LINEARE ALGEBRA

Enrico Leuzinger

Version vom 9. August 2016

Institut für Algebra und Geometrie

Karlsruher Institut für Technologie (KIT)

E-Mail: enrico.leuzinger@kit.edu

Copyright © 2016 Enrico Leuzinger. All rights reserved.

Inhaltsverzeichnis

Teil I

Einführung

1 Gebrauchsanweisung für dieses Skript

Die Lehrveranstaltung *Lineare Algebra* hat drei Bestandteile:

- **Vorlesung**
- **Übung**
- **Tutorium**.

Die *Vorlesung* ist eine „Führung durch die Theorie“: der Lern-Stoff wird präsentiert, die Theorie erklärt und kommentiert.

Das *Skript* erspart Ihnen das Mitschreiben in der Vorlesung und schafft so Raum für das Mitdenken. Den größten Nutzen haben Sie, wenn Sie sich mit dem Abschnitt, der jeweils gerade in der Vorlesung behandelt wird, schon vorher vertraut machen (Zeitaufwand: 30-60 Minuten). In der Vorlesung können Sie dann gezielt Notizen machen oder Fragen stellen. Übrigens: Wenn Sie einen mathematischen Text (z.B. dieses Skript) „lesen“, sollten Sie das nicht passiv, sondern aktiv mit Stift und Papier tun. Notieren Sie sich Definitionen stichwortartig. Eine neue Definition können Sie sich viel besser merken, wenn Sie ein (möglichst einfaches) Beispiel/Gegenbeispiel dazu kennen. Notieren Sie sich auch diese Beispiele. Machen Sie Sich den Inhalt von (Lehr-)Sätzen ebenfalls immer an eigenen Beispielen klar. Rechnen Sie die Beispiele im Text selber durch.

In diesem Skript sind Definitionen, Beispiele und Sätze durchnummeriert. Das soll das Verweisen in der Vorlesung erleichtern: Sie werden jederzeit genau wissen, welche Stelle gerade besprochen wird.

Die *Übungen* dienen dazu, das Verständnis zu vertiefen und die Theorie auf konkrete (mathematische) Probleme anzuwenden. Wie beim Erlernen eines Instruments oder eines Handwerks gilt auch in der Mathematik: **Die Beherrschung dieser Wissenschaft ist nur durch konstante Anstrengung und eigene Aktivität möglich**. Genau dazu sind die Übungen da. In den *Tutorien* besteht die Möglichkeit, in kleineren Gruppen gemeinsam zu üben, zu lernen und Erfahrungen auszutauschen.

2 How to solve it?

Das Lösen von (mathematischen) Problemen ist eine Kunst, die neben Erfolgserlebnissen auch mit Frustrationen verbunden ist. Gerade für Studienanfänger stellt sich immer wieder die Frage: *Wie findet man die Lösung einer Aufgabe?* Leider gibt es dafür kein Patentrezept. Wie so oft braucht es neben Talent auch Ausdauer und Erfahrung. Der Mathematiker Georg Polya hat sich dennoch überlegt, wie eine erfolgreiche Problemlösungs-Strategie aussehen könnte. Hier seine Tipps (vgl. [18]), die Ihnen vielleicht helfen, weiter zu kommen:

1. Vorbereitung: die Aufgabe verstehen.

- Verstehen Sie die Fragestellung? Kennen Sie die vorkommenden Begriffe und Konzepte?
- Was ist gesucht? Was ist gegeben? Wie lauten die Voraussetzungen oder Bedingungen, wie die Behauptung?
- Ist es möglich, die Bedingung zu befriedigen? Ist die Bedingung ausreichend, um die Unbekannte zu bestimmen? Oder genügt sie nicht? Ist sie eventuell sogar widersprüchlich?
- Zeichen Sie Figuren und machen Sie Skizzen! Führen Sie passende Bezeichnungen ein!
- Trennen Sie die verschiedenen Teile der Voraussetzung! Können Sie sie hinschreiben?

2. Brainstorming: Einen Zusammenhang zwischen Gegebenem und Gesuchtem finden und einen Plan für die Lösung ausdenken.

- Haben Sie die Aufgabe schon früher gesehen? Oder haben Sie dasselbe Problem in einer ähnlichen Form gesehen?
- Kennen Sie eine verwandte Aufgabe? Kennen Sie einen Lehrsatz, der hilfreich sein könnte?
- Betrachten Sie die Voraussetzungen! Versuchen Sie, sich auf eine Ihnen bekannte Aufgabe zu besinnnen, die dieselben oder ähnliche Voraussetzungen hatte.
- Hier ist eine Aufgabe, die der Ihren verwandt ist und deren Lösung Sie kennen. Können Sie ihre Methode verwenden? Würden Sie irgend ein Hilfsmittel einführen, damit Sie sie verwenden können?
- Können Sie die Aufgabe anders ausdrücken? Können Sie sie auf noch verschiedenere Weise ausdrücken? Gehen Sie auf die Definition zurück!

- Wenn Sie die vorliegende Aufgabe nicht lösen können, so versuchen Sie, zuerst eine verwandte Aufgabe zu lösen. Können Sie Sich eine zugänglichere, verwandte Aufgabe denken? Eine allgemeinere Aufgabe? Eine analoge Aufgabe? Können Sie einen Teil der Aufgabe lösen? Behalten Sie nur einen Teil der Bedingungen bei und lassen Sie den andern weg; wie weit ist die Unbekannte/Behauptung dann bestimmt, wie kann man sie verändern? Können Sie etwas Nützliches aus den Daten ableiten? Können Sie sich andere Daten denken, die geeignet sind, die Unbekannte zu bestimmen? Können Sie die Unbekannte ändern oder die Daten oder, wenn nötig, beides, so dass die neue Unbekannte und die neuen Daten einander näher sind?

- Haben Sie alle Daten benutzt? Haben Sie die ganze Bedingung benutzt? Haben Sie alle wesentlichen Begriffe in Betracht gezogen, die in der Aufgabe enthalten sind?

3. Ausarbeitung und Kontrolle: Den Plan ausführen und die Lösung prüfen.

- Wenn Sie Ihren Plan der Lösung durchführen, so kontrollieren Sie jeden Schritt. Können Sie deutlich sehen, dass der Schritt richtig ist? Können Sie beweisen, dass er richtig ist?

- Können Sie das Resultat kontrollieren? Können Sie den Beweis kontrollieren?

- Können Sie das Resultat auf verschiedene Weise ableiten? Können Sie es auf den ersten Blick sehen?

- Können Sie das Resultat oder die Methode für irgend eine andere Aufgabe gebrauchen?

3 Was ist lineare Algebra?

Die Frage „Was ist Mathematik?“ ist schwierig zu beantworten und verschiedene Mathematiker haben verschiedene Antworten gegeben. Ein (etwas verstaubter) Klassiker ist Courant-Robbins [4]. Moderner und spannender sind Devlin [6], Davis-Hersh [5] und Hersh [14]. Siehe auch Gowers [11] und Otte [17]. Gegenüber anderen Wissenschaften zeichnen sich die Begriffssysteme und Theorien, die in der Mathematik entwickelt werden, durch drei spezifische Merkmale aus:

1. **Abstraktheit**: Gegenstand der Mathematik sind Systeme von Objekten mit fixierten strukturellen Beziehungen untereinander. Diese Strukturen oder Muster stehen im Vordergrund; von allen weiteren Eigeschaften der Objekte wird abgesehen (abstrahiert).

2. **Genauigkeit**: Ist eine mathematische Struktur (axiomatisch) fixiert, so sind alle Aussagen über diese Struktur durch formales, logisches Schließen aus den einmal gemachten Annahmen ableitbar. Wie man das konkret macht, ist allerdings eine

Kunst, die neben dem Beherrschen der mathematischen Techniken vor allem Intuition und Einsicht in das Wesen der Sache erfordert (also etwas ganz anderes als Logik); siehe dazu z.B. die Bücher von Hadamard [12] und Ruelle [19].

3. **Allgemeinheit**: Ausgangspunkt für den Abstraktionsprozess und die Entwicklung einer mathematischen Struktur ist zwar oft ein konkretes (z.B. physikalisches) Problem oder Phänomen. Alle Aussagen, die über eine Struktur gewonnen werden, sind aber später in allen Situationen anwendbar, in denen Strukturen mit den gleichen Bedingungen vorliegen. Darauf beruht die universelle Anwendbarkeit und Effizienz von Mathematik in andern Wissenschaften.

Diese Besonderheiten sind natürlich auch ein Grund dafür, weshalb das Erlernen von Mathematik nicht so ganz einfach ist.

Wie die Frage „Was ist Mathematik?“ lässt sich auch die Frage „Was ist lineare Algebra?“ zu Beginn des Studiums nur sehr unvollständig und vage beantworten; etwa so: „Lineare Algebra ist die Theorie linearer Gleichungssysteme“. In diesem einleitenden Kapitel begegnen wir solchen Gleichungen, einem grundlegenden Konzept dieser Vorlesung, zum ersten Mal. Am Ende dieses Teils sollten Sie dann wissen, was lineare Gleichungssysteme sind und wie man diese systematisch lösen kann.

3.1 Lineare Gleichungen: Beispiele

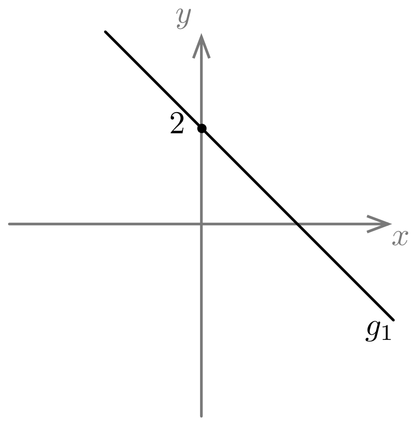
In der Mathematik treten Gleichung in verschiedender Form auf. So sind etwa **Identitäten** allgemeingültig:

- Für den Umfang U eines Kreises vom Radius R gilt immer $U = 2\pi R$.
- Für ein rechtwinkliges Dreieck mit Kathetenlängen a, b und Hypothenusenlänge c gilt immer der Satz von Pythagoras $a^2 + b^2 = c^2$.
- Für die Zahlen $0, 1, e, \pi$ und die imaginäre Einheit $\mathrm{i} = \sqrt{-1}$ gilt die Eulersche Identität $\mathrm{e}^{\pi \mathrm{i}} + 1 = 0$.

Dagegen gelten **Bestimmungsgleichungen** jeweils nur für gewisse Werte, eben die **Lösungen**, aus einer vorgegebenen Grundmenge:

- $x^2 = 2$ hat keine Lösung in der Grundmenge der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \ldots\}$, aber die Lösungen $+\sqrt{2}$ und $-\sqrt{2}$ in der Grundmenge $\mathbb{R}$ der reellen Zahlen.
- $x^2 + y^2 = 1$ gilt für alle Punkte (x, y) auf dem Kreis mit Radius 1 und Zentrum $(0, 0)$ in der xy-Ebene.

Zentraler Gegenstand der linearen Algebra sind Bestimmungsgleichungen von relativ einfacher Bauart, sogenannte **lineare Gleichungen**, wie etwa $x + y = 2$. Geometrisch ist die Menge der Lösungen dieser Gleichung die Gerade g_1 in der xy-Ebene.



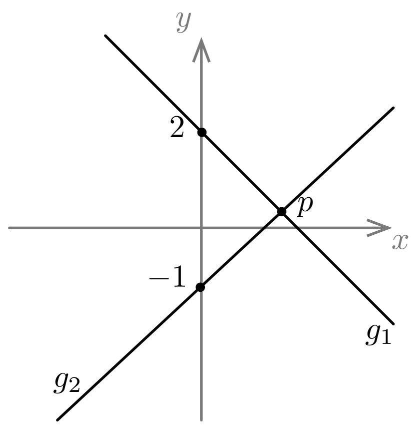
Solche Gleichungen treten in vielen alltäglichen Situationen auf. Zum Beispiel bei der Frage: In welchem Verhältnis muss man eine 20%-ige Lösung und eine 70%-ige Lösung mischen, um eine 30%-ige Lösung zu erhalten?

Ein **(lineares) Gleichungssystem** besteht aus mehreren linearen Gleichungen.

Das Gleichungssystem

$$x + y = 2 \tag{3.1}$$
$$x - y = 1 \tag{3.2}$$

beschreibt die Geraden $g_1$und g_2.



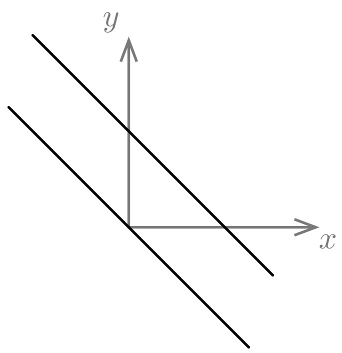
Die Lösungsmenge ist die Menge aller Punkte der xy-Ebene, die simultan beide Gleichungen erfüllen, also sowohl auf g_1 als auch auf g_2 liegen. Aus der Abbildung sieht man, dass die Lösungsmenge $\mathcal{L}$ nur aus dem Punkt p besteht: $\mathcal{L} = \{p\}$.

Um p zu bestimmen, kann man formal so vorgehen: Aus (3.2) folgt $y = x - 1$. Eingesetzt in (3.1) erhalten wir $x + (x - 1) = 2$, also $2x = 3$ oder $x = \frac{3}{2}$ und damit $y = x - 1 = \frac{3}{2} - 1 = \frac{1}{2}$, d.h. $p = (\frac{3}{2}, \frac{1}{2})$.

Zwei Geraden in der Ebene können auch parallel sein, z.B. sind

$$\begin{aligned} x + y &= 2 \\ x + y &= 0 \end{aligned}$$

parallel.



Es gibt also keine Schnittpunkte, was wiederum bedeutet, dass das Gleichungssystem *keine* Lösung hat: $\mathcal{L} = \emptyset$.

Für das System

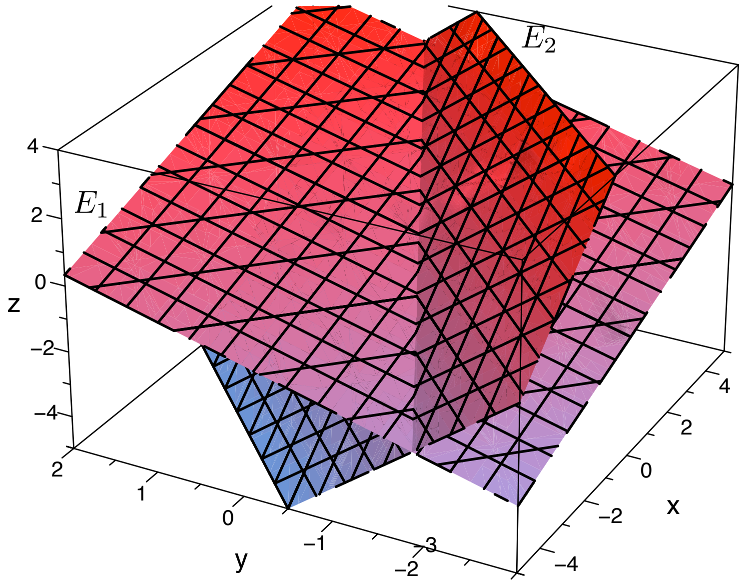
$$\begin{aligned} x + y &= 2 \\ 3x + 3y &= 6 \end{aligned}$$

fallen beide Geraden zusammen und alle Punkte der Geraden sind „Schnittpunkte“: das Gleichungssystem hat unendlich viele Lösungen.

Anstatt lineare Gleichungen mit *zwei* Unbekannten (oder Variablen) können wir natürlich auch solche mit *drei* Unbekannten x, y und z betrachten, etwa

$$x + y + z = -6 \tag{3.3}$$
$$x + 2y + 3z = -10. \tag{3.4}$$

Geometrisch sind das zwei Ebenen im xyz-Raum.



Aus der Abbildung sieht man, dass sich diese Ebenen in einer Geraden schneiden.

Wie kann man diese Schnittgerade, also die Lösungsmenge des Systems (3.3) und (3.4), formal bestimmen?

Aus (3.4) folgt $x = -2y-3z-10$, in (3.3) eingesetzt also $(-2y-3z-10)+y+z = -6$ oder vereinfacht $-y-2z = 4$, also $y = -2z-4$ und $x = -2(-2z-4)-3z-10 = z-2$. Dabei ist die Variable z beliebig wählbar. Wir erhalten eine **Parametrisierung** der Lösungsmenge (oder, geometrisch, der Schnittgeraden):

$$\mathfrak{L} = \{(t-2,\ -2t-4,\ t) \mid t \text{ eine beliebige reelle Zahl}\}.$$

Zwei Ebenen können auch parallel sein. Das Gleichungssystem hat dann keine Lösung, d.h. $\mathfrak{L} = \emptyset$, z.B.

$$\begin{aligned} x+y+z &= -6 \\ x+y+z &= 0. \end{aligned}$$

Oder die Ebenen können zusammenfallen und man hat unendlich viele Lösungen, z.B.

$$\begin{aligned} x+y+z &= -6 \\ -x-y-z &= 6. \end{aligned}$$

Damit man genau eine Lösung (also, geometrisch, genau einen Schnittpunkt) erhält, benötigt man drei Ebenen, z.B. hat

$$\begin{aligned} x + y + z &= -6 \\ x + 2y + 3z &= -10 \\ 2x + 3y + 6z &= -18 \end{aligned}$$

die Lösungsmenge $\mathfrak{L} = \{(-3, -2, -1)\}$.

In der Praxis hat man es oft mit linearen Gleichungssystemen mit sehr vielen (z.B. 200) Unbestimmten zu tun und es stellt sich die Frage, ob, und wenn ja, wie man solche Systeme lösen kann. Ziel der linearen Algebra ist es deshalb, die allgemeinen Strukturen, die solchen linearen Gleichungssystemen zugrunde liegen, zu finden und systematisch zu analysieren. Dazu machen wir einen (in der Mathematik typischen) allgemeinen Ansatz.

3.2 Lineare Gleichungssysteme: allgemein

In den vorhergehenden Beispielen konnten wir die Lösungen der Gleichungssysteme mit zwei Unbestimmten x, y bzw. drei Unbestimmten x, y, z geometrisch als Punkte in der „Ebene“ $\mathbb{R}^2$ bzw. im „Raum“ $\mathbb{R}^3$ auffassen. Für Gleichungssysteme mit n Unbestimmten $x_1, \ldots, x_n$ definieren wir den reellen **Standardraum** $\mathbb{R}^n$ als die Menge aller reellen n-Tupel,

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

Die (rellen) Lösungen für Gleichungssysteme in n Unbestimmten sind dann Elemente oder Punkte im Standardraum $\mathbb{R}^n$.

Definition 3.1 Ein **lineares Gleichungssystem** (LGS) mit m Gleichungen und n Unbestimmten $x_1, \ldots, x_n$ ist gegeben durch

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \tag{3.5}$$

Die a_{ij}, b_i für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ heißen **Koeffizienten** und sind gegebene reelle Zahlen. Die x_j für $j = 1, \ldots, n$ heißen **Unbestimmte** (oder **Unbekannte** oder **Variablen**) und sind gesucht.

Sind in (3.5) alle $b_i = 0$ $(i = 1, \ldots, m)$, so heißt das LGS **homogen**, und sonst **inhomogen**.

Definition 3.2 Die **Lösungsmenge** des reellen linearen Gleichungssystems (3.5) ist die Teilmenge $\mathfrak{L}$ von $\mathbb{R}^n$ bstehend aus allen n-Tupeln $(x_1, \ldots, x_n)$, die bei gegebenen Koeffizienten a_{ij}, b_i ($i = 1, \ldots, m$ und $j = 1, \ldots, n$) alle m Gleichungen in (3.5) simultan erfüllen.

Wie soll man nun vorgehen, um Lösungen des LGS (3.5) zu finden? Dazu definieren wir zunächst einfache Manipulationen des Systems:

Definition 3.3 Elementar-Operationen für das LGS (3.5) sind Umformungen der folgenden Art

(I) Vertauschen von zwei Gleichungen.

(II) Ersetzen einer Gleichung durch ihr λ-faches mit $\lambda \in \mathbb{R}$ und $\lambda \neq 0$.

(III) Ersetzen der i-ten Gleichung durch die Summe der i-ten und dem λ-fachen der j-ten Gleichung ($i \neq j$, $\lambda \in \mathbb{R}$).

Die Nützlichkeit dieser Umformungen liegt in folgender Tatsache

Satz 3.4 *Die Lösungsmenge $\mathfrak{L}$ des LGS (3.5) wird bei einer (und damit auch endlich vielen) Elementar-Operation nicht geändert.*

Wie immer in der Mathematik muss man eine solche Behauptung *beweisen*!

Beweis: Es reicht zu zeigen, dass eine einzige Zeilenumformung vom Typ (I), (II) oder (III) die Lösungsmenge nicht ändert, denn dann ändern auch wiederholte derartige Umformungen nichts.

Für Typ (I) ist dies klar, denn die Reihenfolge der Gleichungen ändert nichts an der Tatsachen, dass alle simultan erfüllt sein müssen.

Typ (II): Erfüllt $x = (x_1, \ldots, x_n)$ die Gleichung

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i,$$

so auch

$$\lambda a_{i1}x_1 + \cdots + \lambda a_{in}x_n = \lambda b_i.$$

Gilt umgekehrt für $x = (x_1, \ldots, x_n)$ die Gleichung

$$\lambda a_{i1}x_1 + \cdots + \lambda a_{in}x_n = \lambda b_i,$$

so kann man durch λ dividieren (hier braucht man $\lambda \neq 0$) und sieht, dass $x = (x_1, \ldots, x_n)$ auch die ursprüngliche Gleichung

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i$$

erfüllt.

Bei einer Umformung vom Typ (III) sind nur die Gleichungen i und j betroffen. Daher genügt es, zu zeigen, dass die beiden Systeme

$$\begin{array}{ccccccccc} a_{i1}x_1 & + & a_{i2}x_2 & + & \ldots & + & a_{in}x_n & = & b_i \\ a_{j1}x_1 & + & a_{j2}x_2 & + & \ldots & + & a_{jn}x_n & = & b_j \end{array} \qquad (*)$$

und

$$\begin{array}{ccccccccc} (a_{i1}+\lambda a_{j1})x_1 & + & (a_{i2}+\lambda a_{j2})x_2 & + & \ldots & + & (a_{in}+\lambda a_{jn})x_n & = & b_i+\lambda b_j \\ a_{j1}x_1 & + & a_{j2}x_2 & + & \ldots & + & a_{jn}x_n & = & b_j \end{array} \qquad (**)$$

die gleiche Lösungsmenge haben. Erfüllt aber $x = (x_1, \ldots, x_n)$ die Gleichungen $(*)$, so erfüllt x auch die zweite Gleichung von $(**)$. Durch Addition des λ-fachen der zweiten Gleichung von $(*)$ zur ersten Gleichung folgt, dass x auch die erste Gleichung von $(**)$ erfüllt. Umgekehrt folgt durch Subtraktion des λ-fachens der zweiten Gleichung aus $(**)$ von der ersten aus $(**)$ auch die erste Gleichung von $(*)$. Damit folgt, dass ein x, das $(**)$ erfüllt auch $(*)$ erfüllt. ■

Nach Satz 3.4 kann man (mindestens im Prinzip) ein „kompliziertes“ LGS in ein „einfacheres“ umformen.

3.3 Wie man ein LGS lösen kann: Der Gaußsche Algorithmus

Ein systematisches Verfahren (Algorithmus) zur Lösung eines allgemeinen linearen Gleichungssystems geht auf Carl Friedrich Gauß (1777-1855) zurück. Das Prinzip war aber chinesischen Mathematikern schon vor mehr als 2000 Jahren bekannt.

3.3.1 Zuerst ein Beispiel

Wir führen das Gaußsche Verfahren zunächst anhand von Beispielen vor.

Beispiel 3.5 Wir betrachten folgendes reelles LGS, das einen Parameter $a \in \mathbb{R}$ enthält.

$$\begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & 3x_3 & + & x_4 & = & 1 \\ 2x_1 & + & x_2 & + & x_3 & - & x_4 & = & 0 \\ & & 2x_2 & - & 13x_3 & + & x_4 & = & -1 \\ 2x_1 & - & x_2 & + & 14x_3 & - & 2x_4 & = & a \end{array}$$

1. Schritt: Wir addieren das (-2)-fache der ersten Gleichung zur zweiten und vierten

Gleichung und erhalten

$$\begin{array}{rcrcrcrcr}
x_1 & + & x_2 & - & 3x_3 & + & x_4 & = & 1 \\
 & - & x_2 & + & 7x_3 & - & 3x_4 & = & -2 \\
 & & 2x_2 & - & 13x_3 & + & x_4 & = & -1 \\
 & - & 3x_2 & + & 20x_3 & - & 4x_4 & = & a-2
\end{array}$$

2. Schritt: Wir addieren die oben angegebenen Vielfachen der zweiten Gleichung zu den anderen Gleichungen und multiplizieren die zweite Gleichung schließlich noch mit -1:

$$\begin{array}{rcrcrcr}
x_1 & & & + & 4x_3 & - & 2x_4 & = & -1 \\
 & & x_2 & - & 7x_3 & + & 3x_4 & = & 2 \\
 & & & & x_3 & - & 5x_4 & = & -5 \\
 & & & - & x_3 & + & 5x_4 & = & a+4
\end{array}$$

3. Schritt: Wir addieren die angegebenen Vielfachen der dritten Gleichung zu den anderen Gleichungen:

$$\begin{array}{rrrcrcr}
x_1 & & & + & 18x_4 & = & 19 \\
 & x_2 & & - & 32x_4 & = & -33 \\
 & & x_3 & - & 5x_4 & = & -5 \\
 & & & & 0x_4 & = & a-1\,.
\end{array}$$

Damit ist das Verfahren beendet. Nach Satz 3.4 hat das LGS, von dem wir ausgegangen sind, dieselbe Lösungsmenge wie das zuletzt erhaltene LGS. Aus der letzten Gleichung ergibt sich, dass das LGS für $a \neq 1$ *unlösbar* ist. Für $a = 1$ ist das LGS *lösbar*; die Lösungsmenge lässt sich aus

$$\begin{aligned}
x_1 &= 19 - 18x_4 \\
x_2 &= -33 + 32x_4 \\
x_3 &= -5 + 5x_4
\end{aligned}$$

unmittelbar ablesen. Man sieht, dass x_4 beliebig wählbar ist, während x_1, x_2, x_3 nach Wahl von x_4 eindeutig bestimmt sind. Schreiben wir noch t anstelle von x_4, so lässt sich jedes Element x der Lösungsmenge $\mathcal{L}$ folgendermaßen darstellen:

$$(x_1, x_2, x_3, x_4) = (19, -33, -5, 0) + t(-18, 32, 5, 1)$$

oder

$$x = u + t\,v, \qquad t \in \mathbb{R}.$$

Beobachtung: $u = (19, -33, -5, 0)$ ist eine Lösung des LGS und $v = (-18, 32, 5, 1)$ eine Lösung des zugehörigen homogenen LGS.

3.3.2 Die wesentlichen Daten: Matrizen

Die durchgeführten Elementaroperationen verändern lediglich die Koeffizienten des LGS. Wenn also die Zugehörigkeit der Koeffizienten zu den Variablen klar ist, kann man sich das Schreiben der Variablen $x_1, \ldots, x_n$ ersparen. Zu diesem Zweck führen wir die ökonomische Matrixschreibweise ein.

Definition 3.6 Eine **Matrix** mit m **Zeilen** und n **Spalten** ist ein rechteckiges Schema von m mal n Zahlen a_{ij} mit $i = 1, \ldots, m$ und $j = 1, \ldots, n$ der Form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Merkregel für die Reihenfolge der Indizes: **Z**eile **z**uerst, **Sp**alte **sp**äter.

Einem linearen Gleichungssystem kann man wie folgt eine Matrix zuordnen: Im „Schnittpunkt" der i-ten Zeile mit der j-ten Spalte hat die **Matrix des LGS** (3.5) den Eintrag a_{ij}.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{3.6}$$

Die **erweiterte Matrix des LGS** (3.5) enthält als letzte Spalte zusätzlich $b_1, \ldots, b_m$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}. \tag{3.7}$$

Beispiel 3.7 Wir betrachten das reelle LGS

$$\begin{array}{ccccccccccccc} & & 2x_2 & + & 4x_3 & - & 2x_4 & + & x_5 & + & 7x_6 & = & -1 \\ x_1 & & & + & x_3 & + & 3x_4 & & & - & x_6 & = & 1 \\ x_1 & + & x_2 & + & 3x_3 & + & 2x_4 & & & + & x_6 & = & 1 \\ & & x_2 & + & 2x_3 & - & x_4 & - & x_5 & - & x_6 & = & 1 \\ 3x_1 & + & 2x_2 & + & 7x_3 & + & 7x_4 & - & x_5 & - & 2x_6 & = & a \end{array}$$

mit der erweiterten Matrix

$$\begin{pmatrix} 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 1 & 1 & 3 & 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 3 & 2 & 7 & 7 & -1 & -2 & a \end{pmatrix}$$

1. Schritt: Wir addieren das (-1)-fache der zweiten Gleichung (bzw. Matrix-Zeile) zur dritten und das (-3)-fache der zweiten Gleichung (bzw. Matrix-Zeile) zur letzten. Schließlich vertauschen wir noch die ersten beiden Gleichungen, damit die Eins links oben steht, und erhalten folgende Matrix:

$$\begin{pmatrix} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 0 & 1 & 2 & -1 & 0 & 2 & 0 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 0 & 2 & 4 & -2 & -1 & 1 & a-3 \end{pmatrix}$$

2. Schritt: Wir addieren die angegebenen Vielfachen der dritten Gleichung zur zweiten, vierten und fünften Gleichung. Dann vertauschen wir noch die zweite und dritte Gleichung, damit die Eins links oben im „Kästchen“ steht, und erhalten

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

3. Schritt: Wegen den Nullen in der dritten und vierten Spalte können wir die dritte und vierte Variable überspringen. Wir addieren die dritte Gleichung zur vierten und fünften Gleichung und bekommen

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

Das Verfahren ist damit beendet. Das zugehörige LGS

$$\begin{array}{rcrcrcrcrcrcr} x_1 & & & + & x_3 & + & 3x_4 & & & - & x_6 & = & 1 \\ & & x_2 & + & 2x_3 & - & x_4 & & & + & 2x_6 & = & 0 \\ & & & & & & & & x_5 & + & 3x_6 & = & -1 \\ & & & & & & & & & & 0x_6 & = & 0 \\ & & & & & & & & & & 0x_6 & = & a-4 \end{array}$$

hat dieselbe Lösungsmenge wie das Ausgangssystem und ist für $a \neq 4$ *unlösbar*, für $a = 4$ *lösbar*. Die Lösungsmenge lässt sich (für $a = 4$) aus

$$\begin{array}{rcrcrcrcr} x_1 & = & 1 & - & x_3 & - & 3x_4 & + & x_6 \\ x_2 & = & & - & 2x_3 & + & x_4 & - & 2x_6 \\ x_5 & = & -1 & & & & & - & 3x_6 \end{array}$$

ablesen: Setzen wir $x_3 = t_1$, $x_4 = t_2$, $x_6 = t_3$, so bekommt man

$$\begin{array}{rcrcrcrcr} x_1 & = & 1 & - & t_1 & - & 3t_2 & + & t_3 \\ x_2 & = & & - & 2t_1 & + & t_2 & - & 2t_3 \\ x_3 & = & & & t_1 & & & & \\ x_4 & = & & & & & t_2 & & \\ x_5 & = & -1 & & & & & - & 3t_3 \\ x_6 & = & & & & & & & t_3 \end{array},$$

und die Lösungsmenge besteht aus allen Elementen $x = (x_1, \dots, x_6) \in \mathbb{R}^6$, die sich darstellen lassen als

$$x = u + t_1\, v_1 + t_2\, v_2 + t_3\, v_3 \quad \text{mit} \quad t_1, t_2, t_3 \in \mathbb{R}$$

mit

$$u = (1, 0, 0, 0, -1, 0), \quad v_1 = (-1, -2, 1, 0, 0, 0),$$

$$v_2 = (-3, 1, 0, 1, 0, 0), \quad v_3 = (1, -2, 0, 0, -3, 1).$$

3.3.3 Das allgemeine Vorgehen

Gegeben sei das relle LGS (3.5) mit $m, n \in \mathbb{N}$ und reellen Koeffizienten a_{ik}, b_i und der erweiterten Matrix (3.7).

Ziel ist es, die erweiterte Matrix $(A \mid b)$ durch elementare Zeilenoperationen möglichst zu vereinfachen, d.h. möglichst viele Einträge zu Null (oder Eins) zu machen.

Der Fall, dass alle a_{ik} Null sind, ist uninteressant: Dann ist nämlich entweder (3.5) unlösbar (falls es ein $b_i \neq 0$ gibt), oder die Lösungsmenge ist $\mathbb{R}^n$ (falls alle $b_i = 0$ sind). Wir werden also im Folgenden annehmen, dass es mindestens ein $a_{ik} \neq 0$ gibt.

1. Schritt: Ist ein Element a_{i1} in der ersten Spalte von (3.5) von Null verschieden, so lässt sich (nötigenfalls durch eine Vertauschung (I)) erreichen, dass $a_{11} \neq 0$. Weiter kann man durch Elementaroperationen (II) und (III) erreichen, dass $a_{11} = 1$ und $a_{i1} = 0$: Man multipliziert dazu die 1. Zeile mit $\frac{1}{a_{11}}$ und addiert zur i-ten Zeile das $-a_{i1}$-fache der ersten Zeile ($i = 2, \dots, m$). Sind dagegen alle Elemente der ersten Spalte Null und kommt in der k-ten Spalte zum ersten Mal ein von Null verschiedenes

Element vor, so kann man entsprechend $a_{1k} = 1,\ a_{ik} = 0 \quad (i = 2, \ldots, m)$ erreichen. (3.5) geht somit im ersten Schritt über in

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & a'_{1,k+1} & \cdots & a'_{1n} & b'_1 \\ \vdots & & \vdots & 0 & a'_{2,k+1} & \cdots & a'_{2n} & b'_2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & a'_{m,k+1} & \cdots & a'_{mn} & b'_m \end{pmatrix}. \tag{3.8}$$

2. Schritt: Ist mindestens eins der a'_{ij} mit $i \geq 2$ und $j \geq k+1$ von Null verschieden, so verfährt man wie beim ersten Schritt und erhält eine erweiterte Matrix der Form

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & * & * & \cdots & * & * \\ \vdots & & \vdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & * & \cdots & * & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * & * \end{pmatrix}. \tag{3.9}$$

Gibt es noch von Null verschiedene Koeffizienten in den Zeilen $3, 4, \ldots$ (mit Ausnahme der Elemente in der letzten Spalte), so folgt in entsprechender Weise ein **3. Schritt** usw.

Das Verfahren ist beendet, wenn entweder in den letzten Zeilen nur noch Nullen stehen (bis auf die Elemente in der letzten Spalte) oder wenn man mit der zuletzt erhaltenen Eins die letzte Spalte oder Zeile der einfachen (d.h. nicht erweiterten) Matrix erreicht hat. Die Endgestalt der Matrix hat schließlich folgende **Zeilen-Stufen-Form**:

$$\left(\begin{array}{ccccccccccccc|c} 0 & \cdots & 0 & 1 & * & \cdots & * & * & *\cdots* & * & * & \cdots & * & c_1 \\ \vdots & & \vdots & 0 & 0 & \cdots & 0 & 1 & *\cdots* & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \ddots & * & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & 1 & * & \cdots & * & c_r \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & 0 & 0 & \cdots & 0 & c_{r+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & c_m \end{array}\right). \tag{3.10}$$

Aus (3.10) liest man ab:

Folgerung 3.8 *Das zu (3.10) gehörige LGS und damit nach Satz 3.4 auch das LGS (3.5) ist genau dann lösbar, wenn gilt* $c_{r+1} = c_{r+2} = \ldots = c_m = 0$.

Durch weitere Zeilenumformungen kann man erreichen, dass oberhalb der Einsen überall Nullen stehen. So erhält man schließlich die **Gaußsche Normalform** des LGS (3.5):

$$\left(\begin{array}{ccccccccccccc|c}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & *\cdots* & 0 & * & \cdots & * & d_1 \\
\vdots & & \vdots & 0 & 0 & \cdots & 0 & 1 & *\cdots* & \vdots & \vdots & & \vdots & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \ddots & 0 & \vdots & & \vdots & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & 1 & * & \cdots & * & d_r \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & 0 & 0 & \cdots & 0 & d_{r+1} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_m
\end{array}\right). \tag{3.11}$$

Parametrisierung der Lösungsmenge: Falls das zu (3.11) bzw. (3.5) gehörige LGS lösbar ist (also $d_{r+1} = \ldots = d_m = 0$), so lassen sich alle Lösungen von (3.5) an (3.11) ablesen.

Um die Darstellung zu vereinfachen, nehmen wir an, dass die Gaußsche Normalform folgende Gestalt hat

$$\left(\begin{array}{ccccc|ccc|c}
1 & 0 & 0 & \cdots & 0 & a''_{1,r+1} & \cdots & a''_{1,n} & d_1 \\
0 & 1 & 0 & & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 1 & & \vdots & \vdots & & \vdots & \vdots \\
\vdots & & & \ddots & 0 & \vdots & & \vdots & \vdots \\
0 & & & 0 & 1 & a''_{r,r+1} & & a''_{r,n} & d_r \\
\vdots & & & \vdots & 0 & 0 & \cdots & 0 & 0 \\
\vdots & & & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & 0
\end{array}\right). \tag{3.12}$$

Durch eine Umordnung der Spalten von A, d.h. eine andere Numerierung der Unbekannten des LGS, kann man das stets erreichen.

Man wählt dann (wie im Beispiel) $t_1, \ldots, t_{n-r} \in \mathbb{R}$ als **Parameter** und setzt

$$x_{r+1} := t_1,\ x_{r+2} := t_2,\ \ldots,\ x_n := t_{n-r}.$$

Aus (3.12) erhält man dann für die restlichen r Unbekannten:

$$\begin{array}{lcllllllll} x_1 & = & d_1 & - & t_1\, a''_{1,r+1} & - & \cdots & - & t_{n-r}\, a''_{1,n} \\ & \vdots & & & & & & & \\ x_r & = & d_r & - & t_1\, a''_{r,r+1} & - & \cdots & - & t_{n-r}\, a''_{r,n} \\ & & & & & & & & \\ x_{r+1} & = & & & t_1 & & & & \\ & \vdots & & & & & \ddots & & \\ x_n & = & & & & & & & t_{n-r} \end{array} \tag{3.13}$$

Durchlaufen $t_1, \ldots, t_{n-r}$ jeweils alle reellen Zahlen, so erhält man mit (3.13) alle Lösungen von (3.5). Für $t_1 = \ldots = t_{n-r} = 0$ ergibt sich speziell die Lösung $x = (d_1, \ldots, d_r, 0, \ldots, 0)$.

Folgerung 3.9 *Ein homogenes LGS mit mehr Unbekannten als Gleichungen ($n > m$) ist immer nichttrivial lösbar (d.h. hat nicht nur die Null-Lösung).*

3.4 Einige weiterführende Fragen

- Wir haben in diesem Abschnitt bereits die Begriffe Menge, Teilmenge, Lösungsmenge verwendet und sind „intuitiv" damit umgegangen. Wie lassen sich diese Begriffe präzisieren, welche Schreibweisen gibt es dafür und welche Operationen kann man mit Mengen ausführen?

- Wie kann man das logische Schließen (etwa im Beweis von Satz 3.4) systematisieren und übersichtlich darstellen? Was für logische Operationen gibt es? Was für Beweis-Methoden gibt es?

- Gibt es noch weitere „Zahlbereiche", mit denen man formal wie mit den reellen oder den rationalen Zahlen rechnen kann?

- Kann man herausfinden, ob ein gegebenes lineares Gleichungssystem eine Lösung hat, ohne den Gaußschen Algorithmus durchzuführen? Kann man a priori etwas über die mögliche Anzahl der Lösungen sagen? (Gibt es z.B. ein LGS, dessen Lösungsmenge genau zwei Elemente enthält?)

- Was sind die algemeinen Eigenschaften (Struktur) der Lösungsmenge eines LGS?

Teil II

Grundlegende Begriffe

In diesem Kapitel führen wir einige Begiffe und Bezeichnungen ein, die nicht nur für die Lineare Algebra, sondern für die gesamte Mathematik grundlegend sind: *Logische Begriffe* sind unentbehrlich, um mathematische Aussagen präzise zu fassen und neue deduktiv herzuleiten. Die Objekte der Mathematik lassen sich zweckmäßig als *Mengen* beschreiben. Mittels *Abbildungen* kann man Beziehungen zwischen einzelnen Mengen beschreiben.

Unser Ziel ist eine kurze Vorstellung der Konzepte und die Festlegung von Sprechweise und Notation anhand von Beispielen. Wir verzichten auf eine systematische Einführung in die Gebiete „Logik“ und „Mengenlehre“ und verweisen z.B. auf die Bücher von Tarski [20] und Halmos [13].

4 Logik und Mengenlehre: ein Steilkurs

4.1 Logik

In der **Aussagenlogik** werden aus „elementaren“ Aussagen und logischen Verknüpfungen neue Aussagen zusammengesetzt.

Beispiel 4.1 Zwei Beispiele für Aussagen sind: Es ist Nacht und 3 ist eine natürliche Zahl.

Logische Verknüpfungen sind

Symbol	Name	Sprechweise
$\wedge$	Konjunktion	„und“
$\vee$	Disjunktion	„oder“
$\neg$	Negation	„nicht“
$\Rightarrow$	Implikation	„daraus folgt“
$\Leftrightarrow$	Äquivalenz	„ist äquivalent zu“

Durch Negation einer „wahren“ Aussage erhält man eine „falsche“ und durch Negation einer falschen Aussage erhält man eine wahre.

Beispiel 4.2 Bezeichnet A die Aussage -1 ist eine natürliche Zahl, so ist A falsch, ihre Negation $\neg A$ (gesprochen „nicht A“) ist eine wahre Aussage. $\neg A$ lässt sich umgangssprachlich formulieren als -1 ist keine natürliche Zahl.

Beispiel 4.3 Im Satz In der Nacht sind alle Katzen grau lassen sich zwei Teil-Aussagen erkennen, nämlich $N :=$ Es ist Nacht und $K :=$ Alle Katzen sind grau. (Das Zeichen $:=$ bedeutet, dass der links stehende Ausdruck durch den rechts stehenden Ausdruck definiert wird.)

Diese beiden Aussagen sind durch eine Implikation verknüpft, was man deutlicher sieht, wenn man den Satz umformuliert in Wenn es Nacht ist, dann sind alle Katzen grau. Mit Hilfe der logischen Verknüpfung $\Rightarrow$ (gesprochen „daraus folgt“ oder „impliziert“) lässt sich der Satz also folgendermaßen schreiben:

$$N \Rightarrow K.$$

Wir haben hier aus den beiden elementaren Aussagen N und K mit Hilfe der logischen Verknüpfung $\Rightarrow$ eine neue, zusammengesetzte Aussage erzeugt.

Weitaus weniger gebräuchlich als die fünf oben genannten Verknüpfungen ist das Zeichen $\veebar$ für „entweder-oder“.

Wenn man mehr als zwei elementare Aussagen zu zusammengesetzten Aussagen verknüpft, muss man auf korrekte Klammerung der einzelnen Aussagen achten, wie man an folgendem Beispiel beobachten kann.

Beispiel 4.4 Zu den oben eingeführten elementaren Aussagen N und K nehmen wir noch eine weitere Aussage hinzu: $R :=$ Es regnet. Mit diesen Aussagen bilden wir die beiden Aussagen

$$N \wedge (R \Rightarrow K) \quad \text{und} \quad (N \wedge R) \Rightarrow K. \tag{4.1}$$

Die beiden Aussagen sind sehr verschieden: die erste kann man lesen als Es ist Nacht und wenn es regnet, sind alle Katzen grau. Die zweite Aussage lautet etwa In regnerischen Nächten sind alle Katzen grau. Dass die beiden Aussagen wirklich verschieden sind, werden wir in Beispiel 4.6 noch genauer verstehen.

Der Wahrheitswert von zusammengesetzten Aussagen wird aus den Wahrheitswerten der einzelnen elementaren Aussagen abgeleitet. Das geschieht mittels **Wahrheitstafeln**, die angeben, in welchen Fällen eine zusammengesetzte Aussage den Wahrheitswert „wahr“ (w) oder „falsch“ (f) annimmt. Die Wahrheitstafeln für die einzelnen Verknüpfungen lauten wie folgt:

D	E	$D \wedge E$
w	w	w
w	f	f
f	w	f
f	f	f

D	E	$D \vee E$
w	w	w
w	f	w
f	w	w
f	f	f

E	$\neg E$
w	f
f	w

D	E	$D \Rightarrow E$
w	w	w
w	f	f
f	w	w
f	f	w

D	E	$D \Leftrightarrow E$
w	w	w
w	f	f
f	w	f
f	f	w

D	E	$D \veebar E$
w	w	f
w	f	w
f	w	w
f	f	f

Die erste Wahrheitstafel gibt beispielsweise an, dass die Aussage $D \wedge E$ nur dann wahr ist, wenn sowohl D als auch E wahr sind. Die Disjunktion von D und E ist hingegen nur dann falsch, wenn sowohl D als auch E falsch sind. Damit $D \vee E$ wahr ist, muss mindestens eine der beiden Aussagen wahr sein. Im Gegensatz dazu ist die Aussage $D \veebar E$ nur wahr, wenn genau eine von beiden wahr ist, da sich die Aussagen gegenseitig ausschließen.

Bemerkung 4.5 Beachten Sie, dass die Aussage $D \Rightarrow E$ wahr ist, auch wenn D falsch ist und zwar unabhängig vom Wahrheitswert von E. Umgangssprachlich formuliert: „Aus einer falschen Aussage kann man alles folgern".

Beispiel 4.6

1. Ist A die Aussage -1 ist eine natürliche Zahl und B die Aussage 3 ist eine natürliche Zahl, dann ist die Aussage $A \Rightarrow B$ (Wenn -1 eine natürliche Zahl ist, dann ist 3 eine natürliche Zahl) wahr, denn eine Implikation $D \Rightarrow E$ hat den Wahrheitswert w, falls D den Wahrheitswert f hat. Die Aussage $A \wedge B$ (also: -1 und 3 sind beides natürliche Zahlen) ist falsch (da mindestens eine der beiden Aussagen falsch ist, in diesem Fall A) und die Aussage $A \vee B$ ist wahr (da mindestens eine der beiden Aussagen wahr ist, in diesem Fall B.)

2. Wenn die Aussagen N und R im obigen Beispiel falsch sind (Es ist nicht Nacht bzw. Es regnet nicht), dann ist die Aussage $N \wedge (R \Rightarrow K)$ falsch (da eine Konjunktion $D \wedge E$ den Wahrheitswert f hat, wenn eine der beiden Aussagen den Wahrheitswert f hat). Die Aussage $(N \wedge R) \Rightarrow K$ ist in diesem Fall jedoch wahr (da eine Implikation $D \Rightarrow E$ den Wahrheitswert w hat, falls D den Wahrheitswert f hat). Die Aussagen in (4.1) sind also tatsächlich verschieden.

Eine Verallgemeinerung der Aussagenlogik ist die **Prädikatenlogik**.

Hier betrachtet man allgemeine **Aussageformen**, die nach dem Einsetzen eines Elementes aus einer gegebenen Menge zu Aussagen im Sinne der Aussagenlogik werden.

Beispiel 4.7

1. $A_1(x) := x$ ist eine natürliche Zahl ist eine Aussageform auf der Menge $\mathbb{Z}$ der ganzen Zahlen. Die Größe x bezeichnet man hier als **Variable** der Aussageform A_1. Setzt man eine ganze Zahl für x ein, so erhält man eine Aussage, z.B. ist $A_1(3)$ die Aussage 3 ist eine natürliche Zahl und $A_1(-1)$ die Aussage -1 ist eine natürliche Zahl.

2. $A_2(x) := (x + x = 2x)$ ist eine Aussageform auf der Menge der ganzen Zahlen $\mathbb{Z}$, die beim Einsetzen eines beliebigen Elementes von $\mathbb{Z}$ für x immer eine wahre Aussage ergibt. Eine solche Aussageform nennt man **allgemeingültig**.

3. $A_3(x) := (3 \leq x) \wedge (x \leq 5)$ ist eine Aussageform auf $\mathbb{Z}$, die zwar nicht allgemeingültig, aber immerhin **erfüllbar** ist, d.h. es gibt mindestens ein Element der Grundmenge, für das die Aussage wahr ist. In diesem Beispiel etwa ist $A_3(4)$ eine wahre und $A_3(1)$ eine falsche Aussage.

4. $G(n, k) :=$ In der Nacht n ist Katze k grau ist eine (zweistellige) Aussageform auf der Grundmenge, die aus allen Paaren (n, k) aus Nächten n und Katzen k besteht.

5. $T(x, y) := x$ ist ein Teiler von y ist eine Aussageform auf der Menge aller Paare von natürlichen Zahlen. Z.B. ist $T(4, 12)$ eine wahre Aussage und $T(1, y)$ eine allgemeingültige Aussageform auf der Menge der natürlichen Zahlen.

6. Sind die Koeffizienten a_{ij} und b_i mit $i = 1, \ldots, m$ und $j = 1, \ldots, n$ fest vorgegeben, so ist $A(x) := x$ ist Lösung des LGS (3.5) eine Aussageform auf der Menge $\mathbb{R}^n$ aller reellen n-Tupel $x = (x_1, \ldots, x_n)$. In dieser Sprechweise ist das LGS genau dann lösbar, wenn $A(x)$ eine erfüllbare Aussageform ist.

Die Variablen in einer Aussageform werden oft **quantifiziert** mit Hilfe des **Existenzquantors** $\exists$ (gesprochen „Es gibt ein...“) und des **Allquantors** $\forall$ (gesprochen „Für alle...“).

Beispiel 4.8

1. $\exists x \in \mathbb{Z} : A_3(x)$ liest sich als Es gibt eine ganze Zahl x, so dass gilt: $3 \leq x$ und $x \leq 5$ und ist eine wahre Aussage, da beispielsweise $A_3(4)$ wahr ist. Die Aussage $\forall x \in \mathbb{Z} : A_3(x)$ liest sich als Für alle ganzen Zahlen x gilt: $3 \leq x$ und $x \leq 5$ und ist eine falsche Aussage.

2. $\forall x \in \mathbb{N}\ \exists y \in \mathbb{N} : T(x, y)$ ist eine wahre Aussage, da es zu jeder natürlichen Zahl x mindestens eine Zahl y gibt, deren Teiler sie ist; man setze z.B.

$y := 2x$. Die Aussage $\exists\, y \in \mathbb{N}\ \forall\, x \in \mathbb{N} : T(x, y)$ ist eine falsche Aussage; in Umgangssprache formuliert lautet sie Es gibt eine natürliche Zahl y, die von allen natürlichen Zahlen geteilt wird.

Bemerkung 4.9 Anhand des letzten Beispiel kann man sehen, dass die Reihenfolge der einzelnen Quantifizierungen der Variablen entscheidend ist: $\forall\, x \in \mathbb{N}\ \exists\, y \in \mathbb{N} : T(x, y)$ ist eine ganz andere Aussage als $\exists\, y \in \mathbb{N}\ \forall\, x \in \mathbb{N} : T(x, y)$.

Ein weiterer Quantor ist $\exists_1$, der „Es gibt genau ein …“ bedeutet. Für eine Aussageform $E(x)$ auf der Grundmenge M ist $\exists_1 x \in M : E(x)$ genau dann eine wahre Aussage, wenn es genau ein Element x in M gibt, für das die Aussage $E(x)$ wahr ist.

4.2 Mengen

Bei der Untersuchung von mathematischen Strukturen werden aus gegebenen Konzepten neue aufgebaut. Verfolgt man diesen Prozess zurück, so stößt man zwangsläufig auf Grundbegriffe, die mathematisch nicht weiter erklärt werden können. Man kann solche Begriffe nur dadurch festlegen, dass man den Umgang mit ihnen durch Gesetze (sogenannte **Axiome**) regelt.

Grundlegend für die gesamte Mathematik ist der Begriff der **Menge**. Der Begründer der Mengenlehre, Georg Cantor (1845–1918), hatte noch definiert:

Unter einer „Menge“ verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die „Elemente“ von M genannt werden) zu einem Ganzen.

In der modernen Mathematik verzichtet man auf eine Definition des Begriffs „Menge“ und verwendet ihn als Grundbegriff. Um Widersprüche zu vermeiden wird gefordert, dass eine Menge sich nicht selbst als Element enthalten darf. Mehr über den axiomatischen Aufbau der Mengenlehre und dabei mögliche Widersprüche findet man in dem Buch von Halmos [13].

Ist ein „Objekt“ a in einer Menge M enthalten, schreiben wir $a \in M$ (lies „a Element M“), andernfalls $a \notin M$ (lies „a nicht Element M“).

Mengen kann man beschreiben durch Auflisten ihrer Elemente, z.B. $M = \{1, 2, 3, 4, 5\}$ oder durch Auswahl bestimmter Elemente einer Grundmenge G mit Hilfe einer Aussageform $A(x)$ auf G, z.B. $G = \mathbb{N}$ und $M = \{n \in \mathbb{N} \mid 1 \leq n \leq 5\}$. Die allgemeine Schreibweise ist $M = \{x \in G \,|\, A(x)\}$ mit einer Grundmenge G und einer Aussageform $A(x)$ auf G.

Beispiel 4.10

1. die **leere Menge** $\emptyset = \{\}$, die keine Elemente enthält.
2. die **natürlichen Zahlen** $\mathbb{N} = \{1, 2, 3, \ldots\}$. Nehmen wir die Null hinzu, so schreiben wir $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
3. die **ganzen Zahlen** $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
4. die **rationalen Zahlen** $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$.
5. die **reellen Zahlen** $\mathbb{R}$, deren Konstruktion in der Vorlesung „Analysis I" detailiert behandelt wird.
6. die **komplexen Zahlen** $\mathbb{C} = \{x + \mathrm{i}y \mid x, y \in \mathbb{R}, \mathrm{i} = \sqrt{-1}\}$. Sie werden später in Abschnitt 5.3.1 näher vorgestellt.
7. Die Lösungsmenge $\mathcal{L}$ des LGS (3.5) bei vorgegebenen (reellen) Koeffizienten a_{ij} und b_i lässt sich mit Hilfe der Aussageform $A(x) = x$ ist Lösung des LGS (3.5) ausdrücken als $\mathcal{L} = \{x \in \mathbb{R}^n \mid A(x)\}$.

A heißt **Teilmenge** von B, wenn jedes Element von A auch in B liegt, wenn also aus $x \in A$ folgt $x \in B$. Die Menge B heißt dann **Obermenge** von A. Wir schreiben $A \subset B$ oder $B \supset A$. Dabei kann A echte oder unechte Teilmenge von B sein, je nachdem, ob $A \neq B$ oder $A = B$ ist. Man nennt $\subset$ das **Inklusionszeichen**.

Zwei Mengen M_1 und M_2 sind **gleich**, wenn sie die gleichen Elemente besitzen, d.h. wenn für jedes x gilt:

$$\text{Aus } x \in M_1 \text{ folgt } x \in M_2 \text{ und aus } x \in M_2 \text{ folgt } x \in M_1 \,.$$

Es gilt also $M_1 = M_2$ genau dann, wenn $M_1 \subset M_2$ und $M_2 \subset M_1$.

Die Menge aller Teilmengen einer Menge M heißt **Potenzmenge**

$$\mathcal{P}(M) := \{A \mid A \subset M\}.$$

Der Name erklärt sich aus folgendem Beispiel:

Ist M eine endliche Menge mit k Elementen, so ist $\mathcal{P}(M)$ eine Menge mit 2^k Elementen. Z.B. ist die Potenzmenge der Menge $M = \{1, 2, 3\}$ die Menge

$$\mathcal{P}(M) = \{\ \emptyset,\ \{1\},\ \{2\},\ \{3\},\ \{1,2\},\ \{1,3\},\ \{2,3\},\ \{1,2,3\}\ \}$$

mit $2^3 = 8$ Elementen.

Der **Durchschnitt** der Mengen A, B ist die Menge

$$A \cap B := \{x \mid x \in A \wedge x \in B\}.$$

Ein Element liegt also genau dann im Durchschnitt von A und B, wenn es sowohl in A als auch in B liegt. Ist der Durchschnitt $A \cap B$ leer, so heißen A und B **disjunkt**.

Die **Vereinigung** der Mengen A, B ist die Menge

$$A \cup B := \{x \mid x \in A \vee x \in B\}.$$

Ein Element liegt also in der Vereinigungsmenge $A \cup B$, wenn es wenigstens in einer der beiden Mengen liegt.

Bemerkung 4.11 Eigenschaften von $\cup$, $\cap$:

- $A \cap B = B \cap A$ und $A \cup B = B \cup A$ (Kommutativgesetze)
- $(A \cap B) \cap C = A \cap (B \cap C)$ und $(A \cup B) \cup C = A \cup (B \cup C)$ (Assoziativgesetze)
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributivgesetze)

Unter dem (cartesischen) **Produkt** der Mengen A, B versteht man die Menge

$$A \times B := \{(x, y) \mid x \in A \wedge y \in B\}.$$

Dabei ist (x, y) ein geordnetes Paar, und $(x, y) = (x', y')$ gilt genau dann, wenn $x = x'$ und $y = y'$. Ein Beispiel ist $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

Die **Differenz** der Mengen A, B ist die Menge

$$A \backslash B := \{x \mid x \in A \wedge x \notin B\}.$$

Ist insbesondere A die Grundmenge G, so nennt man $B^c := G \backslash B$ das **Komplement**:

$$B^c = \{x \in G \mid x \notin B\}.$$

Bemerkung 4.12 Es gelten die Formeln

$$A \backslash A = \emptyset, \quad A \cap A^c = \emptyset, \quad A \cup A^c = G, \quad (A^c)^c = A,$$

sowie die **Regeln von de Morgan**

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c.$$

Durchschnitt und Vereinigung lassen sich auch von mehr als zwei Mengen bilden, indem man die obigen Definitionen sinngemäß überträgt. Sei $\mathcal{M}$ eine Menge von Mengen, z.B. $\mathcal{M} \subset \mathcal{P}(A)$ für eine Menge A.

Der **Durchschnitt** aller Mengen B des Mengensystems $\mathcal{M}$ ($\mathcal{M} \neq \emptyset$) ist die Menge

$$\bigcap_{B \in \mathcal{M}} B := \{x \mid \text{für alle } B \in \mathcal{M} \text{ gilt } x \in B\} = \{x \mid \forall B \in \mathcal{M} : x \in B\}.$$

Sie besteht aus denjenigen Elementen x, die zu *allen* Mengen $B \in \mathcal{M}$ gehören.

Die **Vereinigung** aller Mengen $B \in \mathcal{M}$ ist die Menge

$$\bigcup_{B \in \mathcal{M}} B := \{x \mid \text{es gibt ein } B \in \mathcal{M} \text{ mit } x \in B\} = \{x \mid \exists B \in \mathcal{M} : x \in B\}.$$

Sie besteht aus denjenigen Elementen x, die zu *mindestens einer* Menge $B \in \mathcal{M}$ gehören.

4.3 Beweismethoden

Mathematische (Lehr-)Sätze sind wenn-dann-Aussagen. Aus einer gegebenen Aussage V (der **Voraussetzung**) wird mittels logischer Gesetze eine andere Aussage B (die **Behauptung**) abgeleitet; die Darstellung dieser Ableitung ist der Beweis. Formal hat also jede mathematische Aussage die Gestalt $V \Rightarrow B$ und der Zweck des Beweises ist, diese Implikation mit den Mitteln der Logik nachzuweisen. Dafür gibt es verschiedene Methoden; die gebräuchlichsten sind

- **direkter Beweis**: Aus der Voraussetzung wird die Behauptung „direkt“ bewiesen. Ein Beispiel ist Satz 3.4.

- **indirekter Beweis**: Hier benutzt man die Tatsache, dass die Implikation $V \Rightarrow B$ gleichwertig ist mit der Implikation $\neg B \Rightarrow \neg V$. Anstatt die Aussage „Aus V folgt B“ nachzuweisen, kann man genauso gut die Aussage „Aus nicht B folgt nicht V“ zeigen (und ist dann fertig!). Praktisch formuliert man einen indirekten Beweis meistens als **Widerspruchsbeweis**: „Angenommen, die Behauptung B ist falsch, dann (so muss man zeigen) ist auch die Voraussetzung V falsch“.

- **Ringschlüsse**: Mathematische Sätze sind oft Äquivalenzaussagen: verschiedene Behauptungen sind gleichwertig; wenn eine gilt, so gelten auch alle anderen. Hier kann man so vorgehen: Wenn etwa $A \Leftrightarrow B \Leftrightarrow C$ zu zeigen ist, genügt es, $A \Rightarrow B$, $B \Rightarrow C$ und $C \Rightarrow A$ nachzuweisen.

- **vollständige Induktion**: Hier muss man Aussagen A_n für *für alle natürlichen Zahlen* $n \in \mathbb{N}$ beweisen. Dazu geht man so vor:

  INDUKTIONS-VERANKERUNG: Man zeigt, dass etwa A_1 gilt.

  INDUKTIONS-SCHRITT: Sei dann $k \geq 1$ beliebig. Man nimmt an, dass $A_1, A_2, \ldots, A_k$ gelten. Unter dieser Voraussetzung zeigt man dann, dass auch A_{k+1} gilt.

4.4 Abbildungen

Definition 4.13 Gegeben seien zwei Mengen A und B. Eine **Abbildung** von A in B ordnet jedem Element von A genau ein Element von B zu. Wir schreiben

$$f : A \to B, \quad a \mapsto f(a)$$

A heißt **Definitionsmenge** und B **Zielmenge** von f. Die Menge $f(A) := \{f(a) \mid a \in A\} \subset B$ heißt **Bildmenge** von f. Die Menge $\{(a, f(a)) \mid a \in A\} \subset A \times B$ heißt **Graph** der Abbildung f.

Ist die Zielmenge $\mathbb{R}$ oder $\mathbb{C}$, so sagt man statt Abbildung auch **Funktion** .

Eine Abbildung $f : A \to A$ einer Menge A in sich heißt **Selbstabbildung** der Menge A. Insbesondere ist die **identische Abbildung** von A

$$\mathrm{id}_A : A \to A, \quad x \mapsto x$$

eine Selbstabbildung.

Beispiel 4.14

1. $f : \mathbb{N} \to \mathbb{N}, x \mapsto x^2$.

2. $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, x \mapsto x^2$, wobei $\mathbb{R}_{>0}$ die Menge der positiven reellen Zahlen bezeichnet.

3. $f : \mathbb{R} \to \mathbb{R}, x \mapsto \sin(x)$.

4. $f : \mathbb{N} \to \{0, 1\}, x \mapsto f(x) = \begin{cases} 0 & \text{für} \quad x \text{ gerade} \\ 1 & \text{für} \quad x \text{ ungerade} \end{cases}$

5. Ist B die Menge der Bücher der Universitätsbibliothek Karlsruhe und U die Menge der Bibliotheksbenutzer, so ist die Zuordnung $L : B \rightsquigarrow U$, die jedem Buch seine Leser zuordnet, keine Abbildung (wieso nicht?).

An den Beispielen zeigen sich einige typische Eigenschaften von Abbildungen, die wir in den folgenden Definitionen präzisieren. Für die Abbildung $f : A \to B$ sagen wir:

Definition 4.15 **(a)** f heißt **surjektiv**, wenn $f(A) = B$.

Jedes $b \in B$ kommt hier als Bildelement $f(a)$ vor. Man sagt auch: f ist eine Abbildung von A <u>auf</u> B.

(b) f heißt **injektiv**, wenn gilt:

$$\forall x_1, x_2 \in A \ : \ x_1 \neq x_2 \quad \Longrightarrow \quad f(x_1) \neq f(x_2)\,.$$

Bei injektiven Abbildungen haben also verschiedene Elemente auch verschiedene Bilder. Dazu äquivalent ist

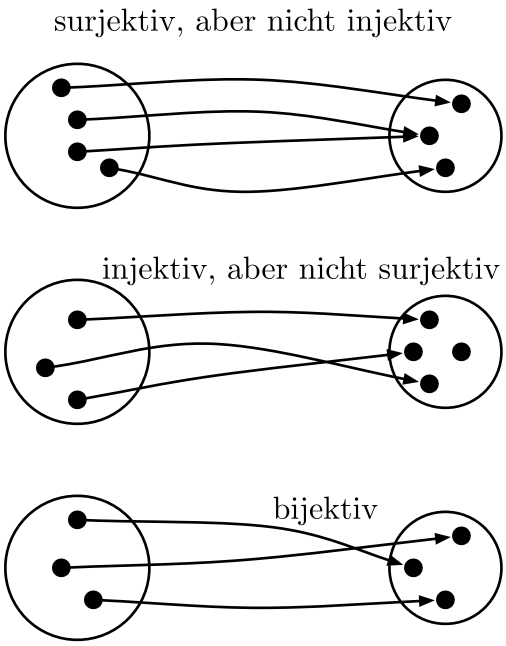
$$\forall x_1, x_2 \in A \ : \ f(x_1) = f(x_2) \quad \Longrightarrow \quad x_1 = x_2.$$

Eine solche injektive Abbildung besitzt eine **Umkehrabbildung**, nämlich

$$f^{-1} : f(A) \to A, \quad y \mapsto f^{-1}(y) \quad \text{mit} \quad f^{-1}(y) = x, \text{ wenn } f(x) = y.$$

Es ist $f^{-1}(f(x)) = x$ für alle $x \in A$ und $f(f^{-1}(y)) = y$ für alle $y \in f(A)$.

(c) f heißt **bijektiv**, wenn f injektiv und surjektiv ist.



Eine bijektive Selbstabbildung einer endlichen Menge heißt **Permutation** von A.

In Beispiel 4.14 ist 3. weder surjektiv noch injektiv, 4. ist surjektiv, aber nicht injektiv, 1. ist injektiv, aber nicht surjektiv, 2. ist eine bijektive Selbstabbildung der Menge $\mathbb{R}_{>0}$.

Definition 4.16 **(a)** Zwei Abbildungen $f : A \to B$ und $f' : A' \to B'$ sind **gleich**, wenn $A = A', B = B'$ und $f(x) = f'(x)$ für alle $x \in A = A'$.

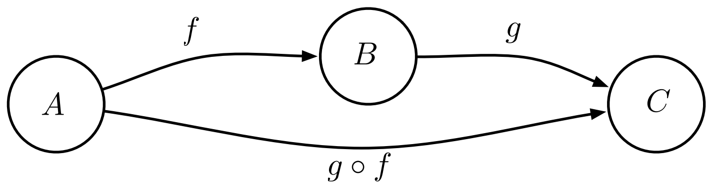
(b) Es seien $f : A \to B$ und $g : A' \to B$ zwei Abbildungen mit $A' \subset A$, und für jedes $x \in A'$ sei $f(x) = g(x)$. Dann heißt g die **Einschränkung** von f auf A' (Schreibweise: $g = f|_{A'}$). Umgekehrt heißt f eine **Fortsetzung** von g auf A.

Unter geeigneten Bedingungen kann man Abbildungen „nacheinander“ ausführen oder „verketten“:

(c) Es seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Dann heißt die Abbildung

$$h : A \to C, \quad x \mapsto h(x) := g(f(x))$$

die **Verkettung** von f und g. Schreibweise: $h = g \circ f$ (gelesen: g nach f).



Im Allgemeinen ist $g \circ f \neq f \circ g$. Jedoch gilt das Assoziativgesetz für Verkettungen:

Hilfssatz 4.17 *Für die Abbildungen $f : A \to B$, $g : B \to C$, $h : C \to D$ ist $h \circ (g \circ f) = (h \circ g) \circ f$.*

Beweis: Die Verkettungen sind alle ausführbar, Definitionsmenge ist jeweils A, Zielmenge jeweils D, und es gilt für alle $x \in A$

$$\begin{aligned} (h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) \\ ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))). \end{aligned}$$

■

4.5 Relationen

Definition 4.18 A und B seien zwei Mengen. Eine **Relation** ist eine Teilmenge $R \subset A \times B$ des cartesischen Produkts $A \times B$. Für $(x, y) \in R$ schreibt man auch xRy und sagt: „x steht in der Relation R zu y“.

Beispiel 4.19

1. A = Menge der Männer, B = Menge der Frauen, $R := \{(x,y) \in A \times B \mid x \text{ ist verheiratet mit } y\}$.

2. A = Menge der Punkte, B = Menge der Geraden in der Ebene, $R := \{(x,y) \in A \times B \mid \text{Der Punkt } x \text{ liegt auf der Geraden } y\}$.

4.5.1 Ordnungsrelationen

Es sei $A = B$ und $R \subset A \times A$. Wir verwenden hier anstatt R das Zeichen $\leq$.

Definition 4.20 Eine Relation $\leq$ heißt **Ordnungsrelation** in A und $(A, \leq)$ heißt (partiell) **geordnete Menge**, wenn für alle $a, b, c \in A$ gilt:

O1 $a \leq a$ (reflexiv)

O2 $a \leq b \wedge b \leq a \implies a = b$ (antisymmetrisch)

O3 $a \leq b \wedge b \leq c \implies a \leq c$ (transitiv).

Eine Menge A mit Ordnungsrelation $\leq$ heißt **total geordnet**, wenn für alle $a, b \in A$ gilt:

$$a \leq b \ \vee \ b \leq a.$$

Beispiel 4.21

1. Für eine beliebige Menge M ist die Inklusion $\subset$ eine Ordnungsrelation in der Potenzmenge $\mathcal{P}(M)$ und $(\mathcal{P}(M), \subset)$ ist partiell geordnet.

2. $(\mathbb{N}, \leq)$ ist eine total geordnete Menge.

In Beispiel 2 sind je zwei Elemente **vergleichbar**: Für beliebige $x, y \in \mathbb{N}$ ist $x \leq y$ oder $y \leq x$. In Beispiel 1 gilt das nicht: Man kann bei einer Menge mit mindestens zwei Elementen stets Teilmengen X, Y finden, für die weder $X \subset Y$ noch $Y \subset X$ gilt.

4.5.2 Äquivalenzrelationen

Es sei wieder $A = B$ und $R \subset A \times A$. Für R verwenden wir jetzt das Zeichen $\sim$.

Definition 4.22 $\sim$ heißt **Äquivalenzrelation**, wenn für alle $a, b, c \in A$ gilt:

Ä1 $a \sim a$ (reflexiv)

Ä2 $a \sim b \implies b \sim a$ (symmetrisch)

Ä3 $a \sim b \,\wedge\, b \sim c \implies a \sim c$ (transitiv).

Äquivalenzrelationen sind die vielleicht wichtigsten Relationen. Sie kommen in allen Bereichen der Mathematik vor.

Beispiel 4.23

1. A sei die Menge der Geraden in einer Ebene. $g \sim h$ gelte genau dann, wenn die Geraden g, h parallel sind (d.h. keinen Schnittpunkt haben oder zusammenfallen). Man sieht leicht ein, dass **Ä1**, **Ä2** und **Ä3** erfüllt sind.

2. A sei die Potenzmenge $\mathcal{P}(M)$ einer Menge M. Für zwei Teilmengen X, Y von M gelte $X \sim Y$ genau dann, wenn es eine bijektive Abbildung von X auf Y gibt. X und Y heißen dann **gleichmächtig**. Gleichmächtigkeit ist eine Äquivalenzrelation.

Es sei $\sim$ eine Äquivalenzrelation in A. Zu jedem $a \in A$ bilden wir die Menge

$$K_a := \{x \in A \mid x \sim a\},$$

der Elemente aus A, die zu a äquivalent sind. K_a heißt (Äquivalenz-)**Klasse** von a; und a ist ein **Repräsentant** der Klasse K_a.

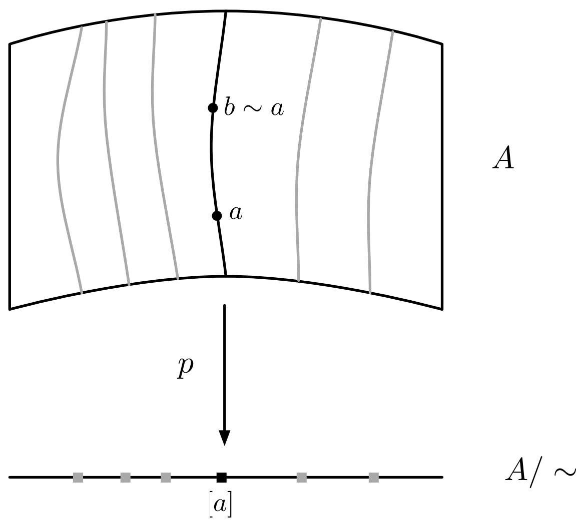
Satz 4.24 (Äquivalenzklassen-Zerlegung) *Ist $\sim$ eine Äquivalenzrelation in der Menge A, so ist A die disjunkte Vereinigung der Äquivalenzklassen von $\sim$.*

Beweis: Wegen der Reflexivität **Ä1** ist $a \in K_a$, also liegt jedes a in *wenigstens* einer Klasse. Damit haben wir $A \subset \bigcup_{a \in A} K_a \subset A$. Bleibt zu zeigen, dass zwei beliebige Klassen K_b und K_c entweder gleich oder disjunkt sind. Nehmen wir also an, dass $K_b \cap K_c \neq \emptyset$. Sei dann etwa $a \in K_b \cap K_c$. Nach Definition einer Äquivalenzklasse gilt $a \sim b$ und $a \sim c$. Mit **Ä2** und **Ä3** folgt dann aber $b \sim c$. Ist jetzt $x \in K_b$, also $x \sim b$, so folgt mit $b \sim c$ wegen **Ä3** $x \sim c$, d.h. $x \in K_c$, also $K_b \subset K_c$. Entsprechend folgt $K_c \subset K_b$, und damit schließlich $K_b = K_c$. ■

Die Menge der Klassen einer Äquivalenzrelation in A nennen wir **Faktormenge** $\tilde{A}$ von A bezüglich $\sim$. Die Abbildung

$$f : A \to \tilde{A} = A/\sim, \quad a \mapsto f(a) = K_a =: \tilde{a},$$

die jedem $a \in A$ seine Klasse $K_a = \tilde{a} \in \tilde{A}$ zuordnet, heißt zugehörige **natürliche** (oder **kanonische**) **Projektion**. Eine andere übliche Schreibweise für die Äquivalenzklassen ist $[a] = K_a$.



4.5.3 Beispiel: Die Menge $\mathbb{Z}/n\mathbb{Z}$ der Restklassen modulo n

Wir betrachten die Menge $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ der ganzen Zahlen und wählen eine natürliche Zahl n. Mit diesem n definieren wir auf $\mathbb{Z}$ die Relation $\sim$ durch

$$a \sim b \iff n \text{ teilt } b - a. \tag{4.2}$$

Dabei bedeutet „n teilt $b - a$" wie üblich $\exists z \in \mathbb{Z} : b - a = zn$, bzw.

$$\exists z \in \mathbb{Z} : b = a + zn\,. \tag{4.3}$$

Also sind a und b äquivalent, wenn beide bei Division durch n den gleichen Rest ergeben. Für (4.2) bzw. (4.3) schreibt man kürzer $b \equiv a \pmod{n}$. Sprechweise: „b kongruent a modulo n".

Beispielsweise ist $19 \equiv 9 \pmod 5, 19 \equiv 4 \pmod 5, 19 \equiv -1 \pmod 5$. Die Relation $\sim$ in (4.2) ist eine Äquivalenzrelation in $\mathbb{Z}$: Für alle $a, b, c \in \mathbb{Z}$ gilt nämlich

Ä1: $a \equiv a \pmod{n}$, denn n teilt $a - a = 0$.

Ä2: Gilt $b \equiv a \pmod{n}$, so gibt es ein $z \in \mathbb{Z}$ mit $b = a + zn$. Also ist $a = b + (-z)n$ mit $-z \in \mathbb{Z}$ und somit $a \equiv b \pmod{n}$.

Ä3: Gilt $b \equiv a \pmod{n}$ und $c \equiv b \pmod{n}$, so gibt es $z_1, z_2 \in \mathbb{Z}$ mit $b = a + z_1 n, c = b + z_2 n$ und damit $c = a + (z_1 + z_2)n$. Wegen $z_1 + z_2 \in \mathbb{Z}$ ist also $c \equiv a \pmod{n}$.

Die von $a \in \mathbb{Z}$ erzeugte Klasse ist

$$\tilde{a} = \{x \in \mathbb{Z} \mid x \sim a\} = \{x \in \mathbb{Z} \mid \exists z \in \mathbb{Z} : x = a + zn\} = a + n\mathbb{Z}.$$

Um die Faktormenge für dieses Beispiel explizit zu beschreiben, benutzen wir die sogenannte **Division mit Rest** in $\mathbb{Z}$: Zu je zwei ganzen Zahlen s, t mit $t \neq 0$ gibt es genau zwei weitere ganze Zahlen q, r mit $s = qt + r$ und $0 \leq r < |t|$.

Ist nun $b \equiv a \pmod{n}$, so gibt es ein $z \in \mathbb{Z}$ mit $b = zn + a$. Nach „Division mit Rest" können wir also genau einen **Repräsentanten** r von $\tilde{a}$ wählen mit $0 \leq r < n$. Die Klasse $\tilde{a} = r + n\mathbb{Z}$ besteht dann aus allen $x \in \mathbb{Z}$, die bei Division durch n den Rest r haben. $\tilde{a}$ heißt daher auch **Restklasse mod** n. Die Faktormenge $\mathbb{Z}/\sim$, die wir auch mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnen, können wir dann schreiben als

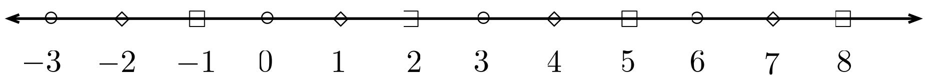
$$\mathbb{Z}/n\mathbb{Z} = \{\tilde{0}, \tilde{1}, \ldots, \widetilde{n-1}\}.$$

Für $n = 3$ sind die Klassen von $\mathbb{Z}/3\mathbb{Z}$ in der folgenden Abbildung skizziert:

$\circ \equiv 0 \mod 3$

$\diamond \equiv 1 \mod 3$

$\square \equiv 2 \mod 3$



5 Algebraische Grundbegriffe

5.1 Worum es geht: das Beispiel der ganzen Zahlen

Was macht man eigentlich, wenn man „rechnet“? Mit welchen Objekten kann man rechnen? Welche Gesetze müssen gelten, damit man Gleichungen formulieren und lösen kann?

Wir betrachten dazu zunächst einmal das Modell-Beispiel der Menge der ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Ganze Zahlen kann man addieren, subtrahieren und multiplizieren. Wenn man hingegen eine ganze Zahl durch eine andere dividiert, erhält man im Allgemeinen eine rationale Zahl; die Division „führt aus der Menge der ganzen Zahlen heraus“.

Diese Tatsachen kann man mit Hilfe des Abbildungsbegriffes präzisieren. Wir fassen die Addition zweier ganzer Zahlen als eine Abbildung mit zwei Argumenten auf, und ordnen diesem Paar eine weitere ganze Zahl zu:

$$+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (x, y) \mapsto x + y,$$

wobei wir wie üblich $x+y$ statt $+(x, y)$ schreiben. Eine Abbildung, die zwei Elementen einer Menge ein Element aus derselben Menge zuordnet, heißt *Verknüpfung*.

In $\mathbb{Z}$ gibt es ein Element, das bezüglich der Addition vor allen anderen ausgezeichnet ist: die Null. Denn diese hat als einziges Element die Eigenschaft, dass man sie zu allen Elementen $a \in \mathbb{Z}$ hinzuaddieren kann, ohne dass sich die Zahl a dadurch ändert: $a + 0 = a$ für alle $a \in \mathbb{Z}$. Man sagt „0 ist das neutrale Element bezüglich der Addition“.

Das Addieren einer Zahl $a \in \mathbb{Z}$ zu einer weiteren Zahl lässt sich rückgängig machen durch das Subtrahieren von a, was das Gleiche ist wie das Addieren von $-a \in \mathbb{Z}$. Man sagt: $-a$ ist das *inverse Element* von a bezüglich der Addition. Das inverse Element $-a$ von a zeichnet sich dadurch aus, dass gilt

$$a + (-a) = 0,$$

d.h. addiert man zu einer Zahl $a \in \mathbb{Z}$ ihr inverses Element $-a$, so erhält man das neutrale Element 0.

Durch diese Struktur sind wir in der Lage, Gleichungen der Form $a + x = b$ nach x aufzulösen: wir addieren auf beiden Seiten der Gleichung das inverse Element $-a$ von a und erhalten $x = b + (-a) := b - a$ als eindeutige Lösung.

Betrachten wir nun die Multiplikation auf $\mathbb{Z}$. Auch diese schreiben wir als Abbildung

$$\cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (x, y) \mapsto x \cdot y.$$

Bei dieser Verknüpfung gibt es ebenfalls ein neutrales Element: die Eins. Es gilt nämlich $a \cdot 1 = 1 \cdot a = a$ für alle $a \in \mathbb{Z}$. Jedoch lässt sich die Multiplikation nicht umkehren (jedenfalls nicht, ohne die Menge $\mathbb{Z}$ zu verlassen). Z.B. lässt sich die Multiplikation einer Zahl $a \in \mathbb{Z}$ mit 2 nicht rückgängig machen, denn dafür müsste man mit der *rationalen* Zahl $\frac{1}{2}$ multiplizieren. Da $\frac{1}{2}$ aber nicht in $\mathbb{Z}$ liegt, gibt es in $\mathbb{Z}$ kein inverses Element von 2 bezüglich der Multiplikation:

$$2 \cdot x = 1 \quad \text{gilt für keine Zahl}\, x \in \mathbb{Z}.$$

Zwei weitere Eigenschaften der Addition und der Multiplikation haben wir stillschweigend verwendet. Bei der Durchführung mehrerer Additionen bzw. mehrerer Multiplikationen kommt es nicht auf die Reihenfolge an: für beliebige $a, b, c \in \mathbb{Z}$ gilt $(a+b)+c = a+(b+c)$ und $a+b = b+a$ (entsprechend für die Multiplikation). Diese Eigenschaften sind natürlich für das „Rechnen“ mit ganzen Zahlen entscheidend. Im Folgenden definieren wir die in diesem Beispiel aufgetretenen Konzepte allgemein und untersuchen ihre Beziehungen.

5.2 Gruppen: die wichtigsten algebraischen Objekte

Definition 5.1 Gegeben sei eine Menge A. Eine (innere) **Verknüpfung** $*$ auf A ist eine Abbildung

$$* : A \times A \to A, \quad (x, y) \mapsto x * y$$

Bemerkung 5.2 Bei Verknüpfungen schreibt man $x * y$ für das Bild $*(x, y)$ von (x, y) unter der Abbildung $*$. Statt $*$ verwendet man auch häufig die Verknüpfungszeichen $+, -, \cdot$ usw.

Beispiel 5.3

1. Die Addition $+$ und die Multiplikation $\cdot$ sind Verknüpfungen auf $\mathbb{Z}$, aber nicht die Division :.

2. Die Subtraktion $-$ ist keine Verknüpfung auf $\mathbb{N}$ (da $2-4 \notin \mathbb{N}$) und die Division : keine Verknüpfung auf $\mathbb{R}$ (da die Division durch 0 in $\mathbb{R}$ nicht erklärt ist). Die Division : ist aber eine Verknüpfung auf $\mathbb{R}\backslash\{0\}$.

3. Auf der Menge $\mathrm{Abb}(M, M) = \{f : M \to M\}$ aller Selbstabbildungen einer nichtleeren Menge M ist die Verkettung eine Verknüpfung.

Definition 5.4 Wir nennen eine Verknüpfung $*$ auf einer Menge A **assoziativ**, wenn

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

gilt, und **kommutativ**, wenn

$$\forall a, b \in A \; : \; a * b = b * a$$

gilt.

Beispiel 5.5

1. Auf $\mathbb{Z}$ ist $+$ assoziativ und kommutativ.

2. Auf $\mathbb{Z}$ ist $-$ weder assoziativ noch kommutativ.

3. Auf $\mathbb{Z}$ ist die Verknüpfung $\diamond : \; (a, b) \mapsto |a - b|$ nicht assoziativ, aber kommutativ.

4. Die Verkettung von Abbildungen auf der Menge $\mathrm{Abb}(M, M)$ aller Selbstabbildungen von A ist stets assoziativ, aber i.Allg. nicht kommutativ.

5. In $\mathbb{Z}/n\mathbb{Z}$ (vgl. Abschnitt 4.5.3) definieren wir eine Verknüpfung $+$ durch

$$+ : \; \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}; \quad (\tilde{a}, \tilde{b}) \mapsto \tilde{a} + \tilde{b} := \widetilde{a + b}\,.$$

(Beachten Sie, dass das Zeichen $+$ hier in zwei verschiedenen Bedeutungen benutzt wird: einmal ist $+$ die „gewöhnliche" Addition in $\mathbb{Z}$ und einmal die neu definierte Addition in $\mathbb{Z}/n\mathbb{Z}$.) Damit obige Definition sinnvoll ist, hat man die Unabhängigkeit der Summenklasse $\widetilde{a + b}$ von der Repräsentantenauswahl zu prüfen, damit wirklich jedem Paar $(\tilde{a}, \tilde{b})$ genau eine Klasse $\widetilde{a + b}$ als Bild zugeordnet wird (Wohldefiniertheit):

Haben wir $\tilde{a}_0 = \tilde{a}$, $\tilde{b}_0 = \tilde{b}$, also $a_0 \sim a$, $b_0 \sim b$, so gilt $a_0 \equiv a \pmod n$, $b_0 \equiv b \pmod n$. Es gibt also $z_1, z_2 \in \mathbb{Z}$ mit $a_0 = a + z_1 n$, $b_0 = b + z_2 n$, woraus $a_0 + b_0 = (a + b) + (z_1 + z_2)n$, also

$$a_0 + b_0 \equiv a + b \pmod n$$

folgt. Es gilt also tatsächlich $\widetilde{a_0 + b_0} = \widetilde{a + b}$. Damit haben wir gezeigt, dass die auf $\mathbb{Z}/n\mathbb{Z}$ definierte Addition tatsächlich eine Verknüpfung auf $\mathbb{Z}/n\mathbb{Z}$ ist. Sie ist assoziativ und kommutativ. Die **Verknüpfungstafel** für die Addition $+$ etwa auf $\mathbb{Z}/3\mathbb{Z}$ sieht folgendermaßen aus

$+$	$\tilde{0}$	$\tilde{1}$	$\tilde{2}$
$\tilde{0}$	$\tilde{0}$	$\tilde{1}$	$\tilde{2}$
$\tilde{1}$	$\tilde{1}$	$\tilde{2}$	$\tilde{0}$
$\tilde{2}$	$\tilde{2}$	$\tilde{0}$	$\tilde{1}$

Definition 5.6 **(a)** Ist $*$ eine Verknüpfung auf A und gibt es ein Element $e \in A$ mit

$$\forall a \in A \ : \ e * a = a = a * e,$$

so heißt e **neutrales Element** bezüglich $*$.

(b) Ist $*$ eine Verknüpfung auf A mit neutralem Element e und gibt es zu einem Element $a \in A$ ein $a^{-1} \in A$ mit

$$a^{-1} * a = e = a * a^{-1},$$

so heißt a^{-1} **inverses Element von** a.

Bemerkung 5.7 **(a)** Es gibt *höchstens ein* neutrales Element für eine Verknüpfung $*$ auf A. Denn sind e_1 und e_2 neutrale Elemente bezüglich $*$, so ist nach Definition $e_1 * e_2 = e_1$, aber auch $e_1 * e_2 = e_2$, also $e_1 = e_2$.

(b) Unter der zusätzlichen Voraussetzung, dass $*$ assoziativ ist, lässt sich zeigen, dass es zu einem $a \in A$ *höchstens ein* Inverses a^{-1} gibt! Denn sind a_1^{-1} und a_2^{-1} inverse Elemente von a, so gilt

$$a_1^{-1} = a_1^{-1} * e = a_1^{-1} * (a * a_2^{-1}) = (a_1^{-1} * a) * a_2^{-1} = e * a_2^{-1} = a_2^{-1}.$$

Beispiel 5.8 Die Addition $+$ auf $\mathbb{Z}$ hat das neutrale Element 0. Das inverse Element von $z \in \mathbb{Z}$ ist $-z$.

Besonders wichtig und reichhaltig sind assoziative Verknüpfungen mit neutralem Element, bei der *jedes* Element ein Inverses besitzt:

Definition 5.9 Eine **Gruppe** ist ein Paar $(G, *)$ bestehend aus einer (nichtleeren) Menge G und einer Verknüpfung $*$ auf G mit folgenden Eigenschaften:

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

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

G3 (**inverses Element**): $\forall a \in G \ \exists a^{-1} \in G : a^{-1} * a = e = a * a^{-1}$.

Gilt zusätzlich

G4 $\forall a, b \in G : a * b = b * a$,

so heißt die Gruppe G **abelsch**.

Bemerkung 5.10 Nach der Bemerkungen 5.7 ist das neutrale Element einer Gruppe *eindeutig* bestimmt und zu jedem Gruppenelement a gibt es *genau ein* Inverses a^{-1}.

Beispiel 5.11

1. $(\mathbb{Z},+),(\mathbb{Q},+),(\mathbb{R},+),(\mathbb{R}\backslash\{0\},\cdot)$ sind abelsche Gruppen. $(\mathbb{Z},\cdot)$ und $(\mathbb{Q},\cdot)$ sind keine Gruppen.

2. $(\mathbb{Z}/n\mathbb{Z},\ +)$ ist eine abelsche Gruppe. $\tilde{0}$ ist das neutrale Element und $\widetilde{n-r}$ ist das zu $\tilde{r}$ inverse Element $(0 \leq r < n)$.

3. Für die Menge $\mathrm{Abb}(M,M)$ der Selbstabbildungen $f : M \to M$ ist die Verkettung assoziativ mit der Identität id_M als neutralem Element; $\mathrm{Abb}(M,M)$ ist aber im Allgemeinen keine Gruppe, weil die Gruppeneigenschaft G3 nicht erfüllt ist. Beschränkt man sich jedoch auf die Teilmenge S_M der bijektiven Selbstabbildungen von M, so ist $(S_M,\circ)$ eine Gruppe. Für den Fall einer endlichen Menge M werden wir uns mit solchen Gruppen im nächsten Abschnitt noch genauer beschäftigen.

Hilfssatz 5.12 (Multiplikation mit Inversen) *In einer Gruppe* $(G,*)$ *sind die Gleichungen* $a*x=b$ *und* $x*c=d$ *eindeutig nach* x *lösbar.*

Beweis: $x=a^{-1}*b$ ist Lösung von $a*x=b$, denn

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

Diese Lösung ist die einzige, denn sind x_1,x_2 zwei Lösungen von $a*x=b$, so gilt

$$\begin{aligned} a*x_1=a*x_2 \quad &\Longrightarrow\quad a^{-1}*(a*x_1)=a^{-1}*(a*x_2)\\ &\Longrightarrow\quad (a^{-1}*a)*x_1=(a^{-1}*a)*x_2\\ &\Longrightarrow\quad e*x_1=e*x_2\\ &\Longrightarrow\quad x_1=x_2. \end{aligned}$$

Entsprechend hat $x*c=d$ die eindeutige Lösung $x=d*c^{-1}$. ■

Bemerkung 5.13 Man kann zeigen, dass eine Menge G mit einer assoziativen Verknüpfung $*$ bereits dann eine Gruppe ist, wenn gilt:

$$\exists e \in G\ \forall a \in G : a*e=a \quad (e \text{ ist rechtsneutral})$$

und

$$\forall a \in G\ \exists a^{-1} \in G : a*a^{-1}=e \quad (a^{-1} \text{ ist rechtsinvers}).$$

5.2.1 Beispiel: Die symmetrische Gruppe

Definition 5.14 Es sei M eine endliche Menge. Eine bijektive Selbstabbildung von M heißt **Permutation**. Die Menge S_M der Permutationen von M ist eine Gruppe bezüglich der Verkettung $\circ$ von Abbildungen und heißt **symmetrische Gruppe** von M.

Jede endliche Menge mit m Elementen ist bijektiv zur Menge $M = \{1, 2, \ldots, m\}$. Es genügt also, dieses spezielle M zu betrachten. Statt S_M schreiben wir dann S_m.

