

# Lineare Algebra 1

Hans Joachim Pflug (pflug@itc.rwth-aachen.de)

Benno Willemsen (willemsen@itc.rwth-aachen.de)

Matthias Grajewski (grajewski@fh-aachen.de)

16. September 2015



# Inhaltsverzeichnis

<b>1</b>	<b>Motivation und Vorbereitung</b>	<b>5</b>
1.1	Inhalt und Anwendungen der Linearen Algebra . . . . .	5
1.2	Mathematisches Arbeiten und Problemlösen . . . . .	8
1.3	Vektoren im $\mathbb{R}^n$ . . . . .	10
1.4	Addition und Multiplikation im $\mathbb{R}^n$ . . . . .	13
1.5	Lineare Gleichungssysteme und Matrizen . . . . .	15
1.6	Lösung eines linearen Gleichungssystems . . . . .	18
1.6.1	Das Gaußsche Eliminationsverfahren . . . . .	18
1.6.2	Unter- und überbestimmte Gleichungssysteme . . . . .	22
<b>2</b>	<b>Analytische Geometrie</b>	<b>27</b>
2.1	Skalarprodukt und Norm . . . . .	27
2.1.1	Einheitsvektoren . . . . .	32
2.1.2	Orthogonale (senkrechte) Vektoren . . . . .	33
2.1.3	Winkel zwischen Vektoren . . . . .	35
2.1.4	Das Vektorprodukt (Kreuzprodukt) . . . . .	39
2.2	Geraden und Ebenen . . . . .	43
2.2.1	Definition und grundlegende Eigenschaften . . . . .	43
2.2.2	Umrechnen zwischen verschiedenen Darstellungsformen . . . . .	50
2.2.3	Lagebeziehungen von Geraden und Ebenen . . . . .	52
2.2.4	Schnittmengen zwischen Geraden und Ebenen im $\mathbb{R}^2$ und $\mathbb{R}^3$ . . . . .	54
2.2.4.1	Schnittmenge einer Gerade oder Ebene mit einer Hyperebene . . . . .	54
2.2.4.2	Bestimmung des Schnittpunkts zweier Geraden im $\mathbb{R}^3$ . . . . .	56
2.2.5	Abstandsbestimmung im $\mathbb{R}^2$ und $\mathbb{R}^3$ . . . . .	57
2.2.5.1	Einführung . . . . .	57
2.2.5.2	Der Abstand eines Punktes von einer Geraden im $\mathbb{R}^3$ . . . . .	60
2.3	Die Determinante im $\mathbb{R}^2$ und $\mathbb{R}^3$ . . . . .	61
2.3.1	Berechnung und geometrische Deutung . . . . .	61

2.3.2	Lineare $3 \times 3$ -Gleichungssysteme . . . . .	65
<b>3</b>	<b>Algebraische Strukturen</b>	<b>67</b>
3.1	Gruppen . . . . .	67
3.2	Körper . . . . .	72
3.3	Vektorräume . . . . .	74
3.4	Lineare Unabhängigkeit, Basis, Dimension . . . . .	77
3.4.1	Lineare Unabhängigkeit . . . . .	77
3.4.2	Nachweis linearer Unabhängigkeit . . . . .	81
3.4.3	Basis und Dimension . . . . .	84
3.4.4	Exkurs: Nicht endlich erzeugte Vektorräume . . . . .	88
3.4.5	Exkurs: Hyperebenen im $\mathbb{R}^n$ . . . . .	90
3.5	Polynome . . . . .	94
3.6	Skalarprodukt, euklidische und unitäre Räume . . . . .	101
3.7	Orthogonalität in unitären Vektorräumen . . . . .	107
3.8	Das Verfahren von Gram-Schmidt und Anwendungen . . . . .	114
	<b>Index</b>	<b>123</b>
	<b>Literatur</b>	<b>127</b>

# Kapitel 1

## Motivation und Vorbereitung

### 1.1 Inhalt und Anwendungen der Linearen Algebra

Die Lineare Algebra ist keine eigenständige mathematische Disziplin wie etwa die Analysis oder die Geometrie und ist doch ein unverzichtbarer Bestandteil jeden Studiums, in dem mathematische Inhalte eine wesentliche Rolle spielen. Dies liegt daran, dass sie grundlegende mathematische Techniken und Begriffe bereitstellt, die die gesamte Mathematik durchsetzen und auf die man deswegen im weiteren Studium immer wieder zurückgreifen wird. Drei wesentliche Themenbereiche prägen die Lineare Algebra:

1. Lineare Gleichungssysteme
2. analytische Geometrie
3. algebraischen Strukturen

Wir werden zur Motivation drei Beispiele skizzieren, ohne dabei zu viel Wert auf mathematische Strenge zu legen.

**Beispiel 1.1:** Zur Eisenerzeugung wird im Hochofen Schrott eingeschmolzen, um die Kosten der Verhüttung zu senken und vorhandenes Material wiederzuverwenden. Bei Schrott handelt es sich nicht um chemisch reines Eisen, weil u. a. Metalle wie Kupfer oder Zink enthalten sind. Zudem wird in der Praxis selten reines Eisenerz erzeugt, sondern gleich eine Legierung hergestellt. Beim Betrieb einer Eisenhütte ergibt sich also die Frage, welche Schrottsorten man in welcher Menge beimischen darf, damit am Ende die gewünschte Legierung entsteht. Wir gehen vereinfachend von drei Schrottsorten  $S_1, S_2$  und  $S_3$  aus, die nach Massenanteil wie folgt zusammengesetzt sind:

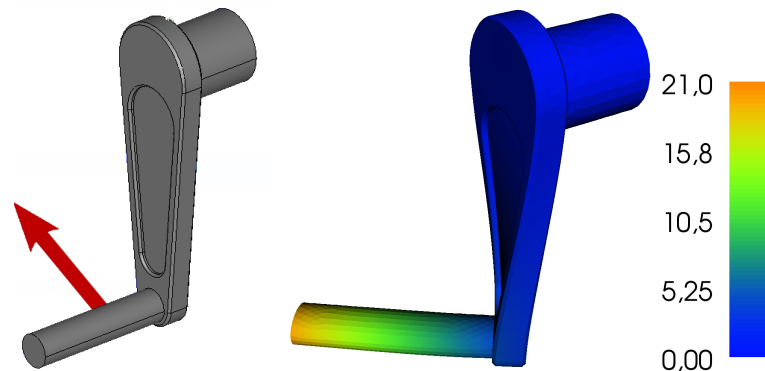


Abbildung 1.1: Handkurbel mit angreifender Kraft (links; als Pfeil dargestellt); resultierende Verformung (rechts; Verformung zur Verdeutlichung überhöht dargestellt)

	$S_1$	$S_2$	$S_3$
Eisen	0,8	0,8	0,8
Kupfer	0	0,2	0,1
Zink	0,2	0	0,1

Das Endprodukt soll 80% Eisen, 12% Kupfer und 8% Zink enthalten. Wir bezeichnen die Massenanteile der drei Legierungen mit  $a, b$  und  $c$  und stellen die Massenbilanzen für alle drei Elemente getrennt auf. Für Eisen erhält man  $0,8a + 0,8b + 0,8c = 0,8$  und analog für Kupfer  $0a + 0,2b + 0,1c = 0,12$  bzw. für Zink  $0,2a + 0b + 0,1c = 0,08$ . Da alle drei Gleichungen zugleich gelten müssen, werden wir auf ein Lineares Gleichungssystem der Form

$$\begin{aligned}
 0,8a + 0,8b + 0,8c &= 0,8 \\
 0a + 0,2b + 0,1c &= 0,12 \\
 0,2a + 0b + 0,1c &= 0,08
 \end{aligned} \tag{1.1}$$

geführt. Man überprüft leicht, dass die Werte  $a = 0$ ,  $b = 0,2$  und  $c = 0,8$  das System lösen, indem man sie in die Formel (1.1) einsetzt und sich von der Gleichheit der beiden Seiten überzeugt.

**Beispiel 1.2:** Ein wesentliches Anwendungsgebiet der Mathematik ist die Numerische Simulation, d.h. die Verhaltensvorhersage eines physikalischen Systems mithilfe mathematischer Methoden. Ein physikalisches System kann ein Auto, die Erde, ein Atom oder ein mechanisches Bauteil wie die in Abb. 1.1 gezeigte Kurbel sein. Soll in einem Industrieunternehmen eine solche Kurbel produziert werden, wird man nach den Regeln und Verfahrensweisen des Maschinenbaus die Kurbel entwerfen, dann einen Prototypen bauen und diesen testen. Bei der Kurbel wird es wahrscheinlich

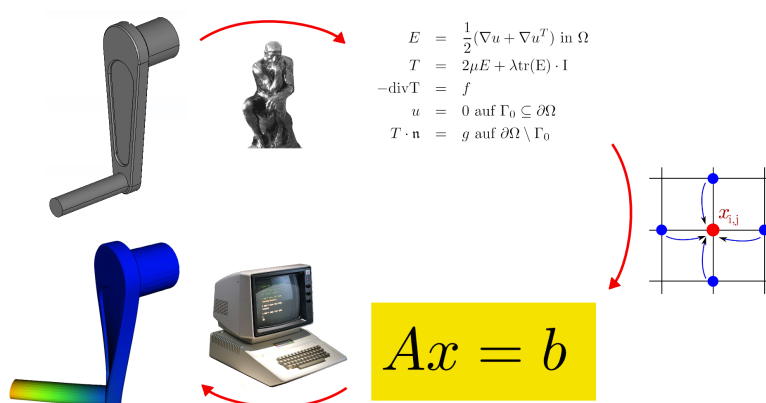


Abbildung 1.2: vereinfachtes Schema einer Numerischen Simulation

um die Frage gehen, ob sie allen Kräften standhält, die im Betrieb erwartet werden. Besteht sie alle Tests, kann sie in Serie gefertigt werden, ansonsten wird der Entwurf solange modifiziert und erneut getestet, bis er alle Anforderungen erfüllt. Mithilfe der Numerischen Simulation lassen sich die Tests der Prototypen im Entwicklungsprozess durch Berechnungen ersetzen. Dies spart Zeit und Geld. Diese Ersparnis mag bei der Kurbel unbedeutend sein, bei der Entwicklung eines Verkehrsflugzeuges ist sie es nicht. Numerische Simulation hat sich als unverzichtbares Werkzeug in der Produktentwicklung erwiesen.

Eine numerische Simulation, z. B. mit der Methode der finiten Elemente, ist das Ergebnis des Zusammenwirkens verschiedenster mathematischer Disziplinen wie Analysis, Numerik, Funktionalanalysis und Bereichen der Informatik sowie von Erkenntnissen aus Physik und Maschinenbau. Am Ende aller Überlegungen steht sehr häufig ein Lineares Gleichungssystem der Form  $Ax = b$ , das es zu lösen gilt (Abb. 1.2). Umfasste das Gleichungssystem in Beispiel 1.1 3 Bedingungen an 3 Unbekannte, sind bei Numerischer Simulation 1000000 Bedingungen an 1000000 Unbekannte nicht unüblich. Eine Handrechnung zur Lösung derartiger Systeme verbietet sich. Mit der effizienten Lösung mithilfe von Computern befasst sich die Numerische Lineare Algebra, die auf den Ergebnissen der Linearen Algebra aufbaut. Numerische Simulation setzt also ein tiefes Verständnis Linearer Gleichungssysteme notwendig voraus.

**Beispiel 1.3:** Der Begriff der Gruppe als einfache algebraische Struktur ist aus der Vorlesung Mathematische Grundlagen bekannt. Es handelt sich dabei um das Paar einer Menge  $M$  zusammen mit einer auf  $M$  definierten

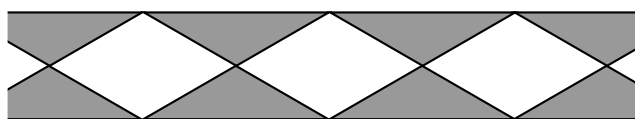


Abbildung 1.3: Beispiel eines Frieses

Verknüpfung  $\circ$ , das gewisse algebraische Bedingungen wie z. B. Assoziativität erfüllen muss. Wir betrachten ein Fries, also eine in einer Raumrichtung unendlich ausgedehnte ebene Struktur, die sich dieser Richtung wiederholt (Abb. 1.3). Wir betrachten als Menge  $M$  alle Starrkörperbewegungen (alle längen- und winkelerhaltenden bijektiven Abbildungen der Ebene in sich), die das Fries wieder in sich überführen und als Verknüpfung  $\circ$  auf  $M$  die Hintereinanderausführung von Abbildungen. Dann bildet  $(M, \circ)$  eine Gruppe, die sog. *Automorphismengruppe* des Frieses, kurz Friesgruppe. Man kann zeigen, dass zu jedem Fries-Typ genau eine Friesgruppe existiert und umgekehrt und dass es im Wesentlichen 7 Friesgruppen gibt. Es gibt also (bis auf Designmerkmale wie Farbe, Größe etc.) genau 7 verschiedene Frieze.

Analoge Betrachtungen lassen sich für unendliche periodische räumliche Objekte anstellen, die als mathematische Modelle für den atomaren Aufbau von Kristallen dienen. Fedorov<sup>1</sup> konnte 1891 zeigen, dass genau 230 sogenannter Raumgruppen existieren und damit den Aufbau von Kristallen klassifizieren. Dies war eine der ersten bedeutenden außermathematischen Anwendungen der Gruppentheorie und gleichzeitig ein großer Durchbruch in der Kristallographie.

## 1.2 Mathematisches Arbeiten und Problemlösen

Kern allen mathematischen Arbeitens ist die systematische Beschäftigung mit mathematischen Ideen, Problemen und ihrer Lösung. Konkrete Rechnungen sind dagegen häufig zweitrangig. Ein mathematisches Problem gilt als gelöst, wenn man eine entsprechende Aussage formulieren und streng beweisen kann bzw. ein geeignetes Gegenbeispiel angibt. Folgerichtig wird in der Vorlesung Lineare Algebra (und nicht nur da!) der Formulierung mathematischer Aussagen und ihrem Beweis breiter Raum gegeben. Dies unterscheidet sich bisweilen stark vom mathematischen Vorgehen außerhalb der Hochschulen. Dieser Paradigmenwechsel zu Beginn des Studiums ist der Erfahrung nach eine der wesentlichen Hürden für die Studierenden. Es ist normal, dass die Lösung eines mathematischen Problems nicht auf der Hand liegt und sollte nicht entmutigen. Wir skizzieren im Folgenden

<sup>1</sup>Evgraf S. Fedorov (1853-1919); russischer Kristallograph und Mineraloge



sehr kurz einige Techniken zur Lösung mathematischer Probleme. Für weitergehende Beschäftigung mit diesem Thema verweisen wir auf das sehr lesenswerte Buch von Grieser [3] und die klassische Abhandlung von Pólya [5].

Wir gehen davon aus, dass Ihnen eine mathematische Aussage vorgelegt wird und Sie diese beweisen oder widerlegen sollen. Dies ist eine Standardanforderung im Studium und zugleich eine große Vereinfachung, denn in der mathematischen Praxis ist es häufig viel schwieriger, die richtigen (d. h. zielführenden) Fragen zu stellen als die richtigen Antworten zu finden. Die folgenden Hinweise können Ihnen bei Ihrer Beschäftigung mit mathematischen Problemen helfen.

### 1. Versuchen Sie, das Problem zu verstehen!

- Sind Ihnen alle verwendeten mathematischen Begriffe in der Formulierung klar? Könnten Sie einem Kommilitonen alle verwendeten mathematischen Begriffe erklären? Wenn nicht, wiederholen Sie zunächst die entsprechenden Inhalte der Vorlesung.
- Sollte die Aussage eine Formel sein, versuchen Sie, dieselbe Aussage für Sie selbst als Text zu formulieren und umgekehrt.
- Betrachten Sie Beispiele und Spezialfälle. Kann man an Beispielen bereits erkennen, warum die Aussage wahr oder falsch sein sollte?
- Fertigen Sie, wann immer das Problem es ermöglicht, eine Skizze an!

### 2. Kennen Sie ähnliche Probleme?

- Kennen Sie bei vergleichbaren Aussagen sogar Beweis oder Gegenbeispiel?
- Kann man die gegebene Aussage auf Bekanntes zurückführen?
- Wenn nicht, was genau ist anders?

### 3. Vorwärtsarbeiten

- Jeder Beweisversuch lebt von den Voraussetzungen. Was lässt sich mit den gegebenen Voraussetzungen anfangen? Welche Aussagen lassen sich damit zeigen?

### 4. Rückwärtsarbeiten

- Unter welchen zusätzlichen Voraussetzungen könnten Sie denn die gewünschte Aussage beweisen?
- Wie kann man sich im zweiten Schritt von diesen zusätzlichen Voraussetzungen befreien?

### 5. Zwischenziele formulieren

- Bei komplexeren Sachverhalten kann es helfen, dass man das Gesamtproblem in Teilprobleme zerlegt, die man dann getrennt bearbeitet.
- Für jede Etappe lassen sich die oben skizzierten Problemlösestrategien verwenden.

### 6. Problemlösestrategien kombinieren und ausprobieren

- In vielen Fällen bringt erst eine Kombination der obigen Strategien den Erfolg.
- Wenn eine Strategie nicht weiterführt, muss man eine andere ausprobieren. Mathematik bedeutet manchmal auch hartnäckiges Herumprobieren!

### 7. Zum Schluss: richtig Aufschreiben!

- Schreiben Sie ihre Argumentation detailliert und nachvollziehbar auf. Oftmals wird umgekehrt die Argumentation erst bei ihrer Formulierung wirklich klar.
- Vermeiden Sie umständliche Prosa, sondern bedienen Sie sich mathematischer Formeln und Formulierungen.
- Kontrollieren Sie zum Schluss: Ist Ihre Beweisführung lückenlos? Ist das Gegenbeispiel wirklich eins?
- Für weiterführende Literatur zur Formulierung mathematischer Gedanken verweisen wir auf Beutelspacher [1].

## 1.3 Vektoren im $\mathbb{R}^n$

Ein epochaler Fortschritt in der Geometrie wurde mit der Einführung von Koordinaten durch René Descartes<sup>2</sup> erzielt (ihm zu Ehren spricht man vom "kartesischen Koordinatensystem"), weil hierdurch die seit alters her bestehende Kluft zwischen Geometrie und Arithmetik beseitigt werden konnte. Nun konnte man Aussagen der Geometrie mit Methoden der Arithmetik untersuchen und umgekehrt.

**Definition 1.4:** 1. Für  $n \in \mathbb{N}$  sei  $x$  das  $n$ -Tupel

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_i)_{i=1}^n$$

---

<sup>2</sup>1596-1650; franz. Mathematiker, Philosoph, Naturwissenschaftler

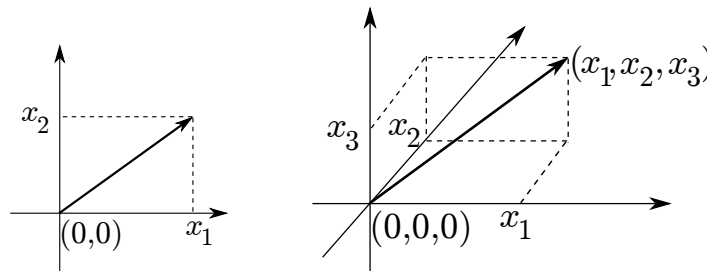


Abbildung 1.4: Beschreibung von Punkten in der Ebene (links) und im Raum (rechts) mithilfe eines Koordinatensystems

mit  $x_i \in \mathbb{R}$  für  $1 \leq i \leq n$ . Dann heißt  $x$  *Vektor*, die Zahl  $x_i$  die *i-te Koordinate*.

2. Zu  $x$  wie oben sei

$$x^T := (x_1, \dots, x_n)$$

der *transponierte Vektor*. Weiter definieren wir  $(x^T)^T := x$ .

3. Zwei Vektoren  $x, y$  sind *gleich*, wenn alle ihre Koordinaten gleich sind.

4. Der Zahlraum

$$\mathbb{R}^n = \{(x_1, \dots, x_n)^T \mid x_1, \dots, x_n \in \mathbb{R}\}$$

sei die Menge aller geordneten  $n$ -Tupel reeller Zahlen.

Wir betrachten den Spezialfall  $n = 2$ . Durch Wahl eines Koordinatensystems gelingt es, jeden Punkt der Ebene durch ein Zahlenpaar  $(x_1, x_2)$  zu parametrisieren (Abb. 1.4 links). Umgekehrt lässt sich jedes Zahlenpaar  $(x_1, x_2)$  als Punkt in der Ebene veranschaulichen. Daher kann man die Ebene mit dem Zahlraum  $\mathbb{R}^2$  identifizieren. Analog lässt sich durch die Wahl eines Koordinatensystems jeder Punkt des Raumes mit einem Zahlentripel  $(x_1, x_2, x_3) \in \mathbb{R}^3$  identifizieren (Abb. 1.4 rechts).

Der allgemeine Zahlraum  $\mathbb{R}^n$  eignet sich auch zur Beschreibung nicht-geometrischer Zusammenhänge.

**Beispiel 1.5:** In den Wirtschaftswissenschaften definiert man Räume mit Koordinaten, die die Gesamtproduktion von Waren oder Dienstleistungen wiedergeben, wobei als Maßstab der Preis (z. B. in 1.000.000 Dollar) verwendet wird. Wir betrachten 7 Wirtschaftszweige:

1.	Stahlindustrie		2.	Automobilindustrie
3.	Landwirtschaft		4.	Fischerei
5.	Chemische Industrie		6.	Textilindustrie
7.	Transportgewerbe			

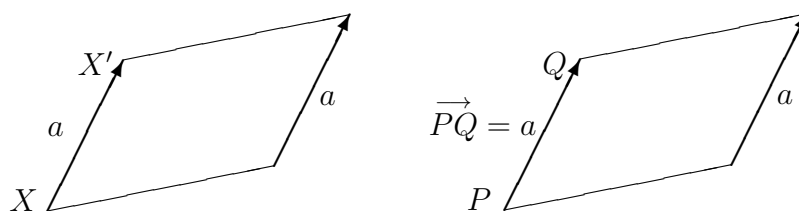


Abbildung 1.5: Verschiebungsvektor (links) und gebundener Vektor (rechts)

und modellieren die wirtschaftlichen Vorgänge mit dem  $\mathbb{R}^7$ . Das entsprechende Verfahren heißt *Input-Output-Analyse* oder *Leontief-Modell*<sup>3</sup>. Ordnen wir die 7 Koordinaten in der obigen Reihenfolge an, bedeutet der Punkt

$$(1000, 800, 550, 300, 700, 200, 900)^T,$$

dass die Stahlindustrie Waren in einem Wert von einer Milliarde Dollar und die Chemische Industrie von 700 Millionen Dollar produziert hat.

Wir kehren nun zur geometrischen Deutung von Vektoren im  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$  zurück. Die Interpretation eines Vektors als Beschreibung eines Punktes wie oben erfordert, dass man einen Nullpunkt festgelegt hat, weil die Koordinaten eines Punktes relativ zum Nullpunkt angegeben werden. Da aber in der Ebene bzw. im Raum alle Punkte gleichberechtigt sind, ist die Festlegung eines Koordinatenursprungs ein Akt der Willkür und daher geometrisch fragwürdig. Wir deuten vor diesem Hintergrund ein  $n$ -Tupel alternativ als *Verschiebung*. Die Koordinaten des Tupels geben an, wie man von irgendeinem Punkt  $X$  zu seinem Bildpunkt  $X'$  kommt (Abb. 1.5 links). Verschiebungen werden als *Verschiebungsvektor* oder allgemein als *Vektor* bezeichnet. Ein Verschiebungsvektor hat also eine bestimmte Richtung und Länge, aber keine bestimmte Lage, benötigt also auch nicht die Festlegung eines Koordinatenursprungs. Dies bringt die Definition des Vektors 1.4 mit der folgenden in der Geometrie verbreiteten in Einklang:

Ein *Vektor* ist eine gerichtete Strecke. Vektoren sind gleich, wenn sie durch Parallelverschiebung ineinander überführt werden können, d.h. gleiche Vektoren sind parallel, gleich lang und gleich gerichtet. Ein Vektor hat also eine bestimmte Richtung und Länge, aber keine bestimmte Lage.

Einen Vektor mit festem Anfangspunkt  $P$  und Endpunkt  $Q$  nennt man *gebunden* und bezeichnet ihn mit  $\vec{PQ}$  (Abb. 1.5 rechts). Ist  $P$  der Ursprung  $0$ , spricht man vom *Ortsvektor* und schreibt  $\vec{0A}$  oder einfach  $a$ . Der Endpunkt

<sup>3</sup>Wassily Leontief (1905-1999); russischer Wirtschaftswissenschaftler, Nobelpreis 1973

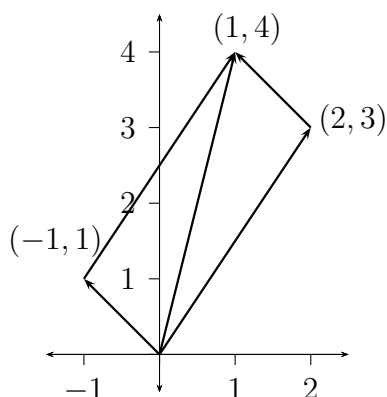


Abbildung 1.6: Addition von Vektoren

$A$  ist dabei das gleiche  $n$ -Tupel wie  $a$ . Ob ein  $n$ -Tupel als Ortsvektor oder Verschiebungsvektor gedeutet werden soll, erschließt sich nur aus dem Zusammenhang.

## 1.4 Addition und Multiplikation im $\mathbb{R}^n$

Wir betrachten die Vektoren  $a = (2, 3)$  und  $b = (-1, 1)$  und deuten sie als Verschiebungen in der Ebene. Verschiebt man zunächst längs  $a$  und dann längs  $b$ , entspricht das einer Verschiebung um den Vektor  $c = (1, 4)$  (Abb. 1.6), der komponentenweise der Summe der Komponenten von  $a$  und  $b$  entspricht. Ein analoger Zusammenhang gilt auch im  $\mathbb{R}^3$ . Weil man also die Verschiebung durch den Vektor  $b$  der Verschiebung um den Vektor  $a$  hinzufügt, liegt es nahe, folgende Vektoraddition zu definieren.

**Definition 1.6:** Seien  $a = (a_1, \dots, a_n)^T, b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ . Dann ist

$$a + b := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}.$$

Wird der Vektor  $a = (2, 1)$  in der Ebene um den Faktor 3 verlängert, erhält man den Vektor  $(6, 3)$  (Abb. 1.7). Dies entspricht genau einer Skalierung der Komponenten von  $a$  mit dem Faktor 3 und motiviert die folgende Definition.

**Definition 1.7:** Für  $\lambda \in \mathbb{R}$  und  $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$  sei

$$\lambda a := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}.$$

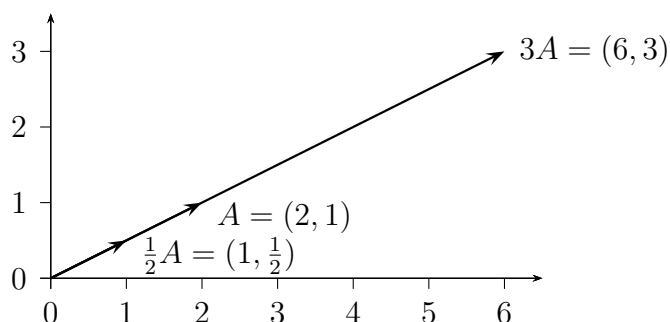


Abbildung 1.7: Multiplikation eines Vektors mit einem Skalar

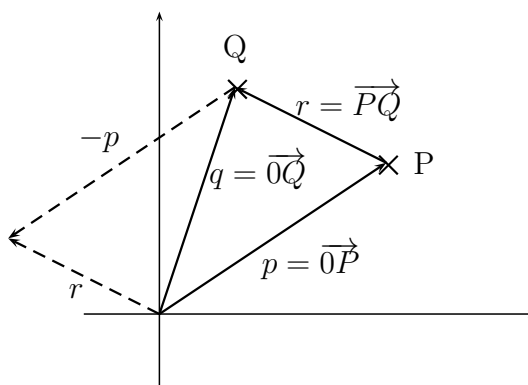


Abbildung 1.8: Differenz zweier Vektoren

Für  $0 < \lambda \in \mathbb{R}$  entspricht also  $\lambda a$  einem Punkt mit derselben Richtung wie  $a$  zum Ursprung, jedoch mit einem  $\lambda$ -fachen Abstand. Anhand einer Skizze erkennt man, dass die Multiplikation von  $a$  mit einer negativen Zahl die Richtung von  $a$  umkehrt.

**Definition 1.8:** Für  $a, b \in \mathbb{R}^n$  sei

$$a - b := a + (-1)b$$

der *Differenzvektor* von  $a$  und  $b$ .

Abb. 1.8 veranschaulicht folgenden Zusammenhang:

$$\begin{aligned} q = \vec{OQ} &= \vec{OP} + \vec{PQ} = p + \vec{PQ} \\ \Leftrightarrow \vec{PQ} &= q - p \end{aligned}$$

Interpretiert man die Vektoren  $p$  und  $q$  als Punkte, so gibt die Differenz der beiden Vektoren den Verbindungsvektor der Punkte an. Da die Subtraktion

nicht kommutativ ist, gibt  $q - p$  den Verbindungsvektor von  $P$  nach  $Q$  an, während  $p - q$  den Verbindungsvektor von  $Q$  nach  $P$  angibt und somit genau der um  $180^\circ$  gedrehte Vektor ist.

**Definition 1.9:** Zwei Vektoren  $a, b \neq 0$  heißen *parallel* (Schreibweise  $a \parallel b$ )  $:\Leftrightarrow \exists \alpha \in \mathbb{R} : a = \alpha b$ . Gilt  $\alpha > 0$ , haben parallele Vektoren die *gleiche Richtung*, im Fall von  $\alpha < 0$  *gegensätzliche Richtung*. Der Nullvektor ist zu jedem Vektor parallel.

**Beispiel 1.10:** Die Vektoren  $a = (1, 2)$  und  $b = (2, 4)$  sind parallel und haben die gleiche Richtung; der Vektor  $c = (1, 3)$  ist weder zu  $a$  noch zu  $b$  parallel.

## 1.5 Lineare Gleichungssysteme und Matrizen

In der Einleitung wurde die zentrale Rolle Linearer Gleichungssysteme betont. Daher beginnen wir hier mit ihrer systematischen Behandlung. Wir betrachten erneut das Lineare Gleichungssystem aus Beispiel 1.1,

$$\begin{aligned} 0,8a + 0,8b + 0,8c &= 0,8 \\ 0a + 0,2b + 0,1c &= 0,12 \\ 0,2a + 0b + 0,1c &= 0,08 \end{aligned} \quad (1.2)$$

Offenkundig hängt der Wert der Lösung nicht von der Bezeichnung der Variablen (hier:  $a, b$  und  $c$ ) ab, sondern vielmehr von den Vorfaktoren vor den Variablen im Linearen Gleichungssystem, den *Koeffizienten*. Zusammen mit der rechten Seite enthalten sie alle relevanten Informationen. Man kann somit ein Gleichungssystem kompakt schreiben, indem man nur die Koeffizienten in Form einer *Koeffizientenmatrix* schreibt. Zu dem Linearen Gleichungssystem (1.2) gehört die Koeffizientenmatrix

$$\begin{pmatrix} 0,8 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 \\ 0,2 & 0 & 0,1 \end{pmatrix}.$$

Wir werden also bei der Untersuchung von Linearen Gleichungssystemen auf Matrizen geführt.

**Definition 1.11:** Eine  $m \times n$ -Matrix  $A$  ist ein rechteckiges Schema von reellen oder komplexen Zahlen (den *Elementen* der Matrix) mit  $m$  Zeilen und  $n$  Spalten,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \quad (1.3)$$

Die Matrixelemente werden mit  $a_{ij}$  bezeichnet, wobei  $i$  der Zeilenindex und  $j$  der Spaltenindex ist. Die Matrix selbst wird manchmal anstatt mit einem Großbuchstaben auch mit  $(a_{ij})$  bzw.  $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  bezeichnet. Die Menge aller  $m \times n$ -Matrizen mit reellen bzw. komplexen Elementen wird mit  $\mathbb{R}^{m \times n}$  bzw.  $\mathbb{C}^{m \times n}$  bezeichnet. Soll nicht zwischen  $\mathbb{R}$  und  $\mathbb{C}$  unterschieden werden, schreiben wir  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ .

**Definition 1.12:**

1. Seien  $A = (a_{ij}), B = (b_{ij}) \in \mathbb{K}^{m \times n}$ . Man setzt  $A = B \Leftrightarrow a_{ij} = b_{ij} \forall 1 \leq i \leq m, 1 \leq j \leq n$ .
2. Eine Matrix, deren Elemente alle den Wert 0 annehmen, heißt *Nullmatrix*.
3. Zu  $A$  wie oben sei

$$A^T := (a_{ji}) \in \mathbb{K}^{n \times m}$$

die *transponierte Matrix*. Man erhält also  $A^T$ , indem man die Spalten von  $A$  als Zeilen von  $A^T$  verwendet.

Die Koeffizientenmatrix des allgemeinen Linearen Gleichungssystems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots + \dots + \dots + \dots &= \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{1.4}$$

von  $m$  linearen Gleichungen und  $n$  Unbekannten ist genau die Matrix  $A$  aus Formel (1.3). Sie allein beschreibt ein Lineares Gleichungssystem aber nicht vollständig, weil die rechte Seite nicht vorkommt. Wird  $A$  rechts um eine Spalte  $b$  ergänzt, die die Werte der rechten Seite enthält, gelangt man zur erweiterten Koeffizientenmatrix  $(A, b)$ . Um die Sonderrolle der letzten Spalte zu betonen, wird sie oftmals durch einen senkrechten Strich von  $A$  getrennt.

**Beispiel 1.13:** Die erweiterte Koeffizientenmatrix des Linearen Gleichungssystems aus Beispiel 1.1 lautet

$$\left( \begin{array}{cccc} 0,8 & 0,8 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 & 0,12 \\ 0,2 & 0 & 0,1 & 0,08 \end{array} \right) \text{ bzw. } \left( \begin{array}{ccc|c} 0,8 & 0,8 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 & 0,12 \\ 0,2 & 0 & 0,1 & 0,08 \end{array} \right).$$

Als besonders interessant werden sich Matrizen erweisen, die gleich viele Zeilen und Spalten haben.

**Definition 1.14:**



1. Eine  $n \times n$ -Matrix heißt *quadratisch*.
2. Eine quadratische Matrix  $A$  mit Elementen  $a_{ij}$  wird häufig in der Form  $A = (a_{ij})_{i,j=1}^n$  geschrieben.
3. Sei  $A = (a_{ij})_{i,j=1}^n$  quadratisch. Die Elemente von  $A$  mit  $i = j$  bilden die *Hauptdiagonale* von  $A$ .
4. Eine quadratische Matrix, bei der alle Elemente oberhalb der Hauptdiagonalen gleich 0 sind, heißt *untere Dreiecksmatrix*.
5. Eine quadratische Matrix, bei der alle Elemente unterhalb der Hauptdiagonalen gleich 0 sind, heißt *obere Dreiecksmatrix*.

**Beispiel 1.15:** Die Matrix

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 2 & 9 & 3 \end{pmatrix}$$

ist eine untere Dreiecksmatrix.

**Definition 1.16:**

1. Sei  $A \in \mathbb{K}^{n \times n}$ . Gilt  $a_{ij} = 0$  für alle  $i \neq j$ , heißt  $A$  *Diagonalmatrix*.
2. Die Diagonalmatrix

$$E_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{K}^{n \times n}$$

heißt *Einheitsmatrix*.

**Bemerkung 1.17:** Die  $m$  Werte der rechten Seite eines reellen linearen Gleichungssystems (vgl. Formel (1.4)) lassen sich zu einem Vektor

$$b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$$

zusammenfassen und als Punkt im  $\mathbb{R}^m$  deuten, ebenso bilden die  $n$  Komponenten der Lösung einen Vektor  $x \in \mathbb{R}^n$ .

## 1.6 Lösung eines linearen Gleichungssystems

### 1.6.1 Das Gaußsche Eliminationsverfahren

Wir stellen nachfolgend ein Standard-Verfahren zum Lösen von Linearen Gleichungssystemen vor, das *Gaußsche Eliminationsverfahren*, das ein gegebenes Gleichungssystem in ein anderes überführt, dessen Lösung unmittelbar ablesbar ist. Diese Lösung ist zugleich die Lösung des ursprünglich gegebenen Linearen Gleichungssystems.

**Definition 1.18:**

1. Zwei Gleichungssysteme heißen *äquivalent*, falls sie die gleiche Lösungsmenge haben.
2. Zwei Matrizen  $A$  und  $B$  heißen äquivalent ( $A \sim B$ ), falls die entsprechenden Gleichungssysteme die gleiche Lösungsmenge besitzen.

**Bemerkung 1.19:** Äquivalenz im Sinne von Definition 1.18 ist eine *Äquivalenzrelation* (vgl. die Vorlesung "Mathematische Grundlagen").

**Beispiel 1.20:** Das Gleichungssystem

$$\begin{aligned} 2x + 3y - 4z &= 8 \\ -4x + 2y + 3z &= -5 \\ 3x + y + 2z &= 13 \end{aligned} \quad (a)$$

ist äquivalent zu den Gleichungssystemen

$$\begin{aligned} 2x + 3y - 4z &= 8 \\ 8y - 5z &= 11 \\ z &= 1 \end{aligned} \quad (b)$$

und

$$\begin{aligned} x &= 3 \\ y &= 2 \\ z &= 1 \end{aligned} \quad (c)$$

Auch die drei zugehörigen erweiterten Matrizen  $A, B, C$  sind dann äquivalent:

$$A = \begin{pmatrix} 2 & 3 & -4 & 8 \\ 3 & 1 & 2 & 13 \\ -4 & 2 & 3 & -5 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 & -4 & 8 \\ 0 & 8 & -5 & 11 \\ 0 & 0 & 1 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Gleichungssystem (b) heißt *Stufenform*. Gleichungssystem (c) heißt *reduzierte Stufenform*. In vorliegendem Fall hat die zugehörige Matrix  $B$  unterhalb der Hauptdiagonalen nur Nullen (da sie nicht quadratisch ist, ist sie keine

Dreiecksmatrix). Die Matrix  $C$  hat Einsen auf der Hauptdiagonalen und Nullen darunter und darüber. Nur in der letzten Spalte stehen beliebige Werte.

