$ und $w_3 = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}$.

Da es drei verschiedene Eigenwerte gibt, stehen die zugehörigen Eigenvektoren paarweise orthogonal aufeinander (nach 7.1.7). Also ergibt normieren der w_1, w_2, w_3 eine Orthonormalbasis des $\mathbb{R}^3$:

$$v_1 = \frac{1}{3}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \ v_2 = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \ v_3 = \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}; \ Q := \frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}.$$

Wegen $Q^T AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ können wir A tatsächlich mittels einer orthogonalen Matrix Q diagonalisieren.

7.2.6 Normalform für unitäre Matrizen. Sei A eine unitäre Matrix. Wir wissen aus 7.1.4, dass A normal ist, also können wir unseren Hauptstatz 7.2.1 anwenden, um zu erkennen, dass es eine andere unitäre Matrix Q gibt, so dass

$$A' = Q^* AQ \quad \text{eine Diagonalmatrix ist}\,.$$

Die Einträge λ_j auf der Diagonalen von A' sind die Eigenwerte von A. Aus 6.4.1 (ii) erhalten wir $|\lambda_j| = 1$. Also ist λ_j eine komplexe Zahl der Form

$$\lambda_j = e^{i\gamma_j} = \cos\gamma_j + i\sin\gamma_j.$$

Deshalb gilt

$$Q^* AQ = A' = \begin{pmatrix} e^{i\gamma_1} & & & 0 \\ & e^{i\gamma_2} & & \\ & & \ddots & \\ 0 & & & e^{i\gamma_n} \end{pmatrix}.$$

Die Spalten $v_1, v_2, \ldots, v_n$ von Q bilden eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren zu den Eigenwerten $e^{i\gamma_1}$, $e^{i\gamma_2}$, $\ldots$, $e^{i\gamma_n}$.

7.2.7 Normalform für orthogonale Matrizen. Nun sei A eine (reelle) orthogonale $n \times n$-Matrix. Wir behaupten, es gibt eine orthogonale Matrix Q mit

$$Q^TAQ = A' = \begin{array}{c} \begin{array}{ccc} \overbrace{\hspace{4em}}^{p} & \overbrace{\hspace{5em}}^{q} & \overbrace{\hspace{6em}}^{2r} \end{array} \\ \begin{pmatrix} \boxed{\begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix}} & & \text{\Large 0} \\ & \boxed{\begin{matrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{matrix}} & \\ \text{\Large 0} & & \begin{matrix} \boxed{D_1} & & \\ & \ddots & \\ & & \boxed{D_r} \end{matrix} \end{pmatrix} \end{array}$$

wobei gilt:

- p ist die Vielfachheit des Eigenwertes $+1$ von A;
- q ist die Vielfachheit des Eigenwertes -1 von A;
- r ist die Anzahl von Paaren komplex konjugierter Eigenwerte $e^{i\gamma_j}$, $e^{-i\gamma_j}$ von A.

Jedes D_j ist eine Matrix der Form

$$D_j = \begin{pmatrix} \cos\gamma_j & -\sin\gamma_j \\ \sin\gamma_j & \cos\gamma_j \end{pmatrix}$$

Zum Nachweis dieser Behauptung argumentieren wir wie folgt:

Eine orthogonale Matrix ist ein Spezialfall einer unitären Matrix. Wir gehen deshalb zunächst wie für unitäre Matrizen in 7.2.6 vor:

- Bestimme die komplexen Eigenwerte von A:

  Sei p die Vielfachheit des Eigenwertes $+1$ und sei q die Vielfachheit des Eigenwertes -1. (Ist $+1$ oder -1 kein Eigenwert von A, sei $p = 0$ oder entsprechend $q = 0$.)

  Die anderen Eigenwerte haben die Form $e^{i\gamma} = \cos\gamma + i\sin\gamma$. Da A eine reelle Matrix ist, hat das charakteristische Polynom $P(\lambda)$ reelle Koeffizienten. Ist $\lambda = a + ib$ eine komplexe Wurzel von $P(\lambda)$, dann ist $\overline{\lambda} = a - ib$ ebenso eine Wurzel von $P(\lambda)$. Also ist für jeden Eigenwert von A mit der Form

  $$e^{i\gamma_j} = \cos\gamma_j + \sin\gamma_j$$

  die komplex konjugierte Zahl

  $$e^{-i\gamma_j} = \cos\gamma_j - i\sin\gamma_j$$

ebenfalls ein Eigenwert von A. Wir ordnen diese Eigenwerte nach Paaren von komplex konjugierten Zahlen. Die Eigenwerte ordnen wir wie folgt

$$\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{e^{i\gamma_1},e^{-i\gamma_1},\dots,e^{i\gamma_r},e^{-i\gamma_r}}_{2r}$$

nach ihrer Vielfachheit.

- Wie in 7.2.6 können wir in $\mathbb{C}^n$ eine Orthonormalbasis von Eigenvektoren zu diesen Eigenwerten finden:

$$v_1,\dots,v_p,v_{p+1},\dots,v_{p+q},w_1,\overline{w}_1,\dots,\overline{w}_r,\overline{w}_r$$

Man beachte, dass die Eigenvektoren v_1, $\dots$, v_p, v_{p+1}, $\dots$, v_{p+q} zu den reellen Eigenwerten ± 1 reell gewählt werden können. Man kann sogar für die komplex konjugierten Eigenwerte die Eigenvektoren komplex konjugiert wählen: Aus $Aw=\lambda w$ folgt $\overline{Aw}=\overline{\lambda w}$, da aber die Matrix A reell ist und da $A=\overline{A}$ gilt, folgern wir $A\overline{w}=\overline{\lambda}\overline{w}$.

In dem zweidimensionalen komplexen Vektorraum, der von w_j und $\overline{w}_j$ aufgespannt wird, bilden wir die Vektoren

$$u_j := \frac{1}{\sqrt{2}}(w_j+\overline{w}_j)\,,\ u'_j=\frac{i}{\sqrt{2}}(w_j-\overline{w}_j)$$

Natürlich sind u_j und u'_j *reelle* Vektoren, sie haben die Länge 1 und sind orthogonal:

$$\begin{aligned}\langle u_j|u'_j\rangle &= \frac{i}{2}(\underbrace{\langle w_j|w_j\rangle}_{1}+\underbrace{\langle \overline{w}_j|w_j\rangle}_{0}-\underbrace{\langle w_j|\overline{w}_j\rangle}_{0}-\underbrace{\langle \overline{w}_j|\overline{w}_j\rangle}_{1})\\ &=0\end{aligned}$$

Bzgl. u_j,u'_j gehört zur Wirkung von A die Matrix (vgl. 8.4.9)

$$D_{\gamma_j}=\begin{pmatrix}\cos\gamma_j & -\sin\gamma_j\\ \sin\gamma_j & \cos\gamma_j\end{pmatrix}.$$

7.3. Die Adjungierte eines Endomorphismus

Jede Operation auf Matrizen korrespondiert mit Operationen auf linearen Abbildungen. Wir kennen bereits die Adjungierte A^* einer Matrix A. Was ist die Adjungierte einer linearen Abbildung?

Im Folgenden sei V ein euklidischer [unitärer] Vektorraum und $\varphi: V\to V$ ein Endomorphismus.

7.3.1 Die Adjungierte eines Endomorphismus. Eine Abbildung $\varphi^*: V\to V$ heißt *adjungiert zu* φ, wenn

$$\langle\varphi^*(v)|w\rangle=\langle v|\varphi(w)\rangle \qquad \text{für alle } v,w\in V$$

gilt. Wie wir in 7.1.3 gesehen haben, ist die Adjungierte einer linearen Abbildung $\varphi:\mathbb{C}^n\to\mathbb{C}^n$, die durch eine Matrix A gegeben ist, diejenige lineare Abbildung, die durch die adjungierte Matrix A^* gegeben ist.

Es ist i.A. nicht offensichtlich, dass eine gegebene lineare Abbildung $\varphi : V \to V$ eine Adjungierte besitzt. Besitzt sie eine Adjungierte, so ist nicht offensichtlich, ob die Adjungierte linear und eindeutig ist:

7.3.2 Eindeutigkeit und Linearität der Adjungierten. *Hat ein Endomorphismus $\varphi : V \to V$ eine Adjungierte φ^*, dann ist diese Adjungierte φ^* linear und durch φ eindeutig bestimmt.*

Beweis.

(i) *Eindeutigkeit:* Angenommen, φ^* und φ' sind zwei Adjungierte von φ, d.h. es gilt

$$\langle \varphi^*(v)|w\rangle = \langle v|\varphi(w)\rangle \qquad \text{und} \qquad \langle \varphi'(v)|w\rangle = \langle v|\varphi(w)\rangle$$

für alle $v, w \in V$. Dann folgern wir $\langle \varphi^*(v)|w\rangle - \langle \varphi'(v)|w\rangle = 0$, d.h.

$$\langle \varphi^*(v) - \varphi'(v), w\rangle = 0$$

Da dies für alle $v, w \in V$ gilt, können wir $w = \varphi^*(v) - \varphi'(v)$ wählen und erhalten

$$\langle \varphi^*(v) - \varphi'(v), \varphi^*(v) - \varphi'(v)\rangle = 0$$

Es folgt $\varphi^*(v) - \varphi'(v) = 0$, d.h. $\varphi^*(v) = \varphi'(v)$ für alle $v \in V$. Damit gilt $\varphi' = \varphi^*$.

(ii) *Linearität:* Wir müssen $\varphi^*(u+v) = \varphi^*(u) + \varphi^*(v)$ und $\varphi^*(\lambda v) = \lambda\varphi^*(v)$ zeigen. Es gilt für alle $u, v, w \in V$

$$\begin{aligned}\langle \varphi^*(u+v)|w\rangle &= \langle u+v|\varphi(w)\rangle = \langle u|\varphi(w)\rangle + \langle v|\varphi(w)\rangle \\ &= \langle \varphi^*(u)|w\rangle + \langle \varphi^*(v)|w\rangle \\ &= \langle \varphi^*(u) + \varphi^*(v)|w\rangle\end{aligned}$$

Da dies insbesondere für alle $w \in V$ gilt, folgern wir $\varphi^*(u+v) = \varphi^*(u) + \varphi^*(v)$. Analog zeigt man $\varphi^*(\lambda v) = \lambda\varphi^*(v)$. $\square$

Auf unendlichdimensionalen euklidischen Vektorräumen V hat nicht jeder Endomorphismus eine Adjungierte. Dafür muß man sich auf Hilberträume einschränken. Wir zeigen hier, daß auf endlichdimensionalen Räumen die Adjungierte immer existiert:

7.3.3 Existenz der Adjungierten. *Sei V ein endlich dimensionaler euklidischer [unitärer] Vektorraum. Dann hat jeder Endomorphismus φ von V eine Adjungierte φ^*.*

Ist $\mathcal{B} : v_1, \ldots, v_n$ eine Orthonormalbasis *von V und A die Matrix von φ bzgl. $\mathcal{B}$, dann ist A^* die Matrix von φ^* bzgl. $\mathcal{B}$.*

Beweis. Sei $[v]_\mathcal{B}$ der Koordinatenvektor von v bzgl. der Orthonormalbasis $\mathcal{B}$ und A die Matrix von φ bzgl. $\mathcal{B}$.

Wir definieren φ^* als diejenige lineare Abbildung, deren Matrix (bzgl. der Basis $\mathcal{B}$) die Adjungierte A^* ist. Dann gilt

$$\begin{aligned}\langle \varphi^*(v)|w\rangle &= \langle A^*[v]_\mathcal{B} \,|\, [w]_\mathcal{B}\rangle = \langle [v]_\mathcal{B} \,|\, A[w]_\mathcal{B}\rangle \\ &= \langle v, \varphi(w)\rangle \qquad\qquad \text{für alle } v, w \in V.\end{aligned}$$

Damit ist φ^* adjungiert zu φ. $\square$

7.3.4 Normale Endomorphismen. Ein Endomorphismus φ von V wird *normal* genannt, wenn $\varphi \circ \varphi^* = \varphi^* \circ \varphi$ gilt.

Sei V endlichdimensional. Nach 7.3.3 ist φ genau dann normal, wenn dessen Matrix A bzgl. einer Orthonormalbasis $\mathcal{B}$ normal ist. Wir erhalten aus 7.2.1:

Satz. Für einen Endomorphismus φ eines endlichdimensionalen unitären Vektorraumes sind die folgenden Eigenschaften äquivalent:

(i) φ ist normal

(ii) Es gibt in V eine Orthonormalbasis von Eigenvektoren für φ

(iii) Die Matrix von φ ist diagonal bzgl. einer Orthonormalbasis $\mathcal{B}$.

7.3.5 Selbstadjungierte Endomorphismen. Ein Endomorphismus φ eines euklidischen [unitären] endlichdimensionalen Vektorraums heißt *selbstadjungiert*, falls $\varphi = \varphi^*$.

Nach 7.3.3 ist φ genau dann selbstadjungiert, wenn die Matrix A von φ bzgl. irgendeiner Orthonormalbasis symmetrisch [hermitesch] ist.

Aus 7.2.2 folgt, dass alle Eigenwerte eines selbstadjungierten Endomorphismus reell sind.

Wir folgern aus 7.2.4, dass ein selbstadjungierter Endomorphismus eines endlichdimensionalen euklidischen oder unitären Vektorraumes eine Orthonormalbasis von Eigenvektoren hat und durch eine (reelle) Diagonalmatrix (bzgl. dieser Basis) dargestellt werden kann.

KAPITEL 8

Linearformen, Bilinearformen, Quadratische Formen

In diesem Kapitel werden wir die Diagonalisierungssätze aus den vorigen Kapiteln weitreichend anwenden. Wir beschränken uns auf reelle Vektorräume, wenn nicht ausdrücklich etwas anderes gesagt wird. Das meiste läßt sich auf komplexe Vektorräume verallgemeinern.

8.1. Bilinearformen

Sei V ein reeller Vektorraum.

8.1.1 Bilinearformen. Eine *Bilinearform* auf V ist eine Abbildung

$$F : V \times V \to \mathbb{R}$$

die bilinear ist, d.h. für alle $v, v', w, w' \in V$ und alle $\lambda \in \mathbb{R}$ gilt:

(B1) $$F(v + v', w) = F(v, w) + F(v', w) \quad , \quad F(\lambda v, w) = \lambda F(v, w)$$

(B2) $$F(v, w + w') = F(v, w) + F(v, w') \quad , \quad F(v, \lambda w) = \lambda F(v, w)$$

Beispiele.

a) Das Standardskalarprodukt auf $\mathbb{R}^n$, allgemeiner das Skalarprodukt auf einem Euklidischen Vektorraum ist nach Definition eine Bilinearform.

b) Setze für $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ in $\mathbb{R}^2$

$$F(v, w) = x_1 y_1 - x_1 y_2 + 2 x_2 y_1 - x_2 y_2$$

Dies ist eine Bilinearform, denn für feste x_1, x_2 hängt der Ausdruck linear von y_1, y_2 ab, und entsprechend für festes y_1, y_2 linear von x_1, x_2. Unter Benutzung der Matrix

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

können wir diese Bilinearform in der folgenden Form schreiben:

$$F(v, w) = v^T A w = (x_1 \ x_2) \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

c) Sei allgemeiner $A = (\alpha_{ij})$ eine $n \times n$-Matrix mit Einträgen $\alpha_{ij} \in \mathbb{R}$.

Für alle $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ in $\mathbb{R}^n$ sei $\boxed{F(v, w) = v^T A w = \sum_{i,j=1}^{n} \alpha_{ij} x_i y_j}$.

Dann ist F eine Bilinearform auf $\mathbb{R}^n$. Dies erklärt auch den Namen *Bilinearform*: Der F definierende Ausdruck ist linear in den x_i's und analog in den y_j's.

Wir werden sehen, daß jede Bilinearform auf $\mathbb{R}^n$ durch eine Matrix A in der obigen Weise dargestellt werden kann:

8.1.2 Matrix einer Bilinearform. Sei $F(v,w)$ eine Bilinearform auf einem endlichdimensionalen Vektorraum V, und sei $\mathcal{B} : v_1, \ldots, v_n$ eine Basis von V. Wir setzen

$$\alpha_{ij} := F(v_i, v_j) \quad \text{für alle } \; i,j = 1, \ldots, n.$$

Die Matrix $A := (\alpha_{ij}) = \big(F(v_i, v_j)\big)$ wird *Matrix der Bilinearform F bzgl. der Basis $\mathcal{B}$* genannt.

Die Bilinearform F ist durch ihre Matrix A vollständig bestimmt: Seien nämlich $v = \sum_{i=1}^n x_i v_i$ und $w = \sum_{j=1}^n y_j v_j$ beliebige Vektoren in V. Dann gilt

$$F(v,w) = F(\sum_i x_i v_i \,, \sum_j y_j v_j) \underset{(B1)}{=} \sum_i x_i F(v_i \,, \sum_j y_j v_j)$$

$$\underset{(B2)}{=} \sum_i x_i \cdot \sum_j y_j \cdot F(v_i, v_j) = \sum_{i,j=1}^n x_i y_j \cdot \alpha_{ij} = \sum_{i,j=1}^n \alpha_{ij} x_i y_j$$

Also ist $\boxed{F(v,w) = [v]_{\mathcal{B}}{}^T \; A \; [w]_{\mathcal{B}}}$.

8.1.3 Basiswechsel. Sei F eine Bilinearform auf V. Wir betrachten zwei Basen $\mathcal{B} : v_1, \ldots, v_n$ und $\mathcal{B}' : v_1', \ldots, v_n'$ von V. Sei $A = (\alpha_{ij})$ die Matrix von F bzgl. $\mathcal{B}$ und $A' = (\alpha_{ij}')$ die Matrix von F bzgl. $\mathcal{B}'$, d.h. es ist

$$\alpha_{ij} = F(v_i, v_j) \qquad \text{und} \qquad \alpha_{ij}' = F(v_i', v_j')$$

Wir behaupten, daß die folgende Transformationsformel gilt:

$$\boxed{A' = S^T A S}$$

wobei $S = (s_{ij})$ die Transformationsmatrix wie in 4.6.1 ist, d.h. die j-te Spalte $s_j = [v_j']_{\mathcal{B}}$ von S wird durch die Koordinaten der neuen Basisvektoren v_j' bzgl. der alten Basis $\mathcal{B}$ gebildet:

$$v_j' = \sum_{i=1}^n s_{ij} v_i \,.$$

Beweis.

$$\alpha_{ij}' = F(v_i', v_j') = F(\sum_{k=1}^n s_{ki} v_k \,, \sum_{l=1}^n s_{lj} v_l) = \sum_{k=1}^n \sum_{l=1}^n s_{ki} s_{lj} \underbrace{F(v_k, v_l)}_{\alpha_{kl}} = \sum_{k,l=1}^n s_{ki} \alpha_{kl} s_{lj}$$

□

Alternativer Beweis. Aus 8.1.2 folgt

(1) $$F(v,w) = [v]_{\mathcal{B}'}{}^T \; A' \; [w]_{\mathcal{B}'} \qquad \text{und}$$

(2) $$F(v,w) = [v]_{\mathcal{B}}{}^T \; A \; [w]_{\mathcal{B}}$$

Da $[v]_{\mathcal{B}} = S\ [w]_{\mathcal{B}'}$ und $[w]_{\mathcal{B}} = S\ [w]_{\mathcal{B}'}$ nach 4.6.1 gilt, ergibt die letzte Gleichung

$$F(v,w) = (S\ [v]_{\mathcal{B}'})^T A (S\ [w]_{\mathcal{B}'}) = [v]_{\mathcal{B}'}{}^T\ S^T A S\ [w]_{\mathcal{B}'}$$

Vergleicht man diese Beziehung mit (1), folgt direkt $A' = S^T A S$. □

Bemerkung Man beachte, dass die Transformationsformel für die Matrix einer Bilinearform verschieden ist von der Transformationsformel für die Matrix einer linearen Abbildung. Nur wenn $S^{-1} = S^T$ gilt, sind die beiden Transformationsformeln gleich. Dies liegt genau dann vor, wenn die Matrix S orthogonal ist, d.h. für Koordinatentransformationen zwischen Orthonormalbasen $\mathcal{B}$ und $\mathcal{B}'$.