Bemerkung 5.15 Mittels vollständiger Induktion beweist man: Es gibt $1 \cdot 2 \cdot 3 \cdot \cdots \cdot m = m!$ Permutationen der Menge $\{1, 2, \ldots, m\}$; die Gruppe S_m hat also $m!$ Elemente.

Eine Permutation $\pi \in S_m$ schreiben wir schematisch folgendermaßen:

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & m \\ \pi(1) & \pi(2) & \cdots & \pi(m) \end{pmatrix}.$$

Wir setzen also unter jedes $i \in \{1, 2, \ldots, m\}$ das entsprechende Bild $\pi(i)$. Zum Beispiel hat die symmetrische Gruppe S_3 von $M = \{1, 2, 3\}$ die $3! = 1 \cdot 2 \cdot 3 = 6$ Elemente

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \pi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

und die Gruppentafel

$(S_3, \circ)$	π_1	π_2	π_3	π_4	π_5	π_6
π_1	π_1	π_2	π_3	π_4	π_5	π_6
π_2	π_2	π_3	π_1	π_6	π_4	π_5
π_3	π_3	π_1	π_2	π_5	π_6	π_4
π_4	π_4	π_5	π_6	π_1	π_2	π_3
π_5	π_5	π_6	π_4	π_3	π_1	π_2
π_6	π_6	π_4	π_5	π_2	π_3	π_1

Dabei steht beispielsweise in der 2. Zeile und 5. Spalte der Tafel die Verkettung $\pi_2 \circ \pi_5 = \pi_4$, in der 5. Zeile und 2. Spalte dagegen $\pi_5 \circ \pi_2 = \pi_6$. Die Gruppe S_3 ist also *nicht abelsch*.

Für $m = 1$ besteht die symmetrische Gruppe S_1 nur aus der identischen Abbildung von $M = \{1\}$. Wir wollen im Folgenden stets $m \geq 2$ voraussetzen und zeigen, dass sich jedes $\pi \in S_m$ als Verkettung von gewissen „einfachen“ Permutationen darstellen lässt. Eine **Transposition** ist eine Permutation aus S_m, bei der zwei verschiedene, fest gewählte Zahlen $i, k \in \{1, 2, \ldots, m\}$ vertauscht werden, während alle übrigen Zahlen fest bleiben, also

$$\begin{aligned} \pi(i) &= k && (i \neq k), \\ \pi(k) &= i && (i \neq k), \\ \pi(l) &= l && \text{für alle } l \neq i, k. \end{aligned}$$

Man schreibt für diese Transposition auch kurz $(i\ k)$. Zum Beispiel ist für $m = 3$

$$\pi_4 = (2\ 3), \quad \pi_5 = (3\ 1), \quad \pi_6 = (1\ 2).$$

Für π_1, π_2, π_3 gilt

$$\pi_1 = (1\ 2) \circ (1\ 2), \quad \pi_2 = (1\ 3) \circ (1\ 2), \quad \pi_3 = (2\ 3) \circ (1\ 2),$$

oder auch

$$\pi_3 = (2\ 3) \circ (1\ 3) \circ (2\ 3) \circ (1\ 3).$$

Bemerkung 5.16 Ist $\tau = (i\ k)$ eine Transposition, so gilt $\tau \circ \tau = \mathrm{id}$; insbesondere also $\tau^{-1} = \tau$.

Allgemein gilt der

Satz 5.17 (S_m wird von Transpositionen erzeugt) *Jede Permutation $\pi \in S_m$ (für $m \geq 2$) lässt sich als Verkettung von Transpositionen darstellen.*

BEWEIS: Wir beweisen die Aussage durch vollständige Induktion: Die Aussage des Satzes ist für $m = 2$ richtig, denn für die S_2 gilt

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (1\ 2) \circ (1\ 2), \qquad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (1\ 2).$$

Unter der Annahme, dass der Satz für $m = k \geq 2$ gilt, zeigen wir jetzt, dass er auch für $m = k + 1$ richtig ist.

1. FALL: Sei $\pi \in S_{k+1}$. Wenn $\pi(1) = 1$, so lässt sich

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & k+1 \\ 1 & \pi(2) & \cdots & \pi(k+1) \end{pmatrix}$$

als Permutation der k Zahlen $2, 3, \ldots, k+1$ nach Induktionsannahme als Verkettung von Transpositionen darstellen.

2. FALL: Wenn $\pi(1) = i \neq 1$, so gilt

$$\begin{aligned}
\pi &= \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & k+1 \\ i & \pi(2) & \cdots & \pi(i-1) & \pi(i) & \pi(i+1) & \cdots & \pi(k+1) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & k+1 \\ \pi(i) & \pi(2) & \cdots & \pi(i-1) & i & \pi(i+1) & \cdots & \pi(k+1) \end{pmatrix} \circ (1\ i)
\end{aligned}$$

und π ist wieder als Verkettung von Transpositionen darstellbar, weil in der vorletzten Permutation i fest ist. ■

Definition 5.18 Es sei $\pi \in S_m$ eine Permutation. Die **Fehlstandszahl** $F(\pi)$ von π ist die (eindeutige) Anzahl der Fälle, in denen für $i < k$ gilt $\pi(i) > \pi(k)$. Die Permutationen mit gerader Fehlstandszahl $F(\pi)$ heißen **gerade**, die Permutationen mit ungerader Fehlstandszahl heißen **ungerade**.

Beispielsweise ist die Fehlstandszahl für

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$$

gleich 5, weil 2 vor 1, 4 vor 1, 4 vor 3, 5 vor 1 und 5 vor 3 steht.

Die Anzahl der Transpositionen in der Darstellung einer Permutation ist *nicht eindeutig* bestimmt. Zum Beispiel gilt

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix} = (1\ 4) \circ (1\ 2) \circ (3\ 5) = (2\ 3) \circ (2\ 5) \circ (1\ 3) \circ (2\ 3) \circ (2\ 4).$$

Hingegen gilt der

Hilfssatz 5.19 (Anzahl Transpositionen) *Sei $\pi \in S_m$ $(m \geq 2)$ eine Permutation. Die Anzahl der Transpositionen in allen Darstellungen von π ist für π gerade stets gerade und für π ungerade stets ungerade.*

BEWEIS: Wir überlegen zuerst wie sich die Fehlstandszahl ändert, wenn man eine Permutation π mit einer Transposition verkettet.

1. FALL: (Transposition vertauscht zwei benachbarte Ziffern): Bei

$$(\pi(i) \;\; \pi(i+1)) \circ \begin{pmatrix} 1 & \cdots & i & i+1 & \cdots & m \\ \pi(1) & \cdots & \pi(i) & \pi(i+1) & \cdots & \pi(m) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cdots & i & i+1 & \cdots & m \\ \pi(1) & \cdots & \pi(i+1) & \pi(i) & \cdots & \pi(m) \end{pmatrix}$$

ändert sich die Fehlstandszahl gegenüber $F(\pi)$ um $+1$, falls $\pi(i) < \pi(i+1)$ bzw. um -1, falls $\pi(i) > \pi(i+1)$.

2. FALL: (Transposition vertauscht zwei nicht benachbarte Ziffern): Bei

$$(\pi(i) \;\; \pi(k)) \circ \begin{pmatrix} 1 & \cdots & i & \cdots & k & \cdots & m \\ \pi(1) & \cdots & \pi(i) & \cdots & \pi(k) & \cdots & \pi(m) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cdots & i & \cdots & k & \cdots & m \\ \pi(1) & \cdots & \pi(k) & \cdots & \pi(i) & \cdots & \pi(m) \end{pmatrix}$$

können wir die Vertauschung durch schrittweises Vertauschen benachbarter Ziffern erreichen, denn es ist (wir schreiben π_j für $\pi(j)$)

$$\begin{aligned} (\pi_i \;\; \pi_k) \circ \pi &= (\pi_{i+1} \; \pi_k) \circ \cdots \circ (\pi_{k-2} \; \pi_k) \circ (\pi_{k-1} \; \pi_k) \circ \\ &\quad \circ (\pi_i \; \pi_k) \circ (\pi_i \; \pi_{k-1}) \circ \cdots \circ (\pi_i \; \pi_{i+2}) \circ (\pi_i \; \pi_{i+1}) \circ \pi. \end{aligned}$$

Bei jedem der $k-i+k-1-i = 2(k-i)-1$ Schritte ändert sich F um ± 1, insgesamt also um eine ungerade Zahl.

FAZIT: Bei Verkettung von π mit einer Transposition τ gilt für die Fehlstandszahl

$$F(\tau \circ \pi) = F(\pi) + n \qquad \text{mit ungeradem } n.$$

Nach Satz 5.17 ist die Permutation π als (nicht eindeutige) Verkettung von, sagen wir r, Transpositionen $\tau_1, \tau_2, \ldots, \tau_r$ darstellbar. Wir können also schreiben

$$\pi = \tau_r \circ \cdots \circ \tau_1 \circ \text{ id}.$$

Ausgehend von der identischen Abbildung id, die die Fehlstandszahl 0 hat, ändert sich auf der rechten Seite obiger Gleichung bei jedem Schritt die Fehlstandszahl um eine ungerade Zahl, so dass

$$\begin{aligned} F(\pi) &= 0 + n_1 + n_2 + \cdots + n_r \\ &= (2z_1 + 1) + (2z_2 + 1) + \cdots + (2z_r + 1) \\ &= 2z + r. \end{aligned}$$

Ist nun π eine gerade Permutation, also die durch π eindeutig bestimmte Fehlstandszahl $F(\pi)$ gerade, so muss nach obiger Formel auch die Anzahl r der Transpositionen gerade sein. Ist π (und damit auch $F(\pi)$) ungerade, dann auch r. ■

Folgerung 5.20 *Die geraden Permutationen von S_m $(m \geq 2)$ bilden bezüglich $\circ$ eine Gruppe $(A_m, \circ)$, die sogenannte* **alternierende Gruppe**.

Bemerkung 5.21 Die Teilmenge B_m der ungeraden Permutationen von S_m $(m \geq 2)$ ist bezüglich $\circ$ *keine Gruppe*, denn $\circ$ ist keine Verknüpfung in B_m (wieso nicht?).

Die Anzahl der geraden Permutationen von S_m $(m \geq 2)$ ist ebenso groß wie die Anzahl der ungeraden Permutationen, nämlich $\frac{1}{2}m!$. Begründung: Die Abbildung

$$f : A_m \to B_m, \quad \pi_g \mapsto \pi_u = (1\ 2) \circ \pi_g$$

ist bijektiv; A_m und B_m sind also gleichmächtig.

5.2.2 Untergruppen

Die eben angetroffene Situation, dass die Teilmenge A_m von S_m selbst wieder eine Gruppe bezüglich der von S_m übernommenen Verknüpfung ist, motiviert folgende Definition:

Definition 5.22 Gegeben sei eine Gruppe $(G, *)$ und eine Teilmenge $U \subset G$, die bezüglich der von G induzierten Verknüpfung $*$ ebenfalls eine Gruppe ist. Dann heißt $(U, *)$ **Untergruppe** von $(G, *)$.

Beispiel 5.23

1. Jede Gruppe $(G, *)$ hat mindestens zwei Untergruppen: $(\{e\}, *)$ und $(G, *)$.
2. Die alternierende Gruppe $(A_m, \circ)$ ist eine Untergruppe von $(S_m, \circ)$.
3. $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Q}, +)$.

Bemerkung 5.24 Das neutrale Element e' einer Untergruppe $(U, *)$ von $(G, *)$ stimmt mit dem neutralen Element e von $(G, *)$ überein. Denn nach Hilfssatz 5.12 ist die Gleichung $e' * x = e'$ in G eindeutig lösbar; die Lösungen e und e' sind also gleich. Ebenso sieht man, dass für ein Element $a \in U \subset G$ das inverse Element in $(G, *)$ und in $(U, *)$ dasselbe ist.

Der folgende Satz zeigt, dass man nicht alle Gruppeneigenschaften nachprüfen muss, um festzustellen, ob eine Untergruppe vorliegt.

Satz 5.25 (Untergruppen-Kriterium) *Sei $(G, *)$ eine Gruppe. Eine Teilmenge $U \subset G$ ist Untergruppe von G, wenn gilt:*

UG1 $U \neq \emptyset$

UG2 $\forall a, b \in U : a * b^{-1} \in U$.

Beweis: Wegen **UG1** gibt es mindestens ein $a \in U$. Wegen **UG2** liegt mit jedem $a \in U$ auch $a * a^{-1} = e$ in U, also gilt für U die Eigenschaft **G2**. Mit e und a liegt nach **UG2** auch $e * a^{-1} = a^{-1}$ in U, also gilt **G3**. Wenn a, $b \in U$, so auch $b^{-1} \in U$ und damit nach **UG2** auch $a * b = a * (b^{-1})^{-1} \in U$, so dass $*$ eine Verknüpfung in U ist. $(U, *)$ ist assoziativ, d.h. es gilt **G1**, da $*$ auf $G \supset U$ assoziativ ist. ■

Definition 5.26 Sei $(G, *)$ eine Gruppe und $M \subset G$ eine beliebige Teilmenge. Dann heißt die kleinste Untergruppe von G, die M enthält, die **von M erzeugte Untergruppe** $\langle M \rangle$. Eine von einem einzigen Element $a \in G$ erzeugte Untergruppe heißt **zyklisch** (Schreibweise: $U = \langle a \rangle$).

Beispiel 5.27 In $(\mathbb{Z}, +)$ ist $n\mathbb{Z} = \{\ldots, -3n, -2n, -n, 0, n, 2n, 3n, \ldots\}$ die von $n \in \mathbb{Z}$ erzeugte zyklische Untergruppe: $\langle n \rangle = n\mathbb{Z}$.

5.2.3 Homomorphismen

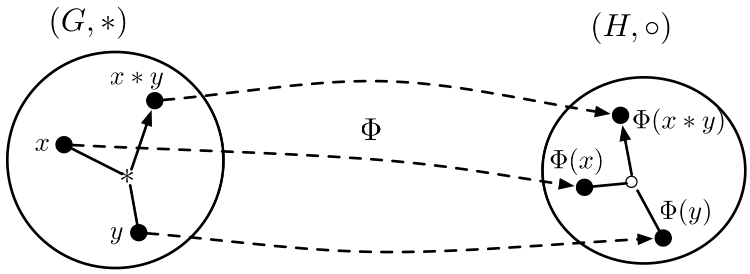
Wenn man zwei Mengen vergleichen will, auf denen Verknüpfungen definiert sind, interessiert man sich besonders für Abbildungen zwischen diesen Mengen, die mit der Verknüpfungsstruktur der Mengen „verträglich“ sind. Man nennt solche **strukturerhaltende Abbildungen** auch **Homomorphismen**. Einen bijektiven Homomorphismus nennt man **Isomorphismus**.

Welche speziellen Eigenschaften eine solche Abbildung haben muss, hängt jeweils von den gegebenen Verknüpfungen ab.

Definition 5.28 Seien $(G, *)$ und $(H, \circ)$ zwei Gruppen und $\Phi : G \to H$ eine Abbildung. Dann heißt Φ ein **(Gruppen-)Homomorphismus**, wenn gilt

$$\forall x, y \in G : \Phi(x * y) = \Phi(x) \circ \Phi(y).$$

Stimmen die beiden Gruppen $(G, *)$ und $(H, \circ)$ überein (ist Φ also eine Selbstabbildung), so spricht man von **Endomorphismen** statt von Homomorphismen. Einen bijektiven Endomorphismus nennt man auch **Automorphismus**.



Beispiel 5.29

1. Die Abbildung $\Phi_1 : \mathbb{Z} \to \mathbb{Q}$, $x \mapsto 3x$ ist ein Homomorphismus der Gruppe $(\mathbb{Z}, +)$ in die Gruppe $(\mathbb{Q}, +)$.

   Die Abbildung $\Phi_2 : \mathbb{Z} \to \mathbb{Q}$, $x \mapsto x^2$ ist kein Homomorphismus (wieso nicht?).

2. Die Abbildung $\Phi_3 : \mathbb{R} \to \mathbb{R}$, $x \mapsto 3x$ ist ein Endomorphismus der Gruppe $(\mathbb{R}, +)$, da für alle $x, y \in \mathbb{R}$ gilt $3(x + y) = 3x + 3y$. Da Φ bijektiv ist, ist Φ_3 sogar ein Automorphismus von $(\mathbb{R}, +)$.

3. Die Abbildung $\Phi_4 : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto 2x$ ist ein Endomorphismus der Gruppe $(\mathbb{Z}, +)$, aber kein Automorphismus.

4. Die Exponential-Abbildung $\exp : \mathbb{R} \to \mathbb{R}_{>0}$, $x \mapsto \mathrm{e}^x$ ist ein Isomorphismus der additiven Gruppe der reellen Zahlen $(\mathbb{R}, +)$ in die multiplikative Gruppe der positiven reellen Zahlen $(\mathbb{R}_{>0}, \cdot)$, denn exp ist bijektiv und für alle $x, y \in \mathbb{R}$ gilt $\mathrm{e}^{x+y} = \mathrm{e}^x \cdot \mathrm{e}^y$.

5.3 Ringe und Körper: die Verallgemeinerungen von $\mathbb{Z}$ und $\mathbb{R}$

Die Menge $\mathbb{Z}$ der ganzen Zahlen ist bezüglich der Addition $+$ eine (abelsche) Gruppe. Auf $\mathbb{Z}$ ist durch die Multiplikation $\cdot$ noch eine zweite Verknüpfung erklärt. Wie wir in 3.1 gesehen haben, ist $(\mathbb{Z}, \cdot)$ jedoch *keine* Gruppe. Die Multiplikation ist aber assoziativ und zudem sind Addition und Multiplikation durch Distributivgesetze verbunden. Diese Struktur verallgemeinern wir in folgender Definition.

Definition 5.30 Ein **Ring** ist eine Menge R zusammen mit zwei Verknüpfungen $+$ und $\cdot$ mit folgenden Eigenschaften:

R1 $(R, +)$ ist eine abelsche Gruppe,

R2 $\cdot$ ist assoziativ,

R3 Distributivgesetze: für alle $a, b, c \in R$ gilt:

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ und } (b + c) \cdot a = b \cdot a + c \cdot a.$$

Wenn die Verknüpfung $\cdot$ kommutativ ist, nennt man den Ring **kommutativ**. Das neutrale Element in $(R, +)$ bezeichnet man mit 0 (**Nullelement**), das zu a inverse Element mit $-a$. Die Differenz $b - a$ ist durch $b - a := b + (-a)$ erklärt. Hat der Ring auch ein neutrales Element ($\neq 0$) bezüglich der Multiplikation $\cdot$, so schreibt man dafür 1 und nennt es **Einselement**; R heißt dann **Ring mit Eins**.

Beispiel 5.31

1. $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit Eins.
2. $(\mathbb{R}, +, \cdot)$ ist ebenfalls ein kommutativer Ring mit Eins (aber noch viel mehr, siehe später).

Bemerkung 5.32 Einige allgemeine Eigenschaften von Ringen:

1. Für alle $a \in R$ gilt $a \cdot 0 = 0 = 0 \cdot a$.
2. Für alle $a, b \in R$ gilt $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$ und $(-a) \cdot (-b) = a \cdot b$.
3. Für alle $a, b, c \in R$ gilt $a \cdot (b - c) = a \cdot b - a \cdot c$.

Beweis:

1. Es ist $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Da in $(R, +)$ die Gleichung $c + x = c$ die eindeutig bestimmte Lösung $x = 0$ hat, folgt $a \cdot 0 = 0$. Entsprechend gilt $0 \cdot a = 0$.
2. Es ist $a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0$. Da $c + x = 0$ die eindeutig bestimmte Lösung $x = -c$ hat, ist $(-a) \cdot b = -(a \cdot b)$. Entsprechend folgt $a \cdot (-b) = -(a \cdot b)$. Weiter ist $(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b))$. Da in $(R, +)$ stets $-(-c) = c$ gilt, folgt schließlich $(-a) \cdot (-b) = a \cdot b$.
3. $a \cdot (b - c) = a \cdot (b + (-c)) = a \cdot b + a \cdot (-c) = a \cdot b + (-a \cdot c) = a \cdot b - a \cdot c$. ∎

Beispiel 5.33 In 5.2 haben wir auf der Menge der Restklassen modulo n eine Addition definiert durch

$$+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, \quad (\tilde{a}, \tilde{b}) \mapsto \tilde{a} + \tilde{b} := \widetilde{a + b} \quad \text{für } a \in \tilde{a},\, b \in \tilde{b}.$$

$(\mathbb{Z}/n\mathbb{Z}, +)$ ist dann eine abelsche Gruppe. Wir definieren eine weitere Verknüpfung auf $\mathbb{Z}/n\mathbb{Z}$ durch

$$\cdot : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, \quad (\tilde{a}, \tilde{b}) \mapsto \tilde{a} \cdot \tilde{b} := \widetilde{ab} \quad \text{für } a \in \tilde{a},\, b \in \tilde{b}.$$

Auch hier müssen wir die Wohldefiniertheit überprüfen. Dazu seien $\tilde{a}_0 = \tilde{a}, \tilde{b}_0 = \tilde{b}$, also $a_0 \equiv a \pmod{n}, b_0 \equiv b \pmod{n}$. Dann gibt es $z_1, z_2 \in \mathbb{Z}$ mit $a_0 = a + z_1 n, b_0 = b + z_2 n$ und es gilt $a_0 b_0 = ab + (a z_2 + b z_1 + z_1 z_2 n)n$. Wegen $a z_2 + b z_1 + z_1 z_2 n \in \mathbb{Z}$ gilt $a_0 b_0 \equiv ab \pmod{n}$, also tatsächlich $\widetilde{a_0 b_0} = \widetilde{ab}$.

Die Multiplikationstafel für das Beispiel $(\mathbb{Z}/3\mathbb{Z}, \cdot)$ sieht so aus:

$$\begin{array}{c|ccc} (\mathbb{Z}/3\mathbb{Z},\,\cdot) & \tilde{0} & \tilde{1} & \tilde{2} \\ \hline \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} \\ \tilde{1} & \tilde{0} & \tilde{1} & \tilde{2} \\ \tilde{2} & \tilde{0} & \tilde{2} & \tilde{1} \end{array}$$

Die Multiplikation $\cdot$ ist also eine Verknüpfung auf $\mathbb{Z}/n\mathbb{Z}$. Man prüft leicht nach, dass $\cdot$ assoziativ und kommutativ ist und das Einselement $\tilde{1}$ besitzt. Wegen der Kommutativität von $\cdot$ braucht man nur ein Distributivgesetz zu prüfen: Für alle $\tilde{a}, \tilde{b}, \tilde{c} \in \mathbb{Z}/n\mathbb{Z}$ gilt

$$\begin{aligned} \tilde{a} \cdot (\tilde{b} + \tilde{c}) &= \tilde{a} \cdot \widetilde{(b+c)} = \widetilde{a(b+c)} \\ &= \widetilde{ab+ac} = \widetilde{ab} + \widetilde{ac} \\ &= \tilde{a} \cdot \tilde{b} + \tilde{a} \cdot \tilde{c}. \end{aligned}$$

Also ist $(\mathbb{Z}/n\mathbb{Z},\,+,\,\cdot)$ ein kommutativer Ring mit Eins.

Definition 5.34 Ein Element $a \neq 0$ eines Rings R heißt (linker) **Nullteiler**, wenn es ein $b \in R$, $b \neq 0$ gibt mit $ab = 0$.

Beispiel 5.35 Im Ring $\mathbb{Z}/3\mathbb{Z}$ gibt es keine Nullteiler (vgl. obige Multiplikationstafel). Im Ring $\mathbb{Z}/6\mathbb{Z}$ hingegen ist z.B. $\tilde{2}$ ein linker Nullteiler, denn es ist $\tilde{2} \cdot \tilde{3} = \tilde{6} = \tilde{0}$ mit $\tilde{2} \neq \tilde{0}$ und $\tilde{3} \neq \tilde{0}$.

Definition 5.36 Sind $(R_1, +, \cdot)$ und $(R_2, +, \cdot)$ zwei Ringe, dann nennt man eine Abbildung $\Phi : R_1 \to R_2$ **(Ring-)Homomorphismus**, wenn für alle $x, y \in R_1$ gilt

$$\Phi(x + y) = \Phi(x) + \Phi(y) \quad \text{und} \quad \Phi(x \cdot y) = \Phi(x) \cdot \Phi(y).$$

Beispiel 5.37 Die kanonische Projektion $k : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $x \mapsto \tilde{x}$, die jedem Element von $\mathbb{Z}$ seine Äquivalenzklasse im Restklassenring $\mathbb{Z}/n\mathbb{Z}$ zuordnet, ist ein Ring-Homomorphismus. Das folgt unmittelbar aus der Wohldefiniertheit (also Repräsentanten-Unabhängigkeit) der Addition und Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$.

Im Gegensatz zu $(\mathbb{Z}\backslash\{0\}, \cdot)$ ist $(\mathbb{R}\backslash\{0\}, \cdot)$ eine (abelsche) Gruppe. Solche Ringe sind von besonderer Bedeutung in der linearen Algebra.

Definition 5.38 Ein Ring $(\mathbb{K}, +, \cdot)$, für den $(\mathbb{K}\backslash\{0\}, \cdot)$ eine abelsche Gruppe ist, heißt **Körper**.

Ein Körper ist also ein kommutativer Ring mit Eins, in dem jedes von Null verschiedene Element ein multiplikativ Inverses hat.

Ein Körper hat insbesondere stets ein Einselement $1 \neq 0$ und zu jedem $a \neq 0$ ein eindeutig bestimmtes Inverses a^{-1} bezüglich der Multiplikation. Jede Gleichung $a \cdot x = b$ ist für $a \neq 0$ durch $x = a^{-1} \cdot b = b \cdot a^{-1}$ eindeutig lösbar. Aus $u \cdot v = 0$ folgt also $u = 0$ oder $v = 0$; die Gleichung $u \cdot v = 0$ kann für $u \neq 0$ und $v \neq 0$ nicht gelten. Ein Körper ist also notwendigerweise „nullteilerfrei“.

Beispiel 5.39 $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ sind Körper, ebenso $(\mathbb{Z}/3\mathbb{Z}, +, \cdot)$. Hingegen ist der Ring $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$ kein Körper, denn er hat Nullteiler.

Bemerkung 5.40 Sie können nachprüfen, dass in den Abschnitten 3.2 und 3.3 nur die Körpereigenschaften der reellen Zahlen $(\mathbb{R}, +, \cdot)$ benutzt wurden. Die Begriffe und Ergebnisse aus diesen Abschnitten übertragen sich deshalb wörtlich auf lineare Gleichungssysteme *über beliebigen* Körpern $\mathbb{K}$. Deshalb gilt auch in diesem allgemeinen Kontext die Invarianz der Lösungsmenge unter Elementaroperationen und der Gaußsche Algorithmus.

Definition 5.41 Sind $(\mathbb{K}_1, +, \cdot)$ und $(\mathbb{K}_2, +, \cdot)$ zwei Körper, dann heißt eine Abbildung $\Phi : \mathbb{K}_1 \to \mathbb{K}_2$ ein **(Körper-)Homomorphismus**, wenn für alle $x, y \in \mathbb{K}_1$ gilt

$$\Phi(x + y) = \Phi(x) + \Phi(y) \quad \text{und} \quad \Phi(x \cdot y) = \Phi(x) \cdot \Phi(y).$$

Beispiel 5.42 Die Einbettung von $\mathbb{Q}$ in $\mathbb{R}$, $\Phi : \mathbb{Q} \to \mathbb{R}$, $x \mapsto x$ ist ein injektiver Körperhomomorphismus.

Definition 5.43 Ist $(\mathbb{K}, +, \cdot)$ ein Körper und gibt es eine natürliche Zahl m, sodass

$$\underbrace{1 + 1 + \cdots + 1}_{m \text{ mal}} = 0$$

gilt, so heißt die kleinste solche Zahl p die **Charakteristik** (char $\mathbb{K}$) von $\mathbb{K}$. Gibt es kein solches m, so hat $\mathbb{K}$ per Definition die Charakteristik 0.

Beispiel 5.44 In $(\mathbb{Z}/3\mathbb{Z}, +, \cdot)$ ist $\tilde{1}$ das Einselement, und es gilt $\tilde{1} + \tilde{1} + \tilde{1} = \tilde{0}$. $\mathbb{Z}/3\mathbb{Z}$ hat also die Charakteristik char $\mathbb{K} = 3$. Dagegen ist in $(\mathbb{Q}, +, \cdot)$ niemals $1 + 1 + \cdots + 1 = 0$. Es gilt also char $\mathbb{Q} = 0$.

Bemerkung 5.45 Ist die Charakteristik $p \neq 0$, so ist die p-fache Summe $a + a + \cdots + a = 0$ für *alle* $a \in \mathbb{K}$ und p ist eine Primzahl.

Beweis: Es ist $\underbrace{a + \cdots + a}_{p \text{ mal}} = a \cdot 1 + \cdots a \cdot 1 = a \cdot (\underbrace{1 + \cdots + 1}_{p \text{ mal}}) = a \cdot 0 = 0$. Wegen $1 \neq 0$ kann p nicht 1 sein in $\mathbb{K}$. Wenn $p > 1$ keine Primzahl wäre, so gäbe es eine

Darstellung $p = p_1 p_2$ mit natürlichen Zahlen p_1, p_2, die beide $< p$ sind. Wegen des in $\mathbb{K}$ geltenden Distributivgesetzes haben wir dann

$$\underbrace{1 + 1 + \cdots + 1}_{p_1 p_2 \text{ mal}} = \underbrace{(1 + \cdots + 1)}_{p_1 \text{ mal}} \cdot \underbrace{(1 + \cdots + 1)}_{p_2 \text{ mal}} = 0.$$

Da $\mathbb{K}$ als Körper nullteilerfrei ist, folgt also $\underbrace{1 + \cdots + 1}_{p_1 \text{ mal}} = 0$ oder $\underbrace{1 + \cdots + 1}_{p_2 \text{ mal}} = 0$, im Widerspruch zur Definition der Charakteristik als kleinste derartige Zahl. ■

Bemerkung 5.46 Man kann zeigen: Der Ring $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ist genau dann ein Körper, wenn p eine Primzahl ist. In diesem Fall ist $\operatorname{char} \mathbb{Z}/p\mathbb{Z} = p$. Zu jeder Primzahl p gibt es also einen Körper

$$\mathbb{Z}/p\mathbb{Z} = \{\tilde{0}, \tilde{1}, \ldots, \widetilde{p-1}\}$$

mit p Elementen. $\mathbb{Z}/p\mathbb{Z}$ heißt daher ein **endlicher Körper**. $\mathbb{Z}/2\mathbb{Z} = \{\tilde{0}, \tilde{1}\}$ ist der kleinste (endliche) Körper.

Man kann weiter zeigen, dass es zu jeder Primzahl p und jeder natürlichen Zahl k einen Körper $\mathbb{F}_{p^k}$ gibt mit p^k Elementen und $\operatorname{char} \mathbb{K} = p$.

Beispiel 5.47 (Ein Körper mit 4 Elementen) Auf dem cartesischen Produkt $\mathbb{F}_4 := \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ erklären wir zwei Verknüpfungen

$$\begin{aligned} x + y &= (\tilde{x}_1, \tilde{x}_2) + (\tilde{y}_1,\ \tilde{y}_2) = (\tilde{x}_1 + \tilde{y}_1,\ \tilde{x}_2 + \tilde{y}_2) \\ x \cdot y &= (\tilde{x}_1, \tilde{x}_2) \cdot (\tilde{y}_1, \tilde{y}_2) = (\tilde{x}_1 \cdot \tilde{y}_1 + \tilde{x}_2 \cdot \tilde{y}_2, \tilde{x}_1 \cdot \tilde{y}_2 + \tilde{x}_2 \cdot \tilde{y}_1 + \tilde{x}_2 \cdot \tilde{y}_2) \end{aligned}$$

mit $\tilde{x}_1, \tilde{x}_2, \tilde{y}_1, \tilde{y}_2 \in \mathbb{Z}/2\mathbb{Z}$.
Setzen wir noch $0 := (\tilde{0}, \tilde{0}), u := (\tilde{1}, \tilde{0}), v := (\tilde{0}, \tilde{1}), w := (\tilde{1}, \tilde{1})$, so erhalten wir die Verknüpfungstafeln

$$\begin{array}{c|cccc} + & 0 & u & v & w \\ \hline 0 & 0 & u & v & w \\ u & u & 0 & w & v \\ v & v & w & 0 & u \\ w & w & v & u & 0 \end{array} \quad \text{und} \quad \begin{array}{c|cccc} \cdot & 0 & u & v & w \\ \hline 0 & 0 & 0 & 0 & 0 \\ u & 0 & u & v & w \\ v & 0 & v & w & u \\ w & 0 & w & u & v \end{array}.$$

Hieraus ergibt sich, dass $(\mathbb{F}_4, +, \cdot)$ ein Körper ist mit 4 Elementen und $\operatorname{char} \mathbb{F}_4 = 2$. Das Nullelement in $\mathbb{F}_4$ ist 0 und das Einselement ist u. Die additive Gruppe ist die sogenannte Kleinsche Vierergruppe $\mathcal{V}_4$ und die multiplikative Gruppe $(\mathbb{F}_4 \backslash \{0\}, \cdot) = \{u, v, w\} = \{v, v^2, v^3\}$ ist die von v erzeugte zyklische Gruppe.

5.3.1 Beispiel: Der Körper $\mathbb{C}$ der komplexen Zahlen

Ausgehend vom Körper $\mathbb{R}$ betrachten wir das cartesische Produkt $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ aller geordneten Paare (a, b) reeller Zahlen und definieren für diese Menge zwei Verknüpfungen

$$\begin{array}{lrcl} \text{Addition:} & (a,b) + (a',b') & = & (a+a', b+b'), \\ \text{Multiplikation:} & (a,b) \cdot (a',b') & = & (aa' - bb', ab' + a'b). \end{array}$$

Mit diesen Verknüpfungen wird $\mathbb{C}$ zu einem Körper; seine Elemente heißen **komplexe Zahlen**.

Bemerkung 5.48 $(1, 0)$ ist das Einselement in $\mathbb{C}$ und $(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$ ist das multiplikative Inverse von $(a, b) \neq (0, 0)$.

Wir wollen jetzt die üblichen Schreibweise für komplexe Zahlen einführen und betrachten dazu die Abbildung

$$h : \mathbb{R} \to \mathbb{C}, \quad a \mapsto (a, 0).$$

Dann ist h ein injektiver Körperhomomorphismus, sodass wir $\mathbb{R}$ mit dem Teilkörper $h(\mathbb{R}) \subset \mathbb{C}$ identifizieren können. Das Element $a \in \mathbb{R}$ wird also mit $(a, 0) \in \mathbb{C}$ identifiziert. In diesem Sinne ist dann $\mathbb{R}$ in den Körper $\mathbb{C}$ „eingebettet“: $\mathbb{R} \subset \mathbb{C}$.

Schreiben wir i für die komplexe Zahl $(0, 1)$, so lässt sich jetzt die komplexe Zahl $z = (a, b)$ eindeutig in der Form $z = (a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1) \cdot (b, 0)$, also

$$z = a + \mathrm{i}b \quad \text{mit } a, b \in \mathbb{R} \tag{5.1}$$

schreiben. Man nennt a den **Realteil** ($a = \mathrm{Re}\ z$) und b den **Imaginärteil** ($b = \mathrm{Im}\ z$) der komplexen Zahl z. Weiter nennt man

$$\overline{z} = (a, -b) = a - \mathrm{i}b$$

die zu $z = a + \mathrm{i}b$ **konjugiert komplexe Zahl** und

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}$$

den (Absolut-)**Betrag** von z.

Die Addition bzw. Multiplikation in der neuen Schreibweise (5.1) lauten jetzt

$$\begin{array}{rcl} z_1 + z_2 & = & a_1 + \mathrm{i}b_1 + a_2 + \mathrm{i}b_2 = a_1 + a_2 + \mathrm{i}(b_1 + b_2) \\ z_1 z_2 & = & (a_1 + \mathrm{i}b_1)(a_2 + \mathrm{i}b_2) = a_1a_2 - b_1b_2 + \mathrm{i}(a_1b_2 + b_1a_2), \end{array}$$

Man rechnet also „wie gewohnt“ unter Berücksichtigung der Vorschrift $\mathrm{i}^2 = -1$.

5.4 Matrizen

Definition 5.49 Es seien $m, n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Eine **Matrix** über $\mathbb{K}$ mit m Zeilen und n Spalten, kurz eine $m \times n$-Matrix, ist ein rechteckiges Schema der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jk} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} & \cdots & a_{mn} \end{pmatrix}, \tag{5.2}$$

mit Einträgen $a_{jk} \in \mathbb{K}$ für $j = 1, \ldots, m$ und $k = 1, \ldots, n$. Man schreibt auch kurz

$$A = (a_{jk})$$

und nennt die a_{jk} die **Komponenten** der $m \times n$-Matrix A. Die Menge aller $m \times n$-Matrizen über $\mathbb{K}$ bezeichnen wir mit $\mathbb{K}^{m \times n}$.

5.4.1 Matrizen-Addition

Zwei $m \times n$ Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ kann man **komponentenweise addieren**:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix},$$

d.h. $(a_{jk}) + (b_{jk}) := (a_{jk} + b_{jk})$. Mit dieser Addition wird $\mathbb{K}^{m \times n}$ zu einer abelschen Gruppe. Das neutrale Element bezüglich der Addition ist die **Nullmatrix**

$$O = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

deren Komponenten alle Null sind. Das additive Inverse $-A$ von A ist

$$-A = \begin{pmatrix} -a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{m1} & \cdots & -a_{mn} \end{pmatrix}.$$

5.4.2 Matrizen-Multiplikation

Wir wollen nun zwei geeignete Matrizen A, B auch multiplizieren. Dabei müssen wir *voraussetzen*, dass

1. die Anzahl q der Spalten von A mit der Anzahl q der Zeilen von B übereinstimmt und
2. dass $A \in \mathbb{K}^{p\times q}$ und $B \in \mathbb{K}^{q\times r}$ ist, dass also A, B beides Matrizen über dem selben Körper $\mathbb{K}$ sind.

Definition 5.50 Es seien A eine $p \times q$-Matrix und B eine $q \times r$-Matrix über $\mathbb{K}$. Unter dem **(Matrizen-)Produkt** $C = AB$ verstehen wir dann die $p \times r$-Matrix $C = (c_{jk}) \in \mathbb{K}^{p\times r}$ mit

$$c_{jk} := a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jq}b_{qk} = \sum_{s=1}^{q} a_{js}b_{sk}\,; \quad j = 1, \ldots, p;\ k = 1, \ldots, r. \tag{5.3}$$

Die Komponente c_{jk} der Produktmatrix AB wird also gemäß (5.3) gebildet, indem man in A die j-te Zeile, in B die k-te Spalte auswählt, nacheinander die Produkte der an gleicher Stelle stehenden Zeilen- bzw. Spaltenelemente bildet und addiert:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ \multicolumn{4}{c}{\dotfill} \\ a_{j1} & a_{j2} & \cdots & a_{jq} \\ \multicolumn{4}{c}{\dotfill} \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{pmatrix} \begin{pmatrix} b_{11} & \vdots & b_{1k} & \vdots & b_{1r} \\ b_{21} & \vdots & b_{2k} & \vdots & b_{2r} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{q1} & \vdots & b_{qk} & \vdots & b_{qr} \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots\cdot & c_{jk} & \cdots \\ & \vdots & \end{pmatrix}.$$

Beispiel 5.51

1. $\begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & -2 & 1 \\ 5 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 9 & 8 & 8 \\ 22 & 12 & 10 & 12 \end{pmatrix}.$

2. $\begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ 22 \end{pmatrix}.$

3. $\begin{pmatrix} 1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \end{pmatrix}.$

4. $$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \\ 6 & 0 & -3 \\ 8 & 0 & -4 \end{pmatrix}.$$

Insbesondere lassen sich **quadratische Matrizen**, d.h. Matrizen, bei denen die Zeilen- und Spaltenzahl übereinstimmt, stets miteinander multiplizieren. Die Matrizen-Multiplikation ist also eine Verknüpfung auf der Menge $\mathbb{K}^{n\times n}$ der quadratischen $n \times n$-Matrizen. Das neutrale Element ist die **Einheitsmatrix**

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Die Matrizen-Multiplikation ist aber keine Verknüpfung auf der Menge $\mathbb{K}^{m\times n}$ mit $m \neq n$.

Satz 5.52 *Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Bezeichnet $+$ die komponentenweise Addition und $\cdot$ die Matrizen-Multiplikation, dann ist $(\mathbb{K}^{n\times n}, +, \cdot)$ ein Ring mit Eins.*

BEWEIS: Neben dem Beweis, dass $(\mathbb{K}^{n\times n}, +)$ eine abelsche Gruppe ist, bleibt zu zeigen, dass das Assoziativgesetz

$$\forall\, A, B, C \in \mathbb{K}^{n\times n} : (AB)C = A(BC)$$

und die beiden Distributivgesetze

$$\forall A, B, C \in \mathbb{K}^{n\times n} : A(B + C) = AB + AC \quad \text{und} \quad (A + B)C = AC + BC$$

gelten. Mit $A = (a_{jk})$, $B = (b_{jk})$, $C = (c_{jk})$ gilt für die Matrix $M = A(B + C) = (m_{il})$ nach Definition der Addition und Matrizen-Multiplikation

$$m_{il} = \sum_{s=1}^{n} a_{is}(b_{sl} + c_{sl}) = \sum_{s=1}^{n} a_{is}b_{sl} + \sum_{s=1}^{n} a_{is}c_{sl}\ ; \quad i,\, l = 1, \dots, n,$$

also $M = AB + AC$. Damit ist das 1. Distributivgesetz bewiesen, das 2. beweist man analog. ■

Bemerkung 5.53 Die Matrizen-Multiplikation ist im Allgemeinen *nicht kommutativ*! Zum Beispiel gilt:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

An diesem Beispiel sieht man auch, dass $\mathbb{K}^{n\times n}$ Nullteiler hat. Der Matrizenring $(\mathbb{K}^{n\times n}, +, \cdot)$ ist also im Allgemeinen weder kommutativ noch nullteilerfrei.

Bemerkung 5.54 Ein LGS (3.5) über dem Körper $\mathbb{K}$ lässt sich als Matrixgleichung schreiben: Sei dazu

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m \times n}$$

die Matrix des LGS, vgl. (3.6), und weiter

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^{n \times 1} \qquad \text{und} \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^{m \times 1}.$$

Dann lässt sich das LGS (3.5) nach Definition der Matrizen-Multiplikation schreiben als

$$A \cdot x = b.$$

Es gilt nämlich $\sum_{k=1}^{n} a_{ik} x_k = b_i$ für $i = 1, \ldots, m$.

5.4.3 Inverse Matrizen

Definition 5.55 Gibt es zu einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ über dem Körper $\mathbb{K}$ ein inverses Element bezüglich der Matrizen-Multiplikation, d.h. eine Matrix $A^{-1} \in \mathbb{K}^{n \times n}$ mit $AA^{-1} = A^{-1}A = E$, so heißt A **invertierbar** und A^{-1} ihre **Inverse** oder **inverse Matrix**.

Satz 5.56 *Die Menge* $\mathbf{GL}(n, \mathbb{K})$ *aller invertierbaren* $n \times n$*-Matrizen über dem Körper* $\mathbb{K}$ *ist bezüglich der Matrizen-Multiplikation eine Gruppe.*[1]

Beweis: Nach Satz 5.52 ist die Matrizen-Multiplikation assoziativ und hat als neutrales Element die Einheitsmatrix E. Nach Voraussetzung hat jede Matrix ein inverses Element. Es bleibt also nur noch zu zeigen, dass $\mathbf{GL}(n, \mathbb{K})$ bezüglich der Matrizen-Multiplikation abgeschlossen ist. Seien dazu $A, B \in \mathbf{GL}(n, \mathbb{K})$. Dann ist auch AB invertierbar, die Inverse von AB ist nämlich gerade $B^{-1}A^{-1}$ wegen

$$B^{-1}A^{-1}A\,B = B^{-1}EB = E \quad \text{und} \quad A\,B\,B^{-1}A^{-1} = A\,E\,A^{-1} = E.$$

■

[1] $\mathbf{GL}(n, \mathbb{K})$ steht für *general linear group*.

5.4.4 Wie berechnet man die inverse Matrix?

Die inverse Matrix einer gegebenen Matrix $A \in \mathbb{K}^{n\times n}$ lässt sich - falls sie existiert - mit dem Gaußschen Algorithmus berechnen. Die Inverse $A^{-1} = (x_{jk}) \in \mathbb{K}^{n\times n}$ existiert genau dann, wenn die Matrizengleichung $AA^{-1} = E$ lösbar ist. Da das inverse Element in einer Gruppe eindeutig bestimmt ist, ist A^{-1} dann auch eindeutig. Wir bezeichnen die k-te Spalte der gesuchten Matrix A^{-1} mit x_k, also

$$x_k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix} \in \mathbb{K}^{n\times 1}.$$

Die Matrizengleichung $A\cdot A^{-1} = E$ ist (nach Definition der Matrizen-Multiplikation) genau dann lösbar, wenn die n Gleichungssysteme

$$A\cdot x_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A\cdot x_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, A\cdot x_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

mit den zugehörigen erweiterten Matrizen

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 1 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right), \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right), \ldots, \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & \vdots \\ \vdots & \vdots & \ddots & \vdots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 1 \end{array}\right)$$

lösbar sind. Wenn wir auf diese n linearen Gleichungssysteme den Gaußschen Algorithmus anwenden, ergibt sich aus dem k-ten Gleichungssystem

$$\text{mit Matrix} \quad \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & 0 \\ \vdots & & \vdots & \vdots \\ \vdots & \ddots & \vdots & 1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 \end{array}\right) \quad \text{die Endgestalt} \quad \left(\begin{array}{ccc|c} 1 & \cdots & 0 & x_{1k} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & x_{nk} \end{array}\right),$$

aus deren letzter Spalte sich die Lösung x_k ablesen lässt. Da jedesmal die Matrix A vorkommt, wird das Verfahren zweckmäßigerweise so durchgeführt, dass man die Elementaroperationen für alle n Gleichungssysteme simultan vornimmt:

$$\left(\begin{array}{ccc|cccc} a_{11} & \cdots & a_{1n} & 1 & & & 0 \\ \vdots & & \vdots & 0 & \ddots & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & 0 \\ a_{n1} & \cdots & a_{nn} & 0 & & & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & \cdots & \cdots & 0 & x_{11} & \cdots & \cdots & x_{1n} \\ \vdots & \ddots & & \vdots & x_{21} & \ddots & & x_{2n} \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & x_{n1} & \cdots & \cdots & x_{nn} \end{array}\right).$$

Beispiel 5.57 Es sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 1 & 0 & 2 \end{pmatrix}$. Der Gaußsche Algorithmus liefert

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & -1 & 4 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -5 & -2 & -2 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array}\right) \mid -1$$

$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -4 & 0 & -2 & 0 & 3 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & 2 & 1 & 1 & 0 & -1 \end{array}\right) \mid -1 \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -4 & 11 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 2 & -5 \end{array}\right).$$

Also ist

$$A^{-1} = \begin{pmatrix} -2 & -4 & 11 \\ 0 & -1 & 2 \\ 1 & 2 & -5 \end{pmatrix}.$$

Bestätigen Sie durch direktes Nachrechnen, dass $A\,A^{-1} = A^{-1}A = E$ ist!

Bemerkung 5.58 Ist $A \in \mathbb{K}^{n\times n}$ eine invertierbare Matrix, so lässt sich das lineare Gleichungssystem $A \cdot x = b$ mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ eindeutig lösen. Die Lösung ist $x = A^{-1} \cdot b$.

5.4.5 Transponierte Matrizen

Aus einer gegebenen $m \times n$-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m\times n}$$

über $\mathbb{K}$ kann man eine $n \times m$-Matrix dadurch bilden, dass man die Zeilen (unter Beibehaltung der Reihenfolge) in die Spalten (und umgekehrt) schreibt. Man erhält so die **transponierte Matrix**

$$A^\top := \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ a_{12} & \cdots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{n\times m}.$$

Satz 5.59

1. *Für alle* $A, B \in \mathbb{K}^{m\times n}$ *gilt* $(A+B)^\top = A^\top + B^\top$.
2. *Für alle* $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{n\times q}$ *gilt* $(AB)^\top = B^\top A^\top$.
3. *Für alle* $A \in \mathbb{K}^{m\times n}$ *gilt* $(A^\top)^\top = A$.
4. *Für alle invertierbaren* $A \in \mathbb{K}^{n\times n}$ *gilt* $\left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top$.

Beweis: 1., 2. und 3. überprüft man durch direktes Nachrechnen. Zum Beweis von 4.: Aus $E^\top = E$ folgt zuerst wegen 2.

$$E = AA^{-1} = \left(AA^{-1}\right)^\top = \left(A^{-1}\right)^\top A^\top$$

und somit die Behauptung $\left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top$. ■

5.5 Polynome

Gegeben sei ein beliebiger Körper $\mathbb{K}$.

Definition 5.60 Ein **Polynom** ist eine formale Summe der Form

$$f = a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + a_nX^n, \quad a_i \in \mathbb{K}.$$

Formal bedeutet hier, dass die Unbestimmte X nur als Symbol aufzufassen ist, aber nicht ein konkretes Element aus $\mathbb{K}$ repräsentieren soll. Die Menge aller Polynome über $\mathbb{K}$ bezeichnen wir mit $\mathbb{K}[X]$. Der **Grad** des Polynoms f ist definiert als

$$\deg f := \begin{cases} n & \text{falls } a_n \neq 0 \text{ und } a_k = 0 \text{ für alle } k > n, \\ -\infty & \text{falls } a_k = 0 \text{ für alle } k \geq 0. \end{cases}$$

Auf $\mathbb{K}[X]$ können wir eine Addition koeffizientenweise definieren:

Für $f = a_0 + a_1X^1 + \cdots + a_nX^n$ und $g = b_0 + b_1X + \cdots + b_nX^n$ setzen wir

$$f + g := (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_{n-1} + b_{n-1})X^{n-1} + (a_n + b_n)X^n.$$

Wir nehmen hier ohne Einschränkung an, dass $m = n$ ist. Denn wir können z.B. im Fall $m < n$ die Koeffizienten $b_{m+1}, \ldots, b_n$ einfach gleich 0 wählen.

Die Multiplikation ist etwas komplizierter: wir setzen

$$f \cdot g = c_0 + c_1X + \cdots + c_{m+n-1}X^{m+n-1} + c_{m+n}X^{m+n},$$

wobei die Koeffizienten c_i gegeben sind durch

$$\begin{aligned} c_0 &:= a_0 b_0, \\ c_1 &:= a_1 b_0 + a_0 b_1, \\ c_2 &:= a_2 b_0 + a_1 b_1 + b_2 a_0, \\ &\vdots \\ c_{m+n} &:= a_n b_m, \end{aligned}$$

oder allgemein

$$c_k := \sum_{i=0}^{k} a_i b_{k-i}.$$

D.h wir erhalten das Produkt von f und g, indem wir beide Ausdrücke unter Verwendung des Distributivgesetzes multiplizieren und die Koeffizienten gleichen Grades sammeln.

Satz 5.61 $(\mathbb{K}[X], +, \cdot)$ *ist ein kommutativer Ring mit Eins.*

BEWEIS: Ein Polynom $f = \sum_i a_i X^i$ ist durch die endliche Folge $(a_i)_{i\in\mathbb{N}_0}$ seiner Koeffizienten vollständig bestimmt. Die Menge $\mathbb{K}[X]$ lässt sich also äquivalent definieren als die Menge aller Folgen $(a_i)_{i\in\mathbb{N}_0}$ mit $a_i \in \mathbb{K}$, in denen alle bis auf endliche viele a_i gleich 0 sind. Addition und Multiplikation sind dann wie oben über die Koeffizienten definiert.

Das Nullelement (also das neutrale Element bezüglich der Addition) ist das Nullpolynom $0 := (0, 0, 0, \ldots)$. Die Assoziativität von $+$ überträgt sich komponentenweise von $\mathbb{K}$ auf $\mathbb{K}[X]$. Zu $(a_0, a_1, a_2, a_3, \ldots)$ ist $(-a_0, -a_1, -a_2, -a_3, \ldots)$ das additive Inverse. Durch direktes Nachrechnen erhält man die Assoziativität der Multiplikation und die Distributivgesetze. Das Einselement ist $1 := (1, 0, 0, 0, \ldots)$, wie man leicht nachprüft. Die Kommutativität folgt so:

$$(a_i)\cdot(b_i) = \left(\sum_{k=0}^{i} a_k b_{i-k}\right)_{i\in\mathbb{N}_0} \overset{l:=i-k}{=} \left(\sum_{l=0}^{i} a_{i-l} b_l\right)_{i\in\mathbb{N}_0} = \left(\sum_{l=0}^{i} b_l a_{i-l}\right)_{i\in\mathbb{N}_0} = (b_i)\cdot(a_i).$$

■

Bemerkung 5.62 (a) Für $f, g \in \mathbb{K}[X]$ ist

$$\deg(fg) = \deg f + \deg g.$$

(b) Die Abbildung

$$\Phi : \mathbb{K} \to \mathbb{K}[X], \quad a \mapsto (a, 0, 0, \ldots)$$

ist ein Ring-Homomorphismus, d.h. es gilt für alle $a, b \in \mathbb{K}$:

$$\Phi(a+b) = \Phi(a) + \Phi(b), \quad \Phi(a\,b) = \Phi(a) \cdot \Phi(b) \quad \text{und} \quad \Phi(1) = 1.$$

Außerdem ist Φ injektiv. Man kann deshalb a mit $(a, 0, 0, \ldots)$ identifizieren und erhält die „Einbettung“ $\mathbb{K} \subset \mathbb{K}[X]$. Insbesondere kann man das Einselement in $\mathbb{K}[X]$ mit $1 \in \mathbb{K}$ identifizieren.

Bemerkung 5.63 Aus einem Polynom $p = a_0 + a_1 X + a_2 X^2 + \cdots$ mit $a_i \in \mathbb{R}$ oder $\mathbb{C}$ erhält man durch **Einsetzen** neue Objekte.

Ersetzt man z.B. die Unbestimmte X in einem (formalen) reellen Polynom p durch eine reelle Zahl, so erhält man eine **Polynom-Funktion** $p(t) := a_0 + a_1 t + a_2 t^2 + \cdots$ mit $t \in \mathbb{R}$, also eine Abbildung von $\mathbb{R}$ nach $\mathbb{R}$.

Insbesondere versteht man unter den **Nullstellen eines Polynoms** die Nullstellen der zugehörigen Polynom-Funktion, also alle t mit $p(t) = 0$.

Analog erhält man durch Ersetzen von X etwa durch eine reelle (2×2)-Matrix A eine Abbildung von $\mathbb{R}^{2\times 2}$ nach $\mathbb{R}^{2\times 2}$ gegeben durch $p(A) := a_0 E + a_1 A + a_2 A^2 + \cdots$ mit $A \in \mathbb{R}^{2\times 2}$.

5.6 *Kryptographie

Dieser Abschnitt über Kryptographie soll die Bedeutung von endlichen Körpern illustrieren. Er gehört nicht zum Prüfungsstoff.

Das Wort Kryptographie setzt sich aus den griechischen Worten „$\kappa\varrho\upsilon\pi\tau o\varsigma$(*kryptos*) = versteckt, geheim“ und „$\gamma\varrho\alpha\varphi\varepsilon\iota\nu$(grafein) = schreiben“ zusammen. Die Grundidee der Kryptographie ist es, gegebene Zeichen durch andere Zeichen zu ersetzen. Die Entschlüsselung muss dann diesen Vorgang wieder rückgängig machen.

Schon Cäsar soll schriftliche Befehle verschlüsselt haben. Er ersetzte dazu jeden Buchstaben durch den im Alphabet drei Positionen weiter hinten stehenden Buchstaben, also an Stelle von „a“ setzte er „d“, statt „b“ schrieb er „e“ usw. Wer das wusste, konnte diese Nachrichten dann wieder entschlüsseln.

Dieses einfache Verfahren bietet natürlich im Zeitalter moderner Computer keinen Schutz vor unberechtigtem Lesen der Nachricht. Man beschränkt sich heute auch nicht auf die 26 Zeichen des Alphabets, sondern fasst mehrere Zeichen zu einer Zeichenfolge zusammen und ordnet dieser eine Zahl a zu. Die Aufgabe der Kryptographie besteht darin, diese in eine Zahl $\mathrm{ch}(a)$ zu verschlüsseln - ein Vorgang, der durch die Dechiffrierung wieder rückgängig gemacht werden soll. An dieser Stelle kommt die Kongruenzrechnung modulo einer natürlichen Zahl n ins Spiel. Die entsprechenden Klassen haben einen Repräsentanten im Bereich $0, \dots, n-1$, die wir als geeignete Kandidaten für die Kryptographie kennenlernen werden.

Wir haben gesehen, dass $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ im Allgemeinen keine Gruppe ist, da nicht jedes Element ein Inverses besitzen muss. Zum Beispiel besitzt in $\mathbb{Z}/6\mathbb{Z}$ die Klasse $\tilde{3}$ mit Repräsentant 3 kein Inverses. Denn Multiplikation von 3 mit einer geraden Zahl g führt auf ein Vielfaches von 6, womit $\widetilde{g \cdot 3} = \tilde{0}$ gilt; Multiplikation von 3 mit einer ungeraden Zahl u führt auf $\widetilde{u \cdot 3} = \tilde{3}$, so dass es keine Zahl $z \in \mathbb{Z}$ gibt mit $\widetilde{z \cdot 3} = \tilde{1}$. Der Grund liegt darin, dass $\tilde{3}$ ein *Nullteiler* ($\underbrace{\tilde{2}}_{\neq \tilde{0}} \cdot \underbrace{\tilde{3}}_{\neq \tilde{0}} = \tilde{0}$) in $\mathbb{Z}/6\mathbb{Z}$ ist. Obwohl $\tilde{3} \neq \tilde{1}$ ist, gilt die Gleichung $\tilde{3} \cdot \tilde{3} = \tilde{3}$.

5.6.1 *Teilbarkeit

Um die Struktur von $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ besser verstehen zu können, beginnen wir mit folgenden Begriffsbildungen.

Definition 5.64 Seien $a, b \in \mathbb{Z}\backslash\{0\}$. Dann heißt b **Teiler** von a, wenn es eine ganze Zahl $n \in \mathbb{Z}$ gibt mit $a = nb$. Man nennt dann a durch b teilbar und schreibt $b \mid a$.

Der Begriff der Teilbarkeit lässt sich noch für andere Ringe außer $(\mathbb{Z}, +, \cdot)$ in natürlicher Weise einführen, etwa für den Ring der Polynome $\mathbb{K}[X]$ über einem Körper $\mathbb{K}$. Der im Folgenden vorgestellte *Euklidische Algorithmus* zur Bestimmung des größten gemeinsamen Teilers lässt sich für Polynome in analoger Weise durchführen.

Definition 5.65 Seien $a, b \in \mathbb{Z}\backslash\{0\}$. $g \in \mathbb{N}$ heißt **größter gemeinsamer Teiler** von a und b, geschrieben $\mathrm{ggT}(a, b)$, falls gilt:

(i) $g \mid a$ und $g \mid b$

(ii) g ist die größte Zahl mit dieser Eigenschaft.

Gilt $\mathrm{ggT}(a, b) = 1$, so heißen a und b **teilerfremd**.

Bemerkung 5.66 Berechnen lässt sich der größte gemeinsame Teiler $\mathrm{ggT}(a, b)$ für $|a| > |b|$ mit Hilfe des **Euklidischen Algorithmus**, den wir hier kurz vorstellen. Zunächst gibt es zu zwei ganzen Zahlen $a, b \in \mathbb{Z}\backslash\{0\}$ mit $|a| \geq |b|$ stets eine ganze Zahl k_0 und eine natürliche Zahl r_0 mit der folgendenden Eigenschaft (**Division mit Rest**):

$$a = k_0 \cdot b + r_0 \quad \text{mit} \quad 0 \leq r_0 < |b| \tag{$*$}$$

Gilt $r_0 = 0$, so ist offensichtlich $|b|$ ein Teiler von a, und damit gilt $\mathrm{ggT}(a, b) = |b|$.

Die grundlegende Idee ist es nun zu sehen, dass für $r_0 > 0$ auf Grund der Gleichung $(*)$ gilt

$$g := \mathrm{ggT}(a, b) = \mathrm{ggT}(b, r_0) =: g_0.$$

Denn es ist $g \leq g_0$, da g die Zahlen a und b ohne Rest teilt, also nach Gleichung $(*)$ auch b und r_0. Nimmt man nun an, dass $g_0 > g$ gilt, so ist wieder nach Gleichung $(*)$ g_0 ein Teiler von b und von a, der größer als $g = \mathrm{ggT}(a, b)$ wäre. Dies wäre ein Widerspruch zur Maximalität von g.

Wir können also an Stelle von $\mathrm{ggT}(a, b)$ den $\mathrm{ggT}(b, r_0)$ der betragskleineren Zahlen b und r_0 berechnen. Division mit Rest führt analog zu oben mit einer ganzen Zahl k_1 und einer natürlichen Zahl r_1 auf die Darstellung

$$b = k_1 \cdot r_0 + r_1 \qquad 0 \leq r_1 < r_0.$$

Gilt in dieser Darstellung $r_1 = 0$, so ist $\mathrm{ggT}(b, r_0) = r_0$. Im Fall $r_1 \neq 0$ ist $\mathrm{ggT}(b, r_0) = \mathrm{ggT}(r_0, r_1)$, wobei auch hier wieder $r_1 < r_0$ gilt.

Setzt man dieses Verfahren weiter fort, so erhält man eine Folge von natürlichen Zahlen r_i, die immer kleiner werden: $r_0 > r_1 > r_2 \cdots$. Da das Verfahren bei einer Zahl $r_0 \neq 0$ begonnen hat, muss irgendwann der Rest 0 auftreten. Es gibt also einen Index j mit der folgenden Eigenschaft:

$$\begin{aligned} r_{j-2} &= k_j \cdot r_{j-1} + r_j \quad , \quad r_j \neq 0 \\ r_{j-1} &= k_{j+1} \cdot r_j \end{aligned}$$

Analog zum oben Gesagten gilt dann $r_j = \mathrm{ggT}(r_{j-1}, r_j) = \mathrm{ggT}(r_{j-2}, r_{j-1})$. Nach dem Prinzip der vollständigen Induktion folgt damit

Hilfssatz 5.67 *Mit den obigen Notationen gilt* $\mathrm{ggT}(a, b) = r_j$.

Beispiel 5.68 Es gilt $\mathrm{ggT}(155, 9) = 1$, d.h. 155 und 9 sind teilerfremd.

$$\begin{aligned} 155 &= 17 \cdot 9 + 2 \\ 9 &= 4 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 \end{aligned}$$

Hilfssatz 5.69 (Lemma von Bézout) *Seien $a, b \in \mathbb{Z}$ und $g = ggT(a, b)$. Dann gibt es Zahlen $s, t \in \mathbb{Z}$ mit*

$$g = s \cdot a + t \cdot b.$$

Beweis: Setzen wir $r_0 := a$ und $r_1 := b$, so liefert der Euklidische Algorithmus eine Folge von Resten

$$r_{i+1} = r_{i-1} - q_i r_i, \quad i = 1, \ldots, n,$$

wobei nach dem n-ten Schritt der Rest $r_{n+1} = 0$ bleibt und $g = r_n$ der $\mathrm{ggT}(a, b)$ ist. Diese Gleichung lässt sich bequem durch Matrizen ausdrücken:

$$\begin{pmatrix} r_i \\ r_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \cdot \begin{pmatrix} r_{i-1} \\ r_i \end{pmatrix}.$$

Somit lässt sich der Euklidische Algorithmus durch eine Folge von Matrizen-Multiplikationen ausdrücken. Setzt man

$$Q_i := \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \qquad \text{und} \qquad S := Q_n Q_{n-1} \cdots Q_1,$$

so erhält man

$$\begin{pmatrix} g \\ 0 \end{pmatrix} = \begin{pmatrix} r_n \\ r_{n+1} \end{pmatrix} = Q_n Q_{n-1} \cdots Q_1 \cdot \begin{pmatrix} r_0 \\ r_1 \end{pmatrix} = S \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

Ist $S = \begin{pmatrix} s & t \\ u & v \end{pmatrix}$, so erhält man sofort die gesuchte Gleichung

$$\mathrm{ggT}(a, b) = g = s \cdot a + t \cdot b$$

aus der ersten Zeile von S. ∎

Bemerkung 5.70 Speziell für *teilerfremde* Zahlen $a, b \in \mathbb{Z}$ folgt daraus: Es gibt $s, t \in \mathbb{Z}$ mit $1 = s \cdot a + t \cdot b$.

Beispiel 5.71 Mit den in Beispiel 5.68 benutzten Zahlen gilt

$$1 = 9 - 4 \cdot 2 = 9 - 4 \cdot (155 - 17 \cdot 9) = 69 \cdot 9 - 4 \cdot 155.$$

5.6.2 *Die Einheitengruppe von $\mathbb{Z}/n\mathbb{Z}$

Nach diesen Vorarbeiten wenden wir uns wieder dem anfangs gestellten Problem zu.

Satz 5.72 *Die Menge der invertierbaren Elemente*

$$\mathbb{Z}/n\mathbb{Z}^* := \{\tilde{x} \in \mathbb{Z}/n\mathbb{Z} \mid \tilde{x} \text{ ist invertierbar}\}$$

ist bezüglich der Multiplikation $\cdot$ in $\mathbb{Z}/n\mathbb{Z}$ eine kommutative Gruppe.

Beweis: Zunächst ist $\mathbb{Z}/n\mathbb{Z}^* \neq \emptyset$, da das selbstinverse Element $\tilde{1}$ in $\mathbb{Z}/n\mathbb{Z}^*$ liegt. Weiter ist die Verknüpfung $\cdot$ auf $\mathbb{Z}/n\mathbb{Z}^*$ als Teilmenge von $\mathbb{Z}/n\mathbb{Z}$ assoziativ. Es bleibt zu zeigen, dass $\mathbb{Z}/n\mathbb{Z}^*$ abgeschlossen ist. Zunächst besteht $\mathbb{Z}/n\mathbb{Z}^*$ aus allen Elementen, die ein inverses Element haben. Damit gehört neben $\tilde{x} \in \mathbb{Z}/n\mathbb{Z}^*$ auch $\tilde{x}^{-1}$ zu $\mathbb{Z}/n\mathbb{Z}^*$, da deren Inverses wieder $\tilde{x}$ ist. Sind $\tilde{x}, \tilde{y} \in \mathbb{Z}/n\mathbb{Z}^*$, dann ist auch $\tilde{x} \cdot \tilde{y} \in \mathbb{Z}/n\mathbb{Z}^*$, da $\tilde{y}^{-1} \cdot \tilde{x}^{-1}$ Inverses dazu ist. Kommutativ ist die Gruppe, da das Verknüpfungsgebilde $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ kommutativ ist. ∎

Definition 5.73 $(\mathbb{Z}/n\mathbb{Z}^*, \cdot)$ heißt die **Einheitengruppe** von $(\mathbb{Z}/n\mathbb{Z}, \cdot)$. Die Elemente von $\mathbb{Z}/n\mathbb{Z}^*$ heißen **Einheiten** in $\mathbb{Z}/n\mathbb{Z}$.

Satz 5.74 *Es gilt*

$$\mathbb{Z}/n\mathbb{Z}^* = \{\tilde{x} \in \mathbb{Z}/n\mathbb{Z} \mid x \text{ und } n \text{ sind teilerfremd (d.h. } \mathrm{ggT}(x,n) = 1)\}.$$

Beweis:

„$\supset$“ $\mathrm{ggT}(x,n) = 1 \implies \exists s,t \in \mathbb{Z} : 1 = s \cdot x + t \cdot n \implies \exists \tilde{s},\tilde{t} \in \mathbb{Z}/n\mathbb{Z} : \tilde{1} = \tilde{s} \cdot \tilde{x} + \tilde{t} \cdot \tilde{0} \implies \tilde{s} = \tilde{x}^{-1}$, es gibt also ein multiplikatives Inverses $\tilde{s}$ von $\tilde{x}$. Damit ist $\tilde{x} \in \mathbb{Z}/n\mathbb{Z}^*$.

„$\subset$“ Indirekt: Wir betrachten oBdA die Repäsentanten x in $\{0, ..., n-1\}$. Annahme $g := \mathrm{ggT}(x,n) > 1 \implies x = g \cdot u$ und $n = g \cdot l$, wobei u und l teilerfremd sind. Für das Produkt dieser Zahlen gilt $x \cdot l = g \cdot u \cdot l = n \cdot u$, woraus $\tilde{x} \cdot \tilde{l} = \tilde{0}$ folgt. Wegen $g > 1$ ist $\tilde{l} \neq \tilde{0}$ und damit $\tilde{x}$ ein Nullteiler in $(\mathbb{Z}/n\mathbb{Z}, \cdot)$, der nicht invertierbar ist. ■

Beispiel 5.75 Es gilt $\mathbb{Z}/6\mathbb{Z}^* = \{\tilde{1}, \tilde{5}\}$ und $\mathbb{Z}/10\mathbb{Z}^* = \{\tilde{1}, \tilde{3}, \tilde{7}, \tilde{9}\}$.

Bemerkung 5.76 Bemerkung 5.70 kann ausgenutzt werden, um die Inverse einer Zahl modulo n zu bestimmen (vgl. Beweis von Satz 5.74). Nach Beispiel 5.68 gilt

$$1 = 69 \cdot 9 - 4 \cdot 155 \quad \text{bzw.} \quad \tilde{1} = \widetilde{69} \cdot \tilde{9} - \tilde{4} \cdot \underbrace{\widetilde{155}}_{=\tilde{0}} = \widetilde{69} \cdot \tilde{9},$$

woraus sich $\tilde{9}^{-1} = \widetilde{69} \in \mathbb{Z}/155\mathbb{Z}$ ergibt.

Definition 5.77 Die Funktion

$$\varphi : \mathbb{N} \to \mathbb{N} \text{ mit } \varphi(n) := \text{ Anzahl der Elemente von } \mathbb{Z}/n\mathbb{Z}^*$$

heißt **Eulersche φ-Funktion**.

Beispiel 5.78 Für eine Primzahl p gilt $\varphi(p) = p - 1$. Sind p, q verschiedene Primzahlen, so gilt für $n = p \cdot q$ gerade $\varphi(n) = (p-1)(q-1)$.

Wie dieses Beispiel zeigt, lässt sich $\varphi(n)$ für spezielles n leicht berechnen. Die Bedeutung dieser Zahlen zeigt der folgende Satz:

Satz 5.79 (Euler-Fermat) *Für alle Einheiten $\tilde{a} \in \mathbb{Z}/n\mathbb{Z}^*$ gilt $\tilde{a}^{\varphi(n)} = \tilde{1}$.*

Beweis: Die abelsche Gruppe $(\mathbb{Z}/n\mathbb{Z}^*, \cdot)$ besitze die $\varphi(n)$ verschiedenen Elemente $\tilde{x}_1, \ldots, \tilde{x}_{\varphi(n)}$. Dann sind für jedes $\tilde{a} \in \mathbb{Z}/n\mathbb{Z}^*$ die Elemente $\tilde{x}_1 \cdot \tilde{a}, \ldots, \tilde{x}_{\varphi(n)} \cdot \tilde{a}$ paarweise verschieden (Das ergibt sich leicht durch Multiplikation von rechts mit $\tilde{a}^{-1}$) und es gilt

$$\tilde{x}_1 \cdot \ldots \cdot \tilde{x}_{\varphi(n)} = \tilde{x}_1 \cdot \tilde{a} \ldots \cdot \tilde{x}_{\varphi(n)} \cdot \tilde{a} = \tilde{x}_1 \cdot \ldots \cdot \tilde{x}_{\varphi(n)} \cdot \tilde{a}^{\varphi(n)}.$$

Wegen Hilfssatz 5.12 folgt daraus $\tilde{a}^{\varphi(n)} = \tilde{1}$. ■

Bemerkung 5.80 Aus Satz 5.79 folgt für $k \in \mathbb{N}$

$$\tilde{a}^{k\varphi(n)} = (\tilde{a}^{\varphi(n)})^k = \tilde{1}^k = \tilde{1}.$$

Da $\tilde{a}^{\varphi(n)} = \widetilde{a^{\varphi(n)}}$ kann Satz 5.79 für alle diejenigen $a \in \mathbb{Z}$ mit $\tilde{a} \in \mathbb{Z}/n\mathbb{Z}^*$ umgeschrieben werden in die Form

$$a^{k\varphi(n)+1} \equiv a \mod n\,.$$

Man ist nun daran interessiert, diese Darstellung möglichst für *alle* Zahlen $a \in \mathbb{Z}$ zu bekommen. Dazu beschränken wir uns auf bestimmte Gruppen.

Satz 5.81 *Seien $p \neq q$ Primzahlen und $n = p \cdot q$. Dann gilt für alle $a \in \mathbb{Z}$*

$$a^{\varphi(n)+1} \equiv a \mod n\,.$$

Beweis: Nach Bemerkung 5.80 haben wir die Gleichung nur noch für Zahlen $a \in \{0, \ldots, n-1\}$ nachzuweisen, die nicht teilerfremd zu n sind, also p oder q als Teiler haben. Sind p *und* q Teiler von a, so gilt $\tilde{a} = \tilde{0} \in \mathbb{Z}/n\mathbb{Z}$ und es ist $\tilde{a}^{\varphi(n)+1} = \tilde{0} = \tilde{a}$.

Sei p Teiler von a und q kein Teiler von a. Dann gilt modulo q nach Bemerkung 5.70

$$a, q \text{ teilerfremd} \implies a^{p-1}, q \text{ teilerfremd} \implies (a^{p-1})^{q-1} = a^{\varphi(n)} \equiv 1 \mod q.$$

Multiplikation mit a ergibt $a^{\varphi(n)+1} \equiv a \mod q$.

Andererseits gilt modulo p, da a durch p teilbar ist, $a \equiv 0 \mod p$, also auch $a^{\varphi(n)+1} \equiv 0 \mod p$ und damit $a^{\varphi(n)+1} \equiv a \mod p$.

Damit ist $(a^{\varphi(n)+1} - a)$ sowohl durch p als auch durch q teilbar. Da die Primzahlen p und q verschieden waren, muss auch $p \cdot q = n$ die Zahl $(a^{\varphi(n)+1} - a)$ teilen, was eine Umformulierung der Behauptung ist. Der Fall, dass q Teiler von a und p kein Teiler ist, verläuft analog. ■

5.6.3 *Der RSA-Algorithmus

Auf der Darstellung aus Satz 5.81 beruht ein bekanntes Verfahren der Kryptographie. Die Grundidee ist folgende:

Potenziert man eine Zahl a mit dem Exponenten $k \cdot \varphi(n) + 1$, so erhält man modulo n wieder a zurück. Dieses Potenzieren zerlegt man in zwei Schritte, indem man die Zahl $k \cdot \varphi(n) + 1$ als Produkt zweier Zahlen e (*encryption=Verschlüsselung*) und d (*decryption=Entschlüsselung*) schreibt:

$$e \cdot d = k \cdot \varphi(n) + 1.$$

Potenziert man nun ein beliebiges $\tilde{a}$ mit dem Exponenten e, so ergibt sich $a^e \mod n$. Weiteres Potenzieren mit d führt auf

$$\widetilde{a^e}^d = \tilde{a}^{e \cdot d} = \tilde{a}^{k\varphi(n)+1} = \tilde{a}.$$

Damit kann das Potenzieren mit e als Verschlüsselung aufgefasst werden, das weitere Potenzieren mit d als Entschlüsselung.

Bemerkung 5.82 Damit die oben beschriebene Ver- und Entschlüsselung möglich ist, müssen sowohl e als auch d zu $\varphi(n)$ teilerfremd sein. Modulo $\varphi(n)$ gilt also $\tilde{d} = \tilde{e}^{-1}$.

Dieses Verfahren ist der sogenannte **RSA-Algorithmus** aus dem Jahre 1977, der nach seinen Entwicklern Rivest, Shamir und Adleman benannt ist:

1. Wähle verschiedene Primzahlen p und q und setze $n := p \cdot q$. Damit gilt $\varphi(n) = (p-1) \cdot (q-1)$.

2. Wähle e mit $\mathrm{ggT}(e, \varphi(n)) = 1$ als Chiffrierschlüssel. Das Zahlenpaar (n, e) heißt **öffentlicher Schlüssel**.

3. Berechne Dechiffrierschlüssel d mit $d = e^{-1} \mod \varphi(n)$. Das Zahlenpaar (n, d) heißt **privater Schlüssel**.

4. Chiffriere eine natürliche Zahl a mit $0 \leq a < n$ mit dem öffentlichen Schlüssel durch $\mathrm{ch}(a) := a^e \mod n$

5. Dechiffriert wird $\mathrm{ch}(a)$ mit dem privaten Schlüssel durch $a := \mathrm{ch}(a)^d \mod n$.

Wir veranschaulichen den Algorithmus an einem Beispiel. Um einen Text in natürliche Zahlen zu transformieren, verwenden wir der Einfachheit halber nur Großbuchstaben und ordnen jedem Buchstaben die Position im Alphabet zu. Damit gilt die Ersetzung A $\to$ 1, B $\to$ 2,. . . , Y $\to$ 25, Z $\to$ 26. Bei Bedarf können Leerzeichen und Interpunktionszeichen weitere Zahlen zugeordnet werden.

Beispiel 5.83

1. Wähle $p := 11$ und $q := 7$. Damit gilt $n = p \cdot q = 77$ und $\varphi(77) = (p-1)\cdot(q-1) = 10 \cdot 6 = 60$.

2. Wähle e teilerfremd zu $\varphi(77) = 60$, etwa $e := 17$.

3. Bestimme d mit $\tilde{d} = \tilde{e}^{-1} \mod 60$ gemäß Bemerkung 5.76. In diesem Zahlenbeispiel gilt $d = 53$, was man mit $17 \cdot 53 = 901 \equiv 1 \mod 60$ leicht verifiziert.

4. Zur Verschlüsselung mit (77,17) wählen wir das Wort KRYPTOGRAPHIE bzw. die Zahlenfolge
$$11\ 18\ 25\ 16\ 20\ 15\ 7\ 18\ 1\ 16\ 8\ 9\ 5.$$
Wegen $11^4 \equiv 11 \mod 77$ ist
$$\begin{aligned} \mathrm{ch}(11) &= 11^{17} \mod 77 \\ &= ((11^4 \mod 77)^4 \mod 77)(11 \mod 77) \\ &= (11^2 \mod 77) \\ &= 44 \mod 77 \end{aligned}$$
Auch ohne die Zusatzeigenschaft $11^4 \equiv 11 \mod 77$ ist die Verschlüsselung durch folgendes kleines Programm leicht möglich:
```
a := 11;
ch(a) := 1;
for j := 1 to e do ch(a) := ch(a) · a mod n;
```
Analoges Vorgehen für die restlichen Zahlen der Nachricht führt auf die verschlüsselte Nachricht:
$$44, 72, 9, 25, 48, 71, 28, 72, 1, 25, 57, 4, 3.$$

5. Dechiffriert wird mit dem privaten Schlüssel $(77, 53)$. Wegen $44^4 \equiv 44 \mod 77$ und daraus abgeleitet $44^{16} \equiv 44 \mod 77$ gestaltet sich für dieses Ergebnis die Dechiffrierung einfach. Es ist
$$\begin{aligned} 44^{53} \mod 77 &= ((44^{16} \mod 77)^3 \mod 77)(44^4 \mod 77)(44 \mod 77) \\ &= 44^5 \mod 77 \\ &= (44^4 \mod 77)(44 \mod 77) \\ &= 44^2 \mod 77 \\ &= 11 \mod 77 \end{aligned}$$
Analog zur Verschlüsselung liefert eine kleine Schleife mit dem privaten Schlüssel (diesesmal d statt e) und vertauschten Rollen von a und $\mathrm{ch}(a)$ die entschlüsselten Daten.

```
ch(a) := 44;
a := 1;
for j := 1 to d do a := a · ch(a) mod n;
```

Bemerkung 5.84 Allein aus dem Wissen des öffentlichen Schlüssels (e, n), lässt sich der private Schlüssel nicht bestimmen. Denn es geht bei dem Verfahren nicht darum, das inverse Element zu e modulo n zu bestimmen, sondern modulo $\varphi(n)$. Deshalb muss die Zahl $\varphi(n)$ bekannt sein. Diese kann man aber mit gängigen Methoden nur bestimmen, wenn die beiden Primzahlfaktoren von n bekannt sind. Das Knacken des Codes läuft mathematisch auf das Problem hinaus, eine Zahl n in ihre Primzahlen p und q zu faktorisieren. Bei sehr großen Primzahlen kann das auch mit modernsten Rechnern Monate dauern.

Teil III

Vektorräume

6 Definition und Beispiele

Wir betrachten das homogene lineare Gleichungssystem

$$\begin{aligned} x + 2y - 3z &= 0 \\ x - 2y - z &= 0 \end{aligned} \tag{6.1}$$

mit den Variablen $x, y, z \in \mathbb{R}$. Die Lösungsmenge $\mathcal{L}$ dieses LGS hat folgende Eigenschaft: sind $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$ zwei Lösungen von (6.1), so ist auch die Summe

$$(x_1 + x_2,\ y_1 + y_2,\ z_1 + z_2)$$

eine Lösung. Entsprechendes gilt für alle Vielfachen einer Lösung: ist $\lambda \in \mathbb{R}$ eine beliebige reelle Zahl und (x, y, z) eine Lösung von (6.1), dann ist auch $(\lambda x, \lambda y, \lambda z)$ eine Lösung von (6.1).

Solche Mengen V, bei denen mit zwei Elementen $v, w \in V$ auch deren „Summe" $v + w$ und alle ihre „Vielfachen" $\lambda \cdot v$ in V liegen, treten in der Mathematik sehr oft auf.

Im Folgenden werden wir diese „Struktur" präzisieren, indem wir den Begriff des Vektorraums in allgemeiner Form einführen und einige wichtige Beispiele kennenlernen. Vektorräume haben sich im Laufe des 19. und 20. Jahrhunderts als eine der wichtigsten mathematischen Strukturen herausgestellt. Sie spielen in praktisch jeder mathematischen Disziplin eine grundlegende Rolle und sind deshalb auch das zentrale Thema in dieser Vorlesung.

6.1 Was ist ein Vektorraum?

Jedem Vektorraum liegt ein gewisser Körper $\mathbb{K}$ zugrunde. Für viele Eigenschaften von Vektorräumen spielt jedoch die spezielle Wahl des Körpers $\mathbb{K}$ keine Rolle.

Definition 6.1 Es sei $\mathbb{K}$ ein Körper. Eine Menge V mit einer **Addition**

$$+ : V \times V \to V, \quad (x, y) \mapsto x + y$$

und einer **skalaren Multiplikation**, d.h. einer Abbildung

$$\cdot : \mathbb{K} \times V \to V, \quad (\lambda, x) \mapsto \lambda \cdot x,$$

heißt $\mathbb{K}$**-Vektorraum** oder ein Vektorraum über $\mathbb{K}$, falls

V1 $(V, +)$ eine abelsche Gruppe ist und

V2 für alle $\lambda, \mu \in \mathbb{K}$ und alle $x, y \in V$ gilt:

(a) $1 \cdot x = x$

(b) $\lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x$

(c) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$

(d) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$.

Die Elemente von V heißen **Vektoren**, die von $\mathbb{K}$ **Skalare**.

Das neutrale Element in $(V, +)$ wird **Nullvektor** genannt und (zumindest im allgemeinen Kontext) mit $0 = 0_V$ bezeichnet und ist vom Nullelement $0 = 0_{\mathbb{K}} \in \mathbb{K}$ zu unterscheiden!

Ist der Skalarkörper $\mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$, so spricht man von einem rationalen, reellen bzw. komplexen Vektorraum.

6.1.1 Erste Eigenschaften

Für die abelsche Gruppe $(V, +)$ eines Vektorraums gelten die für Gruppen hergeleiteten Eigenschaften. Insbesondere ist der Nullvektor eindeutig bestimmt, ebenso zu jedem Vektor v der inverse Vektor $-v$. Eine Gleichung $v + z = w$ hat bei gegebenen $v, w \in V$ genau eine Lösung $z = w + (-v)$, wofür wir wieder $w - v$ schreiben werden.

Satz 6.2 *In einem $\mathbb{K}$-Vektorraum V gilt für alle $\lambda \in \mathbb{K}$ und alle $v \in V$*

$$\text{(a)} \quad \lambda \cdot v = 0 \quad \Longleftrightarrow \quad \lambda = 0_{\mathbb{K}} \ \vee \ v = 0_V;$$

$$\text{(b)} \quad (-\lambda) \cdot v = -(\lambda \cdot v).$$

Beweis: (a):

„$\Leftarrow$“ Nach **V2** (c) ist $(1+0) \cdot v = 1 \cdot v + 0 \cdot v$, also wegen **V2** (a) $v = v + 0 \cdot v$. Nach der eben gemachten Bemerkung gilt somit $0 \cdot v = 0$ für alle $v \in V$. Weiter ist nach **V2** (d) $\lambda \cdot (v + 0) = \lambda \cdot v + \lambda \cdot 0$, also $\lambda \cdot v = \lambda \cdot v + \lambda \cdot 0$. Daraus folgt $\lambda \cdot 0 = 0$ für alle $\lambda \in \mathbb{K}$.

„$\Rightarrow$“ Sei $\lambda \cdot v = 0$ und $\lambda \neq 0$. Dann gilt wegen **V2** (a) und **V2** (b) $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0 = 0$.

(b): Nach Satz 6.2 ist $(\lambda + (-\lambda)) \cdot v = 0 \cdot v = 0$. Andererseits ist nach **V2** (c) $(\lambda + (-\lambda)) \cdot v = \lambda \cdot v + (-\lambda) \cdot v$. Also $0 = \lambda \cdot v + (-\lambda) \cdot v$. Da der inverse Vektor zu $\lambda \cdot v$ eindeutig bestimmt ist, folgt schließlich $-(\lambda \cdot v) = (-\lambda) \cdot v$. ∎

6.2 Beispiele

1. Gegeben sei ein Körper $\mathbb{K}$ und eine natürliche Zahl n. Dann ist die Menge $\mathbb{K}^n$ aller n-Tupel $(v_1, \ldots, v_n)$ mit $v_1, \ldots, v_n \in \mathbb{K}$ ein $\mathbb{K}$-Vektorraum mit der **komponentenweisen Addition**

$$+ : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n, \quad \big((v_1, \ldots, v_n), (w_1, \ldots, w_n)\big) \mapsto (v_1 + w_1, \ldots, v_n + w_n)$$

und der **komponentenweisen Skalarmultiplikation**

$$\cdot : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n, \quad \big(\lambda, (x_1, \ldots, x_n)\big) \mapsto (\lambda x_1, \ldots, \lambda x_n).$$

Man schreibt oft $x = (x_1, \ldots, x_n)$ und bezeichnet $x_1, \ldots, x_n$ als die **Komponenten** von x. Exemplarisch beweisen wir V3: für alle $\lambda, \mu \in \mathbb{K}$ und alle $x \in V$ gilt

$$\begin{aligned}(\lambda + \mu) \cdot x &= (\lambda + \mu) \cdot (x_1, \ldots, x_n) \\ &= \big((\lambda + \mu)x_1, \ldots, (\lambda + \mu)x_n\big) \\ &= (\lambda x_1 + \mu x_1, \ldots, \lambda x_n + \mu x_n) \\ &= (\lambda x_1, \ldots, \lambda x_n) + (\mu x_1, \ldots, \mu x_n) \\ &= \lambda \cdot x + \mu \cdot x.\end{aligned}$$

Man nennt $\mathbb{K}^n$ auch den **Standard-Vektorraum** über $\mathbb{K}$.

2. Statt n-Tupel $(x_1, \ldots, x_n) \in \mathbb{K}^n$ kann man auch unendliche Folgen $(x_i)_{i \in \mathbb{N}_0} = (x_0, x_1, x_2, \ldots)$ von Elementen aus einem Körper $\mathbb{K}$ betrachten. Die Menge $\mathbb{K}^{\mathbb{N}_0}$ der Folgen über $\mathbb{K}$ ist ebenfalls ein $\mathbb{K}$-Vektorraum mit der komponentenweisen Addition und Skalarmultiplikation

$$+ : \begin{cases} \mathbb{K}^{\mathbb{N}_0} \times \mathbb{K}^{\mathbb{N}_0} & \to & \mathbb{K}^{\mathbb{N}_0} \\ \big((x_i), (y_i)\big) & \mapsto & (x_i + y_i) \end{cases} \quad \text{bzw.} \quad \cdot : \begin{cases} \mathbb{K} \times \mathbb{K}^{\mathbb{N}_0} & \to & \mathbb{K}^{\mathbb{N}_0} \\ \big(\lambda, (x_i)\big) & \mapsto & (\lambda x_i). \end{cases}$$

Die Nachweise führt man analog zum Beispiel $\mathbb{K}^n$.

3. Wir können ein Polynom $p \in \mathbb{K}[X]$ mit einer Folge $(a_0, a_1, a_2, \ldots)$ von Körperelementen $a_i \in \mathbb{K}$ identifizieren, bei der nur endlich viele Elemente a_i von Null verschieden sind. Die Menge $\mathbb{K}[X]$ der Polynome ist eine Teilmenge des Vektorraums $\mathbb{K}^{\mathbb{N}_0}$, die mit derselben Addition und skalaren Multiplikation ein $\mathbb{K}$-Vektorraum ist, wie man leicht nachprüft.

Man beachte, dass zwar $\mathbb{K}[X] \subset \mathbb{K}^{\mathbb{N}_0}$ gilt, die beiden Vektorräume aber nicht übereinstimmen: jedes Element $(a_i)_{i \in \mathbb{N}_0}$ von $\mathbb{K}[X]$ hat *nur endlich viele* von Null verschiedene Elemente, ein Element $(x_i) \in \mathbb{K}^{\mathbb{N}_0}$ kann aber *beliebig viele* von Null verschiedene Elemente haben.

4. Sei M eine beliebige, nichtleere Menge. Dann bilden die Abbildungen $f : M \to \mathbb{K}$ einen $\mathbb{K}$-Vektorraum bezüglich der **punktweisen Addition**
$$(f+g)(x) := f(x) + g(x) \quad (x \in M) \quad \text{für alle Abbildungen } f, g : M \to \mathbb{K}$$
und der **punktweisen Skalarmultiplikation**
$$(\lambda \cdot f)(x) := \lambda \cdot f(x) \quad (x \in M) \quad \text{für alle } f : M \to \mathbb{K} \text{ und alle } \lambda \in \mathbb{K}.$$
In Analogie zu den Beispielen 1.-3. bezeichnet man die Menge aller Abbildungen $f : M \to \mathbb{K}$ auch mit $\mathbb{K}^M$. Für endliche Mengen M erhält man Beispiel 1 und für $M = \mathbb{N}_0$ erhält man Beispiel 2 als Spezialfall.
5. Die Menge $\mathbb{K}^{m \times n}$ der $m \times n$-Matrizen über $\mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum mit der Matrizenaddition und der skalaren Multiplikation
$$\lambda \cdot (a_{ij}) = (\lambda a_{ij}) \quad \text{für } \lambda \in \mathbb{K} \text{ und } a_{ij} \in \mathbb{K} \ (i = 1, \ldots, m;\ j = 1, \ldots, n).$$
6. Man kann $\mathbb{K}$ auch als $\mathbb{K}$-Vektorraum (über sich selbst) auffassen. Die Vektoraddition fällt dann mit der Addition in $\mathbb{K}$ zusammen und die skalare Multiplikation mit der Multiplikation in $\mathbb{K}$.
7. Man kann $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum auffassen: die Addition ist die Addition in $\mathbb{R}$, die skalare Multiplikation ist die Multiplikation einer rationalen mit einer reellen Zahl. Man beachte, dass dieser Vektorraum nicht mit dem aus Beispiel 6 übereinstimmt: dort kann man mit allen reellen Zahlen skalar multiplizieren, in diesem Beispiel aber nur mit rationalen!
8. Die Lösungsmenge $\mathcal{L}$ eines beliebigen homogenen LGS über dem Körper $\mathbb{K}$
$$\begin{array}{ccccccc} a_{11}x_1 & +a_{12}x_2 & + & \cdots & +a_{1n}x_n & = & 0 \\ a_{21}x_1 & +a_{22}x_2 & + & \cdots & +a_{2n}x_n & = & 0 \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & + & \cdots & +a_{mn}x_n & = & 0 \end{array}$$
mit Koeffizienten $a_{ik} \in \mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum bezüglich der komponentenweisen Addition und skalaren Multiplikation in $\mathbb{K}^n$.