Es ist leicht zu sehen, dass die Stufenform (Ergebnis des Gauß-Algorithmus) und insbesondere die reduzierte Stufenform (Ergebnis des Gauß-Jordan-Algorithmus) übersichtlicher und die Bestimmung der Lösung deutlich einfacher sind. Bei der erweiterten Stufenform kann man die Lösung sogar einfach ablesen. Darum werden wir versuchen, durch Umformungen ein äquivalentes Gleichungssystem zu bestimmen, das die reduzierte Stufenform besitzt. Solche Umformungen nennt man *Äquivalenzumformungen* der Zeilen. Es gibt davon drei Stück:

**Z1.** Addition eines Vielfachen einer Zeile zu einer anderen.

**Z2.** Vertauschen zweier Zeilen.

**Z3.** Multiplikation<sup>4</sup> einer Zeile mit einem Skalar  $\lambda \neq 0, \lambda \in \mathbb{K}$ .

Um auf die reduzierte Stufenform zu gelangen, muss man zunächst die Stufenform errechnen. Dies erreicht man, indem man (hauptsächlich durch die Operation Z1) die Einträge unterhalb der Hauptdiagonalen auf 0 bringt. Man beginnt in der ersten Spalte und arbeitet sich spaltenweise vor:

$$\begin{array}{r|l}
 2 & 3 & -4 & 8 \\
 -4 & 2 & 3 & -5 & | + 2 \cdot (I) \\
 3 & 1 & 2 & 13 & | - \frac{3}{2} \cdot (I) \\
 \hline
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & -\frac{7}{2} & 8 & 1 & | + \frac{7}{16} \cdot (II) \\
 \hline
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & 0 & \frac{93}{16} & \frac{93}{16}
 \end{array}$$

Die Stufenform ist nicht eindeutig. Es gibt mehrere mögliche (und äquivalente) Stufenformen. Hat man eine Stufenform erreicht, geht man zeilenweise von unten nach oben vor. Zunächst bringt man das Element auf der Hauptdiagonalen auf 1, anschließend alle Elemente darüber auf 0. In

<sup>4</sup>Die dritte Umformung ist für den Gauß-Algorithmus nicht nötig, aber insbesondere für Handrechnungen vorteilhaft.

unserem Beispiel sieht das so aus:

$$\begin{array}{ccc|c}
 2 & 3 & -4 & 8 \\
 0 & 8 & -5 & 11 \\
 0 & 0 & \frac{93}{16} & \frac{93}{16} & | \div \frac{93}{16} \\
 \hline
 2 & -4 & 8 & 8 & | + 4 \cdot (III) \\
 0 & 8 & -5 & 11 & | + 5 \cdot (III) \\
 0 & 0 & 1 & 1 & \\
 \hline
 2 & 3 & 0 & 12 & \\
 0 & 8 & 0 & 16 & | \div 8 \\
 0 & 0 & 1 & 1 & \\
 \hline
 2 & 3 & 0 & 12 & | - 3 \cdot (II) \\
 0 & 1 & 0 & 2 & \\
 0 & 0 & 1 & 1 & \\
 \hline
 2 & 0 & 0 & 6 & | \div 2 \\
 0 & 1 & 0 & 2 & \\
 0 & 0 & 1 & 1 & \\
 \hline
 1 & 0 & 0 & 3 & \\
 0 & 1 & 0 & 2 & \\
 0 & 0 & 1 & 1 & 
 \end{array}$$

Transferiert man das Schema wieder in das eigentliche Gleichungssystem zurück, erhält man die Lösung:

$$\begin{aligned}
 x &= 3 \\
 y &= 2 \\
 z &= 1
 \end{aligned}$$

Dass das Gaußsche Eliminationsverfahren tatsächlich zu äquivalenten linearen Gleichungssystemen führt, wird (mit etwas mehr Theorie) später in Kapitel ?? bewiesen werden.

In vielen Fällen werden die Zahlen im Laufe der Berechnung der Lösung eines linearen Gleichungssystems schnell unhandlich. Dies macht eine Berechnung per Hand aufwändig und fehleranfällig. Daher werden in der Praxis lineare Gleichungssysteme häufig nicht durch Handrechnung gelöst, sondern mithilfe des Computers. Dazu werden oft spezielle Programmpakete verwendet, die primär auf Berechnungen mit Hilfe von Matrizen ausgelegt ist, oftmals aber einen weitaus größeren Funktionsumfang bieten. In der Industrie weit verbreitet ist das kommerzielle Programmpaket MATLAB ([www.mathworks.com](http://www.mathworks.com)), als Alternativen zu MATLAB existieren z. B. die Open-Source-Programme Scilab (<http://www.scilab.org>) oder Octave (<http://www.gnu.org/software/octave/>). Die drei genannten Programmpakete gleichen sich hinsichtlich ihrer Bedienung weitgehend,

so dass wir uns hier prototypisch auf MATLAB beschränken.

In MATLAB kann man interaktiv Anweisungen ausführen lassen. Dazu schreibt man hinter das Prompt-Zeichen `>>` die Eingabezeile. MATLAB gibt dann darunter das Ergebnis der Eingabezeile aus. Die Ausgabe des Ergebnisses kann man unterdrücken, wenn man die Eingabezeile mit einem Semikolon abschließt.

Eine Matrix wird in MATLAB wie folgt angelegt:

```
>> a=[1 2 3;4 5 6;7 8 9]
```

```
a =
```

```
1     2     3
4     5     6
7     8     9
```

Die Eingabezeile ist die oberste Zeile, alle unteren Zeilen sind die Ergebnisausgabe von MATLAB. Zeilen werden durch ein Semikolon getrennt. Es gibt in MATLAB mehrere Arten, lineare Gleichungssysteme zu lösen. Wir verwenden den Befehl `rref` (`rref` ist die Abkürzung von *reduced row echelon form*, der englischen Übersetzung von *reduzierte Stufenform*) und betrachten das untenstehende Beispiel. Die erste Beispiel-Zeile `format rat` veranlasst MATLAB, sein Ergebnis in Form von Brüchen auszugeben, was bei unseren kleinen Beispielaufgaben gut gelingt. MATLAB wie auch Scilab und Octave rechnet intern aber immer mit Fließkommazahlen. In der zweiten Zeile geben wir die erweiterte Matrix zum obigen Beispiel ein, die anschließend in die reduzierte Stufenform umgerechnet wird.

```
>> format rat
```

```
>> a=[2 3 -4 8;-4 2 3 -5;3 1 2 13]
```

```
a =
```

```
2     3     -4     8
-4    2     3    -5
3     1     2    13
```

```
>> rref(a)
```

```
ans =
```

```
1     0     0     3
0     1     0     2
0     0     1     1
```

Die Rückübersetzung des Resultats des `rref`-Kommandos liefert die Lösung

$$x = 3$$

$$y = 2$$

$$z = 1$$

### 1.6.2 Unter- und überbestimmte Gleichungssysteme

Im letzten Kapitel hatten wir uns auf Lineare Gleichungssystem beschränkt, die genauso viele Unbekannte wie Gleichungen aufweisen. Unter gewissen Zusatzvoraussetzungen existiert dann genau eine Lösung, und in diesem Fall ist diese immer mit dem Gauß-Algorithmus berechenbar. Wir betrachten jetzt allgemeine Lineare Gleichungssysteme. Dann sind drei Fälle möglich:

1. Es existiert eine eindeutige Lösung.
2. Es existiert keine Lösung.
3. Es existieren unendlich viele Lösungen.

Im zweiten Fall spricht man von *überbestimmten Linearen Gleichungssystemen*, im dritten Fall von *unterbestimmten Linearen Gleichungssystemen*. Wir werden dieses Phänomen hier nicht systematisch untersuchen, weil die reduzierte Stufenform als Ergebnis des Gaußschen Eliminationsverfahrens immer zeigt, welcher Fall vorliegt. Die folgende Definition vereinfacht die Analyse der reduzierten Stufenform.

**Definition 1.21:** Der erste Nicht-Null-Eintrag jeder Zeile einer Matrix heißt *Pivot-Element*. Eine Spalte, in der ein Pivot-Element vorkommt, heißt *Pivot-Spalte*.

In der reduzierten Stufenform ist das Pivot-Element immer eine 1. Die reduzierte Stufenform ist bisher im allgemeinen Fall nicht eindeutig festgelegt. Daher fordern wir von der reduzierten Stufenform folgende Eigenschaften:

- Es dürfen Null-Zeilen vorkommen, also Zeilen, in denen jeder Eintrag 0 ist. Diese müssen die untersten Zeilen der Matrix sein. Null-Zeilen stehen für die Gleichung  $0 = 0$  und können ignoriert werden.
- Die Pivot-Spalte der Zeile  $i + 1$  muss rechts der Pivot-Spalte der Zeile  $i$  liegen.
- Pivot-Spalten enthalten in der reduzierten Stufenform außer dem Pivot-Element selbst nur Nullen. Diese Bedingung gilt nicht für die letzte Spalte der erweiterten Matrix.

Diese Forderungen lassen sich für jede Matrix durch die Äquivalenzumformungen Z1 - Z3 erfüllen.

**Beispiel 1.22:** Wir bezeichnen ein Pivot-Element symbolisch mit  $P$  und ein beliebiges Element mit  $x$ . Die Matrix

$$\begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \end{pmatrix}$$

liegt in reduzierter Stufenform vor. Alle Spalten mit Ausnahme der letzten sind Pivot-Spalten. Ebenso liegen folgende Matrizen in reduzierter Stufenform vor:

$$\begin{pmatrix} P & 0 & x & 0 & x \\ 0 & P & x & 0 & x \\ 0 & 0 & 0 & P & x \\ 0 & 0 & 0 & 0 & P \end{pmatrix} \begin{pmatrix} P & 0 & x & x & 0 & x & x & x \\ 0 & P & x & x & 0 & x & x & x \\ 0 & 0 & 0 & 0 & P & x & x & x \end{pmatrix} \begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wir nennen im Folgenden die Spalten 1 bis  $n - 1$  der erweiterten Matrix *Koeffizientenmatrix* und die  $n$ -te Spalte *letzte Spalte*. Hat das Gleichungssystem eine eindeutige Lösung, so erhält man immer die Stufenform bzw. die reduzierte Stufenform

$$\left( \begin{array}{ccc|c} P & x & x & x \\ 0 & P & x & x \\ 0 & 0 & P & x \end{array} \right) \quad \text{bzw.} \quad \left( \begin{array}{ccc|c} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \end{array} \right).$$

Falls die letzte Spalte der Stufenform eine Pivot-Spalte ist, ist das Gleichungssystem *nicht lösbar*. Man könnte also z. B. die folgende Stufenform bzw. reduzierte Stufenform erhalten:

$$\left( \begin{array}{cccc|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array} \right) \quad \text{bzw.} \quad \left( \begin{array}{cccc|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array} \right).$$

**Beispiel 1.23:** Die erweiterte Matrix sei

$$\left( \begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right).$$

Die letzte Zeile lautet ausgeschrieben

$$0 \cdot x + 0 \cdot y + 0 \cdot z = 1$$

Diese Gleichung ist unerfüllbar, gleichgültig, welche Werte  $x, y$  und  $z$  annehmen. Daraus folgt, dass das Gleichungssystem nicht lösbar ist.

Erhält man in der (reduzierten) Stufenform weniger Pivotspalten als das Gleichungssystem Unbekannte hat, so hat das Gleichungssystem mehrere Lösungen. Mögliche erweiterte Matrizen sind in diesem Fall z. B.

$$\left( \begin{array}{cccc|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \text{bzw.} \quad \left( \begin{array}{cccc|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Es gibt (mindestens) einen *freien Parameter*, d.h. eine der Variablen kann einen beliebigen Wert annehmen. Befolgt man eine einfache Lösungsregel, dann gehört zu jeder Nicht-Pivot-Spalte der Koeffizientenmatrix eine freie Variable.

**Beispiel 1.24:** Zum Gleichungssystem

$$\begin{aligned} 2 \cdot x - 4 \cdot y + 2 \cdot z &= 8 \\ 1 \cdot x - 1 \cdot y - 7 \cdot z &= 6 \end{aligned}$$

berechnen wir die äquivalente reduzierte Stufenform. Der Bequemlichkeit halber benutzen wir MATLAB:

```
>> format rat
>> a=[2 -4 2 8;1 -1 -7 6];
>> rref(a)
```

ans =

$$\begin{array}{cccc} 1 & 0 & -15 & 8 \\ 0 & 1 & -8 & 2 \end{array}$$

Das Ergebnis bedeutet als Gleichungssystem:

$$\begin{aligned} 1 \cdot x - 0 \cdot y - 15 \cdot z &= 8 \\ 0 \cdot x + 1 \cdot y - 8 \cdot z &= 2 \end{aligned}$$

oder

$$\begin{aligned} x &= 8 + 15 \cdot z \\ y &= 2 + 8 \cdot z \end{aligned}$$

oder in Vektor-Schreibweise:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \end{pmatrix}$$

Diese Lösung ist noch nicht vollständig, denn die vollständige Lösung besteht aus dem Tripel  $(x, y, z)$ . Den Wert von  $z$  können wir aber sofort angeben, denn  $z = 0 + 1 \cdot z$ . Hängen wir die entsprechende Zeile an, so ergibt sich als Endergebnis:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \\ 1 \end{pmatrix}$$



Zusammengefasst ergibt sich bei nicht eindeutigen Gleichungssystemen folgende Vorgehensweise:

- Die Nicht-Pivot-Spalten werden mit der zugehörigen Variablen versehen auf die rechte Seite gebracht, wobei sich alle Vorzeichen umdrehen.
- Zeilen für fehlende Variablen werden in den Lösungsvektor nach dem Schema  $x = x$  eingefügt.

**Beispiel 1.25:** Man berechne die Lösung des Gleichungssystems

$$\begin{aligned} 2a + b - 2c &= -2 \\ -4a - 2b + c + 2d &= 2 \\ -2a - b + d &= 1 \end{aligned}$$

Man erhält die Lösung

```
>> a=[2 1 -2 0 -2;-4 -2 1 2 2;-2 -1 0 1 1];
>> rref(a)
```

ans =

$$\begin{array}{cccccc} 1 & & 1/2 & & 0 & & -1 \\ 0 & & 0 & & 1 & & 0 \\ 0 & & 0 & & 0 & & 1 \end{array}$$

Die zweite Spalte der Matrix ist keine Pivot-Spalte. Das ergibt als Zwischenergebnis

$$\begin{pmatrix} a \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix}$$

und als Endergebnis

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$



## Kapitel 2

# Analytische Geometrie

### 2.1 Skalarprodukt und Norm

Wir werden nun geometrische Konzepte wie die Länge eines Vektors und den Winkel zwischen zwei Vektoren auf den abstrakten Raum  $\mathbb{R}^n$  übertragen. Das wird selbstverständlich so geschehen, dass sich in den Spezialfällen  $\mathbb{R}^2$  und  $\mathbb{R}^3$  die intuitiv bekannten Längen- und Winkelbegriffe ergeben. Entscheidend hierfür ist das Konzept des *Skalarprodukts*, dessen konkrete geometrische Bedeutung wir zunächst nicht betrachten. Vorbereitend definieren wir das *kartesische Produkt* von Mengen.

**Definition 2.1:** Seien  $X_1, \dots, X_n$  mit  $n \in \mathbb{N}$  nichtleere Mengen. Dann versteht man unter dem *kartesischen Produkt*  $X_1 \times \dots \times X_n$  die Menge aller geordneten  $n$ -Tupel:

$$X_1 \times \dots \times X_n := \{(x_1, \dots, x_n)^T \mid x_i \in X_i\}.$$

Für  $X_1 = \dots = X_n = X$  schreibt man kurz  $X^n$  statt  $X_1 \times \dots \times X_n$ .

**Beispiel 2.2:** Offenbar gilt  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$  und allgemeiner  $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$  durch die Wahl von  $X_1 = \dots = X_n = \mathbb{R}$  in Definition 2.1.

**Bemerkung 2.3:**

1. Nach Beispiel 2.2 verallgemeinert das kartesische Produkt die Definition 1.4 des Zahlraums  $\mathbb{R}^n$ .
2. Lässt man statt der *Indexmenge*  $\{1, \dots, n\}$  wie oben für  $i$  alle natürlichen Zahlen zu, wird man auf unendliche kartesische Produkte geführt. Elemente solcher Produktmengen werden *Folgen* genannt und in der Analysis eingehend untersucht.

Von nun an seien im ganzen Kapitel immer  $a = (a_i)_{i=1}^n, b = (b_i)_{i=1}^n \in \mathbb{R}^n$ .

**Definition 2.4:** Eine Abbildung  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  heißt *Skalarprodukt*, wenn folgende Bedingungen erfüllt sind:

**SP1** (Symmetrie):  $\forall a, b \in \mathbb{R}^n : \langle a, b \rangle = \langle b, a \rangle$ .

**SP2**  $\forall a, b, c \in \mathbb{R}^n :$

$$\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$$

**SP3**  $\forall \alpha \in \mathbb{R}$  gilt

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle = \langle a, \alpha b \rangle.$$

**SP4** (positive Definitheit):  $\forall a \in \mathbb{R}^n \setminus \{0\} : \langle a, a \rangle > 0$ , und  $\langle 0, 0 \rangle = 0$ .

Das bei Weitem wichtigste Skalarprodukt ist das *Standardskalarprodukt* oder *euklidische Skalarprodukt*. Wird im Folgenden von *dem* Skalarprodukt gesprochen, ist das euklidische Skalarprodukt gemeint.

**Definition 2.5:** Für  $a, b$  sei ihr *euklidisches Skalar-* bzw. *Punktprodukt*  $\langle a, b \rangle$  definiert als

$$\langle a, b \rangle := a_1 b_1 + \cdots + a_n b_n = \sum_{i=1}^n a_i b_i$$

**Bemerkung 2.6:** Oft wird das euklidische Skalarprodukt als  $a \cdot b$  geschrieben. In der Mathematik wird jedoch am häufigsten  $\langle a, b \rangle$  verwendet. Eine spezielle Schreibweise ist  $a^2 = \langle a, a \rangle$ . Höhere Potenzen von Vektoren wie z. B.  $a^3$  sind nicht definiert.

**Beispiel 2.7:** Für

$$a = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \text{ und } b = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix}$$

gilt

$$\langle a, b \rangle = 1 \cdot (-1) + 3 \cdot 4 + (-2) \cdot (-3) = 17.$$

Aus Gründen der Widerspruchsfreiheit formulieren und beweisen wir folgenden Satz.

**Satz 2.8:** Das euklidische Skalarprodukt ist ein Skalarprodukt im Sinne der Definition 2.4.

**BEWEIS** Seien  $a, b$  beliebig. Man hat  $\langle a, b \rangle = a_1 b_1 + \cdots + a_n b_n \in \mathbb{R}$ , da Produkte und Summen reeller Zahlen wieder reelle Zahlen sind.

**SP1:**

$$\langle a, b \rangle = \sum_{i=1}^n a_i b_i = \sum_{i=1}^n b_i a_i = \langle b, a \rangle,$$

weil es bei der Multiplikation zweier reeller Zahlen nicht auf ihre Reihenfolge ankommt.

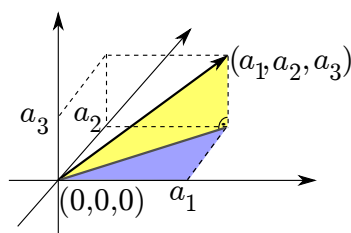


Abbildung 2.1: Berechnung der euklidischen Länge eines Vektors im  $\mathbb{R}^3$  durch doppelte Anwendungen des Satzes von Pythagoras

**SP2:** Sei nun  $c = (c_i)_{i=1}^n \in \mathbb{R}^n$  beliebig. Dann gilt  $b + c = (b_i + c_i)_{i=1}^n$  und deswegen mit dem Distributivgesetz der reellen Zahlen

$$\begin{aligned} \langle a, b + c \rangle &= \sum_{i=1}^n a_i(b_i + c_i) = \sum_{i=1}^n (a_i b_i + a_i c_i) \\ &= \sum_{i=1}^n a_i b_i + \sum_{i=1}^n a_i c_i = \langle a, b \rangle + \langle a, c \rangle \end{aligned}$$

**SP3:** Für  $\alpha \in \mathbb{R}$  gilt

$$\langle \alpha a, b \rangle = \sum_{i=1}^n (\alpha a_i) b_i = \sum_{i=1}^n \alpha (a_i b_i) = \alpha \sum_{i=1}^n a_i b_i = \alpha \langle a, b \rangle$$

**SP4:** Offenbar gilt  $\langle 0, 0 \rangle = 0$ . Sei  $a \neq 0 \in \mathbb{R}^n$ . Dann ist mindestens eine Koordinate  $a_j \neq 0$ . Also gilt wegen  $a_j^2 > 0$

$$\langle a, a \rangle = a_1^2 + \dots + a_n^2 \geq a_j^2 > 0. \quad \blacksquare$$

Wir betrachten nun einen Vektor  $a = (a_1, a_2)^T \in \mathbb{R}^2$ . Nach dem Satz des Pythagoras beträgt seine Länge  $\sqrt{a_1^2 + a_2^2}$ . Analog findet man durch zweimalige Anwendung des Satzes von Pythagoras für die Länge eines Vektors  $b = (b_1, b_2, b_3)^T \in \mathbb{R}^3$  den Ausdruck  $\sqrt{b_1^2 + b_2^2 + b_3^2}$  (Abb. 2.1). Es liegt daher nahe, diesen intuitiven Längenbegriff auf den  $\mathbb{R}^n$  zu verallgemeinern.

**Definition 2.9:** Einem Vektor  $a$  wird die *euklidische Norm* oder *Standardnorm*  $\|a\|$  zugeordnet durch

$$\|a\| := \left( \sum_{i=1}^n a_i^2 \right)^{1/2}. \quad (2.1)$$

**Satz 2.10:** Die in Gleichung (2.1) definierte Norm hat folgende Eigenschaften:

$$\mathbf{N0} : \|a\| \in \mathbb{R}.$$

$$\mathbf{N1} : \|a\| \geq 0.$$

$$\mathbf{N2} : \|a\| = 0 \Leftrightarrow a = 0.$$

$$\mathbf{N3} : \forall \lambda \in \mathbb{R} : \|\lambda a\| = |\lambda| \|a\|$$

$$\mathbf{N4} : (\text{Dreiecksungleichung}) \|a + b\| \leq \|a\| + \|b\|$$

**BEWEIS**

**N0 :** Wegen  $\sum_{i=1}^n a_i^2 \geq 0$  ist die Wurzel dieses Ausdrucks reell.

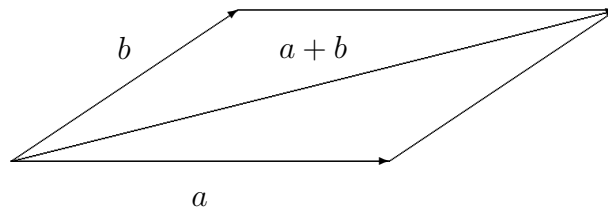
**N1 :**  $\|a\|$  ist die positive Wurzel einer reellen Zahl.

**N2 :**  $\|a\| = 0 \Leftrightarrow \|a\|^2 = a_1^2 + \dots + a_n^2 = 0 \Leftrightarrow a_i^2 = 0 \forall 1 \leq i \leq n \Leftrightarrow a = 0$ , da Quadrate reeller Zahlen nicht negativ sind.

$$\mathbf{N3} : \|\lambda a\| = \sqrt{(\lambda^2 a_1^2) + \dots + (\lambda^2 a_n^2)} = \sqrt{\lambda^2} \|a\| = |\lambda| \|a\|.$$

**N4 :** später ■

**Bemerkung 2.11:** Die untenstehende Abbildung illustriert, warum die Eigenschaft **N4 Dreiecksungleichung** genannt wird: Die Länge einer Seite entspricht höchstens der Summe aus den Längen der beiden anderen Seiten.



Prägend für die Mathematik ist das Streben nach Abstraktion und Allgemeinheit. So haben wir bereits durch Verallgemeinerung auf  $\mathbb{R}^n$  einen Längenbegriff definiert. Wir werden nun das Konzept der Länge erneut verallgemeinern. Dazu deuten wir die Zuordnung einer Länge zu einem Vektor als Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}$ .

**Definition 2.12:** Eine Abbildung  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  heißt *Norm*, wenn sie die Eigenschaften **N0** bis **N4** aus Satz 2.10 besitzt.

Man sieht sofort, dass sich die euklidische Norm nach Definition 2.9 als Wurzel des euklidischen Skalarprodukts eines Vektors mit sich selbst schreiben lässt. Man sagt, *das euklidische Skalarprodukt induziert die euklidische Norm*. Dieser Zusammenhang gilt für jedes Skalarprodukt.

**Satz 2.13:** Sei  $\langle \cdot, \cdot \rangle$  ein beliebiges Skalarprodukt auf  $\mathbb{R}^n$ . Dann wird durch  $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$  eine Norm auf  $\mathbb{R}^n$  induziert.

**BEWEIS** Da  $\langle \cdot, \cdot \rangle \geq 0$ , existiert eine reelle positive Wurzel, und  $\|\cdot\|$  ist in der Tat eine Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}$ . N1 gilt aus demselben Grund.

Offenbar gilt  $\|0\| = \sqrt{\langle 0, 0 \rangle} = 0$  nach SP4. Sei also  $\|a\| = 0$ . Dann gilt auch  $\|a\|^2 = 0 = \langle 0, 0 \rangle$  und damit nach SP4  $a = 0$ , also gilt N2.

Wegen  $\|\lambda a\| = \sqrt{\langle \lambda a, \lambda a \rangle} = \sqrt{\lambda^2} \sqrt{\langle a, a \rangle} = |\lambda| \|a\|$  aufgrund von SP3 gilt auch N3.

N4 beweisen wir später. ■

Es existieren viele Normen, d. h. Längenbegriffe auf dem  $\mathbb{R}^n$ . Wir werden künftig mit „der Norm“ immer die euklidische Norm meinen, mit „einer Norm“ jedoch eine unbestimmte Norm. Es folgen einige gebräuchliche Normen des  $\mathbb{R}^n$ .

- Auf  $\mathbb{R}$  ist der Absolutbetrag eine Norm.
- Für  $p \geq 1$  definiert man die  $l_p$ -Norm durch

$$\|a\|_p := \left( \sum_{i=1}^n |a_i|^p \right)^{1/p} \quad (2.2)$$

Die Dreiecksungleichung der  $l_p$ -Norm ist in der Analysis als *diskrete Minkowskische Ungleichung* bekannt. Für ihren nichttrivialen Beweis verweisen wir auf einschlägige Literatur zur Analysis, z. B. [4]. Der Ausdruck (2.2) ist zwar auch für  $0 < p < 1$  definiert, dann aber gilt die Dreiecksungleichung i. A. nicht mehr.

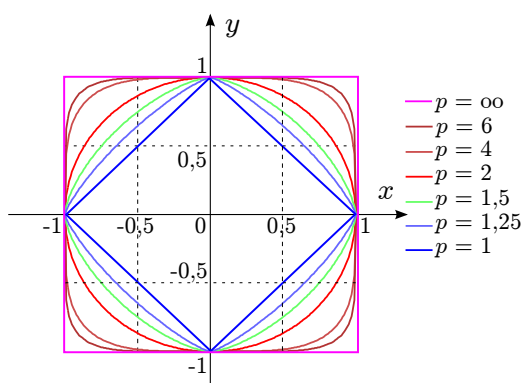
- Der Spezialfall  $p = 1$  wird als *Betragssummennorm* oder *Einernorm* bezeichnet und ist definiert als

$$\|a\|_1 = |a_1| + \dots + |a_n|.$$

- Der Spezialfall  $p = 2$  entspricht der euklidischen Norm.
- Die *Maximumnorm* oder  $l_\infty$ -Norm ist definiert als

$$\|a\|_\infty = \max\{|a_1|, \dots, |a_n|\};$$

der Bezeichnung erklärt sich durch  $\lim_{p \rightarrow \infty} \|a\|_p = \|a\|_\infty$ .

Abbildung 2.2: Einheitskugeln für verschiedene  $l_p$ -Normen

**Beispiel 2.14:** Es sei  $b = (-1, 2, 3)^T$ . Dann ist

$$\begin{aligned}\|b\| &= \sqrt{1 + 4 + 9} = \sqrt{14} \\ \|b\|_1 &= 1 + 2 + 3 = 6 \\ \|b\|_\infty &= \max\{1, 2, 3\} = 3\end{aligned}$$

**Bemerkung 2.15:** Satz 2.13 besagt, dass jedes Skalarprodukt eine Norm induziert. Umgekehrt ist aber nicht jede Vektornorm von einem Skalarprodukt induziert. So z. B. ist *keine* der  $l_p$ -Normen von irgendeinem Skalarprodukt induziert bis auf den Spezialfall  $p = 2$ .

### 2.1.1 Einheitsvektoren

**Definition 2.16:** Sei  $\|\cdot\|$  eine Norm. Ein Vektor  $e \in \mathbb{R}^n$  heißt *Einheitsvektor* (zur Norm  $\|\cdot\|$ ), wenn  $\|e\| = 1$  ist.

**Bemerkung 2.17:** Ob ein gegebener Vektor ein Einheitsvektor ist, hängt entscheidend von der verwendeten Norm ab. Wir betrachten die *Einheitskugel*  $S^1$  im  $\mathbb{R}^2$ ,

$$S^1 := \{x \in \mathbb{R}^2 \mid \|x\| = 1\}.$$

Im Fall der euklidischen Norm handelt es sich bei  $S^1$  um den wohlbekannten Einheitskreis; für bestimmte  $p$ -Normen ergeben sich die in Abb. 2.2 gezeigten Mengen.

Wichtige Einheitsvektoren (bezogen auf die Standardnorm) sind die *kanonischen Einheitsvektoren*, also Einheitsvektoren in Richtung der Koordinaten-



achsen:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

bzw. im  $\mathbb{R}^3$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (2.3)$$

Jeder Vektor  $a \neq 0$  kann *normiert* werden, d. h. man findet zu jedem Vektor  $a$  einen zu  $a$  parallelen und gleichgerichteten Einheitsvektor  $e_a$  durch die Setzung

$$e_a = \frac{1}{\|a\|} a.$$

**Beispiel 2.18:** Sei  $a = (1, 2, -3)^T$ . Aufgrund von  $\|a\| = \sqrt{14}$  ist  $e_a = \frac{1}{\sqrt{14}}(1, 2, -3)^T$ .

**Bemerkung 2.19:** Die Vektoren  $v$  und  $\lambda v$  ( $\lambda > 0$ ) zeigen in die gleiche Richtung. Nach Normierung sind beide Vektoren gleich. Man kann beim Normieren also als ersten Schritt einen positiven Vorfaktor des Vektors streichen, um die Rechnung zu vereinfachen.

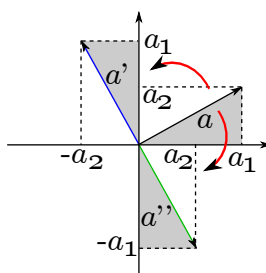
**Beispiel 2.20:** Möchte man den Vektor  $b = \frac{3}{\sqrt{8}}(7, 0, 2)^T$  normieren, kann man statt dessen den Vektor  $b' = (7, 0, 2)^T$  normieren und erhält bei einfacherer Rechnung das gleiche Ergebnis.

## 2.1.2 Orthogonale (senkrechte) Vektoren

**Bemerkung 2.21:** Wir gehen von der Ebene aus und betrachten den Vektor  $a = (a_1, a_2)^T$ . Anhand der Skizze in Abb. 2.3 erkennt man, dass für den um  $90^\circ$  gegen den Uhrzeigersinn gedrehte Vektor  $a'$  gilt:  $a' = (-a_2, a_1)^T$ , und der um  $90^\circ$  im Uhrzeigersinn gedrehte Vektor  $a''$  die Koordinaten  $(a_2, -a_1)^T$  besitzt. Es gilt also mit dem euklidischen Skalarprodukt

$$\langle a, a' \rangle = \langle a'', a \rangle = 0.$$

Umgekehrt wird man durch die Bedingung  $\langle a, b \rangle = 0$  über ein Lineares Gleichungssystem auf das Ergebnis  $b = \lambda(-a_2, a_1)^T$  mit einer beliebigen reellen Zahl  $\lambda$  geführt, so dass  $b$  offenbar parallel zu  $a'$  und  $a''$  verläuft und damit senkrecht auf  $a$  steht. Also steht  $a$  genau dann auf  $b$  senkrecht, wenn  $\langle a, b \rangle = 0$  gilt. Diese Betrachtungen erlauben es, den Begriff "senkrecht" in den allgemeinen  $\mathbb{R}^n$  zu übertragen.



Abbildungung 2.3: Drehung eines Vektors  $a$  um  $90^\circ$  im und entgegen dem Uhrzeigersinn

**Definition 2.22:** Seien  $a, b \in \mathbb{R}^n$  und  $\langle \cdot, \cdot \rangle$  ein beliebiges Skalarprodukt. Die Vektoren  $a$  und  $b$  stehen *senkrecht* (oder auch *orthogonal*) zueinander bzgl.  $\langle \cdot, \cdot \rangle$ , Schreibweise  $a \perp b$ , wenn  $\langle a, b \rangle = 0$  ist.

Für den Spezialfall  $n = 2$  und das euklidische Skalarprodukt entspricht diese Definition der intuitiven Vorstellung des Begriffes "senkrecht".

**Bemerkung 2.23:**

1. Die kanonischen Einheitsvektoren (2.3) stehen senkrecht zueinander im Sinne der Definition 2.22.
2. Ob zwei Vektoren aufeinander senkrecht stehen, hängt entscheidend von der Wahl des Skalarproduktes ab: Man kann zeigen, dass im  $\mathbb{R}^2$  durch  $\langle a, b \rangle' := a_1 b_1 + 2a_2 b_2$  ein Skalarprodukt definiert wird. Die Vektoren  $(-1, 1)^T$  und  $(1, 1)^T$  stehen bezogen auf das euklidische Skalarprodukt senkrecht aufeinander, nicht aber bezogen auf das Skalarprodukt  $\langle \cdot, \cdot \rangle'$ , da  $\langle a, b \rangle' = 1$ .

**Bemerkung 2.24:** Aus  $a \perp b$  folgt  $\alpha a \perp \beta b$  für beliebige reelle Zahlen  $\alpha, \beta \in \mathbb{R}$  wegen

$$\langle \alpha a, \beta b \rangle = \alpha \beta \langle a, b \rangle = 0.$$

Wir beweisen nun den Satz des Pythagoras.

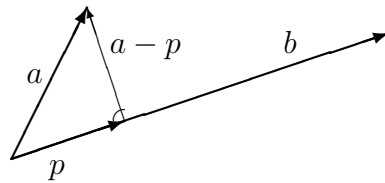
**Satz 2.25:** Seien  $a, b \in \mathbb{R}^n$  mit  $a \perp b$ ,  $\langle \cdot, \cdot \rangle$  ein beliebiges Skalarprodukt und  $\| \cdot \|$  die dadurch induzierte Norm. Dann gilt

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2.$$

**BEWEIS**

$$\|a + b\|^2 = \langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle$$

Mit  $\langle a, b \rangle = 0 = \langle b, a \rangle$  wegen  $a \perp b$  und  $\langle \cdot, \cdot \rangle = \| \cdot \|^2$  nach Definition folgt die Behauptung. ■

Abbildung 2.4: orthogonale Projektion von  $a$  auf  $b$ 

**Bemerkung 2.26:** Dass der Satz des Pythagoras gilt, wird aus der Schule längst bekannt sein. Den Wert der obigen Aussage mögen folgende Aspekte verdeutlichen:

1. Der Satz des Pythagoras gilt nicht nur für das euklidische Skalarprodukt und die von ihm induzierte Norm, sondern für *jedes* Skalarprodukt mit induzierter Norm.
2. Er gilt nicht nur in der Ebene, sondern allgemeiner im  $\mathbb{R}^n$ .
3. Der Beweis mithilfe der Vektorrechnung ist kurz und elegant verglichen mit dem elementargeometrischen Beweis, der auf Flächenvergleichen basiert.
4. Dass der Satz des Pythagoras mit den von der Ebene abstrahierten geometrischen Begriffen und Definitionen immer noch gilt, rechtfertigt diese Abstraktionen.

### 2.1.3 Winkel zwischen Vektoren

Wie schon zuvor gehen wir immer davon aus, dass die Norm  $\|\cdot\|$  durch das gegebene Skalarprodukt induziert wird. Seien  $a, b$  zwei Vektoren und  $b \neq 0$ . Wir möchten die orthogonale Projektion  $p$  von  $a$  in Richtung von  $b$  definieren (Abb. 2.4). Wir bestimmen einen Vektor  $p$ , so dass gilt:

$$\begin{aligned} 1. \quad p \parallel b &\Rightarrow p = \alpha b \\ 2. \quad (a - p) \perp b &\Rightarrow \langle a - p, b \rangle = 0 \end{aligned} \quad (2.4)$$

Daraus folgt:

$$\begin{aligned} \langle a - \alpha b, b \rangle &= 0 \\ \Leftrightarrow \langle a, b \rangle - \alpha \langle b, b \rangle &= 0 \\ \Leftrightarrow \alpha &= \frac{\langle a, b \rangle}{\langle b, b \rangle} \\ \Rightarrow p &= \frac{\langle a, b \rangle}{\langle b, b \rangle} \cdot b \end{aligned} \quad (2.5)$$

$p$  nennt man die *orthogonale Projektion von  $a$  in Richtung  $b$*  oder die *Komponente von  $a$  entlang  $b$* .

Die Norm der Projektion ist:

$$\|p\| = \left\| \frac{\langle a, b \rangle}{\langle b, b \rangle} b \right\| = \left| \frac{\langle a, b \rangle}{\langle b, b \rangle} \right| \|b\| = \frac{|\langle a, b \rangle|}{\|b\|^2} \cdot \|b\| = \frac{|\langle a, b \rangle|}{\|b\|} \quad (2.6)$$

**Bemerkung 2.27:** Ist  $b$  ein Einheitsvektor, so gilt:

$$p = \langle a, b \rangle \cdot b.$$

**Beispiel 2.28:** Es sei

$$a = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Dann ist die Komponente von  $a$  entlang  $b$ :

$$p = \frac{\langle a, b \rangle}{\langle b, b \rangle} b = \frac{1 + 2 - 6}{1 + 1 + 4} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ -1 \end{pmatrix}.$$

**Bemerkung 2.29:** Die Idee der orthogonalen Projektion mag an dieser Stelle unscheinbar anmuten, es handelt sich aber um eines der zentralsten Konzepte der Mathematik, dessen Bedeutung nicht überschätzt werden kann. So basiert die Methode der Finiten Elemente, die zur Numerischen Simulation mechanischer Eigenschaften von Bauteilen weite Anwendung findet (vgl. Beispiel 1.2 aus Kapitel 1.1), auf orthogonalen Projektionen, wenn auch in einem erneut verallgemeinerten Sinn.

Ähnliche Bedeutung in der Mathematik hat die Cauchy<sup>1</sup>-Schwarzsche<sup>2</sup> Ungleichung.

**Satz 2.30 (Cauchy-Schwarzsche Ungleichung):** Für  $a, b \in \mathbb{R}^n$  gilt

$$|\langle a, b \rangle| \leq \|a\| \|b\|. \quad (2.7)$$

**BEWEIS** Für  $b = 0$  oder  $a = 0$  sind beide Seiten der Ungleichung gleich 0, und Formel (2.7) gilt. Es sei jetzt  $b \neq 0$ . Wegen  $p \parallel b$  gilt nach Gleichung (2.4b)  $p - a \perp p$  und daher mit dem Satz des Pythagoras

$$\|a\|^2 = \|p\|^2 + \|p - a\|^2 \geq \|p\|^2$$

<sup>1</sup>Augustin-Louis Cauchy (1789-1857), franz. Mathematiker; bedeutende Beiträge zu den Grundlagen der Analysis, zur Funktionentheorie und Mechanik; Beweis der Cauchy-Schwarzschen Ungleichung für den Spezialfall des Standardskalarprodukts 1821.

<sup>2</sup>Hermann Amandus Schwarz (1843-1921), Beiträge zur Funktionentheorie und Analysis

aufgrund von  $\|p - a\|^2 \geq 0$ . Man erhält also  $\|a\| \geq \|p\|$  aufgrund der Monotonie der Wurzelfunktion. Einsetzen dieser Ungleichung in (2.6) liefert

$$\|a\| \geq \|p\| = \frac{|\langle a, b \rangle|}{\|b\|}$$

und damit nach Multiplikation mit  $\|b\|$  die Behauptung. ■

Wir holen nun den Beweis der Normeigenschaft N4 für eine durch ein Skalarprodukt induzierte Norm nach und vollenden so den Beweis von Satz 2.13.