8.1.4 Beispiel. Sei

$$\begin{aligned} F(v,w) &= 2x_1y_2 + x_2y_1 && \text{für } v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= v^T A w && \text{mit } A = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

im $\mathbb{R}^2$ bzgl. der Standardbasis $\mathcal{B} : e_1, e_2$ und sei $\mathcal{B}'$ die Basis $v_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2' = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Dann ist die zugehörige Transformationsmatrix $S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

Sind $\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$, $\begin{pmatrix} y_1' \\ y_2' \end{pmatrix}$ die Koordinaten von v und w bzgl. der neuen Basis, so ist F gegeben durch $F(v,w) = (x_1'\ x_2') S^T A S \begin{pmatrix} y_1' \\ y_2' \end{pmatrix}$ und es folgt

$$S^T A S = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & -3 \end{pmatrix}.$$

Damit hat man für F in den Koordinaten bezüglich der neuen Basis den Ausdruck $F(v,w) = 3x_1'y_1' + x_1'y_2' - x_2'y_1' - 3x_2'y_2'$.

8.1.5 Symmetrische Bilinearformen. Eine Bilinearform $F(v,w)$ auf einem Vektorraum V wird *symmetrisch* genannt, wenn

$$\text{(S)} \qquad F(v,w) = F(w,v) \text{ für alle } v,w \in V.$$

Eine Bilinearform F ist genau dann symmetrisch, wenn ihre Matrix A bzgl. einer beliebigen Basis $\mathcal{B}$ symmetrisch ist. Ist nämlich F symmetrisch, so folgt

$$\alpha_{ij} = F(v_i, v_j) = F(v_j, v_i) = \alpha_{ji}.$$

Ist umgekehrt A symmetrisch, dann gilt $F(w,v) = [w]_{\mathcal{B}}{}^T\ A\ [v]_{\mathcal{B}} = [w]_{\mathcal{B}}{}^T\ A^T\ [v]_{\mathcal{B}}{}^{TT} = ([v]_{\mathcal{B}}{}^T\ A\ [w]_{\mathcal{B}})^T = [v]_{\mathcal{B}}{}^T\ A\ [w]_{\mathcal{B}} = F(v,w)$.

8.2. Quadratische Formen

Sei V ein reeller Vektorraum.

8.2.1 Quadratische Formen. Sei F eine symmetrische Bilinearform auf V. Die Abbildung $Q : V \to \mathbb{K}$, welche durch $Q(x) = F(x,x)$ für alle $x \in V$ definiert ist, wird die mit F *assoziierte quadratische Form* genannt.

F kann aus Q rekonstruiert werden:

$$\begin{aligned}
&\tfrac{1}{2}(Q(x+y) - Q(x) - Q(y))\\
=&\tfrac{1}{2}(F(x+y,x+y) - F(x,x) - F(y,y))\\
=&\tfrac{1}{2}(F(x,x) + F(x,y) + F(y,x) + F(y,y) - F(x,x) - F(y,y))\\
=&\tfrac{1}{2}(F(x,y) + \underbrace{F(y,x)}_{=F(x,y)}) = F(x,y).
\end{aligned}$$

Beispiel. Die mit dem Standardskalarprodukt assoziierte quadratische Form auf $\mathbb{R}^n$ ist

$$Q(x) = \langle x|x\rangle = \|x\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Sei A die Matrix einer symmetrischen Bilinearform auf $\mathbb{R}^n$ (bzgl. der Standardbasis). Man erinnere sich, daß A eine symmetrische Matrix ist. Wegen

$$F(x,y) = x^T Ay = \sum_{i,j=1}^{n} \alpha_{ij} x_i y_j$$

ist die assoziierte quadratische Form Q gegeben durch

$$Q(x) = F(x,x) = x^T Ax = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j = \sum_{i=1}^{n} \alpha_{ii} x_i^2 + \sum_{i<j} 2\alpha_{ij} x_i x_j.$$

Wir haben hier die Audrücke $\alpha_{ij}x_ix_j + \alpha_{ji}x_jx_i$ zu $2\alpha_{ij}x_ix_j$ für alle $i < j$ zusammengefasst. Das können wir tun, da $\alpha_{ij} = \alpha_{ji}$ aufgrund der Symmetrie von A.

Umgekehrt sei $Q : \mathbb{R}^n \to \mathbb{R}$ gegeben durch einen Ausdruck der Gestalt

$$Q(x) = \sum_{i=1}^{n} a_i x_i^2 + \sum_{i<j} b_{ij} x_i x_j\ ,$$

Dann ist Q eine quadratische Form. Die assoziierte symmetrische Bilinearform hat die Matrix

$$A = \begin{pmatrix} a_1 & \frac{1}{2}b_{12} & \dots & \frac{1}{2}b_{1n} \\ \frac{1}{2}b_{12} & a_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{1}{2}b_{1n} & \dots & \dots & a_n \end{pmatrix}$$

oder in anderen Worten $Q(x) = x^T Ax$.

Beispiel. Für $Q(x) = 2x_1^2 + 4x_1x_2 - x_2^2$ ist die zugehörige Matrix $A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$.

8.3. Hauptachsentransformation

Sei V ein Euklidischer Vektorraum der Dimension n und sei $\mathcal{B} : v_1, \ldots, v_n$ eine Orthonormalbasis von V. (Man denke an $V = \mathbb{R}^n$ mit der Standardbasis $\mathcal{B} : e_1, \ldots, e_n$.)

8.3.1 Satz. *Sei F eine symmetrische Bilinearform auf V und Q die assoziierte quadratische Form $Q(x) = F(x, x)$. Dann gibt es eine Orthonormalbasis*

$$\mathcal{B}' : v'_1, \ldots, v'_n$$

von V, so daß die Matrix A' von F bzgl. der Basis $\mathcal{B}'$ eine Diagonalmatrix ist:

$$A' = \begin{pmatrix} \lambda_1 & & & \text{\Large 0} \\ & \lambda_2 & & \\ & & \ddots & \\ \text{\Large 0} & & & \lambda_n \end{pmatrix},$$

d.h. für $x = x'_1 v'_1 + \cdots + x'_n v'_n$, $y = y'_1 v'_1 + \cdots + y'_n v'_n$ gilt

$$F(x, y) = \lambda_1 x'_1 y'_1 + \cdots + \lambda_n x'_n y'_n$$
$$Q(x) = \lambda_1 {x'_1}^2 + \cdots + \lambda_n {x'_n}^2.$$

Die eindimensionalen linearen Teilräume, die von den neuen Basisvektoren $v'_1, \ldots, v'_n$ aufgespannt werden, heißen *Hauptachsen* von F bzw. Q.

Beweis. Sei A die Matrix von F bzgl. der Basis $\mathcal{B}$. Da F symmetrisch ist, ist die Matrix A eine reelle symmetrische Matrix (vgl. 8.1.5). Aufgrund des Hauptsatzes für reelle symmetrische Matrizen (siehe 7.2.4) gibt es eine Orthonormalbasis $\mathcal{B}' : v'_1, \ldots, v'_n$ aus Eigenvektoren von A. Die Transformationsmatrix Q ist orthogonal. Weiterhin ist $A' := Q^{-1}AQ = Q^T AQ$ eine Diagonalmatrix (die Diagonaleinträge sind die Eigenwerte λ_i von A). Nach 8.1.3 ist $A' := Q^T AQ$ die Matrix von F bzgl. der Basis $\mathcal{B}'$. □

8.3.2 Quadratische Hyperflächen. Für eine quadratische Form Q auf V wird durch die Gleichung

(1) $$Q(x) = 1$$

eine *quadratische Hyperfläche* (oder einfach *Quadrik*) gegeben. Ist $A = (\alpha_{ij})$ die Matrix von Q bzgl. einer Basis $\mathcal{B}$, so kann die Gleichung (1) in der Form

(2) $$\sum_{i=1}^{n} \alpha_{ii} x_i^2 + 2 \sum_{i<j} \alpha_{ij} x_i x_j = 1$$

geschrieben werden, wobei die x_i die Koordinaten von $x \in V$ bzgl. $\mathcal{B}$ sind. Ersetzt man die 1 auf der rechten Seite durch irgendein $a > 0$, so ändert sich nichts Wesentliches.

Nach 8.3.1 gibt es eine Orthonormalbasis $\mathcal{B}' : v'_1, \ldots, v'_n$, so dass die Gleichung (2) auf die Form

(3) $$\lambda_1 {x'_1}^2 + \cdots + \lambda_n {x'_n}^2 = 1$$

gebracht werden kann, wobei die x'_i die Koordinaten von $x \in V$ bzgl. der neuen Basis $\mathcal{B}'$ sind.

Die 1-dimensionalen linearen Teilräume, die von den Vektoren $v'_1, \ldots, v'_n$ aufgespannt werden, heißen *Hauptachsen* der Quadrik.

Die Transformation der Gleichung von der Form (2) in die Form (3) wird *Hauptachsentransformation* genannt.

8.3.3 Hauptachsentransformation. Wir nehmen an, eine Quadrik im $\mathbb{R}^n$ sei durch eine Gleichung der Form

$$(Q) \qquad \sum_{i=1}^{n} a_i x_i^2 + \sum_{1 \le i < j \le n} b_{ij} x_i x_j = 1$$

gegeben, wobei die x_i die Koordinaten von $x \in \mathbb{R}^n$ bzgl. der Standardbasis sind. Wir wollen dieselbe Quadrik bzgl. einer neuen Basis beschreiben, die aus den Hauptachsen der Quadrik besteht.

(1) Bilde die Matrix

$$A = \begin{pmatrix} a_1 & \frac{1}{2}b_{12} & \ldots & \frac{1}{2}b_{1n} \\ \frac{1}{2}b_{12} & a_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}b_{1n} & \ldots & \ldots & a_n \end{pmatrix}$$

Diese Matrix ist symmetrisch und reell.

(2) Gleichung (Q) kann nun in folgender Form geschrieben werden:

$$(Q_1) \qquad x^T A x = 1.$$

(3) Berechne die Eigenwerte $\lambda_1, \ldots, \lambda_n$ von A, jeden mit seiner algebraischen Vielfachheit. (Nach 7.2.2 wissen wir, dass alle Eigenwerte von A reell sind.)

(4) Bestimme eine Orthonormalbasis $\mathcal{B}' : v'_1, \ldots, v'_n$ von Eigenvektoren von A in $\mathbb{R}^n$. (Nach 7.2.4 existiert eine solche Basis.)

(5) Sei $Q = (v'_1, \ldots, v'_n)$ die Transformationsatrix, bei der die Spalten die Vektoren v'_j der Orthonormalbasis $\mathcal{B}'$ sind. Dann ist Q eine orthogonale Matrix. Mit Q wird die Matrix A in die Diagonalmatrix

$$A' := Q^T A Q = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

transformiert und zwar bzgl. der Basis $\mathcal{B}'$.

Seien $x'_1, \ldots, x'_n$ die Koordinaten von $x \in \mathbb{R}^n$ bzgl. der neuen Basis $\mathcal{B}'$. Bzgl. der neuen Basis $\mathcal{B}'$ wird die Quadrik durch die folgende Gleichung gegeben:

$$(Q') \qquad \lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2 + \cdots + \lambda_n {x'_n}^2 = 1.$$

Wir geben nun eine Klassifizierung der quadratischen Formen in den Dimensionen 2 und 3. Dabei hängt der Typ einer Quadrik nur von den Vorzeichen der Eigenwerte λ_i ab. Insbesondere ist der Typ einer Quadrik durch die Eigenwerte der Matrix A vollständig bestimmt.

8.3.4 Klassifizierung der quadratischen Formen. Wir haben gesehen, dass bzgl. einer geeignet gewählten Orthonormalbasis jede quadratische Form im $\mathbb{R}^n$ wie folgt geschrieben werden kann:

$$Q(x) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2.$$

Für eine Klassifikation genügt es, quadratische Formen dieses Typs zu betrachten. Sind einer oder mehrere der Koeffizienten $\lambda_i = 0$, dann heißt die quadratische Form *ausgeartet*. Ist $\lambda_i \neq 0$ für alle i, dann heißt sie *nichtausgeartet*. Zur Klassifizierung betrachten wir die Quadriken mit der Gleichung

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2 = 1\ .$$

Die Ellipse ist gebrochen

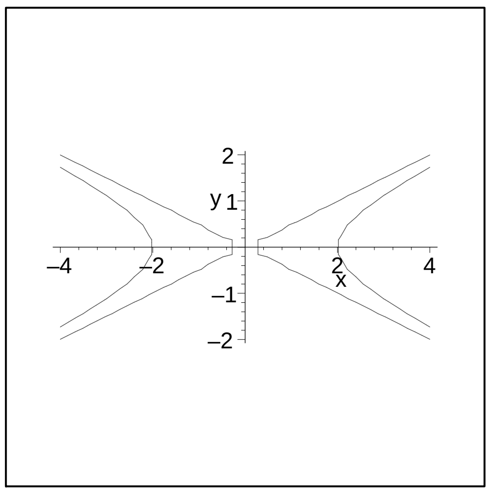


ABBILDUNG 8.1. Ellipse $\frac{x^2}{4} + y^2 = 1$ (links), Hyperbel mit Asymptoten $\frac{x^2}{4} - y^2 = 1$ (rechts)

8.3.5 Fall $\boxed{n = 2}$: Quadratische Kurven. Nach Hauptachsentransformation erhält die quadratische Gleichung die Form:

$$(Q) \qquad \lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$$

Wir müssen folgende Fälle betrachten:

(a) $\underline{\lambda_1 > 0 \text{ und } \lambda_2 > 0:}$

Sei $a = \frac{1}{\sqrt{\lambda_1}}$, $b = \frac{1}{\sqrt{\lambda_2}}$. Dann wird Gleichung (Q) zu

$$\boxed{\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1}$$

Dies ist die Gleichung einer *Ellipse* mit den *Halbachsen* a, b (siehe Zeichnung 8.1 links).

(b) $\underline{\lambda_1 > 0\,,\ \lambda_2 < 0}$:

Sei $a := \frac{1}{\sqrt{\lambda_1}}$, $b := \frac{1}{\sqrt{-\lambda_2}}$. Gleichung (Q) wird zu

$$\boxed{\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1}$$

Dies ist die Gleichung einer *Hyperbel* mit den *Asymptoten* $x_2 = \pm\frac{b}{a}x_1$ (siehe Zeichnung 8.1 rechts).

Die Asymptoten werden auf folgende Art erhalten: Die Gleichung der Hyperbel kann in der Form

$$\frac{x_1^2}{a^2} = \frac{x_2^2}{b^2} + 1$$

geschrieben werden. Sind x_1 und x_2 groß (im Betrag), kann der Wert 1 vernachlässigt werden und es ergibt sich asymptotisch die Gleichung

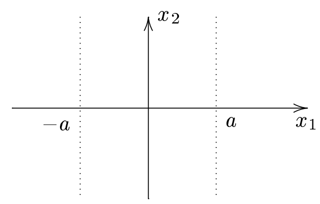
$$\frac{x_1^2}{a^2} \approx \frac{x_2^2}{b^2} \qquad \text{oder} \qquad x_2^2 \approx \frac{b^2}{a^2}x_1^2 \qquad \text{, d.h.} \qquad x_2 \approx \pm\frac{b}{a}x_1$$

(c) $\underline{\lambda_1 < 0\,,\ \lambda_2 < 0}$: Die Gleichung (Q) hat keine reelle Lösung.

(d) $\underline{\lambda_1 > 0\,,\ \lambda_2 = 0}$:

Hier vereinfacht sich (Q) zu $\lambda_1 x_1^2 = 1$, also gilt

$$\boxed{x_1 = \pm\frac{1}{\sqrt{\lambda_1}} = \pm a}$$



In diesem degenerierten Fall besteht die quadratische Kurve aus zwei parallelen Geraden (in der Zeichnung punktiert gezeichnet).

8.3.6 Beispiel. Wir betrachten zunächst die quadratische Gleichung:

(Q)
$$\boxed{x_1^2 + 6x_1x_2 + x_2^2 = 1}$$

Zugehörige Matrix der quadratischen Form: $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$

Charakteristische Gleichung: $\det(A - \lambda E) = \begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 9 = 0$

Eigenwerte: $\lambda_1 = 4$, $\lambda_2 = -2$

Eigenvektoren (normiert) zu den Eigenwerten:

$$v_1' = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad v_2' = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Transformationsmatrix: $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{pmatrix}$

(dies ist eine Drehung um 45°.) Für die Koordinaten bzgl. der neuen Basis v_1', v_2' ergibt sich die Gleichung der quadratischen Kurve zu:

(Q')
$$\boxed{4{x_1'}^2 - 2{x_2'}^2 = 1}$$

oder gleichwertig dazu $\frac{{x'_1}^2}{(\frac{1}{2})^2} - \frac{{x'_2}^2}{(\frac{1}{\sqrt{2}})^2} = 1$. Das bedeutet, es liegt eine Hyperbel mit den Asymptoten $x'_2 = \pm\sqrt{2}x'_1$ vor (siehe Zeichnung 8.2 links).

Die Koordinatentransformation zwischen den Koordinatensystemen wird beschrieben durch

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Q \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \qquad \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = Q^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Das bedeutet

$$x_1 = \frac{1}{\sqrt{2}}(x'_1 - x'_2) \qquad\qquad x'_1 = \frac{1}{\sqrt{2}}(x_1 + x_2)$$
$$x_2 = \frac{1}{\sqrt{2}}(x'_1 + x'_2) \qquad\qquad x'_2 = \frac{1}{\sqrt{2}}(x_1 - x_2)$$

8.3.7 Beispiel. Wir betrachten nun die quadratische Gleichung

(Q)
$$\boxed{7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2 = 16}$$

Zugehörige Matrix der quadratischen Form: $A = \begin{pmatrix} 7 & 3\sqrt{3} \\ 3\sqrt{3} & 13 \end{pmatrix}$

Charakteristisches Polynom:

$$\det(A - \lambda E) = \begin{vmatrix} 7-\lambda & 3\sqrt{3} \\ 3\sqrt{3} & 13-\lambda \end{vmatrix} = (7-\lambda)(13-\lambda) - 27 = \lambda^2 - 20\lambda + 64$$

Eigenwerte: $\lambda_{1/2} = \frac{20\pm\sqrt{400-256}}{2} = \frac{20\pm\sqrt{144}}{2} = \frac{20\pm 12}{2}$, also gilt $\lambda_1 = 4, \lambda_2 = 16$.