Bemerkung 6.3 Sei V die Menge aller Tripel reeller Zahlen (v_1, v_2, v_3) mit der komponentenweisen Addition. Die skalare Multiplikation definieren wir durch
$$\lambda \cdot (v_1, v_2, v_3) = (\lambda v_1, \lambda^2 v_2, \lambda^3 v_3) \quad (\lambda, v_1, v_2, v_3 \in \mathbb{R}).$$
V ist *kein* reeller Vektorraum. Welche Eigenschaften eines Vektorraums sind erfüllt, welche nicht?

6.3 Linearkombinationen

Zur Notation: Skalare werden wir in der Regel mit griechischen Buchstaben bezeichnen. Das Zeichen $\cdot$ für die skalare Multiplikation werden wir meistens weglassen, z.B. nur λx anstatt $\lambda \cdot x$ schreiben.

Definition 6.4 Ein Vektor $x \in V$ eines $\mathbb{K}$-Vektorraums V mit

$$x = \sum_{i=1}^{p} \lambda_i v_i = \lambda_1 v_1 + \cdots + \lambda_p v_p \quad (\lambda_i \in \mathbb{K})$$

heißt **Linearkombination** der Vektoren $v_1, \ldots, v_p \in V$. Man sagt auch: x ist als Linearkombination der v_i **darstellbar**.

Beachten Sie, dass in einer Linearkombination stets nur *endlich* viele Summanden auftreten.

Beispiel 6.5

1. Das Polynom $p = (2, -1, 0, 4, 0, 0, 0, \ldots) \in \mathbb{R}[X]$ ist eine Linearkombination der Monome $p_0 = 1 = (1, 0, 0, \ldots)$, $p_1 = X = (0, 1, 0, 0, \ldots)$ und $p_3 = X^3 = (0, 0, 0, 1, 0, \ldots)$. Es gilt nämlich
$$p = 2 - X + 4X^3 = 2p_0 - 1p_1 + p_3.$$
2. Im Standardvektorraum $\mathbb{R}^2$ ist der Vektor $v = (1, 6)$ als Linearkombination von $v_1 = (1, 2)$, $v_2 = (0, 1)$ und $v_3 = (0, 2)$ darstellbar, denn es gilt $v = v_1 - 2v_2 + 3v_3$ oder auch $v = v_1 + 4v_2$. Es gibt auch noch weitere Möglichkeiten der Darstellung von v als Linearkombination von v_1, v_2, v_3.
3. Die Folge $\left(\frac{1}{(n+1)^2}\right) = \left(1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \ldots\right) \in \mathbb{Q}^{\mathbb{N}_0}$ ist *nicht als endliche* Linearkombination der Folgen $(1, 0, 0, \ldots), (0, 1, 0, 0, \ldots), (0, 0, 1, 0, 0, \ldots), \ldots$ darstellbar.

Den *Nullvektor* kann man immer als Linearkombination von beliebigen gegebenen anderen Vektoren $v_1, \ldots, v_k$ schreiben:

$$0 = \sum_{i=1}^{k} 0 v_i.$$

Weil hier alle Koeffizienten $\lambda_i = 0$ sind, spricht man von der **trivialen Darstellung des Nullvektors**. Dagegen ist es nicht immer möglich, den Nullvektor **nichttrivial** als Linearkombination der v_i darzustellen, d.h. in der Gestalt

$$0 = \sum_{i=1}^{k} \lambda_i v_i, \quad \lambda_i \in \mathbb{K}, \qquad \text{nicht alle } \lambda_i = 0.$$

Beispiel 6.6 Wir betrachten nochmals Beispiel 2 in 6.5. Für die Vektoren $v_1 = (1,2)$, $v_2 = (0,1)$, $v_3 = (0,2) \in \mathbb{R}^2$ gibt es neben der trivialen Darstellung des Nullvektors $0_{\mathbb{R}^2} = (0,0) = 0v_1 + 0v_2 + 0v_3$ auch die nichttriviale Darstellung

$$0_{\mathbb{R}^2} = 0v_1 + 2v_2 - v_3.$$

Dagegen gibt es, wenn man nur die zwei Vektoren $v_1, v_2 \in \mathbb{R}^2$ betrachtet, nur die triviale Darstellung des Nullvektors. Denn es ist

$$\lambda_1 v_1 + \lambda_2 v_2 = \lambda_1(1,2) + \lambda_2(0,1) = (\lambda_1, 2\lambda_1 + \lambda_2),$$

und dieser Vektor ist genau dann der Nullvektor $0_{\mathbb{R}^2} = (0,0)$, wenn $\lambda_1 = 0$ und $\lambda_2 = 0$.

Definition 6.7 Sei V ein $\mathbb{K}$-Vektorraum. *Endlich viele* Vektoren $v_1, \ldots, v_k \in V$ heißen **linear unabhängig**, wenn gilt

$$\sum_{i=1}^{k} \lambda_i v_i = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.$$

Die Vektoren $v_1, \ldots, v_k$ heißen **linear abhängig**, wenn sie nicht linear unabhängig sind, d.h. wenn es eine nichttriviale Linearkombination des Nullvektors aus $v_1, \ldots, v_k$ gibt.

Mit andern Worten: die Vektoren $v_1, \ldots, v_k \in V$ sind genau dann linear unabhängig, wenn man aus ihnen den Nullvektor nur als

$$0_V = 0v_1 + \ldots + 0v_k$$

linear kombinieren kann, und linear abhängig, wenn es noch (mindestens) eine weitere Möglichkeit gibt, den Nullvektor aus $v_1, \ldots, v_k$ linear zu kombinieren. Wir machen auch noch die zweckmäßige Definition: die leere Menge ist linear unabhängig.

Verallgemeinerung: Eine *unendliche Menge* von Vektoren $M \subset V$ heißt **linear unabhängig**, wenn alle endlichen Teilmengen von M linear unabhängig sind, und **linear abhängig**, wenn sie eine endliche, linear abhängige Menge enthält.

Beispiel 6.8

1. Die drei Vektoren v_1, v_2, v_3 aus Beispiel 2 in 6.5 sind linear abhängig, ebenso die zwei Vektoren v_2, v_3. Die zwei Vektoren v_1, v_2 sind dagegen linear unabhängig, ebenso v_1, v_3.

2. Es sei speziell $k = 2$; wir betrachten also zwei Vektoren x, y eines $\mathbb{K}$-Vektorraums V. Wenn x, y linear abhängig sind, so gibt es eine nichttriviale Darstellung

$\lambda x + \mu y = 0$ des Nullvektors ($\lambda, \mu \in \mathbb{K}$, nicht beide Null). Ist dann etwa $\lambda \neq 0$, so folgt $x = (-\lambda^{-1}\mu)y$. Ist $\mu \neq 0$, so folgt $y = (-\mu^{-1}\lambda)x$. Für zwei linear abhängige Vektoren x, y gilt also immer mindestens eine der Gleichungen

$$x = \alpha y \quad \text{oder} \quad y = \beta x \qquad \text{mit gewissen } \alpha, \beta \in \mathbb{K}.$$

Man nennt die Vektoren dann auch **proportional**.

Sind umgekehrt x, y proportional, ist also $x = \alpha y$ oder $y = \beta x$, so ist $1x - \alpha y = 0$ oder $1y - \beta x = 0$ wegen $1 \neq 0$ eine nichttriviale Darstellung von 0. Deswegen sind x, y linear abhängig.

3. Es sei $k = 1$; wir betrachten also jetzt nur einen einzigen Vektor v eines $\mathbb{K}$-Vektorraums V. Der Vektor v ist genau dann linear abhängig, wenn $v = 0$, und also genau dann linear unabhängig, wenn $v \neq 0$: Ist nämlich v linear abhängig und ist $\alpha v = 0$ mit $\alpha \neq 0$ eine nichttriviale Darstellung des Nullvektors, so folgt $v = \alpha^{-1}0 = 0$. Ist umgekehrt $v = 0$, so ist $1v = 1 \cdot 0 = 0$ eine nichttriviale Darstellung des Nullvektors, also der Vektor v linear abhängig.

4. Kommt unter den Vektoren $v_1, \ldots, v_k$ der Nullvektor vor, so sind sie linear abhängig. Denn ist z.B. $v_1 = 0$, so ist $1v_1 + \sum_{i=2}^{k} 0v_i = 0$ eine nichttriviale Darstellung von 0. Man kann auch zeigen, dass $v_1, \ldots, v_k$ linear abhängig sind, wenn zwei proportionale Vektoren vorkommen oder wenn ein Vektor eine Linearkombination der übrigen ist.

5. Im Vektorraum $\mathbb{R}[X]$ der Polynome über $\mathbb{R}$ sind $1 + X$ und $1 - X$ linear unabhängig (nach Beispiel 2). Ebenso sind die Monome $m_0 = 1$, $m_1 = X$, $m_2 = X^2, \ldots, m_k = X^k$ $(k \in \mathbb{N}^0)$ linear unabhängig. Denn die Linearkombination

$$\sum_{i=0}^{k} a_i X^i = a_0 + a_1 X + a_2 X^2 + \cdots + a_k X^k = (a_0, a_1, a_2, \ldots, a_k, 0, 0, \ldots)$$

ist nach Definition genau dann das Nullpolynom $(0, 0, \ldots)$, wenn alle $a_i = 0$ sind.

6. Die Menge der Monome $\{X^i \mid i \in \mathbb{N}_0\}$ im Vektorraum $\mathbb{K}[X]$ der Polynome ist linear unabhängig.

Das folgende Kriterium ist manchmal nützlich.

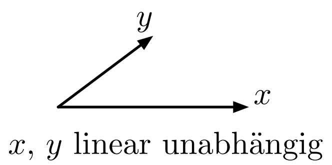
Satz 6.9 *Die Vektoren $v_1, \ldots, v_k$ $(k > 1)$ eines Vektorraums V sind genau dann linear abhängig, wenn es einen Vektor unter ihnen gibt, der sich als Linearkombination der übrigen darstellen lässt.*

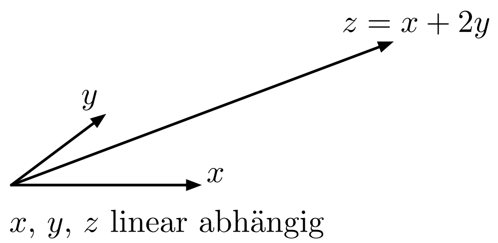
Beweis:

„⇒“ Seien $v_1, \dots, v_k$ linear abhängig und sei $\sum_{i=1}^k \lambda_i v_i = 0$ eine nichttriviale Darstellung von 0. Wenn etwa $\lambda_1 \neq 0$ ist, dann ist $v_1 = \sum_{j=2}^k (-\lambda_1^{-1}\lambda_j) v_j$ eine Linearkombination der übrigen Vektoren.

„⇐“ Sei etwa $v_1 = \sum_{j=2}^k \mu_j v_j$ eine Linearkombination von $v_2, v_3, \dots, v_k$. Dann ist $1v_1 - \sum_{j=2}^k \mu_j v_j = 0$ wegen $1 \neq 0$ eine nichttriviale Darstellung von 0, und $v_1, \dots, v_k$ sind linear abhängig. ∎

Die folgende Abbildung verdeutlicht die Aussage von Satz 6.9.





Hier ist $x + 2y - z = 0$, also lässt sich z durch x und y darstellen. Aber x und y sind linear unabhängig.

Wir beweisen einige weitere Sätze, die wir immer wieder brauchen werden.

Satz 6.10 *Sind $v_1, \dots, v_k, v_{k+1}, \dots, v_n$ $(n > k)$ Vektoren eines Vektorraums V, und sind $v_1, \dots, v_k$ linear abhängig, dann sind auch $v_1, \dots, v_n$ linear abhängig.*

Beweis: Ist $\sum_{i=1}^k \alpha_i v_i = 0$ eine nichttriviale Darstellung von 0, dann auch

$$\sum_{i=1}^k \alpha_i v_i + \sum_{i=k+1}^n 0 v_i = 0.$$

∎

Ein entsprechender Satz für $n < k$ gilt nicht.

Für linear unabhängige Vektoren gilt:

Satz 6.11 *Sind $v_1, \dots, v_k$ linear unabhängige Vektoren eines Vektorraums V, dann sind auch $v_1, \dots, v_m$ linear unabhängig für jedes $m \leq k$.*

Beweis: Wären $v_1, \ldots, v_m$ $(m < k)$ linear abhängig, dann wären auch $v_1, \ldots, v_k$ nach dem letzten Satz linear abhängig im Widerspruch zur Voraussetzung. ■

Auch hier gilt ein entsprechender Satz mit $m > k$ nicht.

Wir denken uns nun k beliebige (linear abhängige oder linear unabhängige) Vektoren $v_1, \ldots, v_k \in V$ gegeben und betrachten $k+1$ Vektoren $w_1, \ldots, w_{k+1}$, die sich als Linearkombinationen der v_i darstellen lassen. Für sie gilt folgender

Satz 6.12 *$k+1$ Linearkombinationen von k Vektoren eines Vektorraums V sind stets linear abhängig.*

Beweis: (Mit vollständiger Induktion)

INDUKTIONS-VERANKERUNG: Für $k = 1$ seien $w_1 = \lambda_1 v_1, w_2 = \lambda_2 v_1$ zwei Linearkombinationen von v_1. Ist $\lambda_1 = \lambda_2 = 0$, so sind $w_1 = 0 = w_2$ linear abhängig. Sind λ_1, λ_2 nicht beide Null, so ist $\lambda_2 w_1 - \lambda_1 w_2 = \lambda_2\lambda_1 v_1 - \lambda_1\lambda_2 v_1 = 0$ eine nichttriviale Darstellung des Nullvektors.

INDUKTIONS-SCHRITT: Die Aussage gelte bereits für k Linearkombinationen von $k-1$ Vektoren. Wir betrachten k Vektoren $v_1, \ldots, v_k$ und wollen zeigen, dass $k+1$ beliebige Linearkombinationen

$$\begin{array}{rcl} w_1 & = & a_{11}v_1 + \cdots + a_{1\ k-1}v_{k-1} + a_{1k}v_k \\ \vdots & & \vdots \quad \ldots \quad \ldots \quad \ldots \quad \vdots \\ w_k & = & a_{k1}v_1 + \cdots + a_{k\ k-1}v_{k-1} + a_{kk}v_k \\ w_{k+1} & = & a_{k+1,\ 1}v_1 + \cdots + a_{k+1,\ k-1}v_{k-1} + a_{k+1,\ k}v_k \end{array}$$

der k Vektoren $v_1, \ldots, v_k$ auch linear abhängig sind. Ohne Einschränkung können wir dabei annehmen, dass die w_i alle von Null verschieden sind (wieso?).

1. FALL: Die Koeffzienten $a_{1k}, \ldots, a_{kk}, a_{k+1\ k}$ vor v_k sind alle Null. In diesem Fall sind die k Vektoren $w_1, \cdots, w_k$ Linearkombinationen von $v_1, \ldots, v_{k-1}$ und somit nach Induktionsannahme linear abhängig. Nach Satz 6.10 sind dann auch die $k+1$ Vektoren $w_1, \cdots, w_k, w_{k+1}$ linear abhängig.

2. FALL: Die $a_{1k}, \ldots, a_{kk}, a_{k+1\,k}$ sind nicht alle Null; ohne Beschränkung der Allgemeinheit sei $a_{k+1\ k} \neq 0$. Dann sind die k Vektoren

$$\begin{array}{rcl} z_1 & := & w_1 - a_{k+1,\ k}^{-1} a_{1k} w_{k+1} \\ \vdots & & \vdots \\ z_k & := & w_k - a_{k+1,\ k}^{-1} a_{kk} w_{k+1} \end{array}$$

Linearkombinationen der $k-1$ Vektoren $v_1, \ldots, v_{k-1}$ und nach Induktionsannahme also linear abhängig. Wieder nach Satz 6.10 sind dann auch die $k+1$ Vektoren

$z_1, \cdots, z_k, w_{k+1}$ linear abhängig, d.h. es gibt $\lambda_1, \ldots, \lambda_{k+1} \in \mathbb{K}$, nicht alle Null und

$$\sum_{i=1}^{k} \lambda_i z_i + \lambda_{k+1} w_{k+1} = 0.$$

Dabei muss sogar mindestens eines der λ_i für ein $1 \leq i \leq k$ von Null verschieden sein. Wäre nämlich $\lambda_i = 0$ für alle $1 \leq i \leq k$, so auch λ_{k+1} (da $w_{k+1} \neq 0$): ein Widerspruch. Setzen wir jetzt die z_i ein, so erhalten wir

$$\sum_{i=1}^{k} \lambda_i w_i + (\lambda_{k+1} - \sum_{i=1}^{k} -\lambda_i a_{k+1\ k}^{-1} a_{ik}) w_{k+1} = 0.$$

Nach obiger Zwischenbemerkung ist dabei mindestens eines der λ_i für ein $1 \leq i \leq k$ von Null verschieden und somit sind die $k+1$ Vektoren $w_1, \ldots, w_k, w_{k+1}$ linear abhängig. ■

6.4 Lineare Hülle einer Teilmenge

Was erhält man, wenn man mit gegebenen Vektoren alle möglichen Linearkombinationen bildet?

Definition 6.13 Sei V ein $\mathbb{K}$-Vektorraum und $M \subset V$ eine beliebige Teilmenge. Die **lineare Hülle** oder der **Spann** $[M]$ von M ist für $M \neq \emptyset$ die Menge aller Linearkombinationen von Vektoren aus M. Für $M = \emptyset$ setzen wir $[M] = \{0\}$. Ist $M = \{v_1, \ldots, v_n\}$, so schreibt man auch $[v_1, \ldots, v_k]$ statt $[\{v_1, \ldots, v_k\}]$.

Beispiel 6.14

1. Die lineare Hülle der Vektoren $v_1 = (1, 2)$ und $v_2 = (0, 1)$ in $\mathbb{R}^2$ ist der gesamte $\mathbb{R}^2$, da sich jeder beliebige Vektor als Linearkombination von v_1 und v_2 darstellen lässt. Es gilt nämlich $(1, 0) = v_1 - 2v_2$ und aus $(1, 0)$ und $v_2 = (0, 1)$ lässt sich jeder beliebige Vektor linear kombinieren: der Vektor (a_1, a_2) lässt sich schreiben als $a_1(v_1 - 2v_2) + a_2 v_2 = a_1 v_1 + (a_2 - 2a_1)v_2$.
2. Die lineare Hülle $[X^0, X^1, X^2, X^3]$ ist gerade die Menge der Polynome vom Grad kleiner gleich 3. Die Menge *aller* Polynome $\mathbb{K}[X]$ ist gerade die lineare Hülle *aller* Monome $X^0, X^1, X^2, \ldots$.

Definition 6.15 Sei V ein $\mathbb{K}$-Vektorraum und $v_1, \ldots, v_k \in V$. Für diese erklären wir folgende **Elementar-Operationen**:

(I) Ersetzen eines Vektors v_i durch $\lambda\, v_i$ mit $\lambda \in \mathbb{K} \backslash \{0\}$.

(II) Ersetzen eines Vektors v_i durch $v_i + v_j$ mit $j \in \{1, \ldots, m\}$ und $j \neq i$.

Bemerkung 6.16

1. Die lineare Hülle $[v_1, \ldots, v_k]$ bleibt bei Elementar-Operationen ungeändert, d.h. es gilt

   $$[v_1, \ldots, v_i{+}v_j, \ldots, v_j, \ldots v_k] = [v_1, \ldots, v_i, \ldots, v_j, \ldots v_k] = [v_1, \ldots, \lambda v_i, \ldots, v_j, \ldots v_k].$$

   Wieso?

2. Eine Menge $\{v_1, \ldots, v_k\}$ von Vektoren bleibt linear unabhängig (bzw. linear abhängig), wenn man Elementar-Operationen auf $v_1, \ldots, v_k$ ausführt

3. Gegeben sei ein LGS mit m Zeilen, n Variablen und der erweiterten Matrix

   $$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

   Interpretiert man die Zeilen der Matrix als Vektoren $z_1, \ldots, z_m \in \mathbb{K}^{n+1}$, so lassen sich die Elementar-Operationen eines LGS (siehe Definition 3.3) interpretieren als Elementar-Operationen auf den Zeilenvektoren der Matrix des LGS. Wenn man den Gaußschen Algorithmus in „Matrixform" durchführt, führt man also geeignete Elementar-Operationen auf der Menge der Zeilenvektoren der Matrix aus.

4. Für die Teilmengen $M \subset V$ eines Vektorraums V und für ihre linearen Hüllen gelten die folgenden Eigenschaften:

   $$M \subset [M] \tag{6.2}$$
   $$M_1 \subset M_2 \implies [M_1] \subset [M_2]. \tag{6.3}$$

Bisher sind wir von einer Teilmenge $M \subset V$ ausgegangen und haben die lineare Hülle $[M]$ gebildet. Nun suchen wir umgekehrt eine Menge M, für die $[M] = V$ gilt. Dies gilt sicher für $M = V$. Gibt es auch *kleinere* Mengen M mit dieser Eigenschaft?

Definition 6.17 Gegeben sei ein $\mathbb{K}$-Vektorraum V. Eine Menge $M \subset V$ mit $[M] = V$ heißt **erzeugende Menge** oder **Erzeugendensystem** von V. Eine erzeugende Menge M von V heißt **minimal**, wenn es keine echte Teilmenge M' von M gibt, für die $[M'] = V$ gilt.

Beispiel 6.18 (Vgl. Beispiel 2 in 6.5). Die Menge $M = \{v_1, v_2, v_3\}$ mit $v_1 = (1, 2)$, $v_2 = (0, 1)$, $v_3 = (0, 2)$ ist eine erzeugende Menge des $\mathbb{R}^2$, denn jedes $v \in \mathbb{R}^2$ ist als Linearkombination von v_1, v_2, v_3 darstellbar. M ist nicht minimal, denn für die echten Teilmengen $M' = \{v_1, v_2\}$ und $M'' = \{v_1, v_3\}$ gilt ebenfalls $[M'] = [M''] = \mathbb{R}^2$. Die Mengen M' und M'' sind minimale erzeugende Mengen von $\mathbb{R}^2$.

Mit minimalen erzeugenden Mengen werden wir uns im folgenden Kapitel näher beschäftigen, wenn wir den Begriff der Basis einführen.

7 Basis und Dimension von Vektorräumen

7.1 Was ist eine Basis?

Definition 7.1 Eine Teilmenge B eines Vektorraumes V heißt **Basis** von V, wenn sie erzeugend und linear unabhängig ist.

Für die Definition von „linear unabhängig“ (insbesondere auch für unendliche Mengen von Vektoren) siehe Definition 6.7.

Beispiel 7.2

1. Für den Standard-Vektorraum $\mathbb{K}^n$ über dem Körper $\mathbb{K}$ bilden die Vektoren

$$\begin{aligned} e_1 &= (1,0,0,\dots,0) \\ e_2 &= (0,1,0,\dots,0) \\ &\vdots \\ e_n &= (0,0,\dots,0,1) \end{aligned}$$

   eine Basis. $B = \{e_1, \dots, e_n\}$ ist erzeugende Menge von $\mathbb{K}^n$, denn für jedes $v = (v_1, \dots, v_n) \in \mathbb{K}^n$ gilt

$$v = \sum_{i=1}^{n} v_i e_i.$$

   B ist auch linear unabhängig. Denn

$$0 = (0,\dots,0) = \sum_{i=1}^{n} \lambda_i e_i \iff \lambda_i = 0 \ \ \forall i.$$

   Man nennt diese Basis auch die **Standardbasis** des $\mathbb{K}^n$.

2. Eine weitere Basis des $\mathbb{K}^n$ ist $B = \{b_1, \dots, b_n\}$ mit

$$\begin{aligned} b_1 &= (1,0,0,0,\dots,0) \\ b_2 &= (1,1,0,0,\dots,0) \\ b_3 &= (1,1,1,0,\dots,0) \\ &\vdots \\ b_n &= (1,1,1,1,\dots,1). \end{aligned}$$

   B ist erzeugend, da man die Standardbasisvektoren $e_1, \dots, e_n$ alle durch die b_i linear kombinieren kann: es gilt $e_1 = b_1$, $e_2 = b_2 - b_1$, $\dots$, $e_n = b_n - b_{n-1}$; somit kann man auch alle Vektoren in $\mathbb{K}^n$ aus Vektoren in B linear kombinieren. B ist auch linear unabhängig (wieso?).

3. Analog zum ersten Beispiel zeigt man, dass im Vektorraum $\mathbb{K}[X]$ aller Polynome über $\mathbb{K}$ die Menge aller Monome $B = \{X^i \mid i \in \mathbb{N}_0\}$ eine Basis ist. Hier ist also die Basis (abzählbar) unendlich.

4. Der Nullraum $\{0\}$ hat die Basis $B = \emptyset$. Denn wir hatten definiert, dass die leere Menge linear unabhängig ist und dass $[\emptyset] = \{0\}$.

In den nächsten zwei Sätzen geben noch weitere Charakterisierungen einer Basis.

Satz 7.3 (Basis = erzeugend + minimal) *Eine Teilmenge B eines $\mathbb{K}$-Vektorraumes V ist eine Basis genau dann, wenn B ein minimales Erzeugendensystem ist.*

Beweis: „$\Longrightarrow$“: Sei B eine Basis. Nach Definition 7.1 ist B erzeugend und linear unabhängig. Wir müssen zeigen, dass B minimal ist. Annahme: B ist nicht minimal. Dann gibt es eine echte Teilmenge B' von B mit $[B'] = V$. Es gibt also einen Vektor $v \neq 0$ mit $v \in B$, $v \notin B'$, der sich wegen $v \in V = [B']$ als Linearkombination

$$v = \sum_{i=1}^{m} \alpha_i b_i' \quad \text{mit gewissen } b_i' \in B' \text{ und } \alpha_i \in \mathbb{K}$$

darstellen lässt. Nach Satz 6.9 sind dann $v, b_1', \dots, b_m'$ linear abhängig. Damit ist aber auch die Menge B linear abhängig. Ein Widerspruch zur Voraussetzung.

„$\Longleftarrow$“: Sei nun umgekehrt B ein minimales Erzeugendensystem von V. Wir müssen zeigen, dass B linear unabhängig ist. Auch hier argumentieren wir indirekt. Annahme: B ist linear abhängig. Nach Definition 6.7 gibt es in B eine endliche Teilmenge $\{b_1, \dots, b_p\}$ von linear abhängigen Vektoren. Nach Satz 6.9 ist dann einer dieser Vektoren, etwa b_k, eine Linearkombination der übrigen. Jeder Vektor $v \in V = [B]$ lässt sich also bereits aus Vektoren aus $B \backslash \{b_k\}$ linear kombinieren, d.h. es gilt $V = [B \backslash \{b_k\}]$. Die Menge B ist also nicht minimal im Widerspruch zur Voraussetzung. ■

Mit ähnlichen Argumenten beweist man:

Satz 7.4 (Basis = linear unabhängig + maximal) *Eine Teilmenge B eines $\mathbb{K}$-Vektorraumes V ist eine Basis genau dann, wenn B maximal linear unabhängig ist.*

Für den Standard-Vektorraum $\mathbb{K}^n$ und für den Raum $\mathbb{K}[X]$ der Polynome konnten wir in Beispiel 7.2 Basen angeben. Hat jeder Vektorraum eine Basis? Die Antwort ist „ja!“ (siehe Satz 7.8) und die Idee ist einfach: Nach den vorhergehenden Sätzen muss man ein Erzeugendensystem zu einer linear unabhängigen Teilmenge verkleinern oder eine linear unabhängige Teilmenge zu einer Basis ergänzen. Dass das geht, besagt der

Satz 7.5 (Basisergänzungssatz) *Es sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum, $V \neq \{0\}$. Weiter sei $E \subset V$ ein endliches Erzeugendensystem von V und $L \subset E$ eine linear unabhängige Menge, $L \neq \emptyset$. Dann gibt es eine Basis B von V mit $L \subset B \subset E$.*

Beweis: Ist L auch ein Erzeugendensystem, so ist $B := L$ eine Basis nach Definition. Andernfalls gibt es einen Vektor $v \in E$, der nicht in der linearen Hülle von L liegt. Wir ergänzen dann L zu der ebenfalls linear unabhängigen Menge $L' = L \cup \{v\}$, und verfahren analog mit L'. Da E endlich ist, muss L um höchstens endlich viele Elemente ergänzt werden, um ein Erzeugendensystem und damit eine Basis zu erhalten. ■

Folgerung 7.6 *Ist V ein endlich erzeugter Vektorraum und $V \neq \{0\}$, so hat V eine endliche Basis.*

Beweis: Nach Voraussetzung gibt es in V ein endliches Erzeugendensystem E und einen Vektor $E \ni v \neq 0$. Die Teilmenge $L := \{v\} \subset E$ ist dann linear unabhängig. Nach Satz 7.5 existiert eine Basis B mit $L \subset B \subset E$. Da E endlich ist, ist auch B endlich. ■

Bemerkung 7.7 (E nicht endlich) Der Basisergänzungssatz gilt auch, wenn E nicht als endlich vorausgesetzt wird. Der Beweis dieses allgemeinen Falles ist aber nicht elementar. Das Problem dabei ist grob gesagt, dass Vektorräume „sehr groß“ sein können und dass eine Basis auch überabzählbar viele Elemente haben kann. Man braucht deshalb das Auswahlaxiom (bzw. das „Lemma von Zorn“). Einzelheiten dazu findet man z.B. in Abschnitt II.6 des Buches [2] von Brieskorn.

Satz 7.8 (Eine Basis existiert immer) *Jeder Vektorraum V hat eine Basis.*

Beweis: Ist $V = \{0\}$, so ist $B = \emptyset$ eine Basis von V. Ist $V \neq \{0\}$, so gibt es einen Vektor $v \neq 0$ in V. Setze $L := \{v\}$, $E := V$. Nach Satz 7.5 und Bemerkung 7.7 gibt es dann eine Basis B von V (die v enthält). ■

Bemerkung 7.9 Die nach dem Basisergänzungssatz mögliche Ergänzung zu einer Basis ist nicht eindeutig bestimmt. Zum Beispiel lassen sich die linear unabhängigen Vektoren $b_1 = (1,0,0,1)$, $b_2 = (0,1,0,0)$ des $\mathbb{R}^4$ durch $b_3 = (0,0,1,0)$, $b_4 = (0,0,0,1)$, aber auch durch $b_3' = (0,1,1,1)$, $b_4' = (1,1,1,1)$ zu einer Basis des $\mathbb{R}^4$ ergänzen. Der Beweis des Basisergänzungssatzes liefert zwar kein praktisches Verfahren zur Basisergänzung, er ist aber ein nützliches theoretisches Hilfsmittel.

7.2 Dimension

Wir betrachten nun die Anzahl der Basisvektoren genauer. In Beispiel 7.2 haben wir zwei Basen für den Standard-Vektorraum $\mathbb{K}^n$ angegeben, die beide gleich viele Elemente haben, nämlich n. Kann man auch eine Basis von $\mathbb{K}^n$ finden, die mehr oder weniger Elemente hat? Dass dies nicht möglich ist, zeigt ganz allgemein der folgende

Satz 7.10 (Anzahl Basiselemente) *Hat ein Vektorraum V eine endliche Basis B mit $n \in \mathbb{N}$ Elementen, so hat jede Basis B' von V ebenfalls n Elemente.*

Beweis: Sei $B = \{b_1, \ldots, b_n\}$ und sei $B' = \{b'_1, \ldots, b'_m\}$ eine weitere Basis von V. Nach Satz 6.12 kann dann B' höchstens n Elemente haben, denn je $n+1$ Elemente von B' wären als Linearkombination der b_i linear abhängig im Widerspruch zur linearen Unabhängigkeit von B'. Also $m \leq n$. Entsprechend schließt man, dass umgekehrt B höchstens so viele Elemente wie B' hat. Also $n \leq m$. Somit ist $m = n$ und hat B' hat ebenfalls n Elemente. ■

Definition 7.11 Ein Vektorraum mit einer endlichen Basis heißt **endlich dimensional**. Die für alle Basen von V übereinstimmende Anzahl $n \in \mathbb{N}$ der Elemente heißt **Dimension** von V. Wir schreiben dann $\dim V = n$. Ein Vektorraum, der keine endliche Basis hat, heißt **unendlich dimensional**.

Beispiel 7.12 Nach Beispiel 7.2 ist $\dim \mathbb{K}^n = n$ und $\dim \mathbb{K}[X] = \infty$. Der Vektorraum aller Polynome vom Grad kleiner gleich g hat die Dimension $g + 1$.

Der nächste Satz ergänzt Satz 7.10.

Satz 7.13 *Für einen n-dimensionalen Vektorraum V gilt:*

a) $n+1$ *Vektoren aus V sind immer linear abhängig.*

b) n *linear unabhängige Vektoren aus V bilden immer eine Basis von V.*

Beweis: a): Nach Voraussetzung gibt es eine Basis $B = \{b_1, \ldots, b_n\}$ von V. Jeder Vektor $v \in V$ ist wegen $[B] = V$ als Linearkombination der b_i darstellbar. Nach Satz 6.12 sind daher je $n+1$ Vektoren aus V linear abhängig.

b): Seien $b'_1, \ldots, b'_n \in V$ linear unabhängig und sei v ein beliebiger Vektor aus V. Nach a) sind die $n+1$ Vektoren $b'_1, \ldots, b'_n, v$ linear abhängig, also gibt es eine nichttriviale Darstellung des Nullvektors:

$$\sum_{i=1}^{n} \lambda_i b'_i + \lambda v = 0 \quad (\lambda_i, \lambda \in \mathbb{K}).$$

Ist $\lambda = 0$, so folgt aus der linearen Unabhängigkeit von $b'_1, \ldots, b'_n$ auch $\lambda_i = 0$ für alle i, im Widerspruch zur Voraussetzung. Also ist $\lambda \neq 0$, und v lässt sich als Linearkombination der b'_i darstellen:

$$v = -\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} b'_i.$$

Somit ist die linear unabhängige Menge $\{b'_1, \ldots, b'_n\}$ eine erzeugende Menge von V, also eine Basis von V. ■

Nach dem letzten Satz bilden n linear unabhängige Vektoren eines n-dimensionalen Vektorraumes stets eine Basis. Hat man weniger als n, etwa p linear unabhängige Vektoren $(0 < p < n)$, so lassen sich diese stets zu einer Basis von V ergänzen nach dem Basisergänzungssatz 7.5.

7.3 Basisdarstellung und Basiswechsel

Definition 7.14 Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Nach der Definition einer Basis als linear unabhängiges Erzeugendensystem hat jeder Vektor $v \in V$ eine **Basisdarstellung**

$$v = \sum_{i=1}^{n} v_i b_i. \tag{7.1}$$

Satz 7.15 (Eindeutige Basisdarstellung) *Sei $B = \{b_1, \ldots, b_n\}$ eine Basis eines n-dimensionalen $\mathbb{K}$-Vektorraumes V. Dann ist die die Darstellung 7.1 eines Vektors $v \in V$ bezüglich der Basis B eindeutig.*

Beweis: Wegen $[b_1, \ldots, b_n] = V$ lässt sich jeder Vektor $v \in V$ als Linearkombination mit geeigneten Koeffizienten $v_i \in \mathbb{K}$ darstellen. Zu zeigen bleibt die *Eindeutigkeit* der Darstellung: Sind

$$v = \sum_{i=1}^{n} v_i b_i \quad \text{und} \quad v = \sum_{i=1}^{n} v'_i b_i$$

zwei Basisdarstellungen von v, so folgt durch Subtraktion

$$0 = \sum_{i=1}^{n} (v_i - v'_i) b_i$$

und daraus wegen der linearen Unabhängigkeit der b_i, dass $v_i = v'_i$ für alle $i \in \{1, 2, \ldots, n\}$ gilt. Die Darstellung ist also eindeutig. ■

Definition 7.16 Sei $B = \{b_1, \ldots, b_n\}$ eine Basis eines $\mathbb{K}$-Vektorraums V und $v \in V$ ein beliebiger Vektor mit der eindeutigen Basisdarstellung $v = \sum_{i=1}^n v_i b_i$. Die Koeffizienten $v_1, \ldots, v_n \in \mathbb{K}$ heißen **Komponenten** von v in der Basis B. Der **Komponentenvektor** von v bezüglich B ist das n-Tupel $\Theta_B(v) := (v_1, \ldots, v_n) \in \mathbb{K}^n$.

Bemerkung 7.17 Aus technischen Gründen werden wir die Komponentenvektoren $\Theta_B(v) = (v_1, \ldots, v_n) \in \mathbb{K}^n$ im Folgenden oft mit $n \times 1$-Matrizen identifizieren, d.h. die äquivalente Schreibweise

$$\Theta_B(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

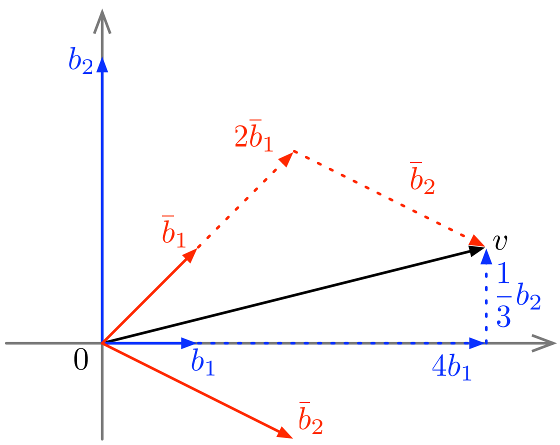
verwenden. Genauso werden wir auch beliebige Elemente von $\mathbb{K}^n$ oft als $n \times 1$-Matrizen auffassen und sie **Spaltenvektoren** nennen.

Beispiel 7.18 (**Verschiedene Basen**) In $\mathbb{R}^2$ seien die beiden Basen $B = \{b_1, b_2\}$ und $\overline{B} = \{\overline{b}_1, \overline{b}_2\}$ mit

$$b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad \overline{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \overline{b}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

gegeben. Weiter sei $v = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$. Dann gilt $v = 4b_1 + \frac{1}{3}b_2 = 2\overline{b}_1 + \overline{b}_2$ und somit

$$\Theta_B(v) = \begin{pmatrix} 4 \\ \frac{1}{3} \end{pmatrix} \quad \text{und} \quad \Theta_{\overline{B}}(v) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$



Wie kommt man nun von der Darstellung eines *beliebigen* Vektors v bzgl. der „alten“ Basis B zur Darstellung von v bzgl. der „neuen“ Basis $\overline{B}$? Um diese Frage zu beantworten, nutzen wir aus, dass die Vektoren b_1 und b_2 bezüglich der Basis $\overline{B}$ darstellbar sind. Es ist

$$b_1 = \frac{1}{3}\overline{b}_1 + \frac{1}{3}\overline{b}_2, \qquad b_2 = 2\overline{b}_1 - \overline{b}_2.$$

Wir können dies in die Darstellung $v = \lambda b_1 + \mu b_2$, also $\Theta_B(v) = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$, einsetzen und erhalten

$$v = \lambda(\frac{1}{3}\overline{b}_1 + \frac{1}{3}\overline{b}_2) + \mu(2\overline{b}_1 - \overline{b}_2) = (\frac{1}{3}\lambda + 2\mu)\overline{b}_1 + (\frac{1}{3}\lambda - \mu)\overline{b}_2,$$

also $\Theta_{\overline{B}}(v) = \begin{pmatrix} \frac{1}{3}\lambda + 2\mu \\ \frac{1}{3}\lambda - \mu \end{pmatrix}$. Dies können wir durch folgende Matrixgleichung ausdrücken:

$$\Theta_{\overline{B}}(v) = \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & -1 \end{pmatrix} \cdot \Theta_B(v)$$

Der Basiswechsel von B nach $\overline{B}$ erfolgt also gerade durch die Matrix, deren Spalten die Komponentenvektoren der „alten“ Basisvektoren b_1, b_2 bezüglich der „neuen“ Basis $\overline{B}$ sind. Aus der Gleichung ist auch sofort ersichtlich, dass man den umgekehrten Basiswechsel, also von der Darstellung $\Theta_{\overline{B}}(v)$ zur Darstellung $\Theta_B(v)$, mittels der inversen Matrix erhält.

Wir wollen nun die Problemstellung des letzten Beispiels auf den allgemeinen Fall übertragen. Dazu betrachten wir folgende Situation:

In einem n-dimensionalen Vektorraum V $(n > 0)$ seien zwei Basen $B = \{b_1, \ldots, b_n\}$ und $\overline{B} = \{\overline{b}_1, \ldots, \overline{b}_n\}$ gegeben. Wir wollen eine Matrix A angeben, die den **Übergang von der „alten“ Basis B zur „neuen“ Basis $\overline{B}$ beschreibt, d.h. die Vektoren der Basis B durch die von $\overline{B}$ ausdrückt**.

Dazu überlegen wir uns zunächst:

Jedes b_i lässt sich nach Satz 7.15 eindeutig als Linearkombination der $\overline{b}_1, \ldots, \overline{b}_n$ mit Koeffizienten aus $\mathbb{K}$ darstellen:

$$\begin{aligned} b_1 &= a_{11}\,\overline{b}_1 + a_{21}\,\overline{b}_2 + \cdots + a_{n1}\,\overline{b}_n \\ b_2 &= a_{12}\,\overline{b}_1 + a_{22}\,\overline{b}_2 + \cdots + a_{n2}\,\overline{b}_n \\ &\vdots \\ b_n &= a_{1n}\,\overline{b}_1 + a_{2n}\,\overline{b}_2 + \cdots + a_{nn}\,\overline{b}_n. \end{aligned}$$

In Summenschreibweise also

$$b_i = \sum_{j=1}^{n} a_{ji}\,\overline{b}_j, \qquad i = 1, \ldots, n; \quad a_{ji} \in \mathbb{K}. \tag{7.2}$$

Die Matrix

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

nennen wir **Übergangsmatrix** des Basiswechsels von der „alten“ Basis B nach der „neuen“ Basis $\overline{B}$. Die Einträge $a_{1i}, a_{2i}, \ldots, a_{ni}$ in der *i-ten Spalte* von A sind die zur Linearkombination von b_i aus den $\overline{b}_1, \ldots, \overline{b}_n$ benötigten Koeffizienten.

Umgekehrt hat jedes $\overline{b}_j$ eine eindeutige Basisdarstellung bezüglich B:

$$\overline{b}_j = \sum_{i=1}^{n} c_{ij}\, b_i, \qquad j = 1, \ldots, n; \quad c_{ij} \in \mathbb{K} \tag{7.3}$$

mit der analog konstruierten, zum Basiswechsel von $\overline{B}$ nach B gehörigen Übergangsmatrix

$$C := \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

Wir sagen, dass die **Basiswechsel** zwischen der „alten“ Basis $B = \{b_1, \ldots, b_n\}$ und der „neuen“ Basis $\overline{B} = \{\overline{b}_1, \ldots, \overline{b}_n\}$ durch (7.2), (7.3) gegeben sind.

Satz 7.19 (Basiswechsel: Übergangsmatrizen) *Seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $B, \overline{B}$ zwei Basen von V. Dann ist die Übergangsmatrix C des Basiswechsels von $\overline{B}$ nach B die Inverse der Übergangsmatrix A des Basiswechsels von B nach $\overline{B}$, d.h. es gilt $C = A^{-1}$.*

Beweis: In (7.2) setzen wir die $\overline{b}_j$ aus (7.3) ein (und ändern den Summationsindex i in k):

$$b_i = \sum_{j=1}^{n} a_{ji} \left(\sum_{k=1}^{n} c_{kj} b_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} c_{kj}\, a_{ji} \right) b_k.$$

Durch Vergleich der Koeffizienten links und rechts erhalten wir wegen der eindeutigen Darstellbarkeit

$$\sum_{j=1}^{n} c_{kj}\, a_{ji} = \delta_{ik} := \begin{cases} 0 & \text{für} \quad i \neq k \\ 1 & \text{für} \quad i = k \end{cases} . \tag{7.4}$$

Die hierdurch definierten δ_{ik} nennt man auch **Kronecker-Symbole**. Setzen wir entsprechend b_i aus (7.2) in (7.3) ein, so erhalten wir die zu (7.4) analoge Beziehung

$$\sum_{j=1}^{n} a_{ij}\, c_{jk} = \delta_{ik}. \tag{7.5}$$

Unter Verwendung der Matrizenmultiplikation lassen sich die Bedingungen (7.4), (7.5) an die Übergangsmatrizen A und C durch das Gleichungspaar

$$CA = E \quad \text{und} \quad AC = E$$

beschreiben. Es gilt also $C = A^{-1}$. ■

Wie transformiert sich nun der Komponentenvektor $\Theta_B(v)$ von $v \in V$ bezüglich B in den Komponentenvektor $\Theta_{\overline{B}}(v)$ bezüglich $\overline{B}$? Das beantwortet der folgende

Satz 7.20 (Basiswechsel: Vektor-Komponenten) *Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $B, \overline{B}$ Basen von V. Für einen gegebenen Vektor $v \in V$ seien $\Theta_B(v)$ und $\Theta_{\overline{B}}(v)$ die Komponentenvektoren bezüglich B bzw. $\overline{B}$. Dann gilt in Matrixschreibweise (vgl. Bemerkung 7.17)*

$$\Theta_{\overline{B}}(v) = A \cdot \Theta_B(v) \quad \textit{und} \quad \Theta_B(v) = A^{-1} \cdot \Theta_{\overline{B}}(v).$$

Dabei ist A die Übergangsmatrix des Basiswechsels von B nach $\overline{B}$.

Beweis: Seien $B = \{b_1, \ldots, b_n\}$ und $\overline{B} = \{\bar{b}_1, \ldots, \bar{b}_n\}$. Es sei $\Theta_B(v) = (v_1, \ldots, v_n)$ und $\Theta_{\overline{B}}(v) = (\overline{v}_1, \ldots, \overline{v}_n)$. Dann gilt nach Definition

$$v = \sum_{j=1}^{n} v_j b_j = \sum_{i=1}^{n} \overline{v}_i \bar{b}_i.$$

Nach Definition der Übergangsmatrix A (vgl. (7.2)) gilt dann auch

$$v = \sum_{i=1}^{n} v_i b_i = \sum_{i=1}^{n} v_i \left(\sum_{j=1}^{n} a_{ji}\, \bar{b}_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} v_i \right) \bar{b}_j.$$

Wegen der eindeutigen Darstellbarkeit folgt nun durch Koeffizientenvergleich

$$\overline{v}_j = \sum_{i=1}^{n} a_{ji} v_i \qquad (j = 1, \ldots, n).$$

Als Matrixgleichung geschrieben erhält man also

$$\Theta_{\overline{B}}(v) = \begin{pmatrix} \overline{v}_1 \\ \vdots \\ \overline{v}_n \end{pmatrix} = A \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A \cdot \Theta_B(v)$$

Nach Satz 7.19 ist A invertierbar. Durch Multiplikation der letzten Gleichung mit A^{-1} erhält man

$$\Theta_B(v) = A^{-1} \Theta_{\overline{B}}(v).$$

■

8 Untervektorräume

8.1 Was ist ein Untervektorraum?

Teilmengen von Vektorräumen, die selber wieder Vektorräume sind (mit derselben Addition und skalaren Multiplikation wie im „umgebenden“ Vektorraum) haben wir schon in den Beispielen des letzten Abschnitts kennengelernt. Solche Teilmengen nennt man Untervektorräume. So ist z.B. die Lösungsmenge eines homogenen LGS mit n Variablen $x_1, \ldots, x_n \in \mathbb{K}$ ein Untervektorraum von $\mathbb{K}^n$. Ebenso ist die Menge der Polynome $\mathbb{K}[X]$ ein Untervektorraum des Vektorraums $\mathbb{K}^{\mathbb{N}_0}$ aller Folgen in $\mathbb{K}$.

Definition 8.1 (**UVR**) Eine Teilmenge U eines $\mathbb{K}$-Vektorraums V heißt **Untervektorraum** von V, wenn U bezüglich der in V erklärten Addition und skalaren Multiplikation ein $\mathbb{K}$-Vektorraum ist.

Insbesondere ist $(U, +)$ eine Untergruppe von $(V, +)$. Deswegen ist der Nullvektor als neutrales Element der Addition in U und V derselbe und auch das Inverse $-u$ eines Vektors u in U stimmt mit dem Inversen $-u$ von u in V überein.

Wegen $0 \in U$ ist jeder Untervektorraum von V eine nichtleere Teilmenge von V. Weiter ist U „abgeschlossen“ bzgl. der Addition, d.h. mit $x, y \in U$ ist auch $x+y \in U$. Ebenso ist U bzgl. der $\mathbb{K}$-Multiplikation abgeschlossen, d.h. mit $\lambda \in \mathbb{K}$ und $x \in U$ gilt auch $\lambda x \in U$.

Diese Eigenschaften genügen nun bereits, um festzustellen, ob eine Teilmenge $U \subset V$ ein Untervektorraum ist:

Hilfssatz 8.2 (UVR-Kriterium) *Eine Teilmenge U eines $\mathbb{K}$-Vektorraums V ist genau dann ein Untervektorraum von V, wenn die folgenden beiden Eigenschaften gelten:*

U1 $U \neq \emptyset$

U2 $\forall x, y \in U$ und $\forall \lambda \in \mathbb{K}$ gilt : $x + y \in U \ \wedge \ \lambda x \in U$.

Beweis:

„⇒“ Wenn U ein Untervektorraum von V ist, so gelten nach der Vorbemerkung die Eigenschaften **U1** und **U2**.

„⇐“ Es sei jetzt U eine Teilmenge von V mit den Eigenschaften **U1** und **U2**. U ist also bzgl. der Addition und $\mathbb{K}$-Multiplikation abgeschlossen. Die in V gültigen „Rechenregeln“ **V2** gelten auch in $U \subset V$. Es bleibt zu zeigen, dass mit x auch $-x$ in U liegt und dass 0 in U liegt: Wegen **U1** ist $U \neq \emptyset$. Also gibt es ein

$x \in U$, und nach **U2** folgt $0x = 0 \in U$. Weiter ist für jedes $x \in U$ stets auch $(-1)x = -(1x) = -x \in U$. ■

Beispiel 8.3

1. Jeder Vektorraum V hat mindestens zwei Untervektorräume: den „Nullraum“ $\{0\}$ und V selbst.
2. Die Teilmenge $U = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\}$ des $\mathbb{R}^3$ ist ein Untervektorraum.
3. Die Lösungsmenge eines homogenen LGS mit n Variablen $x_1, \ldots, x_n \in \mathbb{K}$ ist ein Untervektorraum von $\mathbb{K}^n$ (vgl. Beispiel 8 in Abschnitt 6.2). Man kann sogar zeigen, dass jeder Untervektorraum U von $\mathbb{K}^n$ die Lösungsmenge eines geeigneten homogenen LGS ist.

Hilfssatz 8.4 (Lineare Hülle = UVR) *Für jede Teilmenge M eines Vektorraums V ist die lineare Hülle $[M]$ ein Untervektorraum von V.*

Beweis: Für $M = \emptyset$ ist $[M]$ der Nullraum, also ein Untervektorraum von V. Für $M \neq \emptyset$ ist $[M]$ nach dem UVR-Kriterium 8.2 ein Untervektorraum: zunächst ist $[M] \neq \emptyset$, da $M \subset [M]$ gilt. Sind weiter

$$x = \sum_{i=1}^{k} \lambda_i v_i \quad \text{und} \quad y = \sum_{j=1}^{l} \mu_j v_j' \qquad \text{mit } \lambda_i, \mu_j \in \mathbb{K} \text{ und } v_i, v_j' \in M$$

Linearkombinationen aus $[M]$, so sind auch

$$x + y = \sum_{i=1}^{k} \lambda_i \, v_i + \sum_{j=1}^{l} \mu_j \, v_j' \quad \text{und} \quad \mu x = \sum_{i=1}^{l} (\mu\lambda_i) \, v_i \quad \text{für } \mu \in \mathbb{K}$$

Linearkombinationen aus $[M]$. ■

In den folgenden beiden Abschnitten werden wir einige Möglichkeiten kennenlernen, aus gegebenen Teilmengen oder Untervektorräumen von V neue Untervektorräume zu bilden.

8.2 Durchschnitt und Summe von UVR

Hilfssatz 8.5 (Durchschnitt von UVR) *Es sei $\mathfrak{U}$ eine nichtleere (endliche oder unendliche) Menge von Untervektorräumen eines Vektorraums V. Dann ist der Durchschnitt*

$$D = \bigcap_{U \in \mathfrak{U}} U$$

ein Untervektorraum von V.

Beweis: Wir benutzen das UVR-Kriterium 8.2: Da jedes $U \in \mathfrak{U}$ den Nullvektor enthält, ist $D \neq \emptyset$. Es seien $x, y \in D$ und $\lambda \in \mathbb{K}$. Dann gilt $x, y \in U$ für alle $U \in \mathfrak{U}$. Da jedes U ein UVR ist, gilt auch $x + y \in U$ und $\lambda x \in U$ für alle $U \in \mathfrak{U}$, also $x + y \in D$ und $\lambda x \in D$. ■

Satz 8.6 *Es sei $\mathfrak{U}$ die Menge aller Untervektorräume eines Vektorraums V, die eine gegebene Menge $M \subset V$ enthalten. Dann gilt*

$$\bigcap_{U \in \mathfrak{U}} U = [M].$$

D.h. die lineare Hülle $[M]$ ist der „kleinste" UVR, der M enthält.

Beweis: Für die gegebene Menge $M \subset V$ gilt $M \subset [M]$, und nach Hilfssatz 8.4 ist $[M]$ ein Untervektorraum von V. Also ist $[M] \in \mathfrak{U}$ und damit $[M] \supset \bigcap_{U \in \mathfrak{U}} U$.

Andererseits ist jedes $v \in [M]$ eine Linearkombination von Vektoren aus $M \neq \emptyset$ (bzw. $v = 0$ für $M = \emptyset$). Da $M \subset U$ für alle Untervektorräume $U \in \mathfrak{U}$, liegt auch jedes solche v in U und daher in $\bigcap_{U \in \mathfrak{U}} U$. Also ist auch $[M] \subset \bigcap_{U \in \mathfrak{U}} U$. ■

Bemerkung 8.7 Die *Vereinigung* zweier Untervektorräume U_1, U_2 eines Vektorraumes V ist im allgemeinen *kein* Untervektorraum. So ist die Menge $M := [(1,0)] \cup [(0,1)] \subset \mathbb{R}^2$ kein UVR, da z.B. der Vektor $(2,1) = 2 \cdot (1,0) + (0,1)$ eine Linearkombination der Vektoren $(1,0), (0,1) \in \mathbb{R}^2$ ist, aber nicht in M liegt.

Nach Satz 8.4 ist aber die lineare Hülle $[U_1 \cup U_2]$ zweier Untervektorräume U_1, U_2 ein Untervektorraum. Wir definieren daher allgemein

Definition 8.8 Die **Summe** der Untervektorräume U_1, U_2 des $\mathbb{K}$-Vektorraums V ist der Untervektorraum $U_1 + U_2 := [U_1 \cup U_2]$.

Die Bezeichnung „Summe" ist dadurch gerechtfertigt, dass sich jeder Vektor aus $U_1 + U_2$ als Summe $u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$ schreiben lässt:

Satz 8.9 (Charakterisierung Summe) *Die Summe $U_1 + U_2$ der Untervektorräume U_1, U_2 des $\mathbb{K}$-Vektorraums V ist die Menge aller Vektoren $v = u_1 + u_2$ mit $u_1 \in U_1, u_2 \in U_2$.*

Beweis: Es sei

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

Wir zeigen, dass $W = U_1 + U_2$ gilt: Nach Hilfssatz 8.2 ist W ein Untervektorraum von V. Außerdem gilt $U_1 \cup U_2 \subset W$, also nach Satz 8.6: $U_1 + U_2 = [U_1 \cup U_2] \subset W$. Andererseits ist jedes $v \in W$ eine Linearkombination von Vektoren aus $U_1 \cup U_2$, also $W \subset [U_1 \cup U_2]$. ■

Definition 8.10 Die Summe $U_1 + U_2$ zweier Untervektorräume U_1, U_2 eines Vektorraums V heißt **direkte Summe** $U_1 \oplus U_2$, wenn $U_1 \cap U_2 = \{0\}$.

Für eine direkte Summe ist die Darstellung jedes Vektors aus $U_1 + U_2$ als Summe $w = v_1 + v_2$ sogar eindeutig:

Satz 8.11 (Charakterisierung direkte Summe) *Die Summe U_1+U_2 der Untervektorräume U_1, U_2 eines Vektorraums V ist genau dann direkt, wenn es zu jedem $x \in U_1 + U_2$ genau einen Vektor $u_1 \in U_1$ und genau einen Vektor $u_2 \in U_2$ gibt mit $x = u_1 + u_2$.*

Beweis:

„$\Rightarrow$" Die Summe $U_1 + U_2$ sei direkt, und $x \in U_1 + U_2$ habe zwei Darstellungen $x = u_1 + u_2 = u_1' + u_2'$. Dann gilt $u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2$, und wegen $U_1 \cap U_2 = \{0\}$ folgt $u_1 = u_1'$, $u_2 = u_2'$.

„$\Leftarrow$" Sei $u \in U_1 \cap U_2 \subset U_1 + U_2$. Wir können dann schreiben $u = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$. Da auch gilt $u = 0 + u = u + 0$ folgt aus der Voraussetzung der eindeutigen Darstellung, dass $u_1 = u_2 = 0$. Also ist $u = u_1 + u_2 = 0$ und da u beliebig war, folgt $U_1 \cap U_2 = \{0\}$. ■

Man kann den Begriff der Summe von Untervektorräumen eines Vektorraums V auch auf mehr als zwei Untervektorräume ausdehnen:

Bemerkung 8.12 Es sei $\mathfrak{U}$ eine Menge von Untervektorräumen U_i von V, also $\mathfrak{U} = \{U_i \mid i \in J\}$ mit einer beliebigen Indexmenge $J \neq \emptyset$. Unter der **Summe** der U_i versteht man den Untervektorraum

$$\sum_{i \in J} U_i := \Big[\bigcup_{i \in J} U_i\Big]. \tag{8.1}$$

Die Summe heißt **direkt**, wenn

$$U_i \cap \sum_{j \in J \setminus \{i\}} U_j = \{0\} \quad \text{für alle } i \in J. \tag{8.2}$$

Satz 8.13 (Komplement) *Zu jedem Untervektorraum U_1 eines Vektorraums V gibt es einen* **Komplementärraum** *U_2, d.h. einen Untervektorraum U_2 von V mit $V = U_1 \oplus U_2$.*

Beweis: Nach Satz 7.8 gibt es für U_1 eine Basis B_1, die wir nach dem Basisergänzungssatz durch eine Menge B_2 zu einer Basis $B = B_1 \cup B_2$ von V ergänzen können. Wir setzen $U_2 := [B_2]$. Dann ist $V = U_1 + U_2$, denn jedes $v \in V$ lässt sich als Linearkombination von Vektoren aus $B_1 \cup B_2$ darstellen. Außerdem ist $U_1 \cap U_2 = \{0\}$, denn $B_1 \cup B_2$ ist linear unabhängig. ■

Bemerkung 8.14 Man beachte, dass der Komplementärraum zu einem gegebenen Untervektorraum im Allgemeinen *nicht eindeutig* ist!

8.3 Dimensionssätze

Satz 8.15 *Ist U ein Untervektorraum eines n-dimensionalen Vektorraums V, so ist U endlichdimensional, und es gilt* $\dim U \leq \dim V$. *Das Gleichheitszeichen gilt genau dann, wenn $U = V$.*

Beweis: Für $U = \{0\}$ ist der Satz offensichtlich richtig. Es sei jetzt $U \neq \{0\}$. Wegen $U \subset V$ und nach Satz 7.10 kann es in U höchstens n linear unabhängige Vektoren geben. Seien $b_1, \ldots, b_p$ linear unabhängige Vektoren in U, wobei $1 \leq p \leq n$ die maximale Anzahl linear unabhängiger Vektoren in U ist. Für jedes $v \in U$ sind dann $b_1, \ldots, b_p, v$ für jedes $v \in U$ linear abhängig, und wie im Beweis zu Satz 7.13 sieht man, dass v eine Linearkombination der b_i ist. Der Untervektorraum U wird also von den Vektoren $b_1, \ldots, b_p$ erzeugt. Da diese Vektoren linear unabhängig sind, bilden sie eine Basis von U, und es gilt dim $U = p \leq n =$ dim V.

Für dim $U =$ dim V, also $p = n$, ist nach Obigem $[b_1, \ldots, b_n] = U$. Nach Satz 7.13 ist aber auch $[b_1, \ldots, b_n] = V$, also folgt $U = V$. Umgekehrt gilt mit $U = V$ natürlich auch dim $U =$ dim V. ■

Beispiel 8.16 Gegeben seien die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{pmatrix}, \quad v_4 = \begin{pmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{pmatrix}$$

gegeben, und es sei $U := [v_1, v_2, v_3, v_4] \subset \mathbb{R}^5$ die lineare Hülle.

- Wir wollen unter den Vektoren $v_1, \ldots, v_4$ eine Basis von $U \subset \mathbb{R}^5$ finden. Dazu prüfen wir zunächst nach, ob die Vektoren $v_1, \ldots, v_4$ linear unabhängig sind. Der Ansatz
$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4 = 0$$
führt auf ein lineares Gleichungssystem mit der zugehörigen Matrix
$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{pmatrix}$$

Durch elementare Zeilenumformungen erhalten wir mit dem Gauß-Algorithmus

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & -5 & -10 & 10 \\ 0 & 3 & 6 & -6 \\ 0 & 4 & 8 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ \mid -\frac{1}{5} \\ \mid -\frac{1}{3} \\ \\ \\ \end{matrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Anhand der Normalform sehen wir, dass v_1, v_2 und v_4 linear unabhängig sind, weil das lineare Gleichungssystem

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_4 v_4 = 0$$

nur trivial lösbar ist. Weiter folgt, dass v_3 eine Linearkombination von v_1 und v_2 ist, da das lineare Gleichungssystem

$$\mu_1 v_1 + \mu_2 v_2 = v_3$$

lösbar ist. Also gilt $U = [v_1, v_2, v_4]$ und $\{v_1, v_2, v_4\}$ ist eine Basis von U.

- Wir wollen eine möglichst einfache Basis von U finden, indem wir die Vektoren $v_1, \ldots, v_4$ durch geeignete Linearkombinationen ersetzen. Zur praktischen Durchführung schreiben wir die Vektoren $v_1, \ldots, v_4$ in die *Zeilen* einer Matrix und wenden wieder den Gauß-Algorithmus an.

$$\begin{pmatrix} 1 & 2 & -1 & -1 & -1 \\ 2 & -1 & 1 & 2 & -2 \\ 3 & -4 & 3 & 5 & -3 \\ -1 & 8 & -5 & -6 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 & -1 & -1 \\ 0 & -5 & 3 & 4 & 0 \\ 0 & -10 & 6 & 8 & 0 \\ 0 & 10 & -6 & -7 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & -1 & -1 & -1 \\ 0 & 1 & -\frac{3}{5} & -\frac{4}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 & 0 & -1 \\ 0 & 1 & -\frac{3}{5} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & \frac{1}{5} & 0 & -1 \\ 0 & 1 & -\frac{3}{5} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Also ist

$$U = \Bigg[\underbrace{\begin{pmatrix} 1 \\ 0 \\ \frac{1}{5} \\ 0 \\ -1 \end{pmatrix}}_{=:u_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ -\frac{3}{5} \\ 0 \\ 0 \end{pmatrix}}_{=:u_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{=:u_3} \Bigg].$$

Die Vektoren u_1, u_2, u_3 bilden eine Basis von U, denn sie sind auch linear unabhängig: aus

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 0$$

folgt nämlich $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Sind U_1, U_2 zwei Untervektorräume eines Vektorraumes V, so gilt nach dem letzten Satz 8.15

$$0 \leq \dim(U_1 \cap U_2) \leq \dim U_i \leq \dim(U_1 + U_2) \leq n$$

für $i \in \{1, 2\}$. Eine genauere Aussage liefert der folgende

Satz 8.17 (Dimensionssatz für UVR) *Für zwei Untervektorräume U_1 und U_2 eines n-dimensionalen Vektorraumes V gilt*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Beweis: Ist U_1 der Nullraum, so lautet die Behauptung $\dim U_2 = 0 + \dim U_2 - 0$ und ist also richtig. Wir können im folgenden also $\dim U_1 = p > 0$ und $\dim U_2 = q > 0$ annehmen.

Es sei $d = \dim(U_1 \cap U_2)$ und

$$B_d = \{b_1, \ldots, b_d\}$$

eine Basis von $U_1 \cap U_2$. Für $d = 0$ ist $B_d = \emptyset$. Nach dem Basisergänzungssatz können wir B_d zu einer Basis

$$B' = \{b_1, \ldots, b_d, b'_{d+1}, \ldots, b'_p\}$$

von U_1 ergänzen. Analog sei

$$B'' = \{b_1, \ldots, b_d, b''_{d+1}, \ldots, b''_q\}$$

eine Basis von U_2. Wir wollen zeigen, dass

$$B_s = \{b_1, \ldots, b_d, b'_{d+1}, \ldots, b'_p, b''_{d+1}, \ldots, b''_q\}, \quad (s = p + q - d)$$

eine Basis von U_1+U_2 ist. Zunächst ist $B_s \subset U_1 \cup U_2$, also $[B_s] \subset [U_1 \cup U_2] = U_1+U_2$. Andererseits ist $U_1+U_2 \subset [B_s]$; denn jeder Vektor $v \in U_1+U_2$ lässt sich darstellen in der Gestalt

$$\begin{aligned} v &= u_1 + u_2 \qquad\qquad (u_1 \in U_1,\, u_2 \in U_2) \\ &= \sum_{i=1}^{d} \lambda_i b_i + \sum_{i=d+1}^{p} \lambda_i' b_i' + \sum_{i=1}^{d} \hat{\lambda}_i b_i + \sum_{i=d+1}^{q} \lambda_i'' b_i'' \\ &= \sum_{i=1}^{d} (\lambda_i + \hat{\lambda}_i) b_i + \sum_{i=d+1}^{p} \lambda_i' b_i' + \sum_{i=d+1}^{q} \lambda_i'' b_i'', \end{aligned}$$

sodass $v \in [B_s]$. Damit ist $[B_s] = U_1+U_2$ und B_s ist erzeugende Menge von U_1+U_2.

Weiter ist B_s linear unabhängig, denn aus einer Vektorgleichung

$$\sum_{i=1}^{d} \lambda_i b_i + \sum_{i=d+1}^{p} \lambda_i' b_i' + \sum_{i=d+1}^{q} \lambda_i'' b_i'' = 0 \tag{8.3}$$

folgt

$$\sum_{i=1}^{d} \lambda_i b_i + \sum_{i=d+1}^{p} \lambda_i' b_i' = - \sum_{i=d+1}^{q} \lambda_i'' b_i''. \tag{8.4}$$

Dieser Vektor (8.4) liegt, wie die linke Seite zeigt, in U_1, und wie die rechte Seite zeigt, auch in U_2. Also liegt er in $U_1 \cap U_2$ und lässt sich in der Gestalt

$$\sum_{i=1}^{d} \alpha_i b_i \tag{8.5}$$

darstellen. Für $d = 0$ ist das der Nullvektor. Weil nach Satz 7.15 jeder Vektor aus U_2 eine eindeutige Basisdarstellung bezüglich B'' hat, folgt durch Vergleich von (8.5) mit der rechten Seite von (8.4): $\alpha_1 = \cdots = \alpha_d = 0$ und $\lambda_{d+1}'' = \cdots = \lambda_q'' = 0$. Wegen der linearen Unabhängigkeit von B' folgt weiter aus (8.4): $\lambda_1 = \cdots = \lambda_d = 0$ und $\lambda_{d+1}' = \cdots = \lambda_p' = 0$. In (8.3) sind also alle Koeffizienten Null, und B_s ist linear unabhängig.

Als linear unabhängige und erzeugende Menge von U_1+U_2 ist B_s nach Satz 7.4 Basis von U_1+U_2, und es ist

$$\begin{aligned} \dim(U_1+U_2) &= s = p + q - d \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2). \end{aligned}$$

■

Folgerung 8.18 *Für die direkte Summe von zwei UVR U_1 und U_2 von V gilt wegen $U_1 \cap U_2 = \{0\}$ die Gleichung*

$$\dim U_1 + \dim U_2 = \dim(U_1 \oplus U_2).$$

8.4 UVR in der Praxis: der Rang einer Matrix

Wir betrachten eine $m \times n$-Matrix $A \in \mathbb{K}^{m\times n}$. Die Zeilen (bzw. Spalten) von A spannen einen Untervektorraum von $\mathbb{K}^n$ (bzw. $\mathbb{K}^m$) auf. Beim Rechnen mit Matrizen werden oft elementare Zeilenumformungen durchgeführt (z.B. beim Gauß-Algorithmus). Wie wirken sich solche Transformationen auf den Zeilen- bzw. Spaltenraum aus? Um dies zu klären kommen wir nochmals auf das Beispiel 8.16 zurück. Das dort angegebene Verfahren soll nun allgemein dargestellt werden.

Sei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m\times n}.$$

Zu A kann man zwei Systeme von Vektoren aus $\mathbb{K}^m$ bzw. $\mathbb{K}^n$ betrachten. Zunächst bilden die *Spalten*

$$s_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, s_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

ein System von n Vektoren im $\mathbb{K}^m$. Sie spannen einen Untervektorraum

$$SR := [s_1, \ldots, s_n] \subseteq \mathbb{K}^m$$

auf.

Entsprechend bilden die *Zeilen* von A

$$z_1 = (a_{11}, \cdots, a_{1n}), \ldots, z_m = (a_{m1}, \cdots, a_{mn}),$$

ein System von m Vektoren im $\mathbb{K}^n$. Sie spannen einen Untervektorraum

$$ZR := [z_1, \ldots, z_m] \subseteq \mathbb{K}^n$$

auf.

Für die Matrix A gilt somit

$$A = (s_1 | \cdots | s_n) = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}.$$

Es sei nun $\tilde{A}$ die beim Gauß-Algorithmus durch Anwendung von Zeilenumformungen entstehende Endmatrix, also die Gaußsche Normalform von A, und es seien

$$\tilde{s}_1, \ldots, \tilde{s}_n \quad \text{bzw.} \quad \tilde{z}_1, \ldots, \tilde{z}_m$$

die zugehörigen Spalten- bzw. Zeilenvektoren. Wir wollen überlegen, welcher Zusammenhang zwischen den Spaltenvektoren s_j und $\tilde{s}_j$ $(j = 1, \ldots, n)$ bzw. den Zeilenvektoren z_i und $\tilde{z}_i$ $(i = 1, \ldots, m)$ besteht.

Hilfssatz 8.19 *Es ist*

$$ZR = [\tilde{z}_1, \ldots, \tilde{z}_m]$$

und diejenigen $\tilde{z}_i$, die nicht 0 sind, bilden eine (besonders einfache) Basis von ZR.

Beweis: Da $\tilde{A}$ durch endlich viele elementare Zeilenumformungen aus A entstanden ist, sind die Vektoren $\tilde{z}_1, \ldots, \tilde{z}_m$ Linearkombinationen der ursprünglichen Vektoren $z_1, \ldots, z_m$. Also gilt $[\tilde{z}_1, \ldots, \tilde{z}_m] \subset ZR$.

Da umgekehrt jede angewendete Zeilenumformung wieder rückgängig gemacht werden kann, entsteht auch A durch endlich viele Zeilenumformungen aus $\tilde{A}$. Somit sind die Vektoren $z_1, \ldots, z_m$ Linearkombinationen der Vektoren $\tilde{z}_1, \ldots, \tilde{z}_m$ und es gilt $ZR \subset [\tilde{z}_1, \ldots, \tilde{z}_m]$.

Weiter erkennt man an der Gestalt der Normalform mit k Stufen

$$\tilde{A} = \left(\begin{array}{ccccccccccccccccccccccc}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & & & \cdots & & & 0 & 1 & * & \cdots & * & 0 & * & \cdots & & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & & & & & \cdots & & & & & 0 & 1 & * & \cdots & & 0 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & & & & & & & & & & & & & & \ddots & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & & & & & & & & \cdots & & & & & & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\
0 & & & & & & & & & & & \cdots & & & & & & & 0 & 1 & * & \cdots & * \\
0 & & & & & & & & & & & & \cdots & & & & & & & & & & 0 \\
\vdots & & & & & & & & & & & & & & & & & & & & & & \vdots \\
0 & & & & & & & & & & & & \cdots & & & & & & & & & & 0
\end{array}\right)$$

unmittelbar, dass die ersten k Zeilen linear unabhängig sind. Das sind aber genau diejenigen $\tilde{z}_i$, die von 0 verschieden sind. ■

Nicht ganz so einfach ist der Zusammenhang zwischen den alten und den neuen Spaltenvektoren, denn im Allgemeinen ist der Untervektorraum $[\tilde{s}_1, \ldots, \tilde{s}_n]$ von SR verschieden. Es gilt aber

Hilfssatz 8.20 *Es ist*

$$\dim[\tilde{s}_1, \ldots, \tilde{s}_n] = \dim SR.$$

Diejenigen Vektoren $s_{j_1}, \ldots, s_{j_k}$, deren Indizes $j_1, \ldots, j_k \in \{1, \ldots, n\}$ zu den Treppenstufen in $\tilde{A}$ gehören, die also beim Gauß-Algorithmus in Vektoren der Standardbasis

$$\tilde{s}_{j_1} = e_1, \; \ldots, \; \tilde{s}_{j_k} = e_k$$

übergehen, bilden eine Basis von SR.

Beweis: Wir betrachten die linearen Gleichungssysteme $Ay = 0$ bzw. $\tilde{A}y = 0$, $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$, die wir in der Form

$$y_1 s_1 + \ldots + y_n s_n = 0 \quad \text{bzw.} \quad y_1 \tilde{s}_1 + \ldots + y_n \tilde{s}_n = 0$$

schreiben. Sie haben dieselbe Lösungsmenge. Aus der Gestalt von $\tilde{A}$ ergibt sich, dass die y_j mit $j \notin \{j_1, \ldots, j_k\}$ beliebig gewählt werden können. Damit ist jeder Vektor s_j mit $j \notin \{j_1, \ldots, j_k\}$ Linearkombination von $s_{j_1}, \ldots, s_{j_k}$ und ebenso jeder Vektor $\tilde{s}_j$ mit $j \notin \{j_1, \ldots, j_k\}$ Linearkombination von $\tilde{s}_{j_1}, \ldots, \tilde{s}_{j_k}$. Also gilt

$$SR = [s_{j_1}, \ldots, s_{j_k}] \quad \text{bzw.} \quad [\tilde{s}_1, \ldots, \tilde{s}_n] = [\tilde{s}_{j_1}, \ldots, \tilde{s}_{j_k}],$$

und die Vektoren $\tilde{s}_{j_1} = e_1, \ldots, \tilde{s}_{j_k} = e_k$ sind linear unabhängig.

Wir zeigen jetzt, dass auch die Vektoren $s_{j_1}, \ldots, s_{j_k}$ linear unabhängig sind. Ist $y_{j_1} s_{j_1} + \ldots + y_{j_k} s_{j_k} = 0$, so ist

$$y^* := (0, \ldots, 0, \underbrace{y_{j_1}}_{j_1\text{-te Stelle}}, 0, \ldots, 0, y_{j_2}, 0, \ldots, \ldots, \underbrace{y_{j_k}}_{j_k\text{-te Stelle}}, 0, \ldots, 0)$$

eine Lösung von $Ay = 0$, also auch von $\tilde{A}y = 0$ (da beide LGS die gleiche Lösungsmenge haben!). Aus $\tilde{A}y^* = 0$ folgt aber sofort $y_{j_1} = \ldots = y_{j_k} = 0$. Die Vektoren $s_{j_1}, \ldots, s_{j_k}$ bilden also eine Basis von SR.

Insgesamt haben wir damit gezeigt:

$$\dim SR = k = \dim[\tilde{s}_{j_1}, \ldots, \tilde{s}_{j_k}] = \dim[\tilde{s}_1, \ldots, \tilde{s}_n].$$

■

Fazit: Es seien m Vektoren $x_1, \ldots, x_m \in \mathbb{K}^n$ gegeben und es soll der Untervektorraum $[x_1, \ldots, x_m] \subseteq \mathbb{K}^n$ untersucht werden. Ist man an einer Basis von $[x_1, \ldots, x_m]$ in „Treppenform“ interessiert, so wende man den Gauß-Algorithmus auf die Matrix

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

mit den *Zeilen* $x_1, \ldots, x_m$ an.

Ist man daran interessiert, welche der Vektoren $x_1, \ldots, x_m$ eine Basis von $[x_1, \ldots, x_m]$ bilden, so muss man den Gauß-Algorithmus auf die Matrix

$$(x_1 | \ldots | x_n)$$

mit den *Spalten* $x_1, \ldots, x_m$ anwenden. Zur Bestimmung der Dimension von $x_1, \ldots, x_m$ können beide Verfahren benutzt werden.

Es sei nun wieder eine Matrix $A \in \mathbb{K}^{m \times n}$ gegeben mit den Spalten $s_1, \ldots, s_n$ in $\mathbb{K}^m$ und den Zeilen $z_1, \ldots, z_m$ in $\mathbb{K}^n$.

Definition 8.21 Die Zahl $\dim[s_1, \ldots, s_n]$ heißt der **Spaltenrang** von A, die Zahl $\dim[z_1, \ldots, z_m]$ heißt der **Zeilenrang** von A.

Satz 8.22 (Rang) *Der Zeilenrang und der Spaltenrang einer Matrix $A \in \mathbb{K}^{m\times n}$ sind gleich.*

Wir nennen diese Zahl dann einfach den **Rang** von A und schreiben $\operatorname{Rang} A$ oder $\operatorname{Rg} A$.

BEWEIS: Nach Hilfssatz 8.19 ist der Zeilenrang von A gleich der Anzahl k der Stufen in der Normalform von $\tilde{A}$. Nach Hilfssatz 8.20 ist auch der Spaltenrang gleich k. ■

Folgerung 8.23 *Für alle $A \in \mathbb{K}^{m\times n}$ gilt:* $\operatorname{Rg} A = \operatorname{Rg} A^\top$.

Bemerkung 8.24 Im Kapitel über lineare Abbildungen werden wir noch folgende Charakterisierung von invertierbaren Matrizen sehen: *Für alle $A \in \mathbb{K}^{n\times n}$ gilt: A ist genau dann regulär (d.h. invertierbar), wenn* $\operatorname{Rg} A = n$.

Bemerkung 8.25 Zur Bestimmung des (Zeilen- oder Spalten-)Ranges einer Matrix A kann man nach Satz 8.22 sowohl Zeilen- wie Spaltenoperationen verwenden. Jede Matrix vom Rang $r \geq 0$ lässt sich dann in die Gestalt

$$C = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \\ 0 & \ddots & \ddots & \vdots & \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & \\ \hline \multicolumn{4}{c|}{\underbrace{\qquad\qquad 0 \qquad\qquad}_{r}} & 0 \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\0\\\vdots\\0\end{array}}\right\} r \\ \vphantom{0} \end{array} \tag{8.6}$$

überführen, wobei C für $r = 0$ die Nullmatrix ist.

Man beachte, dass demgegenüber beim Gaußschen Algorithmus nur Zeilenoperationen erlaubt sind!

8.5 Faktorräume

Hilfssatz 8.26 *Sei V ein $\mathbb{K}$-Vektorraum und $U \subset V$ ein Untervektorraum. Dann ist*

$$x \sim y \ :\Longleftrightarrow\ x - y \in U$$

eine Äquivalenzrelation auf V.

BEWEIS:

Ä1 (reflexiv): Für jedes $x \in V$ gilt $x - x = 0 \in U$, da der Nullvektor in jedem Untervektorraum enthalten ist; also nach der Definition von $\sim$: $x \sim x$.

Ä2 (symmetrisch): Für $x, y \in V$ gelte $x \sim y$, d.h. $x - y \in U$. Da U ein Untervektorraum ist, liegt auch $-(x-y) = y - x$ in U; also $y \sim x$.

Ä3 (transitiv): Für $x, y, z \in V$ gelte $x \sim y$ und $y \sim z$, d.h. $x - y$ und $y - z$ liegen in U. Da U ein Untervektorraum ist, liegt auch die Summe $(x-y)+(y-z) = x - z$ in U; also $x \sim z$. ■

Definition 8.27 Sei U ein UVR eines Vektorraumes V. Die Menge der Äquivalenzklassen von V bezüglich der durch U definierten Äquivalenzrelation $\sim$ heißt **Faktorraum** oder **Quotientenraum** und wird mit V/U bezeichnet.

Die Elemente $\tilde{x} \in V/U$ haben die folgende Form:

$$\tilde{x} = \{x + u \mid u \in U\} =: x + U.$$

Insbesondere ist $\tilde{0} = U$.

Die Bezeichnung Faktor*raum* wird gerechtfertigt durch folgenden Satz.

Satz 8.28 *Sei V ein $\mathbb{K}$-Vektorraum und $U \subset V$ ein Untervektorraum. Dann ist der Faktorraum V/U ein $\mathbb{K}$-Vektorraum mit der folgenden Addition und skalaren Multiplikation*

$$\tilde{x} + \tilde{y} := \widetilde{x+y} \quad \textit{bzw.} \quad \lambda \cdot \tilde{x} := \widetilde{\lambda \cdot x} \qquad \textit{für } x, y \in V \textit{ und } \lambda \in \mathbb{K}.$$

Beweis: Die Addition ist wohldefiniert, da die Definition nicht von der Wahl der Repräsentanten x von $\tilde{x}$ und y von $\tilde{y}$ abhängt: seien etwa $x \sim x'$ und $y \sim y'$. Dann gibt es nach Definition von $\sim$ Vektoren $u_1, u_2 \in U$ mit $x - x' = u_1$ und $y - y' = u_2$ und man erhält

$$(x+y) - (x'+y') = (x - x') + (y - y') = u_1 + u_2 \in U$$

da ja mit u_1 und u_2 auch $u_1 + u_2$ in U liegt. Es gilt also $(x+y) \sim (x'+y')$ und damit

$$\tilde{x} + \tilde{y} = \widetilde{x+y} = \widetilde{x'+y'} = \tilde{x}' + \tilde{y}'.$$

Die $\mathbb{K}$-Multiplikation ist ebenfalls wohldefiniert: seien $\lambda \in \mathbb{K}$ und $x, x' \in V$ mit $x \sim x'$. Dann ist mit $x - x' \in U$ auch $\lambda \cdot (x - x') = \lambda \cdot x - \lambda \cdot x' \in U$ und somit $\lambda \cdot x \sim \lambda \cdot x'$. Also gilt

$$\lambda \cdot \tilde{x} = \widetilde{\lambda \cdot x} = \widetilde{\lambda \cdot x'} = \lambda \cdot \tilde{x}'.$$

Die Kommutativität und Assoziativität der Addition überträgt sich von V auf V/U, z.B. gilt

$$\tilde{x} + \tilde{y} = \widetilde{x+y} = \widetilde{y+x} = \tilde{y} + \tilde{x} \quad \text{(entsprechend für die Assoziativität).}$$

Neutrales Element ist $\tilde{0} = \{0 + u \mid u \in U\} = U$ und es gilt $-\tilde{x} = \widetilde{(-x)}$. Die Vektorraum-Eigenschaften V2 übertragen sich ebenfalls von V auf V/U. ■

Beispiel 8.29 In $V = \mathbb{R}^2$ sei der Unterraum $U = [(1,2)]$ gegeben. Dann ist die Klasse $\tilde{x}$ von $x = (x_1, x_2) \in \mathbb{R}^2$ gegeben durch $x + U = \{(x_1, x_2) + \lambda(1,2) \mid \lambda \in \mathbb{R}\}$ gegeben. Geometrisch kann man sich $\tilde{x}$ also als eine Gerade in der Ebene $\mathbb{R}^2$ durch den Punkt $x = (x_1, x_2)$ mit Richtungsvektor $(1,2)$ vorstellen.

Die Summe zweier Klassen $\tilde{x}, \tilde{y}$ in V/U ergibt die Klasse $\widetilde{(x+y)} + U$, also die Gerade mit Richtungsvektor $(1,2)$ durch den Punkt $x+y$. Die Klasse $\widetilde{\lambda x}$ mit $\lambda \in \mathbb{R}$ ist dann gegeben durch die Gerade mit Richtungsvektor $(1,2)$ durch den Punkt λx.

Bemerkung 8.30 Seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum und U ein Untervektorraum von V mit Basis $\{b_1, \ldots, b_d\}$, $0 < d < n$. Weiter sei $\{b_1, \ldots, b_d, b_{d+1}, \ldots, b_n\}$ eine Basis von V. Es gilt dann:

1. $\{\tilde{b}_{d+1}, \ldots, \tilde{b}_n\}$ ist eine Basis des Faktorraums V/U.
2. Es gilt der **Dimensionssatz** $\dim V/U = \dim V - \dim U$.

Teil IV

Lineare Abbildungen und Matrizen

9 Lineare Abbildungen

In diesem Kapitel untersuchen wir Abbildungen zwischen Vektorräumen, die der Vektorraumstruktur besonders gut angepasst sind.

9.1 Definition und Beispiele

Definition 9.1 V und W seien $\mathbb{K}$-Vektorräume. Eine Abbildung $\Phi : V \to W$ heißt **linear**, wenn für alle $x, y \in V$ und alle $\lambda \in \mathbb{K}$ gilt

(**L1**) $\Phi(x + y) = \Phi(x) + \Phi(y)$

(**L2**) $\Phi(\lambda x) = \lambda\Phi(x)$.

Hilfssatz 9.2 *Bei einer linearen Abbildung $\Phi : V \to W$ geht der Nullvektor $0_V \in V$ in den Nullvektor $0_W \in W$ über.*

Beweis: Wegen (**L2**) folgt $\Phi(0_V) = \Phi(0 \cdot 0_V) = 0 \cdot \Phi(0_V) = 0_W$. ■

(**L1**) und (**L2**) lassen sich zu einer einzigen Linearitätseigenschaft (**L**) zusammenfassen:

Hilfssatz 9.3 (linear) *V und W seien $\mathbb{K}$-Vektorräume. Eine Abbildung $\Phi : V \to W$ ist genau dann linear, wenn für alle $x, y \in V$ und alle $\lambda, \mu \in \mathbb{K}$ gilt*

(**L**) $\Phi(\lambda x + \mu y) = \lambda\Phi(x) + \mu\Phi(y)$.

Beweis: (**L1**) folgt aus (**L**), wenn wir $\lambda = \mu = 1$ setzen, und (**L2**) folgt aus (**L**) wegen Hilfsatz 9.2, wenn wir $y = 0$ setzen. Umgekehrt folgt aus (**L1**) $\Phi(\lambda x + \mu y) = \Phi(\lambda x) + \Phi(\mu y)$ und aus (**L2**) $\Phi(\lambda x) + \Phi(\mu y) = \lambda\Phi(x) + \mu\Phi(y)$. ■

Definition 9.4 Eine lineare Abbildung ist „strukturerhaltend“ und heißt deshalb auch **(Vektorraum-)Homomorphismus**. Ein bijektiver Homomorphismus heißt **(Vektorraum-)Isomorphismus**. Gibt es solch einen Isomorphismus $\Phi : V \to W$, so nennt man die Vektorräume V, W **isomorph** und schreibt $V \cong W$.

Ist $V = W$, so nennt man eine lineare Abbildung $\Phi : V \to V$ auch **Selbstabbildung** oder **Endomorphismus** von V. Ein bijektiver Endomorphismus heißt **Automorphismus**.

Beispiele 9.5

1. Die Abbildung

$$\Phi : \begin{cases} \mathbb{R}^3 & \to & \mathbb{R}^4 \\ (x_1, x_2, x_3) & \mapsto & (x_2, x_1 + x_2, x_2 + x_3, x_3) \end{cases}$$

   ist linear und injektiv, aber nicht surjektiv.

2. Für beliebige $\mathbb{K}$-Vektorräume V, W ist die **Nullabbildung**

$$0 : \begin{cases} V & \to & W \\ v & \mapsto & 0 \end{cases}$$

   linear. Dagegen ist die **konstante Abbildung**

$$\Phi : \begin{cases} V & \to & W \\ v & \mapsto & w_0 \end{cases}$$

   für $w_0 \neq 0$ *nicht* linear.

3. Für einen $\mathbb{K}$-Vektorraum V ist die Abbildung

$$\Phi : \begin{cases} V & \to & V \\ v & \mapsto & \lambda v \end{cases}$$

   mit einem festen Parameter $\lambda \in \mathbb{K}$ linear und für $\lambda \neq 0$ sogar ein Automorphismus von V, die **Streckung** um den Faktor λ. Für $\lambda = 1$ erhält man die **Identität** (oder **identische Abbildung**) von V:

$$\mathrm{id}_V : \begin{cases} V & \to & V \\ v & \mapsto & \mathrm{id}_V(v) = v. \end{cases}$$

   Hingegen ist die **Translation** um einen Vektor $v_0 \neq 0$

$$\Phi : \begin{cases} V & \to & V \\ v & \mapsto & v + v_0 \end{cases}$$

   nicht linear.

4. Ein besonders wichtiges Beispiel: Zu jeder Matrix $A \in \mathbb{K}^{m\times n}$ gehört eine lineare Abbildung $\Phi : \mathbb{K}^n \to \mathbb{K}^m$, $x \mapsto Ax$, wenn man $x \in \mathbb{K}^n$ als Spaltenvektor auffasst. Es gilt nämlich für alle $x, y \in \mathbb{K}^n$ und alle $\lambda \in \mathbb{K}$:

$$\Phi(x+y) = A(x+y) = Ax + Ay = \Phi(x) + \Phi(y)$$

   und

$$\Phi(\lambda x) = A(\lambda x) = \lambda Ax = \lambda\Phi(x).$$

   Das ist auch der Grund dafür, dass man Elemente von $\mathbb{K}^n$ (also n-Tuppel oder $(1 \times n)$-Matrizen) manchmal mit Spaltenvektoren (also $(n \times 1)$-Matrizen) identifiziert: Auf diese Weise kann man eine Matrix A mit n Spalten von links an ein Element $v \in \mathbb{K}^n$ multiplizieren und den Vektor v so in den Vektor Av abbilden.

5. Insbesondere gehört zu jedem linearen Gleichungssystem (3.5) eine lineare Abbildung, da man die Matrix $A \in \mathbb{K}^{m\times n}$ des LGS (3.6) nach dem vierten Beispiel als lineare Abbildung auffassen kann. Das LGS $Ax = b$ ist also genau dann lösbar, wenn $b \in \mathbb{K}^m$ in der Bildmenge der zu A gehörigen linearen Abbildung $\Phi : \mathbb{K}^n \to \mathbb{K}^m$, $x \mapsto Ax$ liegt.

6. Ist eine Matrix $A \in \mathbb{K}^{n\times n}$ invertierbar, so ist die zu A gehörige lineare Abbildung $\Phi : \mathbb{K}^n \to \mathbb{K}^n$, $x \mapsto Ax$ ein Isomorphismus. Die Abbildung $\Phi^{-1} : \mathbb{K}^n \to \mathbb{K}^n$, $y \mapsto A^{-1}y$ ist nämlich die Umkehrabbildung von Φ (und Φ somit bijektiv).

7. Hat man in einem n-dimensionalen Vektorraum V zwei Basen $B, \overline{B}$ gegeben, so ist die Transformation des Komponentenvektors $\Theta_B(x)$ eines Vektors $x \in V$ in den Komponentenvektor $\Theta_{\overline{B}}(x)$ ein Isomorphismus von $\mathbb{K}^n$ nach $\mathbb{K}^n$ nach Satz 7.20.

8. Sei $V = \mathbb{K}^{\mathbb{N}_0}$ der Vektorraum der Folgen $(a_0, a_1, a_2, \ldots)$ in $\mathbb{K}$. Dann ist der sogenannte *Shift-Operator*

$$\Phi : \mathbb{K}^{\mathbb{N}_0} \to \mathbb{K}^{\mathbb{N}_0}, \quad (a_0, a_1, a_2, a_3, \ldots) \mapsto (a_1, a_2, a_3, a_4, \ldots)$$

   eine lineare Abbildung.