**Satz 2.31 (Dreiecksungleichung):** Für  $a, b \in \mathbb{R}^n$  gilt

$$\|a + b\| \leq \|a\| + \|b\|.$$

**BEWEIS** Beide Seiten dieser Ungleichung sind nicht negativ. Daher genügt es, zu beweisen, dass ihre Quadrate die gewünschte Ungleichung erfüllen d. h.

$$\langle a + b, a + b \rangle \leq (\|a\| + \|b\|)^2.$$

Man hat

$$\langle a + b, a + b \rangle = \langle a, a \rangle + 2\langle a, b \rangle + \langle b, b \rangle \leq \|a\|^2 + 2\|a\| \|b\| + \|b\|^2$$

aufgrund der Cauchy-Schwarzschen Ungleichung (2.7). Mit der ersten binomischen Formel folgt die Behauptung. ■

**Bemerkung 2.32:** Dass bei einer beliebigen Norm, die eventuell nicht durch ein Skalarprodukt induziert ist, die Dreiecksungleichung gilt, folgt nicht aus Satz 2.31 und muss von Fall zu Fall bewiesen werden.

Wir kommen jetzt zur Definition von Winkeln zwischen zwei Vektoren. Man könnte auch hier von einem allgemeinen Skalarprodukt ausgehen, aber solcherart definierte Winkel haben nur geringe geometrische Bedeutung. Daher betrachten wir von jetzt an nur das euklidische Skalarprodukt und die euklidische Norm. Aus der Cauchy-Schwarzschen Ungleichung (2.7) folgt für zwei Vektoren  $a, b \neq 0$

$$-1 \leq \frac{\langle a, b \rangle}{\|a\| \|b\|} \leq 1.$$

Weil zudem der Cosinus auf dem abgeschlossenen Intervall  $[0, \pi]$  streng monoton fällt, gibt es genau ein  $\varphi \in [0, \pi]$  mit

$$\cos \varphi = \frac{\langle a, b \rangle}{\|a\| \|b\|}. \quad (2.8)$$

Man definiert daher den Winkel zwischen zwei Vektoren wie folgt.

**Definition 2.33:** Seien  $a, b \in \mathbb{R}^n \setminus \{0\}$ . Der Winkel zwischen  $a$  und  $b$ , geschrieben  $\angle(a, b)$ , wird definiert als

$$\angle(a, b) := \arccos \frac{\langle a, b \rangle}{\|a\| \|b\|}. \quad (2.9)$$

Wir werden nun zeigen, dass der Winkel nach Definition 2.33 in der Ebene mit der intuitiven Definition eines Winkels übereinstimmt. Gemäß Gleichung (2.8) ändert sich der Winkel nicht, wenn man statt  $a$  und  $b$  die normierten Vektoren  $\|a\|^{-1}a$  und  $\|b\|^{-1}b$  betrachtet. Wir können also annehmen, dass  $a$  und  $b$  normiert sind. Aus der ebenen Elementargeometrie ist bekannt, dass jeder Punkt auf dem Einheitskreis die Koordinaten  $(\cos \alpha, \sin \alpha)$  für ein eindeutiges  $\alpha \in [0, 2\pi)$  hat. Es gilt also  $a = (\cos \alpha, \sin \alpha)^T$  und  $b = (\cos \beta, \sin \beta)^T$ . Wir gehen davon aus, dass  $\beta > \alpha$ , ansonsten vertauschen wir die Bezeichnungen der Vektoren und Winkel. Weil man unter einem Winkel zwischen zwei Strecken immer den kleineren der beiden möglichen versteht, gilt  $0 \leq \beta - \alpha \leq \pi$ . Für den Winkel  $\varphi$  zwischen  $a$  und  $b$  folgt

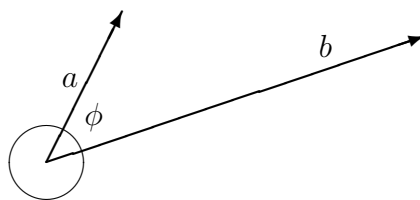
$$\frac{\langle a, b \rangle}{\|a\| \|b\|} = \frac{\cos \alpha \cos \beta + \sin \alpha \sin \beta}{1 \cdot 1} = \cos \varphi$$

Nach einem Additionstheorem aus der Analysis gilt aber

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

und da auf  $[0, \pi]$  der Cosinus invertierbar ist, gilt  $\varphi = \beta - \alpha$ . Damit entspricht der abstrakt definierte Winkel  $\varphi$  im Spezialfall der Ebene genau dem intuitiven Winkel zwischen den Vektoren.

**Bemerkung 2.34:** Wegen  $0 \leq \varphi \leq \pi$  handelt es sich bei dem Winkel im Sinne von Definition 2.33 immer um den kleineren der Winkel zwischen den entsprechenden Ortsvektoren.



**Beispiel 2.35:** Es seien  $A = (1, 2, -3)$  und  $B = (2, 1, 5)$ . Bestimmen Sie den Cosinus des Winkels  $\theta$  zwischen den zugehörigen Ortsvektoren  $a$  und  $b$ . Nach Definition 2.33 gilt

$$\cos \theta = \frac{\langle a, b \rangle}{\|a\| \|b\|} = \frac{2 + 2 - 15}{\sqrt{14}\sqrt{30}} = \frac{-11}{\sqrt{420}} = \frac{-11\sqrt{105}}{210},$$

also  $\theta \approx 122,4^\circ$ .

**Beispiel 2.36:** Bestimmen Sie den Cosinus des Winkels zwischen den zwei lokalisierten Vektoren  $\vec{PQ}$  und  $\vec{PR}$  mit  $P = (1, 2, -3)$ ,  $Q = (-2, 1, 5)$ ,  $R = (1, 1, -4)$ .

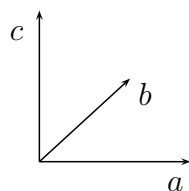
Man errechnet  $A = Q - P = (-3, -1, 8)$  und  $B = R - P = (0, -1, -1)$ . Dann ist der Winkel zwischen  $\vec{PQ}$  und  $\vec{PR}$  der gleiche wie der zwischen  $a = \vec{OA}$  und  $b = \vec{OB}$ . Somit gilt für den gesuchten Winkel  $\theta$

$$\cos \theta = \frac{\langle a, b \rangle}{\|a\| \|b\|} = \frac{0 + 1 - 8}{\sqrt{74} \sqrt{2}} = \frac{-7}{\sqrt{74} \sqrt{2}}.$$

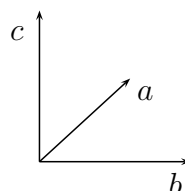
### 2.1.4 Das Vektorprodukt (Kreuzprodukt)

Die Ergebnisse dieses Abschnitts gelten im Gegensatz zu den bisherigen Ergebnissen nur im  $\mathbb{R}^3$  und nur für das euklidische Skalarprodukt bzw. die euklidische Norm.

**Definition 2.37:** Gegeben seien 3 Vektoren  $a, b$  und  $c$ . Die Vektoren  $a$  und  $b$  liegen in einer Ebene und umschließen einen Winkel. Wenn man  $a$  auf kürzestem Wege auf  $b$  dreht, geschieht das entweder im oder gegen den Uhrzeigersinn. Ist der kürzeste Weg im Uhrzeigersinn und zeigt  $c$  von der Ebene aus gesehen nach unten, so bilden  $a, b$  und  $c$  ein *Rechtssystem*, andernfalls ein *Linkssystem*.



Rechtssystem



Linkssystem

Man kann sich Definition 2.37 folgendermaßen veranschaulichen:

1. Rechte-Hand-Regel: Man zeigt mit dem Daumen der rechten Hand in Richtung des ersten Vektors  $a$ , während man versucht mit dem Zeigefinger in Richtung des zweiten Vektors  $b$  zu zeigen. Dann zeigt der Mittelfinger in die Richtung von  $c$ , wenn er senkrecht zu Daumen und Zeigefinger gehalten wird. Bei stumpfen Winkeln zwischen  $a$  und  $b$  widerspricht diese Haltung zwar der menschlichen Anatomie, aber wichtig ist hier nur, dass man mit Daumen und Zeigefinger den kleineren (inneren) der beiden Winkel zwischen  $a$  und  $b$  darzustellen versucht.
2. Für Heimwerker: Man dreht eine Schraube (mit gewöhnlichem Rechtsgewinde) von oben betrachtet im Uhrzeigersinn ein und gegen den Uhrzeigersinn aus (vgl. Abb. 2.5).

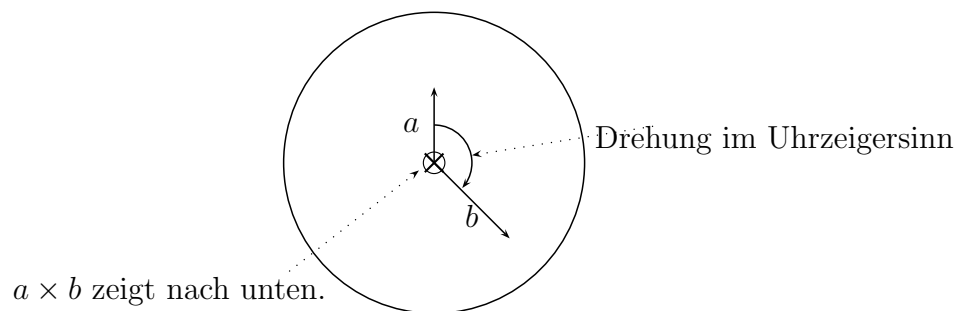
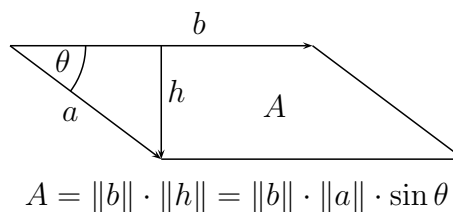


Abbildung 2.5: Veranschaulichung eines Rechtssystems an einer Schraube

Abbildung 2.6: Fläche des durch  $a$  und  $b$  aufgespannten Parallelogramms

**Definition 2.38:** Das *Vektorprodukt*, oder auch *Kreuzprodukt*,  $a \times b \in \mathbb{R}^3$  zweier Vektoren  $a, b \in \mathbb{R}^3$  ist für den Fall  $a = \lambda b$  als der Nullvektor definiert, anderenfalls über folgende Eigenschaften:

1.  $a \times b$  steht sowohl senkrecht auf  $a$  als auch auf  $b$ .
2.  $a, b$  und  $a \times b$  bilden ein Rechtssystem.
3. Die Länge von  $a \times b$  ist gleich dem Flächeninhalt des von  $a$  und  $b$  aufgespannten Parallelogramms (vgl. Abb. 2.6), d. h.

$$\|a \times b\| = \|a\| \|b\| \sin \theta$$

Für den Fall  $a = \lambda b$  gilt  $a \times b = 0$ .

**Satz 2.39:** Aus der Definition des Kreuzprodukts leiten sich folgende Eigenschaften her:

1. Antikommutativität:  $a \times b = -(b \times a)$ .
2. Distributivität:
  - $a \times (b + c) = (a \times b) + (a \times c)$



$$\bullet (b + c) \times a = (b \times a) + (c \times a)$$

$$3. \text{ Für jede reelle Zahl } \alpha \text{ gilt: } (\alpha a) \times b = \alpha(a \times b) = a \times (\alpha b).$$

$$4. (a \times b)^2 = \langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle^2$$

Aus diesen Eigenschaften lässt sich eine einfache Darstellung ableiten:

**Satz 2.40:** Für  $a, b \in \mathbb{R}^3$  gilt

$$a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

**BEWEIS** Da unser kanonisches Koordinatensystem ein Rechtssystem ist, gilt:

$$\begin{aligned} a \times b &= (a_1 e_1 + a_2 e_2 + a_3 e_3) \times (b_1 e_1 + b_2 e_2 + b_3 e_3) \\ &= a_1 b_1 \underbrace{e_1 \times e_1}_0 + a_1 b_2 \underbrace{e_1 \times e_2}_{e_3} + a_1 b_3 \underbrace{e_1 \times e_3}_{-e_2} \\ &\quad + a_2 b_1 \underbrace{e_2 \times e_1}_{-e_3} + a_2 b_2 \underbrace{e_2 \times e_2}_0 + a_2 b_3 \underbrace{e_2 \times e_3}_{e_1} \\ &\quad + a_3 b_1 \underbrace{e_3 \times e_1}_{e_2} + a_3 b_2 \underbrace{e_3 \times e_2}_{-e_1} + a_3 b_3 \underbrace{e_3 \times e_3}_0 \\ &= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \end{aligned}$$

**Beispiel 2.41:** Für  $a = (2, 3, -1)^T$  und  $b = (-1, 1, 5)^T$  gilt

$$a \times b = \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}.$$

Es gibt 2 Schemata, mit denen sich die Berechnung des Kreuzproduktes leicht merken lässt:

Erste Möglichkeit:

<p>1. Komponente</p> $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} + & - \\ - & + \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ $15 - (-1) = 16$	<p>2. Komponente</p> $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} - & + \\ + & - \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ $-10 + 1 = -9$	<p>3. Komponente</p> $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{matrix} + & - \\ - & + \end{matrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ $2 - (-3) = 5$
--	--	--

Entlang der Pfeile wird überkreuz multipliziert und entsprechend der Vorzeichen addiert bzw. subtrahiert.

Zweite Möglichkeit:

$$\ominus \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} \oplus$$

$$\begin{array}{rcccl} -1 & \leftarrow & -1 & \rightarrow & 15 \\ 10 & \leftarrow & 2 & \rightarrow & 1 \\ -3 & \leftarrow & 3 & \rightarrow & 2 \end{array} \Rightarrow \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}$$

Unter beide Vektoren werden die ersten beiden Komponenten des jeweiligen Vektors geschrieben. Dann werden die Komponenten entlang der schrägen Pfeile multipliziert und die Produkte am rechten und linken Rand notiert. Schließlich werden die Produkte am linken Rand von denen am rechten Rand abgezogen.

**Bemerkung 2.42:** Das Vektorprodukt ist i. A. nicht assoziativ. Seien z. B.  $a = (1, 0, 0)^T$ ,  $b = (0, 1, -1)^T$  und  $c = (0, 0, 1)^T$ . Dann gilt  $a \times b = (0, 1, 1)^T$  und  $b \times c = (0, 0, 1)^T$ . Daraus folgt  $(a \times b) \times c = (1, 0, 0)^T$ , aber  $a \times (b \times c) = (0, 0, 0)^T$ . Es gilt jedoch die *Graßmannsche Identität*

$$a \times (b \times c) = \langle a, c \rangle b - \langle a, b \rangle c,$$

die man direkt nachrechnet.

**Beispiel 2.43:** Auch wenn das Vektorprodukt nur für Vektoren aus dem  $\mathbb{R}^3$  definiert ist, so lässt es sich dennoch auch im  $\mathbb{R}^n$  für  $n > 3$  nutzen, um zu  $a$  und  $b$  orthogonale Vektoren zu finden. Mit Hilfe des Kreuzproduktes wird zu den Vektoren

$$a = \begin{pmatrix} 2 \\ 4 \\ 3 \\ -1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 5 \end{pmatrix}$$

mit Hilfe der ersten, dritten und vierten Komponente ein orthogonaler Vektor berechnet, wenn die zweite Komponente auf 0 gesetzt wird:

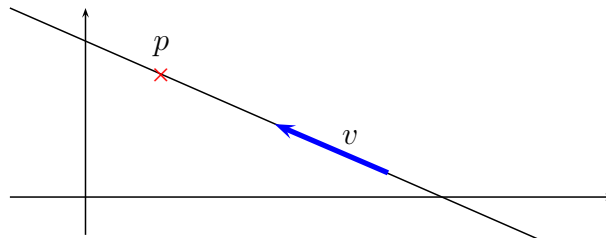
$$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ -9 \\ 5 \end{pmatrix}, \quad \text{also ist} \quad c = \begin{pmatrix} 16 \\ 0 \\ -9 \\ 5 \end{pmatrix}$$

senkrecht zu  $a$  und  $b$ .

## 2.2 Geraden und Ebenen

### 2.2.1 Definition und grundlegende Eigenschaften

Bei der Definition einer Geraden in der Ebene gehen wir direkt von der Anschauung aus. Von jetzt an gelte immer  $\alpha, \beta \in \mathbb{R}$  und  $n, p, v \in \mathbb{R}^2$ .



#### Definition 2.44:

1. Für einen Ortsvektor oder Aufpunkt  $p$  und einen Richtungsvektor  $v \neq 0$  heie

$$G := \{x \in \mathbb{R}^2 \mid \exists \alpha \in \mathbb{R} : x = p + \alpha v\} \quad (2.10)$$

eine Gerade.

2. Die Gleichung

$$x = p + \alpha v$$

aus (2.10) heit *Punkt-Richtungsgleichung* von  $G$ , die reelle Zahl  $\alpha$  aus Formel (2.10) nennt man *Parameter*.

**Bemerkung 2.45:** Wir identifizieren im Folgenden hufig eine Gerade, die ja als Teilmenge von  $\mathbb{R}^2$  definiert ist, mit ihrer Geradengleichung, weil sie die Gerade eindeutig festlegt.

Nach Definition 2.44 legen Aufpunkt und Richtungsvektor eine Gerade eindeutig fest, jeder Punkt  $x$  auf der Geraden gengt der Geradengleichung  $x = p + \alpha v$  fr ein bestimmtes  $\alpha$ . Andererseits sind bei einer gegebenen Gerade Aufpunkt und Richtungsvektor nicht eindeutig bestimmt.

**Satz 2.46:** Sei  $G$  eine Gerade mit dem Richtungsvektor  $v$  und dem Aufpunkt  $p$ .

1. Jeder Vektor der Form  $\tilde{v} = \beta v$  mit  $\beta \neq 0$  ist ebenfalls Richtungsvektor.
2. Jeder Vektor der Form  $\tilde{p} = p + \gamma v$  mit irgendeinem  $\gamma \in \mathbb{R}$  ist ebenfalls Aufpunkt.

BEWEIS 1.) Sei  $G$  gegeben durch  $x = p + \alpha v$  und  $\tilde{G}$  durch  $x = p + \alpha \tilde{v}$ . Wir zeigen  $G = \tilde{G}$ . Man hat

$$\begin{aligned} x \in G &\Leftrightarrow \exists \alpha : x = p + \alpha v \\ &\Leftrightarrow \exists \alpha : x = p + \frac{\alpha}{\beta}(\beta v) \\ &\Leftrightarrow \exists \tilde{\alpha} := \frac{\alpha}{\beta} : x = p + \tilde{\alpha} \tilde{v} \Leftrightarrow x \in \tilde{G}. \end{aligned}$$

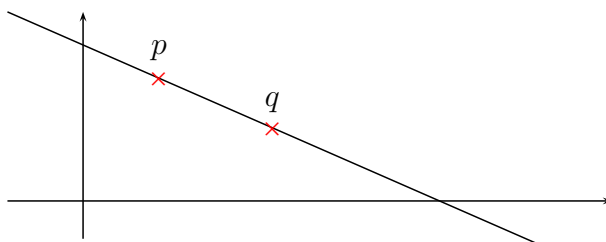
Also sind die beiden Mengen gleich, und es gilt  $G = \tilde{G}$ . Man beachte, dass man hier wegen  $\beta \neq 0$  nicht durch Null dividiert.

2.) Sei  $G$  gegeben durch  $x = p + \alpha v$  und  $\tilde{G}$  durch  $x = \tilde{p} + \alpha v$ . Dann gilt

$$\begin{aligned} x \in G &\Leftrightarrow \exists \alpha : x = p + \alpha v \\ &\Leftrightarrow \exists \alpha : x = (p + \gamma v) + \alpha v - \gamma v \\ &\Leftrightarrow \exists \tilde{\alpha} := \alpha - \gamma : x = \tilde{p} + \tilde{\alpha} v \Leftrightarrow x \in \tilde{G} \end{aligned}$$

und daher  $G = \tilde{G}$ . ■

Wir gehen nun davon aus, dass anstelle eines Aufpunktes und eines Richtungsvektors von einer Geraden zwei verschiedene Punkte bekannt sind.



**Satz 2.47:** Seien  $p, q \in \mathbb{R}^2$  mit  $p \neq q$ . Dann gibt es genau eine Gerade durch  $p$  und  $q$ .

BEWEIS **Existenz** Wir wählen als Aufpunkt  $p$  und als Richtungsvektor  $v = q - p$ , die hierdurch festgelegte Gerade sei  $G$  mit der Geradengleichung

$$x = p + \alpha(q - p).$$

Es gilt  $p = p + 0v \in G$  und  $q = p + 1(q - p) = p + 1v \in G$ , also enthält  $G$  sowohl  $p$  als auch  $q$ .

**Eindeutigkeit** Sei  $\tilde{G}$  noch eine Gerade mit der Geradengleichung  $x = \tilde{p} + \alpha \tilde{v}$ , die  $p$  und  $q$  enthält. Weil nach Satz 2.46.2 sich jeder Punkt auf  $\tilde{G}$ , also auch  $p$ , als Aufpunkt eignet, wird  $\tilde{G}$  auch durch die Geradengleichung  $x = p + \alpha \tilde{v}$  festgelegt. Wegen  $q \in \tilde{G}$  gilt  $q = p + \alpha \tilde{v}$  für ein bestimmtes  $\alpha \in \mathbb{R}$ . Aufgrund von  $p \neq q$  erhält man  $\alpha \neq 0$  und durch Subtraktion von  $q$  auf beiden Seiten:

$$v = q - p = \alpha \tilde{v}.$$

Beide Richtungsvektoren sind somit Vielfache voneinander; aus Satz 2.46.1 folgt  $G = \tilde{G}$ , also die Eindeutigkeit. ■

**Definition 2.48:**

1. Seien  $p, q$  zwei verschiedene Punkte auf der Geraden  $G$ . Dann heißt die Geradengleichung

$$x = p + \alpha(q - p)$$

*Zweipunktform* von  $G$ . Man sagt, "die Gerade liegt in der Zwei-Punkte-Form vor".

2. Alle Gleichungen, die einen reellen Parameter enthalten und eine Gerade beschreiben, nennt man *Parametergleichungen* oder *Parameterformen* einer Geraden.

**Bemerkung 2.49:**

1. Nach Satz 2.46 ist man frei, statt eines vorliegenden Richtungsvektors ein Vielfaches davon zu wählen. Das kann man in der Praxis nutzen, um einen Richtungsvektor mit "möglichst einfachen Zahlen" zu finden. Ersetzt man z.B. den Richtungsvektor  $(\pi/3, \pi/6)^T$  durch  $(2, 1)^T$ , wird das in vielen Fällen Berechnungen wesentlich handhabbarer werden lassen. Analoges gilt für die Wahl des Aufpunkts.
2. Die Umrechnung von der Zwei-Punkte-Form in die Punkt-Richtungsform und umgekehrt ist einfach möglich: Liegt die Zwei-Punkte-Form vor, so findet man z.B. durch  $q - p$  einen Richtungsvektor; als Aufpunkt kann man z. B.  $p$  oder  $q$  wählen. Umgekehrt findet man ausgehend vom Richtungsvektor  $v$  und dem Aufpunkt  $p$  z. B. durch  $p + v$  oder  $p - v$  einen weiteren Punkt auf der Geraden.
3. Dass man durch zwei verschiedene Punkte genau eine Gerade legen kann, hätte man sicherlich auch ohne höhere Mathematik anhand einiger Experimente mit Papier und Lineal einsehen können. Entscheidend ist aber, dass man diese Tatsache jetzt lückenlos bewiesen hat, ohne an eine wie auch immer geartete Einsicht appellieren zu müssen. Zudem ist es ein Indiz dafür, dass die Definition 2.44 des geometrischen Objekts "Gerade" mittels einer algebraischen Gleichung zielführend ist, weil typische Eigenschaften einer Geraden in der Tat aus der Definition folgen.

Wir betrachten nun alternative Beschreibungen von Geraden in der Ebene, bei denen auf einen Parameter verzichtet wird. Vorbereitend zeigen wir folgendes Lemma.

**Lemma 2.50:** Seien  $u, v, w \in \mathbb{R}^2 \setminus \{0\}$  beliebig,  $v \perp u$  und  $w \perp v$ . Dann gilt  $u \parallel w$ .

**BEWEIS** Sei  $u = (u_1, u_2)^T$ . Dann muss nach Bemerkung 2.21  $v = (v_1, v_2)^T = \lambda(-u_2, u_1)^T$  gelten und abermals nach Bemerkung 2.21  $w = \delta(-v_2, v_1)^T = \delta\lambda(u_1, u_2)^T$ . ■

**Satz 2.51:** Sei  $G$  eine Gerade mit Richtungsvektor  $v$  und Aufpunkt  $p$ . Dann existiert ein Vektor  $n \perp v$  mit  $n \neq 0$ , so dass gilt

$$x \in G \Leftrightarrow \langle x, n \rangle = \langle p, n \rangle.$$

Der Vektor  $n$  ist bis auf Skalierung eindeutig bestimmt.

**BEWEIS** “ $\Rightarrow$ ”: Sei  $x \in G$ . Dann existiert  $\alpha : x = p + \alpha v$ . Es folgt mit den Rechenregeln des Skalarprodukts:

$$\langle x, n \rangle = \langle p, n \rangle + \alpha \underbrace{\langle v, n \rangle}_{=0}$$

“ $\Leftarrow$ ”: Erfülle  $x$  die Gleichung  $\langle x, n \rangle = \langle p, n \rangle$ . Es folgt  $\langle x - p, n \rangle = 0$ , daher  $x - p \perp n$  und nach Lemma 2.50  $x - p \parallel v$ , also nach Definition 1.9  $x - p = \alpha v$  für ein  $\alpha \in \mathbb{R}$ . Es folgt durch Addition von  $p$  auf beiden Seiten  $x = p + \alpha v$ .

Die Eindeutigkeit von  $n$  bis auf Skalierung folgt sofort aus Bemerkung 2.21. ■

**Definition 2.52:** Sei  $G$  eine Gerade in der Ebene und  $p$  der Aufpunkt von  $G$ .

1. Ein Vektor  $n$  wie in Satz 2.51 heißt *Normalenvektor* von  $G$ .
2. Die Gleichung

$$\langle x, n \rangle = \langle p, n \rangle$$

heißt *Normalform* von  $G$ .

Wir haben oben bereits gesehen, dass jeder Punkt auf der Geraden als Aufpunkt gewählt werden kann. Es bleibt noch zu zeigen, dass auch die Normalform einer Geraden nicht von der Wahl des Aufpunkts abhängt.

**Bemerkung 2.53:** Seien  $p, q$  zwei Punkte auf der Geraden  $G$  mit Normalenvektor  $n$ . Dann gilt

$$\langle p, n \rangle = \langle q, n \rangle.$$

**BEWEIS** Wir gehen von der Punkt-Richtungsform von  $G$  aus und wählen  $p$  als Aufpunkt. Dann gilt  $q = p + \alpha v$  für ein  $\alpha$  und daher  $q - p = \alpha v$  und damit  $\langle q - p, n \rangle = 0 = \langle q, n \rangle - \langle p, n \rangle$  wegen  $\alpha v \perp n$ . ■

**Bemerkung 2.54:**

1. Seien  $G$ ,  $n$  und  $p$  wie zuvor. Mit  $n = (a, b)^T \in \mathbb{R}^2$ ,  $x = (x_1, x_2)^T$  und  $c := \langle p, n \rangle \in \mathbb{R}$  folgt die *allgemeine Geradengleichung*

$$ax_1 + bx_2 = c.$$

2. Gilt  $b \neq 0$ , hat also der Richtungsvektor eine  $x_1$ -Komponente, setzt man  $m = -a/b$  und  $n = c/b$  (dieses  $n$  ist an dieser Stelle *nicht* der Normalenvektor, sondern eine reelle Zahl) und erhält so die vielleicht aus der Schule bekannte Geradengleichung

$$y = mx + n,$$

wobei hier  $x := x_1$  und  $y := x_2$  gesetzt wird, um auch in der Notation den Bezug zu vielleicht Bekanntem herzustellen. Es handelt sich also um einen Spezialfall einer Normalform einer Geraden.

**Beispiel 2.55:** Es sei  $p = (2, 1)^T$  und  $n = (5, 1)^T$ . Dann gilt für  $x = (x_1, x_2)^T$ :

$$\begin{aligned} \langle x, n \rangle &= \langle p, n \rangle \\ \Leftrightarrow \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \end{pmatrix} \right\rangle \\ \Leftrightarrow 5x_1 + 1x_2 &= 11 \end{aligned}$$

Wir erhalten also eine allgemeine Geradengleichung zur Beschreibung der Geraden. Man beachte, dass die Vorfaktoren von  $x_1$  und  $x_2$  genau den Komponenten des Normalenvektors entsprechen.

Die Normalform ist nur bis auf einen Faktor genau bestimmt. Etwas eindeutiger wird die Darstellung in der Hesseschen Normalform.

**Definition 2.56:** Sei  $G$  eine Gerade mit Normalenvektor  $n$ . Gilt  $\|n\| = 1$ , so heißt die damit gebildete Normalform *Hessesche Normalform*.

**Bemerkung 2.57:** Man erhält die Hessesche Normalform aus einer beliebigen Normalform, indem man die Normalform durch  $\|n\|$  teilt:

$$\frac{\langle x, n \rangle}{\|n\|} = \frac{\langle p, n \rangle}{\|n\|}.$$

Damit liegt der Normalvektor bis auf das Vorzeichen eindeutig fest.

**Beispiel 2.58:** Die Gerade

$$5x + y = 11$$

hat den Normalenvektor  $n = (5, 1)^T$ , was sich direkt aus den Vorfaktoren der linken Seite ablesen lässt. Damit ist

$$\|n\| = \sqrt{25 + 1} = \sqrt{26},$$

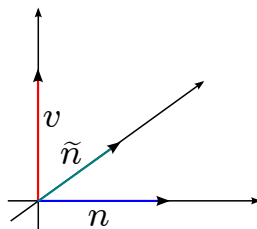
und die Hessesche Normalform lautet

$$\frac{1}{\sqrt{26}}(5x + y) = \frac{11}{\sqrt{26}}.$$

Wir kommen nun zu Geraden im Raum. Wir definieren sie völlig analog zu Geraden in der Ebene und verweisen dazu auf den Wortlaut von Definition 2.44, nur dass jetzt die Vektoren  $p, q, x$  Elemente von  $\mathbb{R}^3$  sind statt von  $\mathbb{R}^2$  wie zuvor. Die Sätze 2.46 und 2.47 gelten dann wortgleich mit wortgleichen Beweisen (es wurde dort niemals benutzt, dass Vektoren nur aus zwei Komponenten bestanden), ebenso lässt sich Definition 2.48 wortgleich auf den vorliegenden Fall übertragen.

Im Gegensatz zur Ebene existiert keine parameterlose Beschreibung einer Geraden im Raum. Dies liegt daran, dass die Richtung eines Normalenvektors zu einem Richtungsvektor nicht eindeutig festliegt.

**Beispiel 2.59:** Sei  $v = (0, 0, 1)^T$  der Richtungsvektor einer Geraden. Dann ist  $n = (1, 0, 0)^T$  ein Normalenvektor aufgrund von  $\langle v, n \rangle = 0$ , aber genauso auch  $\tilde{n} = (0, 1, 0)^T$  und damit nach den Rechenregeln des Skalarprodukts auch alle Vektoren der Form  $\alpha n + \beta \tilde{n}$ , also jeder Vektor, der in der  $(x, y)$ -Ebene liegt.



Wir gehen nun noch kurz auf Ebenen im Raum ein. Den Vektorraum  $\mathbb{R}^2$ , der ja sich geometrisch als Ebene deuten lässt, kann man mit der Teilmenge

$$\tilde{E} := \left\{ x \in \mathbb{R}^3 \mid x = (x_1, x_2, 0)^T; x_1, x_2 \in \mathbb{R} \right\}$$

identifizieren. Es gilt offensichtlich

$$\tilde{E} := \left\{ x \in \mathbb{R}^3 \mid x = (0, 0, 0)^T + x_1(1, 0, 0)^T + x_2(0, 1, 0)^T, x_1, x_2 \in \mathbb{R} \right\};$$

die Ebene  $\tilde{E}$  besteht daher genau aus allen Vektoren der Form  $0 + \alpha(1, 0, 0)^T + \beta(0, 1, 0)^T$  mit den reellen Parametern  $\alpha$  und  $\beta$ . Man beachte, dass die beiden Vektoren unterschiedliche Richtungen aufweisen. Verallgemeinernd definieren wir:



**Definition 2.60:** Seien  $p, v, w \in \mathbb{R}^3$ ,  $v \neq 0$  und  $w \neq 0$ , und seien  $v$  und  $w$  nicht parallel. Dann heißt

$$E := \{x \in \mathbb{R}^3 \mid x = p + \alpha v + \beta w, \alpha, \beta \in \mathbb{R}\}$$

*Ebene*; die Vektoren  $v$  und  $w$  heißen *Richtungsvektoren*. Die Gleichung  $x = p + \alpha v + \beta w$  wird *Punkt-Richtungsgleichung* genannt; man sagt, die "Ebene liegt in Punkt-Richtungsform vor".

**Bemerkung 2.61:**

1. Ebenso wie bei einer Geraden in der Ebene in Parameterform eignet sich jeder Punkt der Ebene als Aufpunkt; beliebige zwei nichtparallele Vektoren in der Ebene ungleich dem Nullvektor eignen sich als Richtungsvektoren.
2. Drei paarweise verschiedene Punkte, die nicht auf einer Geraden liegen, legen eine Ebene eindeutig fest.

**BEWEIS** Die Beweise werden ähnlich zu den Beweisen der entsprechenden Aussagen für Geraden in der Ebene geführt, sind aber etwas komplizierter. Wir verzichten auf Details. ■

Ebenso ohne Beweis geben wir folgenden Satz an.

**Satz 2.62:** Sei  $E$  eine Ebene mit Aufpunkt  $p$  und den Richtungsvektoren  $v$  und  $w$ . Dann existiert ein Vektor  $n \neq 0$  mit  $v \perp n$ ,  $w \perp n$  und

$$x \in E \Leftrightarrow \langle x, n \rangle = \langle p, n \rangle.$$

Der Vektor  $n$  ist bis auf Skalierung eindeutig bestimmt.

**Definition 2.63:** Der Vektor  $n$  aus Satz 2.62 heißt *Normalenvektor* der Ebene, die Gleichung

$$\langle x, n \rangle = \langle p, n \rangle \tag{2.11}$$

heißt *Normalform* oder *Normalgleichung* der Ebene.

**Bemerkung 2.64:**

1. Die Normalform hängt nicht von der Wahl des Aufpunkts  $p$  ab.
2. Sei  $n = (a, b, c)^T$  und  $d = \langle p, n \rangle \in \mathbb{R}$ . Dann erhält man aus (2.11) sofort die *allgemeine Ebenengleichung*

$$ax_1 + bx_2 + cx_3 = d.$$

3. Man erklärt die Hessesche Normalform einer Ebene sinngemäß zu Definition 2.56; sie ist bis auf das Vorzeichen eindeutig.

Wir fassen unsere Ergebnisse in der folgenden Tabelle zusammen.

Raum	Objekt	Punkt-Richtungs-Gl.	Normalgleichung
$\mathbb{R}^2$	Gerade	$x = p + \alpha v$	$\langle x, n \rangle = \langle p, n \rangle$ $ax_1 + bx_2 + c = 0$
$\mathbb{R}^3$	Gerade	$x = p + \alpha v$	Gibt es nicht!
$\mathbb{R}^3$	Ebene	$x = p + \alpha v + \beta w$	$\langle x, n \rangle = \langle p, n \rangle$ $ax_1 + bx_2 + cx_3 + d = 0$

## 2.2.2 Umrechnen zwischen verschiedenen Darstellungsformen

Wir diskutieren nun die Umrechnung zwischen der Parameterform einer Geraden und ihrer Normalform als parameterfreie Darstellung. Da es in  $\mathbb{R}^3$  keine Normalform von Geraden gibt, beziehen wir uns auf Geraden in  $\mathbb{R}^2$ . Zur sehr einfachen Umrechnung zwischen der Punkt-Richtungsform und der Zwei-Punkte-Form verweisen wir auf Bemerkung 2.49 und betrachten nur die Punkt-Richtungsform.

Sei eine Gerade  $G$  gegeben durch  $x = p + \alpha v$  mit  $v = (v_1, v_2)^T$ . Einen Normalenvektor findet man durch

$$n := \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix}$$

Dies ist nach Bemerkung 2.21 auch bis auf Skalierung die einzig mögliche Wahl. Durch Ausrechnen von  $\langle p, n \rangle$  erhält man die rechte Seite der Normalform.

**Beispiel 2.65:** Eine Gerade sei gegeben durch

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Wir lesen  $p = (1, 2)^T$  und  $v = (-1, 3)^T$  ab, erhalten  $n = (3, 1)^T$  und daher mit  $x = (x_1, x_2)^T$  eine Normalform

$$\begin{aligned} \langle x, n \rangle &= \langle p, n \rangle \\ \Leftrightarrow 3x_1 + x_2 &= 5. \end{aligned}$$

Liegt die Gerade in der Normalform  $ax_1 + bx_2 = c$  vor, so braucht man einen Richtungsvektor, der senkrecht zu  $n = (a, b)^T$  steht. Dieser errechnet sich zu  $v = (b, -a)^T$ . Einen Aufpunkt  $p$  erhält man, in dem man z. B.  $x_1$  oder  $x_2$  gleich 0 wählt und aus der parameterlosen Form die andere Komponente errechnet.

**Beispiel 2.66:** Eine Gerade in  $\mathbb{R}^2$  ist durch ihre Normalform  $3x + y = 5$  gegeben. Damit ist  $n = (3, 1)^T$  und  $v = (1, -3)^T$ . Setzt man  $x_1 = 0$  ergibt sich aus der Normalform  $x_2 = 5$  und damit  $p = (0, 5)^T$ . Zusammengefasst erhält man

$$x = \begin{pmatrix} 0 \\ 5 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Wir kommen nun zur Umrechnung von Ebenendarstellungen. Zur Umrechnung der Punkt-Richtungsform in die Normalform benötigt man einen Normalenvektor. In  $\mathbb{R}^3$  kann man dazu das Kreuzprodukt der beiden Richtungsvektoren verwenden. Nach den Überlegungen zum Kreuzprodukt in Kapitel 2.1.4 ist das bis auf Skalierung der einzig mögliche Normalenvektor. Als Aufpunkt lässt sich jeder Punkt der Ebene verwenden, insbesondere der Vektor  $p$  aus der Punkt-Richtungsform.

**Beispiel 2.67:** Eine Ebene in  $\mathbb{R}^3$  sei durch ihre Punkt-Richtungsform

$$x = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix}$$

definiert. Wir erhalten einen Normalenvektor  $n$  durch

$$n = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$$

und schließlich mit  $x = (x_1, x_2, x_3)^T$  die Normalform

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix} \right\rangle \Leftrightarrow -x_1 + x_2 + 4x_3 = 1$$

Zur Umrechnung einer Normalform einer Ebene in die Punkt-Richtungsform benötigen wir zwei nicht parallele Richtungsvektoren, die beide senkrecht auf  $n$  stehen müssen. Aus der Normalgleichung  $ax_1 + bx_2 + cx_3 = d$  liest man den Normalenvektor  $n = (a, b, c)^T$  ab. Mindestens eine Komponente  $n_i$  von  $n$  ist ungleich 0. Wir vertauschen  $n_i$  mit einer anderen Komponente  $n_j$  und verändern das Vorzeichen von  $n_i$  im so erzeugten Vektor. Weil es zwei Möglichkeiten gibt,  $j \neq i$  zu wählen, erhalten wir zwei Vektoren  $v$  und  $w$ . Man erkennt sofort  $v \perp n$  und  $w \perp n$ , und weder  $v$  noch  $w$  sind der Nullvektor. Aufgrund der Position der Nullen können  $v$  und  $w$  nicht parallel sein. Es fehlt noch ein Aufpunkt. Er wird errechnet, indem man zwei der drei Koordinaten von  $x = (x_1, x_2, x_3)^T$  z. B. den Wert 0 zuweist und dann aus der Normalform den Wert der fehlenden Koordinate errechnet. Sollte dies nicht möglich sein, wähle man ein anderes Koordinatenpaar. Es gibt immer zwei Koordinaten in  $x$ , mit denen obige Rechnung möglich ist.