Eigenvektoren (normiert) zu diesen Eigenwerten:

$$v'_1 = \frac{1}{2}\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} \qquad v'_2 = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}$$

Transformationsmatrix: $Q = \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} = \begin{pmatrix} \cos 30° & -\sin 30° \\ \sin 30° & \cos 30° \end{pmatrix}$

(dies ist eine Drehung um 30°). Bzgl. der neuen Koordinaten hat die quadratische Kurve die Gleichung $4{x'_1}^2 + 16{x'_2}^2 = 16$ oder

(Q')
$$\boxed{\frac{{x'_1}^2}{4} + {x'_2}^2 = 1}$$

Dies ist eine Ellipse, die um 30° um die Halbachse $a = 2$, $b = 1$ gedreht wurde (siehe Zeichnung 8.2 rechts).

Wir schreiben wieder die Hauptachsentransformation explizit an:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Q \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \qquad \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = Q^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

oder

$$x_1 = \frac{1}{2}(\sqrt{3}\,x'_1 - x'_2) \qquad\qquad x'_1 = \frac{1}{2}(\sqrt{3}\,x_1 + x_2)$$

$$x_2 = \frac{1}{2}(x_1' + \sqrt{3}\, x_2') \qquad\qquad x_2' = \frac{1}{2}(-x_1 + \sqrt{3}\, x_2)$$

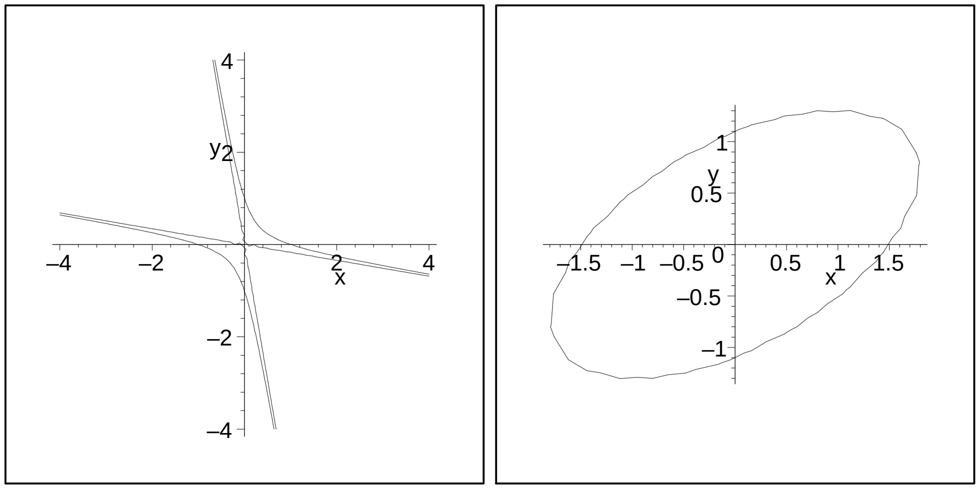


ABBILDUNG 8.2. Hyperbel $x^2 + 6xy + y^2 = 1$ (links) und Ellipse $7x^2 - 6\sqrt{3}xy + 13y^2 = 16$ (rechts)

8.3.8 Fall $\boxed{n = 3}$: Quadratische Flächen. (nichtdegenerierte Fälle)

Wir betrachten die quadratische Gleichung

$$(Q) \qquad \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^3 = 1$$

(a) $\underline{\lambda_1 > 0,\ \lambda_2 > 0,\ \lambda_3 > 0}$:

Sei $a := \frac{1}{\sqrt{\lambda_1}}$, $b := \frac{1}{\sqrt{\lambda_2}}$, $c := \frac{1}{\sqrt{\lambda_3}}$. Dann wird Gleichung (Q) zu

$$\boxed{\tfrac{x_1^2}{a^2} + \tfrac{x_2^2}{b^2} + \tfrac{x_3^2}{c^2} = 1}$$

Dies ist die Gleichung eines *Ellipsoids* mit den Halbachsen a, b, c.

(b) $\underline{\lambda_1 > 0,\ \lambda_2 > 0,\ \lambda_3 < 0}$:

Nun sei $a := \frac{1}{\sqrt{\lambda_1}}$, $b := \frac{1}{\sqrt{\lambda_2}}$, $c := \frac{1}{\sqrt{-\lambda_3}}$. Dann wird (Q) zu

$$\boxed{\tfrac{x_1^2}{a^2} + \tfrac{x_2^2}{b^2} - \tfrac{x_3^2}{c^2} = 1}$$

Dies ist die Gleichung eines *einschaligen Hyperboloids.*

ABBILDUNG 8.3. Ellipsoid $\frac{x^2}{9} + \frac{y^2}{4} + z^2 = 1$ (links), einschaliges Hyperboloid $\frac{x^2}{9} - \frac{y^2}{4} + z^2 = 1$ (Mitte) und zweischaliges Hyperboloid $\frac{x^2}{9} - \frac{y^2}{4} - z^2 = 1$ (rechts)

Analog behandelt man die anderen Fälle, bei denen zwei der $\lambda_1, \lambda_2, \lambda_3$ positiv und der dritte Wert negativ ist.

Unter den Fällen, bei denen zwei der $\lambda_1, \lambda_2, \lambda_3$ negativ und der dritte Wert nicht negativ ist, wählen wir den folgenden Fall aus:

(c) $\underline{\lambda_1 < 0,\ \lambda_2 < 0,\ \lambda_3 > 0}$:
Sei $a := \frac{1}{\sqrt{-\lambda_1}}$, $b := \frac{1}{\sqrt{-\lambda_2}}$, $c := \frac{1}{\sqrt{\lambda_3}}$. (Q) wird dann zu

$$\boxed{-\tfrac{x_1^2}{a^2} - \tfrac{x_2^2}{b^2} + \tfrac{x_3^2}{c^2} = 1}$$

Dies ist ein *zweischaliges Hyperboloid.*

(d) $\underline{\lambda_1 < 0,\ \lambda_2 < 0,\ \lambda_3 < 0}$: Die Gleichung (Q) hat keine reelle Lösung.

Übung. Untersuchen Sie die degenerierten Fälle, bei denen ein oder zwei der Koeffizienten λ_i gleich 0 sind.

Bemerkung. Die beiden Hyperboloide in (b) und (c) haben einen gemeinsamen *asymptotischen Kegel*: Für große x_1, x_2, x_3 kann man die rechte Seite 1 vernachlässigen und in beiden Fällen erhält man asymptotisch die Gleichung

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = \frac{x_3^2}{c^2}$$

Dies ist die Gleichung einer konischen Fläche (siehe Zeichnung 8.4 links). Die Zeichnung rechts zeigt, auf welche Art die beiden Hyperboloide asymptotisch an diesem Kegel anliegen.

Abbildung 8.4. Asymptotischer Kegel (links) eines zweischaligen Hyperboloids (rechts)

Bemerkung. Im Fall $a = b$ ergibt sich abhängig von den Vorzeichen

(a) ein Ellipsoid;

(b) ein einschaliges Hyperboloid oder

(c) ein zweischaliges Hyperboloid.

welches durch Rotation einer Ellipse oder einer Hyperbel um die x_3-Achse entsteht; die Schnittmengen mit horizontalen Ebenen sind Kreise.

8.3.9 Quadriken mit linearen Termen. Wir betrachten nun allgemeiner eine Gleichung der Form

$$\sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\le i<j\le n} a_{ij}x_ix_j + \sum_{i=1}^{n} b_ix_i = c \tag{4}$$

die sich in Matrixschreibweise liest als

$$x^TAx + b^Tx = c \tag{5}$$

mit einer symmetrischen Matrix A und einem Vektor b. Mit $x = Qx'$ erhalten wir

$$x'^TDx' + b'^Tx' = c \tag{6}$$

mit $Q^TAQ = D$ (Diagonalmatrix) und $b' = Q^Tb$.

Nun erinneren wir uns an die Formel $ax^2 + bx = a(x + \frac{b}{2a})^2 - \frac{b^2}{4a}$, die als *quadratische Ergänzung* bekannt ist. Wenden wir diese auf $\lambda_i {x'_i}^2 + b'_i x'_i$ an, so erhalten wir

(7) $$\sum_{i=1}^{n} \lambda_i (x'_i + \frac{b'_i}{2\lambda_i})^2 = c + \sum_{i=1}^{n} \frac{b'^2}{4\lambda_i}$$

Es gilt also $x''^T D x'' = c'$ mit $x'' = x' + v$ für $v_i = \frac{b'_i}{2\lambda_i}$ und $c' = c + \sum_{i=1}^{n} \frac{b'^2}{4\lambda_i}$.

Die Fläche ist also eine der vorigen Flächen, allerdings um v *verschoben*.

Achtung: Dies geht *nur*, wenn die $\lambda_i \neq 0$ sind. Im Allgemeinen bleiben die linearen Terme, für die es keine zugehörigen quadratischen Terme gibt, erhalten.

Ein Beispiel ist die Parabel mit der Gleichung $x_2^2 - x_1 = 0$.

8.3.10 Beispiel. Wir betrachten die quadratische Gleichungen

(Q) $$\boxed{x_1 x_2 + x_1 x_3 - x_2 x_3 = 1}$$

Zugehörige Matrix der quadratischen Form: $A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$

Charakteristische Gleichung:

$$\det(A - \lambda E) = \begin{vmatrix} -\lambda & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\lambda & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\lambda \end{vmatrix} = -\lambda^3 - \frac{1}{8} - \frac{1}{8} + \frac{\lambda}{4} + \frac{\lambda}{4}$$
$$= -\lambda^3 + \frac{3}{4}\lambda - \frac{1}{4} = (\lambda + 1)(-\lambda^2 + \lambda - \frac{1}{4}) = -(\lambda + 1)(\lambda - \frac{1}{2})^2$$

(dabei haben wir die Wurzel -1 geraten, um $\lambda + 1$ ausklammern zu können).

Eigenwerte: $\lambda_1 = -1$, $\lambda_2 = \lambda_3 = \frac{1}{2}$

Eigenvektoren zu den Eigenwerten:

- $\lambda_1 = -1$: Gauß-Jordan-Elimination ergibt

$$\begin{matrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{matrix} \quad \rightsquigarrow \quad \begin{matrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{3}{4} & -\frac{3}{4} \\ 0 & -\frac{3}{4} & \frac{3}{4} \end{matrix} \quad \rightsquigarrow \quad \begin{matrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{3}{4} & -\frac{3}{4} \\ 0 & 0 & 0 \end{matrix}$$

Wir wählen $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ oder normiert $v_1' = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

- $\lambda_2 = \lambda_3 = \frac{1}{2}$: Wir verwenden wieder Gauß-Jordan

$$\begin{matrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{matrix} \quad \rightsquigarrow \quad \begin{matrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

Damit können wir

$$b_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ und } b_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

als Fundamentalsystem von Lösungen wählen. Wir wenden Gram-Schmidt an, um ein Orthonormalsystem von Lösungen zu erhalten:

$$v_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; v_3' = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Damit ergibt sich die neue Orthogonalbasis zu v_1', v_2', v_3'. Wir erhalten die Transformationsmatrix

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

Die Gleichung unserer quadratischen Fläche bzgl. der neuen Basis v_1', v_2', v_3' ist

$$(Q') \qquad \boxed{-{x_1'}^2 + \tfrac{1}{2}{x_2'}^2 + \tfrac{1}{2}{x_3'}^2 = 1}$$

ABBILDUNG 8.5. Gedrehtes einschaliges (links) und zweischaliges (rechts) Hyperboloid

Hier haben wir ein einschaliges Rotationshyperboloid (siehe Zeichnung 8.5 links).

In der x_1', x_3'-Ebene betrachten wir die Hyperbel mit der Gleichung $-{x_1'}^2 + \frac{1}{2}{x_2'}^2 = 1$ und rotieren diese um die x_1'-Achse.

Bemerkung. Wir brauchen nur die Eigenwerte $\lambda_1, \lambda_2, \lambda_3$ von A, um die obige Gleichung zu erhalten, die uns etwas über den Typ der quadratischen Fläche sagt. Die Eigenvektoren v_1', v_2', v_3' werden nur gebraucht, wenn wir wissen wollen, wie die Fläche im $\mathbb{R}^3$ plaziert ist.

8.3.11. Wir betrachten nun

$$(Q) \qquad \boxed{-7x_1^2 - x_2^2 - x_3^2 + 8x_1x_2 + 8x_1x_3 + 16x_2x_3 = 9}$$

Zugehörige Matrix der quadratischen Form: $A = \begin{pmatrix} -7 & 4 & 4 \\ 4 & -1 & 8 \\ 4 & 8 & -1 \end{pmatrix}$

Charakteristisches Polynom:

$$\begin{aligned} \det(A - \lambda E) = \begin{vmatrix} -7-\lambda & 4 & 4 \\ 4 & 1-\lambda & 8 \\ 4 & 8 & -1-\lambda \end{vmatrix} &= (-7-\lambda)(-1-\lambda)(-1-\lambda) + \\ &+ 128 + 128 + 16(1+\lambda) + 16(1+\lambda) + 64(7+\lambda) \\ &= -(7+\lambda)(1+\lambda)^2 + 96\lambda + 736 = \cdots = -\lambda^3 - 9\lambda^2 + 81\lambda + 729 \\ &= -(\lambda - 9)(\lambda + 9)^2 \end{aligned}$$

Eigenwerte: $\lambda_1 = 9$, $\lambda_2 = \lambda_3 = -9$.

Nach Koordinatentransformation lautet die Gleichung $9{x'_1}^2 - 9{x'_2}^2 - 9{x'_3}^2 = 9$ oder

$$(Q') \qquad \boxed{{x'_1}^2 - {x'_2}^2 - {x'_3}^2 = 1}$$

Dies ist ein zweischaliges Hyperboloid. Die Zeichnung 8.5 (rechts) zeigt dessen Gestalt.

Wollen wir das neue Koordinatensystem und die Transformationsmatrix berechnen, müssen wir eine Orthonormalbasis von Eigenvektoren von A bestimmen. (Übung.)

8.3.12 Beispiel. Wir betrachten abschließend

$$(Q) \qquad \boxed{7x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 4x_2x_3 = 3}$$

Zugehörige Matrix der Bilinearform: $A = \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}$

Charakteristisches Polynom:

$$\begin{aligned} \det(A - \lambda E) = \begin{vmatrix} 7-\lambda & -2 & 0 \\ -2 & 6-\lambda & -2 \\ 0 & -2 & 5-\lambda \end{vmatrix} &= (7-\lambda)(6-\lambda)(5-\lambda) - 4(7-\lambda) - 4(5-\lambda) \\ &= (7-\lambda)(6-\lambda)(5-\lambda) - 8(6-\lambda) = (6-\lambda)((7-\lambda)(5-\lambda) - 8) \\ &= (6-\lambda)(\lambda^2 - 12\lambda + 27) = (6-\lambda)(\lambda-3)(\lambda-9) \end{aligned}$$

Eigenwerte: $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$.

Die Gleichung nach Hauptachsentransformation ist $3{x'_1}^2 + 6{x'_2}^2 + 9{x'_3}^2 = 3$ oder

$$(Q') \qquad \boxed{{x'_1}^2 + 2{x'_2}^2 + 3{x'_3}^2 = 1}$$

Dies ist die Gleichung eines Ellipsoids mit den Hauptachsen

$$a = 1\,, \qquad b = \frac{1}{\sqrt{2}}\,, \qquad c = \frac{1}{\sqrt{3}}$$

Die Hauptachsen erhält man durch die Berechnung (normierter) Eigenvektoren zu den Eigenwerten λ_1, λ_2, λ_3:

$$v_1' = \frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix}\,, \qquad v_2' = \frac{1}{3}\begin{pmatrix}2\\1\\-2\end{pmatrix}\,, \qquad v_3' = \frac{1}{3}\begin{pmatrix}2\\-2\\1\end{pmatrix}$$

Die Transformationsmatrix ist $Q = \frac{1}{3}\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & -2\\ 2 & -2 & 1\end{pmatrix}$.

8.4. Definitheit von Formen und von Matrizen

Sei V ein reeller Vektorraum und $F(x,y)$ eine symmetrische Bilinearform auf V. Betrachte die zugehörige quadratische Form $Q(x) := F(x,x)$. Dann ist Q eine Funktion von V nach $\mathbb{R}$. Beachte $Q(0) = 0$.

8.4.1 Definition. Die quadratische Form Q und entsprechend die Bilinearform F heißt

- *positiv definit*, wenn $Q(x) > 0$ für alle $x \neq 0$;
- *negativ definit*, wenn $Q(x) < 0$ für alle $x \neq 0$;
- *positiv semidefinit*, wenn $Q(x) \geq 0$ für alle x;
- *negativ semidefinit*, wenn $Q(x) \leq 0$ für alle x;
- *indefinit*, wenn Q keine der obigen Eigenschaften hat, d.h. wenn Q positive und negative Werte hat.

Nun sei A eine symmetrische $n \times n$-Matrix mit reellen Einträgen. Diese definiert eine symmetrische Bilinearform auf $\mathbb{R}^n$ durch

$$F(x,y) := x^T A y \quad \text{für alle } x, y \in \mathbb{R}^n$$

und eine quadratische Form

$$Q(x) := x^T A x \quad \text{für alle } x \in \mathbb{R}^n\,.$$

Wir nennen die symmetrische Matrix A *positiv* [*negativ*] *(semi-)definit* oder *indefinit*, wenn die zugehörige quadratische Form diese Eigenschaften hat.

Bemerkung. In 3.4.1 haben wir einen euklidischen Vektorraum definiert als einen reellen Vektorraum V mit einer bilinearen, symmetrischen, positiv definiten Form. Ist also A eine symmetrische positiv defnite Matrix, dann definiert

$$F(x,y) = x^T A y = \overbrace{\langle x|Ay\rangle}^{\text{(Standardskalarprodukt)}} \qquad \text{für alle } x, y \in \mathbb{R}^n$$

ein Skalarprodukt auf $\mathbb{R}^n$, das vom Standardskalarprodukt verschieden ist (falls $A \neq E$). Umgekehrt ist jedes Skalarprodukt auf $\mathbb{R}^n$ auf diese Weise durch eine positiv definite symmetrische Matrix definiert.

Übung. Sei B eine $n \times n$-Matrix mit reellen Einträgen. Zeige, daß die Matrix $A := B^T B$ symmetrisch und positiv semidefinit ist. Ist B invertierbar, dann ist A sogar positiv definit.

ABBILDUNG 8.6. Elliptisches (oben links & rechts) und hyperbolisches (unten links) Paraboloid; parabolischer Zylinder (unten rechts)

8.4.2 Beispiele. Sei $V = \mathbb{R}^2$ und $Q(x) = \lambda_1 x_1^2 + \lambda_2 x_2^2$.

(a) $\underline{\lambda_1 > 0,\ \lambda_2 > 0}$: Q ist *positiv definit.*

Die durch die Gleichung
$$z = Q(x)$$
gegebene Fäche is ein *elliptisches Paraboloid* (siehe Zeichnung 8.6 links).