9. Sei V der Vektorraum aller differenzierbaren Funktionen $f : \mathbb{R} \to \mathbb{R}$. Dann ist die Ableitung von Funktionen

$$\Phi : V \to V, \quad f \mapsto f'$$

   eine lineare Abbildung, denn es gilt ja für alle differenzierbaren Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ und alle $\lambda \in \mathbb{R}$ $(f+g)' = f' + g'$ und $(\lambda f)' = \lambda f'$ (vgl. Analysis-Vorlesung).

9.2 Erste Eigenschaften von linearen Abbildungen

Wir überlegen zuerst, dass eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen durch endlich viele Daten vollständig bestimmt ist: Man kennt eine lineare Abbildung $\Phi : V \to W$, wenn man die Bilder einer Basis von V kennt. V und W seien im Folgenden wieder gegebene Vektorräume über demselben Körper.

Satz 9.6 (lineare Fortsetzung) *Sei V ein $\mathbb{K}$-Vektorraum und $\{v_1, \dots, v_n\}$ eine Basis von V. Weiter seien $w_1, \dots, w_n$ beliebig vorgegebene Vektoren eines $\mathbb{K}$-Vektorraumes W. Dann gibt es genau eine lineare Abbildung $\Phi : V \to W$ mit*

$$\Phi(v_i) = w_i, \qquad i = 1, 2, \dots, n. \qquad\qquad (*)$$

Beweis: Um die *Existenz* einer solchen linearen Abbildung zu zeigen, benutzen wir die eindeutige Basisdarstellung eines beliebigen Vektors $x \in V$:

$$x = \sum_{i=1}^{n} x_i v_i.$$

Die gesuchte Abbildung Φ soll linear sein, d.h. es soll gelten

$$\Phi(x) = \Phi\Big(\sum_{i=1}^{n} x_i v_i\Big) = \sum_{i=1}^{n} x_i \Phi(v_i).$$

Wir definieren also einfach

$$\Phi(x) := \sum_{i=1}^{n} x_i w_i,$$

denn die Bilder der Basisvektoren sind ja festgelegt: $\Phi(v_i) = w_i$. Die so definierte Abbildung ist linear und erfüllt $(*)$.

Wir beweisen jetzt die *Eindeutigkeit* von Φ: Ist Ψ eine weitere lineare Abbildung mit der geforderten Eigenschaft $(*)$, so gilt für $x \in V$

$$\begin{aligned}
\Psi(x) &= \Psi\Big(\sum_{i=1}^{n} x_i v_i\Big) \\
&= \sum_{i=1}^{n} x_i \Psi(v_i) \qquad \text{(Linearität von } \Psi) \\
&= \sum_{i=1}^{n} x_i w_i \qquad \text{(nach } (*)) \\
&= \Phi(x).
\end{aligned}$$

Da x beliebig ist, stimmen Φ und Ψ überein. ■

Wir untersuchen nun, wie sich linear abhängige Vektoren bei einer linearen Abbildung verhalten.

Hilfssatz 9.7 *Bei einer linearen Abbildung $\Phi : V \to W$ gehen linear abhängige Vektoren $v_1, \ldots, v_k \in V$ in linear abhängige Vektoren $\Phi(v_1), \ldots, \Phi(v_k) \in W$ über.*

Beweis: Ist $\sum_{i=1}^k a_i v_i = 0$ eine nichttriviale Darstellung des Nullvektors, so folgt daraus $\Phi(\sum_{i=1}^k a_i v_i) = \Phi(0)$, also $\sum_{i=1}^k a_i \Phi(v_i) = 0$ wegen der Linearität von Φ und nach Hilfssatz 9.2. Da nicht alle a_i Null sind, sind die $\Phi(v_i)$ linear abhängig. ■

Dagegen können linear unabhängige Vektoren bei einer linearen Abbildung eventuell auch in linear abhängige Vektoren übergehen. Das ist z.B. bei der Nullabbildung in Beispiel 9.5 der Fall. Es gilt aber folgender

Hilfssatz 9.8 *Bei einer* injektiven *linearen Abbildung $\Phi : V \to W$ gehen linear unabhängige Vektoren $v_1, \ldots, v_k \in V$ in linear unabhängige Vektoren $\Phi(v_1), \ldots, \Phi(v_k) \in W$ über.*

Beweis: Angenommen $\Phi(v_1), \ldots, \Phi(v_k)$ sind linear abhängig. Dann gibt es eine nichttriviale Darstellung $\sum_{i=1}^k a_i \Phi(v_i) = 0$ des Nullvektors in W, woraus dann $\Phi(\sum_{i=1}^k a_i v_i) = \Phi(0)$ folgt. Wegen der Injektivität von Φ ergibt sich daraus $\sum_{i=1}^k a_i v_i = 0$, also eine nichttriviale Darstellung des Nullvektors in V. Das ist ein Widerspruch, da die v_i linear unabhängig sind. ■

9.3 Kern und Bild einer linearen Abbildung

Auch in diesem Abschnitt sind V und W stets wieder Vektorräume über demselben Körper $\mathbb{K}$.

9.3.1 Wann ist eine lineare Abbildung injektiv?

Sei $\Phi : V \to W$ eine lineare Abbildung. Nach Definition ist Φ injektiv, wenn für alle $x, y \in V$ gilt

$$\Phi(x) = \Phi(y) \Longrightarrow x = y.$$

Da Φ linear ist können wir diese Implikation auch schreiben als

$$\Phi(x - y) = \Phi(x) - \Phi(y) = 0 \Longrightarrow x - y = 0.$$

Diese Überlegung motiviert die folgende

Definition 9.9 Sei $\Phi : V \to W$ eine lineare Abbildung. Der **Kern** von Φ ist die Menge aller Vektoren von V, die durch Φ auf den Nullvektor $0 \in W$ abgebildet werden, also

$$\operatorname{Kern} \Phi := \{v \in V \mid \Phi(v) = 0\}.$$

Hilfssatz 9.10 *Der Kern einer linearen Abbildung $\Phi : V \to W$ ist ein Untervektorraum von V.*

Beweis: Seien $x, y \in \operatorname{Kern}\Phi$ und $\lambda, \mu \in \mathbb{K}$ beliebig gewählt. Dann ist auch $\lambda x + \mu y \in \operatorname{Kern}\Phi$, denn $\Phi(\lambda x + \mu y) = \lambda\Phi(x) + \mu\Phi(y) = \lambda 0 + \mu 0 = 0$. Wegen $\Phi(0) = 0$ ist $\operatorname{Kern}\Phi$ nicht leer. Also ist $\operatorname{Kern}\Phi$ nach dem Untervektorraumkriterium 8.2 ein Untervektorraum von V. ■

Satz 9.11 (Kriterium für injektiv) *Eine lineare Abbildung $\Phi : V \to W$ ist genau dann injektiv, wenn* $\operatorname{Kern}\Phi = \{0\} \subset V$ *ist.*

Beweis:

„⇒“ Sei Φ injektiv und $x \in \operatorname{Kern}\,\Phi$. Aus $\Phi(x) = 0 = \Phi(0)$ folgt dann $x = 0$, d.h. $\operatorname{Kern}\,\Phi = \{0\}$.

„⇐“ Sei jetzt $\operatorname{Kern}\,\Phi = \{0\}$. Aus $\Phi(x) = \Phi(y)$ folgt $\Phi(x-y) = 0$, also $x-y \in \operatorname{Kern}\,\Phi$ und damit $x = y$. ■

9.3.2 Wann ist eine lineare Abbildung surjektiv?

Diese Frage ist noch einfacher zu beantworten: Nach Definition ist Φ surjektiv, wenn die Bildmenge $\operatorname{Bild}\Phi = \Phi(V) \subset W$ gleich W ist. Wie der Kern ist auch die Bildmenge ist ein UVR:

Hilfssatz 9.12 *Ist $\Phi : V \to W$ eine lineare Abbildung, so ist die Bildmenge $\Phi(V)$ ein Untervektorraum des Zielraumes W.*

Beweis: Wegen $0 = \Phi(0) \in \Phi(V)$ (Hilfssatz 9.2) ist $\Phi(V)$ nicht leer. Wenn $w_1, w_2 \in \Phi(V)$ Urbilder $v_1, v_2 \in V$ haben, so gilt für alle $\lambda \in \mathbb{K}$

$$\begin{aligned} w_1 + w_2 &= \Phi(v_1) + \Phi(v_2) = \Phi(v_1 + v_2) \in \Phi(V), \\ \lambda w_1 &= \lambda\Phi(v_1) = \Phi(\lambda v_1) \in \Phi(V). \end{aligned}$$

Also ist $\Phi(V)$ nach Hilfssatz 8.2 ein Untervektorraum von W. ■

Wegen dieses Hilfssatzs nennen wir die Bildmenge $\operatorname{Bild}\Phi = \Phi(V)$ genauer auch **Bildraum** von Φ.

Mit einem analogen Argument folgt übrigens auch, dass jeder Untervektorraum von V in einen Untervektorraum von $\Phi(V)$ übergeht (eine lineare Abbildung soll ja auch strukturerhaltend sein).

9.3.3 Der Zusammenhang zwischen Kern und Bild

Wir fragen jetzt allgemeiner nach allen Vektoren aus V, die bei der linearen Abbildung $\Phi : V \to W$ *dasselbe Bild* $w \in W$ haben, betrachten also die **Menge aller Urbilder**

$$\Phi^{-1}(w) := \{x \in V \mid \Phi(x) = w\}.$$

Für $w = 0$ ist $\Phi^{-1}(0) = \operatorname{Kern}\Phi$, also ein Untervektorraum von V. Für $w \neq 0$ ist $\Phi^{-1}(w)$ kein Untervektorraum wegen $0 \notin \Phi^{-1}(w)$.

Für $w \in W$, $w \notin \Phi(V)$ ist $\Phi^{-1}(w) = \emptyset$. Wenn wir uns aber auf Vektoren $w \in \Phi(V)$ beschränken, sind die Mengen $\Phi^{-1}(w)$ nichtleer und disjunkt, und ihre Vereinigungsmenge ist ganz V; sie bilden also die Klassen einer Äquivalenzklassen-Einteilung von V. Diese Klasseneinteilung bzw. die zugehörige Äquivalenzrelation sind uns bereits begegnet: die Klassen sind Elemente des Faktorraums V/U mit $U = \operatorname{Kern}\Phi$. Es gilt nämlich

Hilfssatz 9.13 *Sei $\Phi : V \to W$ eine lineare Abbildung.*

(i) *Zwei Vektoren $x, y \in V$ haben genau dann dasselbe Bild unter Φ, wenn $x - y \in \operatorname{Kern}\Phi$.*

(ii) *Die Menge aller Urbilder von $w \in \operatorname{Bild}\Phi$ ist eine Nebenklasse im Faktorraum $V/\operatorname{Kern}\Phi$:*

$$\Phi^{-1}(\{w\}) = x + \operatorname{Kern}\Phi \quad \textit{für ein}\, x \in V, \textit{für das gilt } \Phi(x) = w.$$

Beweis: (i) Falls $\Phi(x) = \Phi(y)$ so folgt aus der Linearität

$$0 = \Phi(x) - \Phi(y) = \Phi(x - y).$$

Es gilt also $x - y \in \operatorname{Kern}\Phi$. Diese Argumentation kann man umkehren.

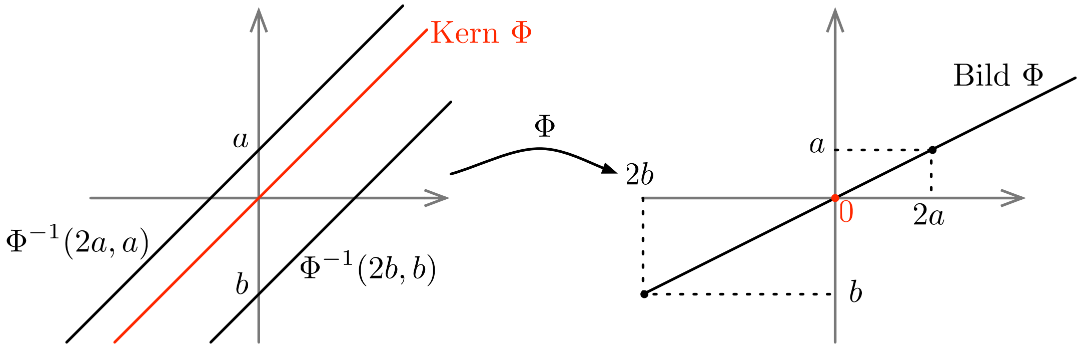
(ii) Nach Hilfssatz 9.10 ist $\operatorname{Kern}\Phi$ ein Untervektorraum; man kann also den Faktorraum $V/\operatorname{Kern}\Phi$ bilden. Ist $x \in \Phi^{-1}(w)$ ein Urbild von $w \in \Phi(V)$, so ist $x + \operatorname{Kern}\Phi = \tilde{x} \in V/\operatorname{Kern}\Phi$ die Menge *aller* Urbilder von w, da sich die Elemente von $\Phi^{-1}(w)$ nach (i) nur um einen Vektor in $\operatorname{Kern}\Phi$ unterscheiden. ■

Beispiel 9.14 Wir betrachten die lineare Abbildung

$$\Phi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -2x_1 + 2x_2 \\ -x_1 + x_2 \end{pmatrix}.$$

Das Bild von Φ ist die Gerade, die vom Vektor $(2, 1)$ aufgespannt wird. Der Kern von Φ ist die Gerade, die von $(1, 1)$ aufgespannt wird,

$$\operatorname{Kern}\,\Phi = \left\{\lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} \,\middle|\, \lambda \in \mathbb{R}\right\}.$$



Wegen $\Phi(-a, 0) = (2a, a)$ ist das Urbild eines Elementes $(2a, a) \in \text{Bild } \Phi$ die affine Gerade

$$\Phi^{-1}\begin{pmatrix} 2a \\ a \end{pmatrix} = \begin{pmatrix} -a \\ 0 \end{pmatrix} + \text{Kern } \Phi = \left\{ \begin{pmatrix} \lambda - a \\ \lambda \end{pmatrix} \;\middle|\; \lambda \in \mathbb{R} \right\}.$$

Die Urbilder verschiedener Elemente von Bild Φ sind parallele Geraden.

Eine lineare Abbildung $\Phi : V \to W$ ist im Allgemeinen weder injektiv noch surjektiv. Wir werden als nächstes sehen, dass man Φ stets als Verkettung $\Phi = \tilde{\Phi} \circ \pi$ einer injektiven linearen Abbildung $\tilde{\Phi}$ und einer surjektiven linearen Abbildung π schreiben kann.

Definition 9.15 Sei $\Phi : V \to W$ eine lineare Abbildung. Als **kanonische Projektion** (zu Φ) bezeichnet man die Abbildung

$$\pi : V \to V/\text{Kern}\,\Phi, \quad x \mapsto \tilde{x},$$

die jedem Vektor v seine Äquivalenzklasse in $V/\text{Kern}\,\Phi$ zuordnet.

Bemerkung 9.16 Wegen der Wohldefiniertheit der Addition und $\mathbb{K}$-Multiplikation im Faktorraum $V/\text{Kern}\,\Phi$ ist π eine lineare Abbildung. π ist surjektiv, da jede Äquivalenzklasse $a \in V/\text{Kern}\,\Phi$ mindestens einen Repräsentanten x hat, es gilt also $a = \tilde{x} = \pi(x)$ für mindestens ein $x \in V$.

Satz 9.17 (Homomorphiesatz) *Es sei $\Phi : V \to W$ eine lineare Abbildung. Dann ist*

$$\tilde{\Phi} : V/\text{Kern}\,\Phi \to W, \quad \tilde{x} \mapsto \Phi(x)$$

eine injektive lineare Abbildung und es gilt $\Phi = \tilde{\Phi} \circ \pi$.

$$\begin{tikzcd} V \arrow[r, "\Phi"] \arrow[d, "\pi"'] & W \\ V/\text{Kern } \Phi \arrow[ur, "\tilde{\Phi}"'] & \end{tikzcd}$$

Beweis: Zunächst zeigen wir, dass die Abbildung $\tilde{\Phi}$ wohldefiniert, d.h. unabhängig von der Wahl der Repräsentanten, ist. Seien dazu $x, y \in V$ mit $\tilde{x} = \tilde{y}$, d.h. es gilt nach Definition des Faktorraumes (bzw. der zugrundeliegenden Äquivalenzrelation) $x \sim y$ bzw. $x - y \in \operatorname{Kern}\Phi$. Daraus folgt mit Hilfssatz 9.13

$$\tilde{\Phi}(\tilde{x}) = \Phi(x) = \Phi(y) = \tilde{\Phi}(\tilde{y}).$$

Zum Nachweis der Linearität seien $\tilde{x}, \tilde{y} \in V/\operatorname{Kern}\Phi$ und $\lambda \in \mathbb{K}$ beliebig vorgegeben. Nach Definition der Addition im Faktorraum und weil Φ linear ist, gilt

$$\tilde{\Phi}(\tilde{x} + \tilde{y}) = \tilde{\Phi}(\widetilde{x+y}) = \Phi(x+y) = \Phi(x) + \Phi(y) = \tilde{\Phi}(\tilde{x}) + \tilde{\Phi}(\tilde{y})$$

und

$$\tilde{\Phi}(\lambda\tilde{x}) = \tilde{\Phi}(\widetilde{\lambda x}) = \Phi(\lambda x) = \lambda\Phi(x) = \lambda\tilde{\Phi}(\tilde{x}).$$

$\tilde{\Phi}$ ist auch injektiv: seien dazu $\tilde{x}, \tilde{y} \in V/\operatorname{Kern}\Phi$ mit $\tilde{\Phi}(\tilde{x}) = \tilde{\Phi}(\tilde{y})$ vorgegeben. Nach Definition von $\tilde{\Phi}$ gilt dann $\Phi(x) = \Phi(y)$ und somit nach Hilfssatz 9.13 $x-y \in \operatorname{Kern}\Phi$, was gleichbedeutend ist mit $x \sim y$ bzw. mit $\tilde{x} = \tilde{y}$.

Zum Nachweis von $\Phi = \tilde{\Phi} \circ \pi$ sei $x \in V$ beliebig. Man erhält

$$(\tilde{\Phi} \circ \pi)(x) = \tilde{\Phi}(\pi(x)) = \tilde{\Phi}(\tilde{x}) = \Phi(x).$$

■

Folgerung 9.18 (Isomorphiesatz) *Ist $\Phi : V \to W$ eine surjektive lineare Abbildung, so ist auch $\tilde{\Phi}$ surjektiv und somit ein Isomorphismus von $V/\operatorname{Kern}\Phi$ nach W. Insbesondere sind die Vektorräume $V/\operatorname{Kern}\Phi$ und $\operatorname{Bild}\Phi$ isomorph.*

9.3.4 Rang einer linearen Abbildung

Es sei V ein n-dimensionaler Vektorraum. Nach Hilfssatz 9.12 ist die Bildmenge $\Phi(V) = \operatorname{Bild}\Phi$ einer linearen Abbildung $\Phi : V \to W$ ein Untervektorraum von W. Jedem UVR können wir eine Zahl zuordnen: seine Dimension. Wir können also definieren:

Definition 9.19 Der **Rang** einer linearen Abbildung $\Phi : V \to W$ ist die Dimension des Bildraumes von Φ, also

$$\operatorname{Rang}\Phi := \dim\Phi(V) = \dim\operatorname{Bild}\Phi.$$

Bemerkung 9.20 (Schranken für Rang) Als Untervektorraum von W hat $\operatorname{Bild}\Phi$ höchstens die Dimension von W, das heißt

$$\operatorname{Rang}\Phi \leq \dim W.$$

Es gilt aber auch
$$\text{Rang } \Phi \leq \dim V.$$

Beweis: Sei $n := \dim V$. Annahme: Rang $\Phi > n$. Dann gibt es im Bild $\Phi(V)$ mindestens $n+1$ linear unabhängige Vektoren $w_1, \ldots, w_{n+1}$ mit Urbildern $v_1, \ldots, v_{n+1} \in V$. Die $n+1$ Vektoren $v_i \in V$ sind linear abhängig, also nach Hilfssatz 9.7 auch ihre Bildvektoren w_i; ein Widerspruch. ■

Neben den eben angegebenen Ungleichungen besteht noch eine wichtige Gleichung, die den Rang von Φ mit der Dimension von Kern Φ in Beziehung setzt.

Satz 9.21 *Für eine lineare Abbildung* $\Phi : V \to W$ *gilt*
$$\text{Rang}\,\Phi = \dim V - \dim \text{Kern}\,\Phi.$$

Anders ausgedrückt: Die Summe der Dimensionen von Kern und Bild von Φ ist die Dimension des Ausgangsraumes.

Beweis: Es sei $n = \dim V$.

Für $n = 0$, also $V = \{0\}$, ist Φ die Nullabbildung und somit Rang $\Phi = 0$ und Kern $\Phi = V = \{0\}$. Also gilt die Behauptung. Auch im Fall $\dim \text{Kern}\,\Phi = n$ ist Φ die Nullabbildung und somit Rang $\Phi = 0$. Sei also im Folgenden $n \geq 1$ und $d := \dim \text{Kern}\,\Phi < n$.

1. Fall: Es sei $d > 0$. Wir wählen eine Basis $\{v_1, \ldots, v_d\}$ des Kerns. Da $d < n$ können wir diese nach Satz 7.5 zu einer Basis $\{v_1, \ldots, v_d, v_{d+1}, \ldots, v_n\}$ von V ergänzen. Ist w ein beliebiger Vektor aus $\Phi(V)$ und $x \in V$ ein Urbildvektor von w, so können wir schreiben
$$x = x_1 v_1 + \cdots + x_d v_d + x_{d+1} v_{d+1} + \cdots x_n v_n$$
und erhalten daraus
$$w = \Phi(x) = 0 + \cdots + 0 + x_{d+1}\Phi(v_{d+1}) + \cdots + x_n \Phi(v_n).$$

Die Vektoren $\Phi(v_{d+1}), \ldots, \Phi(v_n)$ erzeugen also das Bild $\Phi(V)$. Wir wollen noch überlegen, dass sie auch linear unabhängig sind. Aus einer Vektorgleichung
$$a_{d+1}\Phi(v_{d+1}) + \cdots + a_n \Phi(v_n) = 0$$
folgt
$$\Phi(a_{d+1} v_{d+1} + \cdots + a_n v_n) = 0,$$
also liegt $a_{d+1} v_{d+1} + \cdots + a_n v_n$ im Kern von Φ, d.h. es gilt
$$a_{d+1} v_{d+1} + \cdots + a_n v_n = b_1 v_1 + \cdots + b_d v_d \quad (\text{mit } b_i \in \mathbb{K})$$

$$\text{oder, äquivalent,} \quad b_1 v_1 + \cdots + b_d v_d - a_{d+1} v_{d+1} - \cdots - a_n v_n = 0.$$

Da $\{v_1, \ldots, v_d, v_{d+1}, \ldots, v_n\}$ eine Basis ist, folgt hieraus insbesondere $a_{d+1} = \cdots = a_n = 0$.

Somit ist $\{\Phi(v_{d+1}), \ldots, \Phi(v_n)\}$ nach Satz 7.4 eine Basis von $\Phi(V)$, und es gilt

$$\operatorname{Rang}\Phi = \dim\, \Phi(V) = n - d = \dim\, V - \dim \operatorname{Kern}\Phi.$$

2. Fall: Sei jetzt $d = \dim\, \operatorname{Kern}\Phi = 0$. In diesem Fall wählt man im obigen Beweis $\{v_1, \ldots, v_n\}$ als Basis von V und setzt jeweils $d = 0$ bzw. lässt die entsprechenden v_i weg. Der Beweis gilt dann sinngemäß auch in diesem Fall. ■

Als Anwendung geben wir jetzt noch Charakterisierungen von injektiven, surjektiven und bijektiven linearen Abbildungen.

Satz 9.22 (i) *Eine lineare Abbildung $\Phi : V \to W$ ist genau dann injektiv, wenn* $\operatorname{Rang}\Phi = \dim V$.

(ii) *Eine lineare Abbildung $\Phi : V \to W$ ist genau dann surjektiv, wenn* $\operatorname{Rang}\Phi = \dim W$.

Beweis: (i) Nach Satz 9.11 ist Φ genau dann injektiv, wenn $\operatorname{Kern}\, \Phi = \{0\}$, also $\dim \operatorname{Kern}\, \Phi = 0$. Die Behauptung folgt also aus Satz 9.21.

(ii) Nach Satz 8.15 gilt $\operatorname{Rang}\, \Phi = \dim\, W$ genau dann, wenn $\Phi(V) = W$ ist, also wenn Φ surjektiv ist. ■

Satz 9.23 *Zwei endlichdimensionale Vektorräume V, W über $\mathbb{K}$ sind genau dann isomorph, wenn sie gleiche Dimension haben.*

Beweis:

„$\Rightarrow$“ Sind V, W isomorph, so gibt es eine bijektive lineare Abbildung $\Phi : V \to W$, und aus Satz 9.22 folgt $\dim V = \operatorname{Rang}\Phi = \dim W$.

„$\Leftarrow$“ Sei jetzt $\dim V = \dim W := n$. Für $n = 0$ ist $V = W = \{0\}$, der Satz also trivial. Für $n \geq 1$ seien Basen $\{v_1, \ldots, v_n\}$ bzw. $\{w_1, \ldots, w_n\}$ von V bzw. W gewählt. Nach Satz 9.6 gibt es genau eine lineare Abbildung $\Phi : V \to W$, für die $\Phi(v_i) = w_i, i = 1, \ldots, n$, gilt. Wegen $w_i \in \Phi(V)$ ist $\operatorname{Rang}\, \Phi = \dim \Phi(V) = \dim W = n$. Daraus folgt wieder mit Satz 9.22, dass Φ injektiv und surjektiv, also ein Isomorphismus ist. ■

Beispiel 9.24 Jeder $\mathbb{K}$-Vektorraum V mit $\dim V = n$ ist also zu $\mathbb{K}^n$ isomorph. Wir geben einen expliziten Isomorphismus an. Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V.

Die Abbildung

$$\Theta_B : V \longrightarrow \mathbb{K}^n; \quad v = \sum_{i=1}^{n} \lambda_i v_i \longmapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

die jedem Vektor seinen Komponentenvektor bezüglich der Basis B (siehe Definition 7.16) zuordnet, ist linear, injektiv und surjektiv, also ein Vektorraum-Isomorphismus. Ist $E = \{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{K}^n$, also

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

so ist also Θ_B die eindeutige lineare Abbildung $V \to \mathbb{K}^n$ mit $\Theta_B(v_i) = e_i$ für $i = 1, \ldots, n$ (vgl. Satz 9.6).

Bemerkung 9.25 (a) „Isomorph sein“ ist eine Äquivalenzrelation auf der Menge aller endlichdimensionalen $\mathbb{K}$-Vektorräume. Eine Äquivalenzklasse besteht nach Satz 9.23 jeweils aus allen Vektorräumen mit gleicher Dimension n. Ein Repräsentant für jede Klasse mit $n \geq 1$ ist der Standardraum $\mathbb{K}^n$ $(n = 1, 2, \ldots)$.

(b) Es sei $\Phi : V \to W$ eine lineare Abbildung und V endlichdimensional. Dann ist der Faktorraum $V/\mathrm{Kern}\,\Phi$ zum Bildraum $\Phi(V)$ isomorph nach Folgerung 9.18. Aus Satz 9.23 und Satz 9.21 folgt dann:

$$\dim(V/\mathrm{Kern}\ \Phi) = \dim \Phi(V) = \mathrm{Rang}\,\Phi = \dim V - \dim \mathrm{Kern}\ \Phi.$$

9.4 Der Vektorraum Hom(V, W)

In diesem Abschnitt betrachten wir die Menge Hom(V, W) *aller* linearen Abbildungen $\Phi : V \to W$ für Vektorräume V, W über $\mathbb{K}$. Wir werden sehen, dass diese Menge selbst wieder ein Vektorraum ist. Ein Spezialfall ist Hom$(V, \mathbb{K})$, der sogenannte Dualraum von V.

Die Addition und $\mathbb{K}$-Multiplikation in Hom(V, W) ist „punktweise“ definiert:

Definition 9.26 Es seien $\Phi, \Psi \in \mathrm{Hom}(V, W)$ und $\lambda \in \mathbb{K}$. Die **Summe**[2] von Φ und Ψ ist die Abbildung

$$\Phi + \Psi : \begin{cases} V & \to & W \\ v & \mapsto & (\Phi + \Psi)(v) := \Phi(v) + \Psi(v). \end{cases}$$

[2] Man beachte die unterschiedliche Bedeutung des Zeichens $+$ in $\Phi + \Psi$ bzw. in $\Phi(v) + \Psi(v)$

Das λ**-fache**[3] von Φ ist die Abbildung

$$\lambda\Phi : \begin{cases} V & \rightarrow & W \\ v & \mapsto & (\lambda\Phi)(v) := \lambda\Phi(v). \end{cases}$$

Es gilt dann

Hilfssatz 9.27 *Es seien Φ, Ψ lineare Abbildungen von V nach W. Dann sind auch $\Phi + \Psi$ und $\lambda\Phi$ (für $\lambda \in \mathbb{K}$) lineare Abbildungen von V nach W.*

Beweis: Für alle $u, v \in V$ und $\alpha, \beta \in \mathbb{K}$ gilt nach obiger Definition und wegen der Linearität von Φ und Ψ:

$$\begin{aligned} (\Phi + \Psi)(\alpha u + bv) &= \Phi(\alpha u + \beta v) + \Psi(\alpha u + \beta v) \\ &= \alpha\Phi(u) + \beta\Phi(v) + \alpha\Psi(u) + \beta\Psi(v) \\ &= \alpha(\Phi(u) + \Psi(u)) + \beta(\Phi(v) + \Psi(v)) \\ &= \alpha(\Phi + \Psi)(u) + \beta(\Phi + \Psi)(v). \end{aligned}$$

Entsprechend gilt

$$\begin{aligned} (\lambda\Phi)(\alpha u + \beta v) &= \lambda\Phi(\alpha u + \beta v) \\ &= \lambda(\alpha\Phi(u) + \beta\Phi(v)) \\ &= \alpha\lambda\Phi(u) + \beta\lambda\Phi(v) \\ &= \alpha(\lambda\Phi)(u) + \beta(\lambda\Phi)(v). \end{aligned}$$

■

Satz 9.28 $\mathrm{Hom}(V, W)$ *ist bezüglich der erklärten Addition und $\mathbb{K}$-Multiplikation ein Vektorraum über $\mathbb{K}$.*

Beweis: Dass $(\mathrm{Hom}(V, W), +)$ eine abelsche Gruppe ist und dass die Eigenschaften V2 gelten, verifiziert man durch direktes Nachrechnen. Neutrales Element in $(\mathrm{Hom}(V, W), +)$ ist die Nullabbildung $0 : V \rightarrow W, v \mapsto 0$. ■

Wir nehmen jetzt an, dass V und W endlichdimensional sind. Es zeigt sich, dass dann auch der Vektorraum $\mathrm{Hom}(V, W)$ endlichdimensional ist und dass seine Dimension in einfacher Weise mit den Dimensionen von V und W zusammenhängt.

Satz 9.29 *Es seien V und W endlichdimensionale Vektorräume über $\mathbb{K}$. Dann ist auch $\mathrm{Hom}(V, W)$ endlichdimensional, und es gilt*

$$\dim \mathrm{Hom}(V, W) = \dim V \cdot \dim W.$$

[3] Man beachte wieder die unterschiedliche Bedeutung der skalaren Multiplikation in $\lambda\Phi$ bzw. in $\lambda\Phi(v)$

Beweis: Sei $\dim V = n$, $\dim W = m$. Für $n = 0$ oder $m = 0$ ist $\mathrm{Hom}(V,W) = \{\text{Null-Abb}\}$ also auch $\dim \mathrm{Hom}(V,W) = 0$. Für $n > 0$ und $m > 0$ seien $\{v_1, \ldots, v_n\}$ bzw. $\{w_1, \ldots, w_m\}$ Basen von V bzw. von W. Wir werden nun $n \cdot m$ geeignete lineare Abbildungen von V nach W angeben und nachweisen, dass sie eine Basis von $\mathrm{Hom}(V,W)$ bilden; damit ist dann $\dim \mathrm{Hom}(V,W) = n \cdot m$ gezeigt.

Wir definieren $n \cdot m$ lineare Abbildungen Φ_{ij} auf der Basis $\{v_1, \ldots, v_n\}$ durch

$$\Phi_{ij}(v_k) := \delta_{jk} w_i \qquad (1 \leq i \leq m,\ 1 \leq j, k \leq n)$$

und dann auf ganz V durch lineare Fortsetzung nach Satz 9.6. Φ_{ij} bildet also den Basisvektor v_j auf w_i ab, die übrigen v_k $(k \neq j)$ auf den Nullvektor 0. Wir zeigen jetzt, dass die Φ_{ij} linear unabhängig sind und $\mathrm{Hom}(V,W)$ erzeugen, also eine Basis von $\mathrm{Hom}(V,W)$ bilden:

- Die Φ_{ij} sind linear unabhängig: Dazu sei

  $$\sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} \Phi_{ij} = O \qquad (\lambda_{ij} \in \mathbb{K})$$

  eine Darstellung der Nullabbildung O als Linearkombination der Φ_{ij}.

  Einerseits ist dann

  $$\begin{aligned} \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij}\, \Phi_{ij}\Big)(v_k) &= \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} \Phi_{ij}(v_k) \\ &= \sum_{i=1}^{m} (\sum_{j=1}^{n} \lambda_{ij} \delta_{jk}) w_i = \sum_{i=1}^{m} \lambda_{ik} w_i, \end{aligned}$$

  und andererseits

  $$O(v_k) = 0.$$

  Also

  $$\sum_{i=1}^{m} \lambda_{ik} w_i = 0 \qquad (1 \leq k \leq n).$$

  Daraus folgt aber wegen der linearen Unabhängigkeit der w_i, dass alle $\lambda_{ik} = 0$ sind $(1 \leq i \leq m,\ 1 \leq k \leq n)$.

- Die Φ_{ij} erzeugen $\mathrm{Hom}(V,W)$: Dazu sei $\Phi \in \mathrm{Hom}(V,W)$ beliebig gewählt, und $\Phi(v_k)$ sei als Linearkombination der w_i dargestellt:

  $$\Phi(v_k) = \sum_{i=1}^{m} \alpha_{ik} w_i \qquad (1 \leq k \leq n).$$

Wir zeigen, dass dann Φ mit der linearen Abbildung

$$\Psi := \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \Phi_{ij}$$

übereinstimmt. Es ist nämlich für $k = 1, \ldots, n$

$$\Psi(v_k) = \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \Phi_{ij}\Big)(v_k) = \sum_{i=1}^{m} \alpha_{ik} w_i = \Phi(v_k).$$

Hieraus folgt nach Satz 9.6, dass $\Phi = \Psi$ ist. ■

9.4.1 Spezialfall $W = \mathbb{K}$: der Dualraum eines Vektorraums

Wir können den Skalar-Körper $\mathbb{K}$ als eindimensionalen Vektorraum $\mathbb{K}^1$ über sich selbst auffassen. Daraus ergibt sich eine wichtige Spezialisierung des Vektorraums $\mathrm{Hom}(V, W)$: Wir setzen $W = \mathbb{K}^1 = \mathbb{K}$ und erhalten den Vektorraum $\mathrm{Hom}(V, \mathbb{K})$ aller linearen Abbildungen von V in den Körper $\mathbb{K}$. Man nennt $\mathrm{Hom}(V, \mathbb{K})$ den **Dualraum** $V^\star$ von V. Seine Elemente, also die linearen Abbildungen $\Phi : V \to \mathbb{K}$, heißen auch **Linearformen** auf V.

Wenn V n-dimensional ist, so ist nach Satz 9.29 wegen $m = 1$ auch $V^\star$ n-dimensional und somit nach Satz 9.23 zu V isomorph.

Für $n \geq 1$ sei $\{v_1, \ldots, v_n\}$ eine Basis von V. Als Basisvektor von $W = \mathbb{K}$ wählen wir $w = w_1 = 1 \in \mathbb{K}$. Wie im Beweis von Satz 9.29 können wir eine Basis $\{v_1^*, \ldots, v_n^*\}$ von $V^\star$ durch

$$v_j^*(v_k) := \delta_{jk}, \qquad j, k = 1, \ldots, n \tag{9.1}$$

festlegen (v_j^* bildet v_j auf 1, die übrigen v_k mit $k \neq j$ auf 0 ab). Diese Basis $\{v_1^*, \ldots, v_n^*\}$ heißt die zur Basis $\{v_1, \ldots, v_n\}$ gehörige **Dualbasis** von $V^\star$.

Es seien nun $x \in V$ und $\Phi \in V^\star$ beliebig gewählt. Bezüglich der Basen $\{v_1, \ldots, v_n\}$ und $\{v_1^*, \ldots, v_n^*\}$ haben sie die Darstellungen

$$x = \sum_{k=1}^{n} \xi_k v_k, \qquad \Phi = \sum_{j=1}^{n} \alpha_j v_j^*. \tag{9.2}$$

Aus (9.1) und (9.2) ergibt sich

$$\begin{aligned} \Phi(x) &= \Big(\sum_{j=1}^{n} \alpha_j v_j^*\Big)\Big(\sum_{k=1}^{n} \xi_k v_k\Big) = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \xi_k v_j^*(v_k) \\ &= \sum_{k=1}^{n} \alpha_k \xi_k = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \cdots + \alpha_n \xi_n. \end{aligned} \tag{9.3}$$

Der zuletzt gewonnene Ausdruck erklärt die Bezeichnung *Linearform* für $\Phi : V \to \mathbb{K}$. Aus (9.3) und (9.1) ergibt sich noch

$$\Phi(v_k) = \alpha_k, \qquad v_j^*(x) = \xi_j. \tag{9.4}$$

Bemerkung 9.30 Es sei $\dim V = n \geq 1$. Der Dualraum $V^{\star\star} := (V^\star)^\star$ von $V^\star$ heißt **Bidualraum** von V. Der Bidualraum ist isomorph zum ursprünglichen Vektorraum:

$$F : V \to V^{\star\star}, \quad x \mapsto F(x) \quad \text{mit} \quad F(x) : V^\star \to \mathbb{K}, \quad \Phi \mapsto \Phi(x)$$

ist ein VR-Isomorphismus.

10 Darstellungen von linearen Abbildungen durch Matrizen

10.1 Abbildungsmatrizen

Es seien V ein n-dimensionaler und W ein m-dimensionaler Vektorraum über demselben Körper $\mathbb{K}$. Weiter sei $B = \{b_1, \ldots, b_n\}$ eine Basis von V und $C = \{c_1, \ldots, c_m\}$ eine Basis von W.

Einer linearen Abbildung $\Phi : V \to W$ ordnen wir dann in folgender Weise eine Matrix A zu: Es sei

$$\Phi(b_k) = \sum_{i=1}^{m} a_{ik} c_i, \qquad k = 1, \ldots, n, \quad a_{ik} \in \mathbb{K} \tag{10.1}$$

die Basisdarstellung des Bildvektors $\Phi(b_k)$ bezüglich C. Damit ist zu Φ und den gewählten Basen B, C eindeutig eine $m \times n$-Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

definiert, die man als **Abbildungsmatrix** von Φ bezüglich der (geordneten) Basen B, C bezeichnet. Die Konstruktionsvorschrift von A lautet also:

In der k-ten Spalte von A stehen die Komponenten von $\Phi(b_k)$ bezüglich der Basis C gemäß (10.1).

Mit der Abbildungsmatrix A läßt sich die lineare Abbildung $\Phi : V \to W$ folgendermaßen in Komponenten beschreiben: Es ist mit (10.1) zunächst

$$\Phi(x) = \Phi\left(\sum_{k=1}^{n} \xi_k\, b_k\right) = \sum_{k=1}^{n} \xi_k\, \Phi(b_k) = \sum_{i=1}^{m}\left(\sum_{k=1}^{n} a_{ik}\, \xi_k\right) c_i.$$

Setzt man dann

$$y = \Phi(x) = \sum_{i=1}^{m} \eta_i c_i \tag{10.2}$$

und vergleicht mit der obigen Gleichung, so erhält man

$$\eta_i = \sum_{k=1}^{n} a_{ik}\xi_k, \qquad i = 1, \ldots, m$$

oder ausgeschrieben

$$\begin{aligned} \eta_1 &= a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n \\ \eta_2 &= a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n \\ &\vdots \\ \eta_m &= a_{m1}\xi_1 + a_{m2}\xi_2 + \cdots + a_{mn}\xi_n. \end{aligned} \tag{10.3}$$

Mit dem Komponentenvektor $\Theta_B(x)$ von x bzgl. B und dem Komponentenvektor $\Theta_C(y)$ von $y = \Phi(x)$ bzgl. C, also mit

$$\Theta_B(x) = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n, \quad \Theta_C(y) = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{pmatrix} \in \mathbb{K}^m$$

lässt sich (10.3) einfacher in Matrizenschreibweise durch

$$\Theta_C(y) = A\Theta_B(x) \tag{10.4}$$

ausdrücken. Damit haben wir eine lineare Abbildung

$$\hat{\Phi} : \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax \tag{10.5}$$

gefunden, die man auch als die **Darstellung von** Φ mittels der Abbildungsmatrix A bezeichnet. Die lineare Abbildung $\hat{\Phi}$ hat folgende Eigenschaft: Ist $E = \{e_1, e_2, \ldots, e_n\}$ die Standardbasis von $\mathbb{K}^n$, so ist $\hat{\Phi}(e_k) = k$-te Spalte von A ($k = 1, 2, \ldots, n$).

Der Zusammenhang zwischen linearen Abbildungen $\Phi : V \to W$ und ihren Darstellungen bezüglich Basen von V und W wird im folgenden Hilfssatz präzisiert.

Hilfssatz 10.1 (und Definition) *In einem $\mathbb{K}$-Vektorraum V der Dimension n sei eine geordnete Basis B und in einem $\mathbb{K}$-Vektorraum W der Dimension m sei eine geordnete Basis C gewählt. Weiter sei*

$$M_C^B : \mathrm{Hom}(V, W) \to \mathbb{K}^{m\times n}, \quad \Phi \mapsto A = M_C^B(\Phi),$$

die Abbildung die jedem Homomorphismus $\Phi : V \to W$ die Abbildungsmatrix $A \in \mathbb{K}^{m\times n}$ bzgl. B, C gemäß (10.1) *zuordnet. Dann ist M_C^B bijektiv.*

Beweis: Aus den Definitionen ist klar, dass M_C^B eine Abbildung ist. Um zu sehen, dass M_C^B bijektiv ist, betrachten wir eine Matrix $A \in \mathbb{K}^{m\times n}$ und die zugehörige lineare Abbildung gemäß (10.5). Sind dann in V bzw. in W geordnete Basen B bzw. C gewählt, so gibt es genau eine lineare Abbildung $\Phi : V \to W$, die bezüglich B, C die gegebene Matrix A als Abbildungsmatrix hat, denn Φ ist durch die Werte auf der Basis B vollständig bestimmt (vgl. Satz 9.6). ∎

Wir wollen nun noch zeigen, dass die bijektive Abbildung M_C^B in Hilfssatz 10.1 linear und damit ein Isomorphismus zwischen den $\mathbb{K}$–Vektorräumen $\mathrm{Hom}(V, W)$ und $\mathbb{K}^{m\times n}$ ist. Dazu seien Φ und $\Psi \in \mathrm{Hom}(V, W)$ zwei lineare Abbildungen und

$$A = (a_{jk}) = M_C^B(\Phi)\,, \qquad \tilde{A} = (\tilde{a}_{jk}) = M_C^B(\Psi)$$

ihre Abbildungsmatrizen. Wie sehen dann die Abbildungsmatrizen

$$M_C^B(\Phi + \Psi)\,, \quad M_C^B(\lambda\,\Phi)$$

der Summenabbildung $\Phi + \Psi$ und des λ-fachen $\lambda\Phi$, $\lambda \in \mathbb{K}$, aus? Nach der Konstruktionsvorschrift für Abbildungsmatrizen gemäß (10.1) ist

$$\Phi(b_k) = \sum_{i=1}^m a_{ik} c_i, \qquad \Psi(b_k) = \sum_{i=1}^m \tilde{a}_{ik} c_i, \qquad k = 1, \ldots, n$$

und daraus ergibt sich nach Definition 9.26

$$\begin{aligned} (\Phi + \Psi)(b_k) &= \Phi(b_k) + \Psi(b_k) = \sum_{i=1}^m (a_{ik} + \tilde{a}_{ik}) c_i, \\ (\lambda\Phi)(b_k) &= \lambda\Phi(b_k) = \sum_{i=1}^m (\lambda a_{ik}) c_i, \qquad k = 1, \ldots, n. \end{aligned}$$

Also ist die Abbildungsmatrix der Summen-Abbildung $\Phi + \Psi$ die Summe der Matrizen $A + \tilde{A}$, und die Abbildungsmatrix von $\lambda\Phi$ ist das λ-fache λA, d.h. es gilt

$$\begin{aligned} M_C^B(\Phi + \Psi) = M_C^B(\Phi) + M_C^B(\Psi), \\ M_C^B(\lambda\Phi) = \lambda\, M_C^B(\Phi), \end{aligned} \tag{10.6}$$

und M_C^B ist linear. Mit Hilfsatz 10.1 und den Sätzen 9.23 und 9.29 folgt also

Satz 10.2 *Die Abbildung* $M_C^B : \mathrm{Hom}(V,W) \to \mathbb{K}^{m\times n}$ *ist ein Vektorraum-Isomorphismus; insbesondere ist*

$$\dim \mathbb{K}^{m\times n} = \dim \mathrm{Hom}(V,W) = mn. \tag{10.7}$$

Eine Basis des Vektorraums $\mathbb{K}^{m\times n}$ besteht z.B. aus den $m \cdot n$ Matrizen $E_{ij} \in \mathbb{K}^{m\times n}$ der Form

$$E_{ij} = \begin{pmatrix} 0 & \vdots & 0 \\ \cdots & 1 & \cdots \\ 0 & \vdots & 0 \end{pmatrix}, \qquad i = 1,\ldots,m, \quad j = 1,\ldots,n\,, \tag{10.8}$$

die am Kreuzungspunkt der i-ten Zeile mit der j-ten Spalte eine 1 und sonst Nullen enthalten. Diese Basismatrizen sind die Bilder der im Beweis zu Satz 9.29 vorgekommenen Basisvektoren Φ_{ij} von $\mathrm{Hom}(V,W)$ bei einem Isomorphismus der Form M_C^B.

10.1.1 Abbildungsmatrix einer Verkettung

Die Abbildungsmatrix einer *Verkettung* von linearen Abbildungen ist das *Produkt* der einzelnen Abbildungsmatrizen. Genauer gilt:

Satz 10.3 *Es seien* V_1, V_2, V_3 $\mathbb{K}$-*Vektorräume mit* $\dim V_1 = l, \dim V_2 = m$ *und* $\dim V_3 = n$ *mit geordneten Basen* B_1, B_2, B_3. *Weiter seien* $\Phi : V_1 \to V_2$ *und* $\Psi : V_2 \to V_3$ *lineare Abbildungen. Dann gilt*

$$M_{B_3}^{B_1}(\Psi \circ \Phi) = M_{B_3}^{B_2}(\Psi) \cdot M_{B_2}^{B_1}(\Phi).$$

Beweis: Es seien $B_1 = \{x_1, \ldots, x_l\}, B_2 = \{y_1, \ldots, y_m\}, B_3 = \{z_1, \ldots, z_n\}$ und $A = (a_{ij}) := M_{B_2}^{B_1}(\Phi), B = (b_{ij}) := M_{B_3}^{B_2}(\Psi)$ und $C = (c_{ij}) := M_{B_3}^{B_1}(\Psi \circ \Phi)$ die zugehörigen Abbildungsmatrizen. Dann gilt

$$\begin{aligned}(\Psi \circ \Phi)(x_j) = \Psi(\Phi(x_j)) = \Psi(\sum_{i=1}^m a_{ij} y_i) = \sum_{i=1}^m a_{ij} \Psi(y_i) = \\ = \sum_{i=1}^m a_{ij} (\sum_{k=1}^n b_{ki} z_k) = \sum_{k=1}^n (\sum_{i=1}^m b_{ki} a_{ij}) z_k.\end{aligned}$$

Also nach Definition der Abbildungsmatrix (bezüglich gegebener Basen): $c_{kj} = \sum_{i=1}^m b_{ki} a_{ij}$ oder $C = B \cdot A$, also gilt die Behauptung. ■

10.1.2 Abbildungsmatrix der inversen Abbildung

Satz 10.4 *Es seien V, W n-dimensonale $\mathbb{K}$-Vektorräume mit geordneten Basen B, C. Eine lineare Abbildung $\Phi : V \to W$ ist genau dann ein Isomorphismus, wenn die Abbildungsmatrix $A := M_C^B(\Phi)$ invertierbar ist.*

In diesem Fall gilt für die inverse Abbildung $\Phi^{-1} : W \to V$

$$M_B^C(\Phi^{-1}) = (M_C^B(\Phi))^{-1} = A^{-1}.$$

Beweis: Ist Φ ein Isomorphismus, so gilt $\Phi^{-1} \circ \Phi = \mathrm{id}_V$ und $\Phi \circ \Phi^{-1} = \mathrm{id}_W$. Nach Satz 10.3 also

$$M_B^C(\Phi^{-1}) \cdot M_C^B(\Phi) = M_B^B(\mathrm{id}_V) = E_n \text{ und } M_C^B(\Phi) \cdot M_B^C(\Phi^{-1}) = M_C^C(\mathrm{id}_W) = E_n.$$

Also folgt $M_B^C(\Phi^{-1}) = (M_C^B(\Phi))^{-1}$.

Sei nun umgekehrt die Matrix $M_C^B(\Phi)$ invertierbar. Wir betrachten die lineare Abbildung $\Psi : W \to V$, die nach Satz 10.2 zu $(M_C^B(\Phi))^{-1}$ gehört. Es gilt also $(M_C^B(\Phi))^{-1} = M_B^C(\Psi)$ und wieder nach Satz 10.3 folgt

$$E_n = M_B^C(\Psi) \cdot M_C^B(\Phi) = M_B^B(\Psi \circ \Phi) \text{ und } E_n = M_C^B(\Phi) \cdot M_B^C(\Psi) = M_C^C(\Phi \circ \Psi).$$

Daraus folgt nochmals nach Satz 10.2

$$\Psi \circ \Phi = \mathrm{id}_V \text{ und } \Phi \circ \Psi = \mathrm{id}_W,$$

also $\Psi = \Phi^{-1}$. ■

10.1.3 Abbildungsmatrix der dualen Abbildung

Es seien V und W endlich dimensionale $\mathbb{K}$-Vektorräume und $\Phi : V \longrightarrow W$ ein Homomorphismus. Die **duale Abbildung** $\Phi^* : W^* \longrightarrow V^*$ ordnet jeder Linearform $\alpha \in W^*$ die Linearform $\alpha \circ \Phi : V \longrightarrow \mathbb{K}$ zu. Man rechnet leicht nach, dass Φ^* ebenfalls eine lineare Abbildung ist, z.B. gilt für alle $\alpha, \beta \in W^*$ und alle $v \in V$:

$$(\Phi^*(\alpha + \beta))(v) = (\alpha + \beta)(\Phi(v)) = \alpha(\Phi(v)) + \beta(\Phi(v)) = (\Phi^*(\alpha))(v) + (\Phi^*(\beta))(v).$$

Die Abbildung Φ habe bezüglich zweier Basen B bzw. C von V bzw. von W die Abbildungsmatrix $A := M_C^B(\Phi)$. Wie sieht dann die Abbildungsmatrix der dualen Abbildung Φ^* bezüglich der dualen Basen C^* bzw. B^* von W^* bzw. von V^* aus?

Wir bezeichnen die gesuchte Abbildungsmatrix $M_{B^*}^{C^*}(\Phi^*)$ mit $\tilde{A} = (\tilde{a}_{ij})$; d.h. es gilt

$$\Phi^*(c_j^*) = \sum_{l=1}^{n} \tilde{a}_{lj} b_l^*.$$

Daraus folgt einmal

$$\Phi^*(c_j^*)(b_i) = \sum_{l=1}^{n} \tilde{a}_{lj} b_l^*(b_i) = \sum_{l=1}^{n} \tilde{a}_{lj} \delta_{li} = \tilde{a}_{ij}.$$

Nach Definition von Φ^* ist andererseits

$$\Phi^*(c_j^*)(b_i) = c_j^*(\Phi(b_i)) = c_j^*(\sum_{l=1}^{m} a_{li} c_l) = \sum_{l=1}^{m} a_{li} c_j^*(c_l) = \sum_{l=1}^{m} a_{li} \delta_{jl} = a_{ji}.$$

Somit haben wir gezeigt:

Satz 10.5 *Bezüglich der dualen Basen wird die duale Abbildung durch die transponierte Matrix beschrieben:*

$$\tilde{A} = A^\top \quad \text{oder genauer} \quad M_{B*}^{C*}(\Phi^*) = M_C^B(\Phi)^\top.$$

10.1.4 Rang = Rang

Wir wollen den Zusammenhang zwischen einer linearen Abbildung $\Phi : V \to W$ und ihrer Darstellung $\hat{\Phi} : \mathbb{K}^n \to \mathbb{K}^m$ bzgl. Basen B, C noch etwas genauer ansehen. Wir haben die Isomorphismen

$$\begin{aligned} \Theta_B \quad &: \quad V \to \mathbb{K}^n, \quad x = \sum_{k=1}^{n} \xi_k b_k \mapsto \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \\ \Theta_C \quad &: \quad W \to \mathbb{K}^m, \quad y = \sum_{i=1}^{m} \eta_i c_i \mapsto \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} \end{aligned}$$

(vergleiche Beispiel 9.24). Nach (10.5), (10.4) und (10.2) ist dann

$$\Theta_C^{-1}\left(\hat{\Phi}(\Theta_B(x))\right) = \Theta_C^{-1}(A\Theta_B(x)) = \Theta_C^{-1}(\Theta_C(y)) = y = \Phi(x)$$

für alle $x \in V$. Somit gilt

$$\Theta_C^{-1} \circ \hat{\Phi} \circ \Theta_B = \Phi; \tag{10.9}$$

d.h. folgendes Diagramm ist „kommutativ“

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ \Theta_B \Big\downarrow & & \Big\downarrow \Theta_C \\ \mathbb{K}^n & \xrightarrow[\hat{\Phi}]{} & \mathbb{K}^m \end{array}$$

Da Θ_B, Θ_C bijektive Abbildungen sind, ergibt sich aus (10.9) auch

$$\hat{\Phi} = \Theta_C \circ \Phi \circ \Theta_B^{-1}.$$

Wir verwenden (10.9) zum Beweis des folgenden Satzes:

Satz 10.6 *Es sei $\Phi : V \to W$ eine lineare Abbildung und A eine (beliebige) Abbildungsmatrix von Φ. Dann gilt*

$$\text{Rang } \Phi = \text{Rang } A.$$

Beweis: Es sei $n = \dim V$ und $m = \dim W$. Weiter sei $A = M_C^B(\Phi)$ die Abbildungsmatrix von Φ bezüglich Basen B von V und C von W und $\hat{\Phi} : \mathbb{K}^n \to \mathbb{K}^m; x \mapsto Ax$ sei die Darstellung von Φ bzgl. B, C.

Der Bildraum von $\hat{\Phi}$ ist die lineare Hülle der Bilder der Standardbasis von $\mathbb{K}^n$ und wir hatten gesehen, dass $\hat{\Phi}(e_k)$ gerade die k-te Spalte von A (für $k = 1, 2, \ldots, n$) ist. Also ist der Bildraum von $\hat{\Phi}$ gleich dem Spaltenraum von A. Damit haben wir die Gleichheit Rang $\hat{\Phi}$ = Rang A.

Weiter ist nach (10.9) und wegen $\Theta_B(V) = \mathbb{K}^n$

$$\Phi(V) = (\Theta_C^{-1} \circ \hat{\Phi} \circ \Theta_B)(V) = \Theta_C^{-1}(\hat{\Phi}(\mathbb{K}^n)).$$

Da der Isomorphismus Θ_C^{-1} die Dimension von $\hat{\Phi}(\mathbb{K}^n)$ nicht ändert, folgt

$$\text{Rang } \Phi = \dim \Phi(V) = \dim \hat{\Phi}(\mathbb{K}^n) = \text{Rang } \hat{\Phi}.$$

Insgesamt ergibt sich somit die Behauptung: Rang Φ = Rang $\hat{\Phi}$ = Rang A. ■

10.2 Basiswechsel für Homomorphismen

Ein Homomorphismus $\Phi : V \longrightarrow W$ habe bezüglich gegebener Basen $B = \{b_1, \ldots, b_n\}$ von V und $C = \{c_1, \ldots, c_m\}$ von W die Abbildungsmatrix $A := M_C^B(\Phi)$. Wie lässt sich die Abbildungsmatrix $\widetilde{A} := M_{\widetilde{C}}^{\widetilde{B}}$ von Φ bezüglich „neuer“ Basen $\widetilde{B}, \widetilde{C}$ berechnen?

Dazu schreiben wir zunächst die „neuen“ Basisvektoren $\widetilde{b}_j \in \widetilde{B}$ als Linerarkombinationen der „alten“:

$$\widetilde{b}_j = \sum_{i=1}^{n} s_{ij} b_i,$$

fassen also die Koeffizienten von $\widetilde{B}$ bezüglich B in einer invertierbaren Matrix $S = (s_{ij})_{1 \leq i,j \leq n} \in \mathbf{GL}_n(\mathbb{K})$ zusammen (in der Sprechweise von Abschnitt 7.3 ist S die Matrix des Basiswechsels von $\widetilde{B}$ nach B!).

Genauso schreiben wir einen „alten“ Basisvektor $c_k \in C$ bezüglich der „neuen“ Basis $\widetilde{C}$ als

$$c_k = \sum_{l=1}^{m} t_{lk}\widetilde{c}_l$$

und erhalten entsprechend eine Matrix $T = (t_k)_{1\le i,j\le m} \in \mathbf{GL}_m(\mathbb{K})$ (in der Sprechweise von Abschnitt 7.3 ist T die Matrix des Basiswechsels von C nach $\widetilde{C}$). Dann ergibt sich für die Koeffizienten von $\Phi(\widetilde{b}_j)$ bezüglich $\widetilde{C}$ die folgende Gleichung:

$$\begin{aligned}\Phi(\widetilde{b}_j) = \sum_{i=1}^{n} s_{ij}\Phi(b_i) &= \sum_{i=1}^{n} s_{ij}\sum_{k=1}^{m} a_{ki}c_k \\ &= \sum_{i=1}^{n} s_{ij}\sum_{k=1}^{m} a_{ki}\sum_{l=1}^{m} t_{lk}\widetilde{c}_l = \sum_{l=1}^{m}(\sum_{k=1}^{m}\sum_{i=1}^{n} t_{lk}a_{ki}s_{ij})\widetilde{c}_l.\end{aligned}$$

Daraus lesen wir ab, dass die Abbildungsmatrix $\widetilde{A}$ von Φ bezüglich $\widetilde{C}$ und $\widetilde{B}$ gegeben ist durch

$$\widetilde{A} = M_{\widetilde{C}}^{\widetilde{B}}(\Phi) = TAS = TM_C^B(\Phi)S. \tag{10.10}$$

Wir wollen die Gleichung 10.10 noch in einer genaueren Form schreiben. Dazu bemerken wir zunächst, dass die Matrix S des Basiswechsels von $\widetilde{B}$ nach B gegeben ist durch die Gleichungen

$$\widetilde{b}_j = \sum_{i=1}^{n} s_{ij}b_i, \quad j = 1, \dots, n.$$

Da andererseits auch $\widetilde{b}_j = \mathrm{id}_V(\widetilde{b}_j)$ ist, können wir S auch als Darstellungsmatrix der Identität von V schreiben: $S = M_B^{\widetilde{B}}(\mathrm{id}_V)$. Analog gilt für die Matrix T des Basiswechsels von C nach $\widetilde{C}$: $T = M_{\widetilde{C}}^{C}(\mathrm{id}_W)$. Damit ist Gleichung 10.10 also äquivalent zu folgender **Basiswechselformel**

$$M_{\widetilde{C}}^{\widetilde{B}}(\Phi) = M_{\widetilde{C}}^{C}(\mathrm{id}_W) \cdot M_C^B(\Phi) \cdot M_B^{\widetilde{B}}(\mathrm{id}_V).$$

Die obige Formel 10.10 motiviert die folgende

Definition 10.7 Zwei Matrizen $A, \widetilde{A} \in \mathbb{K}^{m\times n}$ heißen **äquivalent**, wenn es invertierbare Matrizen $S \in \mathbb{K}^{n\times n}$ und $T \in \mathbb{K}^{m\times m}$ gibt mit $\widetilde{A} = TAS$.

Satz 10.8 1. *Die Äquivalenz von Matrizen ist eine Äquivalenzrelation auf der Menge $\mathbb{K}^{m\times n}$.*

2. *Durch elementare Zeilen- oder Spaltenumformungen geht eine Matrix $A \in \mathbb{K}^{m\times n}$ in eine äquivalente Matrix A' über.*

3. *Jede Matrix $A \in \mathbb{K}^{m\times n}$ ist zu ihrer Gaußschen Normalform äquivalent.*

4. *Zwei Matrizen $A, B \in \mathbb{K}^{m\times n}$ sind genau dann äquivalent, wenn sie gleichen Rang haben.*

Beweis: Zu 1: Folgt leicht aus den Definitionen.

Zu 2: A ist die Abbildungsmatrix der linearen Abbildung

$$\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m; x \longmapsto Ax$$

bezüglich der Standardbasen. Elementare Zeilenumformungen entsprechen einem Basiswechsel in $\mathbb{K}^m$, elementare Spaltenumformungen einem Basiswechsel in $\mathbb{K}^n$.

Zu 3: Dies folgt direkt aus 2.

Zu 4: Durch elementare Zeilen- und Spaltenumformungen kann jede Matrix A auf die Form

$$A' = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \tag{10.11}$$

mit $r = \operatorname{Rang} A$ gebracht werden (vgl. Bemerkung 8.25). Nach 2. sind A und A' äquivalent. Gilt also $\operatorname{Rang} A = \operatorname{Rang} B$, so sind auch A und B äquivalent. Gilt umgekehrt $B = TAS$ mit regulären Matrizen $S \in \mathbb{K}^{n\times n}$ und $T \in \mathbb{K}^{m\times m}$, so ist A die Abbildungsmatrix von $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m; x \longmapsto Ax$, bezüglich der Standardbasen in $\mathbb{K}^n$ und $\mathbb{K}^m$ und B ist die Abbildungsmatrix von Φ bezüglich der geordneten Basen $(Se_1, \ldots, Se_n)$ in $\mathbb{K}^n$ und $(T^{-1}e_1, \ldots, T^{-1}e_m)$ in $\mathbb{K}^m$. Damit folgt

$$\operatorname{Rang} A = \operatorname{Rang} \Phi = \operatorname{Rang} B.$$

■

10.3 Basiswechsel für Endomorphismen

Wir betrachten eine lineare Selbstabbildung (also einen Endomorphismus) $\Phi : W \longrightarrow W$ und zwei Basen C, $\widetilde{C}$ von W. Weiter seien $A := M_C^C(\Phi)$ bzw. $\widetilde{A} := M_{\widetilde{C}}^{\widetilde{C}}(\Phi)$ die Abbildungsmatrizen des Endomorphismus Φ bezüglich C bzw. $\widetilde{C}$. Ist T die Matrix des Basiswechsels von C nach $\widetilde{C}$, so gilt nach den Ergebnissen des vorherigen Abschnittes,

$$\widetilde{A} = TAT^{-1}.$$

Definition 10.9 Zwei Matrizen $A, \widetilde{A} \in \mathbb{K}^{n\times n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $T \in \mathbb{K}^{n\times n}$ gibt mit $\widetilde{A} = TAT^{-1}$.

„Ähnlichkeit“ ist wie „Äquivalenz“ von Matrizen eine Äquivalenzrelation. Wir hatten gesehen, dass jede Matrix $A \in \mathbb{K}^{n \times n}$ äquivalent ist zu einer Matrix

$$A' = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}; \tag{10.12}$$

es gibt also nur endlich viele Äquivalenzklassen mit sehr einfachen Repräsentanten. Im Gegensatz dazu gibt es - wie wir sehen werden - im Allgemeinen unendlich viele Äquivalenzklassen von ähnlichen Matrizen in $\mathbb{K}^{n \times n}$. Ein Hauptproblem ist es, möglichst einfache Repräsentanten dieser Klassen zu finden (vgl. z.B. das Kapitel über die Jordansche Normalform).

11 Nochmals lineare Gleichungssysteme

Wir beantworten hier nochmals die grundlegenden Fragen über lineare Gleichungssysteme: Existieren Lösungen? Was ist die Struktur der Lösungsmenge? Diesmal aber vom Standpunkt der Theorie linearer Abbildungen aus. Ein konkretes Verfahren zur systematischen Berechnung der Lösungen eines LGS hatten wir mit dem Gauß-Algorithmus bereits in Abschnitt 3.3 kennengelernt.

Gegeben sei das LGS

$$\sum_{k=1}^{n} a_{ik}\, x_k = b_i \qquad (i = 1, \ldots, m) \tag{11.1}$$

mit $m, n \in \mathbb{N}$ und $a_{ik}, a_i \in \mathbb{K}$. Die beiden Matrizen

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (A \mid b) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{11.2}$$

heißen die zum LGS (11.1) gehörige **einfache** bzw. **erweiterte Matrix**. Wir können dann (11.1) auch kurz schreiben als $Ax = b$. Dies wiederum definiert eine lineare Abbildung $\Phi : \mathbb{K}^n \to \mathbb{K}^m$, $x \mapsto Ax$ (vgl. 9.5).

11.1 Wann ist ein LGS lösbar?

Hilfssatz 11.1 (a) *Das LGS $Ax = b$ ist genau dann lösbar, wenn für die lineare Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m;\ x \mapsto Ax$ gilt, dass $b \in \Phi(\mathbb{K}^n)$.*

(b) *Sind $s_1, \ldots, s_n \in \mathbb{K}^m$ die Spaltenvektoren der Matrix A, so gilt*

$$\Phi(\mathbb{K}^n) = [s_1, \ldots, s_n].$$

Beweis: (a) Wenn es eine Lösung $x \in \mathbb{K}^n$ von (11.1) gibt, so ist $\Phi(x) = b \in \Phi(\mathbb{K}^n)$. Liegt umgekehrt b im Bildraum $\Phi(\mathbb{K}^n)$, so gibt es ein Urbild $x \in \mathbb{K}^n$ mit $\Phi(x) = b$, also eine Lösung von (11.1).

(b) Mit Hilfe der Spaltenvektoren $s_1, \ldots, s_n$ kann man das LGS (11.1) als Vektorgleichung

$$x_1 s_1 + x_2 s_2 + \cdots + x_n s_n = b$$

und die zugehörige lineare Abbildung Φ in der Form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \Phi(x) = x_1 s_1 + x_2 s_2 + \cdots + x_n s_n$$

schreiben. Also folgt

$$\Phi(\mathbb{K}^n) \subset [s_1, \ldots, s_n].$$

Da für die Standard-Basis $\{e_1, \ldots, e_n\}$ von $\mathbb{K}^n$ gilt

$$\Phi(e_i) = s_i \qquad i = 1, \ldots, n,$$

liegen die s_i im Bildraum $\Phi(\mathbb{K}^n)$. Also haben wir auch $[s_1, \ldots, s_n] \subset \Phi(\mathbb{K}^n)$. ■

Wir erhalten jetzt den

Satz 11.2 (Lösbarkeitskriterium) *Das LGS $Ax = b$ ist genau dann lösbar, wenn gilt* $\operatorname{Rang} A = \operatorname{Rang}(A \mid b)$.

Beweis: „⇒“: Nach Hilfsatz 11.1 gilt

$$b \in [s_1, \ldots, s_n];$$

also $[s_1, \ldots, s_n] = [s_1, \ldots, s_n, b]$ und somit insbesondere

$$\operatorname{Rang} A = \dim[s_1, \ldots, s_n] = \dim[s_1, \ldots, s_n, b] = \operatorname{Rang}(A \mid b).$$

„⇐“: Aus

$$\dim[s_1, \ldots, s_n] = \dim[s_1, \ldots, s_n, b]$$

folgt nach Satz 8.15, dass $[s_1, \ldots, s_n] = [s_1, \ldots, s_n, b]$ und somit $b \in \Phi(\mathbb{K}^n)$, d.h. $Ax = b$ ist lösbar. ■

Ein Verfahren zur Bestimmung des (Zeilen-) oder (Spalten-)Ranges einer Matrix haben wir in Abschnitt 8.4 vorgestellt.

Beispiel 11.3 Es sei $\mathbb{K} = \mathbb{R}$ und $\alpha \in \mathbb{R}$ beliebig. Wir betrachten das LGS

$$\begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & 3x_3 & + & x_4 & = & 1 \\ 2x_1 & + & x_2 & + & x_3 & - & x_4 & = & 0 \\ & & 2x_2 & - & 13x_3 & + & x_4 & = & -1 \\ 2x_1 & - & x_2 & + & 14x_3 & - & 2x_4 & = & \alpha \end{array} \tag{11.3}$$

mit der zugehörigen einfachen bzw. erweiterten Matrix

$$\left(\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 2 & -13 & 1 & -1 \\ 2 & -1 & 14 & -2 & \alpha \end{array}\right).$$

Durch Elementaroperationen für die Spaltenvektoren erhält man

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ \star & 1 & 0 & 0 & 0 \\ \star & \star & 1 & 0 & 0 \\ \star & \star & \star & 0 & \alpha - 1 \end{array}\right).$$

Dabei ist dort, wo der Wert eines Elements nicht näher interessiert, das Zeichen $\star$ gesetzt. Der Rang der einfachen Matrix (ohne die letzte Spalte) ist also 3. Die erweiterte Matrix (mit 5 Spalten) hat für $\alpha = 1$ ebenfalls den Rang 3, für $\alpha \neq 1$ den Rang 4.

Nach Satz 11.2 ist das LGS (11.3) für $\alpha = 1$ lösbar; eine Lösung ist z. B.$(1, -1, 0, 1)$. Für alle $\alpha \neq 1$ ist (11.3) dagegen unlösbar.

Bemerkung 11.4 Ein homogenes LGS ist nach Satz 11.1 wegen $b = 0 \in \Phi(\mathbb{K}^n)$ stets lösbar. Sie besitzen stets die **triviale Lösung** $0 = (0, 0, \dots, 0) \in \mathbb{K}^n$. Die Frage, wann **nichttriviale** Lösungen existieren, wird im nächsten Abschnitt beantwortet. Dass inhomogene LGSe *nicht immer* lösbar sind, zeigt das obige LGS (11.3).

11.2 Struktur der Lösungsmenge eines LGS

Voraussetzung: In diesem Abschnitt nehmen wir immer an, dass das LGS (11.1) lösbar ist.

Die Menge $\mathfrak{L}$ aller Lösungen von $Ax = \Phi(x) = b$ ist dann

$$\mathfrak{L} = \{x \in \mathbb{K}^n \,|\, \Phi(x) = b\} = \Phi^{-1}(\{b\}).$$

Nach Hilfssatz 9.13 ist $\mathfrak{L}$ eine Restklasse von $\mathbb{K}^n$ modulo dem Kern von Φ: Wenn $x_0 \in \mathfrak{L}$ eine beliebig gewählte Lösung von (11.1) ist, so gilt also nach Hilfssatz 9.13 für jedes $x \in \mathfrak{L}$, dass $x - x_0 = v \in$ Kern Φ, und umgekehrt ist für jedes $v \in$ Kern Φ der Vektor $x = x_0 + v$ ein Element der Restklasse $\mathfrak{L}$. Damit haben wir gezeigt

Satz 11.5 *Die Lösungsmenge $\mathfrak{L}$ eines lösbaren LGS ist eine Restklasse modulo dem Kern der zugehörigen linearen Abbildung Φ. Genauer gilt: Ist $x_0 \in \mathfrak{L}$ eine beliebig gewählte Lösung von (11.1), so gilt für die Lösungsmenge*

$$\mathfrak{L} = \{x \in \mathbb{K}^n \mid \exists\, v \in \text{ Kern } \Phi :\, x = x_0 + v\} = x_0 + \text{Kern } \Phi.$$

Neben dem gegebenen LGS (11.1) betrachtet man oft auch *das zu (11.1) gehörige homogene LGS*:[4]

$$\sum_{k=1}^{n} a_{ik}\, x_k = 0 \qquad (i = 1, \dots, m). \tag{11.4}$$

[4](11.4) ist natürlich mit (11.1) identisch, wenn (11.1) schon homogen ist.

Die zu (11.4) gehörige lineare Abbildung Φ stimmt mit der zu (11.1) überein. Das homogene System (11.4) hat die triviale Lösung $0 \in \mathbb{K}^n$ und die Lösungsmenge $\mathfrak{L}_h$ von (11.4) ist gerade der Kern von Φ. Wir können daher Satz 11.5 auch so formulieren:

Satz 11.6 *Es seien $\mathfrak{L}$ die Lösungsmenge eines lösbaren LGS (11.1), $\mathfrak{L}_h$ die Lösungsmenge des zugehörigen homogenen LGS (11.4) und $x_0 \in \mathfrak{L}$ eine beliebig gewählte Lösung von (11.1). Dann gilt*

$$\mathfrak{L} = \{x \in \mathbb{K}^n \mid \exists\, v \in \mathfrak{L}_h : x = x_0 + v\} = x_0 + \mathfrak{L}_h.$$

Man erhält also alle Lösungen eines LGS, wenn man zu einer beliebig gewählten Lösung x_0 alle Lösungen des zugehörigen homogenen LGS addiert. Das kann man kurz so schreiben:

$$\mathfrak{L} = x_0 + \mathfrak{L}_h = x_0 + \text{Kern } \Phi.$$

Aus Satz 11.6 ergibt sich noch unmittelbar

Folgerung 11.7 *Ein lösbares LGS besitzt nur eine einzige Lösung genau dann, wenn das zugehörige homogene LGS nur die triviale Lösung besitzt.*

11.3 Homogene und inhomogene Gleichungssysteme

Ein homogenes LGS hat stets die triviale Lösung $x_0 = 0 \in \mathbb{K}^n$.

Nichttriviale Lösungen gibt es nach Satz 11.5 genau dann, wenn Kern $\Phi \neq \{0\}$, oder äquivalent, wenn dim Kern $\Phi > 0$. Wegen Rang $\Phi = \dim \mathbb{K}^n - \dim$ Kern Φ (Satz 9.21) ist dies äquivalent zu Rang $\Phi < \dim \mathbb{K}^n = n$.

Ist $d = \dim$ Kern $\Phi > 0$ und $\{v_1, \ldots, v_d\}$ eine Basis von Kern Φ, so lassen sich alle Lösungen des homogenen LGS als Elemente von Kern Φ in der Gestalt

$$x = \sum_{i=1}^{d} \lambda_i\, v_i \qquad (\lambda_i \in \mathbb{K}) \tag{11.5}$$

darstellen. Da noch Rang Φ = Rang A gilt, wobei A die (einfache) Matrix des gegebenen homogenen LGS ist, haben wir gezeigt

Satz 11.8 *Ein homogenes LGS (11.4) mit der Matrix A ist genau dann nichttrivial lösbar, wenn* Rang $A < n$ *ist. Ist $d = n -$* Rang $A > 0$, *dann gibt es d linear unabhängige Lösungen $v_1, \ldots, v_d \in \mathbb{K}^n$ von (11.4), und die Lösungsmenge $\mathfrak{L}_h$ von* (11.4) *besteht aus allen Linearkombinationen (11.5) der $v_1, \ldots, v_d$.*

Beispiel 11.9 Wir kommen auf das LGS (11.3) zurück. Für das zugehörige homogene LGS

$$\begin{array}{rcrcrcrcl} x_1 & + & x_2 & - & 3x_3 & + & x_4 & = & 0 \\ 2x_1 & + & x_2 & + & x_3 & - & x_4 & = & 0 \\ & & 2x_2 & - & 13x_3 & + & x_4 & = & 0 \\ 2x_1 & - & x_2 & + & 14x_3 & - & 2x_4 & = & 0 \end{array} \tag{11.6}$$

ist, wie in Beispiel 11.3 festgestellt, Rang $A = 3$. Wegen $n = 4$ ist also Rang $A < n$, und (11.6) ist nichttrivial lösbar. Eine Lösung ist z.B. $v = (-18, 32, 5, 1)^\top$, wie man durch Einsetzen bestätigt. Wegen $d = 4 - 3 = 1$ ist $\mathfrak{L}_h$ der von v aufgespannte eindimensionale Untervektorraum

$$\mathfrak{L}_h = \text{ Kern } \Phi = [\begin{pmatrix} -18 \\ 32 \\ 5 \\ 1 \end{pmatrix}].$$

Hat man noch (etwa durch Probieren) für das inhomogene LGS (11.3) mit $\alpha = 1$ eine Lösung $x_0 = (1, -1, 0, 1)^\top$ gefunden, so ist dessen Lösungsmenge $\mathfrak{L}$ nach Satz 11.6 von der Form

$$\mathfrak{L} = \{x \in \mathbb{R}^4 \mid x = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -18 \\ 32 \\ 5 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R}\}.$$

Teil V

Endomorphismen

12 Determinanten

Die Determinante ist:

- ein wichtiges Hilfsmittel, um Matrizen und lineare Selbstabbildungen zu untersuchen (z.B. gilt: A ist regulär $\iff$ $\det A \neq 0$)
- eine „Invariante" eines Endomorphismus (da ähnliche Matrizen gleiche Determinante haben)

12.1 Das Signum einer Permutation

Definition 12.1 Es sei $\sigma \in S_n$ eine Permutation der Zahlen $1, 2, \ldots, n$ mit Fehlstandszahl $F(\sigma)$ (vergleiche Abschnitt 5.2.1 für die Definitionen dieser Begriffe). Das **Signum** von σ ist dann definiert durch

$$\operatorname{sign} \sigma := (-1)^{F(\sigma)}.$$

Satz 12.2 *Eine Permutation $\sigma \in S_n$ sei das Produkt von r Transpositionen. Dann gilt:* $\operatorname{sign} \sigma := (-1)^r$.

Beweis: Nach Hilfssatz 5.19 (bzw. dessen Beweis) gilt für die Fehlstandszahl von σ: $F(\sigma) = 2z + r$ für eine ganze Zahl $z \in \mathbb{Z}$. Daraus folgt die Behauptung. ■

Satz 12.3 *Sind $\sigma_1 \in S_n$ und $\sigma_2 \in S_n$ zwei Permutationen der Zahlen $1, 2, \ldots, n$, so gilt:*

$$\operatorname{sign}(\sigma_1 \circ \sigma_2) = \operatorname{sign} \sigma_1 \cdot \operatorname{sign} \sigma_2.$$

Beweis: Nach Satz 5.17 können wir σ_1 und σ_2 als Produkt von Transpositionen schreiben

$$\sigma_1 = \tau_r' \circ \cdots \circ \tau_1', \qquad \sigma_2 = \tau_s'' \circ \cdots \circ \tau_1''.$$

Dann gilt $\sigma_1 \circ \sigma_2 = \tau_r' \circ \cdots \circ \tau_1' \circ \tau_s'' \circ \cdots \circ \tau_1''$. Nach Satz 12.2 ergibt sich daraus

$$\operatorname{sign}(\sigma_1 \circ \sigma_2) = (-1)^{r+s} = (-1)^r \cdot (-1)^s = \operatorname{sign} \sigma_1 \cdot \operatorname{sign} \sigma_2.$$

■

Folgerung 12.4 *Es sei $\sigma \in S_n$. Dann gilt:* $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}\sigma$.

Beweis: Nach Satz 12.3 ist

$$\operatorname{sign}\sigma \cdot \operatorname{sign}\sigma^{-1} = \operatorname{sign}(\sigma \circ \sigma^{-1}) = \operatorname{sign}\mathrm{id} = (-1)^0 = 1.$$

■

12.2 Definition der Determinantenfunktion

Definition 12.5 Eine Funktion D, die jedem geordneten n–Tupel $(a_1, \dots, a_n)$ von Vektoren aus $\mathbb{K}^n$ einen Skalar $D(a_1, \dots, a_n) \in \mathbb{K}$ zuordnet, heißt (n-dimensionale) **Determinanten-Funktion**, falls sie folgende Eigenschaften besitzt:

D1 $\forall\, \lambda \in \mathbb{K}, \forall\ i \in \{1, 2, \dots, n\}: \ D(a_1, \dots, \lambda a_i, \dots, a_n) = \lambda D(a_1, \dots, a_i, \dots, a_n)$,

D2 $D(a_1, \dots, a_i' + a_i'', \dots, a_n) = D(a_1, \dots, a_i', \dots, a_n) + D(a_1, \dots, a_i'', \dots, a_n)$

D3 Falls $a_i = a_j$ für $i \neq j$, so gilt $D(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = 0$.

D4 Für die Standardbasis $e_1, \dots, e_n$ von $\mathbb{K}^n$ ist $D(e_1, \dots, e_n) = 1$.

Beispiel 12.6

- $n = 1$: $D(\lambda) = \lambda \in \mathbb{K}$.
- $n = 2$: Für $a_1 := (a_{11}, a_{12})$ und $a_2 := (a_{21}, a_{22}) \in \mathbb{K}^2$ ist

$$D(a_1, a_2) := a_{11}a_{22} - a_{12}a_{21}$$

eine Determinanten-Funktion.

Folgerung 12.7

(a) *D ist* **multilinear**, *d.h. linear in jedem Argument: für $a_i = \sum_{k=1}^m \lambda_k b_k$, $i = 1, \dots, n$, gilt*

$$D(a_1, \dots, \sum_{k=1}^m \lambda_k b_k, \dots, a_n) = \sum_{k=1}^m \lambda_k D(a_1, \dots, b_k, \dots, a_n).$$

(b)

$$D(a_1, \dots, a_i + \sum_{\substack{k=1\\k\neq i}}^m \lambda_k a_k, \dots, a_n) = D(a_1, \dots, a_i, \dots, a_n).$$

(c) *D ist* **schiefsymmetrisch***:*

$$D(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -D(a_1, \dots a_j, \dots, a_i, \dots, a_n)$$

(d) *Ist $\sigma \in S_n$ eine Permutation, so gilt*

$$D(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \operatorname{sign} \sigma \; D(a_1, \dots, a_n).$$

Beweis: (a) folgt (mit Induktion nach m) direkt aus aus **D1** und **D2**. (b) ergibt sich aus (a) zusammen mit **D3**. (c): Wegen (b) gilt nacheinander:

$$\begin{aligned} D(a_1, \dots, a_i, \dots, a_j, \dots, a_n) &= D(a_1, \dots, a_i + a_j, \dots, a_j, \dots, a_n) \\ &= D(a_1, \dots, a_i + a_j, \dots, a_j - (a_i + a_j), \dots, a_n) \\ &= D(a_1, \dots, a_i + a_j + (-a_i), \dots, -a_i, \dots, a_n) \\ &= -D(a_1, \dots, a_j, \dots, a_i, \dots, a_n). \end{aligned}$$

(d) folgt aus (c), Satz 5.17 und der Definition des Signums. ■

Satz 12.8 *Sind die Vektoren $a_1, \dots, a_n \in \mathbb{K}^n$ linear abhängig, so ist $D(a_1, \dots, a_n) = 0$.*

Beweis: Sind $a_1, \dots, a_n$ linear abhängig, so ist einer der Vektoren a_i eine Linearkombination der anderen:

$$a_i = \sum_{\substack{k=1 \\ k \neq i}}^{n} \lambda_k a_k.$$

Eingesetzt in D ergibt mit Folgerung 12.7, (a):

$$D(a_1, \dots, a_i, \dots, a_n) = D(a_1, \dots, \sum_{\substack{k=1 \\ k \neq i}}^{n} \lambda_k a_k, \dots, a_n) = \sum_{\substack{k=1 \\ k \neq i}}^{n} \lambda_k D(a_1, \dots, a_k, \dots, a_n) = 0,$$

denn jeder der Summanden ist Null wegen **D3**. ■

12.3 Existenz und Eindeutigkeit der Determinantenfunktion

Wir überlegen zuerst, dass eine Determinatenfunktion (wenn sie denn existiert) eindeutig sein muss. Der Beweis der Eindeutigkeit führt dann auf eine explizite Formel.

Satz 12.9 (Eindeutigkeit) *Für jedes $n \geq 1$ gibt es höchstens eine Determinanten-Funktion von $\mathbb{K}^n$.*

Beweis: Seien D und $\widetilde{D}$ zwei n-dimensionale Determinanten-Funktionen. Dann gilt wegen **D4** für die Standardbasis $e_1, \ldots, e_n$ von $\mathbb{K}^n$

$$D(e_1, \ldots, e_n) = \widetilde{D}(e_1, \ldots, e_n) = 1.$$

Weiter gilt wegen Folgerung 12.7 (d) für eine Permutation $\sigma \in S_n$:

$$\begin{aligned} D(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) &= \text{sign } \sigma\, D(e_1, \ldots, e_n) = \text{sign } \sigma = \\ &= \text{sign}\, \sigma\, \widetilde{D}(e_1, \ldots, e_n) = \widetilde{D}(e_{\sigma(1)}, \ldots, e_{\sigma(n)}). \end{aligned}$$

Es sei nun

$$a_i = \sum_{j=1}^{n} a_{ij} e_j \qquad (i = 1, \ldots, n).$$

Da D multilinear ist, gilt:

$$\begin{aligned} D(a_1, \ldots, a_n) &= \sum_{j_1=1}^{n} a_{1j_1} \sum_{j_2=1}^{n} a_{2j_2} \ldots \sum_{j_n=1}^{n} a_{nj_n} D(e_{j_1}, \ldots, e_{j_n}) = \\ &= \sum_{j_1=1}^{n} \ldots \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} \ldots a_{nj_n} \widetilde{D}(e_{j_1}, \ldots, e_{j_n}) = \widetilde{D}(a_1, \ldots, a_n). \end{aligned}$$

Die vorletzte Gleichheit überlegt man sich wie folgt. Die Summanden, die zu Permutationen $\begin{pmatrix} 1 & \ldots & n \\ j_1 & \ldots & j_n \end{pmatrix}$ gehören, stimmen nach Obigem überein. Wenn $(j_1, \ldots, j_n)$ keine Permutation ist, so gilt $j_i = j_k$ für mindestens ein Paar i, k, $i \neq k$. Wegen **D3** ist dann in diesen Fällen

$$D(e_{j_1}, \ldots, e_{j_n}) = 0 = \widetilde{D}(e_{j_1}, \ldots, e_{j_n}).$$

■

Satz 12.10 (Existenz) *Für jedes $n \geq 1$ existiert eine (und deshalb genau eine) Determinantenfunktion. Sie ist gegeben durch die Formel von Leibniz*

$$D(a_1, \ldots, a_n) = \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{12.1}$$

Beweis: Wir stellen fest, dass eine Determinantenfunktion, falls es überhaupt eine gibt, durch die Formel 12.1 beschrieben wird. Dies folgt direkt aus dem Beweis des Eindeutigkeitssatzes: Es gilt ja

$$D(e_{j_1}, \ldots, e_{j_n}) = \begin{cases} \text{sign} \begin{pmatrix} 1 & \ldots & n \\ j_1 & \ldots & j_n \end{pmatrix} & \text{für Permutationen} \\ 0 & \text{sonst.} \end{cases}$$

Man kann nun den Satz beweisen, indem man zeigt, dass (12.1) die Eigenschaften **D1** bis **D4** einer Determinantenfunktion besitzt.

Wir wählen aber einen anderen Weg, der gleichzeitig eine weitere Formel zur Berechnung der Determinantenfunktion liefert.

Für eine $n \times n$-Matrix $A = (a_{ij})$ definieren wir die $(n-1) \times (n-1)$–Matrix A_{ij} als diejenige Matrix, die aus A entsteht, wenn man die i-te Zeile und j-te Spalte weglässt:

$$A_{ij} := \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

Wir konstruieren jetzt eine Determinantenfunktion mittels vollständiger Induktion nach n.

INDUKTIONS-VERANKERUNG: Für $n = 1$ und $a \in \mathbb{K}^1 = \mathbb{K}$ definieren wir $D(a) := a$.

INDUKTIONS-SCHRITT: Sei jetzt $n \geq 2$.

Wir nehmen an, die $n-1$-dimensionale Determinantenfunktion sei definiert. Wir wollen $D(a_1, \dots, a_n)$ definieren für $a_i \in \mathbb{K}^n$. Dazu betrachten wir die Matrix $A = (a_{ij})$, deren *Zeilen* gerade die Vektoren $a_1, \dots, a_n$ sind. Die Zeilen der Matrix A_{ij} sind dann Vektoren aus $\mathbb{K}^{n-1}$. Für diese ist nach Induktions-Annahme eine $(n-1)$-dimensionale Determinantenfunktion D definiert. Wir setzen $D_{ij} := D(A_{ij})$ und definieren

$$D(a_1, \dots, a_n) := (-1)^{1+1} a_{11} D_{11} + (-1)^{1+2} a_{21} D_{21} + \ldots + (-1)^{1+n} a_{n1} D_{n1} \qquad (12.2)$$

Wir weisen nun die Eigenschaften **D1** bis **D4** für (12.2) nach:

D4 $D(e_1, \dots, e_n) = D \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = (-1)^{1+1} D_{11} = 1.$

D3 Es sei $a_i = a_j$ mit $i < j$. Falls $r \neq i$ und $r \neq j$ ist, so hat A_{r1} zwei gleiche Zeilen. Also gilt nach Induktions-Voraussetzung $D_{r1} = 0$. Somit ist

$$D(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = (-1)^{1+i} a_{i1} D_{i1} + (-1)^{1+j} a_{j1} D_{j1}.$$

Weiter ist $a_{i1} = a_{j1}$ und A_{j1} entsteht aus A_{i1} durch $j - i - 1$ Vertauschungen benachbarter Zeilen. Wegen der Schiefsymmetrie (Induktionsvoraussetzung) ergibt sich

$$\begin{aligned} D(a_1, \ldots, a_n) = (-1)^{1+i} a_{i1} D_{i1} + (-1)^{1+j} a_{i1}((-1)^{j-i-1} D_{i1}) &= \\ = a_{i1} D_{i1}((-1)^{1+i} + (-1)^{2j-i}) &= 0. \end{aligned}$$

D1

$$\begin{aligned} &D(a_1, \ldots, \lambda a_i, \ldots, a_n) = \\ &= (-1)^{1+1} a_{11} \lambda D_{11} + (-1)^{1+2} a_{21} \lambda D_{21} + \cdots + (-1)^{1+i} \lambda a_{i1} D_{i1} + \cdots + (-1)^{1+n} a_{n1} \lambda D_{n1} \\ &= \lambda D(a_1, \ldots, a_i, \ldots, a_n). \end{aligned}$$

D2 Sei A' die Matrix mit den Zeilen $a_1, \ldots, a_i', \ldots, a_n$, A'' diejenige mit Zeilen $a_1, \ldots, a_i'', \ldots, a_n$ und A diejenige mit Zeilen $a_1, \ldots, a_i' + a_i'', \ldots, a_n$.

Wir setzen $D_{j1}' := D(A_{j1}')$ und $D_{j1}'' := D(A_{j1}'')$. Dann gilt nach Induktions-Voraussetzung

$$D_{j1} = D_{j1}' + D_{j1}'' \qquad \text{für } j \neq i \ \text{ und } \ D_{i1} = D_{i1}' = D_{i1}''.$$

Damit erhält man

$$\begin{aligned} D(a_1, \ldots, a_i' + a_i'', \ldots, a_n) &= \sum_{j=1}^{n} (-1)^{1+j} a_{j1} D_{j1} = \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{n} (-1)^{1+j} a_{j1} (D_{j1}' + D_{j1}'') + (-1)^{1+i} (a_{i1}' + a_{i1}'') D_{i1} \\ &= D(a_1, \ldots, a_i', \ldots, a_n) + D(a_1, \ldots, a_i'', \ldots, a_n). \end{aligned}$$

■

12.4 Die Determinante einer Matrix

Definition 12.11 Die **Determinante einer quadratischen Matrix** $A \in \mathbb{K}^{n \times n}$ ist definiert durch

$$\det A := D(a_1, \ldots, a_n),$$

wobei $a_1, \ldots, a_n$ die *Zeilen-Vektoren* von A sind. Statt $\det A$ schreiben wir auch $|A|$ bzw.