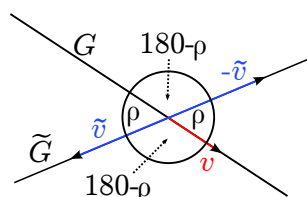


Abbildung 2.7: Winkel zwischen zwei Geraden

**Beispiel 2.68:** Aus der Normalgleichung

$$-x + y + 4z = 1$$

einer Ebene ergibt sich der Normalenvektor  $n = (-1, 1, 4)^T$  und damit als Richtungsvektoren  $v = (1, 1, 0)^T$  und  $w = (4, 0, 1)^T$ . Setzt man  $x_1$  und  $x_3$  gleich 0, erhält man  $x_2 = 1$ . Das führt zu  $p = (0, 1, 0)^T$  und damit zur Punkt-Richtungsform

$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}.$$

### 2.2.3 Lagebeziehungen von Geraden und Ebenen

Wir erklären zunächst Orthogonalität und Winkel zwischen Geraden und Ebenen.

**Definition 2.69:**

1. Zwei Geraden in  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  heißen *parallel*, wenn ihre Richtungsvektoren parallel sind.
2. Zwei sich schneidende Geraden heißen *orthogonal*, wenn ihre Richtungsvektoren orthogonal sind.
3. Seien  $G$  und  $\tilde{G}$  zwei sich schneidende Geraden mit Richtungsvektoren  $v$  bzw.  $\tilde{v}$ . Der *Winkel* zwischen den Geraden  $\angle(G, \tilde{G})$  wird definiert durch

$$\angle G, \tilde{G} := \min\{\angle(v, \tilde{v}), \angle(v, -\tilde{v})\}$$

**Bemerkung 2.70:**

1. Es ist nicht möglich, den Winkel zwischen zwei Geraden einfach als den Winkel zwischen den beiden Richtungsvektoren der Geraden zu definieren, weil diese ja nicht eindeutig festliegen. Ersetzt man

den Richtungsvektor  $\tilde{v}$  durch  $-\tilde{v}$ , dann wird aus dem Winkel  $\rho$  zwischen den Richtungsvektoren der Winkel  $180^\circ - \rho$ , obwohl die Geraden unverändert geblieben sind (vgl. Abb. 2.7). Skalieren des Richtungsvektors mit einer positiven Zahl dagegen ändert den Winkel nicht.

2. Abb. 2.7 zeigt außerdem, dass der Winkel zwischen zwei Geraden niemals größer als  $90^\circ$  werden kann.

3. Für den Winkel zwischen zwei Vektoren  $v, \tilde{v}$  gilt

$$\angle(v, \tilde{v}) = \arccos\left(\frac{\langle v, \tilde{v} \rangle}{\|v\| \|\tilde{v}\|}\right) \quad (2.12)$$

(vgl. Definition 2.33), so dass Winkel größer als  $90^\circ$  gerade einem negativen Wert des Skalarprodukts im Zähler entsprechen. Der Übergang von  $\tilde{v}$  auf  $-\tilde{v}$  entspricht genau dem Umkehren des Vorzeichens des Bruches in Gleichung (2.12). Daher gilt für den Winkel zwischen zwei sich schneidenden Geraden  $G$  und  $\tilde{G}$ :

$$\angle(G, \tilde{G}) = \arccos\left(\frac{|\langle v, \tilde{v} \rangle|}{\|v\| \|\tilde{v}\|}\right) \quad (2.13)$$

**Beispiel 2.71:** Sei  $G$  durch

$$x = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

und  $\tilde{G}$  durch

$$x = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

gegeben. Dann gilt nach Formel (2.13):

$$\cos(\angle(G, \tilde{G})) = \frac{|-1|}{\sqrt{2}\sqrt{6}} = \frac{1}{\sqrt{12}}$$

Der gesuchte Winkel beträgt damit ca.  $73,22^\circ$ . Der Winkel zwischen den beiden Richtungsvektoren beträgt dagegen ca.  $106,78^\circ$ .

**Definition 2.72:**

1. Zwei Ebenen heißen *parallel*, wenn ihre Normalenvektoren parallel sind. Sie heißen *orthogonal*, wenn ihre Normalenvektoren orthogonal sind.

2. Seien  $n$  und  $\tilde{n}$  Normalenvektoren der beiden Ebenen  $E$  und  $\tilde{E}$ . Dann wird der Winkel  $\angle(E, \tilde{E})$  zwischen den beiden Ebenen erklärt durch

$$\angle(E, \tilde{E}) := \min\{\angle(n, \tilde{n}), \angle(n, -\tilde{n})\}$$

Mit einer zum Fall der Geraden analogen Argumentation erkennt man

$$\angle(E, \tilde{E}) = \arccos\left(\frac{|\langle n, \tilde{n} \rangle|}{\|n\| \|\tilde{n}\|}\right). \quad (2.14)$$

**Beispiel 2.73:** Wir betrachten die Ebenen, die durch

$$2x_1 - x_2 + x_3 = 0 \quad \text{und} \quad x_1 + 2x_2 - x_3 = 1.$$

gegeben sind. Normalenvektoren lassen sich ablesen zu

$$n = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \tilde{n} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

Dadurch ergibt sich nach Formel (2.14)

$$\cos(\angle(E, \tilde{E})) = \frac{|\langle n, \tilde{n} \rangle|}{\|n\| \|\tilde{n}\|} = \frac{1}{6} \Rightarrow \angle(E, \tilde{E}) \approx 80,4^\circ.$$

## 2.2.4 Schnittmengen zwischen Geraden und Ebenen im $\mathbb{R}^2$ und $\mathbb{R}^3$

In diesem Unterabschnitt wird ein zweckmäßiger Weg gezeigt, die Schnittmenge zwischen Geraden und Ebenen zu bestimmen. Es gibt darüber hinaus auch andere, gleichwertige Lösungswege. Wir werden Geraden in der Ebene und Ebenen im Raum gemeinsam behandeln, weil beide eine Normalform besitzen. In diesem Zusammenhang verwenden wir verkürzend den Begriff *Hyperebene*, wenn wir entweder eine Gerade in der Ebene oder eine Ebene im Raum meinen.

### 2.2.4.1 Schnittmenge einer Gerade oder Ebene mit einer Hyperebene

Wir starten mit zwei Gleichungen, die eine Gerade oder Ebene beschreiben und parameterbehaftet oder parameterlos sein können.

- Der erste Schritt ist, eine Gleichung in eine Parametergleichung und die andere Gleichung in eine parameterlose Gleichung umzuformen. Das geht immer, wenn eine Hyperebene im Spiel ist.<sup>3</sup> Nur bei zwei Geraden im  $\mathbb{R}^3$  muss man einen anderen Lösungsweg wählen.

<sup>3</sup>also eine Gerade im  $\mathbb{R}^2$  oder eine Ebene im  $\mathbb{R}^3$

- Im zweiten Schritt setzt man die Parametergleichung in die parameterlose ein und erhält den Wert eines Parameters.
- Diesen Wert setzt man in die Parametergleichung ein und erhält das Ergebnis.

**Beispiel 2.74:** Es sei

$$Q = (1, 1, 1) \quad \text{und} \quad P = (1, -1, 2).$$

Weiterhin sei

$$n = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Ermitteln Sie den Schnittpunkt der Geraden durch  $P$  in Richtung von  $n$  mit der Ebene durch  $Q$  senkrecht zu  $n$ .

Naheliegenderweise verwendet man bei der Gleichsetzung für die Ebene die parameterlose und für die Gerade die Parameterform. Die parametrische Gleichung der Geraden durch  $P$  in Richtung von  $n$  ist

$$x = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Also:

$$x = 1 + \gamma; \quad y = -1 + 2\gamma; \quad z = 2 + 3\gamma$$

Die Normalengleichung der Ebene ist:

$$\begin{aligned} \langle x, n \rangle &= \langle q, n \rangle \\ x + 2y + 3z &= 6. \end{aligned}$$

Einsetzen der parameterlosen Gleichung:

$$\begin{aligned} 1 + \gamma + 2(-1 + 2\gamma) + 3(2 + 3\gamma) &= 6 \\ 5 + 14\gamma &= 6 \\ \gamma &= \frac{1}{14} \end{aligned}$$

Das Ergebnis erhält man durch Einsetzen von  $\gamma$  in die Parametergleichung:

$$p + \gamma n = (1, -1, 2) + \frac{1}{14} (1, 2, 3) = \left( \frac{15}{14}, -\frac{12}{14}, \frac{31}{14} \right).$$

### 2.2.4.2 Bestimmung des Schnittpunkts zweier Geraden im $\mathbb{R}^3$

Dieser Sonderfall ist etwas schwieriger. Die beiden Geradengleichungen sind:

$$G1: x = p_1 + \lambda \cdot r_1; \quad G2: y = p_2 + \mu \cdot r_2$$

Der Weg besteht darin, die beiden Gleichung gleichzusetzen und aufzulösen:

$$p_1 + \lambda \cdot r_1 = p_2 + \mu \cdot r_2$$

Jetzt können generell drei Fälle auftreten:

1. Die Geraden haben einen Schnittpunkt. Das Gleichungssystem hat genau eine Lösung.
2. Die Geraden sind parallel. Man erhält entweder mehrere Lösungen (wenn die Geraden genau aufeinander liegen) oder keine Lösung.
3. Die Geraden haben weder einen Schnittpunkt, noch sind sie parallel. In diesem Fall spricht man von *windschiefen* Geraden. Das Gleichungssystem hat keine Lösung.

**Beispiel 2.75:** Finden Sie den Schnittpunkt der Geraden

$$G1: x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad G2: y = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \beta \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

falls ein solcher existiert.

Ansatz:

$$\begin{aligned} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \beta \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \\ \alpha \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} &= \begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix} \end{aligned}$$



Lösung des Gleichungssystems:

$$\begin{array}{r|l}
 -1 & -2 & -1 & \\
 1 & 1 & 0 & +\textcircled{1} \\
 1 & -1 & -2 & +\textcircled{1} \\
 \hline
 -1 & -2 & -1 & \cdot(-1) \\
 0 & -1 & -1 & \cdot(-1) \\
 0 & -3 & -3 & -3 \cdot \textcircled{2} \\
 \hline
 1 & 2 & 1 & -2 \cdot \textcircled{2} \\
 0 & 1 & 1 & \\
 0 & 0 & 0 & \\
 \hline
 1 & 0 & -1 & \\
 0 & 1 & 1 & 
 \end{array}$$

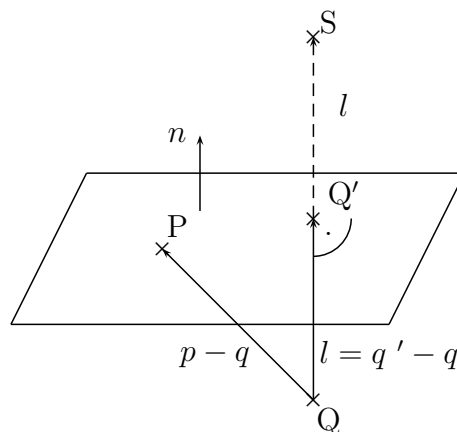
Es wird nur eine der beiden Variablen  $\alpha$  oder  $\beta$  gebraucht. Mit  $\beta = 1$  ergibt sich als Lösung

$$s = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

## 2.2.5 Abstandsbestimmung im $\mathbb{R}^2$ und $\mathbb{R}^3$

### 2.2.5.1 Einführung

**Satz 2.76:** Der kürzeste Verbindungsvektor zwischen einem Punkt, einer Geraden oder einer Ebene stößt immer senkrecht auf die Gerade oder die Ebene.



**BEWEIS** Der Satz wird für die Verbindung zwischen Punkt und Ebene gezeigt. Für andere Kombinationen gilt er analog. Ein Verbindungsvektor  $v$ , der nicht senkrecht auf die Ebene stößt, kann zusammengesetzt werden

aus einer Komponente parallel zur Ebene  $v_p$  und einer Komponente  $v_s$  senkrecht zur Ebene. Da nach Pythagoras

$$\|v\|^2 = \|v_p\|^2 + \|v_s\|^2$$

gilt, kann  $v_s$  nicht länger als  $v$  sein. ■

**Definition 2.77:** Der Punkt  $Q'$  ist derjenige Punkt in der Ebene, der den kürzesten Abstand zu Punkt  $Q$  besitzt. Er wird *Lotfußpunkt* genannt. Der Punkt  $S$  wird *Spiegelpunkt* genannt, da man sich in  $Q$  stehend dort sehen würde, falls die Ebene ein Spiegel ist. Für Geraden anstatt Ebenen sind die Definitionen äquivalent.

Aus Satz 2.76 kann man ein Verfahren zur Abstandsbestimmung herleiten, das für fast alle Kombinationen aus Punkten, Geraden und Ebenen im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  funktioniert. Die Ausnahmen sind:

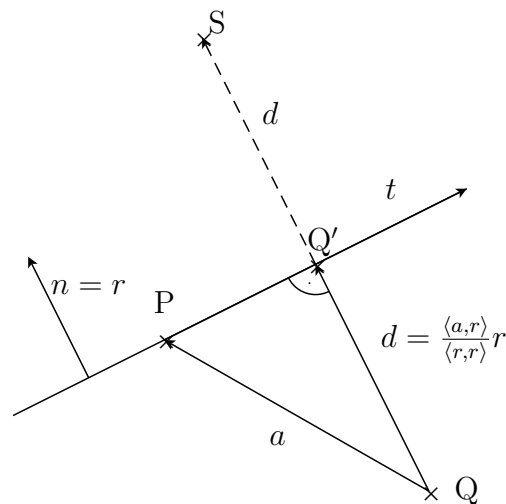
- Der Abstand zwischen zwei Punkten. Hier ergibt sich die Berechnung trivialerweise durch  $\|a - b\|$ .
- Der Abstand zwischen einer Gerade und einem Punkt im  $\mathbb{R}^3$ . Hierzu muss man einen zweiten Rechenweg kennen, der im Anschluss erläutert wird.
- Der Abstand zwischen zwei parallelen Geraden im  $\mathbb{R}^3$  lässt sich auf den Abstand zwischen einem Punkt und einer Geraden zurückführen.

Der Rechenweg:

- Bestimme die Richtung  $r$  des kürzesten Abstandsvektors.
- Bestimme jeweils einen Punkt auf den beiden Objekten und (durch Differenzbildung) den Abstandsvektor  $a$  zwischen den beiden Punkten.
- Der kürzeste Abstandsvektor  $d$  ist die Projektion von  $a$  auf  $r$ . Der kürzeste Abstand ist  $\|d\|$ .

$$d = \frac{\langle a, r \rangle}{\langle r, r \rangle} r; \quad \|d\| = \frac{|\langle a, r \rangle|}{\langle r, r \rangle} \|r\| = \frac{|\langle a, r \rangle|}{\|r\|}$$

Dieser Zusammenhang ist anhand des Abstandes zwischen einem Punkt und einer Geraden im  $\mathbb{R}^2$  an der folgenden Grafik erläutert:



Die einzelnen Fälle unterscheiden sich nur, wie man  $r$  erhält.

- Ist eine Hyperebene im Spiel (Gerade im  $\mathbb{R}^2$  oder Ebene im  $\mathbb{R}^3$ ), dann ist  $r$  gleich dem Normalenvektor  $n$  der Hyperebene.
- Bei zwei Geraden im  $\mathbb{R}^3$  muss  $r$  senkrecht auf beiden Geraden stehen. Daraus ergibt sich  $r$  als Vektorprodukt der beiden Richtungsvektoren. Bei zwei parallelen Geraden funktioniert das nicht, weil das Vektorprodukt 0 ergibt. Hier muss man den zweiten Lösungsweg wählen.

**Beispiel 2.78:** Gegeben ist im  $\mathbb{R}^2$  die Gerade  $g$  und der Punkt  $P = (2, 0)$  mit

$$g : x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Wie groß ist der minimale Abstand zwischen  $g$  und  $P$ ?

Aus dem Richtungsvektor der Geraden  $t = (2, 1)$  folgt der Normalenvektor  $n = (1, -2) = r$ . Ein Abstandsvektor ergibt sich aus

$$a = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Der minimale Abstand ist also:

$$d = \frac{|\langle a, r \rangle|}{\|r\|} = \frac{1}{\sqrt{5}}$$

**Beispiel 2.79:** Gegeben ist im  $\mathbb{R}^3$  der Punkt  $Q = (1, 3, 5)$  und die Ebene

$$-x + y - z = -5.$$

Wie groß ist der minimale Abstand zwischen Punkt und Ebene?

Ein Punkt  $P$  auf der Ebene ist  $P = (5, 0, 0)$ . Damit ist  $a = Q - P = (-4, 3, 5)$ .

Die Richtung des kürzesten Abstands ist gleich dem Normalenvektor der Ebene  $r = n = (-1, 1, -1)$ . Damit ist

$$d = \frac{|\langle a, r \rangle|}{\|r\|} = \frac{|4 + 3 - 5|}{\sqrt{3}} = \frac{2}{\sqrt{3}}$$

Wenn nur der Abstand einer Ebene vom Nullpunkt gesucht ist, dann erhält man die einfachere Formel

$$d = \frac{|\langle p, n \rangle|}{\|n\|},$$

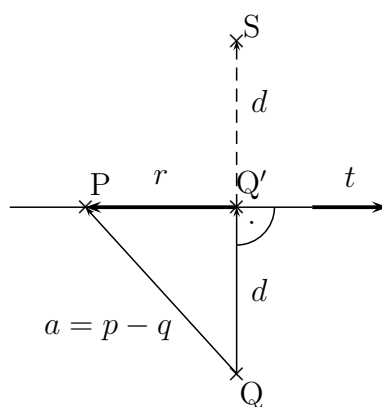
also die Länge der Projektion des Ortsvektors  $p$  zum Punkt  $P$  auf der Ebene mit dem Normalenvektor  $n$ . Der Vergleich mit der Hesseschen Normalform

$$\frac{\langle x, n \rangle}{\|n\|} = \frac{\langle p, n \rangle}{\|n\|}$$

zeigt, dass der Betrag der rechten Seite der Hesseschen Normalform gerade der Abstand der Ebene zum Nullpunkt ist.

### 2.2.5.2 Der Abstand eines Punktes von einer Geraden im $\mathbb{R}^3$

Die Methode aus dem vorigen Abschnitt funktioniert im vorliegenden Fall *nicht*, da die Gerade im  $\mathbb{R}^3$  keine Hyperebene, bzw.  $n$  nicht eindeutig ist. Folgender Ansatz ist recht einfach zu merken:



Man verfolgt nun folgenden Ansatz zur Bestimmung des Abstandes:

1.  $a = d + r$ , also  $d = a - r$ .

2.  $r$  ist die Projektion von  $a$  auf den Richtungsvektor der Geraden  $t$ , also

$$r = \frac{\langle a, t \rangle}{\langle t, t \rangle} t.$$

3. Zusammen ergibt sich:

$$d = a - \frac{\langle a, t \rangle}{\langle t, t \rangle} t.$$

**Beispiel 2.80:** Gegeben sind:

$$g : x = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

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

Damit ist  $t = (2, 2, 1)$  und  $a = (3, 5, 2)$ .  $d$  und  $\|d\|$  errechnen sich zu

$$d = a - \frac{\langle a, t \rangle}{\langle t, t \rangle} t$$

$$= \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} - \frac{6 + 10 + 2}{4 + 4 + 1} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\|d\| = \sqrt{2}$$

Damit ergibt sich für den Lotfußpunkt:

$$q' = q + d = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}.$$

## 2.3 Die Determinante im $\mathbb{R}^2$ und $\mathbb{R}^3$

### 2.3.1 Berechnung und geometrische Deutung

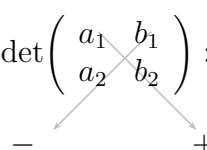
Wir stellen nun unsystematisch den Begriff der *Determinante* für die Spezialfälle von  $(1 \times 1)$ ,  $(2 \times 2)$ - und  $(3 \times 3)$ -Matrizen vor, weil es sich bei der Determinante auch um ein nützliches Rechenwerkzeug handelt, das die Berechnungen im Rahmen der analytischen Geometrie vereinfachen kann. Wir definieren zu einer quadratischen Matrix  $A$  die *Determinante*  $\det(A)$  oder kürzer  $|A|$  als reelle Zahl durch Angabe von Rechenvorschriften:

**Definition 2.81:** Sei  $A \in \mathbb{R}^{n \times n}$ .

**n=1:**  $A = (a_1)$ . Dann gilt

$$\det(A) := a_1$$

**n=2:** Sei  $A = (a, b)$  mit den Spaltenvektoren  $a, b \in \mathbb{R}^2$ . Dann gilt

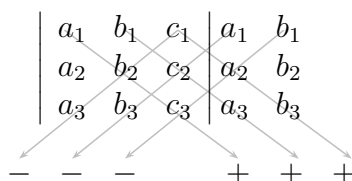
$$\det A = \det(a, b) = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} := a_1 b_2 - b_1 a_2$$


Entlang der Diagonalen wird das Produkt gebildet und entsprechend dem angezeigten Vorzeichen aufsummiert.

**n=3:**

$$\begin{aligned} \det(a, b, c) &= \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \\ &= a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - a_1 c_2 b_3 - b_1 a_2 c_3 - c_1 b_2 a_3 \end{aligned}$$

Die Formel für  $n = 3$  kann durch die Berechnungsmethode nach *Sarrus*<sup>4</sup> veranschaulicht werden:



Entlang der Diagonalen werden Produkte gebildet und diese mit den abgebildeten Vorzeichen versehen aufsummiert.

Wir deuten nun die Determinante zunächst für den Fall  $n = 2$  geometrisch. Man hat

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc = \left\| \begin{pmatrix} 0 \\ 0 \\ ad - bc \end{pmatrix} \right\| = \left\| \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \times \begin{pmatrix} c \\ d \\ 0 \end{pmatrix} \right\|.$$

Nach Definition 2.38 gibt die euklidische Norm des Kreuzprodukts zweier Vektoren im  $\mathbb{R}^3$  aber genau den Flächeninhalt des aufgespannten Parallelogramms an. Somit entspricht der Betrag der Determinante genau der Fläche des von  $a$  und  $b$  aufgespannten Parallelogramms.

Zur geometrischen Deutung der Determinante für  $n = 3$  wird folgende Definition benötigt.

<sup>4</sup>Pierre Frédéric Sarrus (1798-1861), franz. Mathematiker

**Definition 2.82:** Für drei Vektoren  $a, b, c \in \mathbb{R}^3$  nennt man

$$\langle a, b \times c \rangle \in \mathbb{R}$$

das *Spatprodukt* der drei Vektoren  $a, b, c$ .

**Satz 2.83:** Es gilt im  $\mathbb{R}^3$ :

$$\det(a, b, c) = \|a\| \cdot \|b\| \cdot \|c\| \sin \phi \cos \psi \quad (2.15)$$

BEWEIS Seien  $a((a_1, a_2, a_3)^T)$ ,  $b$  und  $c$  entsprechend. Nach Definition 2.81 der Determinante gilt dann

$$\begin{aligned} \det(a, b, c) &= a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 \\ &\quad - a_3 b_2 c_1 - b_3 c_2 a_1 - c_3 a_2 b_1 \\ &= c_1(a_2 b_3 - a_3 b_2) + c_2(a_3 b_1 - a_1 b_3) + c_3(a_1 b_2 - a_2 b_1) \\ &= \left\langle c, \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \right\rangle \end{aligned}$$

Mit der Definition des Kreuzprodukts folgt daraus

$$\begin{aligned} \det(a, b, c) &= \langle c, a \times b \rangle \\ &= \|a \times b\| \cdot \|c\| \cos \psi \\ &= \|a\| \cdot \|b\| \cdot \|c\| \sin \phi \cos \psi \end{aligned}$$

Die Vektoren  $a, b, c$  bilden die Kanten eines Körpers im dreidimensionalen Raum, eines *Parallelepipeds* oder *Spats* (Abb. 2.8). Es entspricht also nach Formel (2.15) der Betrag der Determinante dem Volumen des durch die drei Spaltenvektoren aufgespannten Spats.

**Bemerkung 2.84:** Das Vorzeichen der Determinante zeigt die Orientiertheit der drei Spaltenvektoren: ist die Determinante positiv, handelt es sich um ein Rechtssystem, ist sie negativ, um ein Linkssystem. Daher spricht man in diesem Zusammenhang auch vom *orientierten Volumen* eines Spats.

**Satz 2.85:** Die Determinante hat folgende Eigenschaften.

**D1** :  $\det(a, b, c) = \det(c, a, b) = \det(b, c, a)$ .

**D2** :  $\det(a, b, c) = -\det(b, a, c)$ .

**D3** :  $\det(a, a, c) = 0$ .

**D4** : Für  $\alpha \in \mathbb{R}$  gilt  $\det(\alpha \cdot a, b, c) = \alpha \cdot \det(a, b, c)$ .

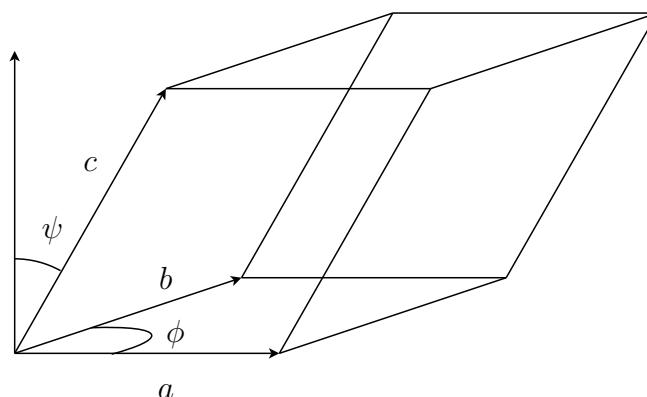


Abbildung 2.8: 3 Vektoren spannen einen Spat auf

$$\mathbf{D5} : \det(a, b, c + d) = \det(a, b, c) + \det(a, b, d).$$

$$\mathbf{D6} : \det(A) = \det(A^T)$$

**BEWEIS** Im Fall  $n = 2$ : direktes Nachrechnen, im Fall  $n = 3$  verwendet man Satz 2.83 zusammen mit den Eigenschaften von Kreuz- und Skalarprodukt. ■

Erläuterungen, Beispiele und Folgerungen:

- Beispiel zu **D4**:

$$\det \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & -1 \\ 5 & 6 & 2 \end{pmatrix} = 2 \cdot \det \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2 & -1 \\ 5 & 3 & 2 \end{pmatrix}$$

- Folgerung aus **D4**:

$$\det(0, b, c) = 0 \cdot \det(0, b, c) = 0$$

Ist in einer Matrix eine Spalte (oder eine Zeile, siehe **D6**) gleich 0, dann ist auch deren Determinante gleich 0.

- Beispiel zu **D5**:

$$\det \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & -1 \\ 5 & 6 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 0 \\ 3 & 1 & -1 \\ 5 & 1 & 2 \end{pmatrix} + \det \begin{pmatrix} 1 & 1 & 0 \\ 3 & 3 & -1 \\ 5 & 5 & 2 \end{pmatrix}$$



**Bemerkung 2.86:** Es gilt:

$$\begin{aligned} \det(a, b, c + \alpha a) &\stackrel{D5}{=} \det(a, b, c) + \det(a, b, \alpha a) \\ &\stackrel{D4}{=} \det(a, b, c) + \alpha \cdot \det(a, b, a) \\ &\stackrel{D1}{=} \det(a, b, c) + \alpha \cdot \det(a, a, b) \\ &\stackrel{D3}{=} \det(a, b, c) \end{aligned}$$

### 2.3.2 Lineare $3 \times 3$ -Gleichungssysteme

Eine Anwendung von Determinanten ist die Untersuchung von linearen Gleichungssystemen auf eindeutige Lösbarkeit. Durch

$$(LG) \begin{cases} (1) a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 = d_1 \\ (2) b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 = d_2 \\ (3) c_1 \cdot x_1 + c_2 \cdot x_2 + c_3 \cdot x_3 = d_3 \end{cases}$$

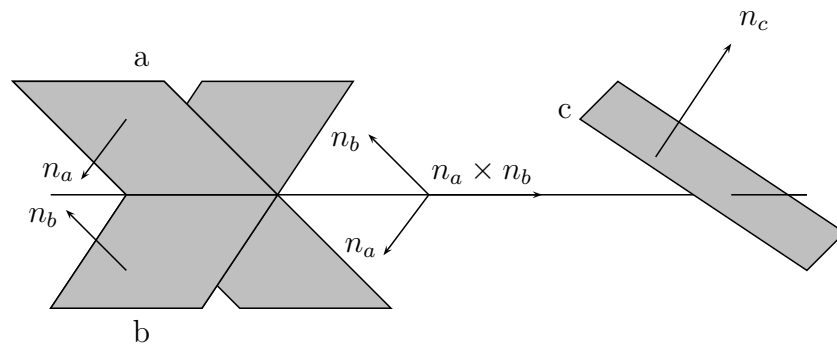
ist ein *lineares Gleichungssystem* mit den Unbekannten  $(x_1, x_2, x_3)$  gegeben. Jede Gleichung ist bekanntlich eine Ebene im dreidimensionalen Raum. Eindeutige Lösbarkeit bedeutet also die Eindeutigkeit des Schnittpunkts der drei Ebenen. Der Vektor

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

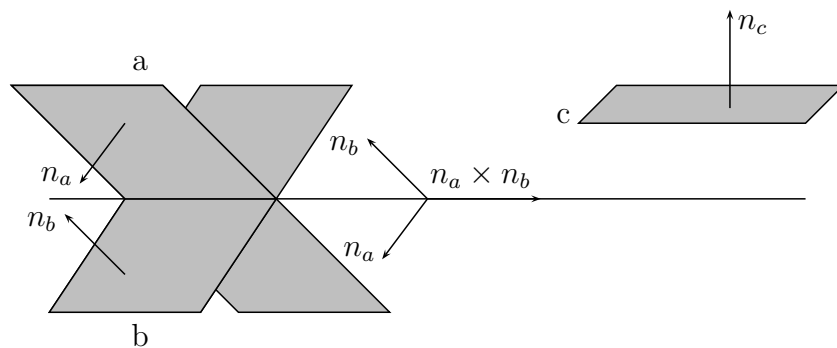
ist der Normalenvektor der Ebene (1). Entsprechend hat man die Normalenvektoren  $b$  und  $c$  der Ebenen (2) und (3). Bei einem eindeutigen Schnittpunkt dürfen die Ebenen (1) und (2) nicht parallel sein, d. h. es muss gelten  $a \times b \neq 0$ . Der Vektor  $a \times b$  zeigt in Richtung der Schnittgeraden der beiden Ebenen (1) und (2). Die Schnittgerade darf nicht parallel zur dritten Ebene, d. h. nicht senkrecht zum Normalenvektor  $c$ , sein. Es muss also gelten

$$\langle (a \times b), c \rangle = \det(a, b, c) \neq 0.$$

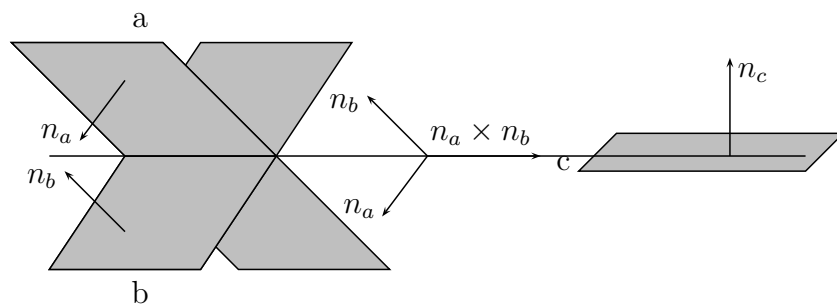
Damit ist die Bedingung für die eindeutige Lösbarkeit des obigen Gleichungssystems gegeben.



Eindeutiger Schnittpunkt:  $\langle n_a \times n_b, n_c \rangle \neq 0$



Kein Schnittpunkt:  $\langle n_a \times n_b, n_c \rangle = 0$



Schnittgerade, kein eindeutiger Schnittpunkt:  $\langle n_a \times n_b, n_c \rangle = 0$

**Satz 2.87:** Notwendig und hinreichend dafür, dass das lineare Gleichungssystem  $LG$  eine eindeutige Lösung besitzt, ist die Bedingung

$$\det(a, b, c) \neq 0.$$

# Kapitel 3

## Algebraische Strukturen

### 3.1 Gruppen

Die Mathematik ist traditionell eine Strukturwissenschaft, d.h. sie ersinnt abstrakte Objekte und untersucht dann ihre Eigenschaften. Dieses Vorgehen prägt insbesondere die Algebra, die als Teil der reinen Mathematik algebraische Strukturen betrachtet. Die Resultate der Algebra haben aber, obwohl ursprünglich aus intellektuelle Neugier heraus erzielt, Eingang in die gesamte Mathematik gefunden und auch die angewandte Mathematik entscheidend beeinflusst. Daher gehen wir hier kurz auf einige wesentliche algebraische Strukturen ein. Eine der einfachsten algebraischen Strukturen ist die aus der Vorlesung Mathematische Grundlagen bekannte *Gruppe*.

**Definition 3.1:**

1. Sei  $M$  eine Menge, und  $\circ : M \times M \rightarrow M$  eine Abbildung,  $(x, y) \rightarrow \circ(x, y)$ . Eine solche Abbildung heie *Verknpfung*; wir schreiben zur Abkrzung statt  $\circ(x, y)$  einfacher  $x \circ y$ .

2. Fr  $M$  und  $\circ$  wie oben heit das Paar  $(M, \circ)$  eine *Gruppe*, wenn gilt:

**G1** (Assoziativitt): Die Verknpfung ist assoziativ, d.h. es gilt:

$$\forall x, y, z \in M : (x \circ y) \circ z = x \circ (y \circ z)$$

**G2** (Neutralement): Es existiert ein neutrales Element  $e \in M$  so dass

$$\forall x \in M : x \circ e = x \tag{3.1}$$

**G3** (Inverses Element):  $\forall x \in M \exists x'$  mit

$$x \circ x' = e.$$

3. Gilt für eine Gruppe  $G = (M, \circ)$ , dass  $x \circ y = y \circ x \forall x, y \in M$ , dann heißt  $G$  *abelsche Gruppe* oder *kommutative Gruppe*.

**Bemerkung 3.2:** Zum Nachweis, dass  $(M, \circ)$  eine Gruppe ist, gehört implizit auch der Nachweis, dass  $\circ$  tatsächlich eine Verknüpfung im Sinne der Definition 3.1 ist. Man muss also zeigen, dass

1.  $x \circ y$  für alle  $x, y$  existiert und eindeutig festgelegt ist.
2. das Ergebnis  $x \circ y$  in  $M$  liegt.

Ist dies der Fall, sagt man, die Abbildung  $\circ : M \times M \rightarrow M$  sei *wohldefiniert*, also sinnvoll definiert.

**Beispiel 3.3:**

1.  $(\mathbb{Z}, +)$ ,  $(\mathbb{Q}, +)$  und  $(\mathbb{R}, +)$  sind Gruppen.
2. Die Menge  $M = \{-5, -4, \dots, 0, \dots, 4, 5\}$  bildet zusammen mit der gewöhnlichen Addition *keine* Gruppe. Natürlich erfüllt  $(M, +)$  die Bedingungen **G1** - **G3** aus Definition 3.1, aber aufgrund z. B.  $4 + 4 = 8 \notin M$  ist "+" *keine* Verknüpfung auf  $M$ , und daher kann  $(M, +)$  keine Gruppe bilden.

**Satz 3.4:** Sei  $G = (M, \circ)$  eine Gruppe. Dann gilt

1.  $M \neq \emptyset$
2. Das inverse Element kommutiert, d.h.  $\forall x \in M : x' \circ x = x \circ x'$ .
3. Das neutrale Element kommutiert, also  $\forall x \in M : x \circ e = e \circ x$ .
4. Das inverse Element ist eindeutig, d.h.  $\forall x$  existiert genau ein  $x' \in M : x' \circ x = e$ .
5. Das neutrale Element  $e$  ist eindeutig.

**BEWEIS**

1. Nach G2 gilt  $e \in M$ .
2. Es gilt:  $x \circ x' = e$ , d.h.  $x'$  ist ein inverses Element von  $x$ . Auch  $x'$  hat nach G3 ein inverses Element  $z$ :  $x' \circ z = e$ . Dann ist

$$\begin{aligned}
 e &= x' \circ z \\
 &= (x' \circ e) \circ z \\
 &= x' \circ (e \circ z) \\
 &= x' \circ ((x \circ x') \circ z) \\
 &= x' \circ (x \circ (x' \circ z)) \\
 &= x' \circ (x \circ e) \\
 &= x' \circ x
 \end{aligned}$$

3. Nach obiger Rechnung gilt

$$\begin{aligned} e \circ x &= (x \circ x') \circ x \\ &= x \circ (x' \circ x) \\ &= x \circ e \\ &= x \end{aligned}$$

4. Es seien  $y$  und  $y'$  zwei inverse Elemente zu  $x$ . Dann folgt:

$$y' = y' \circ e = y' \circ (x \circ y) = (y' \circ x) \circ y = e \circ y = y$$

5. Für ein weiteres neutrales Element  $e'$  folgt  $e' = e' \circ e = e$  nach G2. ■

**Bemerkung 3.5:**

1. Da das inverse Element zu  $a$  eindeutig ist, spricht man von *dem* inversen Element und bezeichnet es mit  $a^{-1}$ .
2. Für eine Gruppe  $G = (M, \circ)$  schreibt man oft kurz  $a \in G$ , wenn man eigentlich  $a \in M$  meint. Ist die Verknüpfung aus dem Zusammenhang klar, identifiziert man häufig  $M$  und  $G$ .

**Beispiel 3.6:**

1. Die *triviale Gruppe*  $(\{1\}, \cdot)$  ist die kleinste aller Gruppen.
2.  $(\mathbb{N}, +)$  ist keine Gruppe, weil z. B. zur 1 das additiv Inverse  $-1$  keine natürliche Zahl ist. Gäbe es ein weiteres Inverses in  $\mathbb{N}$ , wäre dies auch Inverses von 1 in  $\mathbb{Z}$ . Dies widerspräche der Eindeutigkeit des inversen Elements in  $\mathbb{Z}$ .
3. Der Zahlenraum  $\mathbb{R}^n$  bildet mit der Vektoraddition eine Gruppe (Übungsaufgabe!).
4.  $(\mathbb{R}, \cdot)$  mit der gewöhnlichen Multiplikation ist keine Gruppe, weil es zu 0 keine reelle Zahl  $a$  gibt mit  $0a = 1$ , also G3 nicht erfüllt ist.
5. Vektoren im  $\mathbb{R}^3$  bilden bzgl. des Vektorproduktes keine Gruppe, weil das Vektorprodukt nach Bemerkung 2.42 nicht assoziativ ist.