(b) $\underline{\lambda_1 < 0,\ \lambda_2 < 0}$: Q ist *negativ definit.*

Die durch die Gleichung
$$z = Q(x)$$
gegebene Fläche ist wieder ein elliptischs Paraboloid (nach unten geöffnet) (siehe Zeichnung 8.6 Mitte links).

(c) $\underline{\lambda_1 > 0,\ \lambda_2 < 0}$: Q ist *indefinit.*

In der Tat ist $Q\begin{pmatrix}1\\0\end{pmatrix} = \lambda_1 > 0$, $Q\begin{pmatrix}0\\1\end{pmatrix} = \lambda_2 < 0$. Die durch die Gleichung
$$z = Q(x)$$
gegebene Fläche ist ein *hyperbolisches Paraboloid* (siehe Zeichnung 8.6 Mitte rechts).

(d) $\underline{\lambda_1 > 0,\ \lambda_2 = 0}$: Q ist *positiv semidefinit.*

Die durch die Gleichung
$$z = Q(x) = \lambda_1 x_1^2$$
gegebene Fläche ist ein *parabolischer Zylinder*; man erhält ihn durch Auseinanderziehen der Parabel $z = \lambda_1 x_1^2$ in der x_1-z-Ebene entlang der x_2-Achse (siehe Zeichnung 8.6 rechts).

(e) $\underline{\lambda_1 < 0,\ \lambda_2 = 0}$: Q ist *negativ semidefinit.*

Der parabolische Zylinder ist nun nach unten offen.

8.4.3 Quadratische Formen auf $\mathbb{R}^2$. Wir betrachten auf $\mathbb{R}^2$ eine beliebige quadratische Form

$$(Q) \qquad \boxed{z = Q(x) = a x_1^2 + 2b x_1 x_2 + c x_2^2}$$

(bzgl. der Standardbasis). Die zugehörige Matrix ist $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

Mittels der Hauptachsentransformation können wir eine neue Orthonormalbasis v_1', v_2' finden, so dass (Q) in der Form

$$z_1 = \lambda_1 {x_1'}^2 + \lambda_2 {x_2'}^2$$

geschrieben werden kann (bzgl. der neuen Basis), wobei λ_1 und λ_2 die Eigenwerte der Matrix A sind.

In 8.4.2 haben wir die quadratischen Formen dieses Typs klassifiziert. Man erkennt, dass die Definitheitseigenschaften von Q und A nur vom Vorzeichen der Eigenwerte von A abhängen:

Q ist positiv definit	$\Longleftrightarrow$	die Eigenwerte λ_1, λ_2 von A sind beide > 0;
Q ist negativ definit	$\Longleftrightarrow$	die Eigenwerte λ_1, λ_2 von A sind beide < 0;
Q ist indefinit	$\Longleftrightarrow$	einer der Eigenwerte ist > 0, der andere < 0.
Q ist definit	$\Longleftrightarrow$	$\lambda_1 \cdot \lambda_2 > 0$
und Q ist indefinit	$\Longleftrightarrow$	$\lambda_1 \cdot \lambda_2 < 0$
Q ist positiv semidefinit	$\Longleftrightarrow$	die Eigenwerte von A sind beide ≥ 0;
Q ist negativ semidefinit	$\Longleftrightarrow$	die Eigenwerte von A sind beide ≤ 0.

Da wir A mit Hilfe einer Übergangsmatrix S diagonalisieren können, wissen wir, dass

$$S^T A S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

gilt und damit $\det A = \det(S^T A S) = \lambda_1 \cdot \lambda_2$ ist.

Wir folgern:

- Ist $\underline{\det A < 0}$, dann gilt $\lambda_1 \cdot \lambda_2 < 0$, d.h. Q ist indefinit.
- Ist $\underline{\det A > 0}$, dann gilt $\lambda_1 \cdot \lambda_2 > 0$, d.h. Q ist definit.

  Um zu erkennen, ob positive oder negative Definitheit vorliegt, betrachten wir den Spezialfall $x_1 = 1$, $x_2 = 0$. Es folgt

  $$z = a x_1^2 + 2b x_1 x_2 + c x_2^2 = a$$

  und damit $z \geq 0$ genau dann, wenn $a \geq 0$ ist; ebenso gilt $z \leq 0$ genau dann, wenn $a \leq 0$ ist.

  Also ist Q im ersten Fall positiv definit und negativ definit im zweiten Fall. Beachte, dass $a = 0$ in diesem Fall nicht auftreten kann, da $a = 0$ $\det A = -b^2 \leq 0$ impliziert.

Insgesamt haben wir bewiesen:

8.4.4 Satz. Die quadratische Form $Q(x) = a x_1^2 + 2b x_1 x_2 + c x_2^2$ ist

- positiv definit $\Longleftrightarrow$ $\det A = ac - b^2 > 0$ und $a > 0$;
- negativ definit $\Longleftrightarrow$ $\det A = ac - b^2 > 0$ und $a < 0$;

- indefinit $\iff$ $\det A = ac - b^2 < 0$.

Im Folgenden verallgemeinern wir diese Betrachtungen von $n = 2$ auf beliebiges n; die Bedeutung des Kriteriums besteht darin, dass wir damit die Definitheitseigenschaften einer Matrix bestimmen können, ohne die Eigenwerte berechnen zu müssen:

8.4.5 Kriterium für positive Definitheit. *Sei $A = (\alpha_{ij})$ eine symmetrische $n \times n$-Matrix mit reellen Einträgen und $Q(x) := x^T Ax$ die zugehörige quadratische Form.*

Dann sind die folgenden Eigenschaften äquivalent

(a) *A (und Q) sind positiv definit;*

(b) *$\lambda_i > 0$ für jeden Eigenwert λ_i von A;*

(c) *$\det(A_r) > 0$ für $r = 1, 2, \dots, n$, wobei die Matrizen*

$$A_r := \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1r} \\ \vdots & & \vdots \\ \alpha_{r1} & \dots & \alpha_{rr} \end{pmatrix}$$

die Hauptminoren *von A sind.*

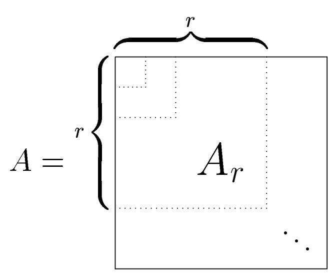
Analog sind die folgenden Eigenschaften äquivalent:

(a') *A (und Q) sind negativ definit;*

(b') *$\lambda_i < 0$ für alle Eigenwerte λ_i von A;*

(c') *$(-1)^r \det(A_r) > 0$ für $r = 1, 2, \dots, n$.*

Wir können die Hauptminoren schematisch wie folgt darstellen:



Beweis. Die Äquivalenz von (a'), (b') und (c') folgt aus der Äquivalenz von (a), (b) und (c) durch Ersetzung von A mit $-A$.

Wir zeigen nun (a) $\iff$ (b), (a) $\implies$ (c) und (c) $\implies$ (b).

normalfont (a) $\iff$ (b): Die Matrix A ist symmetrisch und damit diagonalisierbar. Mittels der Hauptachsentransformation können wir eine Orthonormalbasis von $\mathbb{R}^n$ finden, so dass bzgl. der Koordinaten in dieser neuen Basis gilt:

$$Q(x) = \lambda_1 {x'_1}^2 + \dots + \lambda_n {x'_n}^2.$$

Dabei sind die $\lambda_1, \lambda_2, \dots \lambda_n$ die Eigenwerte von A. In dieser Form erkennt man direkt, dass $Q(x)$ genau dann positiv [negativ] definit ist, wenn alle λ_i positiv [negativ] sind.

(a) $\Longrightarrow$ (c): Ist A symmetrisch und positiv definit, so sind auch alle Matrizen A_r symmetrisch und positiv definit. (Die Werte von $Q_r(y) = y^T A_r^T y$ erhält man aus $Q(x)$ durch Nullsetzen der zusätzlichen Koordinaten.)

Damit ist jede Matrix A_r orthogonal diagonalisierbar zu

$$A_r' = Q^T A_r Q = Q^{-1} A_r Q = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_r \end{pmatrix},$$

wobei die Werte auf der Diagonalen die Eigenwerte von μ_r sind. Mittels obiger Begründung ergibt sich, dass alle μ_i positiv sind, da A_r positiv definit ist.

Die Determinante ist invariant unter Basiswechsel:

$$\begin{aligned} \det A_r &= \det(Q A_r' Q^{-1}) = \det Q \cdot \det A_r' \cdot \det Q^{-1} = \det A_r' \\ &= \mu_1 \cdot \mu_2 \ldots \cdot \mu_r > 0. \end{aligned}$$

Damit gilt $\det A_r > 0$ für alle Hauptminoren.

(c) $\Longrightarrow$ (b): Wir beweisen mittels Induktion nach n, dass eine $n \times n$-Matrix A mit der Eigenschaft, dass die Determinanten ihrer Hauptminoren echt positiv sind, positiv definit ist. Für $n = 1$ ist die Matrix A eine einfache Zahl, die mit ihrer Determinante übereinstimmt. Wir nehmen $n > 0$ an und dass die Implikation bereits für $r < n$ bewiesen wurde. Die Hauptminoren der Matrix A_{n-1} sind unter den Hauptminoren von A, so dass nach der Induktionsannahme A_{n-1} positiv definit ist. Das heißt aber, dass A eingeschränkt auf den Teilraum

$$U = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid x_n = 0 \}$$

positiv definit ist.

Es sei nun W die (direkte) Summe der Eigenräume zu den nichtpositiven (≤ 0) Eigenwerten. Eingeschränkt auf W ist die durch A definierte Form negativ semidefinit. Dies folgt aus der Implikation b)$\Rightarrow$a), da die Einschränkung von A auf W nur nichtpositive Eigenwerte hat. Ist $x \in U \cap W$ und $x \neq 0$, so folgt $x^T A x > 0$ und $x^T A x \leq 0$, was nicht möglich ist, also ist $U \cap W = \{0\}$.

Da U, W beides Teilräume von $\mathbb{R}^n$ sind, gilt $\dim(U+W) \leq n$. Es gilt andererseits wegen $U \cap W = \{0\}$ die Beziehung $\dim(U + W) = \dim(U) + \dim(W) = n - 1 + s$, wobei s die Zahl der nichtpositiven Eigenwerte mit deren algebraischer ($=$ geometrischer) Vielfachheit ist. Also ist $s = 0$ oder $s = 1$. Die Determinante von A ist das Produkt der Eigenwerte (mit deren algebraischer Vielfachheit). Wegen $\det(A) = \det(A_n) > 0$ gilt $s = 0$, also sind alle Eigenwerte von A strikt positiv. $\square$

8.4.6 Beispiel. (für $n = 4$.) Betrachte

$$Q(x) = 3x_1^2 + 2x_2^2 + 2x_3^2 + 6x_4^2 + 4x_1x_2 - 2x_1x_3 + 2x_1x_4 - 2x_3x_4.$$

Die zugehörige Matrix ist

$$A = \begin{pmatrix} 3 & 2 & -1 & 0 \\ 2 & 2 & 0 & 1 \\ -1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 6 \end{pmatrix}$$

und die Hauptminoren sind

$$\det A_1 = 3\,, \quad \det A_2 = \begin{vmatrix} 3 & 2 \\ 2 & 2 \end{vmatrix} = 2\,, \quad \det A_3 = \begin{vmatrix} 3 & 2 & -1 \\ 2 & 2 & 0 \\ -1 & 0 & 2 \end{vmatrix} = 2\,, \quad \det A_4 = \det A = 1\,.$$

Also ist $\det A_r > 0$ für $r = 1, 2, 3, 4$. Nach obigem Kriterium ist $Q(x)$ positiv definit.

8.5. Linearformen und Dualraum, Tensoren

In diesem Abschnitt sei V ein Vektorraum über $\mathbb{R}$. Wir betrachten $\mathbb{R}$ als eindimensionalen Vektoraum mit der Zahl 1 als Basis.

8.5.1 Linearformen. Als *Linearform* auf V bezeichnen wir jede lineare Abbildung $f : V \to \mathbb{R}$.

Wie wir in 4.5.1 gesehen haben, bilden die linearen Abbildungen eines Vektorraums in einen anderen ebenfalls einen Vektorraum. Das gilt insbesondere für Linearformen. Summe und skalares Vielfaches von Linearfomen f und g sind punktweise erklärt:

$$(f+g)(x) = f(x) + g(x) \quad , \quad (\lambda f)(x) = \lambda \cdot f(x) \quad \text{für alle } x \in V \text{ und alle } \lambda \in \mathbb{R}\,.$$

8.5.2 Dualraum. Den Vektorraum $\mathrm{Hom}(V, \mathbb{R})$ aller Linearformen auf V nennen wir den *Dualraum* von V, und wir bezeichnen ihn mit V^*.

Beispiele.

(1) Wähle reelle Zahlen $\alpha_1, \ldots, \alpha_n$. Für jedes $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ definiere

$$f(x) = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

Dann ist f eine Linearform auf $\mathbb{R}^n$. (Dies erklärt die Herkunft des Namens *Linearform*.)

Bezeichnen wir mit $w^T = \begin{pmatrix} \alpha_1 & \ldots & \alpha_n \end{pmatrix}$ den Zeilenvektor der Koeffizienten, so können wir auch schreiben

$$f(x) = \begin{pmatrix} \alpha_1 & \ldots & \alpha_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = w^T x$$

Man kann w^T als die Matrix von f bezüglich der Standardbasis $e_1, \ldots, e_n$ des $\mathbb{R}^n$ betrachten; denn $f(e_j) = \alpha_j$. Da jede lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}$

durch eine $1 \times n$-Matrix gegeben werden kann, sehen wir, daß jede Linearform auf $\mathbb{R}^n$ von der oben angegebenen Gestalt ist.

Wir bezeichnen mit $(\mathbb{R}^n)^*$ den Vektorraum aller Zeilenvektoren w^T. Diesen kann man nach dem eben Gesagten als den Vektorraum aller Linearformen auf $\mathbb{R}^n$ betrachten, also als den Dualraum von $\mathbb{R}^n$. Dieser hat ebenfalls die Dimension n. Die Zeilenvektoren

$$e_1^T = (1,0,0,\dots,0), e_2^T = (0,1,0,\dots,0), \ \dots \ , e_n^T = (0,0,0,\dots,1)$$

bilden eine Basis von $(\mathbb{R}^n)^*$, genannt die zur Standardbasis des $\mathbb{R}^n$ duale Basis. Man bemerke: $e_i^T(e_j) = \delta_{ij}$.

(2) Sei $V = C\big([0,1]\big)$ der Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$.

Wir wählen einen festen Punkt $a \in [0,1]$ und definieren $\delta_a\colon V \in \mathbb{R}$ durch

$$\delta_a(f) = f(a) \text{ für alle } f \in V\,.$$

Dann ist δ_a ein lineares Funktional auf V genannt das Dirac-Maß im Punkte a.

Wählen wir eine feste Funktion $g \in V$, dann ist die durch

$$\varphi(f) = \int_0^1 f(x)g(x)\,dx \text{ für alle } f \in V$$

definierte Funktion φ ebenfalls ein lineares Funktional auf V.

Wir sehen nicht, wie man den Dualraum V^* insgesamt beschreiben könnte. Er besteht nämlich aus den sämtlichen Borelmaßen auf $[0,1]$, und diese sind so etwas wie skalare Vielfache von Wahrscheinlichkeitsverteilungen.

(3) Sei nun V ein beliebiger Euklidischer Vektorraum. Wählt man ein festes $v \in V$, dann ist die durch

$$v^*(x) = \langle v|x\rangle \text{ für alle } x \in V$$

definierte Funktion ein lineares Funktional auf V.

8.5.3 Duale Basis. Wir betrachten einen n-dimensionalen reellen Vektorraum V mit einer Basis

$$\mathcal{B} : v_1, \dots, v_n.$$

Die Koordinaten x^i eines beliebigen Vektors $v \in V$ schreiben wir in diesem Abschnitt mit oberen Indices, d.h.

$$[v]_{\mathcal{B}} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

falls

$$v = x^1 v_1 + x^2 v_2 + \cdots + x^n v_n = \sum_i x^i v_i\,.$$

Der Dualraum V^* hat dieselbe Dimension n wie V, denn allgemein gilt $\dim \mathrm{Hom}(V,W) = \dim V \cdot \dim W$. Wir definieren spezielle Linearformen v^i auf V durch

$$v^i(v_j) = \delta^i_j = \begin{cases} 1 \text{ falls } i = j \\ 0 \text{ sonst} \end{cases}$$

(Wir erinnern uns daran, daß eine lineare Abbildung eindeutig durch ihre Werte auf einer Basis bestimmt ist.) Die Matrix von v^i bezüglich der Basis $\mathcal{B}$ ist dann

$$e_i^T = (0, \dots, 1, \dots, 0)$$

Daraus folgt unmittelbar, daß die $v^1, \dots, v^n$ linear unabhängig sind und folglich eine Basis von V^* bilden. Wir nennen

$$\mathcal{B}^* : v^1, \dots, v^n$$

die zur Basis $\mathcal{B}$ *duale Basis*.

Eine beliebige Linearform $f \in V^*$ hat bezüglich der dualen Basis $\mathcal{B}^*$ eine Darstellung

$$f = y_1 v^1 + \cdots + y_n v^n = \sum_i y_i v^i\,,$$

d.h. bezüglich der dualen Basis hat f die Koordinaten

$$[f]_{\mathcal{B}^*} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

In Koordinaten bezüglich der Basis $\mathcal{B}$ wie oben gilt für ein beliebiges $v \in V$:

$$f(v) = \sum_i y_i v^i \big(\sum_j x^j v_j \big) = \sum_i \sum_j y_i x^j \underbrace{v^i(v_j)}_{\delta^i_j} = \sum_i y_i x^i = [f]^T_{\mathcal{B}^*} [v]_{\mathcal{B}}\,.$$

Merkregel für Indexstellung: In diesem Abschnitt w'ahlen wir wie in der Relativitätstheorie die folgende Konvention:

Basisvektoren in V	Index unten	v_i, $\tilde{v}_i$
Koordinaten von Vektoren in V	Index oben	
x^i, $\tilde{x}^i$		
Basisvektoren in $V*$	Index oben	v^i, $\tilde{v}^i$
Koordinaten von Vektoren in $V*$	Index unten	
y_i, $\tilde{y}_i$		

8.5.4 Des Doppeldual V^{**}. Zu dem Dualraum V^* kann man wiederum den Dualraum V^{**} bilden. Dieser besteht aus allen Linearformen $g \colon V^* \to \mathbb{R}$.

Wir können jedem Vektor $v \in V$ in natürlicher Weise eine Linearform $g_v \colon V^* \to \mathbb{R}$ zuordnen, indem wir setzen

$$g_v(f) = f(v) \text{ für alle } f \in V^*\,.$$

(Die so definierte Abbildung g_v ist in der Tat linear und folglich eine Element von V^{**}; denn für $f, f' \in V^*$ und $\lambda \in \mathbb{R}$ gilt $g_v(f+f') = (f+f')(v) = f(v)+f'(v) = g_v(f)+g_v(f')$ und $g_v(\lambda f) = (\lambda f)(v) = \lambda f(v) = \lambda g_v(f)$.)