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Speziell gilt für die Einheitsmatrix E wegen **D4**:

$$\det E = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{vmatrix} = 1.$$

Bemerkung 12.12 Im Existenzbeweis einer Determinantenfunktion haben wir die *1. Spalte* der Matrix A ausgezeichnet. Wir hätten ebensogut von der k-ten Spalte ausgehen können. Der Nachweis der Eigenschaften **D1**, **D2**, **D3** einer Determinantenfunktion verläuft analog. Damit auch **D4** gilt, muss das Vorzeichen richtig gewählt werden. Wir erhalten dann die **Entwicklung nach der k-ten Spalte**:

$$D(a_1, \dots, a_n) = \sum_{j=1}^{n} (-1)^{k+j} a_{jk} D_{jk}. \tag{12.3}$$

Wegen der Eindeutigkeit beschreiben die Formeln (12.1), (12.2) und (12.3) dieselbe Determinantenfunktion.

Beispiele 12.13

1. $n = 3$: Entwicklung nach der 1. Spalte:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

2. Für eine Dreiecks-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & & a_{2n} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & a_{n-1n} \\ 0 & \cdots & & 0 & a_{nn} \end{pmatrix}$$

gilt nach (12.2) und mit vollständiger Induktion

$$\det A = a_{11} \det \begin{pmatrix} a_{22} & a_{23} & \cdots & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & & a_{3n} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & a_{n-1n} \\ 0 & \cdots & & 0 & a_{nn} \end{pmatrix} = \ldots = a_{11} a_{22} \cdots a_{nn}.$$

3. Ein Beispiel zur Berechnung der Determinante einer Matrix durch elementare Zeilenoperationen unter Berücksichtigung der Eigenschaften D1, D2, D3 und D4 der Determinantenfunktion:

$$
\begin{aligned}
\det \begin{pmatrix} 3 & 2 & 12 & 5 \\ 2 & 1 & 6 & 4 \\ 2 & 0 & 2 & -3 \\ 2 & 2 & 7 & 4 \end{pmatrix} \begin{array}{l} |\cdot 2 \\ |\cdot 3 \\ |\cdot 3 \\ |\cdot 3 \end{array}
&= \frac{1}{2}\left(\frac{1}{3}\right)^3 \det \begin{pmatrix} 6 & 4 & 24 & 10 \\ 6 & 3 & 18 & 12 \\ 6 & 0 & 6 & -9 \\ 6 & 6 & 21 & 12 \end{pmatrix} \\
&= (-1)\frac{1}{2}\left(\frac{1}{3}\right)^3 \det \begin{pmatrix} 6 & 3 & 18 & 12 \\ 6 & 4 & 24 & 10 \\ 6 & 0 & 6 & -9 \\ 6 & 6 & 21 & 12 \end{pmatrix} \\
&= (-1)\frac{1}{2}\cdot\left(\frac{1}{3}\right)^3 \det \begin{pmatrix} 6 & 3 & 18 & 12 \\ 0 & 1 & 6 & -2 \\ 0 & -3 & -12 & 21 \\ 0 & 3 & 3 & 0 \end{pmatrix} \\
&= (-1)\frac{1}{2}\cdot\left(\frac{1}{3}\right)^3 \cdot 6 \det \begin{pmatrix} 1 & 6 & -2 \\ -3 & -12 & 21 \\ 3 & 3 & 0 \end{pmatrix} \\
&= (-1)\left(\frac{1}{3}\right)^2 \det \begin{pmatrix} 1 & 6 & -2 \\ 0 & 6 & -27 \\ 0 & -15 & 6 \end{pmatrix} \begin{array}{l} \\ |\cdot 5 \\ |\cdot 2 \end{array} \\
&= (-1)\left(\frac{1}{3}\right)^2 \cdot \frac{1}{5}\cdot\frac{1}{2} \det \begin{pmatrix} 30 & -135 \\ -30 & 12 \end{pmatrix} \\
&= (-1)\left(\frac{1}{3}\right)^2 \cdot \frac{1}{5}\cdot\frac{1}{2}\cdot 30 \cdot (-123) \\
&= 41.
\end{aligned}
$$

12.5 Rechnen mit Determinanten

Es sei $A^\top$ die zu A transponierte Matrix, $A^\top = (a_{ik})^\top = (a_{ki})$.

Satz 12.14 *Für eine $(n \times n)$-Matrix $A \in \mathbb{K}^{n\times n}$ gilt:* $\det A = \det A^\top$.

Beweis:

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n} \operatorname{sign}\sigma \; a_{1\sigma(1)}\cdots a_{n\sigma(n)} = \sum_{\sigma\in S_n} \operatorname{sign}\sigma \; a_{\sigma^{-1}\circ\sigma(1)\sigma(1)}\cdots a_{\sigma^{-1}\circ\sigma(n)\sigma(n)} = \\
&= \sum_{\sigma\in S_n} \operatorname{sign}\sigma \; a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} = \sum_{\sigma\in S_n} \operatorname{sign}\sigma^{-1} \; a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} = \\
&= \sum_{\sigma^{-1}\in S_n} \operatorname{sign}(\sigma^{-1}) \; a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} = \\
&= \sum_{\sigma\in S_n} \operatorname{sign}\sigma \; a_{\sigma(1)1}\cdots a_{\sigma(n)n} = \det(A^\top).
\end{aligned}$$

Die dritte Gleichheit erhält man durch Umordnen der Summanden, die vierte wegen $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}\sigma$ und die fünfte bzw. sechste weil $\{\sigma \in S_n\} = \{\sigma^{-1} \in S_n\}$.

■

Bemerkung 12.15 Aus den *Zeilen* von A werden beim Transponieren die *Spalten* von $A^\top$. Entwickelt man die Determinante von $A^\top$ nach der i-ten Spalte gemäß (12.3), so ergibt das die **Entwicklung der Determinante von A nach der i-ten Zeile**:

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} D_{ik} \tag{12.4}$$

Satz 12.16 (Determinanten-Multiplikationssatz) *Für das Matrix-Produkt von $A, B \in \mathbb{K}^{n\times n}$ gilt:*

$$\det(A \cdot B) = \det A \cdot \det B.$$

Beweis: Die Spalten von A seien $a_1, \dots, a_n$, $A = (a_1 \mid \cdots \mid a_n)$. Wir können das Matrixprodukt AB dann in Spaltenform schreiben als

$$AB = (b_{11}a_1 + \cdots + b_{n1}a_n \mid \cdots \mid b_{1n}a_1 + \cdots + b_{nn}a_n).$$

Also nach Folgerung 12.7 (a), (d) und **D3**:

$$\begin{aligned}
\det(AB) &= \sum_{i_1=1}^{n} b_{i_1 1} \cdots \sum_{i_n=1}^{n} b_{i_n n} D(a_{i_1}, \dots, a_{i_n}) \\
&= \sum_{\sigma\in S_n} b_{\sigma(1)1}\cdots b_{\sigma(n)n} D(a_{\sigma(1)}, \dots, a_{\sigma(n)}) \\
&= \sum_{\sigma\in S_n} b_{\sigma(1)1}\cdots b_{\sigma(n)n} \operatorname{sign}\sigma \, D(a_1, \dots, a_n) \\
&= \det A \cdot \det B^\top = \det A \cdot \det B.
\end{aligned}$$

■

Satz 12.17 *Eine $n \times n$-Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann regulär, wenn* $\det A \neq 0$.

BEWEIS: „$\Rightarrow$": Ist A regulär, so existiert A^{-1} mit $AA^{-1} = E_n$. Nach Satz 12.16 ist

$$\det A \cdot \det A^{-1} = \det(A \cdot A^{-1}) = \det E_n = 1,$$

also $\det A \neq 0$.

„$\Leftarrow$": Ist $\det A \neq 0$, so sind die Zeilen von A nach Satz 12.8 linear unabhängig, also Rang $A = n$, d.h. A ist regulär. ■

Folgerung 12.18

1. *Für eine reguläre Matrix $A \in \mathbf{GL}(n, \mathbb{K})$ gilt*

$$\det(A^{-1}) = (\det A)^{-1} = \frac{1}{\det A}.$$

2. *Ähnliche Matrizen haben die gleiche Determinante:*

$$\det(T^{-1}AT) = \det A$$

BEWEIS:

1. Folgt aus der Formel im Beweis von Satz 12.17: $\det A \cdot \det(A^{-1}) = 1$.
2. $\det(T^{-1}AT) = \det(T^{-1}) \cdot \det A \cdot \det T = (\det T)^{-1} \cdot \det T \cdot \det A = \det A$. ■

Definition 12.19 Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\Phi : V \to V$ ein Endomorphismus. Weiter sei B eine Basis von V und $A := M_B^B(\Phi)$ die zugehörige Abbildungsmatrix von Φ. Die **Determinante des Endomorphismus** Φ ist

$$\det \Phi := \det A.$$

Dieser Begriff ist wohldefiniert, das heißt unabhängig von der Wahl einer Basis: Falls $\overline{B}$ eine weitere Basis von V ist mit Basiswechsel von $\overline{B}$ nach B, so gilt nach Abschnitt 10.3, dass $\overline{A} = M_{\overline{B}}^{\overline{B}}(\Phi) = T^{-1}AT$ für eine reguläre Matrix T. Also nach Folgerung 11.14 (2) $\det \overline{A} = \det(T^{-1}AT) = \det A$.

Satz 12.20 (Determinante und Inverse einer Matrix) *Es sei $A \in \mathbb{K}^{n \times n}$ eine reguläre Matrix und $d_{jk} := (-1)^{k+j} D_{kj}$. Dann gilt*

$$A^{-1} = \frac{1}{\det A}(d_{jk}).$$

Beweis: Es ist

$$\sum_{j=1}^{n} a_{ij} d_{jk} = \sum_{j=1}^{n} a_{ij} (-1)^{k+j} D_{kj} = \begin{cases} \det A & (k = i) \\ 0 & (k \neq i) \end{cases} .$$

Für $k = i$ steht links nämlich gerade die Entwicklung von A nach der i-ten Zeile. Für $k \neq i$ steht links die Entwicklung nach der k-ten Zeile der Determinante einer Matrix, deren i-te und k-te Zeilen übereinstimmen. ■

12.5.1 Determinanten und lineare Gleichungssysteme

Wir betrachten ein lineares Gleichungssystem mit n Geichungen und n Unbestimmten

$$A\,x = b \tag{12.5}$$

mit $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{K}^{n \times n},\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n,\ b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n.$

Satz 12.21 *Ein homogenes lineares Gleichungssystem (12.5) $Ax = 0$ (also mit n Gleichungen und n Unbekannten) hat genau dann nur die triviale Lösung, wenn* $\det A \neq 0$.

Beweis: Nach Satz 11.8 ist $Ax = 0$ genau dann nur trivial lösbar, wenn $\operatorname{Rang} A = n$ ist. Nach Folgerung 8.23 also genau dann, wenn A^{-1} existiert. Die Behauptung folgt dann mit Satz 12.17. ■

Wir betrachten jetzt ein LGS 12.5 mit $\det A \neq 0$. Es gibt dann genau eine Lösung $x = A^{-1} b$. Der folgende Satz liefert eine Formel für die Lösung x mittels Determinanten. Für $k = 1, \ldots, n$ setzen wir

$$D_k := \det \begin{pmatrix} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{pmatrix} .$$

Satz 12.22 (Cramersche Regel) *Ein LGS (12.5) mit* $\det A \neq 0$ *hat genau eine Lösung $x = (x_1, \ldots, x_n)$ und es gilt*

$$x_k = \frac{D_k}{\det A} \quad (k = 1, \ldots, n).$$

Dabei erhält man die n-reihige Determinante D_k, indem man in $\det A$ *die k-te Spalte durch den Vektor b aus 12.5 ersetzt.*

Beweis: Nach **D1** und Folgerung 12.7 (b) gilt

$$\begin{aligned} x_k \det A &= \det(a_1 \mid \cdots \mid a_{k-1} \mid x_k a_k \mid a_{k+1} \mid \cdots \mid a_n) \\ &= \det(a_1 \mid \cdots \mid a_{k-1} \mid \sum_{j=1}^{n} x_j a_j \mid a_{k+1} \mid \cdots \mid a_n) \\ &= \det(a_1 \mid \cdots \mid a_{k-1} \mid b \mid a_{k+1} \mid \cdots \mid a_n) = D_k. \end{aligned}$$

■

Beispiel 12.23 Für das LGS über $\mathbb{K} = \mathbb{Q}$

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & - & x_3 & = & 0 \\ x_1 & & & + & x_3 & = & 1 \\ 2x_1 & - & x_2 & & & = & 2 \end{array}$$

ist

$$\det\ A = \begin{vmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & -1 & 0 \end{vmatrix} = 4 \neq 0.$$

Wir berechnen

$$D_1 = \begin{vmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & -1 & 0 \end{vmatrix} = 3, \quad D_2 = \begin{vmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 0 \end{vmatrix} = -2, \quad D_3 = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & -1 & 2 \end{vmatrix} = 1$$

$$\begin{array}{ccc} \uparrow & \quad\quad\quad\quad \uparrow & \quad\quad\quad\quad\quad \uparrow \end{array}$$

und erhalten den Lösungsvektor

$$x = \begin{pmatrix} \frac{3}{4} \\ -\frac{1}{2} \\ \frac{1}{4} \end{pmatrix}.$$

13 Eigenwerte und Eigenvektoren

Die einfachsten linearen Selbstabbildungen eines $\mathbb{K}$-Vektorraumes V sind Streckungen, also von der Form

$$\lambda \mathrm{id}_V : V \to V,\ x \mapsto \lambda x \quad (\lambda \in \mathbb{K}).$$

Die Abbildungsmatrix bezüglich einer beliebigen Basis ist dann von der Form

$$A = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Etwas allgemeiner sind Endomorphismen $\Phi : V \to V$, deren Abbildungsmatrix bezüglich einer geeigneten Basis $\{v_1, \ldots, v_n\}$ Diagonalgestalt hat:

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Insbesondere gilt also $\Phi(v_i) = \lambda_i v_i$ für alle i. Ein Ziel dieses Kapitels ist es, solche Abbildungen zu verstehen.

13.1 Definitionen

Definition 13.1 Es seien V ein $\mathbb{K}$-Vektorraum und $\Phi : V \to V$ ein Endomorphismus. Der Skalar $\lambda \in \mathbb{K}$ heißt **Eigenwert** von Φ, falls ein Vektor $x \in V$ mit $x \neq 0$ existiert, so dass

$$\Phi(x) = \lambda x \quad \text{oder, äquivalent,} \quad (\Phi - \lambda\,\mathrm{id}_V)(x) = 0 \tag{13.1}$$

gilt. Der Vektor x heißt **Eigenvektor** von Φ zum **Eigenwert** λ.

Die Menge aller Eigenvektoren von Φ zu einem festen Eigenwert λ bildet *zusammen mit dem Nullvektor* einen Untervektorraum von V. Er heißt der zu λ gehörige **Eigenraum** von Φ und wird mit E_λ bezeichnet. Die Menge aller Eigenwerte von Φ heißt **Spektrum** von Φ.

Bemerkung 13.2 1. Wegen (13.1) gilt

$$E_\lambda = \mathrm{Kern}\,(\Phi - \lambda\,\mathrm{id}_V).$$

2. Es sei $A \in \mathbb{K}^{n\times n}$. Die Matrix A definiert dann die lineare Selbstabbildung $\Phi : V \to V, x \mapsto Ax$. Wir können deshalb auch von *Eigenwerten, Eigenvektoren* und *Eigenräumen* quadratischer Matrizen sprechen.

Sind λ_1, λ_2 zwei verschiedene Eigenwerte der linearen Selbstabbildung Φ von V, und sind x_1 (bzw. x_2) zu λ_1 (bzw. zu λ_2) gehörige Eigenvektoren, so sind x_1, x_2 linear unabhängig. Andernfalls wäre nämlich $x_2 = \alpha x_1$ für ein $\alpha \in \mathbb{K}$, und daraus folgt wegen

$$\begin{aligned} \Phi(x_2) &= \lambda_2 x_2, \\ \Phi(x_2) &= \Phi(\alpha x_1) = \alpha\lambda_1 x_1 = \lambda_1 x_2, \end{aligned}$$

dass $(\lambda_1 - \lambda_2)x_2 = 0$ im Widerspruch zu $\lambda_1 - \lambda_2 \neq 0$ und $x_2 \neq 0$.

FOLGERUNG: $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$.

Allgemeiner gilt

Satz 13.3 *Hat ein Endomorphismus Φ von V r verschiedene Eigenwerte $\lambda_1, \dots, \lambda_r$, so sind zugehörige Eigenvektoren $x_1, \dots, x_r$ linear unabhängig.*

BEWEIS: Für alle $i, k = 1, \dots, r$ gilt

$$(\Phi - \lambda_i \text{ id})(x_k) = \Phi(x_k) - \lambda_i x_k = (\lambda_k - \lambda_i)x_k.$$

Wenden wir $\Phi - \lambda_i$ id auf

$$\sum_{k=1}^{r} \alpha_k x_k = 0 \tag{$*$}$$

an, so folgt entsprechend

$$(\Phi - \lambda_i \text{ id})\left(\sum_{k=1}^{r} \alpha_k x_k\right) = \sum_{k=1}^{r} \alpha_k(\lambda_k - \lambda_i)x_k = 0,$$

und in der Summe fällt der Summand mit x_i weg. Wenden wir also auf $(*)$ nacheinander

$$\Phi - \lambda_1 \text{ id}, \dots, \Phi - \lambda_{j-1} \text{ id},\ \Phi - \lambda_{j+1} \text{ id}, \dots, \Phi - \lambda_r \text{ id}$$

an, so wird

$$\alpha_j(\lambda_j - \lambda_1)\cdots(\lambda_j - \lambda_{j-1})(\lambda_j - \lambda_{j+1})\cdots(\lambda_j - \lambda_r)x_j = 0,$$

und es folgt $\alpha_j = 0$ für $j = 1, \dots, r$. Die $x_1, \dots, x_r$ sind also linear unabhängig. ■

Folgerung 13.4 (a) *Ein Endomorphismus Φ eines n-dimensionalen $\mathbb{K}$-Vektorraumes hat höchstens n Eigenwerte. Jede Matrix $A \in \mathbb{K}^{n\times n}$ hat höchstens n Eigenwerte.*

(b) *Die Summe der Eigenräume von Φ ist direkt, und es gilt $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \subset V$.*

13.2 Berechnung der Eigenwerte: charakteristisches Polynom

Seien V ein n-dimensionaler Vektorraum, $\Phi \in \operatorname{Hom}(V,V)$, $A = M_B^B(\Phi)$ die Abbildungsmatrix von Φ bezüglich einer (geordneten) Basis B und $x_B = \Theta_B(x)$ die Koordinatendarstellung von $x \in V$ bezüglich derselben Basis B (siehe Abschnitt 10.1). Dann ist $\Phi(x) = \lambda x, \lambda \in \mathbb{K}$, äquivalent mit $Ax_B = \lambda x_B$. Es genügt daher, Eigenwerte von Matrizen zu betrachten. Nun gilt:

$$\begin{gathered}
\lambda \text{ ist Eigenwert von } A \in \mathbb{K}^{n\times n} \\
\Longleftrightarrow \text{Es existiert } x \in \mathbb{K}^n, x \neq 0, \text{ mit } Ax = \lambda x \text{ bzw. mit } (A - \lambda E_n)x = 0 \\
\Longleftrightarrow (A - \lambda E_n)x = 0 \text{ ist nichttrivial lösbar} \\
\Longleftrightarrow \operatorname{Rang}(A - \lambda E_n) < n \Longleftrightarrow \det(A - \lambda E_n) = 0
\end{gathered}$$

Nach Definition der Determinante erhalten wir für $A = (a_{ij})$

$$\det(A - \lambda E_n) = \sum_{\sigma \in S_n} \operatorname{sign}\sigma \cdot (a_{1\sigma(1)} - \lambda\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - \lambda\delta_{n\sigma(n)}).$$

Die Definition der Determinante einer Matrix $A = (a_{ij})$ ist nicht nur sinnvoll für Matrizen mit Einträgen a_{ij} aus einem Körper $\mathbb{K}$, sondern auch für Matrizen mit Einträgen aus einem kommutativen Ring R mit Eins. Sind die Einträge z.B. Polynome $a_{ij} \in \mathbb{K}[X]$, so ist $\det A$ ebenfalls ein Polynom. Speziell können wir also das Polynom

$$p_A = \det(A - XE_n) = \sum_{\sigma \in S_n} \operatorname{sign}\sigma \cdot (a_{1\sigma(1)} - X\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - X\delta_{n\sigma(n)}).$$

definieren. p_A heißt **charakteristisches Polynom** von A.

Wir ordnen nach Potenzen von X und erhalten

$$p_A = a_0 + a_1X + a_2X^2 + \cdots + a_{n-1}X^{n-1} + (-1)^n X^n$$

mit $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Dabei ist speziell

$$a_0 = \det A, \quad a_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) =: (-1)^{n-1}\operatorname{Spur} A, \quad a_n = (-1)^n,$$

wobei die **Spur** von A, $\operatorname{Spur} A$, definiert ist als die Summe der Diagonalelemente von A.

Einem formalen Polynom können wir durch Einsetzen eine Polynom-Funktion zuordnen:

$$p = a_0 + a_1X + \cdots a_nX^n \in \mathbb{K}[X] \mapsto f_p(t) = a_0 + a_1t + \cdots a_nt^n \in \mathrm{Abb}(\mathbb{K}, \mathbb{K}).$$

Definition 13.5 Ein Element $\alpha \in \mathbb{K}$ heißt **Nullstelle** von $p \in \mathbb{K}[X]$, falls $f_p(\alpha) = 0$ ist.

Bemerkung 13.6 Der Grundkörper $\mathbb{K}$ ist hier wichtige: Auch wenn p nicht das 0-Polynom ist, kann die zugehörige Polynom-Funktion $f_p(t)$ die Null-Funktion sein!

Beispiel: Sei $p = X - X^2 \in \mathbb{K}[X]$ und somit $f_p(t) = t - t^2 = t(1-t) \in \mathrm{Abb}(\mathbb{K}, \mathbb{K})$. Falls $\mathbb{K} = \mathbb{R}$ so hat $f_p(t)$ die zwei Nullstellen $t_1 = 0$ und $t_2 = 1$; für alle anderen t ist $f_p(t) \neq 0$. Falls $\mathbb{K} = \mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ (der Körper mit 2 Elementen), so ist $f_p(\overline{0}) = \overline{0}$ und $f_p(\overline{1}) = \overline{1}$, also $f_p(t)$ die Null-Funktion.

Zusammenfassend haben wir den

Satz 13.7 *$\lambda \in \mathbb{K}$ ist genau dann Eigenwert der Matrix $A \in \mathbb{K}^{n\times n}$, wenn λ Nullstelle des charakteristischen Polynoms p_A von A ist.*

Ist λ Eigenwert von A, so ist der Eigenraum E_λ gleich dem Lösungsraum des homogenen linearen Gleichungssystems $(A - \lambda E_n)x = 0$.

Satz 13.8 (Fundamentalsatz der Algebra) *Jedes komplexe Polynom mit* grad $\geq$ 1 *hat mindestens eine Nullstelle.*

Einen Beweis dieses Satzes findet man in [7].

Bemerkung 13.9 Ähnliche Matrizen besitzen das gleiche charakteristische Polynom, also insbesondere das gleiche Spektrum.

Beweis: Aus $A, B \in \mathbb{K}^{n\times n}$ mit $B = S^{-1}AS$ und $S \in \mathbb{K}^{n\times n}$ regulär folgt

$$\begin{aligned} p_B &= \det(B - XE_n) = \det(S^{-1}AS - XS^{-1}S) = \det(S^{-1}(A - XE_n)S) = \\ &= \det S^{-1} \det(A - XE_n) \det S = \det(A - XE_n) = p_A. \end{aligned}$$

■

Damit können wir auch das **charakteristische Polynom eines Endomorphismus** $\Phi : V \to V$ definieren als das charakteristische Polynom einer (beliebigen) Abbildungsmatrix A von Φ:

$$p_\Phi := \det(\Phi - X\mathrm{id}_V) = \det(A - XE_n).$$

Satz 13.7 gilt dann entsprechend auch für Endomorphismen.

Beispiel 13.10

1. Sei $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\mathbb{K} = \mathbb{R}$.

   Es ist $p_A = \begin{vmatrix} 1-X & 1 \\ 0 & 1-X \end{vmatrix} = (1-X)^2$, also ist $\lambda = 1$ einziger Eigenwert von A. Für den Eigenraum erhalten wir $E_1 = [\begin{pmatrix} 0 \\ 1 \end{pmatrix}]$.

2. Sei $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{K}^{2\times 2}$.

   Dann ist $p_A = \begin{vmatrix} -X & 1 \\ -1 & -X \end{vmatrix} = X^2 + 1$.

   An diesem Beispiel sehen wir, dass die Existenz von Eigenwerten von dem zugrunde gelegten Körper abhängt.

   Für $\mathbb{K} = \mathbb{R}$ existiert keine Nullstelle von p_A, also auch kein Eigenwert von A.

   Für $\mathbb{K} = \mathbb{C}$ ergeben sich die Eigenwerte $\lambda_1 = \mathrm{i}$ und $\lambda_2 = -\mathrm{i}$. Die zugehörigen Eigenräume sind

   $$E_{\lambda_1} = [\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix}] \quad \text{und} \quad E_{\lambda_2} = [\begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix}].$$

   Für $\mathbb{K} = \mathbb{F}_2 = \{0, 1\}$ ist $\lambda = 1$ einziger Eigenwert mit zugehörigem Eigenraum $[\begin{pmatrix} 1 \\ 1 \end{pmatrix}]$.

3. Gegeben sei die reelle Matrix

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

   Für das charakteristische Polynom ergibt sich:

   $$p_A = \begin{vmatrix} -X & -1 & 1 & 1 \\ -1 & 1-X & -2 & 3 \\ 2 & -1 & -X & 0 \\ 1 & -1 & 1 & -X \end{vmatrix} = (1+X)^2(1-X)(2-X).$$

Also sind $\lambda_1 = -1, \lambda_2 = 1$ und $\lambda_3 = 2$ die Eigenwerte von A. Die zugehörigen Eigenräume sind

$$E_{\lambda_1} = \left[\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right], \quad E_{\lambda_2} = \left[\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right], \quad E_{\lambda_3} = \left[\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right].$$

14 Diagonalisierbare Endomorphismen

Definition 14.1 Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ heißt **diagonalisierbar**, wenn sie zu einer Diagonalmatrix

$$\widetilde{A} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

ähnlich ist. Ein Endomorphismus Φ heißt **diagonalisierbar**, wenn es eine Abbildungsmatrix von Φ gibt, die Diagonalgestalt hat.

Bemerkung 14.2 Da alle Abbildungsmatrizen eines Endomorphismus Φ ähnlich sind, ist im Falle der Diagonalisierbarkeit von Φ *jede* Abbildungsmatrix von Φ diagonalisierbar und außerdem zur *gleichen* Diagonalmatrix ähnlich.

Satz 14.3 (1. Kriterium für Diagonalisierbarkeit eines Endomorphismus) *Für einen Endomorphismus Φ eines n-dimensionalen $\mathbb{K}$-Vektorraumes V sind folgende Aussagen äquivalent:*

(a) *Φ ist diagonalisierbar.*

(b) *In V gibt es eine Basis aus Eigenvektoren von Φ.*

(c) *V ist die direkte Summe der Eigenräume von Φ.*

(d) *Die Summe der Dimensionen der Eigenräume von Φ ist n.*

Beweis: (a) $\Longrightarrow$ (b) Nach Definition der Diagnalisierbarkeit gibt es eine Basis $B = \{v_1, \dots, v_n\}$ von V, bezüglich der die Abbildungsmatrix von Φ Diagonalgestalt

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

(mit nicht notwedig verschiedenenen $\lambda_1, \dots, \lambda_n$) hat. Also gilt $\Phi(v_i) = \lambda_i v_i$ für $i = 1, \dots, n$, d.h. die Basisvektoren v_i sind Eigenvektoren von Φ.

(b) $\Longrightarrow$ (c) Es seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von Φ. Dann ist nach Satz 13.3 die Summe der Eigenräume von Φ direkt, und es gilt $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \subset V$. Sei umgekehrt v ein beliebiger Vektor aus V. Bezüglich der Basis $\{v_1, \dots, v_n\}$ aus

Eigenvektoren von Φ gilt dann die Darstellung $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Fassen wir alle Eigenvektoren zum gleichen Eigenwert λ_i zusammen, so erhalten wir $v = \tilde{v}_1 + \cdots + \tilde{v}_k$ mit $\tilde{v}_i \in E_{\lambda_i}$, $i = 1, \ldots, k$.

(c) $\Longrightarrow$ (d) ist klar.

(d) $\Longrightarrow$ (a) Es seien $E_{\lambda_1}, \ldots, E_{\lambda_k}$ die Eigenräume von Φ und $\dim E_{\lambda_i} = n_i$ für $i = 1, \ldots, k$. In jedem Eigenraum E_{λ_i} wählen wir eine Basis B_i. Dann ist $B := B_1 \cup \cdots \cup B_k$ nach Satz 13.3 linear unabhängig, und wegen $n_1 + \cdots + n_k = n$ ist B sogar eine Basis von V. Bezüglich dieser Basis hat die Abbildungsmatrix von Φ Diagonalgestalt. ∎

Bemerkung 14.4 Satz 14.3 gilt analog, wenn wir Φ durch $A \in \mathbb{K}^{n\times n}$ und V durch $\mathbb{K}^n$ ersetzen.

Folgerung 14.5 *Ein Endomorphismus Φ eines n-dimensionalen $\mathbb{K}$-Vektorraumes (bzw. eine $(n \times n)$-Matrix A) mit n verschiedenen Eigenwerten ist diagonalisierbar.*

Die Frage nach der Diagonalisierbarkeit eines Endomorphismus bzw. einer quadratischen Matrix lässt sich auch mit Hilfe des charakteristischen Polynoms entscheiden.

Satz 14.6 (2. Kriterium für Diagonalisierbarkeit eines Endomorphismus) *Es seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\Phi : V \to V$ ein Endomorphismus. Dann gilt:*

Φ ist genau dann diagonalisierbar, wenn sein charakteristisches Polynom p_Φ in der Form

$$p_\Phi = (-1)^n (X - \lambda_1)^{r_1} \cdots (X - \lambda_k)^{r_k} \qquad (*)$$

darstellbar ist mit $r_i \in \mathbb{N}$ und paarweise verschiedenen $\lambda_i \in \mathbb{K}$ und wenn für $i = 1, \ldots, k$ gilt:

$$\dim \text{Bild}\,(\Phi - \lambda_i \,\text{id}_V) = n - r_i.$$

Bemerkung 14.7 (a) Es gilt:

$$\dim \text{Bild}\,(\Phi - \lambda_i \,\text{id}_V) = n - r_i \Longleftrightarrow \dim E_{\lambda_i} = \dim \text{Kern}\,(\Phi - \lambda_i \,\text{id}_V) = r_i.$$

(b) Besitzt ein Polynom p die Darstellung $(*)$, so sagt man, dass p in **Linearfaktoren zerfällt**. Die Zahl r_i in der Darstellung $(*)$ nennt man **algebraische Vielfachheit** des Eigenwertes λ_i. Die Dimension des Eigenraumes E_{λ_i} heißt **geometrische Vielfachheit** des Eigenwertes λ_i. Satz 14.6 kann man also kurz auch so formulieren:

Ein Endomorphismus Φ ist genau dann diagonalisierbar, wenn das charakteristische Polynom p_Φ in Linearfaktoren zerfällt, und wenn für alle Eigenwerte die geometrische und die algebraische Vielfachheit gleich sind.

Beweis: Es sei Φ diagonalisierbar. Dann besitzt Φ eine Abbildungsmatrix A der Form

$$A = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_k & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{matrix}} \end{pmatrix},$$

wobei die λ_i paarweise verschieden sind mit Vielfachheit (=Kästchengröße) r_i. Somit gilt

$$\det(A - XE_n) = (\lambda_1 - X)^{r_1} \cdots (\lambda_k - X)^{r_k} = (-1)^n (X - \lambda_1)^{r_1} \cdots (X - \lambda_k)^{r_k}$$

und $\dim \operatorname{Bild}(\Phi - \lambda_i \operatorname{id}_V) = \operatorname{Rang}(A - \lambda_i E_n) =$

$$= \operatorname{Rang} \begin{pmatrix} \lambda_1 - \lambda_i & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_1 - \lambda_i & & & & \\ & & & \ddots & & & \\ & & & & \lambda_k - \lambda_i & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_k - \lambda_i \end{pmatrix} = n - r_i.$$

Ist umgekehrt $p_\Phi = (-1)^n (X - \lambda_1)^{r_1} \cdots (X - \lambda_k)^{r_k}$, so sind $\lambda_1, \ldots, \lambda_k$ gerade die Eigenwerte von Φ. Wegen $\dim \operatorname{Bild}(\Phi - \lambda_i \operatorname{id}_V) = n - r_i$ gilt $\dim E_{\lambda_i} = r_i$ (vgl. Bemerkung 14.7) und somit $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = r_1 + \cdots + r_k = \operatorname{grad} p_\Phi = n$. Nach Satz 14.3 ist daher der Endomorphismus Φ diagonalisierbar. ■

Wir setzten jetzt die Beispiele aus 13.10 fort.

Beispiel 14.8 1. Wir betrachten die Matrix

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

mit dem charakteristischen Polynom $p_A = (1-X)^2$. Wegen $\operatorname{Rang}(A - E_2) = 1 \neq 2 - 2 = 0$ ist A nicht diagonalisierbar.

2. Die Matrix
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
ist über $\mathbb{R}$ nicht diagonalisierbar. Über $\mathbb{C}$ gilt $p_A = (\mathrm{i} - X)(-\mathrm{i} - X)$, also hat A zwei verschiedene Eigenwerte und ist über $\mathbb{C}$ diagonalisierbar. Über dem Körper $\mathbb{F}_2$ gilt $p_A = (X+1)^2$. Also ist 1 einziger Eigenwert von A (mit algebraischer Vielfachheit 2) und $\dim E_1 = 1$. Somit ist $A \in \mathbb{F}^{2\times 2}$ nicht diagonalisierbar.

3. Die reelle Matrix
$$A = \begin{pmatrix} 0 & -1 & 1 & 1 \\ -1 & 1 & -2 & 3 \\ 2 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix}$$
hat das charakteristische Polynom $p_A = (1+X)^2(1-X)(2-X)$ und die Eigenwerte $\lambda_1 = -1, \lambda_2 = 1$ und $\lambda_3 = 2$ mit den (algebraischen) Vielfachheiten $r_1 = 2, r_2 = 1$ und $r_3 = 1$. Wir wissen schon, dass $\dim E_{\lambda_1} = 1$ ist. Wegen $1 \neq r_1$ ist also A nicht diagonalisierbar.

4. Gegeben sei die reelle Matrix
$$A = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ 0 & 0 & 2 \end{pmatrix}$$
mit dem charakteristischen Polynom $p_A = (2-X)^2(7-X)$. Wegen
$$\begin{aligned} \lambda_1 &= 2, r_1 = 2, \ \operatorname{Rang}(A - \lambda_1 E) = 1 = 3 - 2, \\ \lambda_2 &= 7, r_1 = 1, \ \operatorname{Rang}(A - \lambda_2 E) = 2 = 3 - 1, \end{aligned}$$
ist A diagonalisierbar.

Bemerkung 14.9 Ist $A \in \mathbb{K}^{n\times n}$ diagonalisierbar und ist $\{v_1, v_2, \ldots, v_n\}$ eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A mit $Av_i = \lambda_i v_i$ für $i = 1, \ldots, n$, so gilt für die reguläre Matrix $S = (v_1 \mid v_2 \mid \cdots \mid v_n)$ mit Spalten v_i
$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Beweis: Die Matrix S beschreibt den Basiswechsel zwischen $\{v_1, \ldots, v_n\}$ und der Standardbasis. Die obige Diagonalmatrix ist gerade die Abbildungsmatrix von $x \mapsto Ax$ bezüglich der (neuen) Basis $\{v_1, v_2, \ldots, v_n\}$. ■

Beispiel 14.10 Im Beispiel 14.8 (4) sind die Eigenräume gegeben durch

$$E_{\lambda_1} = \left[\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right], \quad E_{\lambda_2} = \left[\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right].$$

Also ist

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

eine solche Transformationsmatrix. Es gilt dann

$$S^{-1}AS = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

15 Trigonalisierbare Endomorphismen

Definition 15.1 Ein Endomorphismus $\Phi : V \to V$ eines endlich dimensionalen $\mathbb{K}$-Vektorraumes V heißt **trigonalisierbar**, wenn eine Basis von V existiert bezüglich der Φ durch eine (obere) Dreiecksmatrix dargestellt wird.

Satz 15.2 *Ein Endomorphismus $\Phi \in \mathrm{Hom}\,(V, V)$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom p_Φ in Linearfaktoren zerfällt.*

Beweis: Die Abbildung Φ werde bezüglich einer geeigneten Basis durch eine Dreiecksmatrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

dargestellt. Dann ist $p_\Phi = (a_{11} - X) \cdots (a_{nn} - X)$, d.h. das charakteristische Polynom zerfällt in Linearfaktoren.

Zerfällt umgekehrt das charakteristische Polynom von Φ in Linearfaktoren, so beweisen wir mit Induktion nach n, dass Φ trigonalisierbar ist. Für $n = 1$ ist nichts zu beweisen. Es sei $n \geq 2$. Dann existiert nach Voraussetzung mindestens ein Eigenwert λ_1 von Φ. Es sei v_1 ein Eigenvektor zu λ_1. Ergänzen wir v_1 zu einer Basis $\{v_1, v_2, \ldots, v_n\}$ von V, so stellt sich Φ durch eine Matrix der Form

$$A = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

dar. Wir betrachten nun die lineare Selbstabbildung $\Psi : [v_2, v_3, \ldots, v_n] \to [v_2, v_3, \ldots, v_n]$, die bezüglich der Basis $\{v_2, \ldots, v_n\}$ durch die Matrix

$$A = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ a_{32} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

gegeben ist. Ist p_Ψ das charakteristische Polynom von Ψ, so gilt $p_\Phi = (\lambda_1 - X) p_\Psi$. Da p_Φ in Linearfaktoren zerfällt, gilt dies auch für p_Ψ. Nach Induktionsvoraussetzung existiert in $[v_2, \ldots, v_n]$ eine Basis $\{w_2, \ldots, w_n\}$ bezüglich der Ψ durch eine Dreiecksmatrix dargestellt wird. Bezüglich der Basis $\{v_1, w_2, \ldots, w_n\}$ von V wird dann auch Φ durch eine Dreiecksmatrix dargestellt. ■

Folgerung 15.3 *Jeder Endomorphismus eines endlich dimensionalen* $\mathbb{C}$*-Vektorraumes ist trigonalisierbar.*

Beweis: Dies ist eine direkte Folge des Fundamentalsatzes der Algebra 13.8. ■

Beispiel 15.4 Die lineare Selbstabbildung $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ sei bezüglich der Standardbasis $\{e_1, e_2, e_3\}$ gegeben durch die Matrix

$$\begin{pmatrix} -2 & 1 & 3 \\ 2 & 1 & -1 \\ -7 & 2 & 7 \end{pmatrix}.$$

Man erhält dann $p_\Phi = -(X-2)^3$. Nach Satz 15.2 ist Φ trigonalisierbar, aber nach Satz 14.6 nicht diagonalisierbar, da die geometrische Vielfachheit des Eigenwertes 2 nur 1 beträgt. Der Vektor $v_1 := (1, 1, 1)$ ist Eigenvektor zum Eigenwert 2 und $\{v_1, e_2, e_3\}$ ist eine Basis von $\mathbb{R}^3$. Bezüglich dieser Basis wird Φ dargestellt durch die Matrix

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

Die Matrix

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

beschreibt in dem Unterraum $[e_2, e_3]$ eine Selbstabbildung mit Eigenwert 2. Der Vektor $v_2 := -2e_2 + e_3$ ist Eigenvektor zum Eigenwert 2. Wählt man in $\mathbb{R}^3$ die Basis $\{v_1, v_2, e_3\}$, so stellt sich Φ dar durch die Matrix

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Damit ist Φ triagonalisiert.

16 Der Satz von Cayley-Hamilton

Wir präzisieren zunächst, was unter dem Einsetzen eines Endomorphismus bzw. einer quadratischen Matrix in ein Polynom zu verstehen ist.

Definition 16.1 Es sei V ein $\mathbb{K}$-Vektorraum, Φ ein Endomorphismus von V sowie

$$p = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n \in \mathbb{K}[X]$$

ein Polynom. Dann sei $p(\Phi)$ der Endomorphismus

$$p(\Phi) := a_0 \operatorname{id}_V + a_1 \Phi + a_2 \Phi^2 + \cdots + a_n \Phi^n \in \operatorname{Hom}(V, V),$$

wobei $\Phi^k = \Phi \circ \Phi \circ \cdots \circ \Phi$ (k Faktoren) gilt.
Analog gilt für $A \in \mathbb{K}^{n \times n}$

$$p(A) := a_0 E_n + a_1 A + a_2 A^2 + \cdots + a_n A^n \in \mathbb{K}^{n \times n},$$

wobei $A^k = A \cdot A \cdots A$ (k Faktoren) gilt.

Beispiel 16.2 Ist $p = 3 + X + X^5$, Φ ein Endomorphismus bzw. A eine quadratische Matrix, so ist

$$p(\Phi) = 3 \operatorname{id}_V + \Phi + \Phi^5 \quad \text{bzw.} \quad p(A) = 3E_n + A + A^5.$$

Bemerkung 16.3 Sei $\Phi \in \operatorname{Hom}(V, V)$ ein fest gewählter Endomorphismus. Dann ist die Abbildung

$$\mathbb{K}[X] \to \operatorname{Hom}(V, V), \quad p \mapsto p(\Phi)$$

ein Homomorphismus sowohl bezüglich der Vektorraum- als auch der Ringstruktur, der sogenannte **Einsetzungshomomorphismus**. Denn es gilt für alle $p, q \in \mathbb{K}[X]$ und alle $\lambda \in \mathbb{K}$

$$\begin{aligned} (p+q)(\Phi) &= p(\Phi) + q(\Phi), \\ (\lambda \cdot p)(\Phi) &= \lambda \cdot p(\Phi), \\ (p \cdot q)(\Phi) &= p(\Phi) \circ q(\Phi) = (q \cdot p)(\Phi). \end{aligned}$$

Um die letzte Eigenschaft zu beweisen, genügt es, diese auf der Basis $\{1, X, X^2, \ldots\}$ von $\mathbb{K}[X]$ nachzuprüfen.

Wir wollen jetzt den Endomorphismus Φ in ein spezielles Polynom, nämlich das charakteristische Polynom von Φ, einsetzen.

Satz 16.4 (Cayley-Hamilton) [5] *Es seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum, Φ ein Endomorphismus von V und p_Φ das charakteristische Polynom von Φ. Dann gilt $p_\Phi(\Phi) = 0$ (=Nullabbildung).*

Beweis: Wir schreiben kurz p für das charakteristische Polynom p_Φ. Es ist zu zeigen, dass $p(\Phi)(x) = 0$ für alle $x \in V$ gilt. Für $x = 0$ ist dies klar. Sei also $x \neq 0$ (fest). Wir betrachten nun für jedes $m \in \mathbb{N}_0$ die Vektoren

$$\Phi^0(x) = x, \Phi(x), \Phi^2(x), \ldots, \Phi^m(x),$$

dabei haben wir $\Phi^0 := \mathrm{id}_V$ gesetzt. Für $m = 0$ besteht diese Menge nur aus dem Vektor $x \neq 0$ und ist somit linear unabhängig. Für $m \geq n$ ist sie linear abhängig. Also existiert ein kleinstes $m \in \mathbb{N}$, für das diese Vektoren linear abhängig sind. Dann ist

$$\tilde{B} := \{x, \Phi(x), \Phi^2(x), \ldots, \Phi^{m-1}(x)\}$$

linear unabhängig und $\tilde{B} \cup \{\Phi^m(x)\}$ linear abhängig. Es existieren also (von x abhängige) Skalare $a_0, a_1, \ldots, a_{m-1} \in \mathbb{K}$ mit

$$\Phi^m(x) = a_0 x + a_1 \Phi(x) + \cdots + a_{m-1}\Phi^{m-1}(x).$$

Sei

$$q := a_0 + a_1 X + \cdots + a_{m-1}X^{m-1} - X^m \in \mathbb{K}[X].$$

Dann ist $q(\Phi) \in \mathrm{Hom}\,(V, V)$, und es gilt $q(\Phi)(x) = 0$. Beachten Sie, dass die a_i und damit auch das Polynom q von x abhängig ist.

Wir wollen jetzt zeigen, dass auch $p(\Phi) = 0$ ist. Dazu betrachten wir den Untervektorraum $U := [\tilde{B}]$. Wegen

$$\Phi(\tilde{B}) = \Phi(\{x, \ldots, \Phi^{m-1}(x)\}) = \{\Phi(x), \ldots, \Phi^m(x)\} \subset U$$

ist $\Phi(U) \subset U$, d.h. U ist invariant unter Φ und $\tilde{\Phi} := \Phi\,|_U$ ist ein Endomorphismus von U. Bezüglich der (geordneten) Basis $\tilde{B}$ hat $\tilde{\Phi}$ die Abbildungsmatrix

$$\tilde{A} = \begin{pmatrix} 0 & 0 & \ldots & 0 & a_0 \\ 1 & 0 & \ldots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{m-2} \\ 0 & \ldots & 0 & 1 & a_{m-1} \end{pmatrix}$$

Wir bestimmen das charakteristische Polynom $p_{\tilde{\Phi}}$ von $\tilde{\Phi}$.

[5] Arthur Cayley (1821-1895), William Rowan Hamilton (1805-1865)

$$p_{\tilde{\Phi}} = p_{\tilde{A}} = \det(\tilde{A} - XE_m) = \begin{vmatrix} -X & 0 & \dots & 0 & a_0 \\ 1 & -X & \dots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -X & a_{m-2} \\ 0 & \dots & 0 & 1 & a_{m-1} - X \end{vmatrix}$$

$$= \begin{vmatrix} 0 & \dots & \dots & 0 & a_0 + a_1X + \dots + a_{m-1}X^{m-1} - X^m \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{m-2} + a_{m-1}X - X^2 \\ 0 & \dots & 0 & 1 & a_{m-1} - X \end{vmatrix}.$$

Entwicklung nach der 1. Zeile ergibt $p_{\tilde{A}} = (-1)^{m+1}q$.

Wir ergänzen nun $\tilde{B}$ zu einer Basis B von V. Der Endomorphismus Φ hat bezüglich B eine Abbildungsmatrix A_Φ der Form

$$A_\Phi = \begin{pmatrix} \tilde{A} & C \\ O & D \end{pmatrix}$$

mit geeigneten Matrizen C und D. Für das charakteristische Polynom p von Φ folgt nach dem Kästchenmultplikationssatz für Determinanten (vgl. dazu die Übungsaufgaben zu Determinanten)

$$p = p_\Phi = p_D \cdot p_{\tilde{A}} = p_D \cdot (-1)^{m+1} \cdot q,$$

wobei p_D das charakteristische Polynom von D ist. Setzen wir

$$r := (-1)^{m+1} p_D,$$

so gilt $p = r \cdot q$ und somit $p(\Phi) = r(\Phi) \circ q(\Phi)$. Daraus folgt

$$p(\Phi)(x) = r(\Phi)(q(\Phi)(x)) = 0.$$

Da $x \in V$ beliebig war, gilt $p(\Phi) = 0$ (=Nullabbildung). ■

Bemerkung 16.5 Satz 16.4 gilt analog für quadratische Matrizen $A \in \mathbb{K}^{n \times n}$, d.h. es ist $p_A(A) = O$ (= Nullmatrix).

Folgerung 16.6 *Sei $A \in \mathbb{K}^{2\times 2}$. Dann gilt: $A^2 - (\mathrm{Spur}\, A)A + (\det A)\, E_2 = O$.*

Beweis: Das kann man durch explizites Nachrechnen zeigen. Oder man benutzt den Satz von Cayley-Hamilton, denn das charakteristische Polynom für eine (2×2)-Matrix A ist $p_A = X^2 - (\mathrm{Spur}\, A)\, X + (\det A)\, E_2$. ■

17 Die Jordansche Normalform

Allgemeine Voraussetzung in diesem Kapitel:

V ist ein endlich-dimensionaler *komplexer* Vektorraum.

Ist Φ ein Endomorphismus von V, so zerfällt das charakteristische Polynom von Φ nach dem Fundamentalsatz der Algebra in Linearfaktoren:

$$p_\Phi = (-1)^n (X - \lambda_1)^{r_1} \cdots (X - \lambda_k)^{r_k}.$$

Einige Ergebnisse dieses Kapitels gelten allgemeiner für Endomorphismen von $\mathbb{K}$-Vektorräumen, deren charakteristisches Polynom in Linearfaktoren zerfällt. Wie wir gesehen haben, sind solche Endomorphismen trigonalisierbar, d.h. es gibt eine Abbildungsmatrix $A \in \mathbb{C}^{n\times n}$ von Φ der Form

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & \lambda_r \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r \end{pmatrix} + \begin{pmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{pmatrix} =: D + N.$$

Dabei ist also D eine Diagonalmatrix (deren Einträge gerade die Eigenwerte von Φ sind) und N ist eine obere Dreicksmatrix mit Nullen auf der Diagonalen.

Das Ziel dieses Kapitels ist es, genauer zu untersuchen, was man über die Matrix N, also die Einträge oberhalb der Diagonalen sagen kann. Dies führt schließlich auf eine vollständige Übersicht (Klassifikation) aller komplexen Endomorphismen bzw. aller komplexen $n \times n$ Matrizen durch Normalformen (vgl. Satz 17.12).

Als Motivation für die folgenden allgemeinen Definitionen weisen wir noch auf eine spezielle Eigenschaft von oberen Dreicksmatrizen mit Nullen auf der Diagonalen hin. Für

$$N = \begin{pmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

ist

$$N^2 = N \cdot N = \begin{pmatrix} 0 & 0 & \star & \cdots & \star \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \star \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}, \quad \ldots \quad , \quad N^n = 0.$$

Definition 17.1 Ein Endomorphismus Φ eines $\mathbb{K}$-Vektorraumes (bzw. eine Matrix $A \in \mathbb{K}^{n\times n}$) heißt **nilpotent**, falls ein $k \in \mathbb{N}$ existiert, so dass $\Phi^k = 0$ (bzw. $A^k = 0$) ist.

17.1 Verallgemeinerte Eigenräume

Es seien V ein n-dimensionaler komplexer Vektorraum und Φ ein Endomorphismus von V. Für einen Eigenwert $\lambda \in \mathbb{C}$ von Φ und $k \in \mathbb{N}_0$ definieren wir

$$K_k := \text{ Kern } (\Phi - \lambda \text{ id})^k.$$

Dabei ist $(\Phi - \lambda \text{ id})^k$ für $k \in \mathbb{N}$ die k-fache Verkettung $(\Phi - \lambda \text{ id}) \circ (\Phi - \lambda \text{ id}) \circ \cdots \circ (\Phi - \lambda \text{ id})$ und $(\Phi - \lambda \text{ id})^0$ die identische Abbildung $\text{id} = \text{id}_V$ von V. Insbesondere ist also $K_0 = \{0\}$, und $K_1 = E_\lambda$, der Eigenraum zum Eigenwert λ von Φ.

Hilfssatz 17.2 *Für die Kerne K_k $(k = 0, 1, 2, \ldots)$ zum Eigenwert λ gilt:*

1. *Alle K_k sind Φ-invariant, d.h. $\Phi(K_k) \subseteq K_k$.*
2. *Es ist $K_0 \subset K_1 \subseteq K_2 \subseteq \cdots$.*

Beweis:

1. Um zu zeigen, dass K_k invariant ist unter Φ, wählen wir $x \in K_k$ beliebig aus. Es gilt also $(\Phi - \lambda \text{ id})^k(x) = 0$. Daraus folgt
$$(\Phi - \lambda \text{ id})\big((\Phi - \lambda \text{ id})^k(x)\big) = (\Phi - \lambda \text{ id})^k\big((\Phi - \lambda \text{ id})(x)\big) = 0.$$
Somit ist $(\Phi - \lambda \text{ id})(x) = \Phi(x) - \lambda x \in K_k$. Da $x \in K_k$ ist, ergibt sich $\Phi(x) \in K_k$.
2. Wenn $x \in K_k$ für ein $k \in \{0, 1, 2, \ldots\}$ ist, so gilt $(\Phi - \lambda \text{ id})^k(x) = 0$ und damit auch
$$(\Phi - \lambda \text{ id})^{k+1}(x) = (\Phi - \lambda \text{ id})\big((\Phi - \lambda\text{id})^k(x)\big) = 0.$$
Also haben wir $x \in K_{k+1}$. Da x beliebig war, folgt $K_k \subseteq K_{k+1}$. ■

In Hilfssatz 17.2 ist $K_0 \neq K_1$ weil $K_1 = E_\lambda \neq \{0\} = K_0$. Es sei q diejenige natürliche Zahl, für die zum ersten Mal $K_q = K_{q+1}$ gilt. Eine solche Zahl muss es geben, denn wären alle K_k voneinander verschieden, so wäre $\dim K_{n+1} \geq n + 1 > \dim V$, was nicht möglich ist.

Hilfssatz 17.3 *Falls für die Kerne K_k zum Eigenwert λ und $1 \leq q \leq n$ gilt*

$$K_0 \subsetneq K_1 \subsetneq \cdots \subsetneq K_q = K_{q+1},$$

so ist

$$K_q = K_{q+j}, \qquad j = 1, 2, 3, \ldots.$$

Beweis: (durch vollständige Induktion nach j): Für $j = 1$ gilt die Behauptung nach Voraussetzung. Es gelte also, dass $K_q = K_{q+1} = \cdots = K_{q+j}$ für ein $j \geq 1$. Wir beweisen $K_{q+j} = K_{q+j+1}$:

Für $x \in K_{q+j+1}$ gilt

$$(\Phi - \lambda \text{ id})^{q+j+1}(x) = (\Phi - \lambda \text{ id})^{q+j}\big((\Phi - \lambda \text{ id})(x)\big) = 0.$$

Also ist $(\Phi - \lambda \text{ id})(x) \in K_{q+j} = K_{q+j-1}$ und somit

$$(\Phi - \lambda \text{ id})^{q+j}(x) = (\Phi - \lambda \text{ id})^{q+j-1}\big((\Phi - \lambda \text{ id})(x)\big) = 0.$$

Demnach ist $x \in K_{q+j}$ und also $K_{q+j+1} \subset K_{q+j}$. Nach Hilfssatz 17.2 ist andererseits $K_{q+j} \subset K_{q+j+1}$. Damit haben wir $K_{q+j} = K_{q+j+1}$. ■

Aufgrund der Hilfssätze 17.2 und 17.3 können wir nun die Begriffe "Eigenvektor" und "Eigenraum" eines Endomorphismus Φ von V wie folgt verallgemeinern:

Definition 17.4 Ein Vektor $v \in V$ mit $v \neq 0$ heißt **Hauptvektor** (oder **verallgemeinerter Eigenvektor**) von Φ zum Eigenwert λ, wenn es eine Zahl $k \in \mathbb{N}$ gibt mit

$$(\Phi - \lambda \text{ id})^k(v) = 0.$$

Für $q \in \mathbb{N}$ wie in Hilfssatz 17.3 heißt $K_q =$ Kern $(\Phi - \lambda \text{ id})^q$ der zum Eigenwert λ gehörige **Hauptraum** (oder **verallgemeinerte Eigenraum**) H_λ von Φ. Diese Zahl q heißt **Index** von H_λ.

Die Menge der Hauptvektoren von Φ zum Eigenwert λ ist also gleich $H_\lambda \backslash \{0\}$ und enthält nach Hilfssatz 17.3 insbesondere alle Eigenvektoren zum Eigenwert λ.

Um den zu einem Eigenwert λ gehörigen Hauptraum H_λ von Φ konkret zu bestimmen, kann man Hilfssatz 17.3 verwenden: Wegen $q \leq n = \dim V$ ist

$$H_\lambda = K_n = \text{ Kern } (\Phi - \lambda \text{ id})^n.$$

Wenn diese Methode bei großem n zu mühsam ist, so betrachtet man die Folge der Kerne $K_0, K_1, K_2, \ldots$ und ermittelt die Zahl q, von der an die Folge konstant ist. Es ist dann

$$H_\lambda = K_q = \text{ Kern } (\Phi - \lambda \text{ id})^q.$$

Beispiel 17.5 Für den Endomorphismus Φ von $\mathbb{C}^3$, der bezüglich der kanonischen Basis von $\mathbb{C}^3$ die Abbildungsmatrix

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

hat, sollen die Haupträume ermittelt werden. Man stellt zunächst fest, dass Φ die beiden Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 4$ besitzt.

Um H_1 zu bestimmen, benutzen wir, dass $H_1 = K_3 =$ Kern $(\Phi - \mathrm{id})^3$. Es ist

$$(B-E)^3 = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}^3 = \begin{pmatrix} 0 & 18 & 39 \\ 0 & 27 & 27 \\ 0 & 0 & 27 \end{pmatrix},$$

also $H_1 = [b_1]$ mit $b_1 := (1, 0, 0)$.

Zur Bestimmung von H_4 berechnen wir

$$B - 4E = \begin{pmatrix} -3 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (B-4E)^2 = \begin{pmatrix} 9 & -6 & -7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$(B-4E)^3 = \begin{pmatrix} -27 & 18 & 21 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Seien $b_2 := (2, 3, 0)$ und $b_3 := (7, 0, 9)$. Dann erhalten wir für den Eigenwert $\lambda_2 = 4$

$$\begin{aligned} K_1 &= \text{Kern } (\Phi - 4\ \mathrm{id}) = [b_2] \\ K_2 &= \text{Kern } (\Phi - 4\ \mathrm{id})^2 = [b_2, b_3] \\ K_3 &= \text{Kern } (\Phi - 4\ \mathrm{id})^3 = [b_2, b_3] \end{aligned}$$

Also ist $q = 2$ und $H_4 = [b_2, b_3]$.

17.2 Die Hauptraum-Zerlegung

Hilfssatz 17.6 *Sei H_λ ein Hauptraum von Φ mit dem Index q. Der Bildraum $B_\lambda :=$* Bild $(\Phi - \lambda \mathrm{id})^q$ *ist ein Vektorraum-Komplement von $H_\lambda =$* Kern $(\Phi - \lambda\mathrm{id})^q$, *d.h. es gilt*

$$V = H_\lambda \oplus B_\lambda.$$

Außerdem ist B_λ invariant unter Φ.

Beweis: (i) B_λ ist Φ-invariant:

Ist $x \in B_\lambda$, so gibt es ein $y \in V$ mit $x = (\Phi - \lambda\ \mathrm{id})^q(y)$. Daraus folgt

$$(\Phi - \lambda\ \mathrm{id})(x) = (\Phi - \lambda\ \mathrm{id})^q((\Phi - \lambda\ \mathrm{id})(y)) \in B_\lambda.$$

Aus $(\Phi - \lambda\ \mathrm{id})(x) = \Phi(x) - \lambda x \in B_\lambda$ erhält man dann $\Phi(x) \in B_\lambda$, also $\Phi(B_\lambda) \subset B_\lambda$.

(ii) $H_\lambda \cap B_\lambda = \{0\}$:

Für $x \in H_\lambda \cap B_\lambda$ gilt wegen $x \in H_\lambda$ einerseits $(\Phi - \lambda\ \mathrm{id})^q(x) = 0$. Andererseits gibt es wegen $x \in B_\lambda$ ein $y \in V$ mit

$$x = (\Phi - \lambda\ \mathrm{id})^q(y). \qquad (*)$$

Also folgt $0 = (\Phi - \lambda\ \mathrm{id})^q(x) = (\Phi - \lambda\ \mathrm{id})^{2q}(y)$, d.h. $y \in H_\lambda$ nach Hilfssatz 17.3, und aus $(*)$ erhält man $x = 0$.

(iii) $V = H_\lambda + B_\lambda$:

Für $B_\lambda = \{0\}$, ist $(\Phi - \lambda\ \mathrm{id})^q$ die Nullabbildung und $H_\lambda = K_q = V$. Ist $B_\lambda \neq \{0\}$, so können wir eine Basis $\{v_1, \dots, v_s\}$ von H_λ und eine Basis $\{b_1, \dots, b_t\}$ von B_λ wählen. Die Menge $\{v_1, \dots, v_s, b_1, \dots, b_t\}$ ist linear unabhängig, denn sonst wäre $H_\lambda \cap B_\lambda \neq \{0\}$. Der Untervektorraum $H_\lambda + B_\lambda$ hat also die Dimension $s + t$. Weiter gilt nach Satz 9.21, dass auch

$$\dim V = \dim \mathrm{Kern}\,(\Phi - \lambda\ \mathrm{id})^q + \dim \mathrm{Bild}\,(\Phi - \lambda\ \mathrm{id})^q = s + t.$$

Also ist $H_\lambda + B_\lambda = V$. ■

Hilfssatz 17.7 *Seien Φ ein Endomorphismus von V und*

$$p_\Phi = (-1)^n (X - \lambda_1)^{r_1} (X - \lambda_2)^{r_2} \cdots (X - \lambda_k)^{r_k}$$

das charakteristische Polynom von Φ. Weiter sei $V = H_{\lambda_1} \oplus B_{\lambda_1}$ wie in Hilfssatz 17.6. Dann gilt für die charakteristischen Polynome p_1 von $\Phi\,|_{H_{\lambda_1}}$ bzw. q_1 von $\Phi\,|_{B_{\lambda_1}}$:

$$p_1 = (-1)^{r_1} (X - \lambda_1)^{r_1} \quad \textit{und} \quad q_1 = (-1)^{n - r_1} (X - \lambda_2)^{r_2} \cdots (X - \lambda_k)^{r_k}.$$

Beweis: Da V direkte Summe von H_{λ_1} und B_{λ_1} ist, gilt $p_\Phi = p_1 \cdot q_1$. Es genügt also zu zeigen, dass λ_1 der einzige Eigenwert von $\Phi\,|_{H_{\lambda_1}}$ ist und dass die Eigenwerte von $\Phi\,|_{B_{\lambda_1}}$ alle von λ_1 verschieden sind.

Sei λ ein Eigenwert von $\Phi\,|_{H_{\lambda_1}}$, also $\Phi(x) = \lambda x$ für ein $x \neq 0$. Da $x \in H_{\lambda_1}$ haben wir (für den Index s_1 von H_{λ_1})

$$0 = (\Phi - \lambda_1 \mathrm{id})^{s_1}(x) = (\lambda - \lambda_1)^{s_1} x,$$

also, da $x \neq 0$, $\lambda = \lambda_1$.

Sei jetzt noch $x \in B_{\lambda_1}$ ein Eigenvektor zum Eigenwert λ_1. Da $E_{\lambda_1} \subset H_{\lambda_1}$ folgt

$$x \in E_{\lambda_1} \cap B_{\lambda_1} \subset H_{\lambda_1} \cap B_{\lambda_1} = \{0\},$$

ein Widerspruch, da $x \neq 0$ als Eigenvektor. ■

Satz 17.8 (Hauptraumzerlegung) *Sei Φ ein Endomorphismus eines komplexen Vektorraums V mit charakteristischem Polynom*

$$p_\Phi = (-1)^n (X-\lambda_1)^{r_1} \cdots (X-\lambda_k)^{r_k}.$$

Dann ist V die direkte Summe der zugehörigen Haupträume von Φ:

$$V = H_{\lambda_1} \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_k}.$$

Außerdem gilt für $1 \leq i \leq k$, dass $\dim H_{\lambda_i} = r_i$.

Beweis: Wir beweisen den Satz mittels vollständiger Induktion über die Zahl $k \geq 1$ der verschiedenen Eigenwerte von Φ. Nach Hilfssatz 17.6 haben wir eine direkte Zerlegung

$$V = H_{\lambda_1} \oplus B_{\lambda_1}.$$

Nach Hilfsatz 17.2 und Hilfssatz 17.6 ist diese Zerlegung Φ-invariant. Ist $B_{\lambda_1} = \{0\}$, so sind wir fertig. Anderfalls ist die Restriktion $\Phi_1 := \Phi|_{B_{\lambda_1}}$ ein Endomorphismus von B_{λ_1}. Nach Hilfssatz 17.7 ist das charakteristische Polynom von Φ_1

$$p_{\Phi_1} = (-1)^{n-r_1} (X-\lambda_2)^{r_2} \cdots (X-\lambda_k)^{r_k}$$

und wir können die Induktionsannahme auf Φ_1 anwenden. Damit folgt die erste Behauptung.

Das charakteristische Polynom von $\Phi|_{H_{\lambda_1}}$ ist gleich $(-1)^{r_1}(X-\lambda_1)^{r_1}$ nach Hilfssatz 17.7. Also ist $\dim H_{\lambda_1} = r_1$ und mit Induktion folgt auch die zweite Behauptung des Satzes. ■

Bemerkung 17.9 Wählt man in jedem Hauptraum H_{λ_i} $(i = 1, \ldots, k)$ eine Basis $\{b_i^1, \ldots, b_i^{r_i}\}$ mit $r_i = \dim\, H_{\lambda_i}$, so ist die Vereinigung dieser k Teilbasen eine Basis von V. Bezüglich dieser geordneten Basis hat dann die Abbildungsmatrix von Φ die Gestalt

$$A = \begin{pmatrix} \boxed{A_{\lambda_1}} & & & 0 \\ & \boxed{A_{\lambda_2}} & & \\ & & \ddots & \\ 0 & & & \boxed{A_{\lambda_k}} \end{pmatrix}. \tag{17.1}$$

Die quadratischen $(r_i \times r_i)$-Matrizen A_{λ_i} längs der Hauptdiagonalen sind die Abbildungsmatrizen der Restriktionen $\Phi|_{H_{\lambda_i}}$ bezüglich der gewählten Teilbasen.

In der Matrix A sind die Anzahl k der Blöcke A_{λ_i} und die Anzahl r_i der Zeilen bzw. Spalten dieser Blöcke durch Φ eindeutig bestimmt (k = Anzahl der verschiedenen Eigenwerte, $r_i = \dim\, H_{\lambda_i}$); willkürlich wählbar sind noch die Teilbasen der H_{λ_i} und die Reihenfolge der Blöcke.

Beispiel 17.10 Wir führen das Beispiel 17.5 fort. Der dort betrachtete Endomorphismus von $\mathbb{C}^3$ hatte die Haupträume $H_1 = [b_1], H_4 = [b_2, b_3]$ mit

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 7 \\ 0 \\ 9 \end{pmatrix}.$$

Mit dieser Basis $\{b_1, b_2, b_3\}$ erhält man

$$\Phi(b_1) = b_1, \quad \Phi(b_2) = 4b_2, \quad \Phi(b_3) = 3b_2 + 4b_3$$

und damit die Blockmatrix

$$A = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 4 & 3 \\ 0 & 0 & 4 \end{array}\right); \quad k = 2, r_1 = 1, r_2 = 2.$$

17.3 Bestimmung der Jordanschen Normalform

Es sei wie bisher Φ ein Endomorphismus eines n-dimensionalen $\mathbb{C}$-Vektorraumes V. Wir zeigen jetzt, wie man durch geeignete Wahl von Teilbasen in den einzelnen Haupträumen H_{λ_i} eine besonders einfache Abbildungsmatrix findet, die sogenannte *Jordansche Normalform* von Φ.

Dazu greifen wir einen Eigenwert $\lambda_i = \lambda$ und den zugehörigen Hauptraum $H_{\lambda_i} = H_\lambda$ von Φ heraus. Nach Definition ist der Hauptraum $H_\lambda = K_q = \text{Kern } (\Phi - \lambda \text{ id})^q$. Die Einschränkung von $(\Phi - \lambda \text{ id})^q : V \to V$ auf H_λ ist also die Nullabbildung, d.h. die Abbildung

$$\Omega := (\Phi - \lambda \text{ id}) \mid_{H_\lambda} : H_\lambda \to H_\lambda$$

ist nilpotent.

Wir geben nun an, auf welche Gestalt sich die zugehörige Blockmatrix $A_\lambda := A_{\lambda_i}$ für eine geeignete Teilbasis bringen lässt. Wir werden zeigen, dass sich H_λ als direkte Summe von gewissen Untervektorräumen schreiben lässt:

$$H_\lambda = K_q = U_{q-1} \oplus U_{q-2} \oplus \cdots \oplus U_0.$$

Für jedes U_i konstruieren wir jeweils auch noch eine passende Basis.

1. SCHRITT: Konstruktion von U_{q-1} und einer Basis für U_{q-1}.

Nach Hilfssatz 17.3 ist $K_{q-1} \subsetneq K_q = H_\lambda$. Wir können also einen Untervektorraum $U_{q-1} \neq \{0\}$ finden, so dass

$$K_q = U_{q-1} \oplus K_{q-1} \tag{17.2}$$

gilt. Obwohl das Komplement U_{q-1} nicht eindeutig bestimmt ist, ist die Dimension $s_1 := \dim U_{q-1} \geq 1$ eindeutig. Wir wählen schließlich noch eine Basis von U_{q-1}:

$$\{b_1^{q-1}, \ldots, b_{s_1}^{q-1}\}.$$

Für b_i^{q-1} gilt also $\Omega^q(b_i^{q-1}) = 0$ und $\Omega^{q-1}(b_i^{q-1}) \neq 0$.

2. SCHRITT: Konstruktion von U_{q-2} und einer Basis für U_{q-2}.

Für den Bildraum $\overline{U} := \Omega(U_{q-1})$ gilt

- $\overline{U} \subset K_{q-1} = \operatorname{Kern} \Omega^{q-1}$:

  Für einen beliebigen Vektor $v \in U_{q-1}$ gilt wegen $v \in K_q$ auch $\Omega^{q-1}(\Omega(v)) = \Omega^q(v) = 0$, also $\Omega(v) \in K_{q-1}$.

- $\overline{U} \cap K_{q-2} = \{0\}$ für $q \geq 2$:

  Jeder von 0 verschiedene Vektor $v \in U_{q-1}$ liegt wegen (17.2) nicht in K_{q-1}, also ist

  $$0 \neq \Omega^{q-1}(v) = \Omega^{q-2}(\Omega(v)),$$

  d.h. jeder von 0 verschiedene Vektor aus $\overline{U}$ liegt nicht in K_{q-2}.

Die Summe $K_{q-2} + \overline{U}$ ist also direkt, und wegen $K_{q-2} \subset K_{q-1}$ haben wir $K_{q-2} \oplus \overline{U} \subset K_{q-1}$. Mit einem passend gewählten Komplement $\overline{\overline{U}}$ ergibt sich dann

$$K_{q-1} = K_{q-2} \oplus \overline{U} \oplus \overline{\overline{U}}.$$

Wir definieren $U_{q-2} := \overline{U} \oplus \overline{\overline{U}}$ und erhalten schließlich

$$K_{q-1} = U_{q-2} \oplus K_{q-2}. \tag{17.3}$$

Wir konstruieren jetzt eine Basis von U_{q-2}. Nach Konstruktion ist $\overline{U} = \Omega(U_{q-1}) \subset U_{q-2}$. In U_{q-1} haben wir die Basis $\{b_1^{q-1}, \ldots, b_{s_1}^{q-1}\}$ gewählt.

Wir zeigen jetzt, dass die s_1 Bildvektoren

$$\Omega(b_1^{q-1}), \ldots, \Omega(b_{s_1}^{q-1})$$

für $q \geq 2$ linear unabhängig sind und damit eine Basis von $\overline{U}$ bilden. Sei also

$$0 = \alpha_1 \Omega(b_1^{q-1}) + \cdots + \alpha_{s_1} \Omega(b_{s_1}^{q-1}) = \Omega(\sum_{j=1}^{s_1} \alpha_j b_j^{q-1}).$$

Wir setzen $v := \sum_{j=1}^{s_1} \alpha_j b_j^{q-1} \in U_{q-1}$. Es gilt dann wegen $\Omega(v) = 0$ auch $\Omega^{q-1}(v) = 0$, d.h. $v \in K_{q-1}$. Wegen (17.2) ist also $v = 0$ und damit $\alpha_1 = \cdots = \alpha_{s_1} = 0$.

Es ist $U_{q-2} = \overline{U} \oplus \overline{\overline{U}} = \Omega(U_{q-1}) \oplus \overline{\overline{U}}$. Für $\overline{U}$ haben wir gerade eine Basis gefunden, nämlich das Bild der Basis von U_{q-1}. Wir wählen noch eine Basis $\{b_1^{q-2}, \ldots, b_{s_2}^{q-2}\}$ von $\overline{\overline{U}}$ und haben insgesamt

$$\left\{ \begin{array}{l} \Omega(b_1^{q-1}), \ldots, \Omega(b_{s_1}^{q-1}), \\ b_1^{q-2}, \ldots, b_{s_2}^{q-2} \end{array} \right\} \quad \text{eine Basis von } U_{q-2}.$$

Insbesondere ist $\dim U_{q-2} = s_1 + s_2$; dabei setzen wir $s_2 = 0$, falls $\overline{\overline{U}} = \{0\}$ ist.

3. SCHRITT: Konstruktion von U_{q-3} und einer Basis für U_{q-3}.

Analog zu (17.2) behandeln wir jetzt die Zerlegung (17.3), also $K_{q-1} = U_{q-2} \oplus K_{q-2}$. Wieder haben wir einen Bildraum $\Omega(U_{q-2}) \subset K_{q-2}$, den wir zu einem Komplement U_{q-3} von K_{q-3} in K_{q-2} ergänzen. Die weitere Konstruktion ist dann völlig analog zum 2. Schritt. Wir erhalten dann eine Zerlegung

$$K_{q-2} = U_{q-3} \oplus K_{q-3} \tag{17.4}$$

und auch

$$\left\{ \begin{array}{l} \Omega^2(b_1^{q-1}), \ldots, \Omega^2(b_{s_1}^{q-1}), \\ \Omega(b_1^{q-2}), \ldots, \Omega(b_{s_2}^{q-2}), \\ b_1^{q-3}, \ldots, b_{s_3}^{q-3} \end{array} \right\} \quad \text{eine Basis von } U_{q-3}$$

mit $\dim U_{q-3} = s_1 + s_2 + s_3$.

4. SCHRITT: Sei $k \geq 3$.

- Aus dem vorhergehenden Schritt hat man jeweils eine Zerlegung $K_{q-k+1} = U_{q-k} \oplus K_{q-k}$.
- Das Bild von U_{q-k} unter Ω ist eine Teilmenge von K_{q-k} und wird zu einem Komplement U_{q-k-1} von K_{q-k-1} in K_{q-k} erweitert.
- Die Basis von U_{q-k} wird durch Ω auf eine linear unabhängige Teilmenge von U_{q-k-1} abgebildet und zu einer Basis von U_{q-k-1} ergänzt.

Im **letzten SCHRITT** haben wir

$$K_1 = U_0 \oplus K_0 = U_0 \oplus \{0\} = U_0, \tag{17.5}$$

und es ist

$$\left\{ \begin{array}{l} \Omega^{q-1}(b_1^{q-1}), \ldots, \Omega^{q-1}(b_{s_1}^{q-1}), \\ \Omega^{q-2}(b_1^{q-2}), \ldots, \Omega^{q-2}(b_{s_2}^{q-2}), \\ \quad \cdots \\ \Omega(b_1^1), \ldots, \Omega(b_{s_{q-1}}^1), \\ b_1^0, \ldots, b_{s_q}^0 \end{array} \right\} \quad \text{eine Basis von } U_0$$

mit $\dim U_0 = s_1 + s_2 + \cdots + s_q$.

Fazit: Aus (17.2), (17.3), (17.4) und (17.5) ergibt sich insgesamt die gesuchte direkte Zerlegung von H_λ

$$H_\lambda = K_q = U_{q-1} \oplus U_{q-2} \oplus \cdots \oplus U_0. \tag{17.6}$$

Dabei gilt für $x \in U_k \backslash \{0\}$ gerade $\Omega^k(x) \neq 0$ und $\Omega^{k+1}(x) = 0$. Durch Vereinigung der einzelnen Teilbasen erhält man eine Basis für H_λ, die aus

$$r_\lambda = s_1 + (s_1 + s_2) + \cdots + (s_1 + s_2 + \cdots + s_q) = qs_1 + (q-1)s_2 + \cdots + s_q$$

Vektoren besteht. Dabei ist

$$1 \leq r_\lambda, q, s_1 \leq n; \qquad s_2, s_3, \ldots, s_q \geq 0.$$

Wir ersetzten jetzt die Abkürzung Ω wieder durch $\Phi - \lambda\,\mathrm{id}$ und schreiben alle Basisvektoren in folgender Reihenfolge nochmals auf:

$$\left\{\begin{array}{ccccc}
(\Phi - \lambda\,\mathrm{id})^{q-1}(b_1^{q-1}), & \cdots & \cdots\,, & (\Phi - \lambda\,\mathrm{id})(b_1^{q-1}), & b_1^{q-1} \\
\vdots & & & \vdots & \\
(\Phi - \lambda\,\mathrm{id})^{q-1}(b_{s_1}^{q-1}), & \cdots & \cdots\,, & (\Phi - \lambda\,\mathrm{id})(b_{s_1}^{q-1}), & b_{s_1}^{q-1} \\
\hline
& (\Phi - \lambda\,\mathrm{id})^{q-2}(b_1^{q-2}), & \ldots, & (\Phi - \lambda\,\mathrm{id})(b_1^{q-2}), & b_1^{q-2} \\
& \vdots & & \vdots & \\
& (\Phi - \lambda\,\mathrm{id})^{q-2}(b_{s_2}^{q-2}), & \ldots, & (\Phi - \lambda\,\mathrm{id})(b_{s_2}^{q-2}), & b_{s_2}^{q-2} \\
\hline
& & \vdots & & \vdots \\
\hline
& & & & b_1^0, \\
& & & & \vdots \\
& & & & b_{s_q}^0.
\end{array}\right\}. \tag{17.7}$$

Die Abbildungsmatrix des Endomorphismus $(\Phi - \lambda\,\mathrm{id})\big|_{H_\lambda}$ bezüglich dieser geordneten Basis lautet dann:

$$\begin{pmatrix}
\boxed{\begin{matrix} \boxed{\begin{matrix}0&1&&0\\&\ddots&\ddots&\\&&\ddots&1\\0&&&0\end{matrix}} & & s_1\text{ q-Blöcke}\\ & \ddots & \\ & & \boxed{\begin{matrix}0&1&&0\\&\ddots&\ddots&\\&&\ddots&1\\0&&&0\end{matrix}}\end{matrix}} & & & 0\\
& \boxed{\begin{matrix} \boxed{\begin{matrix}0&1&&0\\&\ddots&\ddots&\\&&\ddots&1\\0&&&0\end{matrix}} & & s_2\text{ (q-1)-Blöcke}\\ & \ddots & \\ & & \boxed{\begin{matrix}0&1&&0\\&\ddots&\ddots&\\&&\ddots&1\\0&&&0\end{matrix}}\end{matrix}} & & \\
& & \ddots & \\
0 & & s_q\text{ 1-Blöcke} & \boxed{\begin{matrix}\boxed{0}&&\\&\ddots&\\&&\boxed{0}\end{matrix}}
\end{pmatrix}.$$

wobei hier s_1 Diagonalblöcke der Größe $q \times q$, s_2 Diagonalblöcke der Größe $(q-1) \times (q-1)$, $\ldots$, und s_q Diagonalblöcke der Größe 1×1 vorkommen.

Hieraus ergibt sich die gesuchte Abbildungsmatrix für $\Phi|_{H_\lambda}$ durch Addition von λE:

$$A_\lambda = \begin{pmatrix} \boxed{\begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{matrix}} & & & & & & \\ & \ddots & & & & & 0 \\ & & \boxed{\begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{matrix}} & & & & \\ & & & \ddots & & & \\ & & & & \boxed{\lambda} & & \\ & 0 & & & & \ddots & \\ & & & & & & \boxed{\lambda} \end{pmatrix}. \tag{17.8}$$

Definition 17.11 Die Matrix (17.8) heißt **Jordan-Block**[6] zum Eigenwert λ. In (17.8) treten s_1 q-reihige Kästchen, s_2 $(q-1)$-reihige Kästchen, ..., schließlich s_q einreihige Kästchen auf. Dabei sind der Index q sowie die Zahlen $s_1, \dots, s_q$ durch Φ eindeutig festgelegt.

Die Abbildungsmatrix von Φ, die man erhält, wenn man in der Blockmatrix (17.1) für jeden Eigenwert $\lambda_1, \dots, \lambda_k$ den entsprechenden Jordan-Block (17.8) einsetzt, heißt **Jordansche Normalform** von Φ. Die Basis (17.7) zusammen mit den sich für die anderen Eigenwerte ergebenden Basen heißt eine **Jordan-Basis** von V.

Zusammenfassend halten wir fest:

Satz 17.12 *Es sei Φ ein Endomorphismus eines Vektorraums V über $\mathbb{C}$ und $\lambda_1, \dots, \lambda_k$ seien die paarweise verschiedenen Eigenwerte von Φ. Dann gibt es eine Basis von V, so dass die zughörige Abbildungsmatrix die Jordansche Normalform von Φ ist. Die Jordansche Normalform von Φ ist bis auf die Reihenfolge der Blöcke eindeutig bestimmt.*

Folgerung 17.13 (Jordansche Normalform einer komplexen Matrix) *Jede komplexe $n \times n$-Matrix C ist ähnlich zu einer (bis auf die Reihenfolge der Jordan-Blöcke) eindeutig bestimmten Matrix der Gestalt* (17.1), (17.8).

[6] Camille Jordan (1838-1922)

Beweis: Sei $C \in \mathbb{C}^{n \times n}$ und $\Phi : \mathbb{C}^n \to \mathbb{C}^n;\ x \mapsto Cx$. Für die Abbildung Φ existiert dann nach Satz 17.12 eine Basis B bezüglich der die Φ eine Dastellungsmatrix A in Jordanscher Normalform hat. Beschreibt S die Matrix des Basiswechsels von der Standardbasis von $\mathbb{C}^n$ nach B, so ist $S^{-1}CS = A$. ■

Folgerung 17.13 besagt, dass es in der zu $C \in \mathbb{C}^{n \times n}$ gehörigen Äquivalenzklasse $[C]$ bezüglich der Äquivalenzrelation "ähnlich" (vgl. Abschnitt 10.3) einen besonders einfachen Repräsentanten gibt: die Jordansche Normalform.

Bemerkung 17.14 Die Jordansche Normalform $J \in \mathbb{C}^{n \times n}$ eines Endomorphismus bzw. einer Matrix ist gegeben durch die Matrizen (17.1) und (17.8). Etwas ungenauer kann man das auch wie folgt formulieren. Es gibt eine eine Diagonalmatrix $D \in \mathbb{C}^{n \times n}$ und eine nilpotente Matrix $N \in \mathbb{C}^{n \times n}$, so dass gilt

$$J = D + N, \quad \text{(und auch noch } \ DN = ND).$$

17.3.1 Ein Beispiel

Es sei Φ ein Endomorphismus von $\mathbb{C}^6$, der bezüglich der kanonischen Basis von $\mathbb{C}^6$ durch die Abbildungsmatrix

$$B = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 9 & 6 & 0 & 0 & 0 & 0 \\ -5 & -2 & -96 & -88 & -77 & 0 \\ 10 & 4 & 135 & 123 & 105 & 0 \\ -5 & -2 & -27 & -24 & -18 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

gegeben ist. Gesucht ist die Jordansche Normalform von Φ bzw. B.

1. Wir bestimmen zunächst die Eigenwerte von Φ. Dazu berechnen wir das charakteristische Polynom $p = \det(B - X \cdot E)$.

$$
\begin{aligned}
p \;&=\; \begin{vmatrix} -X & -1 & 0 & 0 & 0 & 0 \\ 9 & 6-X & 0 & 0 & 0 & 0 \\ -5 & -2 & -96-X & -88 & -77 & 0 \\ 10 & 4 & 135 & 123-X & 105 & 0 \\ -5 & -2 & -27 & -24 & -18-X & 0 \\ 0 & 0 & 0 & 0 & 0 & 3-X \end{vmatrix} \\
&=\; (3-X)\cdot \begin{vmatrix} -X & -1 & 0 & 0 & 0 \\ 9 & 6-X & 0 & 0 & 0 \\ 0 & 0 & -69-X & -64 & -59+X \\ 0 & 0 & 81 & 75-X & 69-2X \\ -5 & -2 & -27 & -24 & -18-X \end{vmatrix} \\
&=\; (3-X)\cdot \begin{vmatrix} -X & -1 \\ 9 & 6-X \end{vmatrix} \cdot \begin{vmatrix} -69-X & -64 & -59+X \\ 81 & 75-X & 69-2X \\ -27 & -24 & -18-X \end{vmatrix} \\
&=\; (3-X)\cdot (X^2-6X+9)\cdot \begin{vmatrix} -5-X & -64 & 5+X \\ 6+X & 75-X & -6-X \\ -3 & -24 & 6-X \end{vmatrix} \\
&=\; (3-X)^3\cdot \begin{vmatrix} -5-X & -64 & 0 \\ 6+X & 75-X & 0 \\ -3 & -24 & 3-X \end{vmatrix} \\
&=\; (3-X)^4\cdot \begin{vmatrix} -5-X & -64 \\ 6+X & 75-X \end{vmatrix} \\
&=\; (3-X)^4\cdot (X^2-70X-375+64X+384) = (3-X)^6
\end{aligned}
$$

Damit gilt: λ ist Eigenwert von $\Phi \iff p(\lambda)=0 \iff \lambda = 3$

3 ist also der einzige Eigenwert von Φ. Nach Satz 17.8 gibt es demnach auch nur einen Hauptraum H_λ, der dann mit $\mathbb{C}^6$ übereinstimmt, und in der Jordanschen Blockmatrix (17.1) tritt nur ein Jordan-Block A_3 der Gestalt (17.8) auf.

2. Um A_3 zu bestimmen, berechenen wir zunächst den Index q des Hauptraumes H_3.

Wir bilden

$$B - 3E = \begin{pmatrix} -3 & -1 & 0 & 0 & 0 & 0 \\ 9 & 3 & 0 & 0 & 0 & 0 \\ -5 & -2 & -99 & -88 & -77 & 0 \\ 10 & 4 & 135 & 120 & 105 & 0 \\ -5 & -2 & -27 & -24 & -21 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$(B - 3E)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & -1 & 0 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 & 0 & 0 \\ -3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$(B - 3E)^3 = (B - 3E)^4 = \cdots = O.$$

Die Folge der K_k wird also ab $k = 3$ konstant, der Index von H_3 ist somit $q = 3$.

3. Jetzt berechnen wir die Untervektorräume $K_k =$ Kern $(\Phi - 3\,\mathrm{id})^k$, $k = 1, 2, 3$.

   Wegen $(B - 3E)^3 = O$ ist $K_3 = H_3 = \mathbb{C}^6$.

   Aus $(B - 3E)^2$ liest man ab:

$$K_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \in \mathbb{C}^6 \mid 3x_1 + x_2 = 0 \right\},$$

also

$$K_2 = \left[\begin{pmatrix} -1 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right].$$

$K_1 = E_3$ ist der Eigenraum zum Eigenwert $\lambda = 3$ von Φ. Mit Hilfe von elementaren Zeilenumformungen wird $B - 3E$ auf Stufenform gebracht, um den Kern zu bestimmen.

$$\begin{pmatrix} -3 & -1 & 0 & 0 & 0 & 0 \\ 9 & 3 & 0 & 0 & 0 & 0 \\ -5 & -2 & -99 & -88 & -77 & 0 \\ 10 & 4 & 135 & 120 & 105 & 0 \\ -5 & -2 & -27 & -24 & -21 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} 3 \\ \\ 2 \quad -1 \\ \\ \\ \\ \end{array} \sim$$

$$\begin{pmatrix} -3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -5 & -2 & -99 & -88 & -77 & 0 \\ 0 & 0 & -63 & -56 & -49 & 0 \\ 0 & 0 & 72 & 64 & 56 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} (-\frac{5}{3}) \\ \\ \\ -\frac{1}{7} \\ \frac{1}{8} \\ \\ \end{array} \sim$$

$$\begin{pmatrix} -3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & -99 & -88 & -77 & 0 \\ 0 & 0 & 9 & 8 & 7 & 0 \\ 0 & 0 & 9 & 8 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \\ (-1) \quad 11 \\ \\ \end{array} \sim$$

$$\begin{pmatrix} -3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 8 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} \\ \\ (-3) \quad \cdot(-3) \\ \\ \\ \\ \end{array} \sim$$

$$\begin{pmatrix} -3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 8 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} \cdot(-\frac{1}{3}) \\ \\ \\ \\ \\ \\ \end{array} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 8 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Hieran liest man ab:

$$K_1 = \left[\begin{pmatrix} 0 \\ 0 \\ 8 \\ -9 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \\ 0 \\ -9 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right].$$

4. Bei der Konstruktion einer Jordan-Basis gehen wir wie zu Beginn des Abschnitts 17.3 beschrieben vor. Ausgangspunkt ist die Kette verschachtelter Untervektorräume
$$K_1 \subset K_2 \subset K_3 = H_3.$$
Der erste Schritt erfordert die Zerlegung des Hauptraums
$$H_3 = K_3 = U_2 \oplus K_2.$$
Wegen $K_3 = \mathbb{C}^6$ und $\dim K_2 = 5$ hat U_2 die Dimension 1. Es ist also als Basis ein Vektor aus $\mathbb{C}^6$ zu bestimmen, der nicht in K_2 liegt. Wir wählen
$$U_2 := \left[\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}\right] \text{ mit dem Basisvektor } b_2^1 := \begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}.$$
Im zweiten Schritt wird eine Zerlegung von K_2 gesucht: $K_2 = U_1 \oplus K_1$. Dabei ist bekannt, dass $(\Phi - 3\mathrm{id})(b_2^1) \in K_2 \backslash K_1$ gilt. Weil K_2 die Dimension 5 und K_1 die Dimension 3 hat, muss U_1 zweidimensional sein. Als Basis von U_1 benutzen wir - wie in der Konstruktion verlangt - $(\Phi - 3\mathrm{id})(b_2^1)$ und wählen einen (dazu linear unabhängigen) Vektor aus $K_2 \backslash K_1$. Wir erhalten
$$U_1 := \left[\begin{pmatrix}-1\\3\\-2\\4\\-2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix}\right] \text{ Basisvektoren } (\Phi - 3\mathrm{id})(b_2^1) = \begin{pmatrix}-1\\3\\-2\\4\\-2\\0\end{pmatrix} \text{ und } b_1^1 := \begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix}.$$
Im dritten und letzten Schritt ist eine Basis von K_1 zu bestimmen, wobei $(\Phi - 3\mathrm{id})(U_1) \subset K_1$ gilt. Zwei Basisvektoren sind nach Konstruktion durch $(\Phi - 3\mathrm{id})^2(b_2^1)$ und $(\Phi - 3\mathrm{id})(b_1^1)$ bereits vorgegeben. Im dreidimensionalen Untervektorraum K_1 ist somit noch ein weiterer Basisvektor b_0^1 wählbar.
$$(\Phi - 3\mathrm{id})^2(b_2^1) = \begin{pmatrix}-0\\0\\-1\\2\\-1\\0\end{pmatrix}, (\Phi - 3\mathrm{id})(b_1^1) = \begin{pmatrix}0\\0\\-77\\105\\-21\\0\end{pmatrix} \text{ und } b_0^1 := \begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}.$$

Für die Jordanbasis wählen wir die Reihenfolge dieser Vektoren wie folgt:

$$\begin{aligned} b_1 &:= (\Phi - 3\mathrm{id})^2(b_2^1), & b_2 &:= (\Phi - 3\mathrm{id})(b_2^1), & b_3 &:= b_2^1 \\ & & b_4 &:= (\Phi - 3\mathrm{id})(b_1^1), & b_5 &:= b_1^1 \\ & & & & b_6 &:= b_0^1. \end{aligned}$$

Damit gilt

$$\begin{aligned} (\Phi - 3\mathrm{id})(b_1) &= (\Phi - 3\mathrm{id})^3(b_2^1) = 0 & &\text{d.h. } \Phi(b_1) = 3 \cdot b_1 \\ (\Phi - 3\mathrm{id})(b_2) &= (\Phi - 3\mathrm{id})^2(b_2^1) = b_1 & &\text{d.h. } \Phi(b_2) = b_1 + 3 \cdot b_2 \\ (\Phi - 3\mathrm{id})(b_3) &= (\Phi - 3\mathrm{id})(b_2^1) = b_2 & &\text{d.h. } \Phi(b_3) = b_2 + 3 \cdot b_3 \\ (\Phi - 3\mathrm{id})(b_4) &= (\Phi - 3\mathrm{id})^2(b_1^1) = 0 & &\text{d.h. } \Phi(b_4) = 3 \cdot b_4 \\ (\Phi - 3\mathrm{id})(b_5) &= (\Phi - 3\mathrm{id})(b_1^1) = b_4 & &\text{d.h. } \Phi(b_5) = b_4 + 3 \cdot b_5 \\ (\Phi - 3\mathrm{id})(b_6) &= (\Phi - 3\mathrm{id})(b_1^1) = 0 & &\text{d.h. } \Phi(b_6) = 3 \cdot b_6 \end{aligned}$$

Zusammenfassend haben wir die Abbildungsmatrix von Φ bzgl. der Basis $\{b_1, \ldots, b_6\}$ in Jordanscher Normalform:

$$A = \begin{pmatrix} \boxed{\begin{matrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{matrix}} & & 0 \\ & \boxed{\begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix}} & \\ 0 & & \boxed{3} \end{pmatrix}.$$

Bemerkung: Wählt man eine Basis $C := \{c_1, \ldots, c_6\}$ durch

$$c_1 := b_3\,, c_2 := b_2\,, c_3 := b_1\,, c_4 := b_5\, c_5 := b_4\,, c_6 := b_6\,,$$

so erhält man $A^\top$ als zughörige Abbildungsmatrix. Hier stehen die Einsen unterhalb der Diagonale. Auch diese Form wird in der Literatur als Jordansche Normalform bezeichnet.