**Bemerkung 3.7:**  $(\mathbb{R} \setminus \{0\}, \cdot)$  ist eine abelsche Gruppe.

**BEWEIS** Wir zeigen G1 - G3. G1 ist klar, ebenso G2 mit  $e = 1$ . Mit  $a^{-1} = 1/a$  gilt G3. Man beachte, dass  $a \neq 0$  wegen  $a \in \mathbb{R} \setminus \{0\}$  n. V. ■

$\oplus$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

$\oplus$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

Abbildung 3.1: Verknüpfungstabellen von  $(\mathbb{Z}/2\mathbb{Z}, \oplus)$  (links) und  $(\mathbb{Z}/4\mathbb{Z}, \oplus)$  (rechts)

Wir wiederholen sehr kurz einige Inhalte der Vorlesung "Mathematische Grundlagen", um sie im Kontext von Gruppen erneut zu betrachten. Für  $m \in \mathbb{N}$  definiert man  $\bar{k} = \{k + m\mathbb{Z}\}$ , die sog. Restklassen als Äquivalenzklassen der Äquivalenzrelation  $a \sim b :\Leftrightarrow (a - b) \bmod m = 0$  und erhält eine disjunkte Zerlegung von  $\mathbb{Z}$ :  $\mathbb{Z} = \bigcup_{k=0}^{m-1} \bar{k}$ . Auf der Menge  $\mathbb{Z}/m\mathbb{Z} := \{\bar{0}, \dots, \overline{m-1}\}$  wird dann die Verknüpfung  $\oplus$  definiert durch  $\bar{k} \oplus \bar{l} := \overline{k+l}$ . Dann gilt:

**Satz 3.8:** Für  $m \in \mathbb{N}$  ist  $(\mathbb{Z}/m\mathbb{Z}, \oplus)$  eine abelsche Gruppe.

**BEWEIS** Dass  $\oplus$  wohldefiniert ist, ist aus der Vorlesung Mathematische Grundlagen bekannt. Wir zeigen die drei Bedingungen G1 -G3.

G1 : Für  $a, b, c \in \mathbb{Z}/m\mathbb{Z}$  gilt

$$\begin{aligned} (\bar{a} \oplus \bar{b}) \oplus \bar{c} &= \overline{a+b} \oplus \bar{c} = \overline{(a+b)+c} = \overline{a+(b+c)} \\ &= \bar{a} \oplus \overline{b+c} = \bar{a} \oplus (\bar{b} \oplus \bar{c}). \end{aligned}$$

G2 : Offenbar gilt  $\bar{a} \oplus \bar{0} = \overline{a+0} = \bar{a} \forall \bar{a} \in \mathbb{Z}/m\mathbb{Z}$ , also ist  $\bar{0}$  das neutrale Element.

G3 : Für  $\bar{a}$  gilt  $\bar{a} \oplus \overline{m-a} = \overline{a+(m-a)} = \bar{m} = \bar{0}$ , also ist  $\overline{m-a}$  das gesuchte inverse Element.

Weiterhin gilt  $\bar{a} \oplus \bar{b} = \overline{a+b} = \overline{b+a} = \bar{b} \oplus \bar{a}$ , also ist die Gruppe abelsch. ■

Endliche Gruppen werden häufig über ihre *Verknüpfungstafel* definiert, in der man das Ergebnis jeder möglichen Verknüpfung explizit angibt (vgl. Abb. 3.1 für die Gruppen  $\mathbb{Z}/4\mathbb{Z}$  und  $\mathbb{Z}/2\mathbb{Z}$ ). Für abelsche Gruppen ist die Verknüpfungstafel stets spiegelsymmetrisch zur Hauptdiagonalen.

Man sieht leicht, dass die Menge  $m\mathbb{Z} = \{\dots, -2m, -m, 0, m, 2m, \dots\}$  mit der gewöhnlichen Addition eine Gruppe bildet. Weil die Obermenge  $\mathbb{Z}$  mit derselben Addition auch eine Gruppe bildet, liegt hier eine Gruppe innerhalb einer Gruppe vor. Diese Beobachtung motiviert die folgende Definition.

$\oplus$	$\bar{0}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{2}$
$\bar{2}$	$\bar{2}$	$\bar{0}$

Abbildung 3.2: Verknüpfungstafel von  $(\mathbb{Z}_2, \oplus)$ 

**Definition 3.9:** Sei  $G = (M, \circ)$  eine Gruppe und  $M' \subseteq M$ . Bildet  $U = (M', \circ)$  eine Gruppe, so heißt  $U$  *Untergruppe* von  $G$ , Schreibweise  $U \leq G$ .

**Folgerung 3.10:** Für  $G$  und  $U$  wie oben gilt  $U \leq G$  genau dann, wenn gilt:

1.  $M' \neq \emptyset$
2.  $\forall a, b \in M' : a \circ b \in M'$
3.  $\forall a \in M' : a^{-1} \in M'$

**BEWEIS “ $\Rightarrow$ ”:** Ist  $U$  Gruppe, so gilt  $e \in U$ , also  $U \neq \emptyset$ . 2. folgt aus der Anforderung, dass  $\circ$  die Verknüpfung von  $U$  ist und deswegen von  $M' \times M'$  nach  $M'$  abbildet. Nach G3 gilt Eigenschaft 3.

**“ $\Leftarrow$ ”:** Aus 2. folgt analog zu oben die Wohldefiniiertheit der Verknüpfung; G1 gilt, weil es bereits in der Obermenge  $M$  gilt. Wegen 1. ex.  $a \in M'$ ; wegen 3. gilt dann  $a^{-1} \in M'$  und damit nach 2.  $e = a \circ a^{-1} \in M'$ , also G2. G3 folgt aus 3. ■

**Beispiel 3.11:** Offensichtlich gilt  $(m\mathbb{Z}, +) \leq (\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +)$ .

Auf der Suche nach Beispielen für Untergruppen betrachten wir erneut die Verknüpfungstafel von  $\mathbb{Z}/4\mathbb{Z}$  (Abb. 3.1). Es fällt auf, dass  $\mathbb{Z}_2 := (\{\bar{0}, \bar{2}\}, \oplus)$  eine Untergruppe bildet. Ihre Verknüpfungstafel (Abb. 3.2) gleicht bis auf Bezeichnung der Verknüpfungstafel von  $\mathbb{Z}/2\mathbb{Z}$  (Abb. 3.1 links). Die beiden Gruppen sind nicht gleich, weil es sich um verschiedene Restklassen handelt, aber besitzen doch die gleiche algebraische Struktur. Dieses Phänomen führt zu folgender Definition.

**Definition 3.12:** Zwei Gruppen  $G_1 = (M_1, \circ)$  und  $G_2 = (M_2, +)$  heißen *isomorph*, Schreibweise  $G_1 \simeq G_2$ , wenn es eine bijektive Abbildung  $\phi : G_1 \rightarrow G_2$  gibt mit

$$\phi(a \circ b) = \phi(a) + \phi(b) \forall a, b \in G_1. \quad (3.2)$$

**Bemerkung 3.13:** Der Begriff “isomorph” kommt aus dem Griechischen und bedeutet “gleichgestaltig”,  $\phi$  wird Gruppenisomorphismus genannt. Die Abbildung “übersetzt” die Elemente von  $G_1$  in die von  $G_2$  und respektiert dabei die Gruppenstruktur.

**Beispiel 3.14:**

1. Isomorphie ist eine Äquivalenzrelation.
2. Man hat  $\mathbb{Z}/2\mathbb{Z} \simeq \mathbb{Z}_2$  durch den elementweise definierten Gruppenisomorphismus  $\phi(\bar{0}_2) = \bar{0}_4, \phi(\bar{1}_2) = \bar{2}_4$ . Die Indizes an den Restklassen bezeichnen jeweils den Modul  $m$ ; man beachte  $\bar{0}_2 \neq \bar{0}_4$ . Es gilt also  $\mathbb{Z}/2\mathbb{Z} \simeq \mathbb{Z}_2 \leq \mathbb{Z}/4\mathbb{Z}$ .
3. Sowohl  $(\mathbb{R}^4, +)$  als auch  $(\mathbb{R}^2 \times \mathbb{R}^2, +)$  sind offenbar Gruppen, aber dennoch formal nicht gleich. Durch den Gruppenisomorphismus  $\varphi: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^4, \varphi((x_1, x_2), (x_3, x_4)) := (x_1, x_2, x_3, x_4)$  sind aber die beiden Gruppen isomorph. Daher werden wir künftig nicht zwischen  $\mathbb{R}^4$  und  $\mathbb{R}^2 \times \mathbb{R}^2$  o. ä. unterscheiden.
4. Es gilt  $\mathbb{C} \neq \mathbb{R}^2$ , wohl aber (als Gruppen)  $(\mathbb{C}, +) \simeq (\mathbb{R}^2, +)$  durch den Isomorphismus  $\varphi: \mathbb{C} \rightarrow \mathbb{R}^2, \varphi(a + ib) = (a, b)^T$ . Dies rechtfertigt die Veranschaulichung der komplexen Zahlen als komplexe Ebene.

## 3.2 Körper

Wir wollen nun ausgehend von den reellen Zahlen eine allgemeinere algebraische Struktur definieren und untersuchen. Auf den reellen Zahlen ist neben der Addition auch die Multiplikation erklärt. Das Tripel  $(\mathbb{R}, +, \cdot)$  hat folgende Eigenschaften:

1.  $(\mathbb{R}, +)$  bildet eine abelsche Gruppe.
2.  $(\mathbb{R} \setminus \{0\}, \cdot)$  bildet eine abelsche Gruppe (vgl. Bemerkung 3.7).
3. Es gelten die *Distributivgesetze*

$$\begin{aligned} \forall x, y, z \in \mathbb{R} : (x + y) \cdot z &= (x \cdot z) + (y \cdot z) \\ x \cdot (y + z) &= (x \cdot y) + (x \cdot z) \end{aligned}$$

Jedes Tripel  $(M, \oplus, \odot)$ , das die obigen drei Bedingungen erfüllt, wird *Körper* genannt:

**Definition 3.15:** Sei  $M$  eine nichtleere Menge.  $M$  zusammen mit zwei Verknüpfungen  $\oplus, \odot: M \times M \rightarrow M$  heißt *Körper*, wenn gilt:

1.  $(M, \oplus)$  bildet eine kommutative Gruppe.
2. Nach Ausschluss des neutralen Elements der Verknüpfung  $\oplus$  bilden die restlichen Elemente von  $M$  eine kommutative Gruppe bezüglich  $\odot$ .
3. Für die beiden Verknüpfungen  $\oplus$  und  $\odot$  gelten die *Distributivgesetze*:



- a)  $\forall x, y, z \in M \Rightarrow (x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$   
 b)  $\forall x, y, z \in M \Rightarrow z \odot (x \oplus y) = (z \odot x) \oplus (z \odot y)$

Das neutrale Element der Addition  $\oplus$  wird meist als „0“ bezeichnet (Null-Element), das neutrale Element der Multiplikation  $\odot$  meist als „1“ (Eins-Element). Im folgenden Satz werden die Elemente als  $\mathbb{0}$  und  $\mathbb{1}$  geschrieben, um eine Verwechslung mit den Zahlen 0 und 1 zu vermeiden.

Die Voraussetzung, dass die  $\mathbb{0}$  - das neutrale Element bzgl. der Addition - in der Eigenschaft 2. ausgeschlossen wird, ist notwendig aufgrund folgenden Satzes.

**Satz 3.16:** Sei  $K$  ein Körper. Es ist  $a \odot \mathbb{0} = \mathbb{0} \forall a \in K$ .

BEWEIS Es ist

$$a \odot (\mathbb{0} + \mathbb{0}) = a \odot \mathbb{0} + a \odot \mathbb{0} \quad \text{und} \quad a \odot (\mathbb{0} + \mathbb{0}) = a \odot \mathbb{0},$$

also

$$a \odot \mathbb{0} + a \odot \mathbb{0} = a \odot \mathbb{0} \quad \implies \quad a \odot \mathbb{0} = \mathbb{0}. \quad \blacksquare$$

Wenn also  $x = \mathbb{0}^{-1}$  existiert, dann muss gelten  $x \odot \mathbb{0} = \mathbb{1}$ . Es gilt aber  $x \odot \mathbb{0} = \mathbb{0}$ , also  $\mathbb{0} = \mathbb{1}$ . Dann gilt aber auch  $a = \mathbb{1} \odot a = \mathbb{0} \odot a = \mathbb{0}$ , d.h. ein solcher Körper hätte nur ein einziges Element 0 und die Operationen:

$$0 \oplus 0 = 0$$

$$0 \odot 0 = 0$$

Allerdings fordert Eigenschaft 2 eines Körpers, dass die neutralen Elemente bezüglich der Addition und der Multiplikation unterschiedlich sind. Jeder Körper hat also mindestens 2 Elemente. Somit ist  $\mathbb{0} = \mathbb{1}$  ausgeschlossen.

**Beispiel 3.17:**

1. Nach Konstruktion ist  $(\mathbb{R}, +, \cdot)$  ein Körper.
2.  $(\mathbb{Q}, +, \cdot)$  ist ein Körper, ebenso  $(\mathbb{C}, +, \cdot)$  mit der komplexen Multiplikation.
3.  $(\mathbb{Z}, +, \cdot)$  ist kein Körper, da  $(\mathbb{Z} \setminus \{0\}, \cdot)$  keine Gruppe ist, weil z. B.  $2^{-1} = \frac{1}{2} \notin \mathbb{Z}$ , also G3 nicht erfüllt ist.

Es liegt die Frage nahe, ob es analog zu endlichen Gruppen auch endliche Körper gibt. Der folgende Satz liefert eine partielle Antwort.

**Satz 3.18:** Wir definieren auf  $\mathbb{Z}/m\mathbb{Z}$  die Verknüpfung  $\odot$  durch

$$\bar{a} \odot \bar{b} := \overline{a \cdot b} \forall \bar{a}, \bar{b} \in \mathbb{Z}/m\mathbb{Z}.$$

Ist  $m$  eine Primzahl, dann ist  $(\mathbb{Z}/m\mathbb{Z}, \oplus, \odot)$  ein Körper.

Für den Beweis verweisen wir auf Beutelspacher [2]. Endliche Körper haben für die Lineare Algebra nur geringe Bedeutung, daher werden wir sie nicht weiter betrachten.

### 3.3 Vektorräume

In Kapitel 1.3 haben wir den Zahlraum  $\mathbb{R}^n$  definiert und als Menge aller Vektoren charakterisiert. Wir werden nun den Zahlraum zum Vektorraum verallgemeinern, indem wir in Analogie zur Gruppe und zum Körper die wesentlichen algebraischen Eigenschaften des  $\mathbb{R}^n$  sammeln und *jede* Struktur, die diese Eigenschaften hat, *Vektorraum* nennen.

**Definition 3.19:** Es sei  $K$  ein beliebiger Körper. Eine nicht-leere Menge  $V$  zusammen mit den beiden Abbildungen

$$\begin{aligned} \oplus : V \times V &\rightarrow V, & (x, y) &\mapsto x \oplus y \in V \\ \text{und } \odot : K \times V &\rightarrow V, & (\lambda, x) &\mapsto \lambda \odot x \in V \end{aligned}$$

heißt *Vektorraum über  $K$* , wenn folgende Axiome gelten:

1.  $(V, \oplus)$  ist eine kommutative Gruppe.
2.  $\forall \lambda, \mu \in K$  und  $x \in V$  gilt  $\lambda \odot (\mu \odot x) = (\lambda\mu) \odot x$ , wobei mit  $\lambda\mu$  die Multiplikation aus  $K$  gemeint ist.
3.  $\forall x \in V$  gilt  $\mathbb{1} \odot x = x$  ( $\mathbb{1}$  ist das neutrale Element der Multiplikation aus  $K$ ).
4.  $\forall \lambda \in K, x, y \in V$  gilt:  $\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y)$ .
5.  $\forall \lambda, \mu \in K, x \in V$  gilt  $(\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x)$ .

Üblicherweise geht man von  $\mathbb{R}$  oder  $\mathbb{C}$  als Grundkörper aus. Im Fall  $K = \mathbb{R}$  heißt  $V$  *reeller Vektorraum*, für  $K = \mathbb{C}$  heißt  $V$  *komplexer Vektorraum*. Allgemein spricht man von einem *Vektorraum über dem Körper  $K$* . Wie bei Gruppen auch identifizieren wir  $(V, \oplus, \odot)$  mit  $V$ , wenn die Verknüpfungen und der Körper aus dem Zusammenhang heraus klar sind.

**Definition 3.20:** Die Elemente der Menge  $V$  eines Vektorraums heißen *Vektoren*. Die Elemente aus  $K$  heißen *Skalare*.

**Bemerkung 3.21:** Im Vektorraum muss die Multiplikation eines Skalars mit einem Vektor definiert sein, nicht aber die Multiplikation zweier Vektoren.

**Beispiel 3.22:**

1.  $\mathbb{R}^n$  ist nach Konstruktion ein reeller Vektorraum mit der Vektoraddition gemäß Definition 1.6 und der Multiplikation eines Vektors mit einem Skalar nach Definition 1.7.
2.  $(\mathbb{C}^n, +, \cdot)$  ist ein komplexer Vektorraum mit der Addition von Vektoren und der Multiplikation mit einem (komplexen) Skalar wie oben.

3. Jeder Körper ist ein Vektorraum über sich selbst mit den Verknüpfungen des Körpers  $\oplus$  als Vektoraddition und  $\odot$  als Multiplikation eines Vektors mit einem Skalar.

**Beispiel 3.23:** Sei  $K = \mathbb{R}$  und  $V = \mathcal{C}[a, b]$  die Menge der auf dem Intervall  $[a, b]$  definierten reellen stetigen Funktionen. Wie in der Analysis üblich werden Summe und skalares Vielfaches von  $f, g \in \mathcal{C}[a, b]$  punktweise definiert:

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

Dann ist  $V$  ein reeller Vektorraum. Zum Nachweis nutzt man die aus der Analysis bekannte Tatsache, dass Summen und Vielfache stetiger Funktionen wieder stetig sind und führt alle anderen Axiome auf die entsprechenden Aussagen für reelle Zahlen zurück.

**Beispiel 3.24:** Auf der Menge der  $m \times n$ -Matrizen mit Elementen im Körper  $K$  lässt sich analog zur Addition von Vektoren eine komponentenweise Addition definieren: Für  $A = (a_{ij})$  und  $B = (b_{ij})$  sei  $A + B := (a_{ij} + b_{ij})$ . Analog definiert man die Multiplikation eines Skalars mit einer Matrix  $A$ :  $\lambda A := (\lambda a_{ij})$ . Mit diesen Verknüpfungen wird die Menge der  $m \times n$ -Matrizen zu einem  $K$ -Vektorraum und wird mit  $M(m \times n, K)$  bezeichnet. Der Vektorraum unterscheidet sich nur durch die Schreibweise seiner Elemente im Rechteck vom Vektorraum  $K^{m \cdot n}$ . Man rechnet die Vektorraumaxiome leicht nach.

**Beispiel 3.25:** Sei  $\mathcal{M}$  die Menge aller komplexwertigen Folgen  $(x_1, x_2, \dots)$  und für zwei Folgen  $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}}$  eine Addition erklärt durch  $(x_n)_{n \in \mathbb{N}} \oplus (y_n)_{n \in \mathbb{N}} := (x_n + y_n)_{n \in \mathbb{N}}$  sowie eine Multiplikation mit einem komplexen Skalar  $\lambda$  durch  $\lambda \odot (x_n)_{n \in \mathbb{N}} := (\lambda x_n)_{n \in \mathbb{N}}$ . Dann ist  $(\mathcal{M}, \oplus, \odot)$  ein komplexer Vektorraum.

Analog zur Gruppe innerhalb einer Gruppe, also einer Untergruppe, kann es Vektorräume innerhalb von Vektorräumen geben. Man spricht von *Untervektorräumen*.

**Definition 3.26:** Es sei  $V$  ein Vektorraum über einem Körper  $K$ . Eine Teilmenge  $U \subseteq V$  heißt *Untervektorraum* oder *Unterraum* von  $V$ , wenn  $U \neq \emptyset$  (leere Menge) und  $\forall x, y \in U$  und alle  $\lambda \in K$  gilt:

$$x \oplus y \in U, \lambda \odot x \in U.$$

**Bemerkung 3.27:** Es sei  $V$  ein Vektorraum.  $U \subseteq V$  ist genau dann ein Vektorraum, wenn  $U$  ein Untervektorraum von  $V$  in Sinne von Definition 3.26 ist.

**BEWEIS** Sei  $U$  ein Vektorraum. Weil  $(U, +)$  eine abelsche Gruppe bildet, gilt  $U \neq \emptyset$ . Die Abgeschlossenheit von  $U$  in  $V$  ist äquivalent zur Wohldefiniertheit der Vektoraddition und der Multiplikation mit einem Skalar im Vektorraum  $U$ .

Umgekehrt leiten sich aus  $U \subset V$  alle Vektorraumaxiome für  $U$  mit Ausnahme der Abgeschlossenheit ab. ■

**Beispiel 3.28:** Sei  $V = \mathbb{R}^2$  und  $U = \{(x, 0); x \in \mathbb{R}\} \subset V$ . Wir zeigen, dass  $U$  ein Unterraum von  $V$  ist und prüfen dazu die Kriterien der Definition 3.26. Es sei  $u_1 = (x_1, 0)$  und  $u_2 = (x_2, 0)$ . Dann gilt:

$$\begin{aligned} u_1 + u_2 &= (x_1 + x_2, 0) \in U \\ \lambda u_1 &= (\lambda x_1, 0) \in U, \end{aligned}$$

also ist  $U$  ein Unterraum von  $V$  und zugleich ist  $U$  nach Bemerkung 3.27 ein Vektorraum.

**Bemerkung 3.29:** Ist  $U$  ein Unterraum von  $V$ , dann ist der Nullvektor  $0$  von  $V$  und zu jedem  $x \in U$  auch  $-x$  in  $U$  enthalten. Aus der Definition eines Vektorraums folgt nämlich, dass  $0 \cdot x = 0$  und  $(-1) \cdot x = -x$  ist.

**Bemerkung 3.30:** Es sind  $\{0\}$  und  $V$  Unterräume von  $V$ .

**Bemerkung 3.31:** Sind  $U_1$  und  $U_2$  Unterräume von  $V$ , dann ist  $U_1 \cap U_2$  ein Unterraum von  $V$ .

**BEWEIS** Wir prüfen die Kriterien der Definition. Seien  $u, v \in U_1 \cap U_2$ . Dann gilt nach Definition  $u + v \in U_1$  und  $u + v \in U_2$ , da  $U_1$  und  $U_2$  Untervektorräume sind, also  $u + v \in U_1 \cap U_2$ . Weiter gilt für  $u \in U_1 \cap U_2$  und  $\lambda \in K$ , dass  $\lambda u \in U_1$  und  $\lambda u \in U_2$  und damit  $\lambda u \in U_1 \cap U_2$ . ■

Im Gegensatz zu Schnitten sind Vereinigungen von Untervektorräumen i. A. keine Untervektorräume.

**Beispiel 3.32:** Es sind

$$\begin{aligned} U_1 &= \{(x, 0); x \in \mathbb{R}\} \\ U_2 &= \{(0, y); y \in \mathbb{R}\} \end{aligned}$$

Unterräume des  $\mathbb{R}^2$ .  $U_v = U_1 \cup U_2$  ist die Teilmenge des  $\mathbb{R}^2$ , bei der mindestens eine der beiden Koordinaten gleich 0 ist. Mit  $x_1 = (1, 0)^T \in U_v$  und  $x_2 = (0, 1)^T \in U_v$  gilt aber  $x_1 + x_2 = (1, 1)^T \notin U_v$ , also ist  $U_v$  kein Untervektorraum.

Im Gegensatz zur Vereinigung von Mengen erhält die Summe von Untervektorräumen die Vektorraumeigenschaft.

**Definition 3.33:**

1. Es seien  $U_1, U_2$  Unterräume von  $V$ . Die *Summe* von  $U_1$  und  $U_2$  ist dann definiert durch

$$U_1 + U_2 := \{v \in V \mid \exists u_1 \in U_1, u_2 \in U_2 : v = u_1 + u_2\}.$$

2. ein Vektorraum  $W$  heißt *direkte Summe* von  $U_1$  und  $U_2$ , Schreibweise  $W = U_1 \oplus U_2$ , wenn gilt  $W = U_1 + U_2$  und  $U_1 \cap U_2 = \{0\}$ .

**Satz 3.34:** Sei  $V$  ein  $K$ -Vektorraum. Sind  $U_1$  und  $U_2$  Unterräume von  $V$ , ist  $U_1 + U_2$  ein Unterraum von  $V$ .

BEWEIS Übungsaufgabe ■

**Beispiel 3.35:**

$$U_1 = \{(x, 0); x \in \mathbb{R}\}$$

$$U_2 = \{(0, y); y \in \mathbb{R}\}$$

sind Unterräume des  $\mathbb{R}^2$ . Man hat

$$U_1 \oplus U_2 = \{(x + 0, 0 + y); x, y \in \mathbb{R}\} = \mathbb{R}^2$$

**Bemerkung 3.36:** Gilt  $V = U_1 \oplus U_2$ , ist die Darstellung  $v = u_1 + u_2$  mit  $u_1 \in U_1$  und  $u_2 \in U_2$  eindeutig für alle  $v \in V$ .

BEWEIS Sei  $v \in U_1 \oplus U_2$  mit  $v = u_1 + u_2 = w_1 + w_2$  mit  $u_1, w_1 \in U_1$  und  $u_2, w_2 \in U_2$ . Dann gilt  $0 = u_1 - w_1 + u_2 - w_2$ , also  $u_1 - w_1 = -(u_2 - w_2)$ . Aus der Abgeschlossenheit von  $U_1$  folgt  $u_1 - w_1 \in U_1$  und deswegen  $u_2 - w_2 \in U_1$ , aber aus der Abgeschlossenheit von  $U_2$  auch  $u_2 - w_2 \in U_2$ . Wegen  $U_1 \cap U_2 = \{0\}$  n. V. erhält man  $u_2 = w_2$  und damit dann  $u_1 = w_1$ . ■

## 3.4 Lineare Unabhängigkeit, Basis, Dimension

### 3.4.1 Lineare Unabhängigkeit

Man wird intuitiv die Ebene als zweidimensionales Objekt auffassen und damit dem Zahlraum  $\mathbb{R}^2$ , den man ja mit der Ebene identifizieren kann, die Dimension 2 zuordnen wollen. Analog liegt es nahe, dem Zahlraum  $\mathbb{R}^3$  die Dimension 3 zuzuordnen. Es stellt sich die Frage, ob und wie man einem allgemeinen Vektorraum eine Dimension sinnvoll zuordnen kann. Bis diese Frage beantwortet werden kann, ist einige Vorarbeit zu leisten. Wir beginnen mit der einfachen Beobachtung, dass im  $\mathbb{R}^3$  jeder

Vektor  $v = (x, y, z)^T$  eindeutig als gewichtete Summe der drei kanonischen Einheitsvektoren geschrieben werden kann:

$$v = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Wir verallgemeinern dies in folgender Definition.

**Definition 3.37:** Seien  $v_1, \dots, v_r \in V$  Vektoren eines  $K$ -Vektorraums  $V$ :

- (a) Ein Vektor  $v$ , der sich als Addition solcher Vektoren mit Vorfaktoren darstellen lässt,

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r \quad \text{mit} \quad \lambda_i \in K,$$

nennt man *Linearkombination* von  $v_1, \dots, v_r$ .

- (b) Die Menge

$$L(v_1, \dots, v_r) := \{ \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r \mid \lambda_i \in K \} \subseteq V$$

aller Linearkombinationen heißt *Lineare Hülle* der  $r$  Vektoren  $v_1, \dots, v_r$ .

Nach Definition ist jede Lineare Hülle ein Unterraum von  $V$ .

**Beispiel 3.38:** Gegeben seien die Vektoren

$$a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} .$$

Dann ist

$$c = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} = 1 \cdot a + (-1) \cdot b \in L(a, b)$$

eine Linearkombination von  $a$  und  $b$  und somit auch in der Linearen Hülle

$$L(a, b) = \{ \lambda_1 a + \lambda_2 b; \lambda_1, \lambda_2 \in \mathbb{R} \}$$

enthalten.

**Beispiel 3.39:** Beispiele im  $\mathbb{R}^3$ :

$$L(e_1, e_2, e_3) = \{ (\lambda_1, \lambda_2, \lambda_3); \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \} = \mathbb{R}^3$$

$$L(e_1, e_2) = \{ (\lambda_1, \lambda_2, 0); \lambda_1, \lambda_2 \in \mathbb{R} \} \quad (\text{x-y-Ebene})$$

$$L(e_1, e_2, e_1 + e_2) = \{ (\lambda_1 + \lambda_3, \lambda_2 + \lambda_3, 0); \lambda_1, \lambda_2 \in \mathbb{R} \} \quad (\text{x-y-Ebene})$$

Der englische Ausdruck für „Lineare Hülle“ ist *span*. Auch im Deutschen schreibt man oft  $\text{span}(e_1, e_2)$  statt  $L(e_1, e_2)$ . Es gibt auch den deutschen Ausdruck *Spann* für die lineare Hülle, der jedoch eher ungewöhnlich ist. Gängig ist aber ein Satz wie: Die Vektoren  $e_1$  und  $e_2$  *spannen die  $x$ - $y$ -Ebene auf*.

Beispiel 3.39 zeigt, dass man den gesamten Zahlraum  $\mathbb{R}^3$  aus drei Vektoren durch Linearkombination erzeugen kann, namentlich  $e_1, e_2$  und  $e_3$ . Dies führt zu folgender Definition.

**Definition 3.40:**

1. Sei  $V$  ein Vektorraum und  $(v_1, \dots, v_n)$  ein  $n$ -Tupel von Vektoren in  $V$ . Spannen die Vektoren den gesamten Raum auf, d.h. gilt  $L(v_1, \dots, v_n) = V$ , so nennt man  $(v_1, \dots, v_n)$  ein *Erzeugendensystem*.
2.  $V$  heißt *endlich erzeugt*, wenn es endlich viele Vektoren  $v_1, \dots, v_r$  gibt, so dass  $L(v_1, \dots, v_r) = V$ , ansonsten nicht endlich erzeugt.

**Beispiel 3.41:** Das Tripel  $(e_1, e_2, e_3)$  der kanonischen Einheitsvektoren ist ein Erzeugendensystem von  $\mathbb{R}^3$ ; ebenso erzeugen die Vektoren  $(0, 1)^T$  und  $(1, 0)^T$  den Zahlraum  $\mathbb{R}^2$ . Offenbar sind  $\mathbb{R}^2$  und  $\mathbb{R}^3$  endlich erzeugt. Ebenso ist der Zahlraum  $\mathbb{R}^n$  durch  $e_1, \dots, e_n$  endlich erzeugt.

Wir nehmen von jetzt an für das ganze Kapitel an, der Vektorraum  $V$  sei endlich erzeugt.

Man kann für einen Vektorraum viele Erzeugendensysteme angeben, so ist neben  $E = (e_1, e_2, e_3)$  auch  $\tilde{E} = (e_1, e_2, e_3, e_1 + e_2)$  ein Erzeugendensystem von  $\mathbb{R}^3$ , allerdings ein unnötig großes, weil der zusätzliche Vektor  $e_1 + e_2$  die Lineare Hülle nicht mehr vergrößert. Da jeder Vektorraum ein Nullelement enthält, betrachten wir dieses beispielhaft und schreiben die 0 als Linearkombination der Vektoren des Erzeugendensystems. In  $E$  lässt sich die Null nur darstellen als  $0 = 0e_1 + 0e_2 + 0e_3$ , in  $\tilde{E}$  gilt  $0 = 0e_1 + 0e_2 + 0e_3 + 0(e_1 + e_2)$ , aber auch  $0 = (-1)e_1 + (-1)e_2 + 0e_3 + 1(e_1 + e_2)$ . Die Redundanz in  $\tilde{E}$  hängt also mit der Möglichkeit zusammen, die 0 auf verschiedene Weise als Linearkombination zu schreiben.

**Definition 3.42:** Es sei  $V$  ein Vektorraum über  $K$ . Ein  $r$ -Tupel  $(v_1, \dots, v_r) \subset V$  heißt *linear unabhängig*, wenn aus  $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r = 0$  stets folgt, dass  $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$  ist.

Ist die Darstellung des Nullvektors als Linearkombination eindeutig, gilt dies sogar für *jeden* Vektor.

**Bemerkung 3.43:** Alle Vektoren  $v$  aus  $L(v_1, \dots, v_n)$  sind genau dann eindeutig als Linearkombination der Vektoren  $v_1, \dots, v_n$  darstellbar, wenn das  $n$ -Tupel  $(v_1, \dots, v_n)$  linear unabhängig ist.

BEWEIS Es seien

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \\ &= \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n \end{aligned}$$

zwei verschiedene Darstellungen von  $v$ . Es gilt:

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots + (\alpha_n - \beta_n)v_n = 0$$

Die Vektoren  $v_1, \dots, v_n$  sind nach Definition genau dann linear unabhängig, wenn alle Differenzen  $\alpha_n - \beta_n$  gleich 0 sind. ■

**Bemerkung 3.44:** Für linear unabhängige Vektoren gilt:

- Keiner der Vektoren ist eine Linearkombination der übrigen.
- Keiner der Vektoren ist der Nullvektor.

**Beispiel 3.45 ( $\mathbb{R}^n$ ):** Zwei Vektoren  $v_1$  und  $v_2$  sind linear abhängig, wenn die Gleichung

$$\lambda_1 v_1 + \lambda_2 v_2 = 0$$

mit  $\lambda_1$  oder  $\lambda_2 \neq 0$  gilt. Sei o. b. d. A.  $\lambda_1 \neq 0$  (für  $\lambda_2 \neq 0$  ist die Rechnung ähnlich). Dann ist

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2.$$

Diese Gleichung ist erfüllt, wenn entweder  $v_1 = 0$  und  $\lambda_2 = 0$  (aber  $\lambda_1 \neq 0$ ) ist oder  $v_1$  und  $v_2$  parallel sind. Zwei Vektoren des  $\mathbb{R}^n$  sind also genau dann linear abhängig, wenn sie entweder parallel sind oder einer der beiden Vektoren der Nullvektor ist.

Die  $n$  Vektoren  $v_1, \dots, v_n$  sind genau dann linear abhängig, wenn

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

gilt, wobei mindestens einer der Vorfaktoren  $\lambda_1, \dots, \lambda_n \neq 0$  ist. Sagen wir  $\lambda_1 \neq 0$  (für die anderen  $\lambda_i$  ist die Rechnung äquivalent). Dann gilt die Gleichung in zwei Fällen:

1.  $v_1 = 0$ . Dann können alle anderen  $\lambda_i = 0$  sein.
2.  $v_1 = -\frac{1}{\lambda_1}(\lambda_2 v_2 + \dots + \lambda_n v_n)$ . Dann ist  $v_1$  eine Linearkombination der anderen  $v_i$ .

**Beispiel 3.46:** Das  $n$ -Tupel der kanonischen Einheitsvektoren im  $\mathbb{R}^n$  ist linear unabhängig, das 4-Tupel  $\tilde{E} = (e_1, e_2, e_3, e_1 + e_2)$  von oben nicht.



### 3.4.2 Nachweis linearer Unabhängigkeit

Wir betrachten zunächst den Vektorraum  $\mathbb{K}^n$ , der sich als prototypisch für endlich erzeugte reelle oder komplexe Vektorräume erweisen wird. Gegeben seien Vektoren  $v_1, \dots, v_r$ ,  $r \leq n$ , die auf lineare Unabhängigkeit getestet werden sollen. Die Gleichung

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0$$

aus der Definition 3.42 ist ein lineares Gleichungssystem in den Koeffizienten  $\lambda_i$ . Die Vektoren sind genau dann linear unabhängig, wenn man als einzige Lösung den Nullvektor erhält.

**Beispiel 3.47:** Gegeben sind

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}.$$

Sie führen zum Gleichungssystem

$$\begin{cases} \lambda_1 + 2\lambda_2 - \lambda_3 = 0 \\ 2\lambda_1 + 3\lambda_2 + 2\lambda_3 = 0 \\ \lambda_1 + \lambda_2 + 5\lambda_3 = 0 \end{cases}$$

Die Lösung nach dem Gauß-Schema ergibt:

$$\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & 3 & 2 & 0 & | -2 \cdot I \\ 1 & 1 & 5 & 0 & | -I \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & -1 & 6 & 0 & | -II \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & 0 & 2 & 0 \end{array}$$

Daraus folgt  $\lambda_3 = \lambda_2 = \lambda_1 = 0$ . Also sind die drei Vektoren linear unabhängig.

**Beispiel 3.48:** Wir betrachten die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

Wir lösen das korrespondierende Lineare Gleichungssystem mit dem Gaußschen Eliminationsverfahren:

$$\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & 3 & 2 & 0 & | & -2 \cdot I \\ 1 & 1 & 3 & 0 & | & -I \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & -1 & 4 & 0 & | & -II \\ \hline 1 & 2 & -1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Es gibt nur noch zwei Gleichungen mit drei Unbekannten, so dass zur Lösung  $\lambda_1=\lambda_2=\lambda_3=0$  (Trivillösung) noch andere Lösungen hinzukommen. In diesem Fall sind dies:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \alpha \cdot \begin{pmatrix} -7 \\ 4 \\ 1 \end{pmatrix}.$$

Im Spezialfall  $r = n$  erhält man ein  $n \times n$ -Gleichungssystem. Wir haben gesehen, dass sich ein  $3 \times 3$ -Gleichungssystem genau dann eindeutig lösen lässt, wenn die Determinante der Koeffizientenmatrix  $A$  ungleich 0 ist. Da  $\lambda_1 = \dots = \lambda_3 = 0$  auf jeden Fall Lösung ist, folgt:

$$\begin{aligned} \det(A) \neq 0 &\Leftrightarrow \text{Lösung des Gleichungssystems ist eindeutig} \\ &\Leftrightarrow \text{Nur } \lambda_1 = \dots = \lambda_3 = 0 \text{ ist Lösung} \\ &\Leftrightarrow \text{Die Vektoren sind linear unabhängig.} \end{aligned}$$

Es würde in den vorigen Beispielen also schon ausreichen, die Determinanten der Koeffizientenmatrizen zu prüfen. Für das erste, linear unabhängige Beispiel gilt:

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

Für das zweite, linear abhängige Beispiel ist:

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

Wir betrachten nun den Vektorraum  $\mathcal{C}[a, b]$ . Vektoren dieses Vektorraums sind stetige Funktionen. Zum Überprüfen der Linearen Unabhängigkeit geht man hier üblicherweise von der Definition aus. Man muss also zeigen, dass die Nullfunktion nur trivial darstellbar ist bzw. für lineare Abhängigkeit, dass eine Funktion als Linearkombination der anderen dargestellt werden kann.

**Beispiel 3.49:**

1. Die Funktionen

$$f_1(x) = x; \quad f_2(x) = x^2; \quad f_3(x) = x^2 - 2x$$

sind auf einem beliebigen Intervall  $[a, b]$  linear abhängig, denn

$$2f_1(x) - f_2(x) + f_3(x) = 2x - x^2 + x^2 - 2x = 0 \quad \forall x \in \mathbb{R}$$

2. Die Funktionen

$$f_1(x) = \sin^2(x); \quad f_2(x) = \cos^2(x); \quad f_3(x) = 1$$

sind auf einem beliebigen Intervall  $[a, b]$  linear abhängig, denn

$$f_1(x) + f_2(x) - f_3(x) = \cos^2(x) + \sin^2(x) - 1 = 0 \quad \forall x \in \mathbb{R}$$

**Beispiel 3.50:** Die Funktionen

$$f_1(x) = x^2; \quad f_2(x) = x$$

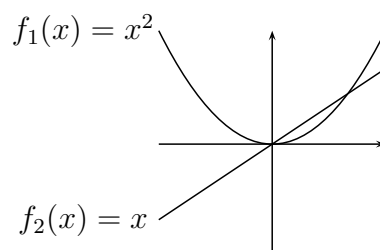
sind auf einem beliebigen Intervall  $[a, b]$  linear unabhängig. Zum Nachweis dessen müssen wir alle Lösungen von

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) = \lambda_1 x^2 + \lambda_2 x = 0$$

berechnen. Wir unterscheiden zwei Fälle.