Damit haben wir eine Abbildung $v \mapsto g_v \colon V \to V^{**}$. Man weist leicht nach, daß diese Abbildung linear ist. Außerdem ist sie injektiv; denn ist $v \neq 0$, dann gibt es ein $f \in V^*$ derart, daß $f(v)\neg 0$. Es folgt $g_v(f) = f(v) = 0$ und folglich $g_v \neq 0$. Wir können also die Elemente $v \in V$ als Elemente von V^{**} betrachten, indem wir v mit der durch v induzierten Linearform $g_v \colon V^* \to \mathbb{R}$ identifizieren.

Für unendlichdimensionale Vektorräume sind nicht alle Elemente von V^{**} von der Form g_v. Das ist jedoch für endlichdimensionale Vektorräume der Fall; denn $\dim V^{**} = \dim V^* = \dim V$.

Sei nämlich V endlichdimensional, und $\mathcal{B} : v_1, \ldots, v_n$ sei eine Basis von V. Wir können die duale Basis $\mathcal{B}^*$ von V^* bilden, und hierzu wieder die duale Basis $\mathcal{B}^{**}$ von V^{**}. Nun gilt $g_{v_j}(v^i) = v^i(v_j) = \delta^i_j$; auf Grund der Definition der dualen Basis bilden also die Linearformen $g_{v_j} \in V^{**}$ die zu $\mathcal{B}^*$ duale Basis $\mathcal{B}^{**}$ in V^{**}, d.h. $v_j^{**} = g_{v_j}$. Wenn wir wie oben die Vektoren $v \in V$ mit den Linearformen g_v identifizieren, haben wir also $v_j^{**} = v_j$.

Zusammenfassend haben wir:

Im endlichdimensionalen Fall kann man V mit dem Doppeldual V^{**} identifizieren; die Basis $\mathcal{B}$ von V entspricht dabei dem Dualen $\mathcal{B}^{**}$ der dualen Basis $\mathcal{B}^*$. Etwas unscharf aber kurz gesagt: $V = V^{**}$ und $\mathcal{B} = \mathcal{B}^{**}$.

8.5.5 Kovariante bzw. kontravariante Koordinatentransformation. Wir betrachten nun in V zwei Basen

$$\mathcal{B} : v_1, \ldots, v_n \quad , \quad \widetilde{\mathcal{B}} : \tilde{v}_1, \ldots, \tilde{v}_n .$$

Wir können die neuen Basisvektoren durch die alten darstellen und umgekehrt:

$$\tilde{v}_j = \sum_i s^i_j v_i \quad , \quad v_l = \sum_k \tilde{s}^k_l \tilde{v}_k \,.$$

Schon früher betrachteten wir die Transformationsmatrizen $S = (s^i_j)$ und $\widetilde{S} = (\tilde{s}^k_l)$. (Der obere Index ist der Zeilenindex, der untere der Spaltenindex.) Wie früher gilt natürlich $\widetilde{S} = S^{-1}$; denn $v_l = \sum_k \tilde{s}^k_l \tilde{v}_k = \sum_k \tilde{s}^k_l \sum_i s^i_k v_k = \sum_i \big(\sum_k \tilde{s}^k_l \cdot s^i_k\big) v_i$ und folglich durch Koeffizientenvergleich

$$\sum_k s^i_k \cdot \tilde{s}^k_l = \delta^i_l \,.$$

Wenden wir uns nun den beiden dualen Basen

$$\mathcal{B}^* : v^1, \ldots, v^n \quad \text{und} \quad \widetilde{\mathcal{B}}^* : \tilde{v}^1, \ldots, \tilde{v}^n$$

zu. Wir behaupten, daß zwischen ihnen die folgenden Beziehungen bestehen:

$$\tilde{v}^j = \sum_i \tilde{s}^j_i v^i \quad , \quad v^l = \sum_k s^l_k \tilde{v}^k \,.$$

Ist nämlich $T = (t_i^j)$ die Transformationsmatrix zwischen den dualen Basen, so hat man eine Darstellung $\tilde{v}^j = t_i^j v^i$. Es gilt dann $\delta_l^j == \tilde{v}^j(\tilde{v}_l) = \big(\sum_i t_i^j v^i\big)\big(\sum_k s_l^k v_k\big) = \sum_{i,k} t_i^j \cdot s_l^k \underbrace{v^i(v_k)}_{=\delta_k^i} = \sum_i t_i^j \cdot s_l^i$, und daraus folgt $TS = E$, d.h. $T = S^{-1} = \widetilde{S} = (\tilde{s}_l^j)$.

Wie transformieren sich die Koordinaten x^j bzw. $\tilde{x}^j$ eines Vektors $v \in V$ bezüglich der beiden Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ ineinander? $v = \sum_j x^j v_j = \sum_i \tilde{x}^i \tilde{v}_i = \sum_i \tilde{x}^i \cdot \sum_j s_i^j v_j = \sum_j \big(\sum_i s_i^j \tilde{x}^i\big) v_j$. Koeffizientenvergleich ergibt

$$x^j = \sum_i s_i^j \tilde{x}^i \quad , \quad \tilde{x}^l = \sum_k \tilde{s}_k^l x^k$$

Die Beziehung zwischen den alten und den neuen Koordinaten wird durch die folgenden Transformationsformeln gegeben:

$$[v]_{\mathcal{B}} = S[v]_{\widetilde{\mathcal{B}}} \quad , \quad [v]_{\widetilde{\mathcal{B}}} = S^{-1}[v]_{\mathcal{B}} \, ,$$

ausführlich geschrieben:

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = S \begin{pmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^n \end{pmatrix} \quad , \quad \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Benutzt man die Transformationsformeln für die dualen Basen von vorhin, so erhält man in ganz analoger Weise für die Koordinaten y_i bzw. $\tilde{y}_i$ einer Linearform $f \in V^*$ bezüglich der dualen Basen $\mathcal{B}^*$ und $\widetilde{\mathcal{B}}^*$ die folgenden Tranformationsformeln

$$y_j = \sum_i \tilde{s}_j^i \tilde{y}_i \quad , \quad \tilde{y}_l = \sum_k s_l^k y_k \, .$$

In Matrixschreibweise

$$[f]_{\mathcal{B}^*} = (S^{-1})^T [f]_{\widetilde{\mathcal{B}}^*} \quad , \quad [f]_{\widetilde{\mathcal{B}}^*} = S^T [f]_{\mathcal{B}^*} \, ,$$

oder ausführlich geschrieben

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (S^{-1})^T \begin{pmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_n \end{pmatrix} \quad , \quad \begin{pmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_n \end{pmatrix} = S^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Man beobachtet ein unterschiedliches Transformationsverhalten der Koordinaten von Vektoren $v \in V$ und von Linearformen $f \in V^*$. Dies drückt man dadurch aus, daß man sagt:

Die Koordinaten von Vektoren in V transformieren sich kontravariant*, während die Koordinaten von in V^* sich* kovariant *transformieren.*

Die beiden Transformationsformeln stimmen genau dann überein, wenn $S^{-1} = S^T$, d.h. wenn die Transformationsmatrix S orthogonal ist. Dies ist dann der Fall, wenn beide Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ Orthonormalbasen sind.

8.5.6 Tensoren über V. Wir betrachten einen endlichdimensionalen Vektorraum V sowie seinen Dualraum V^*.

In V betrachten wir zwei Basen

$$\mathcal{B} : v_1, \ldots, v_n \quad , \quad \widetilde{\mathcal{B}} : \tilde{v}_1, \ldots, v_n \, ,$$

sowie die hierzu dualen Basen in V^*

$$\mathcal{B}^* : v^1, \ldots, v^n \quad , \quad \widetilde{\mathcal{B}}^* : \tilde{v}^1, \ldots, \tilde{v}^n \, .$$

Wie im vorhergehenden Abschnitt haben wir die folgenden Transformationsformeln zwischen diesen Basen, wobei die Matrix $\tilde{S} = (\tilde{s}^i_j)$ die Inverse der Matrix $S = (s^i_j)$ ist:

$$\tilde{v}_j = \sum_i s^i_j v_i \quad , \quad v_l = \sum_k \tilde{s}^k_l \tilde{v}_k \, .$$

$$\tilde{v}^j = \sum_i \tilde{s}^j_i v^i \quad , \quad v^l = \sum_k s^l_k \tilde{v}^k \, .$$

- **Tensoren 0-ter Stufe** Das sind ganz einfach Skalare $\alpha \in \mathbb{R}$.
- **Tensoren 1-ter Stufe**

(1) Vektoren $v \in V$.

Diese können wir nach 8.5.4 als Linearformen $g_v \colon f \mapsto f(v) \colon V^* \to \mathbb{R}$ betrachten. Die Koordinaten von $v \in V$ bezüglich der Basen $\mathcal{B}$ bzw. $\widetilde{\mathcal{B}}$ werden gegeben durch

$$\alpha^i = v^j(v) \quad , \quad \tilde{\alpha}^j = \tilde{v}^j(v) \, ,$$

Wie wir oben gesehen haben lauten die Tranformationsformeln (mit α^j anstelle von x^j)

$$\alpha^j = \sum_i s^j_i \tilde{\alpha}^i \quad , \quad \tilde{\alpha}^l = \sum_k \tilde{s}^l_k \alpha^k$$

(2) Linearformen $f \in V^*$.

Die Koordinaten von f bezüglich der dualen Basen $\mathcal{B}^*$ bzw. $\widetilde{\mathcal{B}}^*$ werden gegeben durch

$$\alpha_j = f(v_j) \quad , \quad \tilde{\alpha}_j = \tilde{f}(v_j) \, ,$$

Wie wir oben gesehen haben lauten die Tranformationsformeln (mit α_j anstelle von y_j)

$$\alpha_j = \sum_i \tilde{s}^i_j \tilde{\alpha}_i \quad , \quad \tilde{\alpha}_l = \sum_k s^k_l \alpha_k \, .$$

Merke: die Koordinaten von Vektoren $v \in V$ transformieren sich kontravariant, die Koordinaten von Linearformen $f \in V^*$ transformieren sich kovariant.

- **Tensoren 2-ter Stufe**. Man unterscheidet hier vier Typen, die sich durch im wesentlichen durch ihr Tranformationsverhalten unterscheiden.
  (1) Jede Bilinearform $F\colon V \times V \to \mathbb{R}$ wird auch *kovarianter* Tensor zweiter Stufe genannt (z.B. Spannungstensor, Trägheitstensor). Koordinatendarstellung (vgl. 8.1) bzgl. der Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$:
  $$\alpha_{ij} = F(v_i, v_j) \quad , \quad \tilde{\alpha}_{ij} = F(\tilde{v}_i, \tilde{v}_j)\,.$$
  Transformationsformeln (vgl. 8.1.3):
  $$\alpha_{ij} = \sum_{k,l} \tilde{s}_i^k \tilde{s}_j^l \tilde{\alpha}_{kl} \quad , \quad \tilde{\alpha}_{k,l} = \sum_{i,j} s_k^i s_l^j \alpha_{ij}\,.$$
  (2) Jede Bilinearform $F\colon V^* \times V^* \to \mathbb{R}$ wird *kontravarianter* Tensor zweiter Stufe genannt. Koordinatendarstellung bzgl. der dualen Basen $\mathcal{B}^*$ und $\widetilde{\mathcal{B}}^*$:
  $$\alpha^{ij} = F(v^i, v^j) \quad , \quad \tilde{\alpha}^{ij} = F(\tilde{v}^i, \tilde{v}^j)\,.$$
  Transformationsformeln:
  $$\alpha^{ij} = \sum_{k,l} s_k^i s_l^j \tilde{\alpha}_{kl} \quad , \quad \tilde{\alpha}^{kl} = \sum_{i,j} \tilde{s}_i^k \tilde{s}_j^l \alpha^{ij}\,.$$
  (3) Jede Bilinearform $F\colon V^* \times V \to \mathbb{R}$ wird auch Tensor zweiter Stufe gemischter Varianz genannt. Koordinatendarstellung:
  $$\alpha_j^i = F(v^i, v_j) \quad , \quad \tilde{\alpha}_j^i = F(\tilde{v}^i, \tilde{v}_j)\,.$$
  Transformationsformeln:
  $$\alpha_j^i = \sum_{k,l} s_k^i \tilde{s}_j^l \tilde{\alpha}_k^l \quad , \quad \tilde{\alpha}_l^k = \sum_{i,j} \tilde{s}_i^k s_l^j \alpha_j^i\,.$$
  (4) Eine Linearform $V \times V^* \to \mathbb{R}$ ist ein Tensor gemischter Varianz analog dem Vorhergehenden.
- Tensoren 3-ter Stufe:

  Trilinearformen $F\colon V \times V \times V \to \mathbb{R}$ (kovariant) bzw. $F\colon V^* \times V^* \times V^* \to \mathbb{R}$ (kontravariant) bzw. $F\colon V^* \times V \times V \to \mathbb{R}$ (gemischter Varianz) usw.

  Eine Trilinearform wird bezüglich einer Basis durch Skalare mit drei Indizes beschrieben. Im Falle einer Trilinearform $F\colon V^* \times V \times V \to \mathbb{R}$ betrachte man
  $$\alpha_{jk}^i = F(v^i, v_j, v_k)$$

  Übung: Stellen sie die Transformationsformeln bei Basiswechsel auf.

KAPITEL 9

Die Jordansche Normalform

Wie wir gesehen haben, können nicht alle Matrizen diagonalisiert werden. Die Obstruktion zur Diagonalisierbarkeit rührt von den nilpotenten Matrizen her, die deshalb in diesem Abschnitt besondere Aufmerksamkeit verdienen. Wir werden zeigen, dass jede Matrix über $\mathbb{C}$ zu einer Matrix in sogenannter *Jordanscher Normalform* ähnlich ist, also einer Matrix mit nur wenigen Einträgen außerhalb der Hauptdiagonalen.

Jordansche Normalformen können zur Berechnung von Matrixpotenzen verwendet werden und damit auch zur Lösung von linearen Rekursionen und gewöhnlichen linearen Differentialgleichungssystemen.

In diesem Kapitel arbeiten wir in einem n-dimensionalen Vektorraum V über dem Körper $\mathbb{C}$ der komplexen Zahlen. Ferner sei $\varphi\colon V \to V$ eine lineare Abbildung und A sei A sein eine $n \times n$-Matrix. Häufig wird $V = \mathbb{C}^n$ sein und A die Matrix von φ bezüglich der Standardbasis.

9.1. Matrizen in Jordanscher Normalform

9.1.1 Jordanblöcke. Ein *Jordanblock* oder eine *elementare Jordanmatrix* ist eine $m \times m$-Matrix der Form

$$J := J_{m,\lambda_0} := \begin{pmatrix} \lambda_0 & 1 & & \text{\Large 0} \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ \text{\Large 0} & & & \lambda_0 \end{pmatrix}$$

Das charakteristische Polynom von J ist $\det(J - \lambda E) = (\lambda_0 - \lambda)^m$, also ist λ_0 der einzige Eigenwert; er hat die algebraische Vielfachheit m.

Die Eigenvektoren v zum Eigenwert λ_0 erfüllen die Gleichung

$$(J - \lambda_0 E)v = 0,$$

woraus $x_2 = 0, x_3 = 0, \ldots, x_m = 0$ und $x_1 \in \mathbb{C}$ beliebig folgt; d.h. bis auf skalare Vielfache ist

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

der einzige Eigenvektor. Also ist die geometrische Vielfachheit des Eigenwertes λ_0 gleich 1. Insbesondere gibt es also keine Basis von Eigenvektoren für J, d.h. J ist nicht diagonalisierbar für $m \geq 2$.

9.1.2 Einige Spezialfälle von Jordanblöcken.

$$J_{1,\lambda_0} = (\lambda_0)\,, \qquad J_{2,\lambda_0} = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}\,, \qquad J_{3,\lambda_0} = \begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$$

$$J_{2,0} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\,, \qquad J_{3,0} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

9.1.3 Jordansche Normalformen. Eine Matrix ist in *Jordanscher Normalform*, wenn sie von der folgenden Gestalt ist

$$A' = \begin{pmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_t} \end{pmatrix}$$

wobei jedes der J_1, J_2, …, J_t ein Jordanblock ist.

9.1.4 Hauptsatz. *Ist φ ein Endomorphismus eines n-dimensionalen komplexen Vektorraums V, dann kann man eine Basis $\mathcal{B} : v_1, v_2, \ldots, v_n$ in V so finden, daß die Matrix A' von φ bzgl. der Basis $\mathcal{B}$ in Jordanscher Normalform ist. (Solch eine Basis $\mathcal{B}$ wird* Jordanbasis *genannt.)*

Ist A eine $n \times n$-Matrix mit komplexen Einträgen, dann gibt es eine invertierbare Matrix S, derart daß

$$A' = S^{-1}AS$$

eine Matrix in Jordanscher Normalform ist. Für die Spalten von S nehme man die Vektoren einer Jordanbasis des durch A (bezüglich der Standardbasis) auf $\mathbb{C}^n$ definierten Endomorphismus. (Die Matrix A' wird Jordansche Normalform *von A genannt.)*

Die Jordansche Normalform einer Matrix A ist eindeutig bis auf eine Permutation der Jordanblöcke.

Die Einträge λ_i in der Diagonalen der Jordanschen Normalform A' sind die (komplexen) Eigenwerte von A; jeder Eigenwert tritt mit seiner algebraischen Vielfachheit s_i auf. Verschiedene Jordanblöcke können zu demselben Eigenwert λ_i gehören: Die Zahl der Jordanblöcke zum Eigenwert λ_i ist die geometrische Vielfachheit t_i des Eigenwertes λ_i.

Der Beweis der Existenz und Eindeutigkeit der Jordanschen Normalform wird in die nächsten Abschnitte verlagert.

9.1.5 Beispiel. Jede Diagonalmatrix ist in Jordanscher Normalform (die Jordanblöcke sind einelementig und enthalten lediglich einen Eintrag λ_i).