Bemerkung: Ist die Berechnung einer Jordan-Basis nicht notwendig, so kann die Jordansche Normalform auch über die Dimensionen der Kerne K_k bestimmt werden. Mit den Bezeichnungen zu Beginn des Abschnitts 17.3 gilt:

	$q = 3$	Größte Kästchenlänge ist 3
$s_1 = \dim U_2 = \dim K_3 - \dim K_2 = 6 - 5 = 1$	$s_1 = 1$	1 Kästchen der Länge 3
$s_1 + s_2 = \dim U_1 = \dim K_2 - \dim K_1 = 2$	$s_2 = 1$	1 Kästchen der Länge 2
$s_1 + s_2 + s_3 = \dim U_0 = \dim K_1 = 3$	$s_3 = 1$	1 Kästchen der Länge 1

Diese Angaben legen die Jordansche Normalform fest.

17.4 Weitere Eigenschaften der Jordanschen Normalform

Im letzten Beispiel haben wir gesehen, dass zum bloßen Aufstellen der Jordanschen Normalform die Berechnung einer Jordan-Basis nicht nötig ist, sondern dass die Kenntnis der Eigenwerte, der Indizes der Haupträume und der Dimensionen s_j ausreicht.

Wir wollen jetzt noch zwei (technische) Eigenschaften ableiten, die beim Aufstellen der Jordanschen Normalform von Φ hilfreich sein können.

Wir setzen voraus, dass sämtliche Eigenwerte von Φ bekannt sind.

Satz 17.15 *Die Anzahl der Kästchen in einem zum Eigenwert λ_j gehörigen Jordanblock ist die Dimension des Eigenraums $K_1^j =$ Kern $(\Phi - \lambda_j \operatorname{id}_V)$.*

Beweis: Die Gesamtzahl der Kästchen zum Eigenwert λ_j ist nach (17.8) gegeben durch $s_1 + s_2 + \cdots + s_q = \dim K_1^j$. ■

Die Anzahl der Kästchen zum Eigenwert λ_j stimmt also überein mit der Anzahl derjenigen Spalten in der Jordanschen Normalform, in denen außer 0 nur λ_j steht.

Weiter lässt sich die Anzahl σ_h^j der $(h \times h)$-Kästchen in einem Jordan-Block zum Eigenwert λ_j direkt ausrechnen unter Verwendung der Ränge

$$\tau_h^j = \text{ Rang } (\Phi - \lambda_j \operatorname{id}_V)^h, \qquad h = 0, 1, \ldots .$$

Es gilt nämlich der

Satz 17.16 *Die Anzahl σ_h^j der $(h \times h)$-Kästchen in einem zum Eigenwert λ_j gehörigen Jordan-Block ist gegeben durch*

$$\sigma_h^j = \tau_{h-1}^j - 2\tau_h^j + \tau_{h+1}^j, \tag{17.9}$$

wobei $j = 1, 2, \ldots, k$; $h = 1, 2, \ldots, q_j$.

Beweis: Wir wählen einen Eigenwert λ (der Index j wird unterdrückt) und erinnern an die Jordan-Teilbasis (17.7) von H_λ. Mit den dortigen Bezeichnungsweisen haben wir:

$$\begin{array}{rrcll} \text{(i)} & K_{q-i} & = & U_{q-i-1} \oplus K_{q-i-1}; & i = 0, \ldots, q-1 \\ \text{(ii)} & \dim U_{q-j} & = & s_1 + s_2 + \cdots + s_j; & j = 1, \ldots, q \\ \text{(iii)} & \sigma_h & = & s_{q-h+1}; & h = 1, \ldots, q. \end{array}$$

Also gilt

$$\begin{aligned} \sigma_h &\overset{\text{(iii)}}{=} s_1 + s_2 + \cdots + s_{q-h+1} - (s_1 + s_2 + \cdots + s_{q-h}) \overset{\text{(ii)}}{=} \dim U_{h-1} - \dim U_h \\ &\overset{\text{(i)}}{=} \dim K_h - \dim K_{h-1} - (\dim K_{h+1} - \dim K_h) \\ &= -\dim K_{h-1} + 2 \dim K_h - \dim K_{h+1} \\ &= -(n - \tau_{h-1}) + 2(n - \tau_h) - (n - \tau_{h+1}) = \tau_{h-1} - 2\tau_h + \tau_{h+1}. \end{aligned}$$

■

Beispiel 17.17 In $\mathbb{C}^5$ sei ein Endomorphismus gegeben durch die Matrix

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

Man rechnet nach: Der einzige Eigenwert von B ist $\lambda = 2$. Wir können daher den oberen Index j in (17.9) weglassen. Wir berechnen die Ränge τ_h und daraus die Zahlen σ_h:

Man findet

$$\tau_1 = \text{Rang}\,(B-2E) = \text{Rang} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 2 \\ 2 & 1 & 0 & 0 & 2 \\ 1 & 2 & 0 & 0 & 0 \end{pmatrix} = 3,$$

$$\tau_2 = \text{Rang}\,(B-2E)^2 = \text{Rang} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 2 & 6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} = 2,$$

$$\tau_3 = \text{Rang}\,(B-2E)^3 = \text{Rang} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = 1,$$

$\tau_h = \text{Rang}\,(B-2E)^h = \text{Rang}\,O = 0$ für $h \geq 4$. Damit ergeben sich insgesamt $5 - \tau_1 = 2$ Jordan-Kästchen (Satz 17.15) und $\sigma_1 = 1, \sigma_2 = 0, \sigma_3 = 0, \sigma_4 = 1, \sigma_5 = 0$

für $h \geq 5$ (Satz 17.16). Es gibt also ein 1-reihiges und ein 4-reihiges Jordan-Kästchen zum Eigenwert 2. Wir erhalten somit

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

als Jordansche Normalform von B.

Teil VI

Vektorräume mit Skalarprodukt

18 Euklidische und unitäre Vektorräume

Ziel dieses Kapitels ist es, in reellen und komplexen Vektorräumen geometrische Konzepte wie "Länge", "Winkel" oder "senkrecht" zu definieren. Dazu benötigt man als Zusatzstruktur ein sogenanntes Skalarprodukt.

18.1 Skalarprodukte

Die Vektorräume, die wir mit der Zusatzstruktur eines Skalarprodukts versehen werden, sind hier stets reelle oder komplexe Vektorräume. Bei der Definition des Skalarprodukts behandeln wir den reellen bzw. den komplexen Fall zunächst getrennt. Später ist es dann oft zweckmäßig, beide Fälle simultan zu betrachten.

18.1.1 Euklidische Vektorräume

Definition 18.1 Es sei V ein *reeller* Vektorraum. Eine **Bilinearform** auf V ist eine Abbildung

$$F : V \times V \to \mathbb{R}, \qquad (a, b) \mapsto F(a, b),$$

die in jedem Argument linear ist; d.h. für alle $a, a_1, a_2, b, b_1, b_2 \in V$ und alle $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ gilt:

$$\begin{aligned} F(\lambda_1 a_1 + \lambda_2 a_2, b) &= \lambda_1 F(a_1, b) + \lambda_2 F(a_2, b), \\ F(a, \mu_1 b_1 + \mu_2 b_2) &= \mu_1 F(a, b_1) + \mu_2 F(a, b_2). \end{aligned}$$

Eine Bilinearform F heißt **symmetrisch**, wenn gilt

$$\forall\, a, b \in V : \quad F(a, b) = F(b, a).$$

Eine Bilinearform F auf V heißt **positiv definit**, wenn gilt:

$$: Skalarprodukte(Forts.) \forall\, a \in V : F(a, a) \geq 0 \ \text{ und } F(a, a) = 0 \Leftrightarrow a = 0.$$

Ein **Skalarprodukt** auf (oder in) einem reellen Vektorraum V ist eine positiv definite, symmetrische Bilinearform auf V. Ein Paar (V, F) bestehend aus einem reellen Vektorraum und einem Skalarprodukt F auf V heißt **euklidischer Vektorraum.**

Beispiel 18.2

1. Das **Standard-Skalarprodukt** in $\mathbb{R}^n$. Im Vektorraum $\mathbb{R}^n$ ist

$$F:\ \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}; \qquad ((\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)) \mapsto \sum_{j=1}^{n} \alpha_j \beta_j$$

ein Skalarprodukt. Insbesondere gilt ja:

$$\sum_{k=1}^{n} \alpha_k^2 \geq 0 \quad \text{für alle} \quad (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n \ \text{ und}$$

$$\sum_{k=1}^{n} \alpha_k^2 = 0 \Longleftrightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

2. In $\mathbb{R}^3$ ist durch

$$F((\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3)) = \alpha_1\beta_1 + 2\alpha_2\beta_2 - \alpha_2\beta_3 - \alpha_3\beta_2 + \alpha_3\beta_3$$

ein Skalarprodukt definiert. Dass F positiv definit ist, folgt aus

$$F(a, a) = \alpha_1^2 + 2\alpha_2^2 - 2\alpha_2\alpha_3 + \alpha_3^2 = \alpha_1^2 + \alpha_2^2 + (\alpha_2 - \alpha_3)^2.$$

3. Die Abbildung

$$F:\ \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}; \qquad ((\alpha_1, \alpha_2), (\beta_1, \beta_2)) \mapsto \alpha_1\beta_2 + \alpha_2\beta_1$$

ist eine symmetrische Bilinearform. F ist aber nicht positiv definit, denn es ist $F((0,1),(0,1)) = 0$. D.h. F ist *kein* Skalarprodukt auf $\mathbb{R}^2$.

4. Es sei $C(I)$ der reelle, unendlichdimensionale Vektorraum aller im Intervall $I = [a, b] \subset \mathbb{R}$ stetigen reellen Funktionen. Dann ist

$$F:\ C(I) \times C(I) \to \mathbb{R}; \qquad (g, h) \mapsto \int_a^b g(t)h(t)\, dt$$

ein Skalarprodukt auf $C(I)$. Denn nach den Grundregeln der Integralrechnung ist F bilinear und symmetrisch. F ist auch positiv definit, denn für alle $g \in C(I)$ ist $F(g, g) \geq 0$ und aus

$$F(g, g) = \int_a^b g^2(t)\, dt = 0$$

folgt $g(t) = 0$ für alle $t \in I$. Wäre nämlich $g(t_0) \neq 0$ für ein $t_0 \in I$, so gäbe es wegen der Stetigkeit von g ein Teilintervall $I' \subset I$ mit $t_0 \in I'$, so dass $g(t) \neq 0$ für alle $t \in I'$, und es wäre, entgegen der Voraussetzung,

$$F(g,g) \geq \int_{I'} g^2(t)\,dt > 0.$$

18.1.2 Unitäre Vektorräume

Versucht man die Begriffe aus Abschnitt 18.1.1 auf *komplexe* Vektorräume zu übertragen, so gibt es eine Schwierigkeit: Bei der Definition von "positiv definit" haben wir benutzt, dass $\mathbb{R}$ ein geordneter Körper ist. Das gilt für $\mathbb{C}$ nicht. Ein Ausweg besteht darin, statt der "Symmetrie" die sogenannte "Hermite-Eigenschaft"[7] zu verlangen. Diese basiert auf folgender Tatsache: Für eine komplexe Zahl $z = x + \mathrm{i}y \in \mathbb{C}$ und ihre komplex **Konjugierte** $\overline{z} = x - \mathrm{i}y$ ist

$$z^2 = zz = (x + \mathrm{i}y)(x + \mathrm{i}y) = (x^2 - y^2) + 2\mathrm{i}xy \in \mathbb{C} \quad \text{hingegen}$$

$$z\overline{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + y^2 \in \mathbb{R}.$$

Definition 18.3 Es sei V ein *komplexer* Vektorraum. Eine Abbildung

$$F: \; V \times V \to \mathbb{C}\,; \qquad (a,b) \mapsto F(a,b)$$

heißt **hermitesche Form** auf V, wenn für alle $a, a_1, a_2, b \in V$ und alle $\lambda_1, \lambda_2 \in \mathbb{C}$ gilt:

$$\begin{aligned} F(\lambda_1 a_1 + \lambda_2 a_2, b) &= \lambda_1 F(a_1, b) + \lambda_2 F(a_2, b) \\ F(b,a) &= \overline{F(a,b)}. \qquad (\textbf{Hermite}\text{-Eigenschaft}) \end{aligned}$$

Hilfssatz 18.4 (Elementare Eigenschaften einer hermiteschen Form) *Sei F eine hermitesche Form auf einem komplexen Vektorraum V. Für alle $a, b_1, b_2 \in V$ und alle $\mu_1, \mu_2 \in \mathbb{C}$ gilt:*

1. $F(a, \mu_1 b_1 + \mu_2 b_2) = \overline{\mu_1}\, F(a, b_1) + \overline{\mu_2}\, F(a, b_2)$,
2. $F(a,a) \in \mathbb{R}, \quad F(0,0) = 0.$

Beweis: Aus Definition 18.3 ergibt sich

$$\begin{aligned} F(a, \mu_1 b_1 + \mu_2 b_2) &= \overline{F(\mu_1 b_1 + \mu_2 b_2, a)} = \overline{\mu_1 F(b_1, a) + \mu_2 F(b_2, a)} \\ &= \overline{\mu_1}\,\overline{F(b_1, a)} + \overline{\mu_2}\,\overline{F(b_2, a)} = \overline{\mu_1}\, F(a, b_1) + \overline{\mu_2}\, F(a, b_2). \end{aligned}$$

[7] Charles Hermite (1822-1901)

Weiter gilt mit der Hermite-Eigenschaft: $F(a,a) = \overline{F(a,a)}$, d.h. $F(a,a)$ ist reell. Dass $F(0,0) = 0$ ist folgt wie im reellen Fall aus der Linearität im ersten Argument. ■

Definition 18.5 Eine hermitesche Form F in einem komplexen Vektorraum V heißt **positiv definit**, wenn gilt:

$$\forall a \in V: \ F(a,a) \geq 0 \ \text{ und } \ F(a,a) = 0 \Leftrightarrow a = 0$$

Ein **Skalarprodukt** in einem komplexen Vektorraum V ist eine positiv definite, hermitesche Form F. Ein Paar (V,F) bestehend aus einem komplexen Vektorraum und einem Skalarprodukt F heißt **unitärer Vektorraum**.

Beispiel 18.6 Das **Standard-Skalarprodukt** in $\mathbb{C}^n$. Die Abbildung

$$F: \ \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}; \qquad ((z_1, \ldots, z_n), (w_1, \ldots, w_n)) \mapsto \sum_{k=1}^{n} z_k \overline{w}_k$$

ist ein Skalarprodukt im Vektorraum $\mathbb{C}^n$. Insbesondere ist F wegen

$$F((z_1, \ldots, z_n), (z_1, \ldots, z_n)) = \sum_{k=1}^{n} z_k \overline{z}_k = \sum_{k=1}^{n} |z_k|^2 > 0$$

für $(z_1, \ldots, z_n) \neq 0$ positiv definit. Der Standard-Vektorraum $\mathbb{C}^n$ versehen mit dem Skalarprodukt F ist also ein n-dimensionaler unitärer Vektorraum.

NOTATION/SPRECHWEISE. Für ein Skalarprodukte F in einem euklidischen oder unitären Vektorraum V schreiben wir

$$\boxed{\langle a,b \rangle := F(a,b) \quad (a,b \in V).}$$

Viele Aussagen gelten sowohl für euklidische als auch für unitäre Vektorraume. In solchen Fällen ist es deshalb zweckmäßig, einfach von **Vektorräumen mit Skalarprodukt** zu sprechen.

18.2 Skalarprodukte und Matrizen

Lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen kann man (via Wahl von Basen) durch Matrizen vollständig beschreiben. Das gilt auch für Skalarprodukte.

18.2.1 Darstellungsmatrizen

Definition 18.7 Sei V ein n–dimensionaler reeller bzw. komplexer Vektorraum und $\langle , \rangle$ ein Skalarprodukt auf V. Wir wählen eine geordnete Basis $B = \{b_1, \ldots, b_n\}$ von V und setzen

$$g_{kj} := \langle b_k, b_j \rangle \qquad 1 \leq k, j \leq n.$$

Die $n \times n$–Matrix $G := (g_{kj}) \in \mathbb{R}^{n\times n}$ bzw. $\mathbb{C}^{n\times n}$ heißt **Matrix des Skalarproduktes** $\langle , \rangle$ bezüglich der Basis B.

Da $\langle b_j, b_k \rangle = \langle b_k, b_j \rangle$ bzw. $\langle b_j, b_k \rangle = \overline{\langle b_k, b_j \rangle}$ gilt $g_{jk} = g_{kj}$ bzw. $g_{jk} = \overline{g_{kj}}$. Für die Matrix eines Skalarprodukts gilt also

$$\boxed{G = G^\top} \text{ falls } V \text{ euklidisch und } \boxed{G = \overline{G}^\top} \text{ falls } V \text{ unitär ist.}$$

Sei jetzt V ein *unitärer Vektorraum* und $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Dann haben wir für $a = \sum\limits_{k=1}^{n} \alpha_k b_k$ und $b = \sum\limits_{j=1}^{n} \beta_j b_j$

$$\begin{aligned}
\langle a, b \rangle &= \sum_{k=1}^{n}\sum_{j=1}^{n} \alpha_k \langle b_k, b_j \rangle \overline{\beta_j} = \sum_{k=1}^{n}\sum_{j=1}^{n} \alpha_k \, g_{kj} \, \overline{\beta_j} \\
&= (\alpha_1, \ldots, \alpha_n) \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} \overline{\beta_1} \\ \vdots \\ \overline{\beta_n} \end{pmatrix}.
\end{aligned}$$

Mit den Komponentenvektoren

$$\Theta_B(a) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \ \Theta_B(b) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

von a und b (vgl. Abschnitt 7.3) können wir also schreiben

$$\langle a, b \rangle = \Theta_B(a)^\top G \, \overline{\Theta_B(b)}; \qquad a, b \in V, \ \ \Theta_C(a), \ \Theta_C(b) \in \mathbb{C}^n.$$

Ganz analog erhalten wir für einen *euklidischen Vektorraum* die Formel

$$\langle a, b \rangle = \Theta_B(a)^\top G \, \Theta_B(b); \qquad a, b \in V, \ \ \Theta_B(a), \Theta_B(b) \in \mathbb{R}^n.$$

Da Skalarprodukte positiv definit sind, ergibt sich als Konsequenz dieser Formel für die Matrix G von $\langle , \rangle$:

Im euklidischen Fall gilt für alle $x \in \mathbb{R}^n, x \neq 0$, $x^\top G x > 0$, und im unitären Fall gilt für alle $z \in \mathbb{C}^n, z \neq 0$, $z^\top G \overline{z} > 0$.

Definition 18.8 Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt **symmetrisch**, falls gilt

$$A^\top = A.$$

Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt **hermitesch**, falls gilt

$$\overline{A}^\top = A.$$

Eine symmetrische (bzw. hermitesche) Matrix A für die gilt $x^\top G x > 0$ für alle $x \in \mathbb{R}^n, x \neq 0$ bzw. $z^\top G \overline{z} > 0$ für alle $z \in \mathbb{C}^n, z \neq 0$, heißt **positiv definit**.

Zusammenfassend haben wir

Satz 18.9 (Beschreibung aller Skalarprodukte) *Es sei V ein euklidischer bzw. unitärer Vektorraum und B eine geordnete Basis von V. Dann ist F genau dann ein Skalarprodukt, wenn eine positiv definite, symmetrische (bzw. hermitesche) Matrix $A \in \mathbb{R}^{n\times n}$ bzw. $A \in \mathbb{C}^{n\times n}$ existiert mit*

$$F(x,y) = \Theta_B(x)^\top A\, \Theta_B(y) \ \text{ bzw. } \ F(x,y) = \Theta_B(x)^\top A\, \overline{\Theta_B(y)} \qquad (x, y \in V). \qquad (*)$$

Beweis: Ist F ein Skalarprodukt, so hat die Matrix von F bezüglich B die behaupteten Eigenschaften. Ist umgekehrt eine positiv definite, symmetrische bzw. hermitesche Matrix $A \in \mathbb{R}^{n\times n}$ bzw. $A \in \mathbb{C}^{n\times n}$ gegeben, so wird durch $(*)$ ein Skalarprodukt auf V *definiert.* ■

Die Frage, wie man nachprüfen kann, ob eine symmetrische Matrix positiv definit ist, werden wir später beantworten. Wir werden zwei dafür Kriterien angeben: Siehe Satz 19.13 und Satz 20.22.

18.2.2 Basiswechsel

Wir wollen nun noch überlegen, wie sich die Matrix G eines Skalarproduktes F ändert, wenn man die Basis $B = \{b_1, \ldots, b_n\}$ durch eine andere Basis $\widetilde{B} = \{\tilde{b}_1, \ldots, \tilde{b}_n\}$ ersetzt. Sei dazu $S = (s_{ij}) \in \mathbb{R}^{n\times n}$ bzw. $S \in \mathbb{C}^{n\times n}$ die (reguläre) Matrix des Basiswechsels von $\widetilde{B}$ nach B, also

$$\begin{array}{rcl} \tilde{b}_1 & = & s_{11}b_1 + s_{21}b_2 + \cdots + s_{n1}b_n \\ \vdots & & \vdots \\ \tilde{b}_n & = & s_{1n}b_1 + s_{2n}b_2 + \cdots + s_{nn}b_n \end{array}.$$

Dann gilt für die Komponenten $\tilde{g}_{jk}$ der Darstellungsmatrix $\widetilde{G}$ von F bzgl. $\widetilde{B}$

$$\tilde{g}_{jk} = F(\tilde{b}_j, \tilde{b}_k) = F\left(\sum_{p=1}^n s_{pj}b_p, \sum_{q=1}^n s_{qk}b_q\right) = \sum_{p=1}^n \sum_{q=1}^n s_{pj}\, g_{pq}\, \overline{s}_{qk}; \qquad 1 \leq j, k \leq n.$$

Damit haben wir gezeigt

Satz 18.10 *Sei V ein euklidischer bzw. unitärer Vektorraumes mit Basen B und $\widetilde{B}$. Weiter sei S die Matrix des Basiswechsels von $\widetilde{B}$ nach B. Dann gelten für die Matrizen des Skalarproduktes die Transformationsformeln*

$$\begin{aligned} \widetilde{G} &= S^\top G\, S, \qquad S \in \mathbf{GL}(n, \mathbb{R}) \quad \text{(euklidischer Fall)},\\ \widetilde{G} &= S^\top G\, \overline{S}, \qquad S \in \mathbf{GL}(n, \mathbb{C}) \quad \text{(unitärer Fall)}. \end{aligned}$$

Bemerkung 18.11 (Erinnerung) Für die Darstellungsmatrix A eines linearen Endomorphismus Φ gilt die Transformationsformel (siehe Abschnitt 10.3)

$$\widetilde{A} = S^{-1} A S, \qquad S \in \mathbf{GL}(n, \mathbb{K}).$$

18.3 Normen

18.3.1 Die Cauchy-Schwarz Ungleichung

Wir beginnen mit einer Ungleichung, die sowohl in euklidischen als auch unitären (endlich- oder unendlich-dimensionalen) Vektorräumen gilt.

Satz 18.12 (Cauchy-Schwarz Ungleichung) [8]
In einem euklidischen oder unitären Vektorraum $(V, \langle , \rangle)$ gilt für alle $a, b \in V$

$$|\langle a, b\rangle|^2 \le \langle a, a\rangle \langle b, b\rangle.$$

Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis: Wir führen den Beweis für einen unitären Vektorraum V durch. Der Beweis im euklidischen Fall ist völlig analog, mit dem einzigen Unterschied, dass "komplex-konjugieren" wegfällt.

Für $b = 0$ ist die Aussage richtig. Sei also $b \neq 0$. Für beliebige $\lambda \in \mathbb{C}$ gilt

$$0 \le \langle a - \lambda b, a - \lambda b\rangle = \langle a, a\rangle - \lambda \langle b, a\rangle - \overline{\lambda}\langle a, b\rangle + \lambda\overline{\lambda}\langle b, b\rangle.$$

TRICK: Wir setzen $\lambda := \frac{\langle a,b\rangle}{\langle b,b\rangle}$ und erhalten:

$$0 \le \langle a, a\rangle - \frac{\langle a, b\rangle \overline{\langle a, b\rangle}}{\langle b, b\rangle} - \frac{\overline{\langle a, b\rangle}\langle a, b\rangle}{\langle b, b\rangle} + \frac{\langle a, b\rangle \overline{\langle a, b\rangle}}{\langle b, b\rangle^2}\langle b, b\rangle.$$

Somit gilt wegen $\langle b, b\rangle > 0$

$$0 \le \langle a, a\rangle\langle b, b\rangle - \langle a, b\rangle\overline{\langle a, b\rangle} = \langle a, a\rangle\langle b, b\rangle - |\langle a, b\rangle|^2.$$

[8] Augustin Cauchy (1789-1857), Hermann Amandus Schwarz (1843-1921)

Gleichheit gilt genau dann, wenn $b = 0$ oder wenn $\langle a - \lambda b, a - \lambda b\rangle = 0$, d.h. wenn $a = \lambda b$ für ein gewisses $\lambda \in \mathbb{C}$. Also gilt Gleichheit genau dann, wenn a und b linear abhängig sind. ■

Beispiel 18.13 1. Es seien $\alpha_1, \ldots, \alpha_n$ und $\beta_1, \ldots, \beta_n$ reelle Zahlen. Dann gilt

$$(\sum_{i=1}^{n} \alpha_i \beta_i)^2 \leq (\sum_{i=1}^{n} \alpha_i^2)(\sum_{i=1}^{n} \beta_i^2).$$

Beweis: Wir setzen $a := (\alpha_1, \ldots, \alpha_n), b := (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$ und verwenden Cauchy-Schwarz für das Standard-Skalarprodukt in $\mathbb{R}^n$. ■

2. Für stetige Funktionen f, g auf dem Interval $[a, b]$ gilt

$$|\int_a^b f(t)g(t)dt|^2 \leq \int_a^b |f(t)|^2 \cdot \int_a^b |g(t)|^2.$$

Beweis: Vergleiche Beispiel 18.2. ■

Definition 18.14 Sei V ein reeller oder komplexer Vektorraum. Eine **Norm** auf V ist eine Funktion $\|\ \| : V \to \mathbb{R},\ a \mapsto \|a\|$ mit folgenden Eigenschaften:

Für alle $\lambda \in \mathbb{R}$ (oder $\mathbb{C}$) und : Skalarprodukte (Forts.) alle $a, b \in V$ gilt:

(i) $\|\lambda a\| = |\lambda|\|a\|$ (homogen)

(ii) $\|a + b\| \leq \|a\| + \|b\|$ (Dreiecks-Ungleichung)

(iii) $\|a\| \geq 0$ und $\|a\| = 0 \Longleftrightarrow a = 0$ (definit).

Ein Paar $(V, \|\ \|)$ bestehend aus einem Vektorraum und einer Norm heißt **normierter Vektorraum**.

Satz 18.15 *Es sei $(V, \langle,\rangle)$ ein euklidischer oder unitärer Vektorraum. Dann ist die Funktion $\|\ \| : V \to \mathbb{R}$ definiert durch $\|a\| = \sqrt{\langle a, a\rangle}$ eine Norm.*

Euklidische und unitäre Vektorräume sind also insbesondere auch normierte Vektorräume. Die Umkehrung gilt nicht (vgl. Beispiel 18.18).

Beweis: Wir weisen die drei Eigenschaften einer Norm nach.

(i) homogen:

$$\|\lambda a\| = \sqrt{\langle \lambda a, \lambda a\rangle} = \sqrt{\lambda\overline{\lambda}\langle a, a\rangle} = \sqrt{|\lambda|^2\|a\|^2} = |\lambda| \cdot \|a\|.$$

(ii) Dreiecks-Ungleichung: Zunächst gilt

$$\begin{aligned}
\|a+b\|^2 &= \langle a+b, a+b\rangle \\
&= \|a\|^2 + \langle a,b\rangle + \langle b,a\rangle + \|b\|^2 \\
&= \|a\|^2 + \langle a,b\rangle + \overline{\langle a,b\rangle} + \|b\|^2 \\
&= \|a\|^2 + 2\,\mathrm{Re}(\langle a,b\rangle) + \|b\|^2.
\end{aligned}$$

Weiter ist $\mathrm{Re}(\langle a,b\rangle) \leq |\langle a,b\rangle|$: Setzen wir $\langle a,b\rangle = \alpha + \mathrm{i}\beta$; $\alpha, \beta \in \mathbb{R}$, so gilt $\mathrm{Re}(\langle a,b\rangle) = \alpha \leq \sqrt{\alpha^2+\beta^2} = |\langle a,b\rangle|$. Also folgt mit Cauchy-Schwarz:

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

(iii) definit: Es ist $\|a\|^2 = \langle a,a\rangle \geq 0$ und $\langle a,a\rangle = 0 \iff a = 0$. Zieht man die Quadratwurzel, so folgt $\|a\| \geq 0$ und $\|a\| = 0 \iff a = 0$. ■

Beispiel 18.16 1. Die euklidische Norm: $\mathbb{R}^n$ mit Standard-Skalarprodukt ergibt

$$\|a\| = \|(\alpha_1, \ldots, \alpha_n)\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}.$$

2. Der Hilbertraum ℓ^2 der quadratsummierbaren Folgen.

   In Anlehnung an das vorige Beispiel sei jetzt V die Menge aller reellen Folgen $a = (\alpha_i)$, $\alpha_i \in \mathbb{R}$ mit der Zusatzbedingung, dass die Reihe $\sum\limits_{i=1}^{\infty} \alpha_i^2$ konvergiert. Durch

$$\begin{aligned}
a+b &= (\alpha_i) + (\beta_i) := (\alpha_i + \beta_i), \\
\lambda a &= \lambda(\alpha_i) := (\lambda\alpha_i), \quad \lambda \in \mathbb{R}
\end{aligned}$$

   wird auf ℓ^2 eine Addition und eine skalare Multiplikation definiert, wodurch ℓ^2 zu einem reellen Vektorraum wird. Dass die Addition sinnvoll ist, ergibt sich mit der Dreiecksungleichung in $\mathbb{R}^n$ (mit dem Standardskalarprodukt) und Grenzübergang $n \to \infty$:

$$\begin{aligned}
\sqrt{\sum_{i=1}^{n}(\alpha_i+\beta_i)^2} &= \|(\alpha_1,\ldots,\alpha_n) + (\beta_1,\ldots,\beta_n)\| \\
&\leq \|(\alpha_1,\ldots,\alpha_n)\| + \|(\beta_1,\ldots,\beta_n)\| \\
&= \sqrt{\sum_{i=1}^{n}\alpha_i^2} + \sqrt{\sum_{i=1}^{n}\beta_i^2} \leq \sqrt{\sum_{i=1}^{\infty}\alpha_i^2} + \sqrt{\sum_{i=1}^{\infty}\beta_i^2} < \infty.
\end{aligned}$$

Die Reihe $\sum_{i=1}^{\infty}(\alpha_i + \beta_i)^2$ hat somit eine beschränkte und monoton wachsende Partialsummenfolge, ist also konvergent. Also ist mit a, b auch $a + b$ in ℓ^2.

Um in ℓ^2 ein Skalarprodukt einzuführen, überlegen wir zuerst, dass für $a = (\alpha_i)$, $b = (\beta_i) \in \ell^2$ auch die Reihe $\sum_{i=1}^{\infty} \alpha_i\beta_i$ konvergiert:
Nach der Cauchy-Schwarz Ungleichung, angewandt auf $\mathbb{R}^n$ mit Standardskalarprodukt, gilt nämlich

$$\left(\sum_{i=1}^{n} |\alpha_i\beta_i|\right)^2 = \left(\sum_{i=1}^{n} |\alpha_i||\beta_i|\right)^2 \le \sum_{i=1}^{n} \alpha_i^2 \cdot \sum_{i=1}^{n} \beta_i^2 \le \sum_{i=1}^{\infty} \alpha_i^2 \cdot \sum_{i=1}^{\infty} \beta_i^2 < \infty.$$

Wir können also definieren

$$\langle a, b\rangle = \langle(\alpha_i), (\beta_i)\rangle := \sum_{i=1}^{\infty} \alpha_i\beta_i.$$

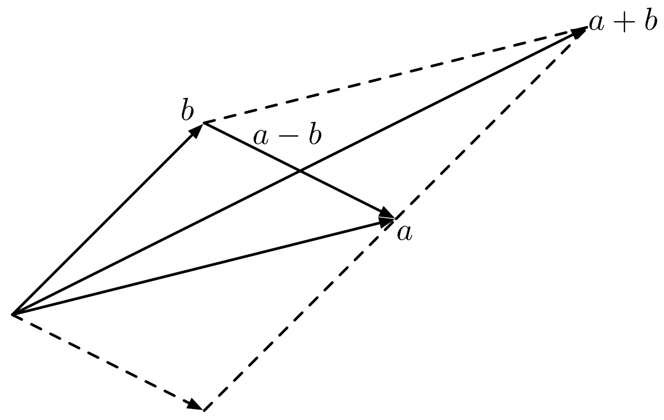
Man prüft nach, dass dadurch ein Skalarprodukt auf ℓ^2 definiert wird (d.h. $\langle , \rangle$ ist bilinear, symmetrisch und positiv definit). Der euklidische Vektorraum $(\ell^2, \langle , \rangle)$ ist zudem "vollständig", d.h. jede Cauchy–Folge konvergiert (vgl. Analysis-Vorlesung); man spricht dann von einem **Hilbert-Raum**.

Wir haben gesehen, dass ein Vektorraum mit Skalarprodukt auch ein normierter Vektorraum ist. Der folgende Satz liefert für *reelle* Vektorräume ein Kriterium für die Umkehrung.

Satz 18.17 (Parallelogramm-Identität) (a) *Sei $(V, \langle , \rangle)$ ein euklidischer oder unitärer Vektorraum mit zugehöriger Norm $\| \ \|$. Dann gilt die* **Parallelogramm-Identität**, *d.h. für alle $a, b \in V$ ist*

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

(b) *Ist umgekehrt $\| \ \|$ eine Norm auf einem* reellen *Vektorraum V, die die Parallelogramm-Identität erfüllt, so existiert ein Skalarprodukt $\langle , \rangle$ auf V mit $\|a\| = \sqrt{\langle a, a\rangle}$ für alle $a \in V$.*



Parallelogrammidentität

Beweis: (a) Sei $(V, \langle , \rangle)$ ein euklidischer oder unitärer Vektorraum. Dann haben wir für alle $a, b \in V$

$$\begin{aligned} & \langle a+b, a+b\rangle + \langle a-b, a-b\rangle = \\ & \quad \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle + \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle \\ = & \quad 2\langle a,a\rangle + 2\langle b,b\rangle . \end{aligned}$$

Mit der Definition der Norm erhält man die Parallelogramm-Identität.

(b) Gilt in einem normierten reellen Vektorraum die Parallelogramm-Identität, so definiert man für $a, b \in V$

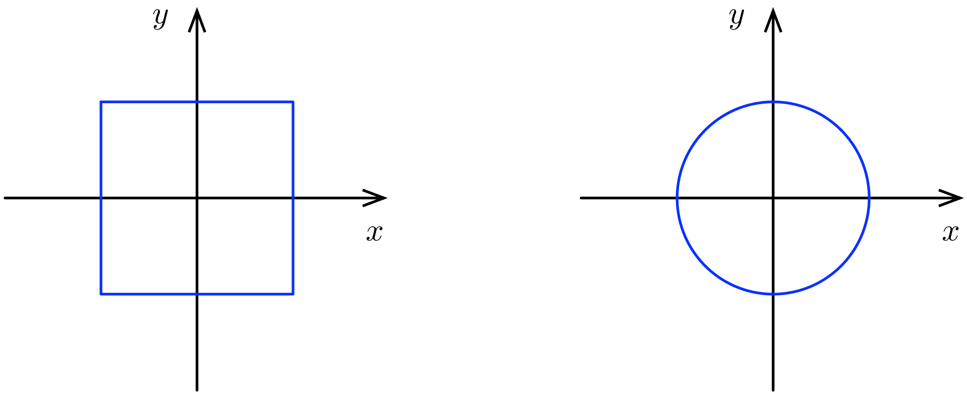
$$\langle a, b\rangle := \frac{1}{2}[\|a+b\|^2 - \|a\|^2 - \|b\|^2].$$

Dass $\langle , \rangle$ symmetrisch und positiv definit ist, folgt direkt aus dieser Formel. Um zu zeigen, dass $\langle , \rangle$ bilinear ist, benötigt man die Parallelogramm-Identität und ein Stetigkeitsargument (siehe [10], Kapitel 9.2). ■

Beispiel 18.18 Wir geben noch ein Beispiel einer Norm, die nicht von einem Skalarprodukt induziert wird. Wir betrachten dazu $\mathbb{R}^2$ versehen mit der Maximum-Norm, d.h. für $(x, y) \in \mathbb{R}^2$ setzen wir $\|(x, y)\| := \max(|x|, |y|)$. Für diese Norm gilt die Parallelogramm-Identität nicht: Für $a = (1, 0)$ und $b = (0, 1)$ ist

$$\|a+b\| = \|(1,1)\| = 1, \; \|a-b\| = \|(1,-1)\| = 1 \text{ also}$$

$$\|a+b\|^2 + \|a-b\|^2 = 2 \neq 4 = 2\|a\|^2 + 2\|b\|^2.$$



Vektoren der Länge 1 für die Maximums-Norm bzw. die Standardnorm in $\mathbb{R}^2$

18.3.2 Metrische Räume

Definition 18.19 Für eine beliebige Menge M heißt eine Funktion $d : M \times M \to \mathbb{R}$ eine **Metrik** (oder **Abstandsfunktion**) , wenn d die folgenden drei Eigenschaften erfüllt:

1. $\forall p, q \in M : d(p, q) = d(q, p) \quad$ (symmetrisch)
2. $\forall p, q, r \in M : d(p, r) \leq d(p, q) + d(q, r) \quad$ (Dreiecks-Ungleichung)
3. $\forall p, q \in M : d(p, q) \geq 0$ und $d(q, p) = 0 \iff p = q \quad$ (positiv).

Das Paar (M, d) heißt dann **metrischer Raum**.

Bemerkung 18.20 Jede Menge kann zu einem metrischen Raum gemacht werden. Durch

$$d(p, q) := \begin{cases} 0 & \text{falls} \quad p = q, \\ 1 & \text{sonst} \end{cases}$$

wird auf M die (nicht besonders interessante) **diskrete Metrik** definiert.

Hilfssatz 18.21 *Ein normierter Vektorraum ist ein metrischer Raum.*

Beweis: Für $x, y \in V$ definieren wir $d(x, y) := \|x - y\| \geq 0$. Da eine Norm homogen ist gilt

$$d(y, x) = \|y - x\| = \| - (x - y)\| = \|x - y\| = d(x, y).$$

Mit der Dreiecks-Ungleichung für die Norm folgt für $x, y, z \in V$

$$d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z).$$

Schließlich ist d positiv, denn aus $d(x, y) = 0$ folgt $\|x - y\| = 0$ also (weil die Norm positiv ist) $x = y$. ∎

Bemerkung 18.22 Nicht jede Metrik auf einem Vektorraum wird durch eine Norm induziert. Ein Beispiel ist die diskrete Metrik (wieso?).

18.3.3 Winkel

Gegeben sei ein euklidischer Vektorraum $(V, \langle , \rangle)$. Nach der Cauchy–Schwarz Ungleichung (18.12) gilt für zwei vom Nullvektor verschiedene Vektoren $a, b \in V$:

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

also

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

Im Intervall $[0, \pi] \subset \mathbb{R}$ ist die Cosinus-Funktion cos streng monoton fallend und bildet $[0, \pi]$ bijektiv auf das Intervall $[-1, 1]$ ab. Damit können wir definieren

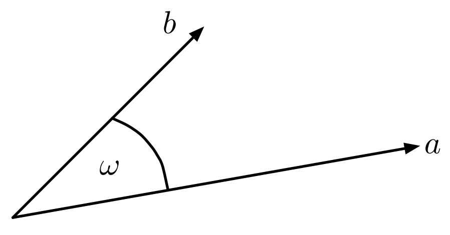
Definition 18.23 Seien $a \neq 0$, $b \neq 0$ zwei Vektoren eines euklidischen Vektorraumes $(V, \langle , \rangle)$. Die eindeutig bestimmte reelle Zahl $\omega(a, b) \in [0, \pi]$ mit

$$\cos \omega(a, b) = \frac{\langle a, b\rangle}{\|a\| \cdot \|b\|}.$$

heißt **Winkel** zwischen a und b.

Diese Definitionsgleichung kann man auch so schreiben:

$$\langle a, b\rangle = \|a\| \cdot \|b\| \cdot \cos \omega(a, b).$$



Die Abbildung

$$\omega : V\backslash\{0\} \times V\backslash\{0\} \to \mathbb{R}; \quad (a, b) \mapsto \omega(a, b)$$

nennen wir **Winkelfunktion**.

Hilfssatz 18.24 (Eigenschaften der Winkelfunktion) *Für alle* $a, b \in V\backslash\{0\}$ *und alle* $\alpha, \beta \in \mathbb{R}\backslash\{0\}$ *gilt*

1. $\omega(a,b) = \omega(b,a)$

2. $\omega(\alpha a, \beta b) = \begin{cases} \omega(a,b) & \textit{für} \quad \alpha\beta > 0 \\ \pi - \omega(a,b) & \textit{für} \quad \alpha\beta < 0 \end{cases}$

3. $\omega(a,b) = 0 \;\Leftrightarrow\; b = \lambda a \quad$ *für ein* $\quad \lambda > 0$

4. $\omega(a,b) = \pi \;\Leftrightarrow\; b = \lambda a \quad$ *für ein* $\quad \lambda < 0.$

Beweis: 1. folgt aus der Definitionsgleichung der Winkelfunktion wegen der Symmetrie des Skalarprodukts.

Weiter ist

$$\cos\omega(\alpha a, \beta b) = \frac{\langle \alpha a, \beta b\rangle}{\|\alpha a\| \cdot \|\beta b\|} = \frac{\alpha\beta}{|\alpha\beta|}\cos\omega(a,b).$$

Für $\alpha\beta > 0$ ist der erste Teil von 2. sofort klar. Für $\alpha\beta < 0$ gilt

$$\cos\omega(\alpha a, \beta b) = -\cos\omega(a,b) = \cos(\pi - \omega(a,b)),$$

woraus der zweite Teil von 2. folgt.

3. und 4. ergeben sich so: Zunächst sind $a, b \neq 0$ nach Cauchy–Schwarz genau dann linear abhängig, wenn $|\langle a,b\rangle| = \|a\| \cdot \|b\|$ ist, und zwar ist

$$\begin{aligned} b = \lambda a \text{ für } \lambda > 0 \quad &\Leftrightarrow \quad \langle a,b\rangle = \|a\| \cdot \|b\|, \\ b = \lambda a \text{ für } \lambda < 0 \quad &\Leftrightarrow \quad \langle a,b\rangle = -\|a\| \cdot \|b\|. \end{aligned}$$

Weiter gilt nun nach nach der Definitionsgleichung der Winkelfunktion

$$\langle a,b\rangle = \|a\| \cdot \|b\| \quad \Leftrightarrow \quad \cos\omega(a,b) = 1 \quad \Leftrightarrow \quad \omega(a,b) = 0$$

und entsprechend

$$\langle a,b\rangle = -\|a\| \cdot \|b\| \quad \Leftrightarrow \quad \cos\omega(a,b) = -1 \quad \Leftrightarrow \quad \omega(a,b) = \pi.$$

■

18.4 Orthogonalität und Orthonormalbasen

Der Winkel $\omega(a,b)$ zwischen zwei Vektoren $a \neq 0, b \neq 0$ eines euklidischen Vektorraumes genau dann $\frac{\pi}{2}$, wenn $\langle a,b\rangle = 0$. Allgemein definieren wir

Definition 18.25 Zwei Vektoren a, b eines euklidischen oder unitären Vektorraumes $(V, \langle , \rangle)$ heißen **orthogonal** oder **senkrecht**, wenn $\langle a, b\rangle = 0$. (Schreibweise $a \perp b$.)

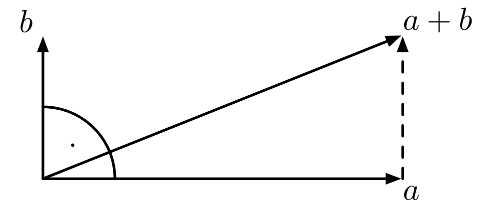
Bemerkung 18.26 In diese Definition ist nun auch der Nullvektor einbezogen. Es ist $\langle a, 0\rangle = 0$ für alle $a \in V$, d. h. 0 ist zu allen $a \in V$ orthogonal. Andererseits ist 0 auch der einzige Vektor, der zu allen $a \in V$ orthogonal ist. Denn aus $\langle a, b\rangle = 0$ für alle $a \in V$ folgt insbesondere $\langle b, b\rangle = 0$, also $b = 0$ wegen der positiven Definitheit des Skalarprodukts.

Satz 18.27 (Pythagoras) *In einem euklidischen oder unitären Vektorraum gilt*

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

In einem euklidischen Vektorraum gilt auch die Umkehrung:

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



Beweis: Die erste Aussage folgt aus

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

Im euklidischen Fall gilt

$$a \perp b \Leftrightarrow 0 = \langle a, b\rangle + \langle b, a\rangle = 2\langle a, b\rangle,$$

und damit die zweite Aussage des Satzes. ∎

Definition 18.28 Eine Teilmenge $S \neq \emptyset$ eines euklidischen oder unitären Vektorraumes $(V, \langle , \rangle)$ heißt **orthogonal**, wenn je zwei verschiedene Elemente von S orthogonal sind: Für $a, b \in S$, $a \neq b$ gilt $a \perp b$.

Eine orthogonale Menge S heißt **orthonormiert**, wenn $\|a\| = 1$ ist für alle $a \in S$. Eine orthonormierte Teilmenge von V, die zugleich eine Basis von V ist, heißt **Orthonormalbasis** (ONB).

Beispiel 18.29 1. Die Standardbasis von $\mathbb{R}^n$ ist eine Orthonormalbasis bezüglich dem Standard-Skalarprodukt.

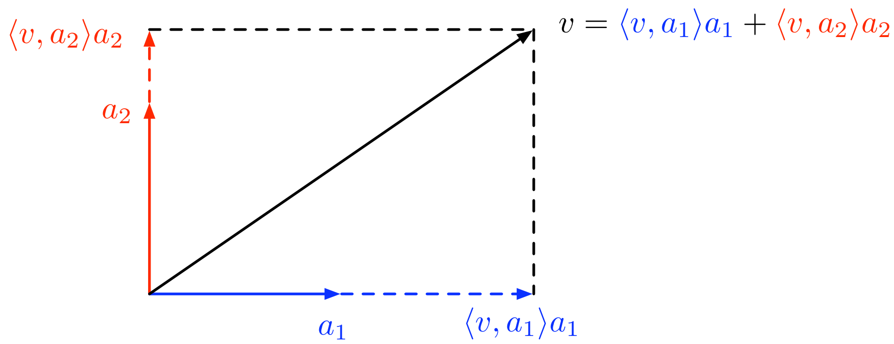
2. Im Vektorraum $C[-\pi, \pi]$ der stetigen Funktionen auf dem Intervall $[-\pi, \pi] \subset \mathbb{R}$ versehen mit dem (euklidischen) Skalarprodukt $\langle f, g \rangle := \int_{-\pi}^{\pi} f(t)g(t)dt$ ist die Menge
$$S = \{1, \sin t, \cos t, \sin 2t, \cos 2t, \sin 3t, \cos 3t, \dots\}$$
orthogonal.

Bemerkung 18.30 $\{a_1, \dots, a_n\}$ ist eine Orthonormalbasis von $(V, \langle , \rangle)$ genau dann, wenn
$$\langle a_i, a_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{für} \quad i \neq j \\ 1 & \text{für} \quad i = j \end{cases}$$

Hilfssatz 18.31 *1. Eine orthogonale Teilmenge $S \subset V \backslash \{0\}$ ist linear unabhängig.*

2. Eine orthogonale Teilmenge $S \subset V \backslash \{0\}$, kann man durch Normieren zu einer orthonormierten Teilmenge machen:
$S^* := \{\|a\|^{-1}a = \frac{a}{\|a\|} \mid a \in S\}$ ist orthonormiert.

3. Sei $\{a_1, \dots, a_n\}$ eine ONB von $(V, \langle , \rangle)$. Dann gilt für alle $v \in V$
$$v = \sum_{i=1}^{n} \langle v, a_i \rangle a_i.$$



Beweis: 1. Sei $\{a_1, \dots, a_r\}$ eine endliche Teilmenge von S und $\sum_{i=1}^{r} \lambda_i a_i = 0$. Dann haben wir für alle $k = 1, \dots, r$
$$0 = \langle \sum_{i=1}^{r} \lambda_i a_i, a_k \rangle = \sum_{i=1}^{r} \lambda_i \langle a_i, a_k \rangle = \lambda_k \langle a_k, a_k \rangle.$$
Da $S \subset V \backslash \{0\}$ ist $a_k \neq 0$ und es folgt $\lambda_k = 0$.

2.
$$\langle \frac{a}{\|a\|}, \frac{b}{\|b\|} \rangle = \frac{1}{\|a\|} \frac{1}{\|b\|} \langle a, b \rangle = \begin{cases} 0 & \text{für} \quad a \neq b \\ 1 & \text{für} \quad a = b \end{cases}$$

3. Ist $v = \sum_{i=1}^n \lambda_i a_i$, so gilt für $k = 1, \ldots, n$
$$\langle v, a_k \rangle = \langle \sum_{i=1}^n \lambda_i a_i, a_k \rangle = \sum_{i=1}^n \lambda_i \langle a_i, a_k \rangle = \sum_{i=1}^n \lambda_i \delta_{ik} = \lambda_k.$$

■

Bemerkung 18.32 Die Darstellungsmatrix eines Skalarproduktes bezüglich einer Orthonormalbasis $B = \{e_1, e_2, \ldots, e_n\}$ ist die Einheitsmatrix (also besonders einfach):
$$g_{ij} = \langle e_i, e_j \rangle = \delta_{ij} \quad \text{also} \quad G = (g_{ij}) = E_n.$$
Insbesondere gilt für Vektoren $x, y \in V$ mit $x = \sum_{i=1}^n x_i e_i$ und $y = \sum_{i=1}^n y_i e_i$:
$$\langle x, y \rangle = (x_1, \ldots, x_n) \cdot E_n \cdot \begin{pmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_n \end{pmatrix} = \sum_{i=1}^n x_i \overline{y_i} = \Theta_B(x)^\top \Theta_B(y),$$
also
$$\begin{aligned} \langle x, y \rangle &= x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n, \quad \|x\|^2 = \sum_{i=1}^n |x_i|^2 \quad \text{im unitären Fall, und} \\ \langle x, y \rangle &= x_1 y_1 + \cdots + x_n y_n, \quad \|x\|^2 = \sum_{i=1}^n x_i^2 \quad \text{im euklidischen Fall.} \end{aligned}$$

Der folgende Satz beschreibt, wie man aus einer gegebenen Basis eine Orthonormalbasis konstruieren kann.

Satz 18.33 (Orthogonalisierungs-Verfahren von Gram–Schmidt) [9]

Sei V Vektorraum mit Skalarprodukt $\langle , \rangle$.

(a) *Wenn $\{b_1, b_2, \ldots, b_l\}$ eine linear unabhängige Teilmenge von V ist, dann ist die Teilmenge $\{a_1, a_2, \ldots, a_l\}$, die rekursiv definiert ist durch*
$$a_1 := b_1, \quad a_{k+1} := b_{k+1} - \sum_{j=1}^k \frac{\langle b_{k+1}, a_j \rangle}{\langle a_j, a_j \rangle} \cdot a_j \quad (\text{für } k = 1, \ldots, l-1),$$
orthogonal.

[9] Jorgen Gram (1850-1916), Erhard Schmidt (1876-1959)

(b) *Die Menge* $\{\frac{1}{\|a_1\|}a_1, \ldots, \frac{1}{\|a_l\|}a_l\}$ *ist orthonormiert.*

(c) *Für die linearen Hüllen gilt*

$$[a_1, \ldots, a_i] = [b_1, \ldots, b_i], \quad i = 1, 2, \ldots, l.$$

(d) *Ist V endlich-dimensional, so besitzt V eine Orthonormalbasis.*

Beweis: (b) folgt aus (a) durch normieren. (d) folgt aus (a), (b) und (c) angewandt auf eine Basis von V.

Für (a) und (c) verwenden wir vollständige Induktion nach l und konstruieren eine explizite orthogonale Menge.

Induktions-Verankerung: Sei $l = 1$. Wir setzen $a_1 := b_1$. Dann ist a_1 (als einzelner, von Null verschiedener Vektor) orthogonal und $[a_1] = [b_1]$.

Induktions-Annahme: Für beliebiges l sei $\{a_1, \ldots, a_l\}$ orthogonal und für die lineare Hülle gelte $[a_1, \ldots, a_l] = [b_1, \ldots, b_l]$.

Induktions-Schluss: Wir machen den Ansatz:

$$a_{l+1} = b_{l+1} + \lambda_1 a_1 + \ldots + \lambda_l a_l \quad (\lambda_i \in \mathbb{C}).$$

Da $\langle a_{l+1}, a_j \rangle = 0$ sein soll für $j = 1, \ldots, l$, ergeben sich die l Bedingungen:

$$\langle b_{l+1}, a_j \rangle + \lambda_j \langle a_j, a_j \rangle = 0, \text{ also } \lambda_j = -\frac{\langle b_{l+1}, a_j \rangle}{\langle a_j, a_j \rangle}.$$

Somit ist notwendigerweise

$$a_{l+1} = b_{l+1} - \sum_{j=1}^{k} \frac{\langle b_{l+1}, a_j \rangle}{\langle a_j, a_j \rangle} \cdot a_j. \qquad (*)$$

Es ist $a_{l+1} \neq 0$. Denn sonst wäre nach $(*)$ und der Induktions-Annahme $b_{l+1} \in [a_1, \ldots, a_l] = [b_1, \ldots, b_l]$; ein Widerspruch dazu, dass die b_j für $j = 1, \ldots, l+1$ (ebenfalls nach Induktions-Annahme) linear unabhängig sind. Weiter ist nach Konstruktion $a_{l+1} \perp a_j$ für $j = 1, \ldots, l$, also ist $\{a_1, \ldots, a_l, a_{l+1}\}$ orthogonal.

Es bleibt noch die Aussage (c) über die lineare Hülle zu zeigen. Nach Konstruktion und Induktions-Anahme haben wir zunächst

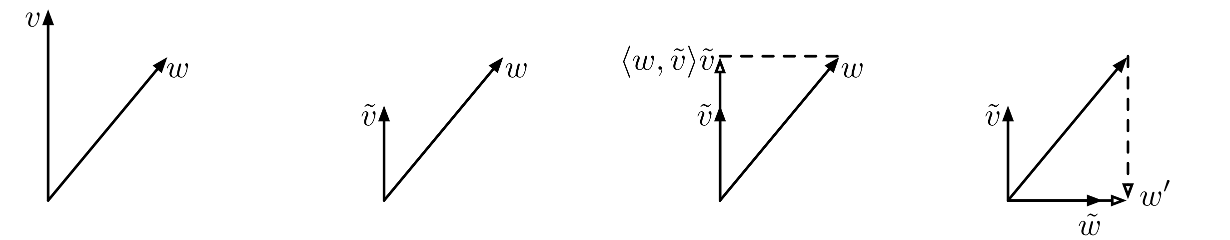
$$[a_1, \ldots, a_{l+1}] \subset [b_1, \ldots, b_{l+1}].$$

Wegen $(*)$ ist $b_{l+1} \in [a_1, \ldots, a_{l+1}]$, also folgt zusammen mit der Induktions-Annahme, dass

$$[a_1, \ldots, a_{l+1}] = [b_1, \ldots, b_{l+1}].$$

Damit ist der Satz bewiesen. ■

Hier nochmals eine schematische Darstellung der einzelnen Schritte im Verfahren von Gram-Schmidt:



v, w sind nicht orthogonal $\dashrightarrow$ normiere v zu $\tilde v$ $\dashrightarrow$ bestimme $\langle w, \tilde v\rangle$ projiziere w auf die $\tilde v$-Achse $\dashrightarrow$ $w' = w - \langle w, \tilde v\rangle \tilde v$ normiere w' zu $\tilde w$ ONB: $\{\tilde v, \tilde w\}$

Beispiel 18.34 Sei V der Vektorraum aller reellen Polynome vom Grad ≤ 3. Für ein Polynom $p = a_0 + a_1 X + a_2 X^2 + a_3 X^3 \in V$ bezeichne $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ die zugehörige Polynomfunktion, $p(t) : \mathbb{R} \to \mathbb{R}$. Auf V ist durch

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

ein Skalarprodukt definiert. Die Matrix dieses Skalarproduktes bezüglich der Basis $B = \{b_1 = 1, b_2 = X, b_3 = X^2, b_4 = X^3\}$ ist

$$G = (g_{ij}) = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix},$$

denn man berechnet beispielsweise

$$g_{44} = \langle X^3, X^3\rangle = \int_0^1 t^3 \cdot t^3 dt = \frac{1}{7} t^7\Big|_0^1 = \frac{1}{7}.$$

Da $G \neq E_4$, ist B keine ONB. Wir bestimmen jetzt eine orthogonale Basis mittels Gram-Schmidt-Orthogonalisierung:

$$\begin{aligned}
a_1 &:= b_1 = 1 \\
a_2 &:= b_2 - \frac{\langle b_2, a_1\rangle}{\langle a_1, a_1\rangle} a_1 = X - \frac{1}{2} \\
a_3 &:= b_3 - \frac{\langle b_3, a_2\rangle}{\langle a_2, a_2\rangle} a_2 - \frac{\langle b_3, a_1\rangle}{\langle a_1, a_1\rangle} a_1 = X^2 - X + \frac{1}{6} \\
a_4 &:= b_4 - \frac{\langle b_4, a_3\rangle}{\langle a_3, a_3\rangle} a_3 - \frac{\langle b_4, a_2\rangle}{\langle a_2, a_2\rangle} a_2 - \frac{\langle b_4, a_1\rangle}{\langle a_1, a_1\rangle} a_1 = X^3 - \frac{3}{2} X^2 + \frac{3}{5} X - \frac{1}{20}.
\end{aligned}$$

Durch normieren, erhält man schließlich eine ONB.

18.5 Orthogonal-Komplemente und Orthogonal-Projektionen

Ist V ein Vektorraum und $U \subset V$ ein Untervektorraum, so hat U im Allgemeinen viele Komplementärräume. Ist hingegen $(V, \langle , \rangle)$ ein euklidischer oder unitärer Vektorraum, so gibt es ein "kanonisches" Komplement.

Definition 18.35 Die Menge $U^\perp := \{x \in V \mid \langle x, u\rangle = 0 \ \forall u \in U\}$ heißt **Orthogonal-Komplement** von U in V.

Gilt $\langle x, u\rangle = 0$ für alle $u \in U$, so schreiben wir kurz $x \perp U$.

Satz 18.36 *Das Orthogonal-Komplement $U^\perp$ eines Untervektorraumes U eines euklidischen oder unitären Vektorraumes V hat folgende Eigenschaften:*

(a) *$U^\perp$ ist Untervektorraum von V,*

(b) $U \cap U^\perp = \{0\}$,

(c) *Ist V endlich-dimensional, so gilt $U \oplus U^\perp = V$.*

Beweis:
(a) gilt nach dem Untervektorraum–Kriterium.

(b) Für $x \in U \cap U^\perp$ gilt $\langle x, x\rangle = 0$. Da $\langle , \rangle$ positiv definit ist, folgt $x = 0$.

(c) Für die "Extremfälle" $U = \{0\}$ bzw. $U = V$ ist $\{0\}^\perp = V$ bzw. $V^\perp = \{0\}$, also gilt (c). In den anderen Fällen wählen wir eine Orthonormalbasis $\{e_1, \ldots, e_k\}$, $1 \leq k < n$, von U und ergänzen diese zu einer Orthonormalbasis $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ von V (das ist möglich nach dem Basis-Ergänzungssatz und Gram-Schmidt). Wir zeigen, dass $U^\perp = [e_{k+1}, \ldots, e_n]$ gilt. Für jedes $x \in [e_{k+1}, \ldots, e_n]$ ist $x \perp U$ nach Konstruktion, d.h. $x \in U^\perp$ und wir haben $U^\perp \supset [e_{k+1}, \ldots, e_n]$. Ist andererseits $x = \sum\limits_{j=1}^{n} \alpha_j e_j$ ein Vektor aus $U^\perp$, so gilt $x \perp U$, also insbesondere $\alpha_p = \langle x, e_p\rangle = 0$ für $p = 1, \ldots, k$. Somit ist $x \in [e_{k+1}, \ldots, e_n]$. Da x beliebig war haben wir auch $U^\perp \subset [e_{k+1}, \ldots, e_n]$. ■

Beispiel 18.37 1. Sei $V = \mathbb{R}^3$ und $U := \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y + 4z = 0\}$ eine Ebene. Dann ist $U^\perp = [(2, 3, 4)]$, denn

$$\langle (a, b, c), (x, y, z)\rangle = 0 \Leftrightarrow ax + by + cz = 0.$$

2. Verallgemeinerung: Sei $V = \mathbb{R}^n$ und $U := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n a_i x_i = 0\}$ eine Hyperebene. Dann ist $U^\perp = [(a_1, \ldots, a_n)]$.

3. Sei $V = C[-1,1]$ der Vektorraum der stetigen, reellwertigen Funktionen auf dem Intervall $[-1,1]$ versehen mit dem Skalarprodukt $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$. U sei der Untervektorraum der *geraden* Funktionen:

$$U = \{g \in V \mid g(-t) = g(t) \quad \forall t \in [-1,1]\}.$$

Das orthogonale Komplement von U ist dann der Untervektorraum der *ungeraden* Funktionen:

$$U^{\perp} = \{u \in V \mid u(-t) = -u(t) \quad \forall t \in [-1,1]\}.$$

Beweis: Jedes $f \in V$ lässt sich (eindeutig) schreiben als Summe einer geraden Funktion f_+ und einer ungeraden Funktion f_-:

$$f(t) = \frac{1}{2}(f(t) + f(-t)) + \frac{1}{2}(f(t) - f(-t)) =: f_+(t) + f_-(t).$$

Weiter gilt für g gerade und u ungerade

$$\begin{aligned}\langle g, u\rangle &= \int_{-1}^{1} g(t)u(t)dt = \int_{-1}^{0} g(t)u(t)dt + \int_{0}^{1} g(t)u(t)dt \\ &= \int_{0}^{1} g(-t)u(-t)dt + \int_{0}^{1} g(t)u(t)dt \\ &= -\int_{0}^{1} g(t)u(t)dt + \int_{0}^{1} g(t)u(t)dt = 0.\end{aligned}$$

■

Bemerkung 18.38 Falls $\dim V = \infty$, so ist im Allgemeinen $U \oplus U^{\perp} \neq V$. Hier ist ein Beispiel:

Es sei $V = \ell^2$ der Hilbertraum der quadrat-summierbaren Folgen reeller Zahlen (vgl. Beispiel 18.16). Weiter sei $U \subset \ell^2$ der Untervektorraum aller Polynome, d.h. aller *endlichen* Folgen. Dann ist $U \subsetneq \ell^2$ und es ist $U^{\perp} = \{0\}$. Denn ist etwa $a = (a_i)_{i\in\mathbb{N}} \in V$ und $a \perp U$, so gilt insbesondere

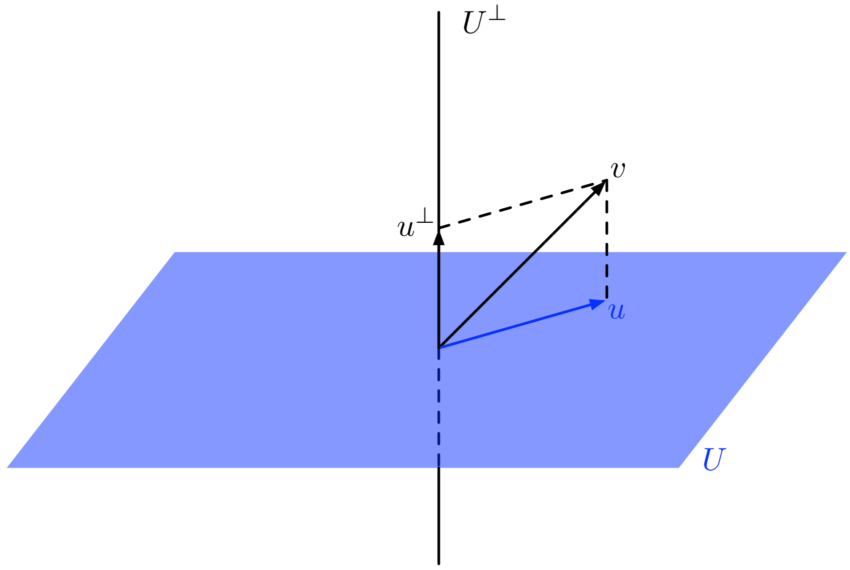
$$\langle a, (0, \ldots, 0, 1, 0, \ldots)\rangle = a_i = 0 \quad \text{für } i = 1, 2, \ldots,$$

und damit $a = 0$.

Für einen endlich-dimensionalen euklidischen oder unitären Vektorraum V und einen Untervektorraum U ist nach Satz 18.36 $V = U \oplus U^{\perp}$. Jeder Vektor $v \in V$ hat also eine eindeutige Darstellung (oder Zerlegung) $v = u + u^{\perp}$ mit $u \in U, u^{\perp} \in U^{\perp}$.

Definition 18.39 Die **Orthogonalprojektion** von V auf U (in Richtung $U^\perp$) ist die Abbildung:

$$\pi_U : V \to U \subset V, \quad v = u + u^\perp \mapsto u.$$



Satz 18.40 (Eigenschaften der Orthogonalprojektion) *Für die Orthogonalprojektion π_U eines Vektorraumes V auf einen Untervektorraum U gilt:*

1. *π_U ist linear und $\pi_U^2 = \pi_U \circ \pi_U = \pi_U$.*
2. Bild $\pi_U = U$, Kern $\pi_U = U^\perp$.
3. *π_U verkürzt Abstände: Für alle $v, w \in V$ ist*

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

Beweis: 1. folgt direkt aus der Definition bzw. durch Nachrechnen. 2. folgt aus den Eigenschaften der direkten Summe und der Definition von π_U. Zu 3.: Sei $v = u + u^\perp \in V$. Zunächst folgt wegen $\langle u, u^\perp \rangle = 0$:

$$\|\pi_U(v)\|^2 = \|u\|^2 = \langle u, u \rangle \leq \langle u, u \rangle + \langle u^\perp, u^\perp \rangle = \langle u + u^\perp, u + u^\perp \rangle = \langle v, v \rangle = \|v\|^2.$$

Die Behauptung folgt jetzt aus der Linearität von π_U. ∎

Bemerkung 18.41 (Formel für Orthogonalprojektion) Es sei $U \neq \{0\}$ ein echter Untervektorraum von V mit Orthonormalbasis $\{e_1, \ldots, e_k\}$. Wir ergänzen zu einer Orthonormalbasis $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ von V; insbesondere ist $\{e_{k+1}, \ldots, e_n\}$

eine Orthonormalbasis von $U^\perp$. Für einen beliebigen Vektor $v \in V$ haben wir $v = u + u^\perp = \sum_{j=1}^{k} \alpha_j e_j + \sum_{j=k+1}^{n} \alpha_j e_j$ mit $\alpha_j = \langle v, e_j \rangle \quad (j = 1, \ldots, n)$. Die orthogonale Projektion π_U auf U läßt sich also auch schreiben als

$$\pi_U(v) = \sum_{j=1}^{k} \langle v, e_j \rangle e_j.$$

Beispiel 18.42 Sei $V = \mathbb{R}^5$, versehen mit dem Standardskalarprodukt. Weiter seien

$$U = [\begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}2\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\\0\end{pmatrix}] \quad \text{sowie} \quad v = \begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix}.$$

Gesucht ist die Orthogonalprojektion von v auf U.

1. SCHRITT: Wir bestimmen eine ONB in U mit Gram-Schmidt:

$$v_1 = \begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix}, \quad \|v_1\| = \sqrt{3},$$

$$v_2 = \begin{pmatrix}2\\0\\1\\1\\0\end{pmatrix} - \begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix}, \quad \|v_2\| = \sqrt{3},$$

$$v_3 = \begin{pmatrix}0\\0\\1\\1\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix} = \frac{2}{3}\begin{pmatrix}-1\\0\\1\\1\\0\end{pmatrix}, \quad \|v_3\| = \frac{2}{\sqrt{3}}.$$

Damit erhalten wir als ONB in U:

$$b_1 := \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix}, b_2 := \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix}, b_3 := \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\0\\1\\1\\0\end{pmatrix}.$$

2. SCHRITT: Es ist

$$\pi_U(v) = \langle v, b_1\rangle b_1 + \langle v, b_2\rangle b_2 + \langle v, b_3\rangle b_3 = \frac{1}{3}\begin{pmatrix} 3 \\ 5 \\ 8 \\ 13 \\ 0 \end{pmatrix}.$$

18.5.1 Abstand eines Vektors zu einem Untervektorraum

Wir beginnen mit einer allgemeinen Definition.

Definition 18.43 Seien (M, d) ein metrischer Raum und $A, B \subset M$ zwei Teilmengen. Der **Abstand** von A und B ist definiert durch

$$d(A, B) := \inf\{d(a, b) \mid a \in A, b \in B\}.$$

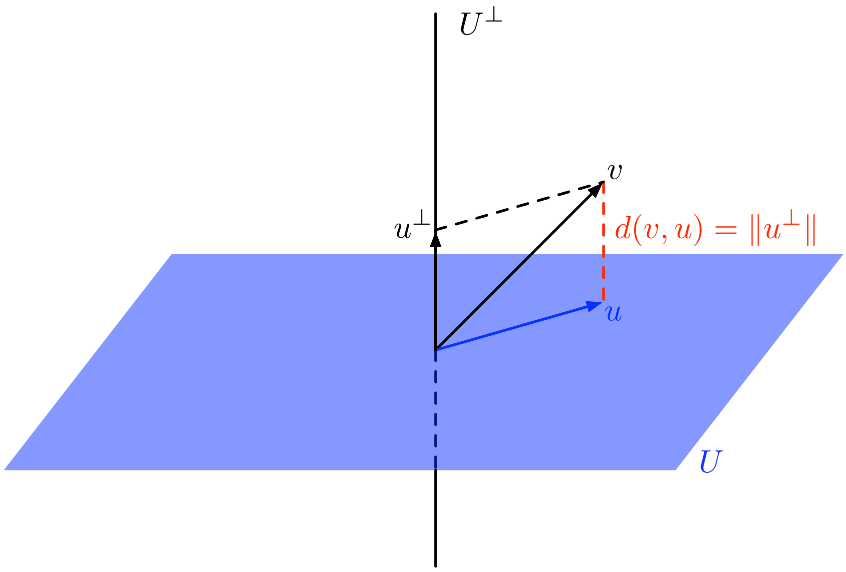
Ist jetzt V ein Vektorraum mit Skalarprodukt, so ist V ein normierter und damit ein metrischer Raum mit Metrik $d(v, w) = \|v - w\|$. Für $v \in V$ und einen Untervektorraum U wollen wir den Abstand von v und U bestimmen, also

$$d(\{v\}, U) = \inf\{\|v - u\| \mid u \in U\}.$$

Man beachte, dass die Menge der reellen Zahlen $\|v-u\|$ mit $u \in U$ nach unten durch 0 beschränkt ist; das Infimum existiert also (das gilt ganz allgemein, da eine Metrik positiv ist).

Satz 18.44 (Kürzester Abstand) *Es seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und U ein Untervektorraum. Weiter sei V zerlegt als $V = U \oplus U^\perp$. Dann gilt für $v = u + u^\perp \in V$ die Gleichung:*

$$d(\{v\}, U) = \min\{\|v - u^*\| \mid u^* \in U\} = \|u^\perp\| = \|\pi_{U^\perp}(v)\|.$$



Beweis: Für $u^* \in U$ gilt nach dem Satz von Pythagoras 18.27

$$d(v,u^*)^2 = \|(u-u^*)+u^\perp\|^2 = \|u-u^*\|^2 + \|u^\perp\|^2 \geq \|u^\perp\|^2.$$

Für $u^* = u$ gilt Gleichheit in dieser Ungleichung, d.h. das Infimum wird angenommen und ist also ein Minimum. Damit folgt die Behauptung. ■

18.5.2 Skalarprodukt und Dualraum

Es sei $(V, \langle , \rangle)$ ein Vektorraum mit Skalarprodukt. Der Dualraum $V^\star$ von V ist der Vektorraum aller Linearformen auf V:

$$V^\star = \{\Phi : V \to \mathbb{K} \mid \Phi \text{ linear}\} \text{ für } \mathbb{K} = \mathbb{R} \text{ oder } \mathbb{C}.$$

Mittels des Skalarproduktes kann man spezielle Linearformen konstruieren:

Sei dazu $a \in V$ fest gewählt. Dann ist die Abbildung

$$\Lambda_a : V \to \mathbb{K};\ \Lambda_a(x) = \langle x, a\rangle$$

eine Linearform auf V, also $\Lambda_a \in V^\star$, da das Skalarprodukt im 1. Argument linear ist.

Bemerkung 18.45 Die geometrische Interpretation dieser Konstruktion wird durch folgenden Spezialfall illustriert. Es sei $e \in V$ ein Einheitsvektor, d.h. $\|e\| = 1$, und $U := [e]$ ein 1-dimensionaler Untervektorraum von V. Weiter sei $\pi_U : V \to U$ die Orthogonalprojektion auf U. Dann gilt für alle $x \in V$:

$$\pi_U(x) = \langle x, e\rangle e \quad \text{und} \quad \|\pi_U(x)\| = |\langle x, e\rangle|.$$

Ist V endlich dimensional, so sind alle Linearformen von obiger Gestalt:

Satz 18.46 (Satz von Riesz) [10] *Es sei V ein n–dimensionaler Vektorraum mit Skalarprodukt und $\Lambda \in V^\star$. Dann existiert genau ein $a \in V$, so dass für alle $x \in V$ gilt*

$$\Lambda(x) = \langle x, a \rangle.$$

Beweis:
Existenz: Es sei $\{e_1, \dots, e_n\}$ eine Orthonormalbasis von V. Wir machen den *Ansatz*: $a = \sum_{k=1}^{n} \alpha_k \, e_k$. Dann gilt dann für alle $x = \sum_{k=1}^{n} x_k e_k \in V$:

$$\langle x, a \rangle \;=\; \langle \sum_{k=1}^{n} x_k e_k, \sum_{r=1}^{n} \alpha_r e_r \rangle = \sum_{k,r=1}^{n} x_k \overline{\alpha_r} \delta_{kr} = \sum_{k=1}^{n} x_k \overline{\alpha_k}.$$

Andererseits haben wir

$$\Lambda(x) = \Lambda(\sum_{k=1}^{n} x_k e_k) = \sum_{k=1}^{n} x_k \Lambda(e_k).$$

Setzen wir also $\alpha_k := \overline{\Lambda(e_k)}$ für $k = 1, \dots, n$, so erhalten wir $\Lambda(x) = \langle x, a \rangle$.

Eindeutigkeit: Aus $\langle x, a_1 \rangle = \langle x, a_2 \rangle$ für alle $x \in V$ folgt $\langle x, a_1 - a_2 \rangle = 0$ für alle $x \in V$. Also insbesondere $\|a_1 - a_2\|^2 = \langle a_1 - a_2, a_1 - a_2 \rangle = 0$, d.h. $a_1 = a_2$. ■

Folgerung 18.47 *Ist V ein n–dimensionaler euklidischer Vektorraum, so ist $V^\star$ isomorph zu V. Genauer gilt: Die Abbildung*

$$\Phi : V \to V^\star;\; a \mapsto \Lambda_a \quad \textit{mit} \quad \Lambda_a : V \to \mathbb{R};\; \Lambda_a(x) = \langle x, a \rangle$$

ist ein $\mathbb{R}$-Vektorraum Isomorphismus. Ist $\{e_1, \dots, e_n\}$ eine Orthonormalbasis von V, so ist $\{\Lambda_{e_1}, \dots, \Lambda_{e_n}\}$ die zugehörige Dualbasis in $V^\star$ (siehe Abschnitt (9.4.1)).

Beweis: Die Abbildung Φ ist linear, da das Skalarprodukt auch im 2. Argument linear ist (hier braucht man "euklidisch"). Die Abbildung Φ ist surjektiv nach dem Satz von Riesz. Um zu zeigen, dass Φ injektiv ist, zeigen wir, dass der Kern von Φ nur aus dem Nullvektor besteht. Sei also $\Phi(a) = \Lambda_a = 0$. Dann gilt $\langle x, a \rangle = 0$ für alle $x \in V$, also $a = 0$. ■

[10] Friedrich Riesz (1880-1956)

18.6 Orthogonale und unitäre Matrizen

Gegeben sei ein Vektorraum V mit Skalarprodukt, eine Orthonormalbasis $B = \{e_1, \dots, e_n\}$ von V und eine Matrix $A = (a_{ij}) \in \mathbf{GL}(n, \mathbb{C})$. Wir definieren eine lineare Abbildung $\Phi_A : V \to V$ durch

$$\Phi_A(e_j) := \sum_{k=1}^{n} a_{kj} e_k, \quad 1 \leq j \leq n.$$

Dann ist $\{e'_j := \Phi_A(e_j) \mid 1 \leq j \leq n\}$ auch eine Basis von V (wieso?).

Frage: Wann ist $B' = \{e'_1, \dots, e'_n\}$ auch wieder eine ONB?

Dazu rechnen wir

$$\langle e'_j, e'_k \rangle = \langle \sum_{r=1}^{n} a_{rj} e_r, \sum_{s=1}^{n} a_{sk} e_s \rangle = \sum_{r,s=1}^{n} a_{rj} \overline{a}_{sk} \delta_{rs} = \sum_{r=1}^{n} a_{rj} \overline{a}_{rk}.$$

Also ist B' genau dann auch eine Orthonormalbasis, wenn gilt:

Für einen euklidischen Vektorraum V:	Für einen unitären Vektorraum V:
$\sum_{r=1}^{n} a_{rj}\, a_{rk} = \delta_{jk},\ j, k = 1, \dots, n$	$\sum_{r=1}^{n} a_{rj}\, \overline{a}_{rk} = \delta_{jk},\ j, k = 1, \dots, n$
$\Longleftrightarrow A^\top A = E_n.$	$\Longleftrightarrow A^\top \overline{A} = E_n.$

Definition 18.48 Eine reelle bzw. komplexe $n \times n$- Matrix A heißt **orthogonal** bzw. **unitär**, falls gilt

$$A^\top A = E_n \quad \text{bzw.} \quad A^\top \overline{A} = E_n.$$

Wir halten fest

Hilfssatz 18.49 *Sind B und B' zwei Orthonormalbasen eines euklidischen (bzw. unitären) Vektorraumes V, so ist die Matrix des Basiswechsels orthogonal (bzw. unitär). Ist umgekehrt $B = \{e_1, \dots, e_n\}$ eine Orthonormalbasis von V und A eine orthogonale (bzw. unitäre) $n \times n$-Matrix, so ist auch $B' := \{\Phi_A(e_1), \dots, \Phi_A(e_n)\}$ eine Orthonormalbasis von V.*

Satz 18.50 (Charakterisierung von orthogonalen Matrizen) *Sei $A \in \mathbb{R}^{n\times n}$. Folgende Aussagen sind äquivalent:*

(a) *A ist eine orthogonale Matrix.*

(b) *A ist regulär und $A^{-1} = A^\top$.*

(c) *Die Spaltenvektoren (bzw. die Zeilenvektoren) von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzgl. des Standardskalarproduktes.*

Analog gilt für unitäre Matrizen

Satz 18.51 (Charakterisierung von unitären Matrizen) *Sei $A \in \mathbb{C}^{n\times n}$. Folgende Aussagen sind äquivalent:*

(a) *A ist eine unitäre Matrix,*

(b) *A ist regulär und $A^{-1} = \overline{A}^\top$;*

(c) *Die Spaltenvektoren (bzw. die Zeilenvektoren) von A bilden eine Orthonormalbasis von $\mathbb{C}^n$ bzgl. des Standardskalarprodukts.*

Beweis: Aus (a), also $A^\top A = E_n$, folgt $(\det A)^2 = 1$. Also existiert A^{-1} und man erhält (b). Dass (a) aus (b) folgt, ist klar. Die Äquivalenz von (a) und (c) sieht man so ein: Sei $B = \{e_1, \ldots, e_n\}$ die Standard-Basis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$. Dann ist B eine Orthonormalbasis bezüglich dem Standard-Skalarprodukt. Die Spalten von A sind dann $\{Ae_1, \ldots, Ae_n\}$. Die Behauptung folgt aus Hilfssatz 18.49 und daraus, dass mit A orthogonal (bzw. unitär) auch $A^\top$ orthogonal (bzw. unitär) ist (und umgekehrt). ■

Folgerung 18.52 (a) *Für eine orthogonale Matrix gilt:* $\det A = \pm 1$.

(b) *Für eine unitäre Matrix gilt:* $|\det A| = 1$.

Beweis: Für eine orthogonale Matrix A gilt

$$1 = \det E_n = \det(A^\top A) = \det A^\top \cdot \det A = (\det A)^2.$$

Für eine unitäre Matrix A gilt

$$1 = \det E_n = \det(A^\top \overline{A}) = \det A \cdot \overline{\det A} = |\det A|^2.$$

■

Bemerkung 18.53 Die Eigenschaft $|\det A| = 1$ ist zwar eine notwendige, aber keine hinreichende Bedingung dafür, dass A eine orthogonale bzw. unitäre Matrix ist. Ein Beispiel ist die Matrix

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

Satz 18.54 (Matrizen-Gruppen) *Für jedes $n \in \mathbb{N}$ bilden die orthogonalen (bzw. unitären) $n \times n$-Matrizen bezüglich der Matrizen-Multiplikation eine Gruppe:*

Die **orthogonale Gruppe**

$$\mathbf{O}(n) := \{A \in \mathbf{GL}(n, \mathbb{R}) \mid A^\top A = E_n\}$$

bzw. die **unitäre Gruppe**

$$\mathbf{U}(n) := \{A \in \mathbf{GL}(n, \mathbb{C}) \mid \overline{A}^\top A = E_n.\}$$

Beweis: Es ist $E_n \in \mathbf{U}(n)$ und mit A auch A^{-1}:

$$\overline{A^{-1}}^\top A^{-1} = (A\overline{A}^\top)^{-1} = (AA^{-1})^{-1} = (E_n)^{-1} = E_n.$$

Aus $A, B \in \mathbf{U}(n)$ folgt auch $A \cdot B \in \mathbf{U}(n)$:

$$(\overline{AB})^\top AB = \overline{B}^\top \overline{A}^\top AB = \overline{B}^\top E_n B = E_n.$$

■

Bemerkung 18.55 Wir erinnern an die Definition der komplexen bzw. reellen speziellen linearen Gruppe:

$$\mathbf{SL}(n, \mathbb{C}) = \{A \in \mathbb{C}^{n\times n} \mid \det A = 1\}, \quad \mathbf{SL}(n, \mathbb{R}) = \mathbf{SL}(n, \mathbb{C}) \cap \mathbf{GL}(n, \mathbb{R}).$$

Die Gruppe

$$\mathbf{SO}(n) = \mathbf{O}(n) \cap \mathbf{SL}(n, \mathbb{R}) \quad \text{bzw.} \quad \mathbf{SU}(n) = \mathbf{U}(n) \cap \mathbf{SL}(n, \mathbb{C})$$

heißt **spezielle orthogonale** bzw. **spezielle unitäre Gruppe**.

Beispiel 18.56

1. $\mathbf{O}(1) = \{r \in \mathbb{R} \mid r^2 = 1\} = \{\pm 1\}$.
2. $\mathbf{U}(1) = \{z \in \mathbb{C} \mid z\overline{z} = 1\} = \{\mathrm{e}^{\mathrm{i}\varphi} \in \mathbb{C} \mid \varphi \in [0, 2\pi]\}$.

3. $\mathbf{O}(2)$: Gegeben sei die orthogonale Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Wegen $a_{11}^2 + a_{21}^2 = 1$ gibt es ein eindeutig bestimmtes $\omega \in [0, 2\pi[$ mit

$$a_{11} = \cos\omega, \quad a_{21} = \sin\omega.$$

Aus $a_{i1}^2 + a_{i2}^2 = 1$ $(i = 1, 2)$ folgt $a_{12}^2 = 1 - a_{11}^2 = \sin^2\omega, a_{22}^2 = 1 - a_{21}^2 = \cos^2\omega$, also

$$a_{12} = \varepsilon \sin\omega, \quad a_{22} = \eta \cos\omega \quad \text{mit} \quad \varepsilon = \pm 1,\ \eta = \pm 1.$$

Wegen $a_{11}a_{12} + a_{21}a_{22} = 0$ ist weiter $(\varepsilon + \eta)\sin\omega \cdot \cos\omega = 0$. Für $\sin\omega\cos\omega \neq 0$ ergibt sich also $\eta = -\varepsilon$:

$$\begin{aligned} A &= \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \quad \text{mit} \quad \det A = +1, & (*) \\ A' &= \begin{pmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{pmatrix} \quad \text{mit} \quad \det A' = -1. & (**) \end{aligned}$$

Ist $\sin\omega\cos\omega = 0$, nimmt also ω einen Wert $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ an, so erhalten wir eine der acht Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

die aber alle bereits in $(*)$ bzw. $(**)$ für $\omega = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ vorkommen.

4. Es ist

$$\mathbf{SU}(2) = \{\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \mid \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1\}.$$

Wir führen noch drei weitere Untergruppen von $\mathbf{GL}(n, \mathbb{R})$ ein. Es sei $\mathbf{B}(n) \subset \mathbf{GL}(n, \mathbb{R})$ die Gruppe der oberen Dreiecksmatrizen mit positiven Diagonalelementen; $\mathbf{A}(n) \subset \mathbf{GL}(n, \mathbb{R})$ bezeichne die (abelsche) Gruppe der Diagonalmatrizen mit positiven Diagonaleinträgen und Determinante 1 und schließlich sei $\mathbf{N}(n) \subset \mathbf{GL}(n, \mathbb{R})$ die Gruppe der oberen Dreieckmatrizen mit Einsen auf der Diagonalen.

Satz 18.57 (Iwasawa–Zerlegung) [11]

[11] Kenkichi Iwasawa (1917-1998)

Es ist

$$\mathbf{GL}(n,\mathbb{R}) = \mathbf{O}(n)\mathbf{B}(n) \quad \textit{und} \quad \mathbf{SL}(n,\mathbb{R}) = \mathbf{SO}(n)\mathbf{A}(n)\mathbf{N}(n).$$

Genauer gilt: Jede Matrix $A \in \mathbf{GL}(n,\mathbb{R})$ kann eindeutig geschrieben werden als Produkt $A = TB$ mit $T \in \mathbf{O}(n)$ und $B \in \mathbf{B}(n)$. Jede Matrix $A \in \mathbf{SL}(n,\mathbb{R})$ kann kann geschrieben werden als Produkt $A = TDN$ wobei $T \in \mathbf{SO}(n)$, $D \in \mathbf{A}(n)$ und $N \in \mathbf{N}(n)$.

Beweis: Sei $A = (a_{ij}) \in \mathbf{GL}(n,\mathbb{R})$, also eine reguläre, reelle $n \times n$-Matrix. Wir betrachten $\mathbb{R}^n$ mit dem Standard-Skalarprodukt $\langle , \rangle$. Die Standardbasis $e_1, \dots, e_n$ ist eine Orthonormalbasis bezüglich $\langle , \rangle$.

Wir orthonormieren jetzt die Vektoren $v_i := Ae_i$ für $i = 1, \dots, n$ (also die Spalten von A) nach Gram-Schmidt und erhalten

$$w_1 := v_1/\|v_1\|, w_2 := (v_2 - \lambda v_1)/\|v_2 - \lambda v_1\| \quad \text{mit } \lambda = \langle v_2, v_1\rangle/\|v_1\|^2 \quad \text{usw.}$$

Wir definieren eine Matrix $S = (s_{ij})$ durch $w_j = \sum_{1 \le i \le j \le n} s_{ij} v_i$. Es ist also $S \in \mathbf{B}(n)$. Insbesondere ist S regulär und $S^{-1} \in \mathbf{B}(n)$. Wir definieren weiter eine orthogonale Matrix $T = (t_{ij}) \in O(n)$ durch $Te_i = w_i$ $(1 \le i \le n)$. Es gilt dann

$$\begin{aligned} t_{jl} &= \langle Te_l, e_j\rangle = \langle w_l, e_j\rangle = \langle \sum_{i \le l} s_{il} v_i, e_j\rangle = \\ &= \sum_{i \le l} s_{il}\langle Ae_i, e_j\rangle = \sum_{i \le l} s_{il} \sum_k a_{ki}\langle e_k, e_j\rangle = \sum_{i \le l} s_{il} a_{ji}. \end{aligned}$$

Also $T = AS$ oder, äquivalent, $A = TS^{-1}$. Die Eindeutigkeit zeigt man so: Wäre $A = T_1S_1 = T_2S_2$ mit $T_i \in \mathbf{O}(n)$ und $S_i \in \mathbf{B}(n)$, so wäre $T_1^{-1}T_2 = S_1S_2^{-1} \in \mathbf{O}(n) \cap \mathbf{B}(n)$. Die Spalten einer oberen Dreiecksmatrix sind aber genau dann orthonormiert, wenn diese Matrix diagonal ist. Da die Diagonaleinträge zudem positiv sind, kann diese Matrix nur die Einheitsmatrix sein. Es folgt $T_1 = T_2$ und $S_1 = S_2$.

Ist $A \in \mathbf{SL}(n,\mathbb{R})$, also $\det A = 1$, so ist $\det T \det S^{-1} = 1$. Da $S^{-1} \in \mathbf{B}(n)$ folgt $\det S^{-1} = \det T = 1$ und wir haben $T \in \mathbf{SO}(n)$. Die Matrix S^{-1} können wir noch weiter zerlegen: $S^{-1} = DN$ mit $D \in \mathbf{A}(n)$ und $N \in \mathbf{N}(n)$. ∎

19 Homomorphismen zwischen Vektorräumen mit Skalarprodukt

In diesem Abschnitt betrachten wir lineare Abbildungen zwischen euklidischen bzw. unitären Vektorräumen, also Vektorräumen mit Skalarprodukt, die diesen Zusatzstrukturen "angepasst" sind.