1. Sei  $\lambda_1 = 0$ . Es folgt  $\lambda_2 x = 0 \forall x$ . Dies ist nur für  $\lambda_2 = 0$  möglich.
2. Sei  $\lambda_1 \neq 0$ . Dann gilt  $x^2 + \frac{\lambda_2}{\lambda_1} x = 0$ . Diese Gleichung ist erfüllt für  $x = 0$  und  $x = -\frac{\lambda_2}{\lambda_1}$ , aber nicht für alle  $x$ . Somit bleibt für alle  $x$  als einzige Lösung  $\lambda_1 = \lambda_2 = 0$ .

Damit sind  $f_1(x)$  und  $f_2(x)$  linear unabhängig.



Man kann in Gegensatz zum  $\mathbb{K}^n$  kein festes Rechenschema angeben, mit dessen Hilfe man in jedem Fall über Lineare Abhängigkeit oder Unabhängigkeit von Funktionen entscheiden kann. Folgende Aussage kann hier nützlich sein.

**Bemerkung 3.51:** Wir bilden zu den Funktionen  $f_1, \dots, f_n$  für paarweise verschiedene  $x_1, \dots, x_n$  die  $n$  Vektoren

$$(f_1(x_1), \dots, f_n(x_1))^T, \dots, (f_1(x_n), \dots, f_n(x_n))^T \in \mathbb{K}^n.$$

Sind diese Vektoren linear unabhängig, dann sind die Funktionen selbst  $f_1, \dots, f_n$  linear unabhängig.

BEWEIS Sei  $0 = \sum_{i=1}^n \lambda_i f_i$ . Diese Gleichung gilt punktweise im gesamten Definitionsbereich, also insbesondere für die  $x_j$ . Aus der linearen Unabhängigkeit der Vektoren  $(f_i(x_j))_{i=1}^n$  folgt  $\lambda_j = 0$  und daraus die Behauptung. ■

Üblicherweise muss man zum Nachweis der linearen Abhängigkeit oder Unabhängigkeit die Besonderheiten der jeweiligen Funktionen ausnutzen und Hilfsmittel aus der Analysis anwenden.

### 3.4.3 Basis und Dimension

**Definition 3.52:** Es sei  $V$  ein endlich erzeugter Vektorraum über  $K$ . Ein  $n$ -Tupel  $\mathcal{B} = (v_1, \dots, v_n) \subseteq V$  heißt *Basis* oder *minimales Erzeugendensystem* von  $V$ , wenn  $\mathcal{B}$  linear unabhängig ist und wenn gilt  $L(\mathcal{B}) = V$ .

**Definition 3.53:** Sei  $(v_1, \dots, v_n)$  eine Basis von  $V$  und  $v \in V$  ein beliebiger Vektor. Dann heißen die Vorfaktoren  $(\lambda_1, \dots, \lambda_n) \in K^n$  der Linearkombination

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

*Koordinaten* des Vektors  $v$  bzgl. der Basis  $(v_1, \dots, v_n)$ .

**Bemerkung 3.54:** Die Koordinaten eines Vektors bzgl. einer gegebenen Basis sind eindeutig.

BEWEIS Wegen  $V = L(v_1, \dots, v_n)$  existiert zu jedem  $v \in V$  mindestens ein  $(\lambda_1, \dots, \lambda_n) \in K^n$  mit

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

Da  $(v_1, \dots, v_n)$  linear unabhängig ist, ist diese Darstellung eindeutig nach Bemerkung 3.43. ■

**Definition 3.55:** Besitzt der Vektorraum  $V$  eine Basis  $(v_1, \dots, v_n)$ , so definieren wir die *Dimension* von  $V$  als  $\dim(V) := n$ ; besitzt der Vektorraum keine endliche Basis, dann setzt man  $\dim(V) := \infty$ .

Der Dimensionsbegriff ist nur dann wohldefiniert, wenn alle Basen gleichviele Vektoren enthalten. Es ist nicht einmal klar, ob überhaupt in jedem Vektorraum eine Basis existiert. Wir beschränken uns jetzt erneut auf endlich erzeugte Räume. Um den Dimensionsbegriff zu rechtfertigen, benötigen wir 2 Aussagen:

- Jeder endlich erzeugte Vektorraum besitzt eine Basis.
- Alle Basen bestehen aus gleich vielen Vektoren (Eindeutigkeit der Basislänge).

Wir beginnen mit der Existenz von Basen.

**Satz 3.56 (Basisergänzungssatz):** Sei  $V$  ein  $K$ -Vektorraum und seien

$$v_1, \dots, v_r, w_1, \dots, w_s$$

Vektoren in  $V$ . Ist  $(v_1, \dots, v_r)$  linear unabhängig und ist

$$L(v_1, \dots, v_r, w_1, \dots, w_s) = V,$$

dann kann man  $(v_1, \dots, v_r)$  durch evtl. Hinzunahme geeigneter Vektoren aus  $\{w_1, \dots, w_s\}$  zu einer Basis von  $V$  ergänzen.

**BEWEIS** Im Fall  $L(v_1, \dots, v_r) = V$  ist die Aussage offensichtlich. Sei also

$$L(v_1, \dots, v_r) \neq V.$$

Dann existiert mindestens ein  $w_i$  mit  $w_i \notin L(v_1, \dots, v_r)$ , denn wären alle  $w_i \in L(v_1, \dots, v_r)$ , dann müsste  $L(v_1, \dots, v_r) = L(v_1, \dots, v_r, w_1, \dots, w_s) = V$  gelten, was der Annahme  $L(v_1, \dots, v_r) \neq V$  widerspricht.

Das Tupel  $(w_i, v_1, \dots, v_r)$  ist linear unabhängig, denn aus  $\sum_{j=1}^r \lambda_j v_j + \lambda w_i = 0$  folgt  $\lambda = 0$ , da  $w_i \notin L(v_1, \dots, v_r)$ , und weiter folgt dann  $\lambda_j = 0 \forall j$ , da alle  $v_j$  linear unabhängig sind.

Möglicherweise ist  $(w_i, v_1, \dots, v_r)$  aber noch keine Basis. Dann wird der vorige Schritt wiederholt und es werden solange weitere  $w_i$  dazugenommen, bis sie  $(v_1, \dots, v_r)$  zu einer Basis von  $V$  ergänzen. Dies ist nach endlich vielen Schritten möglich wegen  $L(v_1, \dots, v_r, w_1, \dots, w_s) = V$ . ■

**Folgerung 3.57:** Jeder endlich erzeugte Vektorraum  $V$  hat eine Basis.

**BEWEIS** Weil  $V$  endlich erzeugt ist, gibt es  $v_1, \dots, v_n \in V$  mit  $L(v_1, \dots, v_n) = V$ . Streicht man alle Nullvektoren im  $n$ -Tupel  $(v_1, \dots, v_n)$ , so erhält man evtl. nach Umindizierung ein  $m$ -Tupel  $(v_1, \dots, v_m)$  mit  $L(v_1, \dots, v_m) = V$ . Aufgrund von  $v_1 \neq 0$  ist  $(v_1)$  linear unabhängig. Nach dem Basisergänzungssatz 3.56 lässt sich  $(v_1)$  mit Vektoren aus  $v_2, \dots, v_m$  zu einer Basis von  $V$  ergänzen. ■

**Beispiel 3.58:** Es sei  $V$  der  $\mathbb{R}^3$ . Weiterhin sei

$$v_1 = (1, 0, 0); \quad v_2 = (0, 1, 0).$$

Die Vektoren  $v_1$  und  $v_2$  sind linear unabhängig. Außerdem sei

$$w_1 = (1, 1, 0); \quad w_2 = (0, 0, 1).$$

Da  $L(v_1, v_2, w_1, w_2) = \mathbb{R}^3$ , sind die Voraussetzungen für den Basisergänzungssatz erfüllt. Dieser besagt: Da  $(v_1, v_2)$  keine Basis des  $\mathbb{R}^3$  ist, ist mindestens eines der Tripel  $(v_1, v_2, w_1)$ ,  $(v_1, v_2, w_2)$  oder  $(v_1, v_2, w_1, w_2)$  eine Basis des  $\mathbb{R}^3$ , d. h. linear unabhängig und erzeugend. In diesem Beispiel ist dies das Tripel  $(v_1, v_2, w_2)$ . Alle vier Vektoren bilden keine Basis des  $\mathbb{R}^3$ , da sie linear abhängig sind.

**Satz 3.59 (Austauschlemma):** Sind  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Basen eines  $K$ -Vektorraums  $V$ , dann gibt es zu jedem  $v_i$  ein  $w_j$ , so dass aus  $(v_1, \dots, v_n)$  wieder eine Basis entsteht, wenn man in ihr  $v_i$  durch  $w_j$  ersetzt.

**BEWEIS** Es seien  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  zwei Basen von  $V$ . Aus der ersten Basis wird der Vektor  $v_i$  gestrichen. Für das verkürzte Tupel  $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$  gilt

$$L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \neq V,$$

denn wäre  $L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) = V$ , dann wäre auch

$$v_i \in L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n),$$

ließe sich also durch eine Linearkombination der restlichen Vektoren darstellen. Damit wären  $(v_1, \dots, v_n)$  linear abhängig und keine Basis, was in Widerspruch zur Voraussetzung steht.

Damit existiert nach dem Basisergänzungssatz 3.56 ein  $w_j$  mit

$$w_j \notin L(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n),$$

und das Tupel  $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n, w_j)$  ist linear unabhängig.

Handelt es sich nicht um eine Basis, könnte man es durch einen der Vektoren  $v_1, \dots, v_n$  zu einer Basis ergänzen. Alle Vektoren außer  $v_i$  kommen dazu trivialerweise nicht in Frage. Aber auch  $v_i$  ist nicht möglich. In diesem Fall wäre die Basis nämlich  $(v_1, \dots, v_n, w_j)$ . Da aber bereits  $(v_1, \dots, v_n)$  eine Basis ist, ist  $w_j$  linear abhängig von diesen Vektoren und  $(v_1, \dots, v_n, w_j)$  ist keine Basis. Damit muss  $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n, w_j)$  bereits eine Basis von  $V$  sein. ■

**Beispiel 3.60:** Wir betrachten die Vektoren

$$\begin{array}{lll} v_1 = (1, 0, 0); & v_2 = (0, 1, 0); & v_3 = (0, 0, 1) \\ w_1 = (2, 0, 0); & w_2 = (0, 2, 0); & w_3 = (0, 0, 3) \end{array}$$

Hierbei sind  $(v_1, v_2, v_3)$  und  $(w_1, w_2, w_3)$  Basen des  $\mathbb{R}^3$ . Aus der ersten Basis wird  $v_3$  entfernt. Nach dem Austauschlemma kann man dafür einen der drei Vektoren  $w_i$  einsetzen, so dass wieder eine Basis entsteht. Im Beispiel ist dies  $w_3$ . Das Tripel  $(v_1, v_2, w_3)$  ist eine Basis des  $\mathbb{R}^3$ .

Wir zeigen nun die Eindeutigkeit der Basislänge in endlich erzeugten Vektorräumen und damit die Wohldefiniertheit des Dimensionsbegriffs.

**Satz 3.61 (Eindeutigkeit der Basislänge):** Sind  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Basen eines  $K$ -Vektorraums  $V$ , dann gilt  $n = m$ .

**BEWEIS** Nimmt man  $n < m$  an (der Fall  $n > m$  ist äquivalent), dann kann man mit Hilfe des Austauschlemmas aus der Basis  $(w_1, \dots, w_m)$  eine Basis  $(v_1, \dots, v_n, w_{n+1}, \dots, w_m)$  erzeugen. Nun ist aber  $(v_1, \dots, v_n)$  bereits eine Basis, so dass  $(v_1, \dots, v_n, w_{n+1}, \dots, w_m)$  linear abhängig sein müssen. Dies ist ein Widerspruch zur linearen Unabhängigkeit der Basiselemente. ■

Damit können wir folgende Bemerkung aufstellen, die uns bei der Suche nach einer geeigneten Basis für einen vorgegeben Vektorraum einen entscheidenden Hinweis liefert:

**Bemerkung 3.62:** Ist  $\dim(V) = n$  und sind  $v_1, \dots, v_n$   $n$  linear unabhängige Vektoren in  $V$ , so ist  $(v_1, \dots, v_n)$  eine Basis von  $V$ .

**BEWEIS** Andernfalls könnte man  $(v_1, \dots, v_n)$  nach dem Basisergänzungssatz 3.56 um  $k \geq 1$  Vektoren aus einer Basis  $(w_1, \dots, w_n)$  zu einer Basis ergänzen. Dann jedoch gäbe es eine Basis der Dimension  $n + k$ , was ein Widerspruch zur Eindeutigkeit der Basislänge ist. ■

**Beispiel 3.63:** Die Vektoren  $v_1 = (1, 0)$  und  $v_2 = (0, 1)$  bilden eine Basis des  $\mathbb{R}^2$ .

$$\lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0$$

Es sind somit  $v_1$  und  $v_2$  linear unabhängig und bilden wegen  $\dim(\mathbb{R}^2) = 2$  eine Basis.

Als Anwendung der in diesem Kapitel entwickelten Theorie untersuchen wir abschließend die Dimension von Summen von Untervektorräumen.

**Satz 3.64:** Seien  $U, W$  Untervektorräume eines endlich erzeugten Vektorraums. Dann gilt

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) \quad (3.3)$$

**BEWEIS** Sei  $\dim(U \cap W) = k$ ,  $\dim(U) = n$  und  $\dim(W) = m$ . Sei  $(v_1, \dots, v_k)$  eine Basis von  $U \cap W$ . Nach dem Basisergänzungssatz 3.56 ergänzen wir zu einer Basis  $(v_1, \dots, v_k, u_1, \dots, u_{n-k})$  von  $U$  und andererseits zu einer Basis  $(v_1, \dots, v_k, w_1, \dots, w_{m-k})$  von  $W$ . Sei

$$\mathcal{B} = (v_1, \dots, v_k, u_1, \dots, u_{n-k}, w_1, \dots, w_{m-k}).$$

Wir zeigen, dass  $\mathcal{B}$  eine Basis von  $U + W$  bildet.

**Lineare Unabhängigkeit:** Man hat  $u_i \notin U \cap W$ , weil sonst  $\dim(U \cap W) > k$  gälte aufgrund der linearen Unabhängigkeit von  $v_1, \dots, v_k, u_1, \dots, u_{n-k}$ . Analog erkennt man  $w_i \notin U \cap W$ . Es sind weiter die Vektoren  $u_1, \dots, u_{n-k}$  und  $w_1, \dots, w_{m-k}$  linear unabhängig: Die Vektoren  $u_1, \dots, u_{n-k}$  und  $w_1, \dots, w_{m-k}$  sind für sich genommen linear unabhängig, weil sie aus einer Basis stammen. Wir nehmen jetzt, an, für irgendein  $u_j$  gälte  $u_j \in L(w_1, \dots, w_{m-k})$ . Dann gälte insbesondere  $u_j \in U \cap W$ , aber das ist nach obigen Überlegungen nicht möglich. Also liegt kein  $u_j$  in  $L(w_1, \dots, w_{m-k})$ , und daher sind die Vektoren linear unabhängig.

**Erzeugend:** Sei jetzt  $v \in U + W$  beliebig. Dann existieren  $u \in U$  und  $w \in W$  mit  $v = u + w$ . Weil  $\mathcal{B}$  Basen von  $U$  und von  $W$  enthält, lassen sich  $u$  und  $w$  als Linearkombinationen von Vektoren aus  $\mathcal{B}$  schreiben und damit auch  $v$  (Linearkombinationen von Linearkombinationen sind wiederum Linearkombinationen). ■

**Folgerung 3.65:** Sei  $V$  ein Vektorraum mit  $\dim(V) = n$ ,  $U_1, U_2$  Unterräume und  $V = U_1 \oplus U_2$ . Dann gilt  $\dim(U_1) + \dim(U_2) = n$ .

### 3.4.4 Exkurs: Nicht endlich erzeugte Vektorräume

Wir betrachten den Vektorraum der stetigen Funktionen auf dem Intervall  $[0, 1]$ , also  $\mathcal{C}[0, 1]$ . Es seien stetige Funktionen  $f_n$  definiert durch

$$f_n(x) = \left\{ \begin{array}{l} 0, \quad x \notin [\frac{1}{n+1}, \frac{1}{n}] \\ 1, \quad x = (\frac{1}{n+1} + \frac{1}{n})/2 \\ \text{linear auf den Zwischenintervallen} \end{array} \right\}, n \in \mathbb{N}$$

Das  $n$ -Tupel  $(f_1, \dots, f_n)$  ist linear unabhängig für jedes  $n \in \mathbb{N}$ . Abbildung 3.3 zeigt die Funktionen  $f_1, f_2, f_3$ , wobei die "Dreiecke" von rechts nach links jeweils die Graphen der Funktionen  $f_1, f_2, f_3$  darstellen. Außerhalb sind die Funktionswerte jeweils gleich Null.

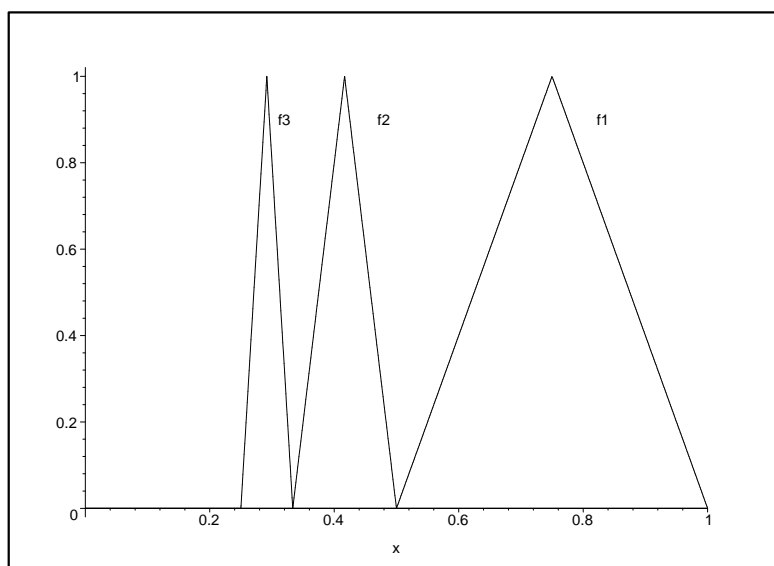
**Satz 3.66:** Der Vektorraum  $\mathcal{C}[0, 1]$  ist nicht endlich erzeugt.

**BEWEIS** sonst: Sei  $\mathcal{B}$  eine Basis mit  $m$  Vektoren. Nach dem Austauschlemma lässt sich aus jeder endlichen Basis  $\mathcal{B}$  eine Basis  $\mathcal{B}'$  erzeugen, die die Funktionen  $f_1, \dots, f_{m+1}$  enthält. Dann aber enthielte  $\mathcal{B}'$  und damit auch  $\mathcal{B}$  mindestens  $m + 1$  Vektoren. Dies ist ein Widerspruch. ■

Es erscheint aufgrund von Satz 3.66 wünschenswert, den Basisbegriff auf nicht endlich erzeugte Vektorräume auszudehnen. Dazu erweitern wir zunächst den Begriff des  $n$ -Tupels.

**Definition 3.67:** Für eine beliebige geordnete Indexmengen  $I$  (z. B. die reellen Zahlen) und einer Menge  $X \neq \emptyset$  bezeichnet man  $(a_i)_{i \in I}$  mit  $a_i \in X \forall i \in I$  als Familie.



Abbildung 3.3: Funktionen  $f_1, f_2$  und  $f_3$ 

Im Fall einer endlichen Indexmenge handelt es sich bei einer Familie um ein  $n$ -Tupel, im Fall einer abzählbar unendlichen Indexmenge  $I$  um eine Folge. Wir definieren nun durch Zurückführung auf den endlich erzeugten Fall:

**Definition 3.68:**

1. **(Lineare Hülle)** Sei  $(v_i)_{i \in I}$  eine Familie von Vektoren. Dann sei  $L(v_i)_{i \in I}$  die Menge aller Vektoren, die sich als Linearkombination von endlich vielen Vektoren  $v_1, \dots, v_r \in (v_i)_{i \in I}$  darstellen lassen.
2. **(Erzeugendensystem)** Sei  $V$  ein Vektorraum und  $(v_1, \dots, v_n)$  ein  $n$ -Tupel von Vektoren in  $V$ . Spannen die Vektoren den gesamten Raum auf, d.h. gilt  $L(v_1, \dots, v_n) = V$ , so nennt man  $(v_i)_{i \in I}$  ein *Erzeugendensystem*.
3. **(Lineare Unabhängigkeit)** Eine beliebige Familie heißt linear unabhängig, wenn jede endliche Teilfamilie linear unabhängig ist.

**Bemerkung 3.69:** Die eindeutige Darstellbarkeit eines Vektors als Linearkombination von Basisvektoren nach Bemerkung 3.43 gilt sofort in beliebigen Vektorräumen durch Auswahl einer endlichen Teilfamilie.

**Beispiel 3.70:** Die Funktionenfolge  $(f_i)_{i \in \mathbb{N}}$  von oben ist offenbar linear unabhängig im Sinne von Definition 3.68, aber kein Erzeugendensystem von  $\mathcal{C}[0,1]$ , weil z. B. die (stetige) konstante Funktion  $g(x) = 1$  nicht

dargestellt werden kann: Auf  $[1/2, 1]$  ist nur die Funktion  $f_1$  ungleich der Nullfunktion, zugleich aber ist kein Vielfaches von  $f_1$  die Funktion  $g$ .

Wir haben jetzt die Konzepte zur Verfügung, um den Basisbegriff auf beliebige Vektorräume ausdehnen zu können.

**Definition 3.71:** Es sei  $V$  ein Vektorraum über  $K$ . Eine Familie  $\mathcal{B} = (v_i)_{i \in I} \subseteq V$  heißt *Basis* oder *minimales Erzeugendensystem* von  $V$ , wenn  $\mathcal{B}$  linear unabhängig ist und wenn gilt  $L(\mathcal{B}) = V$ .

**Bemerkung 3.72:** Die Familie  $f_1, \dots$  ist kein Erzeugendensystem und daher auch keine Basis von  $\mathcal{C}[0, 1]$ .

Sei  $\mathcal{C}^k[0, 1]$  die Menge aller  $k$ -fach stetig differenzierbaren Funktionen. Dann ist nach der Summen- und Faktorregel der Analysis  $\mathcal{C}^k[0, 1]$  ein Vektorraum, und es ist  $\mathcal{C}^{k+1}[0, 1]$  ein Unterraum  $\mathcal{C}^k[0, 1]$  für alle  $k$ . Keiner dieser Räume hat endliche Dimension; man kann zum Nachweis analoge Funktionenfolgen zu  $f_1, \dots$  wie oben konstruieren, keine dieser Funktionenfolgen ist eine Basis. Es gilt aber:

**Satz 3.73:** Jeder Vektorraum hat eine Basis.

Zum Beweis dieses Satzes lassen sich die Techniken aus dem endlich erzeugten Fall nicht anwenden. Der Beweis verwendet stattdessen Hilfsmittel aus der Mengenlehre und wird hier ausgelassen. Zudem ist er nicht konstruktiv, er gibt also keinen Hinweis, wie man eine Basis eines eventuell unendlichdimensionalen Vektorraums konstruieren könnte.

**Bemerkung 3.74:** Man könnte dazu neigen, den recht abstrakten Begriff der Familie durch den eher vertrauten Begriff einer Folge zu ersetzen. Damit aber ließe man nur abzählbare Basen zu, was a priori nicht zu rechtfertigen wäre. In der Tat kann man zeigen, dass für alle  $k \in \mathbb{N}$  die Vektorräume  $\mathcal{C}^k[a, b]$  überabzählbare Basen besitzen. Konkret angeben lässt sich keine von ihnen.

### 3.4.5 Exkurs: Hyperebenen im $\mathbb{R}^n$

Mit der jetzt zur Verfügung stehenden Theorie können wir Geraden in  $\mathbb{R}^2$  und Ebenen in  $\mathbb{R}^3$  zur *Hyperebene* in allgemeinen endlich erzeugten Vektorräumen verallgemeinern. Wir gehen von der Beobachtung aus, dass man eine Gerade in der Ebene als verschobenen Untervektorraum deuten kann. Der Unterraum besitzt die Dimension  $2 - 1 = 1$ . Analog lässt sich eine Ebene in  $\mathbb{R}^3$  als verschobenen Untervektorraum der Dimension  $3 - 1 = 2$  deuten. Dies führt zur folgenden Definition.

**Definition 3.75:** Sei  $V$  ein  $K$ -Vektorraum,  $\dim(V) = n$ ,  $U$  ein Untervektorraum mit  $\dim(U) = n - 1$  und  $p \in V$  beliebig.

1. Dann heißt

$$H(U; p) := p + U := \left\{ x \in V \mid x = p + u, u \in U \right\}$$

*Hyperebene*,  $p$  heißt *Aufpunkt* der Hyperebene.

2. Existieren linear unabhängige Vektoren  $v_1, \dots, v_{n-1}$ , so dass

$$H = \left\{ x \in V \mid x = p + \sum_{i=1}^{n-1} \alpha_i v_i, \alpha_i \in K \right\} \quad (3.4)$$

gilt, dann heißen  $v_1, \dots, v_{n-1}$  *Richtungsvektoren* von  $H$ , die Darstellung einer Hyperebene in Formel (3.4) heißt *Parameterform* der Hyperebene.

**Bemerkung 3.76:** Seien  $U, V, p$  wie oben.

1. Jede Hyperebene  $H$  besitzt eine Parameterdarstellung.
2. Richtungsvektoren sind nicht eindeutig.
3. Sei  $H = p + U$  eine Hyperebene. Dann ist jeder Vektor  $p' := p + u$  mit  $u \in U$  ebenfalls Aufpunkt.

**BEWEIS**

1. Nach Folgerung 3.57 besitzt  $U$  eine Basis, deren Basisvektoren die Richtungsvektoren von  $U$  bilden.
2. Wir wählen eine Basis  $\mathcal{B}$  von  $U$  und wählen irgendeinen Vektor  $u \in U \setminus \mathcal{B}$ . Nach dem Austauschlemma 3.59 kann man dann eine neue Basis bilden, indem man dann einen geeignet gewählten Vektor in  $\mathcal{B}$  durch  $u$  ersetzt, erhält so eine neue Basis von  $U$  und damit andere Richtungsvektoren von  $H$ .
3. Offenbar gilt  $u + U = U$  aufgrund der Abgeschlossenheit von  $U$ . Also hat man

$$H(U; p) = p + U = p + (u + U) = (p + u) + U = H(p'; U). \quad \blacksquare$$

**Beispiel 3.77:**

1. Sei  $V = \mathbb{R}^2$ . Dann wird jede Hyperebene in  $V$  durch einen Richtungsvektor  $v$  und einen Aufpunkt  $p$  erzeugt, so dass

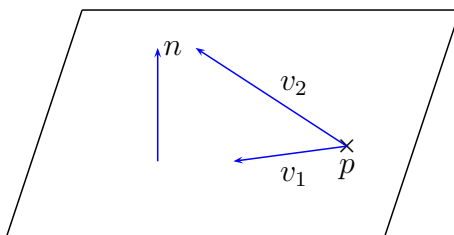
$$H = \{ x \in \mathbb{R}^2 \mid x = p + \alpha v, \alpha \in \mathbb{R} \}$$

gilt und wir wie vorgesehen eine Gerade in der Ebene erhalten.

2. Eine Hyperebene in  $\mathbb{R}^3$  ist gegeben durch

$$H = \{x \in \mathbb{R}^3 \mid x = p + \alpha_1 v_1 + \alpha_2 v_2, \alpha_1, \alpha_2 \in \mathbb{R}\}.$$

Ebenen in  $\mathbb{R}^3$  sind also Spezialfälle von Hyperebenen; die lineare Unabhängigkeit der beiden Richtungsvektoren bedeutet, dass die beiden die Ebene aufspannenden Vektoren nicht in die gleiche Richtung zeigen (in dem Fall erhielte man ja eine Gerade und keine wirkliche Ebene).



**Bemerkung 3.78:** Hyperebenen sind genau dann Unterräume von  $V$ , wenn sie den Nullvektor enthalten, im Allgemeinen also nicht. Daher besitzen sie keine Dimension im Sinne unserer Definition 3.55.

Wir beschränken uns im folgenden auf den Fall  $V = \mathbb{R}^n$ .

**Satz 3.79:** Sei  $V = \mathbb{R}^n$  und  $H$  eine Hyperebene. Dann existiert ein *Normalenvektor*  $w \in \mathbb{R}^n \setminus \{0\}$  mit

$$H = \{\langle x, w \rangle = \langle p, w \rangle\}$$

Der Vektor  $w$  ist bis auf seine Länge eindeutig festgelegt.

**BEWEIS** Wir holen den Beweis nach, wenn die erforderliche Theorie entwickelt worden ist. ■

**Definition 3.80:** Die Gleichung  $\langle x, w \rangle = \langle p, w \rangle$  heißt *Normalgleichung* oder auch *Normalform* der Hyperebene. Eine Normalform mit  $\|w\| = 1$  heißt *Hessesche Normalform*.

**Bemerkung 3.81:**

1. Sei  $w = (w_i)_{i=1}^n$  ein Normalenvektor der Hyperebene  $H$  und  $p \in H$ . Sei weiter  $\langle p, w \rangle = c \in \mathbb{R}$ . Dann lautet die Normalform von  $H$ :

$$\sum_{i=1}^n w_i x_i = c.$$

2. Nach 1. besteht ein Lineares Gleichungssystem mit  $m$  Gleichungen und  $n$  Unbekannten aus  $m$  Hyperebenen in  $\mathbb{R}^n$ , die in Normalform vorliegend die Zeilen des Linearen Gleichungssystems bilden. Da ein Lösungsvektor  $x = (x_i)_{i=1}^n$  alle Gleichungen gleichzeitig erfüllen muss, erfüllt er alle Normalgleichungen dieser Hyperebenen zugleich, liegt also in deren Schnittmenge. Die Charakterisierung der Menge aller Lösungen eines Linearen Gleichungssystems als Schnitt von  $m$  Hyperebenen in  $\mathbb{R}^n$  verallgemeinert die geometrische Deutung von  $3 \times 3$ -Systemen aus Kapitel 2.3.2 auf beliebige Lineare Gleichungssysteme.

Soll eine Normalform einer Hyperebene aus einer Parameterform berechnet werden, die ja parameterfrei ist, betrachtet man die Parametergleichung als Lineares Gleichungssystem und versucht, alle Parameter zu eliminieren. Das Ergebnis ist eine parameterfreie Darstellung, also eine Normalform. Aufgrund von Satz 3.79 ist dies immer möglich. Zur vereinfachten Berechnung von  $w$  im Fall  $\mathbb{R}^2$  und  $\mathbb{R}^3$  verweisen wir auf Kapitel 2.2.2.

**Beispiel 3.82:** Wir demonstrieren das allgemeine Vorgehen anhand von  $\mathbb{R}^3$ . Gegeben sei eine Ebene in Parameterform durch

$$x = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix}.$$

Die Parametergleichungen führen zum Linearen Gleichungssystem

$$\begin{array}{rcl} x_1 & = & 3 + \alpha_1 - \alpha_2 \\ x_2 & = & \alpha_1 + 3\alpha_2 \\ x_3 & = & 1 - \alpha_2 \\ \hline x_1 - x_2 & = & 3 - 4\alpha_2 \\ x_3 & = & 1 - \alpha_2 \\ \hline x_1 - x_2 - 4x_3 & = & -1 \end{array}$$

Zur Umrechnung einer Normalform in eine Parameterform werden  $n - 1$  linear unabhängige Vektoren  $v_i$  benötigt mit  $v_i \perp w$ . Sie können z. B. mit dem folgenden Verfahren erzeugt werden.

Gegeben sei ein Normalenvektor  $w = (w_i)_{i=1}^n$ . Aufgrund von  $w \neq 0$  existiert ein  $w_i \neq 0$ . Für jeden auf  $w$  senkrechten Vektor führe man folgende Schritte durch:

1. Man vertausche  $w_i$  mit einer anderen Komponente  $w_j$ ;  $j \neq i$ .
2. Man ändere das Vorzeichen von  $w_i$  im so erzeugten Vektor.
3. Man setze alle Komponenten  $w_k = 0$  für  $k \neq i, j$ .

Weil man bei einem Vektor der Länge  $n$  genau  $n - 1$  verschiedene Möglichkeiten hat,  $j \neq i$  zu wählen, entstehen so  $n - 1$  Vektoren, die nach Konstruktion senkrecht auf  $w$  stehen. Für ein Beispiel hierzu verweisen wir auf Beispiel 2.68.

### 3.5 Polynome

Als eine einfache Anwendung der zuvor entwickelten Theorie betrachten wir *Polynome*, die zum einen in den verschiedensten Bereichen der Mathematik verwendet werden, und die zum anderen mit algebraischen Mitteln behandelt werden können.

#### Definition 3.83:

1. Ein *Polynom* oder ganzrationale Funktion  $p : \mathbb{K} \rightarrow \mathbb{K}$  ist eine Funktion der Gestalt

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = \sum_{k=0}^n a_kx^k \quad (3.5)$$

mit den *Koeffizienten*  $a_k \in \mathbb{K}, k = 0, 1, 2, \dots, n$ . Der Koeffizient  $a_n$  heißt *Leitkoeffizient*. Im Fall  $a_n = 1$  heißt  $p$  *normiert*.

2. Die Funktion  $p(x) \equiv 0$  heißt *Nullpolynom*.
3. Ist  $a_n \neq 0$ , so heißt  $n$  *Grad* des Polynoms, Schreibweise  $\deg(p) = n$ . Weiter sei  $\deg(0) = -\infty$ .
4. Es sei  $P_n$  die Menge aller Polynome mit einem Grad von höchstens  $n$ .

#### Bemerkung 3.84:

1. In der Algebra wird zwischen einem Polynom als einer endlichen Koeffizientenfolge mit Koeffizienten aus einem abstrakten Körper  $K$  und einer Polynomfunktion  $p : K \rightarrow K$  im Sinne von Definition 3.83 unterschieden (vgl. z. B. [2]). Das ist unerlässlich, wenn man Polynome über endlichen Körpern betrachtet. Da wir uns auf den Körper  $\mathbb{K}$  beschränken, identifizieren wir im Folgenden Polynome und Polynomfunktionen.
2. Durch den Übergang des Laufindex von  $k$  nach  $n - k$  in Formel (3.5) erhält man sofort

$$p(x) = \sum_{k=0}^n a_kx^k = \sum_{k=0}^n a_{n-k}x^{n-k}.$$

Wir werden nachfolgend zeigen, dass jedes Polynom vom Grad  $n$  höchstens  $n$  Nullstellen besitzt. Vorbereitend dafür betrachten wir das Abspalten von Linearfaktoren.

**Satz 3.85:** Sei  $x_0 \in \mathbb{K}$  beliebig und  $p$  ein Polynom mit  $\deg(p) = n \in \mathbb{N}$ . Dann gilt für alle  $x \in \mathbb{K}$ :

$$p(x) = (x - x_0)p_{n-1}(x) + r$$

mit einem Polynom  $p_{n-1}$  vom Grad  $n - 1$  und  $r \in \mathbb{K}$ .

**BEWEIS** Wir führen den Beweis mittels vollständiger Induktion. Für  $n = 1$  gilt

$$\begin{aligned} p_1(x) &= a_1x + a_0 \\ &= a_1(x - x_0) + a_1x_0 + a_0 \\ &= a_1(x - x_0) + r \end{aligned}$$

Die Behauptung gelte nun für  $n$ . Dann errechnet man

$$\begin{aligned} p_{n+1}(x) &= p_n(x)x + a_0 \\ &= ((x - x_0)p_{n-1}(x) + r) \cdot x + a_0 \\ &= x(x - x_0)p_{n-1}(x) + rx + a_0 + \overbrace{rx_0 - rx_0}^{=0} \\ &= x(x - x_0)p_{n-1}(x) + r(x - x_0) + rx_0 + a_0 \\ &= (x - x_0) \overbrace{(xp_{n-1}(x) + r)}^{\tilde{p}_n(x)} + \overbrace{rx_0 + a_0}^s \\ &= (x - x_0)\tilde{p}_n(x) + s \end{aligned}$$

und erhält die Behauptung für  $n + 1$ . ■

**Bemerkung 3.86:** Seien  $x_0, p, n$  wie in Satz 3.85.

1. Es gilt  $r = p(x_0)$ .
2. Ist  $x_0$  Nullstelle von  $p$ , dann gilt

$$p(x) = (x - x_0)p_{n-1}(x).$$

**BEWEIS**

1. Es gilt  $p(x_0) = (x_0 - x_0)p_{n-1}(x_0) + r = r$ .

2. Man hat dann  $0 = p(x_0) = r$  nach 1.) ■

Eine wiederholte Anwendung von Bemerkung 3.86 auf  $p_{n-1}$  liefert folgenden Satz.

**Satz 3.87:** Hat ein Polynom  $p$  mit  $\deg(p) = n \in \mathbb{N}$  genau  $n$  Nullstellen, dann gilt die *Faktorzerlegung*

$$p_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n)a_n.$$

Daraus folgt direkt:

**Satz 3.88:** Ein Polynom vom Grad  $n \in \mathbb{N}$  hat höchstens  $n$  Nullstellen.

**Satz 3.89:** Hat ein Polynom vom Grad  $\leq n$  mehr als  $n$  Nullstellen, dann ist es das Nullpolynom, d.h. es gilt  $a_0 = a_1 = \dots = a_n = 0$ .

**Bemerkung 3.90:** Satz 3.88 besagt nicht etwa, dass ein reelles Polynom vom Grad  $n$  genau  $n$  reelle Nullstellen aufweisen müsste. So etwa besitzt das Polynom  $p(x) = x^2 + 1$  überhaupt keine reelle Nullstelle. Eine solche Aussage gilt aber sehr wohl für  $\mathbb{K} = \mathbb{C}$ .

**Satz 3.91 (Fundamentalsatz der Algebra):** Jedes nichtkonstante Polynom besitzt mindestens eine komplexe Nullstelle.

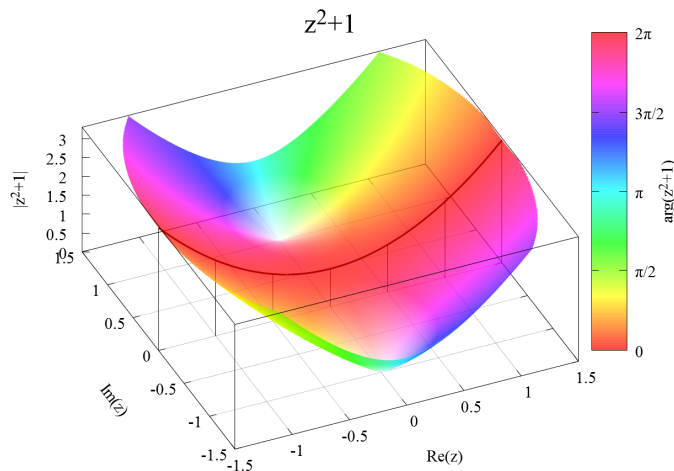
Ein vollständiger Beweis des Fundamentalsatzes wurde zuerst 1799 von C. F. Gauß in seiner Dissertation geführt. Wir verzichten hier aus Gründen des Aufwandes auf einen Beweis.