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Weitere Beispiele von Matrizen in Jordanscher Normalform:

$$\begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{3} \end{pmatrix}, \quad \begin{pmatrix} \begin{array}{|cc|} \hline 1 & 1 \\ 0 & 1 \\ \hline \end{array} & 0 \\ 0 & \begin{array}{|cc|} \hline 2 & 1 \\ 0 & 2 \\ \hline \end{array} \end{pmatrix}, \quad \begin{pmatrix} \begin{array}{|ccc|} \hline 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ \hline \end{array} & \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \\ \begin{array}{ccc} 0 & 0 & 0 \end{array} & \boxed{2} \end{pmatrix}, \quad \begin{pmatrix} \begin{array}{|cc|} \hline 2 & 1 \\ 0 & 2 \\ \hline \end{array} & & 0 \\ & \boxed{2} & \\ 0 & & \boxed{2} \end{pmatrix}$$

Die folgenden Matrizen sind *nicht* in Jordanscher Normalform:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

9.2. Verallgemeinerte Eigenvektoren

9.2.1 Vorbereitende Betrachtungen. Angenommen, $\mathcal{B} : v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ ist eine Basis von V, so daß die Matrix A' von φ bzgl. $\mathcal{B}$ in Jordanscher Normalform ist, wobei die ersten m Basisvektoren $v_1, \ldots, v_m$ zum ersten Jordanblock

$$J_1 = J_{m,\lambda_0}$$

gehören. Dann gilt:

$$\begin{aligned} \varphi(v_1) &= \lambda_0 v_1 & \qquad Av_1 &= \lambda_0 v_1 \\ \varphi(v_2) &= v_1 + \lambda_0 v_2 & Av_2 &= v_1 + \lambda_0 v_2 \\ &\vdots & &\vdots \\ \varphi(v_m) &= v_{m-1} + \lambda_0 v_m & Av_m &= v_{m-1} + \lambda_0 v_m \end{aligned}$$

Wird φ durch eine Matrix A beschrieben:

$$(*) \qquad \begin{aligned} (\varphi - \lambda_0 \,\mathrm{id})v_1 &= 0 & \qquad (A - \lambda_0 E)v_1 &= 0 \\ (\varphi - \lambda_0 \,\mathrm{id})v_2 &= v_1 & (A - \lambda_0 E)v_2 &= v_1 \\ &\vdots & &\vdots \\ (\varphi - \lambda_0 \,\mathrm{id})v_m &= v_{m-1} & (A - \lambda_0 E)v_m &= v_{m-1} \end{aligned}$$

9.2.2 Jordanketten. Eine Folge $v_1, v_2, \ldots, v_m$ von Vektoren wird *Jordankette* für φ (und entsprechend für A) zum Eigenwert λ_0, wenn $(*)$ mit $v_1 \neq 0$ erfüllt ist. In diesem Fall ist v_1 ein Eigenvektor zum Eigenwert λ_0. Weiterhin gilt

$$(\varphi - \lambda_0 \,\mathrm{id})^k v_j = \begin{cases} v_{j-k} & \text{wenn } k < j \\ 0 & \text{sonst} \end{cases} \quad , \quad (A - \lambda_0 E)^k v_j = \begin{cases} v_{j-k} & \text{wenn } k < j \\ 0 & \text{sonst} \end{cases}$$

Insbesondere gilt $(\varphi - \lambda_0 \,\mathrm{id})^m v_m = 0$ und $(A - \lambda_0 E)^m v_m = 0$, und m ist die kleinste natürliche Zahl mit dieser Eigenschaft.

Lemma. *Bilden die Vektoren $v_1, \ldots, v_m$ eine Jordankette für φ, dann sind sie linear unabhängig.*

Beweis. Nehmen wir an, die Vektoren $v_1, \ldots, v_m$ seien linear abhängig. Dann gibt es Skalare α_i, die nicht alle $= 0$ sind, derart daß $\sum_{i=1}^m \alpha_i v_i = 0$. Sei j der größte Index mit $\alpha_j \neq 0$. Nach obigem folgt $(\varphi - \lambda_0 \mathrm{id})^{j-1} v_j = v_1$ und $(\varphi - \lambda_0 \,\mathrm{id})^{j-1} v_i = 0$ für alle $i \leq j$. Mit Hilfe der Linearität folgt nun $0 = (\varphi - \lambda_0 \mathrm{id})^{j-1}\big(\sum_{i=1}^m \alpha_i v_i\big) = \sum_{i=1}^m \alpha_i (\varphi - \lambda_0 \mathrm{id})^{j-1}(v_i) = \lambda_j v_1$. Wegen $v_1 \neq 0$ folgt daraus $\lambda_j = 0$, ein Widerspruch zur Wahl von α_j. □

9.2.3 Verallgemeinerte Eigenvektoren. Ein Vektor $0 \neq v$ heißt *verallgemeinerter Eigenvektor* (oder auch *Hauptvektor*) von φ (bzw. von A) zum Eigenwert λ_i, falls es eine natürliche Zahl m so gibt, daß

$$(\varphi - \lambda_i \,\mathrm{id})^m v = 0 \quad \textit{bzw.} \quad (A - \lambda_i E)^m v = 0$$

Das kleinste m mit dieser Eigenschaft heißt die *Stufe* des verallgemeinerten Eigenvektors v.

Verallgemeinerte Eigenvektoren der Stufe 1 sind gerade Eigenvektoren im bisherigen Sinn. *Die Stufe m eines verallgemeinerten Eigenvektors zum Eigenwert λ_i ist kleiner oder gleich dessen algebraischer Vielfachheit s_i* : Ist nämlich v ein verallgemeinerter Eigenvektor der Stufe m, dann ist

$$v_1 := (\varphi - \lambda_i \,\mathrm{id})^{m-1} v, \qquad \ldots \qquad v_{m-1} := (\varphi - \lambda_i \,\mathrm{id}) v, \qquad v_m := v$$

eine Jordankette, also linear unabhängig. Ergänzt man die $v_1, \ldots, v_m$ zu einer Basis, dann hat die Matrix von φ bezüglich dieser Basis eine Blockgestalt, wo in der linken oberen Ecke der Jordanblock J_{m,λ_i} steht und darunter lauter Nullen. Daraus ergibt sich, daß die algebraische Vielfachheit von λ_i mindestens m ist.

$U^{(\lambda_i)}$ bezeichne die Menge aller verallgemeinerten Eigenvektoren von φ (bzw. A) zum Eigenwert λ_i von φ einschließlich des Nullvektors, d.h.

$$\begin{aligned}
U^{(\lambda_i)} &:= \{v \mid (\varphi - \lambda_i \,\mathrm{id})^m v = 0 \quad \text{für ein } m \in \mathbb{N}\} \\
&= \{v \mid (\varphi - \lambda_i \,\mathrm{id})^{s_i} = 0\} \\
&= \ker\, (\varphi - \lambda_i \,\mathrm{id})^{s_i} \quad \text{bzw.} \\
&= \{v \mid (A - \lambda_i E)^m v = 0 \quad \text{für ein } m \in \mathbb{N}\} = \{v \mid (A - \lambda_i E)^{s_i} v = 0\}
\end{aligned}$$

Als Kern der linearen Abbildung $(\varphi - \lambda_i \,\mathrm{id})^{s_i}$ ist $U^{(\lambda_i)}$ ein linearer Teilraum, genannt der *verallgemeinerte Eigenraum* von φ (bzw. von A) zum Eigenwert λ_i.

9.2.4 Lemma. *$U^{(\lambda_i)}$ ist invariant unter φ (bzw. A), d.h. für alle $v \in U^{(\lambda_i)}$ gilt $\varphi(v) \in U^{(\lambda_i)}$ (bzw. $Av \in U^{(\lambda_i)}$).*

Beweis. Zum Nachweis sei $v \in U^{(\lambda_i)}$. Dann ist $(A - \lambda_i E)^m v = 0$ für ein m. Es folgt $(A - \lambda_i E)^m Av = A(A - \lambda_i E)^m v = A0 = 0$. Also gilt $Av \in U^{(\lambda_i)}$. □

9.2.5 Eindeutigkeit der Jordanschen Normalform. Wir zeigen zunächst, daß die Jordansche Normalform eindeutig bis auf Permutation der Jordanblöcke ist. Das ist gleichwertig zu der Aussage, dass sich zwei ähnliche Jordansche Normalformen nur durch eine Permutation ihrer Blöcke unterscheiden.

Dazu bilden wir für jeden Eigenwert λ_i die (aufsteigende) Folge von ganzen Zahlen

$$\begin{array}{l}\dim(\ker(\varphi - \lambda_i \,\mathrm{id}))\\ \dim(\ker(\varphi - \lambda_i \,\mathrm{id})^2)\\ \dim(\ker(\varphi - \lambda_i \,\mathrm{id})^3)\\ \dots\end{array}$$

Diese Sequenz bestimmt die Größe und Zahl der Jordanblöcke, was man am leichtesten durch ein Beispiel sieht: Es sei für einen Endomorphismus φ

$$\begin{array}{l}\dim(\ker(\varphi - \lambda_i \,\mathrm{id})) = 5,\\ \dim(\ker(\varphi - \lambda_i \,\mathrm{id})^2) = 9,\\ \dim(\ker(\varphi - \lambda_i \,\mathrm{id})^3) = 11,\\ \dim(\ker(\varphi - \lambda_i \,\mathrm{id})^4 = 12 = s_i\end{array}$$

Aus dieser Sequenz von Zahlen ersehen wir, daß wir in unserer Jordanbasis 5 Eigenvektoren, $4 = 9 - 5$ verallgemeinerte Eigenvektoren der Stufe eins, $2 = 11 - 9$ verallgemeinerte Eigenvektoren der Stufe zwei und $1 = 12 - 11$ verallgemeinerte Eigenvektoren der höchsten Stufe drei haben. Der einzige Weg um das zu erreichen, besteht aus einer Jordankette der Länge 4, einer der Länge 3, zwei der Länge 2 und einem einzelnen Eigenvektor.

Allgemeiner ist die Zahl der Jordanketten der Länge t gleich $\dim(\ker(\varphi - \lambda_0 \,\mathrm{id})^t) - \dim(\ker(\varphi - \lambda_0 \,\mathrm{id})^{t-1})$.

Der Beweis des Hauptsatzes über die Existenz der Jordanschen Normalform zerfällt in zwei Teile. Der erste Teil besteht aus dem folgenden Satz, der besagt, daß wir für jede $n \times n$-Matrix über $\mathbb{C}$ eine Basis von $\mathbb{C}^n$ finden können, die aus verallgemeinerten Eigenvektoren von A besteht. (Wir lassen hier den Beweis weg. Hierfür konsultiere man ein beliebiges Buch zur Linearen Algebra). Anschließend werden wir zeigen, dass jede solche Basis in eine Jordanbasis umgeformt werden kann.

9.2.6 Satz. *Sei φ ein Endomorphismus eines endlichdimensionalen komplexen Vektorraums V (bzw. sei A eine komplexe $n \times n$-Matrix). Dann ist V (bzw. $\mathbb{C}^n$) die direkte Summe der verallgemeinerten Eigenräume zu den (paarweise verschiedenen) Eigenwerten $\lambda_1, \dots, \lambda_r$ von φ (bzw. A):*

$$V = U^{(\lambda_1)} \oplus \cdots \oplus U^{(\lambda_r)}$$

Weiterhin ist $\dim U^{(\lambda_i)} = s_i$, die algebraische Vielfachheit von λ_i.

Auf Grund dieses Satzes genügt es, in jedem der verallgemeinerten Eigenräume $U^{(\lambda_i)}$ eine Jordanbasis zu finden. All diese Jordanbasen zusammengenommen ergeben dann eine Jordanbasis des ganzen Raumes. Hierfür wollen wir zunächst einen Spezialfall betrachten.

9.3. Nilptente Endomorphismen bzw. Matrizen

9.3.1 Definition. Ein Endomorphimus ψ eines Vektorraums V wird *nilpotent* genannt, falls es eine natürliche Zahl m so gibt, daß $\psi^m = 0$. Entsprechend wird eine quadratische Matrix B *nilpotent* genannt, falls es eine natürliche Zahl m so gibt, daß $B^m = 0$. Die kleinste natürliche Zahl mit dieser Eigenschaft wird die *Nilpotenzstufe* von ψ bzw. von B genannt.

9.3.2 Beispiel.

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 \\ 1 & 0 & -1 & -1 & -1 \end{pmatrix} \quad , \quad B^2 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \quad , \quad B^3 = 0$$

Also ist B nilpotent der Stufe 3.

Nilpotenz kann wie folgt gekennzeichnet werden:

9.3.3 Satz. *Ein Endomorphismus ψ eines n-dimensionalen Vektorraums V (bzw. eine $n \times n$-Matrix B) ist genau dann nilpotent, wenn 0 der einzige komplexe Eigenwert von ψ (bzw. von B) ist. Die Nilpotenzstufe ist dann $\leq n$.*

Beweis. Für jeden Eigenwert λ und jeden zugehörigen Eigenvektor v von ψ gilt $\psi(v) = \lambda v$, woraus $\psi^2(v) = \psi(\psi(v)) = \psi(\lambda v) = \lambda\psi(v) = \lambda^2 v$, ... , $\psi^m(v) = \lambda^m v$ folgt. Ist $\psi^m = 0$, folgt $0 = \lambda^m v$, also $\lambda^m = 0$ und folglich $\lambda = 0$; d.h., 0 ist der einzige Eigenwert von ψ. Ist umgekehrt 0 der einzige Eigenwert von ψ, so ist dessen algebraische Vielfachheit n und V nach Satz 9.2.6 ist V gleich dem verallgemeinerten Eigenraum $U^{(0)}$. Für jedes $v \in V$ gilt dann $\psi^m(v) = 0$ und folglich $\psi^m = 0$. $\square$

Wir wollen nun zeigen, daß nilpotente Endomorphismen und Matrizen eine Jordansche Normalform besitzen. Da der einzige Eigenwert 0 ist, müssen die Jordanblöcke die folgende Gestalt haben:

$$J = J_{m,0} = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \cdot & \cdot & \vdots \\ 0 & \dots & & 0 & 1 & 0 \\ 0 & \dots & & & 0 & 1 \\ 0 & \dots & & & \dots & 0 \end{pmatrix}$$

Bildet man die Potenzen $J^2, J^3, \dots$, dann rutschen die Einsen in der Nebendiagonalen jeweils um einen Schritt nach außen. Die Potenz J^{m-1} hat nur noch einen Eintrag mit einer 1, nämlich oben rechts und $J^m = 0$.

Eine Jordankette für einen nilpotenten Endomorphismus ψ ist eine Folge von Vektoren

$$v_1, \dots, v_m \text{ mit } \psi(v_j) = v_{j-1} \text{ für } j = 2, \dots, m \text{ und } \psi(v_1) = 0$$

Dabei ist v_1 Eigenvektor (zum Eigenwert 0).

Sei nun ψ ein nilpotenter Endomorphismus und

$$\begin{array}{ll}\mathcal{C}: & v_1^1, v_2^1, \ldots, v_{m_1}^1 \\ & v_1^2, v_2^2, \ \ldots \ , v_{m_2}^2 \\ & \vdots \\ & v_1^s, v_2^s, \ \ \ldots \ \ \ , v_{m_s}^s \end{array}$$

eine Familie von s Jordanketten der Längen $m_1, m_2, \ldots, m_s$ für ψ. Dann gilt:

9.3.4 Lemma. *a) Sind die Eigenvektoren $v_1^1, \ldots, v_1^s$ linear unabhängig, dann ist das System $\mathcal{C}$ aller Vektoren wie oben linear unabhängig.*

b) Sind die Eigenvektoren $v_1^1, \ldots, v_1^s$ linear abhängig, dann kann man eine der Jordanketten verkürzen, ohne den von den Jordanketten erzeugten linearen Teilraum zu verändern.

Beweis. a) Nehmen wir an, die Vektoren des Systems $\mathcal{C}$ seien linear abhängig. Es gibt dann Skalare α_{ij}, die nicht alle $= 0$ sind, derart daß $\sum \alpha_{ij} v_j^i = 0$. Sei k maximal unter den Indices, für die $\alpha_{ik} \neq 0$ ist. Wenden wir ψ^{k-1} auf die obige Linearkombination an, so erhalten wir eine lineare Abhängigkeit zwischen den Eigenvektoren $v_1^1, \ldots, v_1^s$ im Widerspruch zur Voraussetzung.

b) Sind die Eigenvektoren v_1^i linear abhängig, dann gibt es Skalare α_i, die nicht alle $= 0$ sind, derart daß $\sum_{i=1}^s \alpha_i v_1^i = 0$. Wir wählen aus den Jordanketten $v_1^i, \ldots, v_{m_i}^i$, für welche $\alpha_i \neq 0$ ist, eine von minimaler Länge aus. Nach Umnummerierung können wir annehmen, daß $v_1^1, \ldots, v_{m_1}^1$ eine Jordankette minimaler Länge ist, d.h., $\alpha_1 \neq 0$ und $m_1 \leq m_i$, für alle i mit $\alpha_i \neq 0$. Wir zeigen, daß diese Jordankette durch eine kürzere ersetzt werden kann.

Falls $m_1 = 1$ können wir die Kette v_1^1 einfach entfernen, da ihr alleiniges Element eine Linearkombination von Vektoren der anderen Jordanketten ist.

Im anderen Fall beachten wir, däs $v_1^1 + \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} v_1^i = 0$ und wir definieren

$$\tilde{v}_j^1 = v_{j+1}^1 + \sum_i \frac{\alpha_i}{\alpha_1} v_{j+1}^i$$

für $j = 1, \ldots, m_1 - 1$, wobei die Summe nur über die $i > 1$ läuft, für welche $\alpha_i \neq 0$. Man beachte, daß wegen der Minimalitätsannahme v_{j+1}^i in diesem Fall existiert. Nun ist $\tilde{v}_1^1, \ldots, \tilde{v}_{m_1-1}^1$ wieder eine Jordankette (der Länge $m_1 - 1$); denn $\psi(\tilde{v}_j^1) = \psi(v_{j+1}^1) + \sum_i \frac{\alpha_i}{\alpha_1} \psi(v_{j+1}^i) = v_j^1 + \sum_i \frac{\alpha_i}{\alpha_1} v_j^i = \tilde{v}_j^1$ for $j = 2, \ldots, m_1 - 1$ und $\psi(\tilde{v}_1^1) = \psi(v_2^1) + \sum_i \frac{\alpha_i}{\alpha_1} \psi(v_2^i) = v_1^1 + \sum_i \frac{\alpha_i}{\alpha_1} v_1^i = 0$. Eine Ersetzung der alten Kette $v_1^1, \ldots, v_{m_1}^1$ durch $\tilde{v}_1^1, \ldots, \tilde{v}_{m_1-1}^1$ verändert nicht den erzeugten Untervektorraum. $\square$

Dieses Lemma liefert einen Algorithmus für das Auffinden einer Jordanbasis für einen nilpotenten Endomorphismus ψ eines n-dimensionalen Vektorraums V. Insbesondere folgt daraus die Existenz einer Jordanbasis für den nilpotenten Fall. Das Verfahren kann man natürlich in derselben Weise auf nilpotente Matrizen anwenden:

9.3.5 Algorithmus für das Auffinden einer Jordanbasis. Für einen nilpotenten Endomorphismus ψ (ganz analog für eine nilpotente Matrix B) gehe man so vor:

Schritt 1. Wähle eine Basis $u_1, \ldots, u_n$ in V. (Im Falle $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ kann man einfach die Standardbasis $e_1, \ldots, e_n$ nehmen.)

Schritt 2. Für jeden Basisvektor u_i bilde eine Jordankette durch Anwenden von ψ:

$$(\mathcal{B}_i) \qquad v_1^i = \psi^{m_i-1}(u_i)\ ,\ \ldots\ ,\ v_{m_i-1}^i = \psi(u_i)\ ,\ v_{m_i}^i = u_i$$

Da jedes u_i verallgemeinerter Eigenvektor einer Stufe m_i zum Eigenwert 0 ist, bricht diese Kette nach m_i Schritten ab (gerade bevor $\psi^{m_i}(u_i) = 0$ wird). Die Gesamtheit dieser Jordanketten wird nun linear abhängig sein.

Schritt 3. Wende das Verfahren aus Lemma 9.3.4(b) an, um die Jordanketten zu verkürzen, bis die Gesamtlänge der Jordanketten auf n verkürzt ist. Diese bilden dann ein Jordanbasis.

