

Lineare Algebra für Ingenieure

Volker Mehrmann Jörg Rambau Ruedi Seiler

1. Juni 2011

INHALTSVERZEICHNIS

1	Lineare Algebra und Differentialgleichungen: Ein erster Eindruck	5
1.1	Was ist lineare Algebra?	5
1.2	Beispiele	7
1.3	Ziele und Struktur des Kurses	10
1.4	Vom Lösen mathematischer und anderer Probleme	11
1.5	Tipps	12
2	Die Vektorräume \mathbb{R}^n, \mathbb{C}^n, \mathbb{K}^n	13
2.1	Definition und Eigenschaften	13
2.2	Basis und Dimension	15
2.3	Teilräume	18
3	Matrizen	20
3.1	Anwendungen die auf Matrizen führen	20
3.2	Definition und Beispiele von Matrizen	21
3.3	Matrizen als Vektorraum	23
3.4	Matrizenmultiplikation	24
3.5	Lineare Abbildungen und Matrizen	26
4	Der Gaußalgorithmus	30
4.1	Beispiele	30
4.2	Lineare Gleichungssysteme	32
4.3	Die Zeilenstufenform und die Normierte Zeilenstufenform	34
4.4	Der Gaußalgorithmus	36
4.5	Was man aus der NZSF alles berechnen kann	39
5	Vektorräume über dem Körper \mathbb{K}	41
5.1	Definition, Beispiele, Eigenschaften von Vektorräumen	41
5.2	Basis und Dimension	42
5.3	Koordinaten und Koordinatenabbildung	46
6	Lineare Abbildungen	48
6.1	Definition und Beispiele	48

6.2	Der Vektorraum der linearen Abbildungen	50
6.3	Die Komposition linearer Abbildungen	50
6.4	Kern und Bild einer linearen Abbildung, Lösungsraum linearer Gleichungen	51
7	Koordinaten und darstellende Matrizen	54
7.1	Die Darstellung von Vektoren durch Koordinatenvektoren	54
7.2	Die Transformation der Koordinatenvektoren bei Basiswechsel . .	56
7.3	Die Darstellung linearer Abbildungen durch Matrizen	59
7.4	Die Transformation der darstellenden Matrix bei Basiswechsel . .	61
8	Euklidische und unitäre Räume	63
8.1	Norm, Beispiele	63
8.2	Skalarprodukt, Beispiele	66
8.3	Eigenschaften von Skalarprodukten und Normen	68
8.4	Orthonormalbasis, Gram-Schmidtsches Verfahren	69
8.5	Das Vektorprodukt	70
9	Orthogonale und unitäre Abbildungen	71
9.1	Definitionen der grundlegenden Begriffe und Beispiele	71
9.2	Die QR -Zerlegung	76
10	Determinante	81
10.1	Einleitung	81
10.2	Definition und Eigenschaften der Determinante	82
11	Eigenwerte, Eigenvektoren und Charakteristisches Polynom	90
11.1	Definition, Beispiele und grundlegende Eigenschaften	91
11.2	Definition und elementare Eigenschaften charakteristischer Poly- nome	96
11.3	Berechnung der Eigenwerte als Nullstellen des charakteristischen Polynoms	97
12	Diagonalisierbarkeit von Matrizen	100
12.1	Definition und weitere Beispiele	101
12.2	Diagonalisierung, Eigenvektoren und Eigenwerte	102
12.3	Algorithmus zur Berechnung der Diagonalisierung	102
12.4	Rechnen mit diagonalisierbaren Matrizen	103
12.5	Klassen diagonalisierbarer und nichtdiagonalisierbarer Matrizen .	105
13	Lineare Differentialgleichungen (Teil 1)	110

13.1	Definition und Beispiele	112
13.2	Differentialgleichungen und Lösungen zeichnen, das Phasenraum- portrait	113
13.3	Die Exp-Methode zur Berechnung von Lösungen	114
14	Lineare Differentialgleichungen (2. Teil)	118
14.1	Die lineare Differentialgleichung zweiter Ordnung	120
14.1(a)	Definitionen und Begriffe	120
14.1(b)	Struktur der Lösungsmenge	121
14.1(c)	Das Anfangswertproblem	122
14.2	Die lineare DGL. 2-ter Ordnung mit konstanten Koeffizienten: Die Schwingungsgleichung	123
14.2(a)	Die homogene Gleichung: Allgemeine Lösung für die freie Schwingung	124
14.2(b)	Die inhomogene Gleichung: Die erzwungene Schwingung .	126