**Folgerung 3.92:** Jedes Polynom vom Grad  $n \in \mathbb{N}$  weist genau  $n$  komplexe Nullstellen auf.

**BEWEIS** Sei  $p$  ein Polynom vom Grad  $n$ . Mit einer Nullstelle  $x_0$  gilt dann nach Bemerkung 3.86  $p(x) = (x - x_0)p_{n-1}(x)$ . Da nach dem Fundamentalsatz der Algebra  $p_{n-1}$  wieder eine Nullstelle besitzt, wendet man Bemerkung 3.86 auf  $p_{n-1}$  an und fährt solange fort, bis die Faktorzerlegung von  $p$  erreicht ist. ■

**Beispiel 3.93:** Das Polynom  $p(z) = z^2 + 1$  besitzt die komplexen Nullstellen  $i$  und  $-i$  (vgl. Abb. 3.4). Die Höhe der Fläche entspricht dem Absolutbetrag von  $p(z)$ , ihre Färbung dem Argument von  $p(z)$ , also dem Winkel in Polarkoordinatendarstellung. Eine komplexe Zahl mit einem Argument von  $0$ ,  $\pi$  und  $2\pi$  entspricht somit einer reellen Zahl, bei einem Argument von  $\pi/2$  oder  $3/2\pi$  ist sie rein imaginär.



Abbildung 3.4:  $z^2 + 1$  als komplexes Polynom

**Bemerkung 3.94:** Dass jedes Polynome (komplexe) Nullstellen besitzt, bedeutet nicht, dass man sie immer berechnen könnte. Die Nullstellen eines Polynoms vom Grad 2 lassen sich mit der bekannten  $pq$ -Formel ausrechnen, die schon in der Antike bekannt war. Für Polynome vom Grad 3 und 4 existieren ziemlich komplizierte Analoga, die sog. *Cardanischen Formeln*, die in der Renaissance von Cardano erstmalig veröffentlicht wurden. Die Suche nach einer derartigen Formel für Polynome vom Grad mindestens 5 verlief über Jahrhunderte erfolglos, bis Abel 1824 mit algebraischen Methoden zeigte, dass es diese nicht geben kann. In der Praxis ist man daher in den allermeisten Fällen auf Näherungsrechnungen zur Nullstellenbestimmung angewiesen.

Wir betrachten nun Polynome als Elemente von Vektorräumen. Polynome werden als Funktionen punktweise addiert, ebenso erklärt man ein skalares Vielfaches eines Polynoms.

**Satz 3.95:** Seien  $p, q \in P_n$ . Dann sind für  $\lambda \in \mathbb{K}$  sowohl  $\lambda p$  als auch  $(p + q)$  Polynome, und es gilt

$$\deg(\lambda p) \leq n, \quad \deg(p + q) \leq n \quad (3.6)$$

**BEWEIS** Die Polynome  $p$  und  $q$  besitzen die Darstellung  $p(x) = \sum_{k=0}^n a_k x^k$  und  $q(x) = \sum_{k=0}^n b_k x^k$ , und deswegen gilt  $(p + q)(x) = \sum_{k=0}^n (a_k x^k + b_k x^k) = \sum_{k=0}^n (a_k + b_k) x^k$ . Es handelt sich also um ein Polynom mit Höchstgrad  $n$ . Die zweite Aussage folgt analog aus der Darstellung  $(\lambda p)(x) = \sum_{k=0}^n \lambda a_k x^k$ .

■

**Folgerung 3.96:** Sei  $+$  die Addition von Funktionen und  $\cdot$  die Multiplikation einer Funktion mit einem Skalar. Dann bildet  $(P_n, +, \cdot)$  einen  $\mathbb{K}$ -Vektorraum.

**BEWEIS** Da Polynome als Produkte und Summen stetiger Funktionen stetig sind, zeigen wir, dass  $P_n$  ein Untervektorraum von  $\mathcal{C}(\mathbb{K})$  bildet. Die Abgeschlossenheit im Sinne der Definition 3.26 ist genau die Aussage von Satz 3.95. ■

Wir werden nun die Dimension von  $P_n$  bestimmen und eine Basis angeben.

**Satz 3.97:** Die Funktionen  $1, x, x^2, \dots, x^n$  bilden eine Basis des Vektorraums  $P_n$ , und es gilt

$$\dim(P_n) = n + 1.$$

**BEWEIS** Nach Konstruktion von  $P_n$  bilden die Funktionen  $1, x, x^2, \dots, x^n$  ein Erzeugendensystem von  $P_n$ . Zu zeigen bleibt die lineare Unabhängigkeit. Der Nullvektor in  $P_n$  ist das Nullpolynom, die Funktion  $p(x) = 0 \forall x \in \mathbb{K}$ . Offenbar ist diese als triviale Linearkombination der Funktionen  $x^k$  darstellbar. Wir müssen nun zeigen, dass dies die einzige Darstellung des Nullpolynoms ist. Wir führen den Beweis mittels vollständiger Induktion über den Polynomgrad  $n$ . Im Fall  $n = 0$  folgt sofort  $a_0 = 0$ . Sei jetzt die Behauptung bewiesen für  $n - 1$ . Wir schreiben

$$p_n(x) = \underbrace{\sum_{k=0}^{n-1} a_k x^k}_{:=p_{n-1}(x)} + a_n x^n \equiv 0 \quad (3.7)$$

mit  $\deg(p_{n-1}) = n - 1$ . Wäre  $a_n \neq 0$ , dann müsste  $p_{n-1}(x) = -a_n x^n$  gelten, also  $\deg(p_{n-1}) = n > n - 1$ . Dies ist ein Widerspruch, also  $a_n = 0$ . Nach Voraussetzung folgt weiter  $a_0 = \dots = a_{n-1} = 0$ . ■

**Bemerkung 3.98:**

1. Die Basis  $(1, x, \dots)$  aus Satz 3.97 wird *Monombasis* genannt.
2. Der Koeffizientenvektor  $(a_0, \dots, a_n)$  bildet die Koordinaten des Polynoms  $p_n(x) = \sum_{k=0}^n a_k x^k$  bezüglich der Monombasis. Aufgrund der Eindeutigkeit der Koordinatendarstellung nach Bemerkung 3.54 sind die Koeffizienten eines Polynoms eindeutig festgelegt.
3. Stimmen zwei Polynome  $p, q \in P_n$  in mindestens  $(n + 1)$  Stellen überein, dann sind sie gleich, d.h. für

$$p_n(x) = \sum_{k=0}^n a_k x^k \quad \text{und} \quad q_n(x) = \sum_{k=0}^n b_k x^k$$

gilt dann

$$a_k = b_k, k = 0, 1, \dots, n.$$

BEWEIS Das *Differenzpolynom*

$$d_n(x) = p_n(x) - q_n(x) = \sum_{k=0}^n (a_k - b_k)x^k$$

hat nach Voraussetzung mindestens  $n + 1$  Nullstellen. Also gilt  $d_n(x) = 0$  nach Satz 3.89. ■

Wir kommen nun zu einer ersten konkreten Anwendung von Polynomen. Wir nehmen an, dass zu bestimmten Zeitpunkten  $t_1 < \dots < t_n$  Messungen irgendeiner physikalischen Größe durchgeführt wurden. Möchte man nun nachträglich vielleicht zu Auswertungszwecken einen Messwert zwischen zwei Messzeitpunkten generieren, muss man aus den vorhandenen Daten eine Funktion  $f$  erstellen, die zu den Messzeitpunkten die gemessenen Werte annimmt. Wertet man dann  $f$  zu einem Zwischenzeitpunkt aus, hofft man so einen brauchbaren Ersatzwert für eine reale Messung, die ja dort nicht stattgefunden hat, zu erhalten. Man spricht von *Interpolation* von Daten. Es liegt nahe,  $f$  als Polynom anzusetzen. Es stellt sich die Frage, ob die Interpolationsaufgabe überhaupt immer eindeutig lösbar ist.

**Satz 3.99:** Gegeben seien die  $n + 1$  Punkte  $(x_k, y_k), 0 \leq k \leq n$  mit paarweise verschiedenen  $x_k$ . Dann existiert genau ein  $p_n \in P_n$  mit  $y_k = p_n(x_k) \forall 0 \leq k \leq n$ . Dies ist das sogenannte *Interpolationspolynom*.

BEWEIS Die Eindeutigkeit folgt sofort aus Bemerkung 3.98. Die Existenz zeigen wir durch Induktion über  $n$ . Für  $n = 0$  wähle man  $p_n(x) = y_0$ . Sei nun die Behauptung gezeigt für  $n - 1$ . Das Polynom  $p_{n-1}$  interpoliere  $(x_1, y_1), \dots, (x_n, y_n)$ . Setze  $p_n(x) = p_{n-1}(x) + q(x)$  mit

$$q(x) = \frac{(x - x_0) \dots (x - x_{n-1})}{(x_n - x_0) \dots (x_n - x_{n-1})} (y_n - p_{n-1}(x_n))$$

Man hat  $q \in P_n$ , nach Folgerung 3.96 also auch  $p_n \in P_n$  und weiter  $q(x_k) = 0$  für  $k \leq n - 1$ , weil dann immer ein Linearfaktor im Nenner den Wert 0 annimmt. Es gilt also  $p_n(x_k) = y_k$  für  $k \leq n$ . Weiterhin gilt  $q(x_n) = y_n - p_{n-1}(x_n)$ , also  $p_n(x_n) = y_n$ . ■

**Beispiel 3.100:** Wir betrachten die drei Punkte  $(-2/1), (-1/-1)$  und  $(1/1)$ . Nach Satz 3.99 legen diese Punkte eine interpolierende Parabel  $p_2$  eindeutig fest. Diese kann man mit der Definition von  $p_n$  aus dem Beweis des Satzes 3.99 bestimmen. Für eine Handrechnung und wenige zu interpolierende Punkte erweist sich folgender Ansatz als ebenfalls geeignet.

Die allgemeine Form des Polynoms ist  $p_2(x) = ax^2 + bx + c$ . Einsetzen der drei Punkte ergibt die Gleichungen

$$\begin{aligned} (1/1) : \quad & 1 = a + b + c \\ (-1/-1) : \quad & -1 = a - b + c \\ (-2/1) : \quad & 1 = 4a - 2b + c \end{aligned}$$

und führt damit zum Gleichungssystem

$$\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 4 & -2 & 1 & 1 \end{array}$$

Man errechnet als Lösung  $a = 1, b = 1, c = -1$  und erhält

$$p_2(x) = x^2 + x - 1.$$

**Bemerkung 3.101:** In vielen Fällen werden Polynome vom Grad  $n$  auf Teilmengen  $M \subset \mathbb{K}$  eingeschränkt. Man bezeichnet solche Polynomräume mit  $P_n(M)$ . Enthält  $M$  mehr als  $n + 1$  Elemente, dann gelten alle Aussagen dieses Kapitels unverändert, weil dann nicht alle Elemente von  $M$  Nullstellen sein können und man deswegen mit Satz 3.89 schließen kann wie oben geschehen. Insbesondere gelten alle Aussagen dieses Kapitels für sämtliche offenen Mengen (sie enthalten alle unendlich viele Elemente), alle echten reellen Intervalle  $[a, b]$  mit  $a < b$  und alle Kreisscheiben  $|z - z_0| \leq r \subset \mathbb{C}$  mit  $r > 0$ .

Wir zeigen abschließend ein Verfahren, mit dem man die lineare Abhängigkeit von Polynomen einfach bestimmen kann. Seien  $p_1(x), \dots, p_n(x) \in P_m$ . Aus

$$\sum_{i=1}^n \lambda_i p_i(x) = 0 \tag{3.8}$$

muss für lineare Unabhängigkeit  $\lambda_1 = \lambda_2 = \dots = 0$  folgen. Jedes Polynom  $p_i(x)$  lässt sich schreiben als

$$p_i(x) = \sum_{k=0}^m a_{ki} x^k.$$

Einsetzen in Gleichung (3.8) und Umordnen der Summanden ergibt

$$0 = \sum_{i=1}^n \lambda_i \left( \sum_{k=0}^m a_{ki} x^k \right) = \sum_{k=0}^m \left( \sum_{i=1}^n \lambda_i a_{ki} \right) x^k.$$

Aus der linearen Unabhängigkeit der Monome folgt

$$\sum_{i=0}^n a_{ki} \lambda_i = 0 \quad \forall 0 \leq k \leq m$$

Dies ist ein lineares Gleichungssystem in  $(\lambda_1, \dots, \lambda_n)$ , dessen Spalten durch die Koeffizientenvektoren der einzelnen Polynome gebildet werden. Gibt es andere Lösungen als  $\lambda_1 = \dots = \lambda_n = 0$ , sind die Polynome linear abhängig, sonst linear unabhängig. Man überprüft also die lineare Unabhängigkeit von Polynomen anhand der linearen Unabhängigkeit ihrer Koeffizientenvektoren.

**Beispiel 3.102:** Die Polynome

$$p_1(x) = (1-x)^2; \quad p_2(x) = (1-x)x; \quad p_3(x) = x^2$$

sollen auf lineare Unabhängigkeit geprüft werden. Wir berechnen zunächst die Koeffizientenvektoren.

$$\begin{aligned} p_1(x) &= (1-x)^2 = 1 - 2x + 1x^2 \Rightarrow a_{01} = 1, a_{11} = -2, a_{21} = 1 \\ p_2(x) &= (1-x)x = 0 + 1x - x^2 \Rightarrow a_{02} = 0, a_{12} = 1, a_{22} = -1 \\ p_3(x) &= x^2 \Rightarrow a_{03} = 0, a_{13} = 0, a_{23} = 1 \end{aligned}$$

Das zu lösende Gleichungssystem lautet somit:

$$\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \quad | +2 \cdot I \\ 1 & -1 & 1 & 0 \quad | -I \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \quad +II \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}$$

Damit ergibt sich als einzige Lösung  $\lambda_1 = \lambda_2 = \lambda_3 = 0$ , und die Polynome sind linear unabhängig.

### 3.6 Skalarprodukt, euklidische und unitäre Räume

In Kapitel 2 hatte sich erwiesen, dass sich wesentliche geometrische Zusammenhänge in  $\mathbb{R}^n$  wie Lagebeziehungen von Geraden und Ebenen, Abstandsberechnungen, Winkelbestimmung u.s.w. unter Rückgriff auf Norm und Skalarprodukt untersuchen lassen. In Kapitel 3.3 wurde der Vektorraum

$\mathbb{R}^n$  zum allgemeinen  $K$ -Vektorraum erweitert, so dass es naheliegt, die geometrischen Betrachtungen in  $\mathbb{R}^n$  auf allgemeine Vektorräume auszudehnen. Wir gehen daher von  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$  aus und erweitern zunächst die Definition 2.4 des Skalarprodukts von  $\mathbb{R}^n$  auf allgemeine  $\mathbb{K}$ -Vektorräume, indem wir wie schon in ähnlichen Fällen zuvor jede Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$  ein Skalarprodukt nennen, die die in Definition 2.4 geforderten Eigenschaften besitzt. Danach werden wir eine Norm, also einen Abstandsbegriff auf allgemeinen  $\mathbb{K}$ -Vektorräumen definieren. Dann lässt sich analog zu Kapitel 2 fortfahren.

**Definition 3.103:** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Eine Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$  heißt *Skalarprodukt*, wenn folgende Bedingungen erfüllt sind:

**SP1:**  $\forall a, b \in V$ :

$$\langle a, b \rangle = \begin{cases} \langle b, a \rangle & \text{für } \mathbb{K} = \mathbb{R} \\ \overline{\langle b, a \rangle} & \text{für } \mathbb{K} = \mathbb{C}. \end{cases}$$

**SP2**  $\forall a, b, c \in V$  :

$$\begin{aligned} \langle a, (b + c) \rangle &= \langle a, b \rangle + \langle a, c \rangle \\ \langle (a + b), c \rangle &= \langle a, c \rangle + \langle b, c \rangle \end{aligned}$$

**SP3:**  $\forall \alpha \in \mathbb{K}$  gilt

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle = \begin{cases} \langle a, \alpha b \rangle & \text{für } \mathbb{K} = \mathbb{R} \\ \langle a, \bar{\alpha} b \rangle & \text{für } \mathbb{K} = \mathbb{C}. \end{cases}$$

**SP4:** (positive Definitheit)  $\forall a \in V \setminus \{0\} : \langle a, a \rangle > 0$ , und  $\langle 0, 0 \rangle = 0$ .

**Bemerkung 3.104:** Um die positive Definitheit **SP4** formulieren zu können, muss  $\langle a, a \rangle$  immer eine reelle Zahl sein, auch wenn  $a$  komplexe Anteile enthält, weil nur  $\mathbb{R}$  die Ordnungsrelation " $>$ " besitzt. Man erreicht dies für  $\mathbb{K} = \mathbb{C}$ , indem man in die Symmetriebedingung an das Skalarprodukt wie in **SP1** modifiziert. Dann gilt nämlich  $\forall a \in V : \langle a, a \rangle = \overline{\langle a, a \rangle}$ , und deswegen  $\text{Im}(\langle a, a \rangle) = 0$ .

**Beispiel 3.105:** Ein Skalarprodukt auf  $\mathbb{R}^n$  nach Definition 2.1 ist auch Skalarprodukt im Sinn von Definition 3.103, und deswegen sind die beiden Definitionen 2.4 und 3.103 miteinander konsistent.

Wir werden zunächst in Analogie zum euklidischen Skalarprodukt das Standardskalarprodukt auf  $\mathbb{C}^n$  definieren.

**Definition 3.106:** Auf  $\mathbb{C}^n$  sei für  $a = (a_i)_{i=1}^n, b = (b_i)_{i=1}^n$  das Standardskalarprodukt definiert durch

$$\langle a, b \rangle := \sum_{i=1}^n a_i \overline{b_i}. \quad (3.9)$$

**Beispiel 3.107:**

1. Seien  $a = (1, -1), b = (1 + i, 2 - i) \in \mathbb{C}^2$ . Dann gilt

$$\langle a, b \rangle = 1 \cdot \overline{1 + i} + (-1) \cdot \overline{2 - i} = (1 - i) - (2 + i) = -1 - 2i \in \mathbb{C}$$

und weiter  $\langle a, a \rangle = 2$  sowie  $\langle b, b \rangle = (1 - i)(1 + i) + (2 - i)(2 + i) = 2 + 5 = 7 \in \mathbb{R}$ .

2. Im Fall von Vektoren mit reellen Komponenten entspricht das Skalarprodukt auf  $\mathbb{C}^n$  genau dem euklidischen Skalarprodukt auf  $\mathbb{R}^n$ .

Wir definieren nun ein abstraktes Skalarprodukt auf dem Vektorraum der stetigen Funktionen, der sich ja wesentlich von  $\mathbb{K}^n$  unterscheidet.

**Beispiel 3.108:** Seien  $f, g \in \mathcal{C}[a, b]$ . Auf  $\mathcal{C}[a, b]$  wird durch

$$\langle f, g \rangle := \int_a^b f(x) \cdot g(x) dx \quad (3.10)$$

ein Skalarprodukt definiert.

**BEWEIS** Auf einem beschränkten und abgeschlossenen Intervall nehmen stetige Funktionen nach einem Satz von Weierstraß ihr Minimum und Maximum an. Es existiert also  $\alpha = \max_{x \in [a, b]} (f \cdot g)(x) \in \mathbb{R}$  und  $\beta = \min_{x \in [a, b]} (f \cdot g)(x) \in \mathbb{R}$ , weil  $f \cdot g$  stetig ist. Damit gilt

$$\begin{aligned} \langle f, g \rangle &\leq \int_a^b \alpha dx = \alpha(b - a) < \infty \\ \langle f, g \rangle &\geq \int_a^b \beta dx = \beta(b - a) > -\infty, \end{aligned}$$

und daher ist die Abbildung  $\langle \cdot, \cdot \rangle : V \rightarrow \mathbb{R}$  wohldefiniert, weil ja das Integral immer einen endlichen Wert annimmt.

**SP1:**

$$\begin{aligned} \langle f, g \rangle &= \int_a^b f(x) \cdot g(x) dx \\ &= \int_a^b g(x) \cdot f(x) dx = \langle g, f \rangle \end{aligned}$$

**SP2:**

$$\begin{aligned}
\langle f, g+h \rangle &= \int_a^b f(x) \cdot (g(x) + h(x)) \, dx \\
&= \int_a^b (f(x) \cdot g(x)) + (f(x) \cdot h(x)) \, dx \\
&= \int_a^b (f(x) \cdot g(x)) \, dx + \int_a^b (f(x) \cdot h(x)) \, dx = \langle f, g \rangle + \langle f, h \rangle
\end{aligned}$$

**SP3:**

$$\begin{aligned}
\langle \lambda f, g \rangle &= \int_a^b \lambda f(x) \cdot g(x) \, dx \\
&= \lambda \cdot \int_a^b f(x) \cdot g(x) \, dx = \lambda \cdot \langle f, g \rangle
\end{aligned}$$

**SP4:**

$$\langle f, f \rangle = \int_a^b f^2(x) \, dx$$

Im Fall  $f = 0$  gilt  $\langle f, f \rangle = \int_a^b 0 \, dx = 0$ . Sei  $f \neq 0$ . Dann gilt  $f^2 \geq 0$ , und es existiert  $\tilde{x} \in [a, b]$  mit  $f(\tilde{x}) \neq 0$ , also  $f^2(\tilde{x}) > 0$ . Wir nehmen zunächst  $x \neq a$  und  $x \neq b$  an. Sei  $f^2(\tilde{x}) = \varepsilon > 0$ . Aufgrund der Stetigkeit von  $f^2$  existiert ein  $\delta > 0$ :  $f^2(y) > \varepsilon/2 \forall y$  im offenen Intervall  $(\tilde{x} - \delta, \tilde{x} + \delta)$ . Daher gilt

$$\langle f, f \rangle = \int_a^b f^2(x) \, dx \geq \int_{\tilde{x}-\delta}^{\tilde{x}+\delta} f^2(x) \, dx \geq \varepsilon/2 \cdot 2\delta > 0.$$

Im Fall  $x = a$  gilt mit derselben Schlussweise  $f^2(y) > \varepsilon/2$  auf  $[a, a + \delta)$  und deswegen

$$\langle f, f \rangle = \int_a^b f^2(x) \, dx \geq \int_a^{a+\delta} f^2(x) \, dx \geq \varepsilon/2 \cdot \delta > 0.$$

Analog argumentiert man im Fall  $x = b$  mit dem Intervall  $(b - \delta, b]$ . Es folgt damit **SP4**. ■

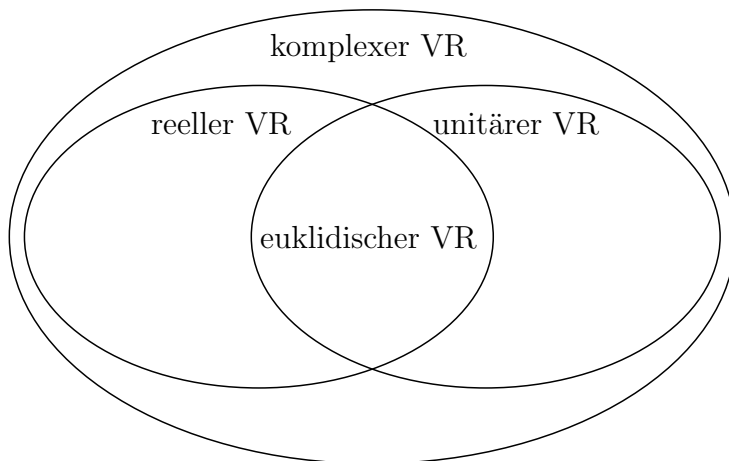
**Beispiel 3.109:** Gegeben seien  $f(x) = 1$  und  $g(x) = x$  auf dem Intervall  $[-1, 1]$ . Sowohl  $f$  als auch  $g$  sind stetige Funktionen. Man errechnet

$$\langle 1, x \rangle = \int_{-1}^1 1 \cdot x \, dx = \frac{1}{2} x^2 \Big|_{-1}^1 = 0 \quad (3.11)$$

**Definition 3.110:** Ein reeller Vektorraum gemeinsam mit einem Skalarprodukt heie *Euklidischer Vektorraum*, ein komplexer Vektorraum mit einem Skalarprodukt heie *Unitärer Vektorraum*.



Grafisch kann man die Zusammenhänge als Mengendiagramm wie folgt darstellen:



**Beispiel 3.111:** Die Vektorräume  $\mathbb{R}^n$  mit dem euklidischen Skalarprodukt und  $\mathcal{C}[a, b]$  mit dem Skalarprodukt aus Beispiel 3.108 sind euklidische Vektorräume, ebenso natürlich jeder Unterraum dieser Vektorräume wie z.B.  $P_n([a, b])$ .

Zur Definition einer Norm auf einem allgemeinen Vektorraum  $V$  gehen wir in der bewährten Weise vor: Wir gehen von der Definition 2.12 einer Norm auf  $\mathbb{R}^n$  aus und nennen *jede* Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$ , die die in Definition 2.12 Eigenschaften aufweist, eine Norm auf  $V$ . Die Konsistenz beider Definitionen ist damit nach Konstruktion gesichert.

**Definition 3.112:** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $a, b \in V$ . Eine Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$  heißt *Norm* genau dann, wenn

$$\mathbf{N0} : \|a\| \in \mathbb{R}.$$

$$\mathbf{N1} : \|a\| \geq 0.$$

$$\mathbf{N2} : \|a\| = 0 \Leftrightarrow a = 0.$$

$$\mathbf{N3} : \forall \lambda \in \mathbb{K} : \|\lambda a\| = |\lambda| \|a\|$$

$$\mathbf{N4} : (\text{Dreiecksungleichung}) \|a + b\| \leq \|a\| + \|b\|$$

Wie im Spezialfall  $V = \mathbb{R}^n$  (vgl. Kapitel 2.1) induziert ein Skalarprodukt eine Norm.

**Satz 3.113:** In einem unitären (bzw. euklidischen) Raum induziert das Skalarprodukt eine (Standard-)Norm durch

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

**BEWEIS N0** ist für euklidische Vektorräume klar und folgt für unitäre Vektorräume aus Bemerkung 3.104. Die Eigenschaften **N1** - **N4** wurden bereits in Satz 2.13 für  $\mathbb{R}^n$  gezeigt. Für **N1** - **N3** haben wir dazu nur die Gleichung  $\|x\| = \sqrt{\langle x, x \rangle}$  verwendet und keine der speziellen Eigenschaften von  $\mathbb{R}^n$  benötigt. Daher sind mit dem dortigen Beweis **N1** - **N3** schon gezeigt. Den Beweis von **N4** stellen wir kurz zurück. ■

**Satz 3.114:** In allen unitären Vektorräumen  $V$  gilt die *Cauchy-Schwarzsche Ungleichung*

$$|\langle a, b \rangle| \leq \|a\| \|b\| \quad \forall a, b \in V. \quad (3.12)$$

**BEWEIS** Im Beweis der Cauchy-Schwarzschen Ungleichung auf  $\mathbb{R}^n$  (vgl. Satz 2.30). wurden nur die Eigenschaften **SP2** und **SP4** benutzt. Skalare Größen wurden nur aus dem ersten Argument gezogen. Weil also nur die Eigenschaften genutzt wurden, die im reellen wie komplexen Fall gleichermaßen gelten, gilt der Beweis wortgleich für den allgemeinen Fall. ■

Beim Beweis der Dreiecksungleichung **N4** auf  $\mathbb{R}^n$  wurde in Satz 2.31 die Symmetrie des Skalarproduktes ausgenutzt, so dass sich dieser Beweis unmittelbar nur auf euklidische Vektorräume übertragen lässt. Wir zeigen also die Dreiecksungleichung für unitäre Vektorräume (und damit auch für euklidische Vektorräume).

**BEWEIS** Beide Seiten der Dreiecksungleichung sind reell und insbesondere nicht negativ. Daher genügt es, zu beweisen, dass ihre Quadrate die gewünschte Ungleichung erfüllen d. h.

$$\langle a + b, a + b \rangle \leq (\|a\| + \|b\|)^2.$$

Man hat

$$\langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle$$

Es gilt  $\langle b, a \rangle = \overline{\langle a, b \rangle}$  und deswegen  $\langle a, b \rangle + \langle b, a \rangle = 2\operatorname{Re}\langle a, b \rangle \leq 2|\langle a, b \rangle|$ , weil der Absolutbetrag einer komplexen Zahl mindestens so groß ist wie ihr Realteil. Aufgrund der Cauchy-Schwarzschen Ungleichung (3.12) gilt dann  $\langle a, b \rangle + \langle b, a \rangle \leq 2\|a\| \|b\|$ . Mit der ersten binomischen Formel folgt die Behauptung. ■

**Beispiel 3.115:**

1. Für  $V = \mathbb{C}$  und  $z = a + ib \in \mathbb{C}$  induziert das Standardskalarprodukt (3.9) die Norm

$$\|z\| = \sqrt{z \cdot \bar{z}} = \sqrt{(a + ib)(a - ib)} = \sqrt{a^2 + b^2} = |z|,$$

also den gewöhnlichen Betrag komplexer Zahlen.

2. Für  $\mathbb{R}^n$  erhält man aus dem euklidischen Skalarprodukt die euklidische Norm (vgl. Kapitel 2.1).
3. Auf  $V = \mathcal{C}[a, b]$  und  $f \in V$  induziert das Skalarprodukt aus Beispiel 3.109 die sog.  $L_2$ -Norm

$$\|f\|_{L_2} := \int_a^b f^2(x) dx. \quad (3.13)$$

Sei z. B:  $[a, b] = [-1, 1]$  und  $f(x) = x$ . Dann ist

$$\begin{aligned} \|f\|_{L_2}^2 &= \|x\|_{L_2}^2 = \int_{-1}^1 x \cdot x dx \\ &= \frac{1}{3} x^3 \Big|_{-1}^1 = \frac{2}{3}. \end{aligned}$$

4. Nicht jede Norm auf einem Vektorraum wird durch ein Skalarprodukt induziert (vgl. Bemerkung 2.15). So wird die auf  $\mathcal{C}[a, b]$  in der Analysis sehr gebräuchliche Maximumsnorm

$$\|f\|_{\infty} := \max_{x \in [a, b]} |f(x)|$$

von keinem Skalarprodukt induziert.

**Bemerkung 3.116:** Weil die Cauchy-Schwarzsche Ungleichung gilt, ist es auf euklidischen Vektorräumen problemlos möglich, einen Winkelbegriff wie in Kapitel 2.1.3 zwischen Vektoren einzuführen. In der Praxis haben aber Winkel in beliebigen euklidischen Räumen eine sehr geringe Bedeutung. Winkel in unitären Räumen sind dagegen nicht ohne Weiteres wie in Kapitel 2.1.3 zu definieren, weil das Skalarprodukt in Gleichung (2.8), die zur Winkeldefinition dienen müsste, komplexe Werte annehmen kann.

### 3.7 Orthogonalität in unitären Vektorräumen

Für die analytische Geometrie in  $\mathbb{R}^n$  (vgl. Kapitel 2) hat sich die *orthogonale Projektion* eines Vektors auf einen anderen als fundamental herausgestellt. Man darf also erwarten, dass auch in unitären Vektorräumen, die ja  $\mathbb{R}^n$  verallgemeinern, eine orthogonale Projektion eine wesentliche Rolle spielen wird. Wir definieren zunächst Orthogonalität in Analogie zur Definition 2.22 in  $\mathbb{R}^n$  auf allgemeinen unitären Vektorräumen. Der euklidische Vektorraum ist ein Spezialfall dessen, so dass wir nur unitäre Vektorräume betrachten.

**Definition 3.117:** Sei  $V$  ein unitärer Vektorraum und  $a, b \in V$ . Es stehen  $a$  und  $b$  orthogonal zueinander, falls

$$\langle a, b \rangle = 0$$

gilt. Man schreibt  $a \perp b$ .

**Bemerkung 3.118:** Mit wortgleichem Beweis wie in Satz 2.25 gilt dann der Satz des Pythagoras  $\|a + b\|^2 = \|a\|^2 + \|b\|^2$  für  $a \perp b$  in allen unitären Vektorräumen.

Da bei der Herleitung der orthogonalen Projektion (2.5) in  $\mathbb{R}^n$  nur die Eigenschaften SP2 und SP3 des Skalarprodukts benutzt wurden, die ja in unitären Vektorräumen unverändert gelten, erhält man sofort

**Satz 3.119:** Für die orthogonale Projektion  $p_b(a)$  eines Vektors  $a$  auf  $b$  gilt in jedem unitären Vektorraum

$$p_b(a) = \frac{\langle a, b \rangle}{\|b\|^2} b \quad (3.14)$$

Man erkennt an Gleichung (3.14), dass sich der Wert von  $p_b(a)$  nicht ändert, wenn man  $b$  durch ein skalares Vielfaches (bis auf den Nullvektor natürlich) ersetzt. Alle Vektoren des von  $b$  aufgespannten Untervektorraums ohne den Nullvektor führen also zur gleichen Projektion, so dass es naheliegt,  $p_b(a)$  als eine orthogonale Projektion von  $a$  auf den Untervektorraum  $U = L(b)$  anzusehen. Insbesondere gilt  $a - p_b(a) \perp u \forall u \in U$ . Wir untersuchen jetzt eine solche orthogonale Projektion auf beliebige Untervektorräume.

**Definition 3.120:** Sei  $U$  ein endlich erzeugter Untervektorraum eines unitären Vektorraums  $V$  und  $a \in V$ . Ein Vektor  $p_U(a) \in U$  heißt *orthogonale Projektion* von  $a$  auf  $U$ , wenn

$$a - p_U(a) \perp u \quad \forall u \in U \quad (3.15)$$

gilt.

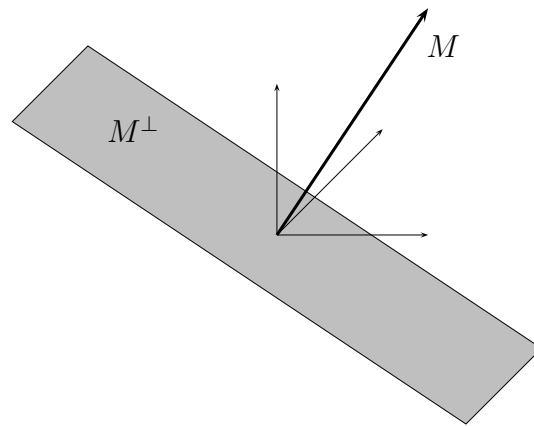
Es stellt sich die Frage nach der Wohldefiniertheit, d.h. ob es immer einen derartigen Vektor  $p_U(a)$  gibt und ob er eindeutig ist. Zur Diskussion der Eindeutigkeit hilft folgender Begriff.

- Gibt es immer ein eindeutiges  $p_U(a)$ ?
- folgender Begriff hilfreich

**Definition 3.121:** Sei  $M$  eine Teilmenge eines unitären Vektorraums  $V$ . Dann heißt

$$M^\perp = \{v \in V \mid v \perp u \forall u \in M\}$$

das *orthogonale Komplement* von  $M$ .

Abbildung 3.5: orthogonales Komplement von  $M = \{v\}$ **Bemerkung 3.122:**

1.  $M^\perp$  ist ein Untervektorraum von  $V$ .
2. Sei  $U$  ein Untervektorraum von  $V$ . Dann gilt  $U \cap U^\perp = \{0\}$ .

**BEWEIS** 1.) Es ist die Abgeschlossenheit zu prüfen. Für  $u \in M$ ;  $x, y \in M^\perp$  und  $\lambda \in \mathbb{R}$  gilt:

$$\begin{aligned}\langle x + y, u \rangle &= \langle x, u \rangle + \langle y, u \rangle = 0 \\ \langle \lambda x, u \rangle &= \lambda \langle x, u \rangle = 0\end{aligned}$$

2.) Sei  $a \in U \cap U^\perp$ . Dann gilt  $\langle a, u \rangle = 0 \forall u \in U$ , da  $a \in U^\perp$ , also insbesondere  $\langle a, a \rangle = 0$  wegen  $a \in U$  und deswegen  $a = 0$ . ■

**Beispiel 3.123:** Es sei  $V = \mathbb{R}^3$ ,  $v = (2, 0, 0)^T$  und  $M = \{v\}$ , eine Menge mit nur einem Vektor. Dann gilt  $M^\perp = \{x \in \mathbb{R}^3 \mid \langle x, v \rangle = 0\}$ , also für die konkrete Wahl von  $n$  genau die  $(y, z)$ -Ebene. Für beliebige Vektoren  $v \neq 0$  entspricht  $M^\perp$  der Ebene mit dem Normalenvektor  $v$  durch den Nullpunkt (vgl. Kapitel 2.2 und Abb. 3.5).

**Folgerung 3.124:** Seien  $U, V$  und  $a$  wie in Definition 3.120. Die orthogonale Projektion von  $a$  auf  $U$  ist eindeutig.

**BEWEIS** Sei  $q \in U$  mit  $a - q \perp u \forall u \in U$ . Dann gilt  $q - p_U(a) \in U$ , aber wegen

$$\begin{aligned}0 &= \langle a - q, u \rangle = \langle a - p_U(a), u \rangle \quad \forall u \in U \\ \Leftrightarrow \langle a, u \rangle - \langle q, u \rangle &= \langle a, u \rangle - \langle p_U(a), u \rangle \quad \forall u \in U \\ \Leftrightarrow \langle q - p_U(a), u \rangle &= 0 \quad \forall u \in U\end{aligned}$$

auch  $q - p_U(a) \in U^\perp$ , also nach Bemerkung 3.122  $q = p_U(a)$ . ■

**Lemma 3.125:** Sei  $U$  wie zuvor und  $(u_1, \dots, u_m)$  eine Basis von  $U$ . Für  $v \in V$  gilt:

$$v \in U^\perp \Leftrightarrow \langle v, u_i \rangle = 0 \quad \forall 1 \leq i \leq m$$

**BEWEIS** " $\Rightarrow$ " ist klar. " $\Leftarrow$ ": Sei  $u \in U$  beliebig. Dann gilt  $u = \sum_{i=1}^m \lambda_i u_i$  mit  $\lambda_i \in \mathbb{K}$  und damit

$$\langle u, v \rangle = \left\langle \sum_{i=1}^m \lambda_i u_i, v \right\rangle = \sum_{i=1}^m \lambda_i \langle v, u_i \rangle = 0. \quad \blacksquare$$

Sei  $(u_1, \dots, u_m)$  eine Basis von  $U$ . Existiert  $p_U(a) \in U$ , dann folgt

$$p_U(a) = \sum_{i=1}^m \alpha_i u_i \quad (3.16)$$

mit noch zu bestimmenden Koeffizienten  $\alpha_i$  und aus Gleichung (3.15) zusammen mit Lemma 3.125:

$$\begin{aligned} \left\langle a - \sum_{i=1}^m \alpha_i u_i, u_j \right\rangle &= 0 \quad \forall 1 \leq j \leq m \\ \Leftrightarrow \sum_{i=1}^m \langle u_i, u_j \rangle \alpha_i &= \langle a, u_j \rangle \quad \forall 1 \leq j \leq m \\ G\alpha &= b. \end{aligned} \quad (3.17)$$

mit dem Koeffizientenvektor  $\alpha = (\alpha_1, \dots, \alpha_m)^T$  und der Matrix  $G = (\langle u_i, u_j \rangle)_{i,j=1}^m$ . Die Matrix  $G$  wird Gram<sup>1</sup> zu Ehren *Gram-Matrix* genannt. Man wird also auf ein lineares Gleichungssystem in den Koeffizienten geführt. Die orthogonale Projektion von  $a$  auf  $U$  existiert daher genau dann, wenn dieses Gleichungssystem eine Lösung besitzt. Wir stellen die Frage der Lösbarkeit zunächst zurück.