Schritt 4. Ist ψ ein nilpotenter Endomorphismus von $\mathbb{C}^n$ und B seine Matrix bezüglich der Standardbasis, dann ist $B' = S^{-1}BS$ die Jordansche Normalform von B, wobei die Spalten von S gerade die Vektoren der Jordanbasis sind.

Beachte: In den Jordanketten $\mathcal{B}_i$ entstehen die Vektoren durch wiederholtes Anwenden von ψ; man muß die so entstehenden Vektoren aber in umgekehrter Reihenfolge von hinten her anordnen, um eine Jordankette zu haben.

9.3.6 Beispiel. Wir wenden obiges Verfahren auf die Matrix B aus Beispiel 9.3.2 an.

Schritt 1. In $\mathbb{C}^5$ wählen wir die Standardbasis e_1, e_2, e_3, e_4, e_5.

Schritt 2. Anwenden von B auf die Basisvektoren führt zu den Jordanketten

$$\begin{pmatrix}1\\-1\\1\\-1\\1\end{pmatrix} = Be_1, \quad e_1$$

$$\begin{pmatrix}1\\-1\\1\\-1\\1\end{pmatrix} = B^2e_2, \quad e_1 = Be_2, \qquad\qquad e_2$$

$$\begin{pmatrix}1\\-1\\1\\-1\\1\end{pmatrix} = B^2e_3 \quad , \begin{pmatrix}0\\1\\-1\\1\\-1\end{pmatrix} = Be_3, \quad e_3$$

$$\begin{pmatrix}0\\0\\0\\1\\-1\end{pmatrix} = Be_4, \quad e_4$$

$$\begin{pmatrix}0\\0\\0\\1\\-1\end{pmatrix} = Be_5, \quad e_5$$

Man bemerke daß man die Vektoren Be_i und B^2e_i in den $i-$ten Spalten von B bzw. B^2 ablesen kann.

Schritt 3. Die jeweils ersten Vektoren der dritten und vierten obiger Jordanketten sind linear unabhängig. Nach 9.3.4(a) sind die fünf Vektoren dieser beiden Jordanketten zusammengenommen linear unabhängig. Also bilden

$$B^2e_3, Be_3, e_3, Be_4, e_4$$

eine Basis des $\mathbb{C}^5$, also eine Jordanbasis für B. (In diesem Fall brauchten wir das Verfahren zur Verkürzung von Jordanketten gar nicht anwenden, da wir aus den im Schritt 2 entstandenen Jordanketten direkt eine Jordanbasis auswählen konnten.)

Schritt 4. Jordansche Normalform

$$B' = \begin{pmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&1\\0&0&0&0&0\end{pmatrix} = S^{-1}BS \quad \text{mit} \quad S = \begin{pmatrix}1&0&0&0&0\\-1&1&0&0&0\\1&-1&1&0&0\\-1&1&0&1&1\\1&-1&0&-1&0\end{pmatrix}$$

9.4. Bestimmung einer Jordanbasis im allgemeinen Fall

Sei nun φ ein beliebiger Endomorphismus von $\mathbb{C}^n$ und A seine Matrix bezüglich der Standardbasis. Um eine Jordanbasis und die Jordansche Normalform sowie die Transformationsmatrix zu bestimmen gehen wir ähnlich wie bei nilpotenten Matrizen (siehe 9.3.5) vor:

Schritt 0. Bestimme die paarweise verschiedenen komplexen Eigenwerte $\lambda_1, \ldots, \lambda_r$ von A sowie ihre algebraischen Vielfachheiten $s_1, \ldots, s_r$.

Wir setzen nun $\psi_i = \varphi - \lambda_i \,\mathrm{id}$ und $B_i = A - \lambda_i E$.

Schritt 1. Bestimme eine Basis $u_1, \ldots, u_{s_i}$ des verallgemeinerten Eigenraums $U^{(\lambda_i)}$ für jeden der Eigenwerte λ_i. Das läuft darauf hinaus ein Fundamentalsystem von Lösungen des homogenen linearen Gleichungssystems

$$(A - \lambda_i E)^{s_i} v = A_i^{s_i} v = 0$$

zu bestimmen.

Da $0 = (\varphi - \lambda_i \,\mathrm{id})^{s_i}(v) = \psi_i^{s_i}(v) = B_i^{s_i} v$, wirkt die Einschränkung von ψ_i und ebenso von B_i auf dem verallgemeinerten Eigenraum $U^{(\lambda_i)}$ als nilpotenter Endomorphismus der Stufe $\leq s_i$. Wir fahren daher fort wie im nilpotenten Fall:

Schritt 2. Erzeuge Jordanketten aus den Basisvektoren $u_1, \ldots, u_{s_i}$ von $U^{(\lambda_i)}$ durch wiederholtes Anwenden von A_i.

Schritt 3. Nach der Methode von Lemma 9.3.4(b) modifiziere diese Jordanketten bis sie eine Jordanbasis von $U^{(\lambda_i)}$ bilden.

Schritt 4. Fasse die Jordanbasen der $U^{(\lambda_i)}$ zusammen. Man erhält dadurch eine Jordanbasis für φ.

9.4.1 Beispiel. Sei

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

Das charakteristische Polynom von A ist $P(\lambda) = (\lambda - 1)^5$, also ist $\lambda_1 = 1$ der einzige Eigenwert mit algebraischer Vielfachheit $s_1 = 5$.

In diesem Fall gilt $\mathbb{C}^5 = U^{(\lambda_1)}$, also bildet z.B.

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, v_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

eine Basis von $U^{(\lambda_1)}$. Wir hätten natürlich auch e_1, e_2, e_3, e_4, e_5 wählen können.

Sei nun $(A-E)v_1 = w$ und $(A-E)v_2 = -w$ mit

$$w = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \end{pmatrix}$$

und $(A-E)v_3 = (A-E)v_4 = e_3$ und ebenso $(A-E)v_5 = 0$.

Die von v_1 und v_2 beginnenden Jordanketten sind also linear abhängig und wir können v_1, w durch

$$\tilde{v}_1 = v_1 + \frac{1}{1}v_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ersetzen, was einen Eigenvektor liefert.

Weiterhin ist $(A-E)v_3 = (A-E)v_4 = v_5$, was ein Eigenvektor ist. Die Jordanketten, die von v_3 und v_4 beginnen, sind also linear abhängig und wir können v_3 durch

$$\tilde{v}_3 = v_3 - \frac{1}{1}v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}$$

ersetzen, was ein Eigenvektor ist.

Wir haben folgendes System von Jordanketten erhalten: $\tilde{v}_1, v_2 \to -w, \tilde{v}_3, v_4 \to v_5$.

Die vier Eigenvektoren sind linear abhängig: $-w - v_5 - \frac{1}{2}\tilde{v}_3 + \frac{1}{2}\tilde{v}_1 = 0$, wir entfernen $\tilde{v}_1$, um ein linear unabhängiges System von Jordanketten: $-w, v_2, v_5, v_4, \tilde{v}_3$ zu erhalten, d.h. die Transformationsmatrix

$$S = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & 2 \end{pmatrix}$$

Beachte, dass die Jordanketten in S eingesetzt werden, indem mit dem Eigenvektor begonnen wird!

Jetzt gilt tatsächlich

$$S^{-1}AS = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Mitunter wird der Vorschlag gemacht, man solle, um eine Jordanbasis zu erhalten, rückwärts von den Eigenvektoren für einen Eigenwert λ_0 arbeiten, d.h. starte von einer linear unabhängigen Menge von Eigenvektoren $v_1^1, \ldots, v_1^t$ (hier wäre t die geometrische Vielfachheit von λ_0) und bilde Jordanketten durch Lösen von z.B. $(A - \lambda_0 E)v_2^i = v_1^i$ und dann von $(A - \lambda_0 E)v_3^i = v_2^i$ etc. bis keine weiteren Lösungen mehr gefunden werden können. Dieses Verfahren führt aber nicht notwendigerweise auf eine Jordanbasis.

Im obigen Beispiel ist die geometrische Vielfachheit 3 und damit gleich der Anzahl der Jordanblöcke. Die Vektoren

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

sind drei linear unabhängige Eigenvektoren von A, wobei keiner von ihnen im Bild von $A - \lambda_0 E$ ist, d.h. das Ende einer Jordankette der Länge > 1. Allgemeiner treten solche Beispiele wie folgt auf: ist v ein Eigenvektor, so dass $(A - \lambda_0 E)x = v$ keine Lösungen hat und v_1 ein Eigenvektor, der eine Jordankette der Länge > 1 terminiert, d.h. $v_1 = (A - \lambda_0 E)v_2$, dann ist $v_1 + v$ ein anderer Eigenvektor, so dass $(A - \lambda_0 E)x = (v_1 + v)$ keine Lösung hat.

9.5. Anwendungen

9.5.1 Matrixpotenzen. Ist J eine Jordansche Normalform, dann können wir $J = D + N$ schreiben, wobei D eine Diagonalmatrix und N eine Matrix mit verschwindender Hauptdiagonale und einigen Einsen oberhalb der Hauptdiagonale ist. Ist J gerade ein Jordanblock, dann gilt $D = \lambda E$ und natürlich $DN = ND$. Ist J eine Jordansche Normalform aus mehreren Blöcken, die möglicherweise zu verschiedenen Eigenwerten gehört, dann gilt $DN = ND$ weiter, wie man anhand von Blockmultiplikation sehen kann.

Es gilt also

$$J^k = (D + N)^k = \sum_{i=0}^{k} \binom{k}{i} D^{k-i} N^i$$

Nur die ersten Terme dieser Summe sind von Null verschieden, genauer diejenigen, bei denen i kleiner als die Größe des größten Jordanblocks in J ist. Die Summe kann also mit vertretbarem Aufwand berechnet werden. Ist z.B. J die Jordansche Normalform aus dem letzten Beispiel, so gilt

$$J^l = \begin{pmatrix} 1 & l & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & l & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Natürlich folgt wie gewöhnlich aus $S^{-1}AS = J$ auch $A^l = SJ^lS^{-1}$.

Man erinnere sich, dass sich die Lösung einer linearen Rekursion erster Ordnung

$$f(0) = g, f(k+1) = Af(k)$$

mit $f : \mathbb{N} \to \mathbb{C}^n, g \in \mathbb{C}^n, A \in \mathbb{C}^{n \times n}$ zu $f(n) = A^n g$ ergibt. Ist A nicht diagonalisierbar, dann können wir die Jordansche Normalform benutzen, um A^n zu berechnen. Betrachte z.B. das folgende System einer linearen Rekursionen erster Ordnung:

$$\begin{aligned} a_{k+1} &= 3a_k + b_k \\ b_{k+1} &= 3b_k \\ c_{k+1} &= -b_k + 3c_k + d_k \\ d_{k+1} &= 3d_k \\ e_{k+1} &= -b_k + 3e_k \\ a_0 = b_0 = c_0 = d_0 = e_0 = 1 & \end{aligned}$$

Schreiben wir $f(k) = (a_k\ b_k\ c_k\ d_k\ e_k)^t$ und $g = (1\ 1\ 1\ 1\ 1)^t$, so folgt $f(k+1) = (A+2E)f(k), f(0) = g$ wobei A die Matrix aus unserem Beispiel ist. Wegen $S^{-1}AS = J$ gilt auch $S^{-1}(A+2E)S = J+2E$, woraus wir die Jordansche Normalform von $A+2E$ ohne weitere Berechnung erhalten.

Aus dem Obigen erkennen wir

$$(J+2E)^k = \begin{pmatrix} 3^k & k3^{k-1} & 0 & 0 & 0 \\ 0 & 3^k & 0 & 0 & 0 \\ 0 & 0 & 3^k & k3^{k-1} & 0 \\ 0 & 0 & 0 & 3^k & 0 \\ 0 & 0 & 0 & 0 & 3^k \end{pmatrix}$$

woraus wir folgende Lösung in geschlossener Form erhalten:

$$f(k) = S(J+2E)^k S^{-1} g$$

oder

$$\begin{aligned} a(k) = e(k) &= 3^{k-1}k + 3^k \\ b(k) = c(k) = d(k) &= 3^k \end{aligned}$$

Damit können wir Matrixpotenzen benutzen, um eine geschlossene Form für $e^A = \sum_{k=1}^{\infty} \frac{A^k}{k!}$ zu erhalten und damit Systeme linearer Differentialgleichungen lösen. Sie werden darüber mehr in späteren Vorlesungen erfahren.

KAPITEL 10

Gruppen, Symmetrien, Symmetriegruppen

Warum spielen Gruppen eine Rolle in der Physik? Gruppen geben Symmetrieeigenschaften wieder. Die Kenntnis der Symmetriegruppe einer Konfiguration erlaubt es die Untersuchung dieser Konfiguration durch Ausnutzen der Symmetrieeigenschaften zu vereinfachen. Physikalische Größen sind invariant unter gegenüber gewissen Transformationsgruppen.

10.1. Gruppen

Wir wiederholen die Definition des Begriffs einer Gruppe.

10.1.1 Der Begriff der Gruppe. Eine *Gruppe* ist ein Paar $(G, *)$ bestehend aus einer Menge G und einer zweistelligen Operation

$$(g, h) \mapsto g * h : \; G \times G \to G$$

welche die folgenden Eigenschaften erfüllt:

(G1) $\forall\, g, h, k \in G. \quad g * (h * k) = (g * h) * k$ (Assoziativgesetz)

(G2) $\exists\, e \in G. \forall\, g \in G. \quad g * e = g = e * g$ (Existenz eines neutralen Elements)

(G3) $\forall\, g \in G.\; \exists\, g^{-1} \in G. \quad g * g^{-1} = e = g'^{-1} g$ (Existenz von Inversen)

Statt $(G, *)$ bezeichnen wir eine Gruppe oft einfach mit G, wenn die Operation aus dem Zusammenhang klar ist. In der Regel benutzen wir die multiplikative Schreibweise $g \cdot h$ für die Gruppenoperation. Oft lassen wir das Operationszeichen ganz weg und schreiben einfach gh statt $g \cdot h$.

10.1.2 Abelsche Gruppen. Eine Gruppe $(G, *)$ wird *kommutativ* (oder auch *abelsche Gruppe*) genannt, falls zus'atlich gilt

(G4) $\forall\, g, h \in G. \quad g * h = h * g$ (Kommutativgesetz)

In kommutativen Grupppen wird die Operation oft durch $+$ bezeichnet (additive Schreibweise). In diesem Fall bezeichnet man das Neutralelement mit 0 statt mit e und das (additiv) Inverse von g mit $-g$ statt mit g^{-1}.

10.1.3 Beispiel. a) **Abelsche Gruppen**: $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{R}^n, +)$. Jeder Vektorraum ist bezüglich der Addition eine abelsche Gruppe. Die von 0 verschiedenen Elemente von $\mathbb{R}$ bzw. $\mathbb{C}$ bilden eine multiplikative abelsche Gruppe.

b) **Allgemeine lineare Gruppen:** Die Menge $(Gl(n,\mathbb{R}),\cdot\ )$ der invertierbaren $n\times n$ - Matrizen mit reellen Einträgen ist eine Gruppe bezüglich der Multiplikation von Matrizen. Man kann $\mathbb{R}$ durch den Körper $\mathbb{C}$ der komplexen Zahlen ersetzen. Das Neutralelement ist hier die Einheitsmatrix

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

Da die Multiplikation von Matrizen i.a. nicht kommutativ ist, liegen hier nicht-kommutative Gruppen vor.

c) **Transformationsgruppen:** Die Menge $\mathcal{S}_X$ aller bijektiven Abbildungen $\sigma : X \to X$ einer Menge X in sich mit der Hintereinanderausführung $\circ$ von Abbildungen als Operation ist eine Gruppe. Die identische Abbildung id_X ist das Neutralelement, die Umkehrabbildung liefert das Inverse σ^{-1}.

Spezialfall: $\mathcal{S}_n$ bezeichne die Gruppe der Permutationen der n-elementigen Menge $X = \{1, 2, \ldots, n\}$; diese wird auch *symmetrische Gruppe* genannt. Auch diese Gruppen sind nicht kommutativ.

Beispiel: $\mathcal{S}_3$ hat die sechs Elemente

$$\underset{id}{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}},\ \underset{\zeta_1}{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}},\ \underset{\zeta_2}{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}}$$

$$\underset{\tau_{12}}{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}},\ \underset{\tau_{23}}{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}},\ \underset{\tau_{13}}{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}}$$

10.2. Untergruppen, Homomorphismen, direkte Produkte

10.2.1 Definition. Eine Teilmenge H einer Gruppe G wird *Untergruppe* genannt, falls sie die folgenden Eigenschaften erfüllt:

(U1) $e \in H$ (additive Schreibweise: $0 \in H$

(U2) $g, h \in H \Rightarrow g \cdot h \in H$ $\quad g, h \in H \Rightarrow g + h \in H$

(U3) $g \in H \quad \Rightarrow g^{-1} \in H$ $\quad g \in H \Rightarrow -g \in H$)

Bemerkung. Jede Untergruppe H einer Gruppe ist selbst auch eine Gruppe bezüglich der induzierten Operation.

Merke: Hat man zu zeigen, daß etwas eine Gruppe ist, versuche man nachzuweisen, daß es eine Untergruppe einer schon bekannten Gruppe ist.

10.2.2 Beispiel. a)

$$\mathbb{T} := \{z \in \mathbb{C} \ : \ |z| = 1\} = \{e^{i\alpha} \mid 0 \leq \alpha \leq 2\pi\}$$

ist eine Untergruppe von $(\mathbb{C}^\times, \cdot\ )$; man nennt sie die *Kreisgruppe*.

b) **Matrizengruppen**, das sind Untergruppen von $\mathrm{GL}(n,\mathbb{R})$ und $Gl(n,\mathbb{C})$, z. B. $\mathrm{O}(n)$, die Gruppe der orthogonalen $n\times n$-Matrizen, die *orthogonale Gruppe* [*orthogonal group*]
$\mathrm{SO}(n)$, die Gruppe der orthogonalen $n\times n$-Matrizen S der Determinante $\det(S)=+1$, die *spezielle orthogonale Gruppe* [*special orthogonal group*], die Gruppe $\mathrm{U}(n)$ der unitären $n\times n$-Matrizen, die *unitäre Gruppe* [*unitary group*] und die Gruppe $\mathrm{SU}(n)$ der unitären $n\times n$-Matrizen der Determinante $+1$, die *spezielle unitäre Gruppe* [*special unitary group*].

c) **Permutationsgruppen**, das sind Untergruppen der symmetrischen Gruppe $\mathcal{S}_n$.

G und H seien Gruppen (mit multiplikativ geschriebener Gruppenoperation).

10.2.3 Homomorphismen und Isomorphismen. Eine Abbildung $f : G \to H$ wird *Homomorphismus* genannt, wenn gilt

(H) $$f(xy) = f(x)f(y) \quad \text{für alle} \quad x,y\in G\,.$$

Ein bijektiver Homomorphismus wird *Isomorphismus* genannt. Wir sagen, die Gruppen G und H seien *isomorph* und wir schreiben dann $G\cong H$, wenn es einen Isomorphismus der Gruppe G auf die Gruppe H gibt.

10.2.4 Eigenschaften von Homomorphismen. Für einen Homomorphismus $f : G\to H$ gilt:

(i) $f(e)=e$.

(ii) $f(x^{-1})=f(x)^{-1}$ für alle $x\in G$.