19.1 Adjungierte Abbildungen

Definition 19.1 Es seien $(V, \langle , \rangle_V)$ und $(W, \langle , \rangle_W)$ zwei Vektorräume mit Skalarprodukt und $\Phi : V \to W$ eine lineare Abbildung. Eine lineare Abbildung $\Phi^* : W \to V$ heißt zu Φ **adjungierte lineare Abbildung**., falls für alle $x \in V$ und alle $y \in W$ gilt:

$$\langle \Phi(x), y \rangle_W = \langle x, \Phi^*(y) \rangle_V.$$

NOTATION: Für eine Matrix $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ definieren wir

$$\boxed{A^* = (a^*_{ij}) := \overline{A}^\top = (\overline{a}_{ji}) \in \mathbb{C}^{n \times m}.}$$

Satz 19.2 (Existenz und Matrixform von Φ^*) *Es seien V, W zwei euklidische bzw. zwei unitäre Vektorräume. Dann gilt:*

(a) *Ist V endlich-dimensional, so existiert zu jeder linearen Abbildung $\Phi : V \to W$ genau eine Adjungierte $\Phi^* : W \to V$.*

(b) *Es seien V und W beide endlich-dimensional. Weiter sei $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V und $C = \{c_1, \ldots, c_m\}$ eine Orthonormalbasis von W. Ist dann A die Abbildungsmatrix von $\Phi : V \to W$ bezüglich B, C, so ist A^* die Abbildungsmatrix von $\Phi^* : W \to V$ bezüglich der Basen C, B. Also $A^* = A^\top \in \mathbb{R}^{n \times m}$ im euklidischen Fall und $A^* = \overline{A}^\top \in \mathbb{C}^{n \times m}$ im unitären Fall.*

Beweis:

(a) Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V. Zu der gegebenen linearen Abbildung $\Phi : V \to W$ definieren wir die Abbildung $\Phi^* : W \to V$ durch

$$\Phi^*(y) := \sum_{r=1}^{n} \langle y, \Phi(e_r) \rangle_W e_r.$$

Die Abbildung Φ^* ist linear und es gilt für alle $x \in V$, $y \in W$

$$\begin{aligned}
\langle x, \Phi^*(y)\rangle_V &= \langle \sum_{k=1}^n x_k e_k \,, \sum_{r=1}^n \langle y, \Phi(e_r)\rangle_W e_r \rangle_V \\
&= \sum_{k=1}^n x_k \sum_{r=1}^n \overline{\langle y, \Phi(e_r)\rangle_W}\, \langle e_k, e_r\rangle_V = \sum_{k=1}^n x_k \langle \Phi(e_k)\,, y\rangle_W \\
&= \langle \Phi \left(\sum_{k=1}^n x_k e_k \right) , y\rangle_W = \langle \Phi(x), y\rangle_W .
\end{aligned}$$

Die Eindeutigkeit gilt ganz allgemein. Man beweist sie so: Sind Φ_1^*, Φ_2^* zwei Adjungierte von Φ, so gilt

$$\forall x \in V \quad \text{und} \quad \forall y \in W : \quad \langle x, \Phi_1^*(y)\rangle = \langle x, \Phi_2^*(y)\rangle .$$

Daraus folgt wegen der Definitheit des Skalarproduktes $\Phi_1^*(y) = \Phi_2^*(y)$ für alle $y \in W$, d.h. $\Phi_1^* = \Phi_2^*$.

(b) Für die $m \times n$–Abbildungsmatrix von Φ bezüglich B, C

$$M_C^B(\Phi) = A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{gilt} \quad \Phi(e_k) = \sum_{s=1}^m a_{sk} c_s, \quad k = 1, \ldots, n.$$

Nach (a) und wegen $\langle c_r, c_s\rangle = \delta_{rs}$ haben wir für die Adjungierte Φ^*

$$\Phi^*(c_r) = \sum_{k=1}^n \langle c_r, \Phi(e_k)\rangle e_k = \sum_{k=1}^n \langle c_r, \sum_{s=1}^m a_{sk} c_s\rangle e_k = \sum_{k=1}^n \overline{a}_{rk} e_k, \quad r = 1, \ldots, m.$$

Also ergibt sich als Abbildungsmatrix von Φ^* bezüglich C, B

$$M_B^C(\Phi^*) = \begin{pmatrix} \overline{a}_{11} & \overline{a}_{21} & \cdots & \overline{a}_{m1} \\ \vdots & \vdots & & \vdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \cdots & \overline{a}_{mn} \end{pmatrix} = \overline{A}^\top = A^* .$$

Im reellen Fall entfällt die komplexe Konjugation, so daß Φ^* die Transponierte $A^\top$ als Abbildungsmatrix besitzt. ■

Für *unendlich-dimensionale* Vektorräume mit Skalarprodukt hat nicht jede lineare Abbildung eine Adjungierte.

Beispiel 19.3 Es sei $W := \ell^2$, der Hilbert–Raum der quadrat-integrierbaren Folgen (siehe Beispiel 18.16) und $V := \mathbb{R}[X] \subset \ell^2$ sei der Untervektorraum der reellen Polynome. Durch das von ℓ^2 induzierte Skalarprodukt wird $\mathbb{R}[X]$ zu einem euklidischen Vektorraum. Wir betrachten die lineare Abbildung

$$\Phi : \mathbb{R}[X] \to \ell^2;\ a = \alpha_0 + \alpha_1 X + \cdots + \alpha_n X^n \mapsto (\alpha_0, \alpha_1, \ldots, \alpha_n, 0, \ldots)$$

und nehmen an, daß die Adjungierte $\Phi^* : \ell^2 \to \mathbb{R}[X]$ existiert. Für $y = (y_j) \in \ell^2 \backslash \mathbb{R}[X]$, $a \in \mathbb{R}[X]$ und $\Phi^*(y) = (y_0', \ldots, y_q', 0, \ldots) \in \mathbb{R}[X]$ hätten wir dann:

$$\sum_{j=0}^{n} \alpha_j y_j = \langle \Phi(a), y \rangle = \langle a, \Phi^*(y) \rangle = \sum_{j=0}^{q} \alpha_j y_j'.$$

Setzen wir für a der Reihe nach die Monome $X^0 = 1, X, X^2$ usw. ein, so ergibt sich $y_j = y_j'$ für $j = 0, \ldots, q$ und $y_j = 0$ für $j > q$ im Widerspruch dazu, dass $y \notin \mathbb{R}[X]$. Die Adjungierte zu Φ existiert also nicht.

Bemerkung 19.4 (Formeln) Falls für die linearen Abbildungen $\Phi : U \to V$ und $\Psi : V \to W$ die Adjungierten $\Phi^* : V \to U$, $\Psi^\star : W \to V$ existieren, so gelten folgende Beziehungen:

$$\begin{aligned}
(\Phi + \Psi)^\star &= \Phi^* + \Psi^\star \\
(\lambda\Phi)^\star &= \overline{\lambda}\Phi^* \quad (\lambda \in \mathbb{C}) \\
(\Psi \circ \Phi)^\star &= \Phi^* \circ \Psi^\star \\
\Phi^{\star\star} &= (\Phi^*)^\star = \Phi \\
\operatorname{Kern} \Phi^\star &= (\operatorname{Bild} \Phi)^\perp.
\end{aligned}$$

Beweis: Wir zeigen exemplarisch die letzte Gleichung:

$$\begin{aligned}
w \in \operatorname{Kern} \Phi^\star &\Leftrightarrow \Phi^\star(w) = 0 \Leftrightarrow \langle \Phi(v), w \rangle = \langle v, \Phi^\star(w) \rangle = 0 \ \text{ für alle } \ v \in V \\
&\Leftrightarrow w \in (\operatorname{Bild} \Phi)^\perp.
\end{aligned}$$

■

19.2 Selbstadjungierte Endomorphismen

Definition 19.5 Ein Endomorphismus $\Phi : V \to V$ eines endlich-dimensionalen Vektorraumes V mit Skalarprodukt heißt **selbstadjungiert**, wenn $\Phi^* = \Phi$ ist, d.h. wenn für alle $x, y \in V$ gilt

$$\langle \Phi(x), y \rangle = \langle x, \Phi(y) \rangle.$$

Beispiel 19.6

1. **Projektionen** sind selbstadjungiert: Es sei $V = U \oplus U^\perp$. Für Vektoren $v, w \in V$ haben wir dann eindeutige Zerlegungen $v = v_1 + v_2$ und

$w = w_1 + w_2$ mit $v_1, w_1 \in U, v_2, w_2 \in U^\perp$. Ist π_U die Orthogonalprojektion auf den Untervektorraum U, so gilt

$$\langle \pi_U(v), w \rangle = \langle v_1, w \rangle = \langle v_1, w_1 \rangle = \langle v, w_1 \rangle = \langle v, \pi_U(w) \rangle.$$

2. Der **Laplace-Operator**. Wir betrachten den Vektorraum $\mathcal{S}(\mathbb{R})$ aller unendlich oft differenzierbaren, reellwertigen Funktionen auf $\mathbb{R}$ mit "beschränktem Träger". Eine C^∞-Funktion f ist also genau dann in $\mathcal{S}(\mathbb{R})$, wenn es ein endliches Intervall $I_f \subset \mathbb{R}$ gibt, so dass f und alle Ableitungen von f außerhalb von I_f gleich Null sind. Wir schreiben kurz $f^{(k)}(\infty) = 0 = f^{(k)}(-\infty), k \in \mathbb{N}_0$. Mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{-\infty}^{\infty} f(t)g(t)dt \quad (f, g \in \mathcal{S}(\mathbb{R}))$$

wird $\mathcal{S}(\mathbb{R})$ zu einem (unendlich-dimensionalen) euklidischen Vektorraum.

Behauptung: Die 2. Ableitung (Laplace-Operator)

$$\Delta : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R}); \quad f \mapsto f''$$

ist selbstadjungiert.

Beweis: Für $f, g \in \mathcal{S}(\mathbb{R})$ folgt mit der Produktregel $(f'g)' = f''g + f'g'$ zunächst

$$\langle \Delta(f), g \rangle = \int_{-\infty}^{\infty} f''(t)g(t)dt = (f'g)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(t)g'(t)dt = -\int_{-\infty}^{\infty} f'(t)g'(t)dt.$$

Da $-f'g' = -(fg')' + fg''$ folgt durch nochmalige Integration

$$-\int_{-\infty}^{\infty} f'(t)g'(t)dt = -(fg')\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f(t)g''(t)dt = \int_{-\infty}^{\infty} f(t)g''(t)dt = \langle f, \Delta(g) \rangle$$

und damit die Behauptung.

Hilfssatz 19.7 (Matrix eines s.a. Endomorphismus) *Ein Endomorphismus Φ ist genau dann selbstadjungiert, wenn die Abbildungsmatrix A von Φ bezüglich einer Orthonormalbasis im euklidischen Fall symmetrisch ($A^\top = A$) und im unitären Fall hermitesch ($\overline{A}^\top = A$) ist.*

Beweis: Gilt $\Phi^* = \Phi$, so folgt die Behauptung einfach daraus, dass die Abbildungsmatrix von $\Phi^\star$ nach Satz 19.2 gleich A^* ist. Ist umgekehrt $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis und im euklidischen Fall $A = A^\top$, so folgt

$$\langle \Phi(e_i), e_j \rangle = \langle \sum_{k=1}^{n} a_{ki} e_k, e_j \rangle = a_{ji} = a_{ij} = \langle e_i, \Phi(e_j) \rangle.$$

Die Behauptung, dass Φ selbstadjungiert ist, folgt dann durch lineare Fortsetzung. Der Beweis im unitären Fall ist analog. ∎

Satz 19.8 *Es sei V ein n–dimensionaler (reeller oder komplexer) Vektorraum mit Skalarprodukt und $\Phi : V \to V$ ein selbstadjungierter Endomorphismus. Dann sind alle Nullstellen des charakteristischen Polynoms von Φ reell. Das charakteristische Polynom von Φ zerfällt also in n Linearfaktoren:*

$$p_\Phi = \pm(X - \lambda_1) \cdots (X - \lambda_n), \quad \lambda_i \in \mathbb{R}.$$

Insbesondere sind alle Eigenwerte eines selbstadjungierten Endomorphismus (bzw. einer symmetrischen oder hermiteschen Matrix) reell.

Beweis: *Unitärer Fall*: Nach dem Fundamentalsatz der Algebra ist

$$p_\Phi = \pm(X - \lambda_1) \cdots (X - \lambda_n), \quad \lambda_i \in \mathbb{C}.$$

Es genügt also zu zeigen, dass alle Eigenwerte von Φ reell sind. Sei also $x \in V$ ein Eigenvektor von Φ zum Eigenwert λ. Dann gilt:

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle \Phi(x), x\rangle = \langle x, \Phi(x)\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle,$$

also, da $x \neq 0$, haben wir $\lambda = \overline{\lambda}$, d.h. $\lambda \in \mathbb{R}$.

Euklidischer Fall: Wir führen den Beweis mittels "Komplexifizierung". Wir wählen eine Orthonormalbasis von V. Die zu Φ gehörige Abbildungsmatrix bezüglich dieser Basis ist reell und symmetrisch und definiert eine Abbildung

$$\tilde{\Phi} : \mathbb{C}^n \to \mathbb{C}^n, \ z \mapsto Az.$$

Da A als reelle, symmetrische Matrix auch komplex, hermitesch ist, ist $\tilde{\Phi}$ nach Hilfssatz 19.7 selbstadjungiert bezüglich dem Standard-Skalarprodukt von $\mathbb{C}^n$. Nach dem oben gezeigten unitären Fall hat $\tilde{\Phi}$ nur reelle Eigenwerte. Mit andern Worten: Das charakteristische Polynom $p_{\tilde{\Phi}}$ hat nur reelle Nullstellen. Nun ist aber $p_\Phi = p_A = p_{\tilde{\Phi}}$, d.h. die Eigenwerte von Φ stimmen mit den Eigenwerten von $\tilde{\Phi}$ überein und sind damit auch reell. ■

Satz 19.9 (Spektralsatz) *Es sei V ein n–dimensionaler Vektorraum mit Skalarprodukt und $\Phi : V \to V$ ein selbstadjungierter Endomorphismus. Dann ist Φ diagonalisierbar. Genauer: Es existiert eine Orthonormalbasis B von V, die aus Eigenvektoren von Φ besteht und die Abbildungsmatrix von Φ bezüglich dieser Orthonormalbasis hat Diagonalform*

$$M_B^B(\Phi) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ die n (reellen) Eigenwerte von Φ sind.

Beweis: Um die Existenz einer Orthonormalbasis aus Eigenvektoren zu beweisen, benutzen wir vollständige Induktion nach $n = \dim V$.

Für $n = 1$ ist nichts zu zeigen.

Sei $n \geq 2$. Nach Satz 19.8 zerfällt das charakteristische Polynom von Φ in Linearfaktoren:

$$p_\Phi = \pm(X - \lambda_1) \cdots (X - \lambda_n), \quad \lambda_i \in \mathbb{R}.$$

Es sei nun a_1 ein Eigenvektor zum Eigenwert λ_1 und $\|a_1\| = 1$. Wir setzen

$$W := [a_1]^\perp = \{x \in V \mid \langle a_1, x\rangle = 0\}.$$

Der Untervektorraum W ist Φ-invariant, d.h. $x \in W \Rightarrow \Phi(x) \in W$, denn wir haben

$$\langle a_1, \Phi(x)\rangle = \langle \Phi(a_1), x\rangle = \langle \lambda_1 a_1, x\rangle = \lambda_1\langle a_1, x\rangle = 0.$$

Die Einschränkung $\Phi|_W$ ist ein selbstadjungierter Endomorphismus von W und $\dim W = \dim[a_1]^\perp = n - 1$. Nach Induktions-Voraussetzung existiert dann eine Orthonormalbasis $\{a_2, \ldots, a_n\}$ von W aus Eigenvektoren von $\Phi|_W$ (und somit auch von Φ). Damit ist $B := \{a_1, a_2, \ldots, a_n\}$ eine Orthonormalbasis von Eigenvektoren von Φ. Dass die Diagonalelemente der Abbildungsmatrix (also die Eigenwerte von Φ) reell sind, folgt aus Satz 19.8. ■

Folgerung 19.10 (Spektralsatz für Matrizen) *Ist A eine reelle, symmetrische (bzw. komplexe, hermitesche) $n \times n$–Matrix, so ist A diagonalisierbar. Genauer gilt: Es gibt eine orthogonale (bzw. unitäre) $n \times n$–Matrix S mit*

$$S^*AS = S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ die n (reellen) Eigenwerte von A sind.

Beweis: Die Matrix A definiert einen selbstadjungerten Endomorphismus $x \mapsto Ax$ von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bezüglich dem Standard-Skalarprodukt. Die Behauptung folgt dann aus Satz 19.9. ■

Beispiel 19.11 Zu der symmetrischen Matrix

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

wird eine orthogonale Matrix S gesucht, so dass $S^\top AS$ eine Diagonalmatrix ist.

Aus dem charakteristischen Polynom $p = -(2+X)^2 X$ von A ergeben sich die Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = -2$. Für die zugehörigen Eigenräume E_{λ_1} und E_{λ_2} erhalten wir:

$$Ax = 0 \Leftrightarrow x = \begin{pmatrix} r \\ 0 \\ r \end{pmatrix}, r \in \mathbb{R}\,; \qquad (A+2E)x = 0 \Leftrightarrow x = \begin{pmatrix} t \\ s \\ -t \end{pmatrix}, s,t \in \mathbb{R}$$

also

$$E_{\lambda_1} = [\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}] \quad \text{und} \quad E_{\lambda_2} = [\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}].$$

Nun sind

$$\{\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}\} \quad \text{und} \quad \{\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\}.$$

Orthonormalbasen von E_{λ_1} bzw. E_{λ_2}. Also ist

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}$$

eine orthogonale Matrix mit

$$S^\top A S = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Bemerkung 19.12 1. Aus dem Spektralsatz folgt, dass die Eigenvektoren eines selbstadjungierten Endomorphismus zu verschiedenen (reellen) Eigenwerten orthogonal sind. Das kann man auch direkt zeigen: Seien $e_i \in E_{\lambda_i}, e_j \in E_{\lambda_j}$ mit $\lambda_i \neq \lambda_j$, dann gilt

$$\lambda_i \langle e_i, e_j \rangle = \langle \lambda_i e_i, e_j \rangle = \langle \Phi(e_i), e_j \rangle = \langle e_i, \Phi(e_j) \rangle = \lambda_j \langle e_i, e_j \rangle,$$

also $\langle e_i, e_j \rangle = 0$.

2. Es seien $\lambda_1, \ldots, \lambda_r$ die r *verschiedenen* Eigenwerte eines selbstadjungierten Endomorphismus $\Phi : V \to V$, und $E_1, \ldots, E_r$ seien die zugehörigen Eigenräume. Dann gilt $E_j \perp E_k$ $(j \neq k)$ nach 1., und mit dem Spektralsatz erhalten wir die orthogonale direkte Zerlegung von V:

$$V = \bigoplus_{j=1}^{r} E_j.$$

Bezeichnet π_{E_j} die orthogonale Projektion von V auf den Untervektorraum E_j;

$$\pi_{E_j} : V \to V; \; x = x_1 + \cdots + x_r \mapsto x_j, \qquad j = 1, \ldots, r$$

mit $x_k \in E_k,\ k = 1, \ldots, r$, so gilt für alle $x \in V$

$$\Phi(x) = \sum_{j=1}^{r} \Phi(x_j) = \sum_{j=1}^{r} \lambda_j x_j = \sum_{j=1}^{r} \lambda_j \pi_{E_j}(x) = \left(\sum_{j=1}^{r} \lambda_j \pi_{E_j} \right)(x),$$

und wir haben

$$\Phi = \lambda_1 \pi_{E_1} + \cdots + \lambda_r \pi_{E_r}.$$

Diese Darstellung heißt **Spektraldarstellung** von Φ. Schließlich gilt noch

$$\pi_{E_1} + \cdots + \pi_{E_r} = \mathrm{id}_V.$$

Die Matrix des Skalarproduktes eines euklidischen Vektorraumes ist (bezüglich irgendeiner Basis) symmetrisch (siehe Abschnitt 18.2). Offen geblieben ist die Frage, wann eine solche Matrix positiv definit ist. Als Anwendung des Spektralsatzes geben wir dafür ein (erstes) Kriterium an. Ein weiteres, praktischeres Kriterium beweisen wir später (siehe Satz 20.22).

Satz 19.13 (Ein Kriterium für "positiv definit") *Eine reelle, symmetrische Matrix A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.*

Beweis: Es sei $A \in \mathbb{R}^{n \times n}$ positiv definit, λ ein Eigenwert und $x \in \mathbb{R}^n$ ein Eigenvektor von A. Dann gilt wegen $x \neq 0$

$$0 < x^\top A x = x^\top(\lambda x) = \lambda x^\top x = \lambda(x_1^2 + \cdots x_n^2),$$

also ist $\lambda > 0$.

Seien umgekehrt $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A und es gelte $\lambda_i > 0$ für alle $1 \leq i \leq k$. Nach Folgerung 19.10 ist A diagonalisierbar, d.h. wir haben insbesondere $\mathbb{R}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Für jeden Vektor $x \in \mathbb{R}^n$ gibt es also eine eindeutige Zerlegung $x = x_1 + \cdots + x_k$ mit $x_i \in E_{\lambda_i}$. Damit erhalten wir

$$x^\top A x = \sum_{i=1}^{k} x_i^\top A x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^{k} x_i^\top A x_j = \sum_{i=1}^{k} \lambda_i \|x_i\|^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{k} \lambda_j \langle x_i, x_j \rangle.$$

Da nach Bemerkung 19.12 Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind, folgt

$$x^\top A x = \sum_{i=1}^{k} \lambda_i \|x_i\|^2.$$

Ist $x \neq 0$, muss mindestens einer der Vektoren x_i auch ungleich Null sein. Da $\lambda_i > 0$ für $i = 1, \ldots, k$ folgt dann insgesamt, dass $x^\top A x > 0$. ■

Beispiel 19.14 1. Gegeben sei die Matrix

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

mit dem charakteristischen Polynom

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

Alle Eigenwerte sind positiv, also ist A positiv definit.

2. Die Matrix

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

mit dem charakteristischen Polynom $p = X^2(X-2)^2$ besitzt den Eigenwert 0, ist also nicht positiv definit.

Folgerung 19.15 ("Wurzel" aus einer positiv definiten Matrix) *Ist $A \in \mathbb{R}^{n\times n}$ eine positiv definite symmetrische Matrix, so existiert eine Matrix $\sqrt{A} \in \mathbb{R}^{n\times n}$ mit $\sqrt{A} \cdot \sqrt{A} = A$.*

Beweis: Nach Satz 19.10 und Satz 19.13 existiert eine orthogonale Matrix S und eine Diagonalmatrix D, so dass

$$S^{-1}AS = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i > 0$ für alle i. Wir können also definieren

$$\sqrt{D} := \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \quad \text{und} \quad \sqrt{A} := S\sqrt{D}S^{-1}.$$

Dann haben wir

$$\sqrt{A} \cdot \sqrt{A} = S\sqrt{D}S^{-1}S\sqrt{D}S^{-1} = SDS^{-1} = A.$$

■

19.3 Lineare Isometrien

Wir beginnen mit einer Definition für allgemeine metrische Räume.

Definition 19.16 Es seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt **Isometrie**, falls f Abstände erhält, falls also für alle $p, q \in X$ gilt

$$d_Y(f(p), f(q)) = d_X(p, q).$$

Wir betrachten jetzt wieder zwei Vektorräume V, W mit Skalarprodukt. Wir interessieren uns hier für "lineare Isometrien", also lineare Abbildungen zwischen V und W, die die Länge von Vektoren nicht ändern. Im wichtigen Spezialfall $V = W$ bilden die linearen Isometrien von V eine Untergruppe der Automorphismengruppe von V.

Definition 19.17 Es seien V, W entweder zwei euklidische oder zwei unitäre Vektorräume. Eine lineare Abbildung $\Phi : V \to W$ heißt **lineare Isometrie**, wenn für alle $x \in V$ gilt

$$\|\Phi(x)\|_W = \|x\|_V,$$

d.h. wenn Φ die Länge jedes Vektors $x \in V$ invariant läßt.

Bemerkung 19.18

1. Eine allgemeine Isometrie f und damit insbesondere eine lineare Isometrie Φ ist injektiv:

$$f(p) = f(q) \Longrightarrow d_X(p, q) = d_Y(f(p), f(q)) = 0,$$

also wegen der Positivität einer Metrik $p = q$.

2. Ein Vektorraum mit Skalarprodukt ist ein normierter Raum und deshalb auch ein metrischer Raum (siehe Abschnitt 18.3). Die Metrik ist gegeben durch

$$d(u, v) := \|u - v\| = \sqrt{\langle u - v, u - v\rangle}.$$

Eine lineare Isometrie ist also auch eine Isometrie im allgemeinen Sinn (für die durch die Skalarprodukte induzierten Metriken).

Beispiel 19.19 1. **Drehungen**. Wir vesehen $\mathbb{R}^2$ mit dem Standard-Skalarprodukt und betrachten Polarkoordinaten

$$\mathbb{R}^2 \ni x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}.$$

Dann haben wir (mit den Additionstheoremen für Sinus und Cosinus)

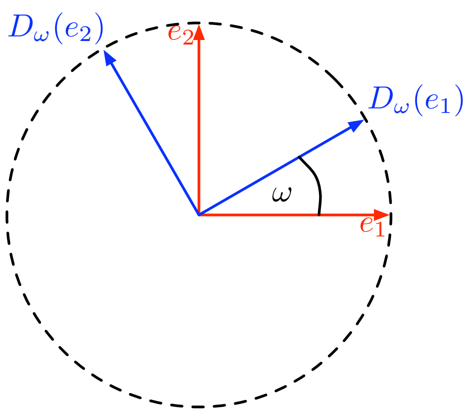
$$D_\omega(x) := \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} r\cos(\alpha+\omega) \\ r\sin(\alpha+\omega) \end{pmatrix}.$$

Für die lineare Abbildung D_ω gilt

$$\|D_\omega(x)\|^2 = r^2(\cos^2(\alpha+\omega) + \sin^2(\alpha+\omega)) = r^2 = \|x\|^2,$$

also ist D_ω eine lineare Isometrie von $\mathbb{R}^2$.

Geometrisch ist D_ω eine Drehung (im Uhrzeigersinn) um den Ursprung von $\mathbb{R}^2$ mit dem Drehwinkel ω.



Entsprechend ist $D_{-\omega}$ eine Drehung (im Gegen-Uhrzeigersinn) um den Ursprung von $\mathbb{R}^2$ mit dem Drehwinkel ω. Es ist $\det D_\omega = \det D_{-\omega} = 1$. Als Spezialfälle haben wir für $\omega = 0, \pi$:

$$D_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad D_\pi = D_{-\pi} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

also die Identität und die Punktspiegelung am Ursprung $(0,0)$.

2. **Verallgemeinerung**: Sei $A \in \mathbb{R}^{n\times n}$ eine orthogonale Matrix, also $A^\top A = E$, dann ist die lineare Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto Ax$ eine lineare Isometrie:

$$\|Ax\|^2 = (Ax)^\top Ax = x^\top A^\top Ax = x^\top x = \|x\|^2.$$

3. **Translationen** sind <u>nicht</u>-lineare Isometrien. Es sei V ein Vektorraum mit Skalarprodukt. Für $a \in V$ ist

$$T_a : V \to V; \quad x \mapsto x + a.$$

die Translation um den Vektor a. Dann ist T_a (für $a \neq 0$) keine lineare Abbildung, aber eine Isometrie im allgemeinen Sinn:

$$d(T_a(x), T_a(y)) = \|(x+a) - (y+a)\| = \|x - y\| = d(x,y).$$

Satz 19.20 *Es seien V, W zwei unitäre oder zwei euklidische Vektorräume. Folgende Aussagen für eine lineare Abbildung $\Phi : V \to W$ sind äquivalent:*

(a) *Φ ist eine lineare Isometrie.*

(b) *$\forall x, y \in V$ gilt $\langle \Phi(x), \Phi(y) \rangle_W = \langle x, y \rangle_V$.*

Beweis: (b) $\Rightarrow$ (a): Setzen wir $y = x$, so folgt aus (b)

$$\|\Phi(x)\|_W^2 = \langle \Phi(x), \Phi(x) \rangle_W = \langle x, x \rangle_V = \|x\|_V^2 \quad \text{für alle } \ x \in V.$$

(a) $\Rightarrow$ (b):

1. FALL: *V, W euklidische Vektorräume.*

Für $x, y \in V$ gilt $\|\Phi(x+y)\|^2 = \|x+y\|^2$, also

$$\begin{aligned} \langle \Phi(x+y), \Phi(x+y) \rangle &= \langle x+y, x+y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &= \langle \Phi(x), \Phi(x) \rangle + 2\langle \Phi(x), \Phi(y) \rangle + \langle \Phi(x), \Phi(x) \rangle. \end{aligned}$$

Wegen $\langle \Phi(x), \Phi(x) \rangle = \langle x, x \rangle$, $\langle \Phi(y), \Phi(y) \rangle = \langle y, y \rangle$ folgt $\langle \Phi(x), \Phi(y) \rangle = \langle x, y \rangle$.

2.Fall: *V, W unitäre Vektorräume.* Die gleiche Rechnung wie oben zeigt:

$$\langle \Phi(x), \Phi(y) \rangle + \overline{\langle \Phi(x), \Phi(y) \rangle} = \langle x, y \rangle + \overline{\langle x, y \rangle}.$$

Analog folgt aus $\|\Phi(x - \mathrm{i}y)\|^2 = \|x - \mathrm{i}y\|^2$, dass

$$\langle \Phi(x), \Phi(y) \rangle - \overline{\langle \Phi(x), \Phi(y) \rangle} = \langle x, y \rangle - \overline{\langle x, y \rangle}.$$

Durch Kombination ergibt sich (b) auch im unitären Fall. ■

Folgerung 19.21 *Eine lineare Isometrie zwischen euklidischen Vektorräumen ist "winkeltreu":*

$$\omega(\Phi(x), \Phi(y)) = \omega(x, y).$$

Beweis: Für $x, y \neq 0$ ist

$$\cos \omega(\Phi(x), \Phi(y)) = \frac{\langle \Phi(x), \Phi(y) \rangle}{\|\Phi(x)\| \cdot \|\Phi(y)\|} = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \cos \omega(x, y).$$ ■

Wir erinnern daran, dass für eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen mit Skalarprodukt stets die Adjungierte existiert. Damit ergibt sich folgende Charakterisierung von linearen Isometrien.

Satz 19.22 (Lineare Isometrien und Adjungierte) (a) *Ein Endomorphismus $\Phi : V \to V$ eines endlich-dimensionalen Vektorraums mit Skalarprodukt ist genau dann eine lineare Isometrie, wenn gilt*

$$\Phi^* \circ \Phi = id_V.$$

(b) *Ein Endomorphismus Φ eines unitären (bzw. euklidischen) Vektorraums V ist genau dann eine lineare Isometrie, wenn für die Abbildungsmatrix A von Φ bezüglich einer Orthonormalbasis von V gilt:*

$$A^* A = E,$$

d.h genau dann, wenn A unitär (bzw. orthogonal) ist.

Beweis: (a) Ist $\Phi : V \to V$ eine lineare Isometrie, so gilt nach Satz 19.20

$$\langle x, (\Phi^* \circ \Phi)(y)\rangle = \langle \Phi(x), \Phi(y)\rangle = \langle x, y\rangle$$

für alle x, $y \in V$. Also gilt für alle $y \in V$: $(\Phi^* \circ \Phi)(y) = y$ und somit $\Phi^* \circ \Phi = \mathrm{id}_V$. Gilt umgekehrt $\Phi^* \circ \Phi = \mathrm{id}_V$, so folgt für alle x, $y \in V$:

$$\langle x, y\rangle = \langle x, (\Phi^* \circ \Phi)(y)\rangle = \langle \Phi(x), \Phi(y)\rangle.$$

Nach Satz 19.20 ist Φ also eine lineare Isometrie.

(b) Die Abbildungsmatrix von Φ^* bezüglich einer Orthonormalbasis ist $A^\star$. Nach (a) ist Φ also genau dann eine lineare Isometrie, wenn

$$A^\star A = E$$

gilt, d.h wenn $\overline{A}^\top A = E$ im unitären Fall bzw. $A^\top A = E$ im euklidischen Fall. ∎

Folgerung 19.23 (a) *Ein Endomorphismus Φ eines n–dimensionalen Vektorraumes V mit Skalarprodukt ist genau dann eine lineare Isometrie von V, wenn $\Phi^* = \Phi^{-1}$ gilt. Insbesondere ist Φ^* auch eine lineare Isometrie von V.*

(b) *Die linearen Isometrien eines n-dimensionalen euklidischen (bzw. unitären) Vektorraumes V bilden bezüglich der Verkettung von Abbildungen eine Gruppe.*

Beweis: Da $\dim V = n < \infty$ und weil Φ als lineare Isometrie injektiv ist, ist Φ auch surjektiv. Es existiert also die inverse Abbildung $\Phi^{-1} : V \to V$, und nach (19.22) gilt

$$\Phi^{-1} = \mathrm{id}_V \circ \Phi^{-1} = \Phi^* \circ (\Phi \circ \Phi^{-1}) = \Phi^*.$$

(b) Die Identität ist eine lineare Isometrie, nach Folgerung 19.23 (a) ist auch Φ^{-1} eine Isometrie. Schließlich folgt etwa mittels Satz 19.22, dass auch die Verkettung von zwei Isometrien wieder eine Isometrie ist. ∎

Beispiel 19.24 (Alle linearen Isometrien der euklidischen Ebene.) Wir betrachten den euklidischen Vektorraum $\mathbb{R}^2$ versehen mit dem Standardskalarprodukt. Die Standardbasis $B = \{e_1 = (1,0), e_2 = (0,1)\}$ ist dann eine Orthonormalbasis. Für eine lineare Isometrie Φ ist die Abbildungsmatrix bezüglich B nach Satz 19.22 orthogonal. Also $M_B^B(\Phi) \in O(2)$. Die Gestalt dieser Matrizen haben wir in Beispiel 18.56 bestimmt. Für $[0, 2\pi]$ haben wir:

$$M_B^B(\Phi) = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \quad \text{oder} \quad M_B^B(\Phi) = \begin{pmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{pmatrix}.$$

Umgekehrt ist durch diese Abbildungsmatrizen für beliebiges ω eine Isometrie von $\mathbb{R}^2$ gegeben.

Es ist

$$\begin{pmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{pmatrix} = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Eine lineare Isometrie von $\mathbb{R}^2$ (versehen mit dem Standard-Skalarprodukt) ist also entweder eine Drehung $D_\omega, \omega \in [0, 2\pi]$, oder das Produkt einer solchen Drehung mit einer Spiegelung (an der x-Achse).

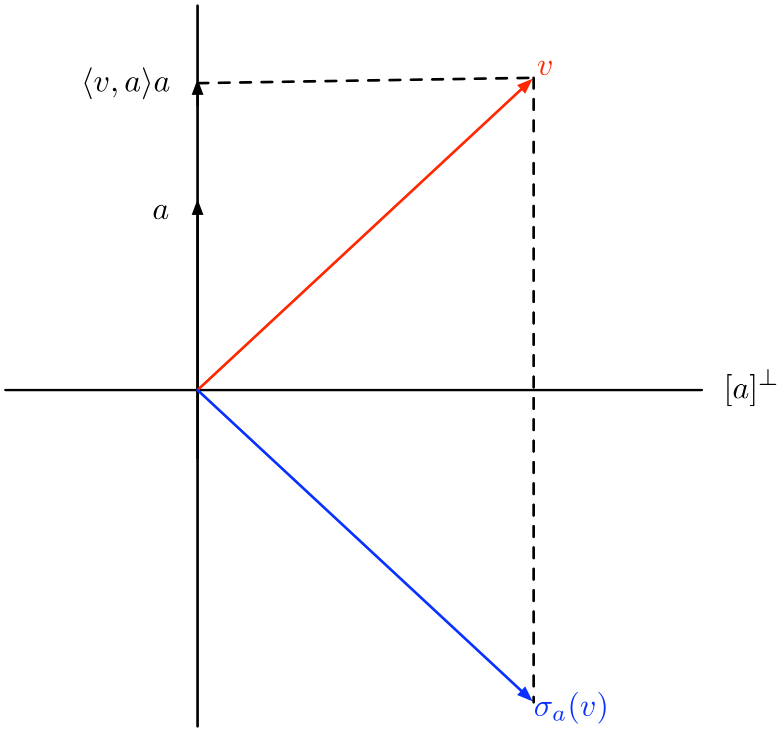
Beispiel 19.25 (Spiegelungen) Es sei $(V, \langle,\rangle)$ ein endlich-dimensionaler Vektorraum mit Skalarprodukt und $a \in V$ ein Einheitsvektor, also $\|a\| = 1$. Wir definieren die Abbildung

$$\sigma_a : V \to V,\ \sigma_a(v) := v - 2\langle v, a\rangle a,$$

von V auf sich selbst. Da ein Skalarprodukt im ersten Argument linear ist, ist σ_a linear. Weiter haben für alle $v, w \in V$

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

d.h. σ_a ist eine lineare Isometrie.



Die wesentlichen **Eigenschaften** von σ_a sind:

- $\sigma_a^2 = \mathrm{id}_V$.
- Für $v \perp a$ ist $\sigma_a(v) = v$, d.h. auf $[a]^\perp$, dem Orthogonalkomplement von a, ist σ_a die Identität.
- Sei $v \in V$ beliebig. Da $V = [a] \oplus [a]^\perp$ können wir v eindeutig zerlegen: $v = \lambda a + w$ mit $w \in [a]^\perp$. Es ist dann $\sigma_a(v) = -\lambda a + w$, also $\sigma_a = -\mathrm{id}\,|_{[a]} + \mathrm{id}\,|_{[a]^\perp}$. Die Abbildungsmatrix von σ_a bezüglich einer Orthonormalbasis der Form $B := \{a, b_2, \dots, b_n\}$ ist also

$$M_B^B(\sigma_a) = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

- Ist V euklidisch und sind $v, w \in V$ zwei Vektoren *gleicher Länge*, so gibt es ein $a \in V$, für das die Spiegelung σ_a die Vektoren v und w vertauscht, d.h. $\sigma_a(v) = w$ und $\sigma_a(w) = v$.

Um diese letzte Eigenschaft einzusehen, setzen wir $a := v - w/\|v - w\|$. Weil v und w gleiche Länge haben gilt $v - w \perp v + w$. Damit erhalten wir

$$\sigma_a(v) = \sigma_a(\frac{1}{2}(v+w) + \frac{1}{2}(v-w)) = \frac{1}{2}(v+w) - \frac{1}{2}(v-w) = w.$$

Die lineare Isometrie σ_a $(a \in V, \|a\| = 1)$ heißt **Spiegelung** von V (an der Spiegel-Hyperebene $[a]^\perp$).

Satz 19.26 *Jede lineare Isometrie Φ eines n-dimensionalen euklidischen Vektorraumes V ist ein Produkt von höchstens n Spiegelungen:*

$$\Phi = \sigma_{a_1} \circ \sigma_{a_1} \circ \cdots \circ \sigma_{a_n}.$$

Beweis: Sei $\{e_1, e_2 \ldots, e_n\}$ eine Orthonormalbasis von V. Wir setzen

$$a_1 := (\Phi(e_1) - e_1)/\|\Phi(e_1) - e_1\|.$$

Wie wir in Beispiel 19.25 gesehen haben, gilt dann $\sigma_{a_1}(\Phi(e_1)) = e_1$. Somit ist $\Phi_1 := \sigma_{a_1} \circ \Phi$ eine lineare Isometrie von V mit $\Phi_1(e_1) = e_1$.
Sei jetzt $a_2 := (\Phi_1(e_2) - e_2)/\|\Phi_1(e_2) - e_2\|$. Dann ist

$$(\sigma_{a_2} \circ \Phi_1)(e_2) = \sigma_{a_2}(\Phi_1(e_2)) = e_2.$$

Weiter ist $a_2 \perp e_1$, denn $e_2 \perp e_1$ und $\Phi_1(e_2) \perp \Phi_1(e_1) = e_1$. Damit haben wir auch

$$(\sigma_{a_2} \circ \Phi_1)(e_1) = \sigma_{a_2}(e_1) = e_1.$$

Nach n solchen Schritten erhalten wir n Spiegelungen $\sigma_{a_1}, \sigma_{a_2}, \cdots, \sigma_{a_n}$ mit

$$(\sigma_{a_n} \circ \sigma_{a_{n-1}} \circ \cdots \circ \sigma_{a_1} \circ \Phi)(e_i) = e_i$$

für alle $1 \leq i \leq n$. Also ist $(\sigma_{a_n} \circ \sigma_{a_{n-1}} \circ \cdots \circ \sigma_{a_1} \circ \Phi)$ die Identität von V und wir haben

$$\Phi = \sigma_{a_1}^{-1} \circ \sigma_{a_2}^{-1} \circ \cdots \circ \sigma_{a_n}^{-1} \circ \mathrm{id}_V = \sigma_{a_1} \circ \sigma_{a_2} \circ \cdots \circ \sigma_{a_n}.$$

■

Zum Schluss geben wir noch zwei notwendige (aber nicht hinreichende) Eigenschaften an, die lineare Isometrien eines endlich-dimensionalen Vektorraumes haben müssen.

Satz 19.27 *Es sei Φ eine lineare Isometrie von V.*

(a) *Es gilt $|\det \Phi| = 1$.*

(b) *Ist $\lambda \in \mathbb{C}$ ein Eigenwert von Φ, so gilt $|\lambda| = 1$.*

Beweis: (a) Nach (19.22) ergibt sich $\det \Phi^* \det \Phi = 1$, und wegen $\det \Phi^* = \overline{\det \Phi}$ wird $|\det \Phi|^2 = \overline{\det \Phi} \det \Phi = 1$.

(b) Ist λ ein Eigenwert von Φ, so existiert ein Eigenvektor x von Φ, es gilt also $\Phi(x) = \lambda x$, $x \neq 0$, und (19.17) ergibt direkt

$$0 < \|x\| = \|\Phi(x)\| = |\lambda| \|x\|,$$

also $|\lambda| = 1$ wegen $\|x\| \neq 0$. ■

Für eine lineare Isometrie Φ eines *euklidischen* Vektorraums gilt also $\det \Phi = \pm 1$ und jeder vorhandene Eigenwert von Φ ist $+1$ oder -1. Man beachte aber, dass es lineare Isometrien etwa von $\mathbb{R}^{2n}$ (mit dem Standardskalarprodukt) ohne Eigenwerte gibt. Ein Beispiel ist die Drehmatrix

$$\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix},$$

deren charakteristisches Polynom $p = X^2 - \sqrt{2}X + 1$ keine reellen Nullstellen hat.

19.4 Normalformen von linearen Isometrien

Unser Ziel ist es, zu einer gegebenen linearen Isometrie eine Basis zu finden, so dass die Abbildungsmatrix bezüglich dieser Basis eine "möglichst einfache" Form hat.

19.4.1 Lineare Isometrien von unitären Vektorräumen

Satz 19.28 (Diagonalisierbarkeit) *Sei V ein endlich-dimensionaler unitärer Vektorraum und $\Phi : V \to V$ eine lineare Isometrie. Dann ist Φ diagonalisierbar. Genauer: Es existiert eine Orthonormalbasis von V, die aus Eigenvektoren von Φ besteht.*

Beweis: Wir argumentieren mit vollständiger Induktion nach $n = \dim V$. Für $n = 1$ ist nichts zu zeigen. Sei also $n \geq 2$. Nach dem *Fundamentalsatz der Algebra* (siehe 13.8) zerfällt das charakteristische Polynom von Φ in Linearfaktoren:

$$p_\Phi = \pm (X - \lambda_1) \cdots (X - \lambda_n), \quad \lambda_i \in \mathbb{C}.$$

Da λ_i Eigenwert ist, gilt nach Satz 19.27 dass $|\lambda_i| = 1$ für $i = 1, \dots, n$.

Sei a_1 Eigenvektor zu λ_1 und $\|a_1\| = 1$; weiter sei $W := [a_1]^\perp$. Dann ist W invariant unter Φ: Für $x \in W$ ist auch $\Phi(x) \in W$, denn

$$0 = \langle a_1, x \rangle = \langle \Phi(a_1), \Phi(x) \rangle = \lambda_1 \langle a_1, \Phi(x) \rangle$$

also, da $|\lambda_i| = 1$, $\Phi(x) \in [a_1]^\perp = W$. Φ induziert also eine lineare Isometrie $\Phi_{|W} : W \to W$. Da $\dim W = n - 1$, existiert nach Induktions-Voraussetzung eine

Orthonormalbasis aus Eigenvektoren von $\Phi_{|W}$ (und damit auch von Φ). Zusammen mit a_1 erhält man die gewünschte Orthonormalbasis von V. ■

Bemerkung 19.29 Der Beweis zeigt, dass der Satz auch für eine lineare Isometrie Φ eines *euklidischen* Vektorraumes gilt, falls das charakteristische Polynom von Φ *über* $\mathbb{R}$ in Linearfaktoren zerfällt.

Folgerung 19.30 (Unitäre Normalform ist diagonal) *Ist A eine unitäre $n\times n$–Matrix, so existiert eine unitäre Matrix T mit*

$$T^\star AT = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

dabei sind die λ_k die Eigenwerte von A und es gilt $|\lambda_k| = 1 \quad (k = 1, \ldots, n)$.

Beweis: Sei $\mathbb{C}^n$ versehen mit dem Standardskalarprodukt. Die Standardbasis B ist eine Orthonormalbasis und wir fassen A als die Abbildungsmatrix einer linearen Isometrie Φ von $\mathbb{C}^n$ bezüglich B auf. Nach dem obigen Satz existiert eine Orthonormalbasis C aus Eigenvektoren von Φ. D.h. die Abbildungsmatrix von Φ bezüglich C hat Diagonalgestalt

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ alle Eigenwerte von Φ sind. Mit der Transformationsmatrix T des Basiswechsels von C nach B gilt dann $D = T^{-1}AT$ und da B, C Orthonormalbasen sind, ist T unitär, also $T^{-1} = T^*$. ■

Die Diagonalmatrix in Satz 19.30 heißt die **(unitäre) Normalform der linearen Isometrie** Φ. Sie ist bis auf die Reihenfolge der Eigenwerte λ_k durch Φ eindeutig festgelegt. Damit ist auch eine **Klassifikation der linearen Isometrien** eines unitären Vektorraumes gegeben: Zur selben Klasse gehören alle linearen Isometrien mit derselben Normalform (einschließlich permutierter Diagonalelemente).

19.4.2 Lineare Isometrien von euklidischen Vektorräumen

Wir betrachten jetzt lineare Isometrien Φ von euklidischen Vektorräumen. Hier ist die Situation wesentlich komplizierter, da das charakteristische Polynom von Φ im Allgemeinen nicht in Linearfaktoren zerfällt.

Satz 19.31 (euklidische Normalform: Dreh-Kästchen) *Ist Φ eine lineare Isometrie eines n–dimensionalen euklidischen Vektorraumes V, so existiert eine Orthonormalbasis von V, bezüglich der die Abbildungsmatrix von Φ die folgende Form hat:*

$$\begin{pmatrix} \boxed{\begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix}} & & & & 0 \\ & \boxed{\begin{matrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{matrix}} & & & \\ & & \boxed{\begin{matrix} \cos\omega_1 & -\sin\omega_1 \\ \sin\omega_1 & \cos\omega_1 \end{matrix}} & & \\ & & & \ddots & \\ 0 & & & & \boxed{\begin{matrix} \cos\omega_r & -\sin\omega_r \\ \sin\omega_r & \cos\omega_r \end{matrix}} \end{pmatrix} \qquad (\star)$$

mit $0 < \omega_j < \pi;\ j = 1, \dots, r$. Die Matrix $(\star)$ ist bis auf die Reihenfolge der Kästchen durch Φ eindeutig bestimmt und heißt die **euklidische Normalform der linearen Isometrie** Φ.

Bemerkung 19.32
1. Der Kästchenform $(\star)$ entspricht eine Zerlegung des Vektorraumes V:
$$V = E_1 \oplus E_{-1} \oplus W_1 \oplus \cdots \oplus W_r$$
dabei sind E_1, E_{-1} die Eigenräume zum Eigenwert 1, -1, (die auch fehlen können) und die W_j sind 2–dimensionale Φ–invariante Untervektorräume (die ebenfalls fehlen können).

2. Eine lineare Isometrie eines euklidische Vektorraumes ist also durch die geometrischen Vielfachheiten p und q der Eigenwerte $+1$ und -1 und die r Winkel ω_j charakterisiert (wobei $p + q + 2r = \dim V$ ist).

Der folgende Beweis liefert auch gleich ein praktisches Verfahren zur Aufstellung der Normalform.

BEWEIS: Wir betrachten die *Hilfsabbildung*
$$\Psi := \Phi + \Phi^* = \Phi + \Phi^{-1}.$$
Ψ ist selbstadjungiert, also existiert nach dem Spektralsatz 19.9 in V eine Orthonormalbasis aus Eigenvektoren von Ψ. Somit gilt
$$V = E_2 \oplus E_{-2} \oplus E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$$

mit $\lambda_i \neq \pm 2$, und alle Eigenräume von Ψ sind paarweise orthogonal. Dabei kommt der Eigenraum E_2 zum Eigenwert 2 genau dann vor, wenn Φ den Eigenwert 1 hat, und E_{-2} kommt genau dann vor, wenn Φ den Eigenwert -1 hat. Wir zeigen dies für den ersten Fall (der zweite folgt analog):

Aus $\Phi(x) = x$ folgt $\Phi^{-1}(x) = x$ und damit $\Psi(x) = 2x$. Ist umgekehrt $\Psi(x) = 2x$, also $(\Phi + \Phi^*)(x) = 2x$, so haben wir $2\langle x, x\rangle = \langle \Phi(x), x\rangle + \langle \Phi^*(x), x\rangle = 2\langle \Phi(x), x\rangle$ und somit $\langle \Phi(x) - x, \Phi(x) - x\rangle = 0$. Also gilt $\Phi(x) = x$.

Alle Eigenräume E_λ von Ψ sind Φ-invariant, denn für $x \in E_\lambda$ folgt

$$\Psi(\Phi(x)) = (\Phi + \Phi^{-1})(\Phi(x)) = \Phi \circ (\Phi + \Phi^{-1})(x) = \Phi(\Psi(x)) = \Phi(\lambda x) = \lambda \Phi(x),$$

also $\Phi(x) \in E_\lambda$. Somit folgt $\Phi(E_\lambda) \subset E_\lambda$. Da Φ eine lineare Isometrie ist, gilt sogar $\Phi(E_\lambda) = E_\lambda$.

Wir können also alle Eigenräume von Ψ getrennt betrachten und dort jeweils geeignete Orthonormalbasen suchen.

1. FALL: $\lambda = \pm 2$. Wie wir gesehen haben, ist $E_2(\Psi)$ gerade der Eigenraum von Φ zum Eigenwert 1 und $E_{-2}(\Psi)$ ist der Eigenraum von Φ zum Eigenwert -1. In jedem dieser Eigenräume können wir beliebige Orthonormalbasen wählen.

2. FALL: $\lambda \neq \pm 2$. Wir behaupten, dass in diesem Fall $\dim E_\lambda$ gerade ist und dass es $l = \frac{1}{2} \dim E_\lambda$ paarweise orthogonale, zweidimensionale, Φ-invariante Untervektorräume $U_1, \ldots, U_l$ gibt mit $E_\lambda = U_1 \oplus \cdots \oplus U_l$.

Beweis: Wäre $\dim E_\lambda$ ungerade, so hätte das charakteristische (reelle) Polynom von $\Phi|_{E_\lambda}$ eine (reelle) Nullstelle. Diese wäre Eigenwert von Φ, also ± 1, und somit gäbe es in E_λ einen Eigenvektor von Ψ zum Eigenwert ± 2. Dies ist ein Widerspruch.

Sei nun $x_1 \in E_\lambda$, $||x_1|| = 1$ beliebig gewählt. Dann sind x_1 und $\Phi(x_1)$ linear unabhängig, da sonst $\Phi(x_1) = \pm x_1$ und somit $\Psi(x_1) = \pm 2x_1$ folgen würde. Der Untervektorraum $U_1 := [x_1, \Phi(x_1)] \subset E_\lambda$ ist also zweidimensional.

U_1 ist Φ-invariant: Aus $\Phi \circ (\Phi + \Phi^{-1})(x_1) = \Phi(\Psi(x_1)) = \Phi(\lambda x_1) = \lambda \Phi(x_1)$ folgt nämlich $\Phi^2(x_1) + x_1 = \lambda \Phi(x_1)$, also $\Phi^2(x_1) \in U_1$. Somit gilt $\Phi(U_1) \subset U_1$ und, weil Φ lineare Isometrie ist, sogar $\Phi(U_1) = U_1$.

$U_1^\perp$ ist ebenfalls Φ-invariant: Seien $x \in U_1^\perp$ und $y \in U_1$ beliebig gewählt. Dann gilt $\Phi^{-1}(y) \in U_1$, und somit $\langle \Phi(x), y\rangle = \langle x, \Phi^*(y)\rangle = \langle x, \Phi^{-1}(y)\rangle = 0$, also $\Phi(x) \in U_1^\perp$.

Wegen $\Phi(E_\lambda) = E_\lambda$ und $\Phi(U_1^\perp) = U_1^\perp$ gilt daher auch $\Phi(U_1^\perp \cap E_\lambda) = U_1^\perp \cap E_\lambda$.

Ist $U_1^\perp \cap E_\lambda \neq \{0\}$, so wählen wir einen beliebigen normierten Vektor x_2 aus dieser Menge und bilden den Untervektorraum $U_2 := [x_2, \Phi(x_2)]$. Dieser ist wieder zweidimensional, Φ-invariant und wegen $U_2 \subset U_1^\perp \cap E_\lambda$ orthogonal zu U_1. Danach wählen wir x_3 aus $(U_1 \oplus U_2)^\perp \cap E_\lambda$, $||x_3|| = 1$, usw. Das Verfahren bricht nach $l = \frac{1}{2} \dim E_\lambda$ Schritten ab, und wir erhalten eine Darstellung $E_\lambda = U_1 \oplus U_2 \oplus \cdots \oplus U_l$. Damit ist die Beauptung des 2. Falles bewiesen.

In jedem dieser l Unterräume $U = [x, \Phi(x)] \subset E_\lambda$ konstruieren wir schließlich von der Basis $\{x, \Phi(x)\}$ ausgehend eine Orthonormalbasis $\{x, y\}$ mit

$$y := \frac{\Phi(x) - \langle \Phi(x), x\rangle x}{||\Phi(x) - \langle \Phi(x), x\rangle x||}.$$

Wir stellen den Vektor y noch in anderer Form dar: Zunächst folgt aus $\Psi(x) = \lambda x$ die Gleichung $\langle \Phi(x), x\rangle = \frac{\lambda}{2}\langle x, x\rangle = \frac{\lambda}{2}$. Ist $\omega \in (0, \pi)$ der Winkel zwischen x und $\Phi(x)$, so gilt

$$\cos\omega = \frac{\lambda}{2}, \quad \sin\omega = \sqrt{1 - \frac{\lambda^2}{4}} \neq 0.$$

Daraus folgt

$$y = \frac{1}{\sin\omega}(\Phi(x) - \cos(\omega)x)$$

und somit

$$\Phi(x) = \cos(\omega)x + \sin(\omega)y.$$

Wegen $\Phi^2(x) = \lambda\Phi(x) - x = 2\cos(\omega)\Phi(x) - x$ gilt

$$\begin{aligned}
\Phi(y) &= \tfrac{1}{\sin\omega}(\Phi^2(x) - \cos(\omega)\Phi(x)) \\
&= \tfrac{1}{\sin\omega}(2\cos(\omega)\Phi(x) - x - \cos(\omega)\Phi(x)) \\
&= \tfrac{1}{\sin\omega}(\cos^2(\omega)x + \cos(\omega)\sin(\omega)y - x) \\
&= -\sin(\omega)x + \cos(\omega)y.
\end{aligned}$$

Bezüglich dieser Basis gilt also

$$A_\Phi|_U = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix}$$

und damit

$$A_\Phi|_{E_\lambda} = \begin{pmatrix} \boxed{\begin{matrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{matrix}} \end{pmatrix}$$

mit $l(\lambda) = \frac{1}{2}\dim E_\lambda$ Kästchen.

Insgesamt erhalten wir so die gesuchte Normalform A_Φ. ∎

Bemerkung 19.33 1. In der Literatur wird manchmal auch eine andere Matrix als Normalform bezeichnet. Sie unterscheidet sich von der obigen Form dadurch, dass die Zweierkästchen die transponierte Form

$$\boxed{\begin{array}{rr} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{array}}$$

besitzen. Jede der Normalformen geht aus der anderen dadurch hervor, dass bei jedem Zweierkästchen jeweils die Reihenfolge der zugehörigen orthonormalen Basisvektoren x, y vertauscht wird oder dass y durch $-y$ ersetzt wird.

2. Ist $A \in \mathbb{R}^{n\times n}$ orthogonal, so liefert das obige Beweisverfahren, angewandt auf die symmetrische Hilfsmatrix $B := A + A^\top$, eine orthogonale Matrix $S \in \mathbb{R}^{n\times n}$, sodass $S^\top AS$ obige Normalform besitzt.

Als Matrixversion von Satz 19.31 ergibt sich

Satz 19.34 *Zu jeder orthogonalen Matrix A gibt es eine orthogonale Matrix S, so daß*

$$B = S^\top AS$$

gilt, wobei B die Gestalt $(\star)$ in Satz 19.31 hat.

Diesen Satz kann man auffassen als eine **Klassifikation der linearen Isometrien** eines n-dimensionalen euklidischen Vektorraumes, wobei alle linearen Isometrien mit derselben Normalform - einschließlich der verschiedenen Diagonalordnungen - zur selben Klasse zählen.

Aus Satz 19.31 ergibt sich die folgende geometrische Beschreibung einer lineare Isometrie $\Phi : V \to V$. Es gibt eine Orthogonalzerlegung

$$V = U_1 \oplus U_2 \oplus W_1 \oplus \cdots \oplus W_k,$$

so dass

- $\Phi|_{U_1} = \mathrm{id}_{U_1}$ ist, also U_1 Fixunterraum ist,
- dass weiter $\Phi|_{U_2} : x \mapsto -x$ gilt, Φ also die Vektoren aus U_2 an $U_2^\perp = U_1 \oplus U_2 \oplus W_1 \oplus \cdots \oplus W_k$ spiegelt,
- und so dass $\Phi|_{W_i}$ für $i = 1, \ldots, k$ Drehungen in den zweidimensionalen Untervektorräumen W_i darstellen.

Ist die Dimensionen von V speziell 2 oder 3, so haben wir:

n=2: Nach (19.31) treten 4 Klassen von linearen Isometrien auf. Diese lassen sich durch die Normalformen

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ B_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ B_3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ B_4 = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix}$$

mit $0 < \omega < \pi$ beschreiben. Die entsprechenden linearen Isometrien sind der Reihe nach die Identität, eine (Geraden-)Spiegelung, die Spiegelung an 0 und eine Drehung um 0 mit Drehwinkel ω.

n=3: Ist $\det\Phi = 1$, so gibt es einen Vektor $v \neq 0$, der Fixvektor ist. Im Orthogonalraum $[v]^\perp$ ist Φ eine Drehung. Φ heißt dann **Drehung** des Vektorraums V mit der **Drehachse** $[v]$ und der Drehebene $[v]^\perp$.

Ist $\det\Phi = -1$, so gibt es ein $v \neq 0$ mit $\Phi(v) = -v$. Im Orthogonalraum $[v]^\perp$ ist Φ eine Drehung. Φ ist dann Produkt einer Drehung um die Achse $[v]$ mit einer Spiegelung an der Drehebene $[v]^\perp$ und heißt deswegen **Drehspiegelung** von V.

Beispiel 19.35 Die Matrix

$$A := \frac{1}{9\sqrt{2}} \begin{pmatrix} 9 & 4 & 8 & 1 \\ -4 & 9 & -1 & 8 \\ -8 & 1 & 9 & -4 \\ -1 & -8 & 4 & 9 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

ist orthogonal, denn es gilt $A^\top A = E_4$. Wir wollen ihre euklidische Normalform $\widetilde{A}$ bestimmen und betrachten hierzu zunächst die symmetrische Matrix

$$A + A^\top = \frac{1}{9\sqrt{2}} \begin{pmatrix} 18 & 0 & 0 & 0 \\ 0 & 18 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix} = \sqrt{2}E_4.$$

$A + A^\top$ hat den 4-fachen Eigenwert $\lambda = \sqrt{2}$. Die Normalform $\widetilde{A}$ von A besteht daher aus zwei Kästchen der Form

$$\boxed{\begin{matrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{matrix}}$$

mit $\cos\omega = \frac{1}{2}\sqrt{2}$. Daraus folgt $\omega = \frac{\pi}{4}$ und $\sin\omega = \frac{1}{2}\sqrt{2}$ sowie

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

Berechnung der orthogonalen Transformationsmatrix S:

Sei v Eigenvektor von $A + A^\top$ zum Eigenwert λ. Wegen $E_\lambda = \mathbb{R}^4$ können wir v beliebig wählen, etwa $v = e_1$. Dann ist

$$Av = \frac{1}{9\sqrt{2}} \begin{pmatrix} 9 \\ -4 \\ -8 \\ -1 \end{pmatrix}.$$

Orthogonalisieren ergibt

$$v_1 := e_1, \quad \tilde{v}_2 = Av_1 - \langle Av_1, v_1 \rangle v_1 = \frac{1}{9\sqrt{2}} \begin{pmatrix} 0 \\ -4 \\ -8 \\ -1 \end{pmatrix},$$

also

$$v_2 := \frac{\tilde{v}_2}{||\tilde{v}_2||} = \frac{1}{9} \begin{pmatrix} 0 \\ -4 \\ -8 \\ -1 \end{pmatrix}.$$

Nun muß $[v_1, v_2]^\perp$ bestimmt werden:

$$[v_1, v_2]^\perp = \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mid x_1 = 0,\ 4x_2 + 8x_3 + x_4 = 0\} = [\begin{pmatrix} 0 \\ 1 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}].$$

In $[v_1, v_2]^\perp$ bestimmen wir ebenfalls eine Orthonormalbasis: Wir wählen einen (beliebigen) Vektor $v_3 \in [v_1, v_2]^\perp$ mit $||v_3|| = 1$, bilden Av_3 und orthogonalisieren:

$$v_3 := \frac{1}{3\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 4 \end{pmatrix}, \quad \tilde{v}_4 = Av_3 - \langle Av_3, v_3 \rangle v_3 = \frac{1}{9} \begin{pmatrix} 0 \\ 7 \\ -4 \\ 4 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 4 \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 0 \\ 11 \\ -5 \\ -4 \end{pmatrix},$$

also

$$v_4 = \frac{\tilde{v}_4}{||\tilde{v}_4||} = \frac{1}{9\sqrt{2}} \begin{pmatrix} 0 \\ 11 \\ -5 \\ -4 \end{pmatrix}.$$

Damit erhalten wir

$$S = (v_1|v_2|v_3|v_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -4/9 & 1/3\sqrt{2} & 11/9\sqrt{2} \\ 0 & -8/9 & -1/3\sqrt{2} & -5/9\sqrt{2} \\ 0 & -1/9 & 4/3\sqrt{2} & -4/9\sqrt{2} \end{pmatrix}.$$

Bemerkung 19.36 Eine Klasse von Endomorphismen, die lineare Isometrien und selbstadjungierte lineare Abbildungen verallgemeinert sind die sogenannten normalen Selbstabbildungen. Ein Endomorphismus $\Phi : V \to V$ eines n-dimensionalen euklidischen oder unitären Vektorraumes V heißt **normal**, wenn gilt

$$\Phi \circ \Phi^* = \Phi^* \circ \Phi.$$

Übersetzt in die Matrizensprache haben wir die folgende Definition. Eine Matrix $A \in \mathbb{C}^{n\times n}$ bzw. $\mathbb{R}^{n\times n}$ heißt **normal**, falls gilt

$$AA^* = A^*A \qquad \text{bzw.} \quad AA^\top = A^\top A.$$

Beispiele:

1. Selbstadjungierte Endomorphismen sind normal. Symmetrische bzw. hermitesche Matrizen sind normal.
2. Lineare Isometrien sind normale Endmorphismen. Orthogonale bzw. unitäre Matrizen sind normal.
3. Schiefsymmetrische bzw. schiefhermitesche Matrizen, $A^\top = -A$ bzw. $A^* = \overline{A}^\top = -A$, sind normal.

Einige Eigenschaften, die wir für lineare Isometrien und selbstadjungierte Endomorphismen nachgewiesen haben, gelten allgemeiner auch für normale Endomorphismen. Weitere Informationen findet man in der angegebenen Literatur.

20 Bilinearformen

Spezielle Bilinearformen, die Skalarprodukte, haben wir in den vorhergehenden Kapiteln ausführlich untersucht. In diesem Kapitel werden wir allgemeine Bilinearformen noch etwas genauer ansehen. Im ganzen Kapitel ist V ein *reeller* Vektorraum.

20.1 Bilinearformen und quadratische Formen

Definition 20.1 Es sei V ein Vektorraum über den reellen Zahlen $\mathbb{R}$. Eine (reelle) **Bilinearform** ist eine Abbildung

$$\beta : V \times V \to \mathbb{R}; \quad (v, w) \mapsto \beta(v, w),$$

die linear in beiden Argumenten ist:
Für alle $v, w, v_1, v_2, w_1, w_2 \in V$ und alle $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ gilt

$$\begin{aligned} \beta(\lambda_1 v_1 + \lambda_2 v_2, w) &= \lambda_1 \beta(v_1, w) + \lambda_2 \beta(v_2, w) \\ \beta(v, \mu_1 w_1 + \mu_2 w_2) &= \mu_1 \beta(v, w_1) + \mu_2 \beta(v, w_2). \end{aligned}$$

Eine Bilinearform β heißt **symmetrisch** (bzw. **schiefsymmetrisch**), falls für alle $v, w \in V$ gilt

$$\beta(w, v) = \beta(v, w) \qquad (\text{bzw. } \beta(w, v) = -\beta(v, w)).$$

Beispiel 20.2

1. Ein reelles Skalarprodukt ist eine symmetrische Bilinearform.
2. Die Bilinearform

   $$\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \beta((x_1, x_2), (y_1, y_2)) := x_1 y_1 - x_2 y_2$$

   ist symmetrisch, aber kein Skalarprodukt.
3. Die Bilinearform

   $$\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad \beta((x_1, x_2), (y_1, y_2)) := x_1 y_2 - x_2 y_1$$

   ist schiefsymmetrisch.
4. Ist $(V, \langle , \rangle)$ ein euklidischer Vektorraum und Φ ein Endomorphismus, so ist

   $$\beta(v, w) := \langle v, \Phi(w) \rangle$$

   eine Bilinearform. Falls Φ selbstadjungiert ist, so ist β symmetrisch.

Falls V endlich-dimensional ist, so können wir auch einer bilinearen Abbildung β eine Matrix zuordnen. Wir wählen dazu eine Basis $\{a_1, \ldots, a_n\}$ von V. Für $x = \sum_{i=1}^n x_i a_i$ und $y = \sum_{i=1}^n y_i a_i$ haben wir

$$\beta(x, y) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \beta(a_i, a_j) =: \sum_{i=1}^n \sum_{j=1}^n x_i y_j a_{ij}. \tag{20.1}$$

Definition 20.3 Die Matrix $A = (a_{ij}) = (\beta(a_i, a_j)) \in \mathbb{R}^{n\times n}$ heißt **Matrix der Bilinearform** β bezüglich der Basis $\{a_1, \ldots, a_n\}$ von V.

Beispiel 20.4 Die Matrix von obigem Beispiel 2 (bzw. 3) bezüglich der Standardbasis von $\mathbb{R}^2$ ist

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (\text{bzw.} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}).$$

Umgekehrt erhält man zu jeder Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ eine Bilinearform β auf V (mit Basis $\{a_1, \ldots, a_n\}$), indem man β durch (20.1) <u>definiert</u>. Die Matrix von β bezügich der gegebenen Basis ist dann gerade A.

Bemerkung 20.5

1. Man kann zeigen: Die Menge aller Bilinearformen auf V ist ein $\mathbb{R}$-Vektorraum, der isomorph ist zu $\mathbb{R}^{n\times n}$, dem Vektorraum aller $(n \times n)$-Matrizen.
2. Eine Bilinearform β ist symmetrisch genau dann, wenn die Matrix von β bezüglich einer (beliebigen) Basis symmetrisch ist
3. Noch allgemeiner als Bilinearformen sind **Multilinearformen**. Diese werden bezüglich Basen nicht mehr durch Matrizen, sondern durch sogenannte "Tensoren" dargestellt (vgl. "Multilineare Algebra").

Wir überlegen jetzt noch, wie sich die Matrizen von Bilinearformen bei Basiswechseln verhalten (vergleiche dazu den entsprechenden Abschnitt 18.2 über Skalarprodukte).

Sei also $\{b_1, \ldots, b_n\}$ eine weitere Basis von V und $S = (s_{ik}) \in \mathbf{GL}(n, \mathbb{R})$ die Matrix des Basiswechsels, also

$$b_k = \sum_{i=1}^n s_{ik} a_i, \quad k = 1, \ldots, n.$$

Sei $B = (\beta(b_k, b_l)) \in \mathbb{R}^{n\times n}$ die Matrix von β bezüglich $\{b_1, \ldots, b_n\}$. Es gilt dann:

$$\begin{aligned} \beta(b_k, b_l) &= \beta(\sum_{i=1}^n s_{ik} a_i, \sum_{j=1}^n s_{jl} a_j) = \\ &= \sum_{i=1}^n \sum_{j=1}^n s_{ik} s_{jl}\, \beta(a_i, a_j) = \sum_{i=1}^n \sum_{j=1}^n s_{ik} a_{ij} s_{jl}, \end{aligned}$$

d.h. wir erhalten (natürlich) die gleiche Transformationsformel wie für Skalarprodukte:

$$B = S^\top AS.$$

Definition 20.6 Es sei $\beta : V \times V \to \mathbb{R}$ eine Bilinearform. Die Funktion

$$Q : V \to \mathbb{R}; \quad Q(v) := \beta(v, v)$$

heißt die zu β gehörige **quadratische Form**.

Beispiel 20.7 1. Ist die Bilinearform ein Skalarprodukt, $\beta = \langle , \rangle$, so ist die zugehörige quadratische Form das Quadrat der Norm: $Q(v) = \langle v, v\rangle = \|v\|^2$.

2. β sei die Bilinearform aus Beispiel 20.2. 3, also

$$\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad \beta((x_1, x_2), (y_1, y_2)) := x_1y_2 - x_2y_1.$$

Dann gilt für die zugehörige quadratische Form: $Q(x) = 0$ für alle $x \in \mathbb{R}^2$.

Hilfssatz 20.8 (Zusammenhang Bilinearform/Quadratische Form) *Es sei $\beta : V \times V \to \mathbb{R}$ eine Bilinearform und $Q : V \to \mathbb{R}$ die zugehörige quadratische Form. Durch*

$$\alpha(v, w) := \frac{1}{2}[Q(v + w) - Q(v) - Q(w)] \quad (v, w \in V)$$

wird eine symmetrische *Bilinearform definiert, zu der ebenfalls die quadratische Form Q gehört und für alle $v, w \in V$ gilt:*

$$\alpha(v, w) = \frac{1}{2}[\beta(v, w) + \beta(w, v)].$$

Ist insbesondere β symmetrisch, so gilt $\alpha = \beta$. In diesem Fall kann man also β aus der quadratischen Form (durch "Polarisieren") zurückgewinnen.

Beweis: Es gilt:

$$\begin{aligned} Q(v + w) &= \beta(v + w, v + w) = \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) \\ &= Q(v) + Q(w) + \beta(v, w) + \beta(w, v). \end{aligned}$$

Also folgt mit der Definition von α, dass

$$\alpha(v, w) = \frac{1}{2}[\beta(v, w) + \beta(w, v)]$$

und α ist bilinear. Es bleibt zu zeigen, dass Q auch zu α gehört:

$$\alpha(v, v) = \frac{1}{2}[\beta(v, v) + \beta(v, v)] = \beta(v, v) = Q(v).$$

■

20.2 Bilinearformen in euklidischen Vektorräumen

Der reelle Vektorraum V soll jetzt mit der Zusatzstruktur eines Skalarprodukts versehen sein, d.h. wir betrachten einen euklidischen Vektorraum $(V, \langle , \rangle)$.

Der folgende Satz ordnet einer (symmetrischen) Bilinearform einen (selbstadjungierten) Endomorphismus zu.

Satz 20.9 *Es sei β eine Bilinearform auf dem endlich-dimensionalen euklidischen Vektorraum $(V, \langle , \rangle)$. Dann gibt es eindeutig bestimmte Endomorphismen $\Phi : V \to V$ und $\Psi : V \to V$ mit*

$$\beta(v, w) = \langle v, \Phi(w) \rangle = \langle \Psi(v), w \rangle \qquad \forall v, w \in V.$$

Weiter ist Φ (und Ψ) selbstadjungiert genau dann, wenn β symmetrisch ist.

Beweis: Es sei $\{v_1, \dots, v_n\}$ eine Orthonormalbasis von V. Weiter sei $B = (b_{ik}) = (\beta(v_i, v_k))$ die Matrix von β bezüglich dieser Orthonormalbasis.

Existenz: Wir definieren $\Phi, \Psi \in \mathrm{Hom}\,(V, V)$ durch:

$$\begin{aligned} \Phi(v_k) &:= \sum_{j=1}^{n} b_{jk} v_j, \qquad k = 1, \dots, n, \\ \Psi(v_i) &:= \sum_{j=1}^{n} b_{ij} v_j, \qquad i = 1, \dots, n. \end{aligned}$$

Dann gilt

$$\begin{aligned} \langle v_i, \Phi(v_k) \rangle &= \sum_{j=1}^{n} b_{jk} \langle v_i, v_j \rangle = \sum_{j=1}^{n} b_{jk} \delta_{ij} = b_{ik} = \beta(v_i, v_k), \\ \langle \Psi(v_i), v_k \rangle &= \sum_{j=1}^{n} b_{ij} \langle v_j, v_k \rangle = \sum_{j=1}^{n} b_{ij} \delta_{jk} = b_{ik} = \beta(v_i, v_k). \end{aligned}$$

Wegen der Bilinearität von β und $\langle , \rangle$ folgt dann durch lineare Fortsetzung:

$$\beta(v, w) = \langle v, \Phi(w) \rangle = \langle \Psi(v), w \rangle, \qquad \forall v, w \in V.$$

Eindeutigkeit: Ist etwa $\tilde{\Phi}$ ein weiterer Endomorphismus mit $\beta(v, w) = \langle v, \tilde{\Phi}(w) \rangle$, so gilt

$$\langle v, \tilde{\Phi}(w) \rangle = \langle v, \Phi(w) \rangle \quad \forall v, w \in V.$$

Also $0 = \langle v, \tilde{\Phi}(w) - \Phi(w) \rangle$ für alle $v, w \in V$. Daraus folgt wegen der Definitheit des Skalarproduktes, dass $\tilde{\Phi}(w) - \Phi(w) = 0$ für alle $w \in V$, d.h. $\tilde{\Phi} = \Phi$.

Analog zeigt man, dass auch Ψ eindeutig ist.

Symmetrie: Ist β symmetrisch, so gilt für alle $v, w \in V$ wegen der Symmetrie des Skalarproduktes:

$$\langle v, \Phi(w)\rangle = \beta(v, w) = \beta(w, v) = \langle w, \Phi(v)\rangle = \langle \Phi(v), w\rangle,$$

d.h. Φ ist selbstadjungiert.

Ist umgekehrt Φ selbstadjungiert, so gilt für alle $v, w \in V$:

$$\beta(v, w) = \langle v, \Phi(w)\rangle = \langle \Phi(v), w\rangle = \langle w, \Phi(v)\rangle = \beta(w, v),$$

d.h. β ist symmetrisch. ■

Bemerkung 20.10

1. Ist in V eine beliebige Basis $\{a_1, \dots, a_n\}$ gegeben, so wird duch $\langle a_i, a_j\rangle := \delta_{ij}$ ein Skalarprodukt definiert. Bezüglich diesem Skalarprodukt ist die gegebene Basis eine Orthonormalbasis. Man erhält also eine Variante des obigen Satzes, indem man statt eines Skalarproduktes eine Basis vorgibt.
2. Die linearen Abbildungen Φ und Ψ sind vom gewählten Skalarprodukt in V abhängig. Die Eigenschaft "Φ selbstadjungiert $\Leftrightarrow$ β symmetrisch" hingegen ist vom gewählten Skalarprodukt unabhängig.
3. Aus dem Beweis folgt: Ist A die Matrix von β bezüglich einer Orthonormalbasis, so ist A (bzw. $A^\top$) die Matrix von Φ (bzw. Ψ) bezüglich dieser Basis.

Definition 20.11 Sei $\beta : V \times V :\to \mathbb{R}$ eine *symmetrische* Bilinearform. Der **Nullraum** von β ist

$$N := \{v \in V \mid \beta(v, w) = 0 \ \ \forall\, w \in V\}.$$

Bemerkung 20.12

1. Der Nullraum ist ein Untervektorraum (wegen der Bilinearität von β).
2. Da β symmetrisch ist, gilt auch

$$N = \{w \in V \mid \beta(v, w) = 0 \ \ \forall\, v \in V\}.$$

Beispiel 20.13

1. Der Nullraum eines Skalarproduktes ist trivial: $N = \{0\}$.
2. Sei $\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ bezüglich der Standardbasis von $\mathbb{R}^2$ gegeben durch die Matrix

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Dann ist der Nullraum von β:

$$N = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2 = 0, \quad \forall \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2\} = [\begin{pmatrix} 1 \\ 1 \end{pmatrix}].$$

Satz 20.14 (Interpretation Nullraum) *Sei $(V, \langle, \rangle)$ ein euklidischer Vektorraum, $\Phi : V \to V$ ein Endomorphismus und $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf V. Ist*

$$\beta(v, w) = \langle v, \Phi(w) \rangle \quad \textit{für alle} \;\; v, w \in V,$$

so gilt:

$$\textit{Nullraum von} \;\; \beta = \textit{Kern}\, \Phi.$$

Beweis: Aus

$$0 = \beta(v, w) = \beta(w, v) = \langle w, \Phi(v) \rangle$$

für alle $w \in V$ folgt $\Phi(v) = 0$ und umgekehrt. ∎

20.3 Hauptachsen-Transformation, Trägheits-Satz

Sei $(V, \langle, \rangle)$ ein euklidischer Vektorraum. Nach dem Orthognalisierungsverfahren von Gram-Schmidt existiert immer eine Orthonormalbasis von V. Bezüglich einer solchen ONB wird das Skalarprodukt $\langle, \rangle$ durch die Einheitsmatrix (also eine Diagonalmatrix) dargestellt.

Wir beweisen in diesem Abschnitt eine Verallgemeinerung für Bilinearformen: Zu einer *symmetrischen* Bilinearform β existiert eine Orthonormalbasis von V bezüglich welcher β durch eine Diagonalmatrix dargestellt wird. Die Bilinearform und das Skalarprodukt sind also "simultan diagonalisierbar".

Satz 20.15 (Hauptachsen-Transformation) *Sei V ein endlich-dimensionaler reeller Vektorraum. Weiter seien $\langle, \rangle$ ein Skalarprodukt und β eine symmetrische Bilinearform auf V. Dann existiert eine Orthonormalbasis $\{a_1, \ldots, a_n\}$ von V bezüglich welcher β durch eine Diagonalmatrix D (und $\langle, \rangle$ durch die Einheitsmatrix E_n) dargestellt wird. Die Diagonaleinträge von D sind gerade die Eigenwerte λ_i der Matrix von β bezüglich einer (beliebigen) Orthonormalbasis.*

Beweis: Da β symmetrisch ist, existiert nach Satz 20.9 ein selbstadjungierter Endomorphismus $\Phi : V \to V$ mit $\beta(v, w) = \langle v, \Phi(w) \rangle$ für alle $v, w \in V$. Nach dem Spektralsatz 19.9 ist Φ diagonalsierbar. Genauer: Es existiert eine Orthonormalbasis $\{a_1, \ldots, a_n\}$ aus Eigenvektoren von Φ. Für diese gilt dann:

$$\beta(a_i, a_j) = \langle a_i, \Phi(a_j) \rangle = \lambda_j \langle a_i, a_j \rangle = \lambda_j \delta_{ij}.$$

Bezüglich der Orthonormalbasis $\{a_1, \ldots, a_n\}$ wird β also durch die Diagonalmatrix

$$D := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

dargestellt. Dabei ist λ_j der Eigenwert, der zum Eigenvektor a_j von Φ gehört.

Wir zeigen noch: λ_j ist auch Eigenwert für die Darstellungsmatrix von β bezüglich einer beliebigen andern Orthonormalbasis. Dazu sei $\{b_1, \ldots, b_n\}$ eine weitere Orthonormalbasis von V. Sei S die Matrix des Basiswechsels. Die Darstellungsmatrix C von β bezüglich dieser Basis ist dann $C = S^\top DS$. Da beide Basen orthonormiert sind, ist S orthogonal, d.h. $S^\top = S^{-1}$. Also sind C und D ähnlich und haben insbesondere dieselben Eigenwerte. ■

Beispiel 20.16 Auf $\mathbb{R}^3$ sei eine symmetrische Bilinearform β bezüglich der Standardbasis gegeben durch

$$A = \begin{pmatrix} 4 & -5 & -2 \\ -5 & 4 & -2 \\ -2 & -2 & -8 \end{pmatrix}.$$

Wählt man in $\mathbb{R}^3$ (als "Hilfs-Skalarprodukt") das Standard-Skalarprodukt, so ist der zu β gehörige selbstadjungierte Endomorphismus gegeben durch

$$\Phi : \mathbb{R}^3 \to \mathbb{R}^3; \quad x \mapsto Ax.$$

Die Eigenwerte von Φ sind die Nullstellen des charakteristischen Polynoms:

$$p_\Phi = p_A = -(X+9)(X-9)X.$$

Bezüglich der Orthonormalbasis von $\mathbb{R}^3$ die aus Eigenvektoren von Φ zu den Eigenwerten $9, -9, 0$ besteht, hat Φ die Abbldungsmatrix

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die zugehörigen Basisvektoren $\{a_1, a_2, a_3\}$, d.h. die normierten Eigenvektoren von Φ sind

$$a_1 = \frac{1}{\sqrt{2}}(1, -1, 0), \ a_2 = \frac{1}{3\sqrt{2}}(1, 1, 4), \ a_3 = \frac{1}{3}(2, 2, -1).$$

Wir kommen jetzt auf Satz 20.15 zurück. Wir nummerieren die dortigen Eigenwerte $\lambda_1, \dots, \lambda_n$ so, dass $\lambda_1, \dots, \lambda_p$ positiv, $\lambda_{p+1}, \dots, \lambda_{p+q}$ negativ und $\lambda_{p+q+1}, \dots, \lambda_n$ gleich Null sind. Dann definieren wir

$$\begin{aligned} b_i &:= \frac{1}{\sqrt{\lambda_i}} a_i,\ i = 1, \dots, p; \qquad b_j := \frac{1}{\sqrt{-\lambda_j}} a_j,\ j = p+1, \dots, p+q; \\ b_k &:= a_k,\ k = p+q+1, \dots, n. \end{aligned}$$

Als eine orthogonale Menge von Vektoren, die den Nullvektor nicht enthält, ist $\{b_1, \dots, b_n\}$ eine Basis von V. Die symmetrische Bilinearform β stellt sich bezüglich dieser neuen Basis durch die folgende Diagonalmatrix dar:

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & 0 & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & 0 & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}, \tag{20.2}$$

dabei ist die Anzahl der Einsen gleich p, die Anzahl der Minus-Einsen ist q und die Anzahl der Nullen ist $n - (p+q)$.

Dieses Resultat wird im folgenden Satz von Sylvester [12] noch präzisiert:

Satz 20.17 (Trägheitssatz von Sylvester) *Es sei β eine symmetrische Bilinearform auf einem endlich-dimensionalen reellen Vektorraum V. Dann existiert eine Basis $\{b_1, \dots, b_n\}$ von V in der β durch die Matrix (20.2) dargestellt wird. Insbesondere hat dann die zughörige quadratische Form bezüglich dieser Basis die Gestalt*

$$Q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2, \quad x = \sum_{i=1}^{n} x_i b_i.$$

Die Anzahl p der positiven bzw. q der negativen Eigenwerte der Matrix (20.2) ist dabei durch β eindeutig bestimmt (also von der Wahl der Basis unabhängig).

Beweis: Nur der 2. Teil des Satzes ist noch zu beweisen. Wir haben schon gezeigt: Der Nullraum N von β ist von der Basis unabhängig (denn $N = \operatorname{Kern} \Phi$ für die zu

[12] James Sylvester (1814-1897)

β gehörige selbstadjungierte lineare Abbildung, vgl. Satz 20.14). Damit ist $p+q = \dim V - \dim N$ durch β eindeutig bestimmt.

Wir benutzen jetzt, dass die symmetrische Bilinearform β durch Q eindeutig bestimmt ist und umgekehrt (vgl. Hilfssatz 20.8).

Es sei $\{b'_1, \dots, b'_n\}$ eine Basis von V, in der Q dargestellt wird durch

$$Q(x) = (x'_1)^2 + \cdots + (x'_{p'})^2 - (x'_{p'+1})^2 - \cdots - (x_{p'+q'})^2, \quad x = \sum_{i=1}^{n} x'_i b'_i.$$

Da $p' + q' = p + q$ ist, genügt es zu zeigen, dass $q' = q$ ist. Dazu betrachten wir die Unterräume U und U', die definiert sind durch:

$$U := [b_{p+1}, \dots, b_{p+q}], \qquad U' := [b'_1, \dots, b'_{p'}, b'_{p'+q'+1}, \dots, b'_n].$$

Dann gilt

$$Q(x) < 0 \ \text{ für } \ x \in U, x \neq 0; \qquad Q(y) \geq 0 \ \text{ für } \ y \in U'.$$

Es folgt $U \cap U' = \{0\}$. Daraus schließen wir:

$$\dim(U + U') = \dim U + \dim U' = q + p' + (n - p' - q') = n - q' + q.$$

Da $U + U'$ ein Untervektorraum von V ist, gilt $n - q' + q \leq n$, also $q \leq q'$.

Ein analoges Argument mit Untervektorräumen W und W' definiert durch:

$$W := [b'_{p'+1}, \dots, b'_{p'+q'}], \qquad W' := [b_1, \dots, b_p, b_{p+q+1}, \dots, b_n]$$

ergibt, dass $q' \leq q$. Also haben wir insgesamt $q' = q$, d.h. q ist eindeutig durch Q (und damit auch β) bestimmt. Da (wie oben gezeigt) auch $p + q$ eindeutig ist, ist auch p eindeutig. ■

Definition 20.18 Es sei Q eine quadratische Form auf einem n-dimensionalen reellen Vektorraum V. Das eindeutig bestimmte Trippel $(p, q, n - p - q)$ in Satz 20.17 heißt **Signatur** von Q.

Eine quadratische Form Q heißt

- **positiv definit**, wenn $p = n$ und $q = 0$ ist;
- **negativ definit**, wenn $p = 0$ und $q = n$ ist;
- **positiv semi-definit**, wenn $p < n$ und $q = 0$ ist;
- **negativ semi-definit**, wenn $p = 0$ und $q < n$ ist;
- **indefinit**, wenn $p > 0$ und $q > 0$ ist.

Bemerkung 20.19 Die Bezeichnungen "Hauptachsen-Transformation" und "Trägheitssatz" stammen aus der Physik. Dort kommen symmetrische Blinearformen (in $\mathbb{R}^3$) im Zusammenhang mit der Dynamik starrer Körper vor ("Trägheitsellipsoide").

Beispiel 20.20

1. Ein Skalarprodukt ist positiv definit.
2. Die quadratische Form Q auf $\mathbb{R}^4$ sei bezüglich der Standardbasis gegeben durch:

$$Q(x) := 2x_1x_2 + 2x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4, \quad x = (x_1, x_2, x_3, x_4).$$

Die Matrix der zugehörigen symmetrischen Bilinearform β ist dann

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

Wir wählen das Standardskalarprodukt auf $\mathbb{R}^n$ (als Hilfsstruktur). Zu β gehört dann ein selbstadjungierter Endomorphismus Φ, der ebenfalls durch die Matrix A dargestellt wird. Das charakteristische Polynom ist $p_\Phi = (X+1)^3(X-3)$. Daraus ergibt sich die Sylvester-Form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Die quadratische Form Q hat Signatur $(1, 3, 0)$ und ist also indefinit.

Folgerung 20.21 (Simultanes Diagonalisieren von quadratischen Formen) *Es sei V ein reeller, endlich-dimensionaler Vektorraum. Auf V seien zwei quadratische Formen Q_1 und Q_2 gegeben. Weiter sei Q_1 positiv definit. Dann existiert eine Basis $\{b_1, \ldots, b_n\}$ von V, so dass für $x = \sum_{i=1}^n x_i\, b_i$ gilt:*

$$Q_1(x) = \sum_{i=1}^n x_i^2 \qquad \text{und} \qquad Q_2(x) = \sum_{i=1}^n \lambda_i x_i^2.$$

Beweis: Die zu Q_1 gehörige symmetrische Bilinearform β_1 ist positiv definit, kann also als Skalarprodukt in V aufgefasst werden. Nach Satz 20.15 über Hauptachsentransformationen existiert eine (bezüglich β_1) orthonormierte Basis $\{b_1, \ldots, b_n\}$ von V, in der Q_2 Diagonalgestalt hat. ■

20.4 Kriterium für “positiv definit”

Die Matrix des Skalarprodukts eines euklidischen Vektorraumes (bezüglich irgendeiner Basis) ist symmetrisch (siehe Abschnitt 18.2). Die Frage, wann eine solche Matrix A positiv definit ist, haben wir schon in Satz 19.13 beantwortet: A ist genau dann positiv definit, wenn alle Eigenwerte positiv sind. Oft ist es aber schwierig, die Eigenwerte einer Matrix A auszurechnen. Dann ist das folgende Kriterium nützlicher.

Satz 20.22 (Jacobi-Hurwitz) [13] *Für eine symmetrische Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ sind folgende Aussagen äquivalent:*

(a) A ist positiv definit.

(b) Die Determinanten aller **Hauptminoren** *von A sind positiv, d.h. für $k = 1, \dots, n$ gilt:*
$$\det\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk}\end{pmatrix} > 0.$$

(c) Es gibt eine obere Dreiecksmatrix $B \in \mathbf{GL}(n, \mathbb{R})$ mit $A = B^\top B$.

Beweis: (a)⇒(b): Ist A positiv definit, so ist auch jede der Matrizen
$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk}\end{pmatrix}, \quad k = 1, \dots, n,$$
positiv definit. Ist nämlich $(x_1, \dots, k_k) \neq (0, \dots, 0)$, so ist $x := (x_1, \dots, x_k, 0, \dots, 0) \neq (0, \dots, 0)$ und es gilt
$$\begin{pmatrix} x_1 & \cdots & x_k\end{pmatrix} A_k \begin{pmatrix} x_1 \\ \vdots \\ x_k\end{pmatrix} = x^\top A x > 0.$$

Nach Satz 19.10 ist jede der symmetrischen Matrizen A_k ähnlich zu einer Diagonalmatrix D_k:
$$A_k = S_k^{-1} D_k S_k.$$

Nach Satz 19.13 sind die Eigenwerte von A_k alle positiv, also gilt für $k = 1, \dots, n$
$$\det A_k = \det D_k > 0.$$

[13] Carl Gustav Jacobi (1804-1851), Adolf Hurwitz (1859-1919)

(b)$\Rightarrow$(c): Für $k = 1, \ldots, n$ konstruieren wir rekursiv obere Dreicksmatrizen $B_k \in \mathbf{GL}(\mathbb{R}, k)$ mit $B_k^\top B_k = A_k$.

Für $k = 1$ setzen wir $B_1 := \sqrt{a_{11}}$. Wir nehmen nun an, dass B_k für ein k mit $1 \leq k \leq n-1$ bereits konstruiert ist und machen für $B_{k+1} \in \mathbf{GL}(\mathbb{R}, k+1)$ den Ansatz

$$B_{k+1} := \begin{pmatrix} B_k & b_k \\ 0 & \beta_k \end{pmatrix}, \quad b_k \in \mathbb{R}^k, \beta_k \in \mathbb{R}.$$

Unser Wunsch ist, dass gilt

$$A_{k+1} = \begin{pmatrix} A_k & a_{k+1} \\ a_{k+1}^\top & a_{k+1,k+1} \end{pmatrix} = B_{k+1}^\top B_{k+1} = \begin{pmatrix} B_k^\top B_k & B_k^\top b_k \\ b_k^\top B_k & b_k^\top b_k + \beta_k^2 \end{pmatrix}.$$

Wir müssen also setzen $b_k := (B_k^\top)^{-1} \cdot a_{k+1}$ (hier haben wir benutzt, dass B_k regulär ist).