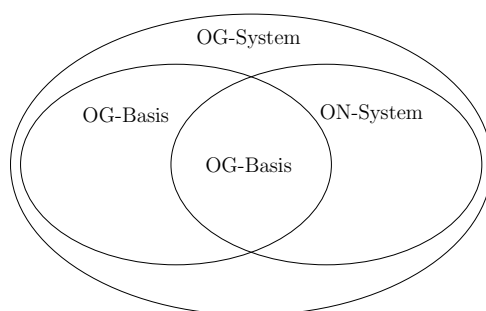
Wie gesehen erfordert die Berechnung der orthogonalen Projektion die eventuell aufwändige Lösung eines linearen Gleichungssystems. Andererseits hängt die Matrix  $G$  dieses Gleichungssystems unmittelbar von der Wahl der Basis von  $U$  ab. Am besten geeignet erscheint eine Basis von  $U$ , für die  $\langle u_i, u_j \rangle = 0$  für  $i \neq j$  und  $\langle u_i, u_i \rangle = 1$  für  $i = j$  gilt, also eine mit paarweise orthogonalen Basisvektoren der Länge eins. Dann nämlich gälte  $A = E$ , und das Lösen des linearen Gleichungssystems (3.17) entfiel.

**Definition 3.126:** Sei  $V$  ein unitärer Vektorraum und  $\mathcal{B} = (v_1, \dots, v_m)$  ein  $m$ -Tupel mit Vektoren in  $V \setminus \{0\}$ .

1.  $\mathcal{B}$  heißt *Orthogonalsystem* in  $V$ , falls sämtliche  $v_i$  paarweise orthogonal sind.

<sup>1</sup>Jørgen Pederson Gram (1850-1916): dänischer Mathematiker; Beiträge zur Mathematik und Forstwirtschaft; Vorstand einer Versicherungsgesellschaft

2. Ein Orthogonalsystem, für das zusätzlich noch  $\|v_i\| = 1 \forall i = 1, \dots, m$  gilt, heißt *Orthonormalsystem*.
3. Ein Orthogonalsystem, das eine Basis von  $V$  bildet, heißt *Orthogonalbasis* von  $V$ .
4. Ein Orthonormalsystem, das eine Basis von  $V$  bildet, heißt *Orthonormalbasis* von  $V$ .



**Bemerkung 3.127:** Mit dem *Kronecker-Symbol*<sup>2</sup>

$$\delta_{k,l} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$$

gilt in jedem Orthonormalsystem  $\langle v_i, v_j \rangle = \delta_{i,j}$ .

**Beispiel 3.128:** Die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

sind zwar eine Basis, aber nicht orthogonal, da  $\langle a_1, a_3 \rangle = 1 \neq 0$ .

**Beispiel 3.129:** Die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

sind eine Basis und orthogonal, aber  $a_3$  ist nicht normiert. Also bilden die drei Vektoren eine Orthogonalbasis.

<sup>2</sup>Leopold Kronecker (1823-1891), dt. Mathematiker; Beiträge vornehmlich zur Algebra und Zahlentheorie

**Beispiel 3.130:** Die Vektoren

$$qa_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

sind orthogonal und normiert, bilden aber keine Basis. Es handelt sich also um ein Orthonormalsystem.

**Beispiel 3.131:** Die Vektoren

$$a_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind eine Basis, orthogonal und normiert. Sie bilden also eine Orthonormalbasis.

**Satz 3.132:** Ein Orthogonalsystem  $(v_1, \dots, v_n)$  ist linear unabhängig.

BEWEIS Sei  $\sum_{i=1}^n \lambda_i v_i = 0$ . Für  $v_j$  folgt

$$\left\langle v_j, \sum_{i=1}^n \lambda_i v_i \right\rangle = 0 \quad \Leftrightarrow \quad \sum_{i=1}^n \lambda_i \langle v_j, v_i \rangle = 0$$

und daraus aufgrund der Orthogonalität  $\lambda_j \langle v_j, v_j \rangle = 0$ . Wegen  $v_j \neq 0$  in jedem Orthogonalsystem gilt auch  $\langle v_j, v_j \rangle = \|v_j\|^2 > 0$  und daher  $\lambda_j = 0$  für jedes  $1 \leq j \leq n$ . ■

**Satz 3.133:** Ist  $\mathcal{B} = (v_1, \dots, v_n)$  eine Orthogonalbasis von  $V$ , dann gilt für jedes  $v \in V$ :

$$v = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k, \quad (3.18)$$

d.h.  $v$  hat bzgl.  $\mathcal{B}$  die Koordinaten  $(\frac{1}{\|v_k\|^2} \langle v, v_k \rangle, 1 \leq k \leq n)^T$ .

BEWEIS Da  $\mathcal{B}$  eine Basis ist, existiert immer eine Darstellung  $v = \sum_{k=1}^n \lambda_k v_k$ . Bildet man auf beiden Seiten das Skalarprodukt mit  $v_l$ , dann erhält man wegen  $\langle v_k, v_l \rangle = \delta_{k,l} \|v_k\|^2$ :

$$\langle v, v_l \rangle = \lambda_l \langle v_l, v_l \rangle = \lambda_l \|v_l\|^2 \quad \forall 1 \leq l \leq n. \quad \blacksquare$$

**Bemerkung 3.134:** Die  $k$ -te Koordinate eines Vektors bezogen auf eine Orthogonalbasis ist genau die Länge der Projektion des Vektors auf den  $k$ -ten Basisvektor.

Ist  $\mathcal{B}$  wie in Satz 3.133 sogar eine Orthonormalbasis, besitzt  $v$  die Koordinaten  $(\langle v, v_k \rangle, 1 \leq k \leq n)^T$  bezogen auf  $\mathcal{B}$ .



**Beispiel 3.135:** Die Vektoren

$$a_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

bilden eine Orthonormalbasis des  $\mathbb{R}^3$ . Der Vektor  $v = (5, 3, 7)^T$  lässt sich also als  $\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$  schreiben. Die Vorfaktoren  $\lambda_1, \lambda_2, \lambda_3$  ergeben sich aus:

$$\begin{aligned} \lambda_1 &= \langle v, a_1 \rangle = \frac{5}{\sqrt{2}} + \frac{3}{\sqrt{2}} = 4 \cdot \sqrt{2} \\ \lambda_2 &= \langle v, a_2 \rangle = \frac{5}{\sqrt{2}} - \frac{3}{\sqrt{2}} = \sqrt{2} \\ \lambda_3 &= \langle v, a_3 \rangle = 7 \end{aligned}$$

Also ist  $v = 4\sqrt{2} \cdot a_1 + \sqrt{2} \cdot a_2 + 7 \cdot a_3$ .

**Satz 3.136:** Sei  $\mathcal{B} = (v_1, \dots, v_m)$  ein Orthogonalsystem in  $V$  und  $U = L(\mathcal{B})$  der von  $\mathcal{B}$  aufgespannte Unterraum.

1. Für jedes  $v \in V$  gilt

$$p_U(v) = \sum_{i=1}^m \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i. \quad (3.19)$$

2. Jedes  $v \in V$  lässt sich eindeutig als Summe  $v = p_U(v) + w$  mit  $w \in U^\perp$  schreiben. Dabei gilt  $w = v - p_U(v)$ .

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

4. Sei  $\dim(V) = n$ . Dann gilt  $\dim(U) + \dim(U^\perp) = n$  für jeden Untervektorraum  $U$ .

**BEWEIS** 1.) Die Koordinaten  $(\alpha_1, \dots, \alpha_m)^T$  von  $p_U(v)$  sind Lösung des Gleichungssystems (3.17). Im Fall eines Orthogonalsystems erhält man

$$\begin{pmatrix} \langle v_1, v_1 \rangle & 0 & \dots & 0 \\ 0 & \langle v_2, v_2 \rangle & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \langle v_m, v_m \rangle \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \langle v, v_1 \rangle \\ \langle v, v_2 \rangle \\ \vdots \\ \langle v, v_m \rangle \end{pmatrix}$$

Es folgt  $(\alpha_1, \dots, \alpha_m)^T = \left( \frac{1}{\|v_1\|^2} \langle v, v_1 \rangle, \dots, \frac{1}{\|v_m\|^2} \langle v, v_m \rangle \right)^T$  (man beachte dabei  $\|v_i\| > 0$  wegen  $v_i \neq 0$  nach Definition eines Orthogonalsystems).

Einsetzen in Formel (3.16) liefert die Behauptung.

2.) Es gilt nach Definition 3.120 der orthogonalen Projektion  $w \perp u \forall u \in U$ , also  $w \in U^\perp$ . Die Eindeutigkeit der orthogonalen Projektion (Folgerung 3.124) impliziert die Eindeutigkeit der Zerlegung von  $v$ .

3.) Folgt aus der Definition 3.33 der direkten Summe und 2.).

4.) erhält man unmittelbar aus Folgerung 3.65. ■

**Beispiel 3.137:**

$$\mathcal{B} = \left( \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \right)$$

ist sogar ein Orthonormalsystem. Seine lineare Hülle ist

$$U = L(\mathcal{B}) = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}; a, b \in \mathbb{R} \right\}.$$

Das Orthogonale Komplement  $U^\perp = \{v \in V; v \perp u \forall u \in U\}$  ist

$$U^\perp = \left\{ \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}; c \in \mathbb{R} \right\}.$$

### 3.8 Das Verfahren von Gram-Schmidt und Anwendungen

Im letzten Kapitel wurde deutlich, dass eine orthogonale Projektion eines abstrakten Vektors immer und relativ einfach möglich ist, wenn von dem Unterraum, auf den projiziert werden soll, eine Orthogonal-, besser noch eine Orthonormalbasis bekannt ist. Es ist aber nicht sofort klar, ob jeder endlich erzeugte unitäre Vektorraum eine Orthonormalbasis besitzt und wie man, wenn ja, eine solche konstruieren kann. Beide Fragen werden mithilfe des *Orthonormalisierungsverfahrens nach Gram-Schmidt* positiv beantwortet. Das Verfahren wurde unabhängig voneinander von Schmidt<sup>3</sup> und Gram (1879) veröffentlicht. Beide gelten als „Entdecker“ des Verfahrens, allerdings wurde das Verfahren schon 1836 von Cauchy benutzt.

Gegeben seien  $m$  linear unabhängige Vektoren  $v_1, \dots, v_m$  eines unitären Vektorraums. Das Verfahren von Gram-Schmidt erzeugt aus diesen Vektoren ein Orthonormalsystem  $w_1, \dots, w_m$  mit

$$L(v_1, \dots, v_m) = L(w_1, \dots, w_m),$$

<sup>3</sup>Erhard Schmidt (1876-1959), dt. Mathematiker; bedeutende Beiträge zur Funktionalanalysis, Entwicklung des nach ihm benannte Verfahrens im Kontext nicht-endlich erzeugter unitärer Vektorräume (1907)

also eine Orthonormalbasis des Untervektorraums  $L(v_1, \dots, v_m)$ . Zur Orthogonalisierung wird die orthogonale Projektion von Basisvektoren verwendet. Wir betrachten zur Veranschaulichung eine beliebige Basis  $(v_1, v_2, v_3)$  von  $\mathbb{R}^3$ . Wir gehen so vor:

1. Man wählt einfach  $w_1 = \frac{1}{\|v_1\|}v_1$ , weil dann offenbar  $\|w_1\| = 1$  gilt.
2. Wir konstruieren einen Vektor  $r_2$ , der senkrecht auf  $w_1$  steht. Dazu projizieren wir  $v_2$  auf den von  $w_1$  erzeugten Unterraum  $L(w_1)$  und setzen  $r_2 := v_2 - p_{L(w_1)}(v_2)$ . Dann gilt nach Definition der orthogonalen Projektion  $v_2 \perp r_2$ . Nach Satz 3.136 erhält man

$$r_2 = v_2 - \langle v_2, w_1 \rangle w_1. \quad (3.20)$$

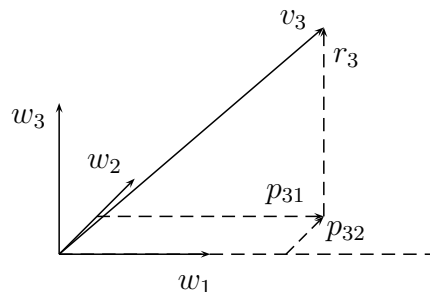
Aufgrund der linearen Unabhängigkeit von  $v_1$  und  $v_2$  gilt  $r_2 \neq 0$  wegen  $v_2 \notin L(w_1)$ . Normierung von  $r_2$  liefert  $w_2$ . Weil  $r_2$  und damit  $w_2$  eine Linearkombination von  $v_1$  und  $v_2$  ist, gilt  $L(w_1, w_2) \subseteq L(v_1, v_2)$ . Es sind aber  $w_1$  und  $w_2$  orthonormal und damit nach Satz 3.132 linear unabhängig, also gilt auch  $\dim(L(w_1, w_2)) = \dim(L(v_1, v_2))$  und daher

$$L(w_1, w_2) = L(v_1, v_2).$$

3. Der Vektor  $r_3 := v_3 - P_{L(v_1, v_2)}(v_3)$  steht nach Konstruktion senkrecht auf  $L(v_1, v_2)$  und daher gilt insbesondere  $r_3 \perp v_1$  und  $r_3 \perp v_2$ . Aufgrund der linearen Unabhängigkeit der  $v_i$  liegt  $v_3$  nicht in  $L(v_1, v_2) = L(w_1, w_2)$ , und daher  $r_3 \neq 0$ . Wir finden  $w_3$  durch Normierung von  $r_3$ . Weil  $(w_1, w_2)$  nach 2.) eine Orthonormalbasis von  $L(v_1, v_2)$  bilden, gilt nach Satz 3.136:

$$r_3 = v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2$$

Das folgende Bild veranschaulicht die Konstruktion von  $w_3$ .



Man kann die obige Konstruktion leicht verallgemeinern.

**Satz 3.138 (Gram-Schmidtsches Orthogonalisierungsverfahren):** Sei  $V$  ein unitärer Vektorraum und  $v_1, \dots, v_m$  linear unabhängig. Seien

$$\begin{aligned} w_1 &:= \frac{v_1}{\|v_1\|} \\ r_{k+1} &:= v_{k+1} - \sum_{i=1}^k \langle v_{k+1}, w_i \rangle w_i \\ w_{k+1} &= \frac{r_{k+1}}{\|r_{k+1}\|} \end{aligned}$$

Dann bilden  $(w_1, \dots, w_m)$  eine Orthonormalbasis von  $L(v_1, \dots, v_m)$ .

**BEWEIS** Man zeigt obigen Satz per vollständiger Induktion. Der Induktionsanfang entspricht Schritt 2.), der Induktionsschluss verläuft analog zu Schritt 3.) mit etwas allgemeineren Indizes. ■

**Beispiel 3.139:**

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix}$$

Berechnung von  $w_1$ :

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Berechnung von  $w_2$ :

$$\begin{aligned} r_2 &= v_2 - \langle v_2, w_1 \rangle w_1 \\ &= \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} - \frac{1}{3}(3 + 8 + 10) \cdot \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} - \frac{7}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \\ &= \frac{1}{3} \left[ \begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix} - \begin{pmatrix} 7 \\ 14 \\ 14 \end{pmatrix} \right] = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \\ w_2 &= \frac{r_2}{\|r_2\|} = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \end{aligned}$$

Berechnung von  $w_3$ :

$$\begin{aligned}
 r_3 &= v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2 \\
 &= \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{3} \cdot 11 \cdot \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{1}{3} \cdot 13 \cdot \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \\
 &= \frac{1}{9} \left[ \begin{pmatrix} 63 \\ 9 \\ 9 \end{pmatrix} - \begin{pmatrix} 11 \\ 22 \\ 22 \end{pmatrix} - \begin{pmatrix} 26 \\ -26 \\ 13 \end{pmatrix} \right] \\
 &= \frac{1}{9} \begin{pmatrix} 26 \\ 13 \\ -26 \end{pmatrix} = \frac{13}{9} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \\
 w_3 &= \frac{r_3}{\|r_3\|} = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}
 \end{aligned}$$

**Folgerung 3.140:** Jeder endlich erzeugte unitäre Vektorraum besitzt eine Orthonormalbasis.

**Bemerkung 3.141:** Folgerung 3.140 ist in nicht endlich erzeugten Vektorräumen i. A. falsch.

Wir holen jetzt den Existenzbeweis der orthogonalen Projektion nach.

**Folgerung 3.142:** Sei  $V$  ein unitärer Vektorraum und  $U$  ein endlich erzeugter Untervektorraum. Dann existiert für jedes  $v \in V$  die orthogonale Projektion  $p_U(v)$  von  $v$  auf  $U$ .

**BEWEIS** Sei  $U = L(v_1, \dots, v_m)$ . Mithilfe des Gram-Schmidt-Verfahrens findet man eine Orthonormalbasis  $(w_1, \dots, w_m)$  von  $U$ . Anwendung von Satz 3.136 ergibt die Behauptung. ■

**Folgerung 3.143:** Sei  $V$  ein endlich erzeugter unitärer Vektorraum und  $U$  irgend ein Untervektorraum. Dann gilt  $V = U \oplus U^\perp$ , und  $\dim(V) = \dim(U) + \dim(U^\perp)$ .

**BEWEIS** Die Aussage wurde bereits in Satz 3.136 unter der Voraussetzung gezeigt, dass  $U$  durch ein Orthogonalsystem erzeugt wird. Mithilfe des Gram-Schmidt-Verfahrens lässt sich dies für jeden Unterraum  $U$  von  $V$  konstruieren, so dass wir jetzt auf die dort genannte Voraussetzung verzichten können. ■

Jetzt holen wir den Beweis dafür nach, dass jede Hyperebene in  $\mathbb{R}^n$  eine Normaldarstellung besitzt (vgl. Satz 3.79).

**Folgerung 3.144:** Jede Hyperebene in  $\mathbb{R}^n$  besitzt eine Normaldarstellung; der Normalenvektor ist bis auf Skalierung eindeutig.

**BEWEIS** Jede Hyperebene in  $\mathbb{R}^n$  besitzt die Darstellung  $H = p + U$  mit  $\dim(U) = n - 1$ . Es gilt somit  $\dim(U^\perp) = 1$ , also existiert  $v \in \mathbb{R}^n \setminus \{0\}$  mit  $U^\perp = L(v)$ . Dies ist der gesuchte Normalenvektor. Jeder andere Normalenvektor muss auch in  $U^\perp$  liegen und ist daher ein Vielfaches von  $v$ . ■

**Bemerkung 3.145:** Das Ergebnis der Gram-Schmidt-Verfahrens  $w_1, \dots, w_m$  hängt von der Reihenfolge der Ausgangsvektoren  $v_1, \dots, v_m$  ab. Vertauscht man diese, erhält man i. A. nicht etwa  $w_1, \dots, w_m$  in vertauschter Reihenfolge. Die Aussagen von Satz 3.138 bleiben natürlich unverändert gültig.

**Satz 3.146:** Sei  $V$  wie oben und  $v_1, \dots, v_m \in V$ . Gelingt es, aus diesen mithilfe des Gram-Schmidt-Verfahrens orthonormale Vektoren  $w_1, \dots, w_m$  zu erzeugen, dann sind  $(v_1, \dots, v_m)$  linear unabhängig.

**BEWEIS** Andernfalls existiert  $v_k$  mit  $v_k \in L(v_1, \dots, v_{k-1})$ . Dann gilt  $v_k = p_{L(v_1, \dots, v_{k-1})}(v_k)$  und daher  $r_k = 0$ . Damit ist  $r_k$  nicht normierbar, und das Gram-Schmidt-Verfahren bricht ab. ■

Es ist also nicht erforderlich, die lineare Unabhängigkeit der Ausgangsvektoren bei der Orthonormalisierung nach Gram-Schmidt zu prüfen, weil das im Laufe der Rechnung ohnehin klar wird.

**Beispiel 3.147:** Wir betrachten  $P_2([-1, 1])$  mit der bekannten Basis  $v_1(x) = 1, v_2(x) = x, v_3(x) = x^2$  und wollen daraus eine Orthonormalbasis bezogen auf das Skalarprodukt (3.10) konstruieren. Wir verwenden dazu das Verfahren von Gram-Schmidt und normieren zunächst  $v_1$ . Man hat  $\|1\|_{L_2}^2 = \int_{-1}^1 1^2 dx = 2$ , und daher

$$w_1(x) = \frac{1}{\sqrt{2}}v_1(x). \quad (3.21)$$

Für  $w_2(x)$  erhält man:

$$\begin{aligned} r_2(x) &= v_2(x) - \langle v_2(x), w_1(x) \rangle w_1(x) \\ &= x - \int_{-1}^1 x \cdot \frac{1}{\sqrt{2}} dx \cdot \frac{1}{\sqrt{2}} \\ &= x - \frac{1}{2} \left[ \frac{1}{2} x^2 \right]_{-1}^1 = x \end{aligned}$$

und weiter  $\|x\|_{L_2}^2 = \int_{-1}^1 x \cdot x dx = \frac{2}{3}$ , also

$$w_2(x) = \frac{r_2(x)}{\|r_2(x)\|_{L_2}} = \sqrt{\frac{3}{2}}x \quad (3.22)$$

Mit einer analogen Rechnung erhält man

$$w_3(x) = \frac{\sqrt{10}}{4}(3x^2 - 1). \quad (3.23)$$

Orthonormale Funktionensysteme wie in Beispiel 3.147 konstruiert spielen in der angewandten Mathematik eine große Rolle.

In der Praxis kommt es häufig vor, dass man eine komplizierte Funktion durch eine einfache Funktion ersetzen möchte, weil diese vielleicht effizienter zu berechnen ist oder in der Praxis besser handhabbar ist (Wie würde man z. B.  $\sin(35^\circ)$  ohne Hilfsmittel praktisch berechnen?). Dabei soll natürlich der Fehler durch diese Ersetzung möglichst klein sein. Man spricht hier von *Approximation*. Die mathematische Disziplin der Approximationstheorie beschäftigt sich mit der möglichst geschickten Konstruktion approximierender Funktionen und derartiger Fehlerschranken. Es liegt nahe, als approximierende Funktionen Polynome aufgrund ihrer Einfachheit zu verwenden. Wir wollen prototypisch die Funktion

$$f(x) = \frac{1}{1+x^2} \quad (3.24)$$

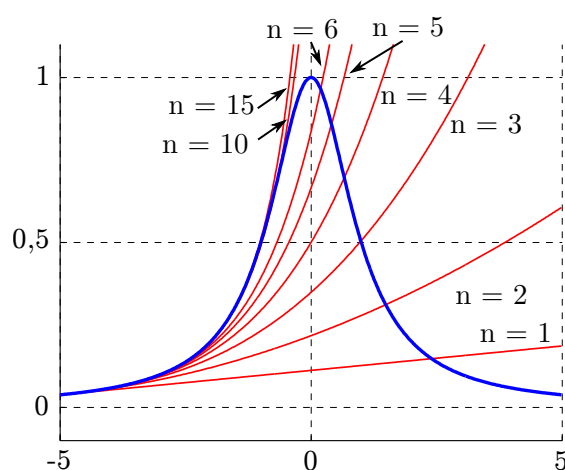
auf dem Intervall  $[-5, 5]$  durch Polynome vom Höchstgrad  $n$  approximieren. Dazu werden wir drei Ansätze vergleichen:

1. Approximation durch Taylorpolynome mit Grad  $n$  (Polynome  $t_n$ )
2. Approximation durch Interpolationspolynome mit Grad  $n$  (Polynome  $i_n$ )
3. Approximation durch orthogonale Projektion auf  $P_n$  (Polynome  $p_n$ )

Wir führen die durchaus umfangreichen Berechnungen in MATLAB aus. Die  $n+1$  Interpolationspunkte werden auf dem Intervall gleichabständig verteilt, für die Taylorpolynome verwenden wir den Entwicklungspunkt  $x_0 = -5$ .

In Tabelle 3.1 werden die Approximationsfehler  $\|f - t_n\|_{L_2}$ ,  $\|f - i_n\|_{L_2}$  und  $\|f - p_n\|_{L_2}$  für die drei Ansätze für einige Polynomgrade  $n$  gezeigt. Man erkennt, dass allein die orthogonale Projektion von  $f$  auf  $P_n$  brauchbare polynomielle Näherungen an  $f$  hervorbringt; insbesondere scheint der Approximationsfehler kleiner zu werden, je größer  $n$  gewählt wird. Es gilt in der Tat  $\|f - p_n\|_{L_2} \rightarrow 0, n \rightarrow \infty$ , ohne hier genau zu erklären, warum das so ist. In diesem Fall lässt sich der Approximationsfehler kleiner als jede vom Anwender gewählte Toleranz machen, wenn man nur den Polynomgrad hoch genug wählt. Solche Aussagen gelten für die anderen beiden Ansätze offenbar nicht. Der Vergleich der verschiedenen Polynome in den Abbildungen 3.6 bis 3.8 zeigt, dass die Taylor-Polynome zwar eine

$n$	$\ f - t_n\ _{L_2}$	$\ f - i_n\ _{L_2}$	$\ f - p_n\ _{L_2}$
1	1,045	1,170	0,9007
2	1,108	1,438	0,9007
3	2,037	0,7712	0,6092
4	4,483	0,8829	0,6092
5	9,463	0,462	0,4103
6	19,02	0,8663	0,4103
10	218,4	1,835	0,1857
15	1074	-	0,0564

Tabelle 3.1:  $L_2$ -Normen der Approximationsfehler an  $f$ Abbildung 3.6:  $f$  (blau dargestellt) und Taylorpolynome  $t_n$ 

exzellente Approximation an  $f$  in der Nähe des Entwicklungspunktes liefern, aber eben nicht auf dem ganzen Intervall  $[-5, 5]$ . Die interpolierenden Polynome zeigen starke Oszillationen, so dass sie zwar den Wert von  $f$  in den Interpolationspunkten genau treffen, ansonsten aber den Verlauf von  $f$  nicht nachbilden. Dieses Oszillationsphänomen tritt in der Praxis eher selten auf und lässt sich durch eine günstigere Wahl von Interpolationspunkten wesentlich reduzieren, aber man wird von einem Verfahren, das nur meistens funktioniert, in der Praxis doch eher Abstand nehmen. Wir werden jetzt die Überlegenheit der orthogonalen Projektion erklären.

**Definition 3.148:** Sei  $V$  ein unitärer Vektorraum und  $v \in V$  sowie  $M \subset V$  eine beliebige nichtleere Menge. Ein Vektor  $v^* \in M$  heißt *Bestapproximation* in  $M$  an  $v$ , falls

$$\|v^* - v\| = \inf_{x \in M} \|x - v\|.$$



3.8. DAS VERFAHREN VON GRAM-SCHMIDT UND ANWENDUNGEN 121

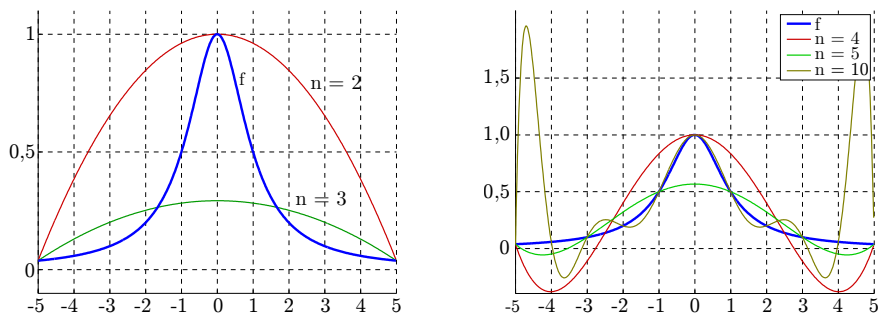


Abbildung 3.7:  $f$  (blau dargestellt) und Interpolationspolynome  $i_n$

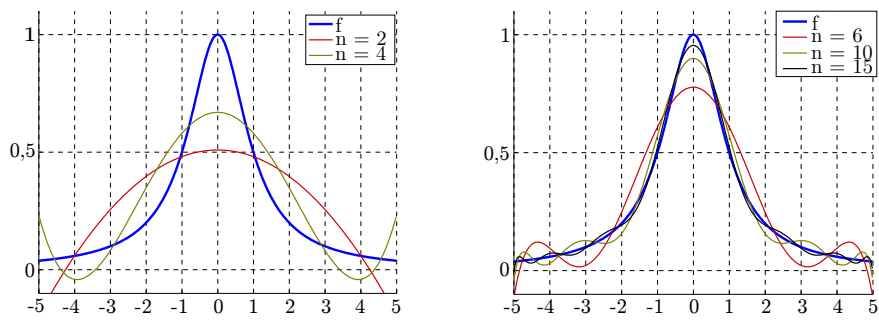


Abbildung 3.8:  $f$  (blau dargestellt) und  $p_n$

Für allgemeine Mengen muss es eine Bestapproximation an  $v$  weder geben, noch muss sie eindeutig sein. Anders ist das, wenn  $M$  ein endlich erzeugter Untervektorraum ist.

**Satz 3.149 (Bestapproximation):** Sei  $V$  wie oben,  $v \in V$  und  $U$  ein endlich erzeugter Untervektorraum von  $V$ . Dann gilt

$$\|v - p_U(v)\| = \min_{u \in U} \|u - v\|, \quad (3.25)$$

die orthogonale Projektion von  $v$  auf  $U$  ist die einzige Bestapproximation an  $v$  in  $U$ .

**BEWEIS** Sei  $u \in U$  beliebig. Dann ist  $u - P_U(v) \in U$  und  $v - P_U(v) \in U^\perp$ , und daher  $u - P_U(v) \perp v - P_U(v)$ . Nach dem Satz des Pythagoras gilt dann

$$\|u - v\|^2 = \|u - p_U(v)\|^2 + \|v - p_U(v)\|^2 \geq \|v - p_U(v)\|^2. \quad (3.26)$$

Somit ist  $p_U(v)$  eine Bestapproximation. Gleichheit in Formel (3.26) gilt nur, wenn  $\|u - P_U(v)\| = 0$ , also  $u = p_U(v)$  gilt. Daher ist  $p_U(v)$  die einzige Bestapproximation in  $U$  an  $v$ . ■

Nach Satz 3.149 ist der Approximationsfehler gemessen in der  $L_2$ -Norm der  $p_n$  an  $f$  bei festem Grad nicht mehr zu unterbieten; die durch orthogonale Projektion gewonnenen  $p_n$  stellen die bestmögliche Wahl dar.

**Folgerung 3.150:** Wegen  $P_n \subset P_{n+1}$  folgt  $\|f - p_{P_{n+1}}(f)\|_{L_2} \leq \|f - p_{P_n}(f)\|_{L_2}$  nach Satz 3.149; erhöht man den Polynomgrad, wird also der Approximationsfehler auf keinen Fall größer (aber auch nicht notwendig kleiner, siehe Tabelle 3.1).

**Bemerkung 3.151:** Man kann Bestapproximation bezüglich beliebiger Normen betrachten, so liegt es hier durchaus nahe, statt der  $L_2$ -Norm die Maximumsnorm auf  $\mathcal{C}[a, b]$  zu wählen. Das Fehlermaß ist dann die maximale punktweise Abweichung von  $p_n$  zu  $f$ . Die Maximumsnorm wird aber nach Bemerkung 3.116 von keinem Skalarprodukt induziert, so dass ein Zugang über Orthogonalität wie oben unmöglich ist. Tatsächlich sind dann sowohl praktische Rechnungen als auch die theoretischen Betrachtungen dazu ungleich komplizierter als bei einer Norm, die von einem Skalarprodukt induziert wird.

# Index

- Abbildung
  - wohldefiniert, 68
- Abstand, 57
- Äquivalenzumformungen, 19
- allgemeine Ebenengleichung, 49
- allgemeine Geradengleichung, 47
- Assoziativität, 67
- Aufpunkt, 43
- Austauschlemma, 86
  
- Basis, 84, 90
- Basisergänzungssatz, 85
- Betragssummennorm, 31
  
- Cauchy-Schwarzsche Ungleichung, 106
  
- Determinante, 61
- Differenzpolynom, 99
- Differenzvektor, 14
- Dimension, 84
- direkte Summe, 77
- Distributivgesetz, 72
- Dreiecksmatrix
  - obere, 17
  - untere, 17
- Dreiecksungleichung, 30, 37, 105
  
- Ebene, 11, 49
  - Normalform, 49
  - Punkt-Richtungsgleichung, 49
  - Richtungsvektor, 49
- Eindeutigkeit der Basislänge, 87
- Einernorm, 31
- Einheitsmatrix, 17
- Einheitssphäre, 32
- Einheitsvektor, 32
- Eisen, 5
  
- Eisenerzeugung, 5
- Entwicklungsprozess, 7
- Erzeugendensystem, 79, 89
  - minimal, 84, 90
- Euklidischer Vektorraum, 104
  
- Faktorzerlegung, 96
- Familie, 88
- Folge, 27
- Fries, 8
  
- Gauß-Algorithmus, 18
- Gauß-Jordan-Algorithmus, 18
- Gerade, 43
  - Aufpunkt, 43
  - Normalform, 46
  - Parameterform, 45
  - Punkt-Richtungsgleichung, 43
  - Richtungsvektor, 43
  - Zweipunktform, 45
- Geradengleichung
  - allgemeine, 47
- Gleichungssystem
  - äquivalent, 18
  - Äquivalenzumformungen, 19
  - linear, 18
  - Lösung mit MATLAB, 20
  - reduzierte Stufenform, 18
  - Stufenform, 18
  - überbestimmtes, 22
  - unterbestimmtes, 22
- Graßmannsche Identität, 42
- Grad, 94
- Gram-Schmidt, 114
- Gruppe, 7, 67
  - abelsch, 68

- Automorphismengruppe, 8
  - kommutativ, 68
  - Raumgruppe, 8
  - triviale, 69
- Gruppenisomorphismus, 71
- Hessesche Normalform, 47
  - Ebene, 50
- Hyperebene, 90, 91
  - Aufpunkt, 91
  - Parameterform, 91
- Indexmenge, 27
- Input-Output-Analyse, 12
- Interpolation, 99
- Interpolationspolynom, 99
- inverses Element, 67
- isomorph (Gruppe), 71
- Körper, 72
- kanonische Einheitsvektoren, 32
- Kartesisches Produkt, 27
- Koeffizienten, 94
- Koeffizientenmatrix, 15
  - erweiterte, 16
- Körper, 72
- komplexe Ebene, 72
- komplexer Vektorraum, 74
- Komponente eines Vektors, 35
- Koordinaten, 84
- Koordinatensystem, 11
- Kreuzprodukt, 40
- Kristallographie, 8
- Kronecker-Symbol, 111
- Kurbel, 6
- Legierung, 5
- Leitkoeffizient, 94
- Lineare Gleichungssysteme, 18
- Lineare Hülle, 78
- Lineare Unabhängigkeit, 79
- Lineares Gleichungssystem, 6, 15, 65
  - überbestimmtes, 22
  - Koeffizienten, 15
  - unterbestimmtes, 22
- Linearkombination, 78
- Linkssystem, 39
- Lotfußpunkt, 58
- $M(m \times n, K)$ , 75
- Matrix, 15
  - äquivalent, 18
  - Äquivalenzumformungen, 19
  - Diagonal-, 17
  - Einheits-, 17
  - Elemente, 16
  - Gleichheit, 16
  - Hauptdiagonale, 17
  - Koeffizienten-, 15
  - MATLAB, 20
  - Null-, 16
  - obere Dreiecks-, 17
  - quadratisch, 17
  - Spaltenindex, 16
  - transponiert, 16
  - untere Dreiecks-, 17
  - Zeilenindex, 16
- Maximumnorm, 31
- minimales Erzeugendensystem, 84, 90
- Monombasis, 98
- Neutralelement, 67
- Norm, 30, 105
  - Betragssummen-, 31
  - Einer-, 31
  - euklidische, 29
  - Maximum-, 31
- Normalenvektor, 46
  - Ebene, 49
- Normalform, 46
  - Ebene, 49
- Nullpolynom, 94
- numerische Lineare Algebra, 7
- Numerische Simulation, 6
- Octave, 20
- Orthogonalbasis, 111
- orthogonale Projektion, 36
- orthogonales Komplement, 108
- Orthogonalität

- Ebene, 53
- von Geraden, 52
- Orthogonalität von Vektoren, 34
- Orthogonalsystem, 110
- Orthonormalbasis, 111
- Orthonormalisierungsverfahren, 114
- Orthonormalsystem, 111
- Ortsvektor, 12
- Parallelepiped, 63
- Parallelität
  - Ebene, 53
  - von Geraden, 52
- Parallelität von Vektoren, 15
- Parameterform, 45
- physikalisches System, 6
- Pivot-Element, 22
- Pivot-Spalte, 22
- Polynom, 94
  - Faktorzerlegung, 96
  - Grad, 94
  - Interpolations-, 99
  - Koeffizienten, 94
- Problemlösestrategien, 8
- Projektion, 35
  - orthogonale, 35
  - senkrechte, 35
- Prototyp, 6
- Punkt-Richtungsgleichung
  - Ebene, 49
  - Gerade, 43
- Punktprodukt, 28
- Rückwärtsarbeiten, 9
- Rechte-Hand-Regel, 39
- Rechtssystem, 39
- reeller Vektorraum, 74
- Restklasse, 70
- Richtungsvektor, 43
- Sarrussche Regel, 62
- Satz des Pythagoras, 34
- Schnittmengen, 54
- Schwarzsche Ungleichung, 36
- Scilab, 20
- Skalar, 74
- Skalarprodukt, 27, 28, 102
  - euklidisches, 28
- span, 79
- Spat, 63
- Spatprodukt, 63
- Spiegelpunkt, 58
- Standardnorm, 29, 105
- Standardskalarprodukt, 28
- Starrkörperbewegungen, 8
- Stufenform, 18
- Summe (von Vektorräumen), 77
- überbestimmtes Gl.-system, 22
- Unitärer Vektorraum, 104
- unterbestimmtes Gl.-system, 22
- Untergruppe, 71
- Unterraum, 75
- Untervektorraum, 75
- Vektor, 10, 11, 74
  - Addition, 13
  - Differenz-, 14
  - Einheits-, 32
  - gebunden, 12
  - Gleichheit, 11
  - Multiplikation mit Skalar, 13
  - normiert, 33
  - orthogonal, 34
  - parallel, 15
  - senkrecht, 34
  - transponiert, 11
  - Winkel, 37
- Vektorprodukt, 40
- Vektorraum, 74
  - endlich erzeugt, 79
  - komplexer, 74
  - reeller, 74
- Verknüpfung, 67
- Verknüpfungstafel, 70
- Verschiebungsvektor, 12
- Volumen, 63
- Vorwärtsarbeiten, 9
- windschief, 56

- Winkel
  - Ebene, [53](#)
  - zwischen Geraden, [52](#)
- Winkel zwischen Vektoren, [37](#), [38](#)
- Wirtschaftswissenschaften, [11](#)
  
- Zahlraum, [11](#)
- Zweipunktform, [45](#)
- Zwischenziele, [10](#)

# Literaturverzeichnis

- [1] A. Beutelspacher. *Das ist o. B. d. A. trivial!* Vieweg, 8 edition, 2006.
- [2] A. Beutelspacher. *Lineare Algebra*. Springer Spektrum, 8 edition, 2014.
- [3] D. Grieser. *Mathematisches Problemlösen und Beweisen*. Springer Spektrum, 2013.
- [4] W. Kabbalo. *Einführung in die Analysis 1*. Spektrum akademischer Verlag, 2 edition, 2000.
- [5] G. Pólya. *Schule des Denkens*. Narr Francke Attempto, 2010. Sonderausgabe der 4. Auflage.