10.2.5 Beispiel. a) Die Exponentialfunktion $x\mapsto e^x$ ist ein Isomorphismus der additiven Gruppe der reellen Zahlen auf die multiplikative Gruppe der positiven reellen Zahlen. Die Gruppen $(\mathbb{R},+\ )$ und $(\mathbb{R}_+^\times,\cdot\ )$ sind also isomorph.

(b) $t\mapsto e^{2\pi i t} : (\mathbb{R},+\ )\to\mathbb{C}^\times$ ist ein Homomorphismus von der additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der von 0 verschiedenen komplexen Zahlen. Das Bild ist die Kreisgruppe $\mathbb{T}$.

(c) **Innere Automorphismen** : Sei g ein beliebiges Element einer Gruppe G. Die Abbildung

$$\varphi_g : G\to G \text{ definiert durch } \varphi_g(x) = gxg^{-1}\ (\forall\ x\in G)$$

ist ein Isomorphismus von G auf sich selbst. Die Abbildungen dieser Gestalt nennt man innere Automorphismen.

(Beweis: $\varphi_g(xy)=gxyg^{-1}=gx(g^{-1}g)yg^{-1}=(gxg^{-1})(gyg^{-1})=\varphi_g(x)\cdot\varphi_g(y)$; also ist φ_g ein Endomorphismus. φ_g ist bijektiv, denn $\varphi_{g^{-1}}$ ist Umkehrabbildung zu φ_g wie man leicht nachrechnet.)

10.2.6 Beispiel. (a) Die Determinante $A\mapsto\det(A)\colon\ \mathrm{GL}(n,\mathbb{K})\to\mathbb{K}^\times$ ist ein Homomorphismus der Gruppe aller invertierbaren $n\times n$-Matrizen in die multiplikative Gruppe der von) verschidenen Elemente des Grundkörpers $\mathbb{K}=\mathbb{R}$ oder $\mathbb{C}$. Der *Kern* dieses Homomorphismus $\mathrm{SL}(n,\mathbb{K}) := \{A\in\mathrm{GL}(n,\mathbb{K}) :$

$\det(A) = 1\}$ ist eine Untergruppe von $\mathrm{GL}(n, \mathbb{K})$, genannt die *spezielle lineare Gruppe* [*special linear group*].

(b) Die Signatur (= das Signum) $\sigma \mapsto \operatorname{sign}(\sigma)\colon \mathcal{S}_n \to \{-1, 1\}$ ist ein Homomorphismus der $\mathcal{S}_n$ in die multiplikative Gruppe $\{-1, +1\}$. Der *Kern* dieses Homomorphismus $\mathcal{A}_n = \{\sigma \in \mathcal{S}_n \,|\, \operatorname{sign}(\sigma) = +1\}$ besteht aus den geraden Permutationen. Die Menge $\mathcal{A}_n$ aller geraden Permutationen ist eine Untergruppe der $\mathcal{S}_n$. Man nennt sie die *alternierende Gruppe* [*alternating group*].

Durch Bildung direkter Produkte kann man neue Gruppen aus schon bekannten konstruieren.

10.2.7 Direkte Produkte. Gegeben seien zwei Gruppen G_1 und G_2. Auf der Menge

$$G_1 \times G_2 \;:=\; \{(x_1, x_2) \mid x_1 \in G_1,\ x_2 \in G_2\}$$

definieren wir eine Operation komponentenweise durch:

$$(x_1, x_2) \cdot (y_1, y_2) \;:=\; (x_1 y_1, x_2 y_2).$$

Mit dieser Operation ist $G_1 \times G_2$ eine Gruppe, genannt das *direkte Produkt* der Gruppen G_1 und G_2. Das Neutralelement von $G_1 \times G_2$ ist (e, e), das Inverse von (x_1, x_2) ist (x_1^{-1}, x_2^{-1}).

10.2.8 Beispiel. (a) Die Symmetriegruppe D_2 eines (nichtquadratischen) Rechtecks ist isomorph zum direkten Produkt zweier zweielementigen Gruppen.

D_2 besteht aus der Identität 1
der Spiegelung σ_1 an der senkrechten Achse g_1
der Spiegelung σ_2 an der waagrechten Achse g_2
der Spiegelung δ am Mittelpunkt

Multiplikationstafel:

	1	σ_1	σ_2	δ
1	1	σ_1	σ_2	δ
σ_1	σ_1	1	δ	σ_2
σ_2	σ_2	δ	1	σ_1
δ	δ	σ_2	σ_1	1

$H_1 = \{1, \sigma_1\}$ und $H_2 = \{1, \sigma_2\}$, sind Untergruppen. Man prüfe nach, daß die Abbildung $(h_1, h_2) \mapsto h_1 h_2\colon H_1 \times H_2 \to D_2$ ein Isomorphismus ist.

(b) Die multiplikative Gruppe $\mathbb{C}^\times$ der von 0 verschiedenen komplexen Zahlen ist isomorph zum direkten Produkt der multiplikativen Gruppe $\mathbb{R}_{>0}$ der positiven reellen Zahlen und der Kreisgruppe $\mathbb{T}$ der komplexen Zahlen z vom Betrag 1. In der Tat prüfe man nach, daß die Abbildung $(r, e^{i\varphi}) \mapsto re^{i\varphi}\colon \mathbb{R}_{>0} \times \mathbb{T} \to \mathbb{C} \setminus \{0\}$ ein Isomorphismus ist.

10.3. Symmetriegruppen

An Ornamenten, Kristallen, usw. fasziniert immer wieder die Regelmäßigkeit, mit der gewisse Muster auftauchen. Es ist Aufgabe der Mathematik, den Begriff der Regelmäßigkeit sauber zu formulieren und Muster zu klassifizieren.

Bei Ornamenten stellt man fest, daß gewisse Verschiebungen (Translationen), Spiegelungen und Drehungen sowie Hintereinanderausführungen von solchen, das Ornament mit sich selber zur Deckung bringen. macht man eine solche Abbildung rückgängig, kommt wieder das Ornament mit sich selbst zur Deckung. Es liegt nahe daß diese Abbildungen eine Gruppe bilden. Welche Abbildungen lassen wir nun zu?

10.3.1 Bewegungen. Wir betrachten den $\mathbb{R}^n$ mit dem Standard-Skalarprodukt und dem zugehörigen Euklidischen Abstand von Punkten $a, b \in \mathbb{R}^n$:

$$d(a,b) = ||b-a|| = \sqrt{\sum_{i=1}^{n}(b_i-a_i)^2}$$

Eine bijektive Abbildung $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ wird eine *Bewegung* oder auch eine *Isometrie* genannt, falls sie Abstände erhält, d.h. falls

$$d(\psi(a), \psi(b)) = d(a,b) \quad \text{für alle } a, b \in \mathbb{R}^n$$

Beispiele für Bewegungen des $\mathbb{R}^2$ bzw. des $\mathbb{R}^3$ sind:

Translationen (Verschiebungen) $\tau_b\colon x \mapsto x + b$
Spiegelungen an einer beliegigen Geraden bzw. beliebigen Ebene
Drehungen um einen beliebigen Punkt bzw. eine beliebige Achse

Da die identische Abbildung eine Bewegung ist, da die Umkehrabbildung einer bewegung wieder eine bewegung ist und die hintereinanderausführung einer Bewegung wieder eine Bewegung ist, ist die Menge $B(n)$ aller Bewegungen des $\mathbb{R}^n$ eine Untergruppe der vollen Transformationsgruppe des R^n, also selbst eine Gruppe:

Satz. *Die Bewegungen des $\mathbb{R}^n$ bilden eine Gruppe $B(n)$ (mit der Hintereinanderausführung von Abbildungen als Operation).*

Nicht ganz so leicht ist die folgende Aussage zu beweisen:

10.3.2 Satz. *Zu jeder Bewegung ψ des $\mathbb{R}^n$ gibt es eine eindeutig bestimmte orthogonale Abbildung $\varphi \in \mathrm{O}(n)$ und ein eindeutig bestimmtes $b \in \mathbb{R}^n$ derart, daß*

$$\psi = \tau_b \circ \varphi$$

d.h.,

$$\psi(x) = \varphi(x) + b \quad \textit{für alle } x \in \mathbb{R}^n$$

Ist A die (orthogonale) Matrix von φ bezüglich der Standardbasis, so hat man die Darstellung

$$\psi(x) = Ax + b \quad \textit{für alle } x \in \mathbb{R}^n$$

Beweis: Übung.

10.3.3 Symmetriegruppen. Sei M eine Teilmenge des $\mathbb{R}^n$. Jede Bewegung ψ, welche das Gebilde M mit sich selbst zur Deckung bringt, nennt man eine Symmetrie von M. Genauer: Eine *Symmetrie* von M ist eine Bewegung ψ mit der Eigenschaft daß $\psi(M) = M$.

Es ist klar, daß die Identität eine Symmetrie von M ist. Sind ψ und ψ' Symmetrien von M, dann auch $\psi \circ \psi'$ und ψ^{-1}. Also bilden die Symmetrien von M eine Untergruppe G der Bewegungsgruppe $B(n)$; wir nennen sie die Symmetriegruppe von M.

Wir wollen zunächst die Symmetrien eines regelmäßigen n-Ecks betrachten.

10.3.4 Die Diedergruppen D_n. Wir betrachten ein regelmäßiges n-Eck im euklidischen $\mathbb{R}^2$. Natürlich ist dann $n \geq 3$. Der Mittelpunkt des n-Ecks sei der Koordinatenursprung. Jede orthogonale Abbildung des $\mathbb{R}^2$ in sich, welche das n-Eck auf sich abbildet, ist eine *Symmetrie* des n-Ecks. (Ein Translationsanteil kann hier nicht auftreten. Man überlegt sich schnell, dass die sämtlichen Symmetrien eines regelmäßigen n-Ecks eine Untergruppe der Gruppe $\mathrm{O}(2)$ aller orthogonalen Abbildungen des $\mathbb{R}^2$ bilden; man bezeichnet sie mit D_n und nennt sie die *Diedergruppe* [*dihedral group*].

Die Gruppe D_n hat $2n$ Elemente:

- die n Drehungen um die Winkel $\frac{2\pi}{n}k,\ k = 0, 1, \cdots, n-1$,
- die n Spiegelungen an den n Symmetrieachsen:
  - Falls n *gerade* ist, gibt es zwei Typen von Symmetrieachsen:
    (1) $\frac{n}{2}$ Geraden durch gegenüberliegende Ecken.
    (2) $\frac{n}{2}$ Geraden durch gegenüberliegende Kantenmitten.
  - Falls n *ungerade* ist, gibt es n Spiegelungen von demselben Typ.

Die n Drehungen bilden eine Untergruppe N von D_n. Sie besteht aus den Potenzen der Drehung δ um den Winkel $\frac{2\pi}{n}$. Insbesondere gilt

(1) $$\delta^n = \mathrm{id}\ .$$

Der Index der Untergruppe N in D_n ist 2.

Für jede Spiegelung σ an einer der Symmetrieachsen gilt

(2) $$\sigma^2 = \mathrm{id}\ .$$

Insbesondere ist $\{\mathrm{id}, \sigma\}$ eine zweielementige Untergruppe. Man überzeugt sich leicht, daß folgende Gleichung gilt:

(3) $$\delta\sigma = \sigma\delta^{-1} = \sigma\delta^{n-1}\,.$$

Wählen wir eine Spiegelungen σ an einer der Symmetrieachsen aus, dann ist die Gruppe D_n die disjunkte Vereinigung der Untergruppe N und der *Nebenklasse* σN. Jedes Element von D_n lässt sich also eindeutig in der Form

$$\sigma^i \delta^k \text{ mit } i = 0, 1;\ k = 0, 1, \cdots, n-1$$

schreiben, wobei δ wie oben die Drehung um den Winkel $\frac{2\pi}{n}$ bezeichnet. Die Gleichungen (1), (2), (3) legen eindeutig fest wie man in dieser Gruppe rechnet.

Nun wenden wir uns den Symmetriegruppen von räumlichen Gebilden zu. Zunächst einige Hilfsbetrachtungen:

10.3.5. O(3) bezeichne die Gruppe aller orthogonalen Abbildungen ($\cong$ orthogonalen Matrizen) des euklidischen $\mathbb{R}^3$ in sich. Für $\varphi \in \mathrm{O}(3)$ gilt $\det \varphi = \pm 1$, und zwar:

$$\det \varphi = \begin{cases} +1 & \text{falls } \varphi \text{ eine Drehung ist.} \\ -1 & \text{falls } \varphi \text{ eine Spiegelung an einer Ebene oder eine Drehspiegelung ist.} \end{cases}$$

SO(3) bezeichnet die Menge aller $\varphi \in \mathrm{O}(3)$ mit $\det \varphi = +1$.

Hilfssatz. *Sei G eine Untergruppe von* O(3)*, sei* $G_+ := \{\varphi \in G : \det \varphi = 1\} = G \cap \mathrm{SO}(3)$. *Setzt man voraus, dass $G \neq G_+$, d.h. dass G ein σ mit* $\det \sigma = -1$ *enthält, dann ist G_+ eine Untergruppe in G, und G ist die disjunkte Vereinigung von G_+ und der* Nebenklasse σG_+.

Beweis. Wegen $\det \varphi\psi = \det \varphi \cdot \det \psi$ ist $\det : G \to \{1, -1\}$ ein surjektiver Homomorphismus von G auf die zweielementige multiplikative Gruppe $\{1, -1\}$. Man prüft leicht nach, daß $G_+ = \{\varphi \in G |\det(\varphi) = 1\}$ eine Untergruppe in G ist. Ferner ist $\sigma G_+ = \{\varphi \in G | \det(\varphi) = -1\}$. Also ist G disjunkt zerlegt in G_+ und σG_+. $\square$

10.3.6 Die Symmetriegruppe eines konvexen Polyeders P. Gegeben sei ein konvexes Polyeder P (z.B. Tetraeder, Würfel, Oktaeder, Dodekaeder, Ikosaeder, Pyramide, Prisma). Wir legen den Koordinatenursprung in den Mittelpunkt (= Schwerpunkt) von P. Symmetrien von P können keinen Translationsanteil enthalten. Also ist

$$G := \{\varphi \in \mathrm{O}(3) : \varphi(P) = P\}$$

die Symmetriegruppe von O(3). Wie oben sei

$$G_+ := \{\varphi \in G : \det \varphi = 1\} = G \cap \mathrm{SO}(3)$$

die *Gruppe der Drehungen* von P. Zur Kenntnis von G genügt im Wesentlichen die Kenntnis von G_+. Denn entweder ist $G = G_+$ oder es gibt ein $\sigma \in G \setminus G_+$. Im zweiten Fall ist nach 10.3.5 die Gruppe G die disjunkte Vereinigung von G_+ und σG_+; d.h. die

Elemente von G sind entweder Drehungen oder die Hintereinanderausführung von σ und einer Drehung (Spiegelungen oder Drehspiegelungen).

10.3.7 Die Symmetriegruppe des Würfels. bezeichnen wir mit $\mathbb{O}$; mit $\mathbb{O}_+$ bezeichnen wir die Gruppe der Drehsymmetrien des Würfels. Die Drehgruppe $\mathbb{O}_+$ hat 24 Elemente, die in 4 Klassen zerfallen:

Klasse	Beschreibung	Anzahl
1	id	1
2	Drehungen um die 3 Achsen durch gegenüberliegende Seitenmittelpunkte, jeweils um $\frac{\pi}{2}$, π, $\frac{3}{2}\pi$.	9
3	Drehungen um die 6 Achsen durch diametral gegenüberliegende Kantenmitten, jeweils um Winkel π.	6
4	Drehungen um die 4 Achsen durch diametral gegenüberliebende Ecken, jeweils um Winkel $\frac{2}{3}\pi$, $\frac{4}{3}\pi$.	8

Man kann sich leicht überlegen, dass der Würfel keine anderen Drehsymmetrien besitzt; denn als Drehachsen kommen aus geometrischen Gründen nur die oben angegebenen in Frage.

Sei σ die Punktspiegelung am Mittelpunkt. Deren Matrix ist $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Also ist $\det \sigma = -1$. Damit ist $\sigma \in \mathbb{O} \setminus \mathbb{O}_+$. Gemäß 10.3.5 hat die volle Symmetriegruppe $\mathbb{O}$ also 48 Elemente: die oben angegebenen Drehungen und die Hintereinanderausführung von σ mit diesen Drehungen. Es gilt sogar:

10.3.8. $\mathbb{O}$ ist isomorph zu dem direkten Produkt von $\mathbb{O}_+$ und der von der Punktspiegelung σ erzeugten zweielementigen Gruppe.

Beweis. Übung. □

10.3.9 Satz. *Die Drehgruppe $\mathbb{O}_+$ des Würfels ist zu der symmetrischen Gruppe $\mathcal{S}_4$ aller Permutationen auf einer vierelementigen Menge isomorph.*

Beweis. Wir nummerieren die Raumdiagonalen des Würfels so wie auf der vorhergehenden Seite (oben angegeben) mit 1,2,3,4. Diametral gegenüberliegende Ecken erhalten dieselbe Nummer. Da bei einer Drehsymmetrie eine Raumdiagonale in eine ebensolche übergeht, induziert jede Drehsymmetrie φ eine Permutation $\pi(\varphi)$ der Raumdiagonalen 1,2,3,4. Es ist klar, dass $\pi(\varphi\psi) = \pi(\varphi)\pi(\psi)$ gilt, d.h. dass

$$\pi : \mathbb{O}_+ \to \mathcal{S}_4$$

ein Gruppenhomomorphismus ist. Da der Kern von π nur aus der Identität besteht, ist π injektiv. Da $\mathbb{O}_+$ und $\mathcal{S}_4$ beide 24 Elemente besitzen, muss π dann auch surjektiv sein. Also ist π ein Isomorphismus. □

10.3.10 Andere Symmetriegruppen.

(a) Die Symmetriegruppe des Oktaeders ist dieselbe wie die Symmetriegruppe des Würfels (Oktaedergruppe), denn der Oktaeder entsteht, wenn man die Mittelpunkte benachbarter Seiten des Würfels verbindet.

(b) Aus demselben Grund wie in (a) ist die Symmetriegruppe $\mathbf{Y}$ des Ikosaeders dieselbe wie die Symmetriegruppe des Dodekaeders. Es gilt:

$$\mathbf{Y}_+ \cong \mathcal{A}_5 \qquad\qquad \mathbf{Y} \cong \mathcal{A}_5 \times \mathbb{Z}_2$$

(c) Es bezeichne $\mathbf{T}$ die Isometriegruppe des Tetraeders. Es gilt:

$$\mathbf{T}_+ \cong \mathcal{A}_4 \qquad\qquad \mathbf{T} \cong \mathcal{S}_4$$

(d) Ist G die Symmetriegruppe eines Prismas, dessen Grundfläche ein regelmäßiges n-Eck ist, so ist:

$$G_+ \cong D_n \qquad\qquad G \cong D_n \times \mathbb{Z}_2$$

Dabei sind $\mathcal{A}_4$ und $\mathcal{A}_5$ alternierende Gruppen (siehe 10.2.6(b)), die D_n sind die Diedergruppen und $\mathbb{Z}_2$ bezeichnet eine zweielementige Gruppe.