Eine weitere notwendige Bedingung liefert uns einen Kandidaten für β_k. Wir benutzen die Voraussetzung, dass $\det A_{k+1} > 0$ ist: Aus dem Ansatz haben wir dann

$$\det A_{k+1} = (\det B_{k+1})^2 = \beta_k^2 (\det B_k)^2.$$

Also setzen wir $\beta_k := \sqrt{\det A_{k+1}} / \det B_k$.

Nach der bisherigen Konstruktion stimmen alle Einträge von A_{k+1} und $B_{k+1}^\top B_{k+1}$ bis eventuell auf den Eintrag an der Stelle $(k+1, k+1)$ überein. Ausserdem stimmen (ebenfalls nach Konstruktion) die Determinanten dieser Matrizen überein. Um auch die Gleicheit der Einträge an der Stelle $(k+1, k+1)$ zu zeigen, entwickeln wir diese Determinanten jeweils nach der letzten Zeile und erhalten eine Gleichung der Form

$$R + a_{k+1,k+1} \det A_k = \det A_{k+1} = \det B_{k+1}^\top B_{k+1} = R' + (b_k^\top b_k + \beta_k^2) \det B_k^\top B_k.$$

Die Terme R und R' stimmen nach dem oben Gesagten überein. Da nach Rekursions-Voraussetzung auch $\det A_k = \det B_k^\top B_k$ gilt, folgt $a_{k+1,k+1} = b_k^\top b_k + \beta_k^2$. Insgesamt haben wir damit $A_{k+1} = B_{k+1}^\top B_{k+1}$. Für $k+1 = n$ ist das die Behauptung (c) des Satzes.

(c)$\Rightarrow$(a): Sei $x \neq 0$. Dann ist (wegen B regulär) $Bx \neq 0$ und es gilt

$$x^\top A x = x^\top B^\top B x = (Bx)^\top (Bx) > 0.$$

Also ist A positiv definit. ■

Bemerkung 20.23 1. Die Zerlegung einer positiv definiten Matrix A in der Form $A = B^\top B$ mit einer regulären Dreiecksmatrix B spielt unter dem Namen **Cholesky-Zerlegung** auch eine wichtige Rolle in der numerischen Mathematik.

2. Der Beweis für die Implikation (c)$\Rightarrow$(a) in Satz 20.22 zeigt, dass man nur braucht, dass B regulär (aber nicht unbedingt eine Dreiecksmatrix) ist. Satz 20.22 liefert also ein Rezept zur Konstruktion von positiv definiten symmetrischen Matrizen: Ist $R \in \mathbf{GL}(n, \mathbb{R})$ (also eine reguläre Matrix), so ist $A := R^\top R$ symmetrisch und positiv definit.

Beispiel 20.24 Für welche $a \in \mathbb{R}$ ist

$$A := \begin{pmatrix} 1 & a & -1 \\ a & 9 & 0 \\ -1 & 0 & 4 \end{pmatrix}$$

positiv definit? Es gilt

$$p = -X^3 + 14X^2 - (48 - a^2)X + 27 - 4a^2.$$

Hier die Nullstellen in Abhängigkeit von a bestimmen zu wollen, ist nicht einfach. Wir wenden deshalb obigen Satz an und erhalten:

$$\begin{aligned}
&\det(A_1) = 1 > 0, \\
&\det(A_2) = 9 - a^2 > 0 \Leftrightarrow |a| < 3, \\
&\det(A_3) = \det(A) = 27 - 4a^2 > 0 \Leftrightarrow |a| < \tfrac{3}{2}\sqrt{3}.
\end{aligned}$$

Somit ist A genau dann positiv definit, wenn $-\frac{3}{2}\sqrt{3} < a < \frac{3}{2}\sqrt{3}$ gilt.

Satz 20.25 (Cartan-Zerlegung) [14] *Zu einer regulären Matrix $A \in \mathbf{GL}(n, \mathbb{R})$ existieren orthogonale Matrizen $O_1, O_2 \in O(n)$ und eine positiv definite Diagonalmatrix D, so dass $A = O_1 D O_2$.*

Beweis: Nach Bemerkung 20.23 ist $B := AA^\top$ positiv definit und es ist $B = S^{-1}D'S$ für eine orthogonale Matrix S und eine positive Diagonalmatrix D'. Nach Folgerung 19.15 ist $P := S^{-1}\sqrt{D'}S = P^\top$ auch positiv definit. Wir behaupten, dass $O := P^{-1}A$ orthogonal ist:

Nach Definition von O haben einmal $O^{-1} = A^{-1}P$. Andererseits ist $O^\top = A^\top(P^{-1})^\top = A^\top P^{-1}$. Wegen $P^2 = AA^\top \Leftrightarrow A^{-1}P = A^\top P^{-1}$ folgt die Behauptung.

Wir haben also $A = PO$ für O orthogonal und P positiv definit symmetrisch. Nach Satz 20.15 und dessen Beweis existiert eine orthogonale Matrix O_1 und eine positive Diagonalmatrix D so, dass $P = O_1 D O_1^\top$. Damit haben wir $A = O_1 D O_1^\top O$ und mit $O_2 := O_1^\top O$ folgt die Behauptung des Satzes. ■

[14] Élie Cartan (1869-1961)

20.5 Extremaleigenschaft der Eigenwerte

Wir wollen hier noch auf eine Eigenschaft der Eigenwerte einer reellen symmetrischen Matrix hinweisen, die für die numerische Behandlung des Eigenwertproblems wichtig ist.

Versucht man, die Eigenwerte und Eigenvektoren einer reellen symmetrischen Matrix mit Hilfe des charakteristischen Polynoms zu bestimmen, so stößt man auf Schwierigkeiten, indem ja i.A. die Nullstellen eines Polynoms nur näherungsweise bestimmbar sind. Um dieses Problem zu umgehen, sind numerische Verfahren entwickelt worden, die auf der im Folgenden dargestellten Extremaleigenschaft der Eigenwerte beruhen.

Wir betrachten den euklidischen Vektorraum $\mathbb{R}^n$ versehen mit dem Standard-Skalarprodukt $\langle , \rangle$. Nach Satz 20.9 gibt es dann eine eineindeutige Beziehung zwischen den selbstadjungierten Endomorphismen $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ und den symmetrischen Bilinearformen $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Dabei gilt

$$\beta(x,x) = \langle x, \Phi(x) \rangle, \quad x \in \mathbb{R}^n.$$

Bezüglich der Standardbasis von $\mathbb{R}^n$ werden β und Φ durch die gleiche reelle symmetrische $(n \times n)$-Matrix beschrieben und umgekehrt definiert jede solche Matrix einen selbstadjungierten Endomorphismus Φ von $\mathbb{R}^n$ und eine symmetrische Bilinearform β auf $\mathbb{R}^n$, so dass die obige Gleichung erfüllt ist.

Wir definieren eine Funktion $q : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ als Quotient

$$q(x) := \frac{\beta(x,x)}{\langle x, x \rangle} = \frac{\langle x, \Phi(x) \rangle}{\langle x, x \rangle}, \quad x \in \mathbb{R}^n \backslash \{0\},$$

und leiten Eigenschaften von q her.

Ist $v \in \mathbb{R}^n$ ein Eigenvektor von Φ zum Eigenwert λ, so gilt

$$q(v) = \frac{\langle v, \Phi(v) \rangle}{\langle v, v \rangle} = \frac{\lambda \langle v, v \rangle}{\langle v, v \rangle} = \lambda.$$

Die Eigenwerte von Φ liegen somit im Wertebereich von q.

Für $\alpha \in \mathbb{R}$, $\alpha \neq 0$, gilt

$$q(\alpha x) = \frac{\langle \alpha x, \Phi(\alpha x) \rangle}{\langle \alpha x, \alpha x \rangle} = \frac{\alpha^2 \langle x, \Phi(x) \rangle}{\alpha^2 \langle x, x \rangle} = q(x).$$

Wir nutzen diese Eigenschaft wie folgt aus: Um das Verhalten der Funktion q zu studieren, genügt es, die Vektoren mit Norm Eins (die "Einheits-Sphäre") zu betrachten. Wir setzen deshalb

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

und betrachten q im Folgenden als reellwertige Funktion auf S.

Führen wir in $\mathbb{R}^n$ die übliche Topologie ein, so ist S als abgeschlossene und beschränkte Teilmenge von $\mathbb{R}^n$ kompakt. Da die Funktion q stetig ist, folgt mit einem Satz von Weierstraß, dass q auf S sowohl ein Maximum als auch ein Minimum annimmt (siehe z.B. [16] Kap. 7.5).

Satz 20.26 *Das Maximum (bzw. Minimum) der Funktion q auf S ist der grösste (bzw. kleinste) Eigenwert. Die Maximalstellen (bzw. Minimalstellen) sind die zugehörigen Eigenvektoren.*

Beweis: Nach Satz 20.15 existiert eine Orthonomalbasis $\{b_1, \dots, b_n\}$ von $\mathbb{R}^n$, so dass für $x = \sum_{i=1}^n x_i b_i$ gilt $\beta(x,x) = \sum_{i=1}^n \lambda_i \; x_i^2$. Dabei sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte von Φ. Sie seien so numeriert, dass λ_1 der grösste und λ_n der kleinste Eigenwert ist: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Dann haben wir

$$\lambda_1(x_1^2 + \cdots + x_n^2) \geq \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2 \geq \lambda_n(x_1^2 + \cdots + x_n^2). \qquad (*)$$

Daraus folgt sofort $\lambda_1 \geq q(x) \geq \lambda_n$.

Wie wir oben gesehen haben, nimmt die Funktion q auf den Eigenvektoren zum Eigenwert λ_1 den Wert λ_1 und auf den Eigenvektoren zum Eigenwert λ_n den Wert λ_n an. Es bleibt also nur noch zu zeigen, dass diese Extremalwerte nur auf den entsprechenden Eigenvektoren angenommen werden. Das folgt aber unmittelbar aus der Ungleichung $(*)$. Gilt nämlich

$$\lambda_1(x_1^2 + \cdots + x_n^2) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2,$$

so folgt daraus $x_i = 0$ für all diejenigen Indizes i mit $\lambda_i \neq \lambda_1$, denn λ_1 ist ja der größte Eigenwert.

Für den kleinsten Eigenwert verläuft der Beweis analog. ■

20.6 Klassifikation von Matrizen: Eine Übersicht

Will man sich eine Übersicht über eine große und vielleicht komplizierte Gesamtheit $\mathcal{M}$ von mathematischen Objekten verschaffen, so ist es oft zweckmäßig (aber auch etwas willkürlich), zwischen "wesentlichen" und "unwesentlichen" Eigenschaften zu unterscheiden. Vergisst man dann die "unwesentlichen" Eigenschaften, so wird die Gesamtheit der Objekte "übersichtlicher" und man kann versuchen, sie aufzulisten, zu klassifizieren. Mathematisch beschreibt man solche Verdichtungen von Information durch *Äquivalenzrelationen* $\sim$. Man erhält als neue ("komprimierte") Gesamtheit die Menge $\mathcal{M}/\sim$ der Äquivalenzklassen. *Klassifizieren* heißt dann, die Menge $\mathcal{M}/\sim$ zu verstehen. Zwei häufige Methoden sind:

- Klassifikation durch charakteristische Daten,
- Klassifikation durch Repräsentanten.

Beispiel: Sei $\mathcal{V}$ die Menge aller endlich-dimensionalen $\mathbb{K}$-Vektorräume. Man kann jetzt z.B. entscheiden, dass zwei Vektorräume $V, W \in \mathcal{V}$ "im wesentlichen gleich" sind ($V \sim W$), wenn es einen Isomorphismus zwischen V und W gibt. Die Klassifikation durch charakteristische Daten geschieht z.B. mittels der *Dimension*:

$$V \sim W \iff \dim V = \dim W.$$

Eine Klassifikation durch Repräsentanten ist z.B.

$$\mathcal{V}/\sim = \{\mathbb{K}^n \mid n \in \mathbb{N}_0\}.$$

Wir wiederholen nochmals einige der verschiedenen Äquivalenzrelationen, die wir für Mengen von Matrizen erklärt haben.

Definition 20.27 1. Für Matrizen $A, B \in \mathbb{K}^{m\times n}$ definieren wir

$$A \overset{1}{\sim} B \ :\iff \ \text{Es gibt reguläre } T \in \mathbf{GL}(m,\mathbb{K}), S \in \mathbf{GL}(n,\mathbb{K}) \text{ mit } B = TAS.$$

2. Für Matrizen $A, B \in \mathbb{C}^{n\times n}$ definieren wir

$$A \overset{2}{\sim} B \ :\iff \ \text{Es gibt eine reguläre Matrix } S \in \mathbf{GL}(n,\mathbb{C}) \text{ mit } B = S^{-1}AS.$$

3. Für *symmetrische* Matrizen $A, B \in \mathbb{R}^{n\times n}$ definieren wir

$$A \overset{3}{\sim} B \ :\iff \ \text{Es gibt eine orthogonale Matrix } S \in O(n) \text{ mit } B = S^\top AS.$$

4. Für *symmetrische* Matrizen $A, B \in \mathbb{R}^{n\times n}$ definieren wir

$$A \overset{4}{\sim} B \ :\iff \ \text{Es gibt eine reguläre Matrix } S \in \mathbf{GL}(n,\mathbb{R}) \text{ mit } B = S^\top AS.$$

Zu jeder dieser Äquivalenzrelationen haben wir einen Klassifikationssatz bewiesen:

Zu $\overset{1}{\sim}$: Das ist die *Äquivalenz* von Matrizen. Die Klassifizierung geschieht durch den *Rang* als charakteristisches Datum. Klassifizierung durch Repräsentanten:

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

siehe Abschnitt 10.3.

Zu $\overset{2}{\sim}$: Das ist die *Ähnlichkeit* von Matrizen. Klassifizierung durch Repräsentanten in *Jordanscher Normalform* (vgl. Satz 17.12).

Zu $\overset{3}{\sim}$: Klassifizierung durch *Diagonalmatrizen* als Repräsentanten oder durch *Eigenwerte mit Vielfachheiten* als charakteristische Daten (vgl. Hauptachsentransformation, Satz 20.15).

Zu $\overset{4}{\sim}$: Klassifizierung durch *Signatur* als charakteristische Daten (vgl. Trägheitssatz von Sylvester 20.17).

Weitere Äquivalenzrelationen bzw. Klassifikationen durch Repräsentanten (Normalformen) haben wir auch für unitäre und orthogonale Matrizen bewiesen. Siehe dazu die Sätze 19.30 und 19.31.

21 Affine und euklidische Geometrie

In diesem Kapitel werden wir die Grundbegriffe der affinen und der euklidischen Geometrie kennen lernen.

21.1 Was ist Geometrie?

Eine mögliche Antwort auf diese Frage hat Felix KLEIN[15] 1872 in seinem "Erlanger-Programm" formuliert:

Eine **Geometrie** *ist eine Menge M zusammen mit einer Gruppe G von Transformationen (d.h. gewissen bijektiven Selbstabbildungen) von M. Eine Eigenschaft oder eine Größe heißt* **geometrisch**, *wenn sie invariant ist unter G.*

Ein (allgemeines) Beispiel: Ist (M, d) ein metrischer Raum und G die Gruppe der Isometrien von M, so ist der Abstand zwischen zwei Punkten eine geometrische Größe (denn das ist gerade die Definition einer Isometrie).

Der Kleinsche Geometrie-Begriff aus dem 19. Jahrhundert setzt (zu)viel Symmetrie voraus und ist deshalb zu restriktiv, um die gesamte moderne "Geometrie" (wie z.B. die Riemannsche Geometrie oder die algebraische Geometrie) zu erfassen.

21.2 Gruppenaktionen

Um affine Räume definieren zu können, benötigen wir den Begriff einer Gruppen-Operation (oder -Aktion). Dieses fundamentale und nützliche Konzept kommt in vielen Bereichen der Mathematik vor.

Definition 21.1 Eine **Operation** (oder **Aktion**) einer Gruppe $(G, \cdot)$ auf einer Menge X ist eine Abbildung

$$\varphi : G \times X \to X; \quad (g, x) \mapsto \varphi(g, x),$$

die folgende Eigenschaften besitzt:

- Ist e das neutrale Element von G, so gilt $\varphi(e, x) = x$ für alle $x \in X$.
- Für alle $g, h \in G$ und alle $x \in X$ gilt $\varphi(g \cdot h, x) = \varphi(g, \varphi(h, x))$.

Die **Bahn** eines Punktes $x \in X$ bzgl. φ ist die Menge $\{\varphi(g, x) \subset X \mid g \in G\}$. Der **Stabilisator** von x bzgl. φ ist die Untergruppe $G_x := \{g \in G \mid \varphi(g, x) = x\}$ von G.

[15] 1849-1925

Die Bahnen einer G-Operation definieren eine Äquivalenzrelation auf der Menge X: $x \sim y \Longleftrightarrow x$ gehört zur Bahn von y. Damit folgt: die Menge X zerfällt in disjunkte Bahnen (Äquivalenzklassen).

Eine G-Aktion heißt **transitiv**, wenn es genau *eine* Bahn gibt. In diesem Fall heißt die Menge X **homogen**.

Eine G-Aktion φ heißt **einfach transitiv**, wenn es zu je zwei Elementen x und y aus X genau ein $g \in G$ gibt, so dass $y = \varphi(g, x)$.

Beispiel 21.2

1. Die Gruppe $SO(2)$ operiert (einfach) transitiv auf dem Einheitskreis $S^1 := \{(x_1, x_2)^\top \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$.
2. Die Gruppe $SO(2)$ operiert (mit unendlich vielen Bahnen) auf der 2-dimensionalen Einheitssphäre $S^2 := \{(x_1, x_2, x_3)^\top \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$ durch Drehungen um die x_3-Achse.
3. Die Gruppe $SO(3)$ operiert transitiv auf S^2.

21.3 Affine Räume

Um lineare Selbstabbildungen eines $\mathbb{K}$-Vektorraumes V zu vestehen, haben wir versucht, diese durch möglichst einfache Matrizen darzustellen. Dazu muss man jeweils geeignete Basen wählen. Die Elemente von V werden dann durch Koordinatenvektoren in $\mathbb{K}^n$ dargestellt. Basiswechsel entsprechen *linearen* Koordinatentransformationen in $\mathbb{K}^n$ (siehe 10.1).

Für gewisse, mehr geometrische Fragestellungen in Vektorräumen ist es nun zweckmäßig, neben linearen Koordinatentransformationen auch die einfachsten nicht-linearen Abbildungen (also die Translationen) zu betrachten. Hier ist ein Beispiel.

Beispiel 21.3 In $\mathbb{R}^2$ sei die folgende Teilmenge gegeben:

$$K := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 2(x + y) + 1 = 0\}.$$

Um zu sehen, was diese Menge K ist, formen wir zuerst die Definitionsgleichung um ("quadratisches Ergänzen"):

$$x^2 + y^2 - 2(x + y) + 1 = x^2 - 2x + 1 + y^2 - 2y + 1 - 1 = (x - 1)^2 + (y - 1)^2 - 1 = 0.$$

Nun machen wir die "(affine) Koordinaten-Transformation"

$$\begin{aligned} \tilde{x} &:= x - 1 \\ \tilde{y} &:= y - 1. \end{aligned}$$

In den neuen Koordinaten lautet die Gleichung für K dann einfach $\tilde{x}^2 + \tilde{y}^2 = 1$, und wir erkennen, dass K ein Kreis ist mit Zentrum $(1, 1)$.

Um solche solche *affinen Transformationen* genauer untersuchen zu können, benötigen wir den Begriff eines affinen Raumes. Die Grundidee ist die folgende. Betrachtet man $\mathbb{R}^n$ als Vektorraum, so ist der Nullvektor ausgezeichnet. Als affiner Raum betrachtet ist $\mathbb{R}^n$ aber "homogen": jeder Punkt ist "gleichwertig".

Definition 21.4 Es seien V ein Vektorraum über einem Körper $\mathbb{K}$, A eine (nichtleere) Menge und $\tau : (V, +) \times A \to A$ eine Operation der (additiven) Gruppe $(V, +)$ auf A. Dann heißt das Tripel (A, V, τ) ein **affiner Raum**, mit **Translationsvektorraum** V, falls die Operation τ einfach transitiv ist. Es gilt also:

A1 Für alle $p \in A$ ist $\tau(0, p) = p$.

A2 Für alle $p \in A$ und alle $v_1, v_2 \in V$ ist $\tau(v_1, \tau(v_2, p)) = \tau(v_1 + v_2, p)$.

A3 Für alle $p, q \in A$ gibt es genau ein $v \in V$ mit $\tau(v, p) = q$.

Der nach **A3** durch p und q eindeutig bestimmte Vektor v wird gelegentlich auch mit $\overrightarrow{pq}$ bezeichnet (der "Ortsvektor von p nach q").

Beispiel 21.5 Ist $A = V$ ein Vektorraum und $\tau = +$ die Vektorraum-Addition, so ist $(V, V, +)$ ein affiner Raum. Ist speziell $V = \mathbb{K}^n$ so heißt $(\mathbb{K}^n, \mathbb{K}^n, +)$ n-dimensionaler **affiner Standardraum**.

Bemerkung 21.6 Sei $p \in A$ beliebig (aber fest gewählt). Dann ist die Abbildung

$$F_p : V \to A, \quad F_p(v) := \tau(v, p)$$

wegen **A3** bijektiv. Heuristisch kann man F_p so beschreiben: Heftet man in $p \in A$ eine Kopie von V an, so erhält man A. Die Umkehrabbildung F_p^{-1} besagt: Wählt man in A einen Ursprung p, so wird A zu einer Kopie von V.

Ein Vektorraum ist also dasselbe wie ein affiner Raum mit einem ausgezeichneten Punkt (Nullvektor als Ursprung).

KONVENTION. Im Folgenden werden wir uns auf die in Beispiel 21.5 beschriebenen affinen Räume $(V, V, +)$ beschränken. Dadurch vereinfacht sich die Darstellungsweise (außerdem kann man zeigen: Jeder affine Raum ist isomorph zu einem affinen Raum dieses Typs).

Die Notation mit Tripeln ist zwar wichtig für die begriffliche Klarheit, aber sehr schwerfällig. Wir werden daher im Folgenden den affinen Raum $(V, V, +)$ einfach mit $\mathbb{A}$ oder auch $\mathbb{A}(V)$ bezeichnen.

Definition 21.7 Es sei $\mathbb{A}(V)$ der zu einem Vektorraum gehörige affine Raum. Eine Teilmenge $A \subset V$ heißt **affiner Unterraum**, falls ein $a \in A$ und ein Untervektorraum U von V existieren, so dass gilt

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

Der Untervektorraum U heißt **Translationsraum** von A.

Die **Dimension** eines affinen Unterraumes $A = a + U$ ist die Dimension seines Translationsraumes: $\dim A = \dim U$.

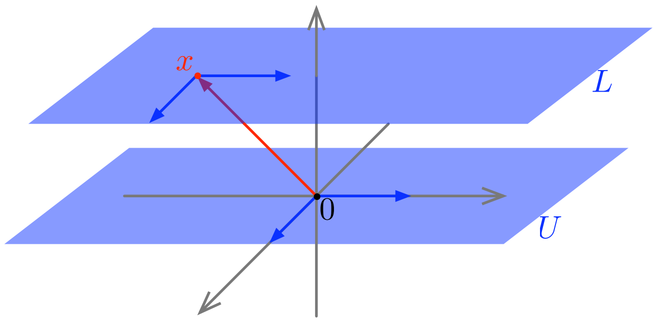
Ein 0-dimensionaler affiner Unterraum heißt **Punkt**. Punkte sind also genau die Mengen $x+\{0\} = \{x\}$, $x \in V$. Ein 1-dimensionaler affiner Unterraum heißt **Gerade**, und ein 2-dimensionalen affiner Unterraum heißt **Ebene**. Ist $\dim A = n-1$, so heißt A **Hyperebene**.

Den Punkt $a \in A = a+U$ nennt man auch **Fußpunkt** oder im Fall $A = a+V$ auch **Ursprung**. Neben a sind dann auch alle Punkte der Form $a + u, u \in U$ Fußpunkte (a ist also im Gegensatz zu U nicht eindeutig).

Beispiel 21.8 1. Sei $\mathbb{A}^n := \mathbb{A}(\mathbb{R}^n)$ der zu $\mathbb{R}^n$ gehörige affine Standard-Raum. Für einen Untervektorraum $U \subset \mathbb{R}^n$ und einen (festen) Punkt $x \in \mathbb{R}^n$ ist

$$A := x + U = \{x + u \mid u \in U\}$$

ein affiner Unterraum von $\mathbb{A}^n$.



2. Ist $\Phi : \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung und $y = \Phi(x) \in \operatorname{Bild}\Phi \subset \mathbb{R}^n$, so ist $\Phi^{-1}(y) = \{v \in \mathbb{R}^m \mid \Phi(v) = y\}$ ein affiner Unterraum von $\mathbb{A}^m$. Denn nach Hilfssatz 9.13 ist $\Phi^{-1}(y) = x + \operatorname{Kern}\Phi$.

3. Die Lösungsmenge $\mathfrak{L}$ eines linearen Gleichungssystems mit n Unbekannten ist entweder leer oder ein affiner Unterraum von $\mathbb{A}^n$. Der zugehörige Untervektorraum (Translationsraum) ist die Lösungsmenge $\mathfrak{L}_h$ des entsprechenden homogenen LGS (siehe 11.2).

4. Manchmal ist es zweckmäßig $\mathbb{A}(\mathbb{R}^n)$ als affinen Unterraum von $\mathbb{A}(\mathbb{R}^{n+1})$ aufzufassen:

$$\mathbb{A}(\mathbb{R}^n) = \{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{pmatrix} \mid x_i \in \mathbb{R}\} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + \{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 0 \end{pmatrix} \mid x_i \in \mathbb{R}\} \subset \mathbb{R}^{n+1}.$$

Hilfssatz 21.9 *Es sei $\mathcal{A}$ ein nichtleeres System affiner Unterräume von $\mathbb{A}(V)$. Dann ist der Schnitt*

$$M := \bigcap_{A\in\mathcal{A}} A$$

entweder leer oder ein affiner Unterraum von $\mathbb{A}(V)$ mit Translationsraum

$$U_M := \bigcap_{A\in\mathcal{A}} U_A,$$

wobei U_A jeweils den Translationsraum von A bezeichnet.

Beweis: Sei $M \neq \emptyset$. Dann gibt es einen Punkt $x \in V$, der in allen affinen Unterräumen $A \in \mathcal{A}$ liegt. Damit kann man jedes Element $A \in \mathcal{A}$ in der Form $A = x + U_A$ darstellen. Es folgt

$$M = \bigcap_{A\in\mathcal{A}} A = \bigcap_{A\in\mathcal{A}} x + U_A = x + \bigcap_{A\in\mathcal{A}} U_A = x + U_M.$$

Als Durchschnitt von Untervektorräumen von V ist U_M auch ein Untervektorraum von V und $x + U_M$ ein affiner Unterraum. ■

Beispiel 21.10 Im affinen Raum $\mathbb{A}^4 := \mathbb{A}(\mathbb{R}^4)$ seien die affinen Unterräume

$$A_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \left[\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right], \quad A_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \left[\begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right]$$

gegeben. Wir wollen ihren Schnitt bestimmen: $x \in A_1 \cap A_2$ ist äquivalent zur Existenz von $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2 \in \mathbb{R}$ mit

$$\begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = x = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta_1 \begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Wir erhalten ein inhomogenes lineares Gleichungssystem mit den Unbekannten α_1, α_2, α_3, β_1, β_2 und der zugehörigen erweiterten Matrix

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

Der Gauß-Algorithmus liefert folgende Stufenform:

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

Daraus lesen wir ab: $\beta_2 = 1$, $\beta_1 \in \mathbb{R}$ beliebig. Somit gilt

$$x = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta_1 \begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \beta_1 \begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix}.$$

Satz 21.11 (Intrinsische Charakterisierung eines affinen Unterraumes) *Es sei $\mathbb{A} = \mathbb{A}(V)$ der zu einem Vektorraum V gehörige affine Raum. Eine nichtleere Teilmenge $A \subset V$ ist genau dann ein affiner Unterraum von V, wenn für $k \geq 2$ gilt*

$$\forall a_1, \ldots,\ a_k \in A,\ \forall\ \lambda_1, \ldots, \lambda_k \in \mathbb{K} \ \ \textit{mit} \ \ \sum_{i=1}^{k} \lambda_i = 1 \Longrightarrow \sum_{i=1}^{k} \lambda_i a_i \in A.$$

Beweis: "$\Rightarrow$": Sei $A = x + U, x \in A$, ein affiner Unterraum. Für $i = 1, \ldots, k$ seien $a_i \in A$ und $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^{k} \lambda_i = 1$. Dann ist $a_i = x + u_i$ für $u_i \in U$ und $i = 1, \ldots, k$. Also

$$\sum_{i=1}^{k} \lambda_i a_i = \sum_{i=1}^{k} \lambda_i (x + u_i) = (\sum_{i=1}^{k} \lambda_i) x + \sum_{i=1}^{k} \lambda_i u_i = x + \sum_{i=1}^{k} \lambda_i u_i \in x + U = A.$$

"$\Leftarrow$": Wir wählen ein $x \in A$ und definieren $U := A - x := \{a - x \in V \mid a \in A\}$. Zu zeigen ist, dass U ein Untervektorraum von V ist. Seien also $a - x, b - x \in U$ und $\lambda \in \mathbb{K}$. Wir haben

$$(a - x) + (b - x) \in U \Longleftrightarrow x + (a - x) + (b - x) = a + b - x \in A.$$

Die rechte Seite ist aber nach Voraussetzung erfüllt, da $1 + 1 - 1 = 1$. Ebenso ist $\lambda(a - x) \in U$, da $x + \lambda(a - x) = \lambda a + (1 - \lambda) x \in A$. ■

Definition 21.12 Sei $\mathbb{A} = \mathbb{A}(V)$ der zu einem Vektorraum V gehörige affine Raum. Sei $X \subset V$ eine beliebige Teilmenge. Die **affine Hülle** $\mathrm{Aff}\,(X)$ von X ist

$$\mathrm{Aff}\,(X) := \{\sum_{i=1}^{k} \lambda_i p_i \mid p_i \in X,\ \lambda_i \in \mathbb{K},\ \sum_{i=1}^{k} \lambda_i = 1\}.$$

Hilfssatz 21.13 *$\mathrm{Aff}(X)$ ist der kleinste affine Unterraum von $\mathbb{A}$, der X enthält.*

Beweis: Dass $\mathrm{Aff}\,(X)$ ein affiner Unterraum ist, folgt aus Satz 21.11. Die zweite Behauptung folgt mit Hilfssatz 21.9. ■

Definition 21.14 Sei $\mathbb{A} = \mathbb{A}(V)$ der zu einem Vektorraum V gehörige affine Raum. Die Elemente einer Menge $X = \{p_0, p_1, \ldots, p_r\} \subset V$ heißen **affin unabhängig** oder **in allgemeiner Lage**, falls gilt

$$\sum_{i=0}^{r} \lambda_i p_i = 0 \ \text{ und } \ \sum_{i=0}^{r} \lambda_i = 0 \Longleftrightarrow \lambda_0 = \lambda_1 = \cdots = \lambda_r = 0.$$

Hilfssatz 21.15 *Eine endliche Menge $\{p_0, p_1, \ldots, p_r\} \subset V$ ist affin unabhängig genau dann, wenn $\{p_1 - p_0, \ldots, p_r - p_0\}$ linear unabhängig ist im Vektorraum V.*

Beweis: Sei $\{p_0, p_1, \ldots, p_r\}$ affin unabhängig. Ist $\sum_{i=1}^{r} \alpha_i(p_i - p_0) = 0$, so folgt $\sum_{i=1}^{r} \alpha_i p_i + (-\sum_{i=1}^{r} \alpha_i)p_0 = 0$ und damit $\alpha_i = 0$ für alle $1 \leq i \leq r$.

Ist umgekehrt $\{p_1 - p_0, \ldots, p_r - p_0\}$ linear unabhängig, so folgt aus

$$\sum_{i=0}^{r} \lambda_i p_i = 0 \ \text{ und } \ \sum_{i=0}^{r} \lambda_i = 0, \ \text{dass}$$

$$0 = \sum_{i=1}^{r} \lambda_i p_i + \lambda_0 p_0 = \sum_{i=1}^{r} \lambda_i p_i + (-\sum_{i=1}^{r} \lambda_i)p_0 = \sum_{i=1}^{r} \lambda_i(p_i - p_0)$$

und damit nach Voraussetzung $\lambda_i = 0$ für alle i. ■

Definition 21.16 Sei $\mathbb{A} = \mathbb{A}(V)$ der zu einem Vektorraum V gehörige affine Raum. Weiter seien A ein k-dimensionaler affiner Unterraum von V und $\{p_0, p_1, \ldots, p_k\}$ affin unabhängige Punkte in A. Dann heißt $\{p_0, p_1, \ldots, p_k\}$ eine **affine Basis** von A und $(p_0; p_1 - p_0, \ldots, p_k - p_0)$ ein **affines Koordinatensystem** von A mit **Ursprung** p_0.

Für einen Punkt $a \in A$ können wir dann schreiben

$$a = p_0 + \sum_{i=1}^{k} \alpha_i(p_i - p_0), \quad \alpha_i \in \mathbb{K}.$$

Die Komponenenten $(\alpha_1, \ldots, \alpha_k)$ heißen **affine Koordinaten** von $a \in A$ bezüglich der affinen Basis $\{p_0, p_1, \ldots, p_k\}$.

Im Gegensatz zum Vektorraum V erhalten wir also im affinen Raum $\mathbb{A}(V)$ für jeden Punkt p und jede Basis B von V eine affine Basis $(p; B)$ von V. Ein Vektorraum ist also "dasselbe" wie ein affiner Raum mit einem ausgezeichneten Ursprung 0.

Bemerkung 21.17 Nach Hilfsatz 21.15 kann man ein affines Koordinatensystem auch wie folgt konstruieren. Es sei $A = x+U$ ein k-dimensionaler affiner Unterraum. Ist $\{u_1, \ldots, u_k\}$ eine Basis von U, so lässt sich jeder Punkt $p \in A$ in der Form

$$p = x + \alpha_1 u_1 + \ldots + \alpha_k u_k, \quad (\alpha_i \in \mathbb{K})$$

darstellen. Setzt man $x_i := x+u_i$ für $i = 1, \ldots k$, so ist $(x, x_1, \ldots, x_k)$ eine affine Basis von A und die affinen Koordinaten von p bezüglich dieser Basis sind $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k$.

21.4 Affine Abbildungen

Wir suchen Abbildungen zwischen affinen Räumen, die strukturerhaltend sind, d.h. zum Beispiel affine Hüllen in affine Hüllen abbilden. Wir definieren deshalb

Definition 21.18 Es seien V und W $\mathbb{K}$-Vektorräume und $\mathbb{A}(V)$ und $\mathbb{A}(W)$ die zugehörigen affinen Räume. Eine Abbildung $f : V \to W$ heißt **affin**, falls für alle $a, b \in V$ und alle $\lambda, \mu \in \mathbb{K}$ mit $\lambda + \mu = 1$ gilt

$$f(\lambda a + \mu b) = \lambda f(a) + \mu f(b).$$

Bemerkung 21.19 Mit vollständiger Induktion folgt:

Eine Abbildung $f : V \to W$ ist affin genau dann, wenn für $k \geq 2$, alle $a_1, \ldots,\ a_k \in V$ und alle $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ mit $\sum_{i=1}^n \lambda_i = 1$ gilt

$$f(\sum_{i=1}^k \lambda_i a_i) = \sum_{i=1}^k \lambda_i f(a_i).$$

Beispiel 21.20 Sei $\mathbb{A}(V)$ der zu einem $\mathbb{K}$-Vektorraum V gehörige affine Raum und $b \in V$.

1. Lineare Abbildungen sind affine Abbildungen.

2. Die **Translation**

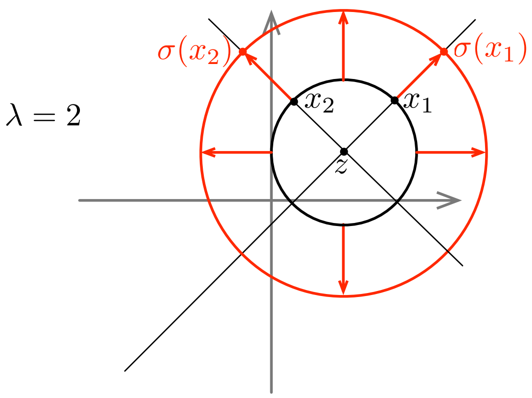
   $$T_b : V \to V; \quad T_b(x) := x + b$$

   ist eine affine Abbildung. Denn für $\lambda, \mu \in \mathbb{K}$ mit $\lambda + \mu = 1$ gilt

   $$T_b(\lambda x + \mu y) = \lambda x + \mu y + b = \lambda x + \mu y + (\lambda + \mu)b = \lambda T_b(x) + \mu T_b(y).$$

   Eine Translation ist nur für $b = 0$ eine lineare Abbildung.

3. Eine **Streckung** mit **Zentrum** $z \in V$ und **Streckungsfaktor** λ bildet z auf sich ab und ordnet jedem von z verschiedenen Punkt $x \in V$ denjenigen Punkt $y \in V$ auf der Gerade durch x und z zu, für den $y - z = \lambda(x - z)$ gilt. Somit hat eine solche Streckung $\sigma : V \to V$ die Darstellung $\sigma(x) = y = \lambda x + (1 - \lambda)z$, ist also eine affine Abbildung.



Für $\lambda = 0$ bildet σ alle Punkte auf z ab, für $\lambda = 1$ ist σ die Identität.

4. Für $\lambda = -1$ gilt für die obige Streckung $z = \frac{1}{2}(x + \sigma(x))$, d.h. z ist der Mittelpunkt der Punkte x und $\sigma(x)$. In diesem Fall ist σ eine **Punktspiegelung** an z.

Satz 21.21 (Allgemeine Form von affinen Abbildungen) *Sei $\mathbb{A}(V)$ der zu einem $\mathbb{K}$-Vektorraum V gehörige affine Raum. Wir wählen einen Ursprung $p_0 \in V$.*

Eine Abildung $f : V \to W$ ist affin genau dann, wenn eine lineare Abbildung $\Phi : V \to W$ existiert, so dass gilt

$$f(p_0 + v) = \Phi(v) + f(p_0).$$

Die lineare Abbildung Φ ist dabei durch f eindeutig bestimmt und unabhängig von der Wahl des Urprungs p_0. Sie heißt der **lineare Anteil** *von f.*

Beweis: "$\Leftarrow$: Ist Φ linear, so ist folgt für $p = p_0 + v, q = p_0 + w \in V$ und $\lambda, \mu \in \mathbb{K}$ mit $\lambda + \mu = 1$

$$f(\lambda p + \mu q) = \Phi(\lambda v + \mu w) + f(p_0) = \lambda(\Phi(v) + f(p_0)) + \mu(\Phi(w) + f(p_0)) = \lambda f(p) + \mu f(q).$$

"$\Rightarrow$: Wir definieren

$$\Phi : V \to W;\ \Phi(v) := f(p_0 + v) - f(p_0).$$

Zu zeigen ist, dass Φ linear ist, d.h. dass für alle $u, v \in V$ und alle $\lambda \in \mathbb{K}$ gilt:

$$\Phi(u+v) = \Phi(u) + \Phi(v), \quad \Phi(\lambda u) = \lambda\Phi(u).$$

Zunächst ist $p_0 + u, p_0 + v \in V$ und auch $p_0 + u + v = (p_0 + u) + (p_0 + v) - p_0 \in V$. Also haben wir mit 21.19

$$\begin{aligned}
\Phi(u+v) &= f(p_0 + u + v) - f(p_0) \\
&= f(p_0 + u) + f(p_0 + v) - f(p_0) - f(p_0) = \Phi(u) + \Phi(v); \\
\Phi(\lambda u) &= f(p_0 + \lambda u) - f(p_0) = f(\lambda(p_0 + u) + (1-\lambda)p_0) - f(p_0) \\
&= \lambda f(p_0 + u) + (1-\lambda) f(p_0) - f(p_0) = \lambda\Phi(u).
\end{aligned}$$

Zur Eindeutigkeit: Sei $p_1 \in V$ ein anderer Ursprung des affinen Raumes. Wir definieren dann wie oben die lineare Abbildung

$$\Psi : V \to W; \quad \Psi(v) := f(p_1 + v) - f(p_1).$$

Wir zeigen, dass $\Psi = \Phi$. Für beliebiges $v \in V$ ist

$$\begin{aligned}
\Psi(v) &= f(p_1 + v) - f(p_1) = f(p_0 + (p_1 - p_0) + v) - f(p_1) \\
&= \Phi((p_1 - p_0) + v) + f(p_0) - f(p_1) = \Phi(v) + \Phi(p_1 - p_0) + f(p_0) - f(p_1) \\
&= \Phi(v) + f(p_1) - f(p_0) + f(p_0) - f(p_1) = \Phi(v).
\end{aligned}$$

■

Bemerkung 21.22 Nach Satz 21.21 ist die Eigenschaft "affine Abbildung", unabhängig vom Ursprung. Wir können also insbesondere den Nullvektor 0 als Ursprung wählen. Wir erhalten dann eine besonders übersichtliche Darstellung der affinen Abbildungen zwischen V und W.

Jede affine Abbildung $f : V \to W$ kann man schreiben als

$$f = T_b \circ \Phi,$$

wobei die lineare Abbildung Φ und der Translationsvektor $b := f(0)$ durch f eindeutig bestimmt sind.

Speziell sind die affinen Abbildungen zwischen den affinen Standardräumen $\mathbb{A}(\mathbb{K}^n)$ und $\mathbb{A}(\mathbb{K}^m)$ von der Form

$$f(x) = Ax + b, \quad \text{wobei} \quad A \in \mathbb{K}^{m\times n}, \ b \in \mathbb{K}^n.$$

Wir benutzen jetzt noch affine Koordinatensysteme, um affine Selbstabbildungen von $\mathbb{A}(V)$ mittels Matrizen zu beschreiben.

Sei $B := \{v_1, \ldots, v_n\}$ eine Basis des Vektorraumes V und p_0 ein Ursprung in V. Diese Wahl definiert die affine Basis $\{p_0; B\}$ von V. Für $x \in V$ haben wir dann die Darstellung

$$x = p_0 + \sum_{j=1}^{n} x_j v_j,$$

d.h. x hat bezüglich $\{p_0; B\}$ die affinen Koordinaten $\hat{x} := (x_1, \ldots, x_n)^\top \in \mathbb{K}^n$. Weiter sei $A := (a_{ij})$ die Abbildungsmatrix des linearen Anteils Φ einer affinen Abbildung $f : V \to V$ bezüglich der Basis B. Schließlich sei $f(p_0) = p_0 + \sum_{i=1}^{n} b_i v_i$, d.h. $f(p_0)$ hat die affinen Koordinaten $\hat{b} := (b_1, \ldots, b_n)^\top$. Dann haben wir

$$f(x) = \Phi(\sum_{i=j}^{n} x_j v_j) + f(p_0) = p_0 + \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij} x_j + b_i) v_i.$$

In affinen Koordinaten bezüglich der affinen Basis $\{p_0; B\}$ wird die affine Selbstabbildung f also dargestellt durch

$$\hat{f} : \mathbb{K}^n \to \mathbb{K}^n; \quad \hat{x} \mapsto A\hat{x} + \hat{b}.$$

Definition 21.23 Eine *bijektive* affine Selbstabbildung $f : V \to V$ heißt **Affinität**.

Bemerkung 21.24

1. Translationen sind bijektiv: $(T_b)^{-1} = T_{-b}$.
2. Eine affine Abbildung $f = T_b \circ \Phi$ ist genau dann injektiv (bzw. bijektiv), wenn der lineare Anteil Φ injektiv (bzw. bijektiv) ist.

Satz 21.25 *Die Affinitäten eines affinen Raumes* $\mathbb{A}(V)$ *bilden bezüglich der Verkettung von Abbildungen eine Gruppe* $\mathrm{Aff}(\mathbb{A}(V))$.

Beweis: Wir wählen den Nullvektor als Ursprung von V. Dann haben Affinitäten die Form $f = T_b \circ \Phi$, wobei der lineare Anteil $\Phi : V \to V$ ein Vektorraum-Isomorphismus ist. Es gilt $\mathrm{id}_V \in \mathrm{Aff}(\mathbb{A}(V))$, da die Identität affin und bijektiv ist. Sind $f = T_b \circ \Phi$ und $g = T_c \circ \Psi$ Affinitäten, so ist auch $f \circ g \in \mathrm{Aff}(\mathbb{A}(V))$:

$$f \circ g = T_b \circ (\Phi \circ T_c) \circ \Psi = T_b \circ (T_{\Phi(c)} \circ \Phi) \circ \Psi = T_{b+\Phi(c)} \circ (\Phi \circ \Psi).$$

Ist $f = T_b \circ \Phi \in \mathrm{Aff}(\mathbb{A}(V))$ so ist auch $f^{-1} \in \mathrm{Aff}(\mathbb{A}(V))$, denn mit Bemerkung 21.24 gilt:

$$f^{-1} = \Phi^{-1} \circ T_b^{-1} = \Phi^{-1} \circ T_{-b} = T_{\Phi^{-1}(-b)} \circ \Phi^{-1}.$$

∎

Bemerkung 21.26 Die affinen Selbstabbildungen von $\mathbb{A}^n = \mathbb{A}(\mathbb{R}^n)$ sind nach Obigem alle von der Form

$$f : \mathbb{R}^n \to \mathbb{R}^n; \quad x \mapsto Ax + b \qquad (A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n).$$

Beschreiben wir $\mathbb{A}^n$ als affinen Unterraum von $\mathbb{R}^{n+1}$ (siehe Beispiel 21.8 4.), so können wir affine Selbstabbildungen auch durch $(n+1)\times(n+1)$-Matrizen beschreiben. Für eine affine Abbildung $f(x) = Ax + b$ definieren wir

$$\widetilde{A} := \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Dann haben wir

$$\widetilde{f} : \mathbb{A}^n \to \mathbb{A}^n; \begin{pmatrix} x \\ 1 \end{pmatrix} \mapsto \widetilde{A} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Ax + b \\ 1 \end{pmatrix}.$$

21.5 Affine Invarianten

In diesem Abschnitt ist $\mathbb{A}(V)$ der zu einem $\mathbb{K}$-Vektorraum V gehörige affine Raum. Wir geben einige Eigenschaften, Begriffe oder Größen an, die bei affinen Abbildungen $f : V \to V$ invariant bleiben (also Gegenstände der affinen Geometrie sind).

Aus der Definition von affinen Abbildungen folgt: Affine Unterräume werden durch affine Abbildungen wieder in affine Unteräume abgebildet. Insbesondere sind affine Bilder von Geraden wieder Geraden. Affine Abbildungen sind also **geradentreu**.

Definition 21.27 Zwei affine Unterräume A_1 und A_2 heißen **parallel**, geschrieben $A_1 \| A_2$, falls für die zugehörigen Translationsräume U_1, U_2 gilt: $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$.

Es seien p, q, r drei kollineare Punkte (d.h. Punkte, die auf einer Geraden liegen) und $p \neq q$. Ist $q = p + v$ und $r = p + tv$, so heißt die Zahl $t \in \mathbb{K}$ das **Teilverhältnis** von r bezüglich p und q.

Bemerkung 21.28 "Parallelität" ist eine reflexive und symmetrische Relation, aber im allgemeinen nicht transitiv. So sind zwei sich schneidende Geraden in einer Ebene zwar zu dieser Ebene parallel, aber sie selbst sind nicht parallel. Beschränkt man sich jedoch auf affine Unterräume *gleicher* Dimension, so ist Parallelität auch transitiv, also eine Äquivalenzrelation.

Hilfssatz 21.29

1. *Affine Abbildungen bilden parallele Unterräume auf parallele Unterräume ab, sie sind* **parallelentreu**.
2. *Affinitäten lassen das Teilverhältnis invariant, sie sind* **teilverhältnistreu**.

Beweis: 1. Sei $f : V \to V$ eine affine Abbildung und Φ der lineare Anteil von f. Der Translationsraum des Bildes ist gleich dem Bild des Translationsraumes unter Φ: Seien $A_1 = x_1 + U_1$ und $A_2 = x_2 + U_2$ parallel. Dann ist $f(A_1) = \Phi(U_1) + f(x_1)$ und $f(A_2) = \Phi(U_2) + f(x_2)$. Daraus folgt die Behauptung.

2. Sei $f : V \to V$ eine Affinität und Φ der lineare Anteil von f. Sind $p, q = p+v$ und $r = p + tv$ kollinear, so sind auch $f(p), f(q), f(r)$ kollinear und $f(p) \neq f(q)$. Weiter ist $f(q) = \Phi(v) + f(p)$ und $f(r) = t\Phi(v) + f(p)$. Damit folgt die Behauptung. ■

Bemerkung 21.30 Es gilt die folgende Charakterisierung von Affinitäten:

Eine bijektive Abbildung $f : V \to V$ ist genau dann eine Affinität, wenn sie geradentreu, parallelentreu und teilverhältnistreu ist.

21.6 Euklidische Isometrien

Es sei V ein Vektrorraum über den reellen Zahlen $\mathbb{R}$.

Im vorherigen Abschnitt haben wir die Geometrie des affinen Raumes $\mathbb{A}(V) = (V, V, +)$ betrachtet. Jetzt versehen wir V zusätzlich mit einem Skalarprodukt: $(V, \langle , \rangle)$ ist dann ein euklidischer Vektorraum und damit insbesondere auch ein metrischer Raum (siehe Abschnitt 18.3.2):

Wir wiederholen die Definition: Die **euklidische Abstandsfunktion** oder **euklidische Metrik** auf dem euklidischen Vektorraum $(V, \langle , \rangle)$ ist definiert durch:

$$d : V \times V \to \mathbb{R};\ d(x, y) := \|x - y\| = \sqrt{\langle x - y, x - y\rangle}.$$

Definition 21.31 Eine **euklidische Isometrie** oder **Bewegung** ist eine Isometrie des metrischen Raumes $\mathbb{E}(V) := (V, d)$, d.h. eine bijektive Abbildung $f : V \to V$ für die gilt:

$$d(f(x), f(y)) = d(x, y) \quad \text{für alle } x, y \in V.$$

Bemerkung 21.32 Für Vektorräume mit Skalarprodukt $(V, \langle , \rangle)$ hatten wir *lineare* Isometrien definiert als lineare Abbildungen $\Phi : V \to V$, für die gilt

$$\langle \Phi(x), \Phi(y)\rangle = \langle x, y\rangle \quad \text{für alle } x, y \in V.$$

Da die euklidische Metrik d auf V mittels des Skalarproduktes definiert ist, sind also lineare Isometrien von $(V, \langle , \rangle)$ spezielle Beispiele von euklidischen Isometrien.

Satz 21.33 *Die euklidischen Isometrien von $\mathbb{E}(V) = (V, d)$ bilden bezüglich der Verkettung von Abbildungen eine Gruppe $\mathcal{I}so(\mathbb{E}(V))$.*

Beispiel 21.34 1. Lineare Isometrien sind euklidische Isometrien, die 0 auf 0 abbilden und die orthogonale Matrizen ($A^\top = A^{-1}$) als Abbildungsmatrizen bezüglich Orthonormalbasen haben.

2. Translationen sind *nicht-lineare* euklidische Isometrien.

3. Verkettungen von Translationen und linearen Isometrien sind euklidische Isometrien.

Der folgende Satz besagt, dass es keine weiteren euklidischen Isometrien gibt (vergleiche dazu den entsprechende Aussage 21.22 für affine Abbildungen).

Satz 21.35 (Struktur von euklidischen Isometrien) *Jede euklidische Isometrie $f : \mathbb{E}(V) \to \mathbb{E}(V)$ ist von der Form $f = T_b \circ \Phi$, wobei Φ eine lineare Isometrie ist. Sowohl Φ als auch der Translationsvektor b sind durch f eindeutig bestimmt.*

Um Satz 21.35 zu beweisen, benötigen wir den folgenden Hilfssatz.

Hilfssatz 21.36 *Seien $a_0, a_1, \ldots, a_n$ affin unabhängige Punkte in $\mathbb{E}(V)$. Eine Isometrie $f : \mathbb{E}(V) \to \mathbb{E}(V)$ ist eindeutig bestimmt durch die Bilder $f(a_0), f(a_1), \ldots, f(a_n)$.*

Beweis: Seien f und g euklidische Isometrien mit $f(a_i) = g(a_i)$ für $0 \le i \le n$. Dann ist $g^{-1}f$ eine Isometrie mit $g^{-1}f(a_i) = a_i$ für alle i. Wir setzen $b_i := T_{-a_0}(a_i)$ für $0 \le i \le n$. Dann ist $b_0 = 0$ und $\{b_1, \ldots, b_n\}$ ist eine Basis von V. Wir werden zeigen, dass $h := T_{-a_0} g^{-1} f T_{-a_0}^{-1}$ die Identität ist. Insbesondere folgt dann die Behauptung $f = g$.

Zunächst ist nach Konstruktion $h(0) = 0$ und $h(b_i) = b_i$ für alle i. Da h eine Isometrie ist, folgt daraus für $x \in V$ und $y := h(x)$ sowie alle $1 \le i \le n$

$$d(0, x) = d(h(0), h(x)) = d(0, y) \quad \text{und} \quad d(b_i, x) = d(h(b_i), h(x)) = d(b_i, y).$$

Das ist äquivalent zu

$$\langle x, x \rangle = \langle y, y \rangle \quad \text{und} \quad \langle x - b_i, x - b_i \rangle = \langle y - b_i, y - b_i \rangle \quad \forall\, 1 \le i \le n.$$

Durch "Ausmultiplizieren" erhält man aus den letzten n Gleichungen

$$\langle x, b_i \rangle = \langle y, b_i \rangle \quad \forall\, 1 \le i \le n,$$

und deshalb, da $\{b_1, \ldots, b_n\}$ eine Basis ist, $\langle x, z \rangle = \langle y, z \rangle \ \forall z \in V$. Also $x = y = h(x)$ und h ist die Identität. ■

Beweis von Satz 21.35: Wir wählen eine Basis $\{a_1, \ldots, a_n\}$ von V. Die Punkte $\{a_0 := 0, a_1, \ldots, a_n\}$ sind dann affin unabhängig und bilden somit eine affine Basis.

Sei $b_0 := f(a_0) = f(0)$ und $g := T_{-b_0} \circ f$. Es ist dann $g(0) = 0$. Weiter setzen wir $b_i := g(a_i)$ für $1 \leq i \leq n$. Wir konstruieren im Folgenden eine *lineare* Isometrie Φ mit den gleichen Bildern der Basis wie g, so dass mit Hilfsatz 21.36 folgt, dass $T_{-b_0} \circ f = g = \Phi$, was dann die Behauptung des Satzes beweist.

Da g eine Isometrie ist, haben wir zunächst für alle $1 \leq i, j \leq n$:

$$d(0, a_i) = d(g(0), g(a_i) = d(0, b_i) \quad \text{und} \quad d(a_i, a_j) = d(g(a_i), g(a_j)) = d(b_i, b_j).$$

Mit der Definition von d durch Norm bzw. Skalarprodukt folgt daraus insbesondere

$$\langle a_i, a_j \rangle = \langle b_i, b_j \rangle \quad \text{für alle} \;\; 1 \leq i, j \leq n. \qquad (*)$$

Sei nun Φ die eindeutige lineare Abbildung mit $\Phi(a_i) = b_i$ für $i = 1, \ldots, n$. Wir zeigen, dass Φ eine euklidische Isometrie ist. Seien dazu $x, y \in \mathbb{E}(V)$ beliebig. Wir schreiben $x - y = \sum_{i=1}^n \lambda_i a_i$. Dann gilt $\Phi(x) - \Phi(y) = \Phi(x - y) = \sum_{i=1}^n \lambda_i b_i$ und wir haben mit $(*)$

$$\begin{aligned} d(\Phi(x), \Phi(y))^2 &= \|\Phi(x) - \Phi(y)\|^2 = \sum_{i,j=1}^n \lambda_i \lambda_j \langle b_i, b_j \rangle = \sum_{i,j=1}^n \lambda_i \lambda_j \langle a_i, a_j \rangle \\ &= \|x - y\|^2 = d(x, y)^2, \end{aligned}$$

d.h. die lineare Abbildung Φ ist eine Isometrie. ■

Zusammenfassend haben wir folgende Situation:

$\mathcal{I}so(\mathbb{E}(V))$	$\subsetneq$	$\mathrm{Aff}(\mathbb{A}(V))$
$f = T_b \circ \Phi$ Φ lineare Isometrie		$f = T_b \circ \Phi$ Φ invertierbare lineare Abbildung
Abbildungsmatrix $A \in O(n)$		Abbildungsmatrix $A \in \mathbf{GL}(n, \mathbb{R})$

Bemerkung 21.37 Euklidische Isometrien heißen auch **Bewegungen**. Ist $f = T_b \circ \Phi$ eine Bewegung so heißt f **eigentlich** bzw. **uneigentlich**, falls $\det \Phi = 1$ bzw. $\det \Phi = -1$.

21.7 Quadriken

In diesem Abschnitt benutzen wir frühere Ergebnisse, um ein geometrisches Klassifikationsproblem zu lösen.

Wir betrachten den Vektorraum $V = \mathbb{R}^n$ versehen mit dem Standard-Skalarprodukt $\langle , \rangle$. Entsprechend haben wir dann den affinen Raum $\mathbb{A}^n = \mathbb{A}(\mathbb{R}^n)$ und den euklidischen Raum $\mathbb{E}^n = \mathbb{E}(\mathbb{R}^n)$.

Definition 21.38 Sei Q eine quadratische Form im Vektorraum $\mathbb{R}^n, b \in \mathbb{R}^n$ und $\delta \in \mathbb{R}$. Eine **Quadrik** in $\mathbb{R}^n$ (oder **Fläche 2. Grades**) ist eine Menge der Form

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid \ Q(x) + 2\langle x, b\rangle = \delta\}.$$

In affinen Koordinaten bezüglich der affinen Basis $\{0, e_1, \ldots, e_n\}$ haben wir

$$\mathcal{F} = \{(x_1, \ldots, x_n)^\top \in \mathbb{R}^n \mid \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + 2 \sum_{k=1}^n b_k x_k = \delta\},$$

wobei $A = (a_{ij})$ eine reelle symmetrische $(n \times n)$-Matrix ist, $(b_1, \ldots, b_n) \in \mathbb{R}^n$ und $\delta \in \mathbb{R}$.

Sei $Q : V \to \mathbb{R}$ die zu A gehörige quadratische Form und $\Phi : V \to V$ der zu Q (bzw. A) gehörige selbstadjungierte Endomorphismus. Setzen wir $b := (b_1, \ldots, b_n)$ und $x := (x_1, \ldots, x_n)$, so wird $\mathcal{F}$ durch folgende Gleichungen beschrieben:

$$x^\top A x + 2b^\top x = \delta, \quad Q(x) + 2\langle b, x\rangle = \delta \quad \text{oder auch} \quad \langle x, \Phi(x) + 2b\rangle = \delta.$$

Eine weitere nützliche Schreibweise für Quadriken ergibt sich, wenn man $\mathbb{R}^n$ als affinen Unterraum von $\mathbb{R}^{n+1}$ auffasst (siehe Beispiel 21.8). Für die Quadrik $\mathcal{F}$ definieren wir die **erweiterte Matrix**

$$\widetilde{A} := \begin{pmatrix} A & b \\ b^\top & -\delta \end{pmatrix}. \tag{21.1}$$

Wir haben dann

$$x \in \mathcal{F} \Longleftrightarrow \begin{pmatrix} x^\top & 1 \end{pmatrix} \widetilde{A} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0.$$

Bemerkung 21.39 Q, b und δ sind durch $\mathcal{F}$ nur bis auf einen gemeinsamen, konstanten Faktor bestimmt.

Beispiel 21.40 Kegelschnitte oder Quadriken in $\mathbb{R}^2$.

1. **Ellipsen**: $\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid a_1^2 x_1^2 + a_2^2 x_2^2 = \delta^2\}$.

2. **Hyperbeln**: $\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid a_1^2 x_1^2 - a_2^2 x_2^2 = \delta^2\}$.

3. **Parabeln**: $\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid a_1^2 x_1^2 - x_2 = 0\}$.

Hilfssatz 21.41 (Quadrik als Begriff der affinen Geometrie) *Das Bild einer Quadrik unter einer Affinität ist wieder eine Quadrik.*

Beweis: Sei $f \in \mathcal{A}ff(\mathbb{A}^n)$. Wir fassen $\mathbb{R}^n$ als affinen Unterraum von $\mathbb{R}^{n+1}$ auf. Dann können wir $x = f^{-1}(y)$ schreiben als

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} M & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix},$$

wobei $M \in \mathbf{GL}(n, \mathbb{R})$ und $c \in \mathbb{R}^n$ ist. Somit haben wir

$$\begin{aligned} x \in \mathcal{F} \quad &\iff \quad 0 = \begin{pmatrix} x^\top & 1 \end{pmatrix} \widetilde{A} \begin{pmatrix} x \\ 1 \end{pmatrix} \\ y = f(x) \in f(\mathcal{F}) \quad &\iff \quad 0 = \begin{pmatrix} y^\top & 1 \end{pmatrix} \begin{pmatrix} M^\top & 0 \\ c^\top & 1 \end{pmatrix} \begin{pmatrix} A & b \\ b^\top & -\delta \end{pmatrix} \begin{pmatrix} M & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix}. \end{aligned}$$

Weiter ist

$$\begin{pmatrix} M^\top & 0 \\ c^\top & 1 \end{pmatrix} \begin{pmatrix} A & b \\ b^\top & -\delta \end{pmatrix} \begin{pmatrix} M & c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} M^\top A M & M^\top A c + M^\top b \\ c^\top A M + b^\top M & c^\top A c + c^\top b + b^\top c - \delta \end{pmatrix} \tag{21.2}$$

Diese Matrix kann man auffassen als erweiterte Matrix (21.1) einer Quadrik. Damit folgt, dass auch $f(\mathcal{F})$ eine Quadrik ist. ■

Mit dem letzten Ergebnis können wir definieren

Definition 21.42 Zwei Quadriken $\mathcal{F}_1$ und $\mathcal{F}_2$ in $\mathbb{R}^n$ sind genau dann **affin äquivalent**, $\mathcal{F}_1 \overset{a}{\sim} \mathcal{F}_2$, wenn eine Affinität $f \in \mathcal{A}ff(\mathbb{A}^n)$ existiert, so dass $\mathcal{F}_2 = f(\mathcal{F}_1)$. Weil $\mathcal{A}ff(\mathbb{A}^n)$ eine Gruppe ist, ist $\overset{a}{\sim}$ eine Äquivalenzrelation.

Zwei Quadriken $\mathcal{F}_1$ und $\mathcal{F}_2$ in $\mathbb{R}^n$ sind genau dann **euklidisch äquivalent**, $\mathcal{F}_1 \overset{e}{\sim} \mathcal{F}_2$, wenn eine euklidische Isometrie $f \in \mathcal{I}so(\mathbb{E}^n)$ existiert, so dass $\mathcal{F}_2 = f(\mathcal{F}_1)$. Weil $\mathcal{I}so(\mathbb{E}^n)$ eine Gruppe ist, ist $\overset{e}{\sim}$ eine Äquivalenzrelation.

Ziel der folgenden beiden Abschnitte ist es, Quadriken in $\mathbb{R}^3$ durch Repräsentanten (für die obigen Äquivalenzrelationen) zu klassifizieren.

21.7.1 Euklidische Klassifikation von Quadriken

Wir untersuchen zuerst, ob der lineare Teil durch eine Translation des Koordinatensystems "weggeschafft" werden kann.

Wir setzen $x = x' + c$ für $c = (c_1, \dots, c_n) \in \mathbb{R}^n$ (c ist dabei als Ortsvektor des Ursprungs des neuen (affinen) Koordinatensystems aufzufassen).

Wir formen dazu die Gleichung für $\mathcal{F}$ wie folgt um:

$$\begin{aligned}
& \delta = \langle x, \Phi(x) + 2b \rangle \\
\Leftrightarrow \quad & \delta = \langle x' + c, \Phi(x' + c) + 2b \rangle \\
\Leftrightarrow \quad & \delta = \langle x', \Phi(x') \rangle + \langle x', \Phi(c) \rangle + \langle x', 2b \rangle + \langle c, \Phi(x') \rangle + \langle c, \Phi(c) \rangle + \langle c, 2b \rangle \\
\Leftrightarrow \quad & \delta = \langle x', \Phi(x') \rangle + 2\langle x', \Phi(c) \rangle + 2\langle x', b \rangle + \langle c, \Phi(c) + 2b \rangle \\
\Leftrightarrow \quad & \langle x', \Phi(x') \rangle + 2\langle x', \Phi(c) + b \rangle = \delta - \langle c, \Phi(c) + 2b \rangle.
\end{aligned}$$

Der lineare Term verschwindet also genau dann, wenn $\Phi(c) = -b$.

Damit ergeben sich 2 FÄLLE:

FALL (I) $-b \in \operatorname{Bild} \Phi$. Das ist genau dann der Fall, wenn das LGS

$$\sum_{k=1}^{n} a_{ik} x_k = -b_i, \quad i = 1, 2, \dots, n \qquad (*)$$

mindestens eine Lösung besitzt.

FALL (II) $-b \notin \operatorname{Bild} \Phi$. Das ist genau dann der Fall, wenn das LGS $(*)$ keine Lösung besitzt (in diesem Fall muss A singulär sein).

Im Folgenden diskutieren wir diese beiden Fälle ausführlich für $\boxed{n = 3}$.

Der Fall (I): Flächen mit Zentrum

Ist c eine Lösung von $(*)$ und setzt man $x = x' + c$ so ist die Fläche durch die Gleichung $\langle x', \Phi(x') \rangle = \delta'$ gegeben. Dabei gilt $\delta' = \delta - \langle c, \Phi(c) + 2b \rangle = \delta - \langle c, b \rangle$. Insbesondere folgt, dass mit x' auch $-x'$ die Gleichung der Fläche erfüllt: Der Ursprung des neuen Koordinatensystems ist ein sogenanntes **Zentrum** der Fläche.

Nach dem Spektralsatz 19.9 existiert eine ONB $\{a_1, a_2, a_3\}$ aus Eigenvektoren der Abbildung Φ mit Eigenwerten $\lambda_1, \lambda_2, \lambda_3$. Bezüglich dieser Basis schreiben wir $x' = y_1 a_1 + y_2 a_2 + y_3 a_3$ (d.h. y_1, y_2, y_3 sind neue affine Koordinaten, die durch eine Translation und eine Rotation aus x_1, x_2, x_3 hervorgehen). Die Gleichung der Fläche lautet dann

$$\boxed{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 = \delta'.}$$

Wieder sind verschiedene Fälle zu unterscheiden.

I.1: Alle $\lambda_i \neq 0$, $i = 1, 2, 3.$

In diesem Fall ist A regulär und das LGS $(*)$ hat genau eine Lösung, also genau ein Zentrum.

I.1.1: $\delta' = 0$

- Haben alle λ_i gleiches Vorzeichen, so ist $\mathfrak{F}$ ein **Punkt**.
- Haben nicht alle λ_i gleiches Vorzeichen, also o.B.d.A. $\lambda_1 > 0, \lambda_2 > 0$ und $\lambda_3 < 0$, so ist $\mathfrak{F}$ ein **(Doppel-)Kegel** mit Spitze in 0 und der 3-Achse als Achse. Wenn $\lambda_1 = \lambda_2$, so ist $\mathfrak{F}$ ein **Kreis-Kegel**.

I.1.2: $\delta' \neq 0$ (o.B.d.A. $\delta' = 1$ und $\lambda_1 \geq \lambda_2 \geq \lambda_3$).

- Sind alle $\lambda_i > 0$, so ist $\mathfrak{F}$ ein **Ellipsoid** mit Halbachsen $1/\sqrt{\lambda_i}$. $\mathfrak{F}$ ist ein **Rotationsellipsoid**, wenn $\lambda_1 = \lambda_2$ oder $\lambda_2 = \lambda_3$. Wenn $\lambda_1 = \lambda_2 = \lambda_3$, so ist $\mathfrak{F}$ eine **Kugel**.
- Falls $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$, so ist $\mathfrak{F}$ ein **Einschaliges Hyperboloid**. Wenn $\lambda_1 = \lambda_2$, so ist $\mathfrak{F}$ rotations-symmetrisch bezüglich der 3-Achse.
- Falls $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$, so ist $\mathfrak{F}$ ein **Zweischaliges Hyperboloid**. Wenn $\lambda_2 = \lambda_3$, so ist $\mathfrak{F}$ rotations-symmetrisch bezüglich der 1-Achse.

I.2: Nicht alle $\lambda_i \neq 0$, $i = 1, 2, 3.$

In diesem Fall hat das LGS $(*)$ unendlich viele Lösungen und $\mathfrak{F}$ hat unendlich viele (Symmetrie-)Zentren. O.B.d.A. sei $\lambda_3 = 0$. Die Fläche $\mathfrak{F}$ wird dann beschrieben durch die Gleichung

$$\boxed{\lambda_1 y_1^2 + \lambda_2 y_2^2 = \delta'.}$$

I.2.1: $\delta' = 0$

- Seien $\lambda_1 \neq 0$ und $\lambda_2 \neq 0$. Haben λ_1 und λ_2 gleiches Vorzeichen, so ist $\mathfrak{F}$ die 3-Achse. Haben λ_1 und λ_2 verschiedenes Vorzeichen, so ist $\mathfrak{F}$ ein **Paar von Ebenen** mit der 3-Achse als Schnittgeraden.
- Sei $\lambda_1 = 0$ oder $\lambda_2 = 0$. O.B.d.A. $\lambda_2 = 0$. Ist $\lambda_1 \neq 0$, so ist $\mathfrak{F}$ die $(2,3)$-Ebene. Ist $\lambda_1 = 0$, so ist $\mathfrak{F} = \mathbb{E}^3$.

I.2.2: $\delta' \neq 0$ (o.B.d.A. $\delta' = 1$ und $\lambda_1 \geq \lambda_2$).

- Falls $\lambda_1 \geq \lambda_2 > 0$, ist $\mathcal{F}$ ein **Elliptischer Zylinder** mit der 3-Achse als Achse.
- Falls $\lambda_1 > \lambda_2 = 0$ ist $\mathcal{F}$ ein **Paar von parallelen Ebenen**.
- Falls $\lambda_1 > 0 > \lambda_2$ ist $\mathcal{F}$ ein **Hyperbolischer Zylinder** mit der 3-Achse als Achse.

Der Fall (II): Flächen ohne Zentrum

In diesem Fall hat das LGS $(*)$ keine Lösung.

Nach dem Spektralsatz 19.9 existiert eine ONB $\{a_1, a_2, a_3\}$ aus Eigenvektoren der Abbildung Φ. Da im vorliegenden Fall A singulär ist, muss mindestens einer der Eigenwerte $\lambda_1, \lambda_2, \lambda_3$ gleich Null sein. O.B.d.A. sei $\lambda_3 = 0$. Für $x = y_1a_1 + y_2a_2 + y_3a_3$ und $b = c_1a_1 + c_2a_2 + c_3a_3$ lautet die Gleichung von $\mathcal{F}$

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + 2(c_1y_1 + c_2y_2 + c_3y_3) = \delta$$

(dabei muss $c_3 \neq 0$ sein, da sonst $(*)$ lösbar wäre).

Wir machen nochmals Fallunterscheidungen.

Wieder sind verschiedene Fälle zu unterscheiden.

II.1: $\lambda_1 \neq 0$ und $\lambda_2 \neq 0$.

In diesem Fall können wir quadratisch ergänzen (Translation):

$$\lambda_1(y_1 + \frac{c_1}{\lambda_1})^2 + \lambda_2(y_2 + \frac{c_2}{\lambda_2})^2 + 2c_3y_3 = \delta + \frac{c_1^2}{\lambda_1} + \frac{c_2^2}{\lambda_2}.$$

Setzt man:

$$z_1 := y_1 + \frac{c_1}{\lambda_1}, \quad z_2 := y_2 + \frac{c_2}{\lambda_2}, \quad z_3 := y_3 - \frac{1}{c_3}(\delta + \frac{c_1^2}{\lambda_1} + \frac{c_2^2}{\lambda_2}),$$

so gilt im neuen Koordinatensystem

$$\lambda_1^* z_1^2 + \lambda_2^* z_2^2 + 2z_3 = 0 \qquad (\lambda_i^* = \frac{\lambda_i}{c_3},\ i = 1, 2).$$

- Haben λ_1 und λ_2 (und damit auch λ_1^*, λ_2^*) gleiches Vorzeichen, so ist $\mathcal{F}$ ein **Elliptisches Paraboloid** (mit 3-Achse als Achse) und rotationssymmetrisch genau dann, wenn $\lambda_1 = \lambda_2$.
- Haben λ_1 und λ_2 (und damit auch λ_1^*, λ_2^*) verschiedenes Vorzeichen, so ist $\mathcal{F}$ ein **Hyperbolisches Paraboloid** (mit 3-Achse als Achse).

II.2: $\lambda_1 = 0$ oder $\lambda_2 = 0$. O.B.d.A. sei $\lambda_2 = 0$.

- Sei $\lambda_1 \neq 0$. Dann haben wir zunächst

$$\lambda_1(y_1 + \frac{c_1}{\lambda_1})^2 + 2(c_2y_2 + c_3y_3) = \delta + \frac{c_1^2}{\lambda_1}.$$

Wir setzen

$$u_1 := y_1 + \frac{c_1}{\lambda_1},\ u_2 := y_2,\ u_3 := y_3,\quad (\text{und } \delta' := \delta + \frac{c_1^2}{\lambda_1}).$$

Damit erhalten wir als Gleichung für $\mathcal{F}$

$$\lambda_1 u_1^2 + 2(c_2u_2 + c_3u_3) = \delta'.$$

Eine Drehung um die u_1-Achse ergibt ein weiteres Koordinatensystem definiert durch

$$\begin{array}{rcl} u_1 & =: & v_1 \\ u_2 & =: & \frac{c_2}{\sqrt{c_2^2+c_3^2}}v_2 - \frac{c_3}{\sqrt{c_2^2+c_3^2}}v_3 \\ u_3 & =: & \frac{c_3}{\sqrt{c_2^2+c_3^2}}v_2 + \frac{c_2}{\sqrt{c_2^2+c_3^2}}v_3 \end{array}$$

ergibt

$$\lambda_1 v_1^2 + 2(\sqrt{c_2^2+c_3^2})v_2 = \delta'.$$

Eine weitere Translation definiert durch

$$z_1 := v_1,\ z_2 := v_2 - \frac{\delta'}{2\sqrt{c_2^2+c_3^2}},\ z_3 := v_3$$

ergibt schließlich

$$\lambda_1^* z_1^2 + 2z_2 = 0 \ \text{ mit } \ \lambda_1^* = \frac{\lambda_1}{\sqrt{c_2^2+c_3^2}}$$

und $\mathcal{F}$ ist ein **Parabolischer Zylinder** mit 3-Achse als Achse.

- Ist $\lambda_1 = 0$, so ist $\mathcal{F}$ eine **Ebene** ($c_3 \neq 0$).

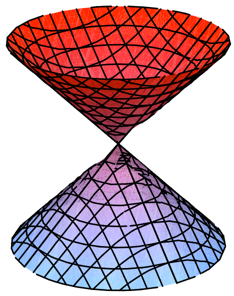
Zusammenfassend haben wir also gezeigt:

Satz 21.43 *Zu einer Quadrik (Fläche 2.Ordnung) $\mathcal{F}$ in $\mathbb{E}^3$ (verschieden von Punkt, Gerade und Ebene) existiert eine Quadrik $\tilde{\mathcal{F}}$ und Translationen bzw. Drehungen um $0 \in \mathbb{E}^3$, $f_1, \ldots, f_r$, so dass gilt*

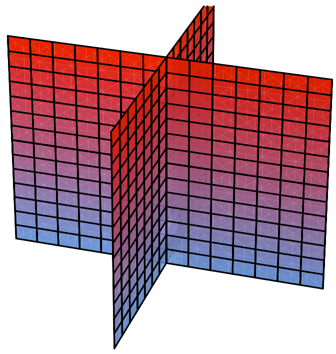
- $\tilde{\mathcal{F}} = f_1 \circ f_2 \circ \cdots \circ f_r(\mathcal{F})$
- $\tilde{\mathcal{F}}$ *wird durch eine Gleichung aus der folgenden Liste beschrieben:*

$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_1^2} - \frac{x_3^2}{a_1^2} = 0$	*Kegel*
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_1^2} = 0$	*Paar sich schneidender Ebenen*
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	*Ellipsoid, Kugel für* $a_1 = a_2 = a_3$
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1$	*Einschaliges Hyperboloid*
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1$	*Zweischaliges Hyperboloid*
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	*Elliptischer Zylinder*
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1$	*Hyperbolischer Zylinder*
$\frac{x_1^2}{a_1^2} = 1$	*Paar paralleler Ebenen*
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 2x_3$	*Elliptisches Paraboloid*
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 2x_3$	*Hyperbolisches Paraboloid*
$\frac{x_1^2}{a_1^2} = 2x_3$	*Parabolischer Zylinder*

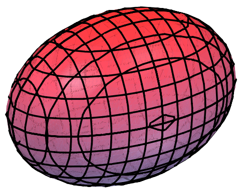
Bemerkung 21.44 Die a_i in Satz 21.43 sind beliebige reelle Zahlen > 0 ("freie Parameter").



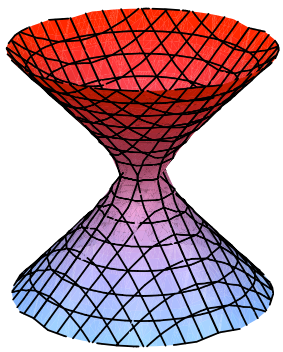
Kegel



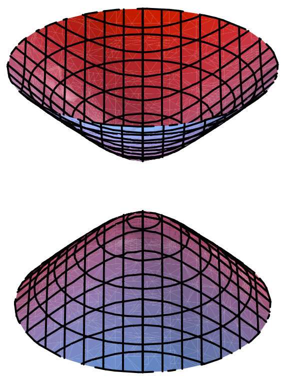
Paar sich schneidender Ebenen



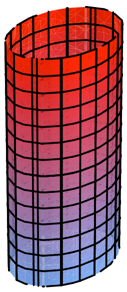
Ellipsoid



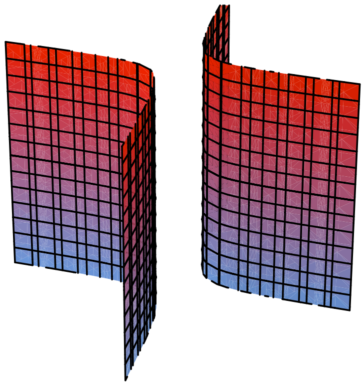
einschaliges Hyperboloid



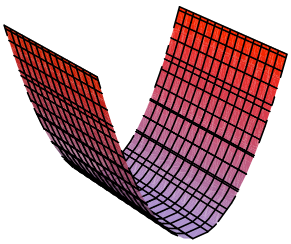
zweischaliges Hyperboloid



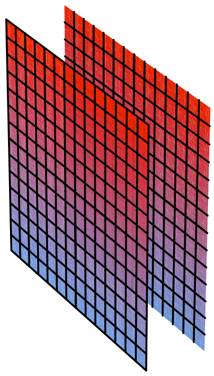
elliptischer Zylinder



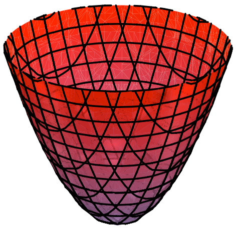
hyperbolischer Zylinder



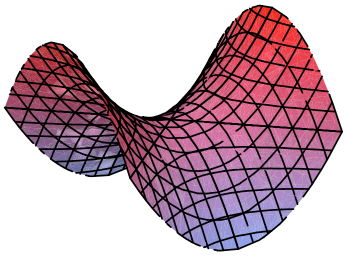
parabolischer Zylinder



Paar paralleler Ebenen



elliptisches Paraboloid



hyperbolisches Paraboloid

Satz 21.43 können wir auch als Klassifikationsresultat (durch Repräsentanten) formulieren.

Satz 21.45 *Jede Quadrik $\mathcal{F}$ in $\mathbb{E}^3$ ($\mathcal{F} \neq$ Punkt, Gerade, Ebene) ist euklidisch äquivalent zu einer Quadrik aus der Liste in Satz 21.43.*

Beweis: Wir müssen noch zeigen, dass keine zwei Quadriken in der Liste von Satz 21.43 euklidisch äquivalent sein können. Dazu sei A bzw. $\widetilde{A}$ die zu einer Quadrik gehörige symmetrische bzw. erweiterte Matrix. Wegen Formel (21.2) sind $\operatorname{Rang} A$ und $\operatorname{Rang} \widetilde{A}$ affine Invarianten einer Quadrik. Außerdem ist nach dem Satz über Hauptachsentransformationen 20.15 die Menge der Eigenwerte von A eine euklidische Invariante. Man prüft nun direkt nach, dass diese drei Invarianten für alle Quadriken in Satz 21.43 verschieden sind. ■

Beispiel 21.46 Für $\delta \in \mathbb{R}$ sei eine Quadrik $\mathcal{F} \subset \mathbb{E}^3$ gegeben durch die Gleichung

$$5x_1^2 + 11x_2^2 + 2x_3^2 - 16x_1x_2 - 20x_1x_3 - 4x_2x_3 + 14x_1 + 10x_2 - 28x_3 = \delta.$$

Die zugehörige symmetrische Matrix $A \in \mathbb{R}^3$ und der Vektor $b \in \mathbb{R}^3$ sind dann

$$A = (a_{ij}) = \begin{pmatrix} 5 & -8 & -10 \\ -8 & 11 & -2 \\ -10 & -2 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 7 \\ 5 \\ -14 \end{pmatrix}.$$

Für die Eigenwerte von A findet man: $\lambda_1 = 9, \lambda_2 = -9, \lambda_3 = 18$.

Das LGS $\Phi(x) = Ax = b$, also

$$\begin{aligned} 5x_1 - 8x_2 - 10x_3 &= -7 \\ -8x_1 + 11x_2 - 2x_3 &= -5 \\ -10x_1 - 2x_2 + 2x_3 &= 14, \end{aligned}$$

hat genau eine Lösung $c_0 = (-1, -1, 1)^\top$, *das* Zentrum der Fläche $\mathcal{F}$.

Um den "Typ" von $\mathcal{F}$ zu bestimmen, muss man δ' berechen:

$$\delta' = \delta - \langle c_0, \Phi(c_0) + 2b\rangle = \delta - \langle c_0, b\rangle = \delta + 26.$$

Damit haben wir folgendes Ergebnis:

$$\begin{aligned} \delta = -26 \quad &\Leftrightarrow \quad \mathcal{F} \text{ ist ein Kegel} \\ \delta > -26 \quad &\Leftrightarrow \quad \mathcal{F} \text{ ist ein einschaliges Hyperboloid} \\ \delta < -26 \quad &\Leftrightarrow \quad \mathcal{F} \text{ ist ein zweischaliges Hyperboloid} \end{aligned}$$

Hauptachsen von $\mathcal{F}$ = Eigenvektoren von A zu den Eigenwerten $\lambda_1, \lambda_2, \lambda_3$:

$$b_1 = \frac{1}{3}(1, 2, -2)^\top, b_2 = \frac{1}{3}(2, 1, 2)^\top, b_1 = \frac{1}{3}(2, -2, -1)^\top.$$

21.7.2 Affine Klassifikation von Quadriken

Manchmal ist man lediglich an der “Sylvesterform” einer Quadrik interessiert (d.h. nicht unbedingt an den expliziten Werten der Parameter a_i in der Liste aus Satz 21.43.

Im Beweis von Satz 21.43 haben wir sukzessive eulidische Isometrien (Translationen und Rotationen) verwendet um neue affine Koordinaten einzuführen, in denen die Quadrik eine immer einfachere Definitionsgleichung besitzt. Will man die Quadriken nur bis auf affine Äquivalenz bestimmen, hat man die größere Gruppe $\mathcal{A}ff(\mathbb{R}^3)$ zur Verfügung. Wir können also insbesondere durch Affinitäten der Form

$$y_i := \frac{x_i}{a_i} \quad i = 1, 2, 3$$

die Gleichungen aus Satz 21.43 weiter vereinfachen. Damit haben wir dann gezeigt

Satz 21.47 *Jede Quadrik $\mathcal{F}$ in $\mathbb{R}^3$ ($\mathcal{F} \neq$ Punkt, Gerade, Ebene) ist affin äquivalent zu einer Quadrik aus folgender Liste:*

$y_1^2 + y_2^2 - y_3^2 = 0$	*Kegel*
$y_1^2 - y_2^2 = 0$	*Paar sich schneidender Ebenen*
$y_1^2 + y_2^2 + y_3^2 = 1$	*Ellipsoid*
$y_1^2 + y_2^2 - y_3^2 = 1$	*Einschaliges Hyperboloid*
$y_1^2 - y_2^2 - y_3^2 = 1$	*Zweischaliges Hyperboloid*
$y_1^2 + y_2^2 = 1$	*Elliptischer Zylinder*
$y_1^2 - y_2^2 = 1$	*Hyperbolischer Zylinder*
$y_1^2 = 1$	*Paar paralleler Ebenen*
$y_1^2 + y_2^2 = 2y_3$	*Elliptisches Paraboloid*
$y_1^2 - y_2^2 = 2y_3$	*Hyperbolisches Paraboloid*
$y_1^2 = 2y_3$	*Parabolischer Zylinder*

Beweis: Wir müssen noch zeigen, dass keine zwei Quadriken in der Liste von Satz 21.47 affin äquivalent sein können. Dazu sei wieder A bzw. $\widetilde{A}$ die zu einer Quadrik gehörige symmetrische bzw. erweiterte Matrix. Wegen Formel (21.2) sind Rang A und Rang $\widetilde{A}$ affine Invarianten einer Quadrik. Außerdem ist nach dem Trägheitssatz von Sylvester 20.17 die Signatur von A eine lineare Invariante. Man prüft nun direkt nach, dass diese drei Invarianten für alle Quadriken in Satz 21.47 verschieden sind. ■

Ist man nur an der affinen Klassifikation interessiert, so gibt es ein einfacheres Verfahren, als das in Beispiel 21.46 durchgeführte, nämlich das **quadratisches Ergänzen**. Dazu zwei Beispiele.

Beispiel 21.48 1. Für $t \in \mathbb{R}$ sei eine Quadrik $\mathcal{F}_t \subset \mathbb{R}^3$ gegeben durch die Gleichung

$$2x_1^2 + 3x_2^2 + \frac{5}{2}x_3^2 + 4x_1x_2 - 6x_1x_3 - 8x_2x_3 + 2tx_3 + t^2 = 0.$$

Wir beginnen mit $2x_1^2$ und betrachten alle Summanden in denen x_1 als Faktor vorkommt. Durch quadratisches Ergänzen ergibt sich:

$$2[x_1^2+2x_1(x_2-\frac{3}{2}x_3)+(x_2-\frac{3}{2}x_3)^2]-2(x_2-\frac{3}{2}x_3)^2+3x_2^2-8x_2x_3+\frac{5}{2}x_3^2+2tx_3+t^2 = 0$$

also

$$2(x_1 + x_2 - \frac{3}{2}x_3)^2 + x_2^2 - 2x_2x_3 - 2x_3^2 + 2tx_3 + t^2 = 0.$$

Jetzt betrachtet man entsprechend x_2^2 und die (restlichen) Summanden in den x_2 vorkommt und ergänzt quadratisch:

$$2(x_1 + x_2 - \frac{3}{2}x_3)^2 + (x_2^2 - 2x_2x_3 + x_3^2) - 3x_3^2 + 2tx_3 + t^2 = 0.$$

Eine letzte quadratische Ergänzung ergibt:

$$2(x_1 + x_2 - \frac{3}{2}x_3)^2 + (x_2 - x_3 + x_3)^2 - 3(x_3 - \frac{1}{3}t)^2 + \frac{4}{3}t^2 = 0.$$

Definiert man eine affine Koordinaten-Transformation durch:

$$\begin{aligned} y_1 &:= \sqrt{2}(x_1 + x_2 - \frac{3}{2}x_3) \\ y_2 &:= x_2 - x_3 \\ y_3 &:= \sqrt{3}(x_3 - \frac{1}{3}t), \end{aligned}$$

so lautet die Gleichung für $\mathcal{F}_t$ in den neuen Koordinaten

$$y_1^2 + y_2^2 - y_3^2 + \frac{4}{3}t^2 = 0.$$

Damit haben wir folgendes Ergebnis:

$$\begin{aligned} t = 0 \quad &\Leftrightarrow \quad \mathcal{F}_0 \text{ ist ein Kegel} \\ t \neq 0 \quad &\Leftrightarrow \quad \mathcal{F}_t \text{ ist ein zweischaliges Hyperboloid} \end{aligned}$$

2. Die Quadrik $\mathcal{F}$ sei gegeben durch

$$x_1x_2 + x_1x_3 + x_2x_3 + 2x_1 - 1 = 0.$$

Um quadratisch ergänzen zu können, brauchen wir quadratische Terme. Diese verschaffen wir uns durch folgenden Trick. Wir machen die (lineare) Koordinatentransformation

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

Damit lautet die Gleichung von $\mathcal{F}$

$$y_1^2 - y_2^2 + 2y_1y_3 + 2y_1 - 2y_2 - 1 = 0.$$

Jetzt kann man quadratisch ergänzen:

$$\begin{aligned} & y_1^2 + 2y_1(y_3+1) + (y_3+1)^2 - (y_3+1)^2 - y_2^2 - 2y_2 - 1 = 0 \\ \Leftrightarrow\ & (y_1 + y_3 + 1)^2 - (y_2+1)^2 - (y_3+1)^2 = 0. \end{aligned}$$

Macht man schließlich noch die affine Koordinatentransformation

$$\begin{aligned} z_1 &:= y_1 + y_3 - 1 \\ z_2 &:= y_2 + 1 \\ z_3 &:= y_3 + 1, \end{aligned}$$

so erhält man als Ergebnis $z_1^2 - z_2^2 - z_3^2 = 0$, d.h. $\mathcal{F}$ ist ein Kegel.

Literatur

[1] A. Beutelspacher
Lineare Algebra
Vieweg Verlag, 1994

[2] E. Brieskorn
Lineare Algebra und Analytische Geometrie I
Vieweg Verlag, 1983

[3] E. Brieskorn
Lineare Algebra und Analytische Geometrie II
Vieweg Verlag, 1985

[4] R. Courant/H. Robbins
Was ist Mathematik?
Springer Verlag, 1967

[5] P. Davis, R. Hersh
Erfahrung Mathematik
Birkhäuser Verlag, 1985

[6] K. Devlin
Muster der Mathematik
Spektrum Verlag, 1998

[7] H. Ebbinghaus et.al.
Zahlen, Grundwissen der Mathematik 1,
Springer Verlag, 1988

[8] G. Fischer
Analytische Geometrie
Vieweg Verlag, 2001

[9] G. Fischer
Lineare Algebra
Vieweg Verlag, 2005

[10] W. Greub
Linear Algebra, GTM 23
Springer Verlag, 1975

[11] T. Gowers
Mathematics, A very short introduction
Oxford University Press, 2002

[12] J. Hadamard
The Psychology of Invention in the Mathematical Field
Princeton University Press, 1945

[13] P.R. Halmos
Naive Mengenlehre
Vandenhoek & Ruprecht, 1976

[14] R. Hersh
What is Mathematics, Really?
Vintage Books, 1997

[15] K. Jänich
Lineare Algebra, 10. Auflage
Springer Verlag, 2008

[16] K. Königsberger
Analysis I, 3. Auflage
Springer Verlag, 1995

[17] M. Otte (Hrsg.)
Mathematiker über die Mathematik
Springer Verlag, 1974

[18] G. Polya
Schule des Denkens (engl. How to solve it)
Sammlung Dalp, 1949

[19] D. Ruelle
The Mathematician's brain
Princeton University Press, 2007

[20] A. Tarski
Einführung in die mathematische Logik
Vandenhoeck & Ruprecht, 1977.

[21] Biographien:
`www-history.mcs.st-and.ac.uk`

Symbole

Index