1. KAPITEL

LINEARE ALGEBRA UND DIFFERENTIALGLEICHUNGEN: EIN ERSTER EINDRUCK

1.1 WAS IST LINEARE ALGEBRA?

Diese erste Vorlesung ist eine Vorschau auf das, was Euch in diesem Kurs über *Lineare Algebra und Differentialgleichungen* erwartet. Wie in jedem Menü muss von Dingen geredet werden, die noch nicht verdaut sind. Dass Ihr deshalb heute einiges noch nicht oder nicht so ganz richtig versteht, ist kein Grund zur Sorge. Die im Folgenden angesprochenen Themen werden in den kommenden Vorlesungen erklärt.

Lineare Algebra ist eine abstrakte Sprache um in mathematisch präziser Form über Konzepte wie Ort, Geschwindigkeit, Kraft, Gerade, Ebene usw. zu sprechen und zu schreiben. Sie ist universell in zweifacher Hinsicht: Sie kann auf sehr viele verschiedene Situationen in Mathematik, Natur- und Ingenieurwissenschaft angewandt werden und sie ist – wie die Mathematik überhaupt – eine gemeinsame Sprache aller Mathematiker, Natur- und Ingenieurwissenschaftler in welchen Ländern und Kulturen sie auch immer leben.

Das wesentlichste Konzept der linearen Algebra ist der Vektor. Vektoren können wie Zahlen addiert und in gewissem Sinne oftmals auch multipliziert werden.

Vektoren sind eine mathematische Abstraktion von Pfeilen, die alle von einem Punkt ausgehen und in irgendwelchen anderen Punkten enden. Dabei stellen wir uns meistens eine Ebene vor, z.B. die Tafel, auf der ein Punkt, eben der Ausgangspunkt der Pfeile, markiert ist. Der markierte Punkt heißt auch Ursprung. Das Addieren von Vektoren entspricht dann dem Aneinandersetzen von Pfeilen.

Entscheidend wichtig für die Lineare Algebra und ihren großen Erfolg ist Folgendes: Dieses eben geschilderte Bild der Pfeile und wie damit umzugehen ist, kann auf die unterschiedlichsten Situationen angewandt werden. Wir werden die verschiedensten mathematischen Objekte als Vektoren interpretieren und mit ihnen so rechnen wie mit Pfeilen. So können beispielsweise Polynome, Matrizen und Funktionen als Vektoren interpretiert werden, um nur einige wenige zu nennen.

Neben den Vektoren spielen noch eine Reihe anderer Konzepte eine für die Lineare Algebra bedeutende Rolle. Ohne weiter auf sie einzugehen, ist hier eine unvollständige Liste:

- Vektorraum (auch linearer Raum genannt)
- Linearkombination von Vektoren
- Lineare Abhängigkeit von Vektoren

- Basis
- Dimension eines Vektorraums
- Matrizen
- Lineare Abbildungen, speziell Projektionen
- Skalarprodukt
- Lineares Gleichungssystem
- Determinante

Die Theorie der „Linearen Differentialgleichungen“ ist eine schöne und wichtige Anwendung der Linearen Algebra. Sie werden uns ganz zum Schluss des Kurses beschäftigen. Das trifft sich zeitlich gut, weil in der Theorie der linearen Differentialgleichungen die Exponentialfunktion von Matrizen und linearen Abbildungen entscheidend wichtig ist. Bis dahin werdet Ihr diesen Begriff in der Analysis gelernt haben.

Die Methoden der Linearen Algebra eignen sich zur Beschreibung einer großen Zahl von Phänomenen der Natur- und Ingenieurwissenschaften. (Man spricht von einer mathematischen Modellierung der Phänomene.) Hier sind einige Beispiele, die ganz oder zu einem guten Teil mit Methoden der Linearen Algebra beschrieben werden können:

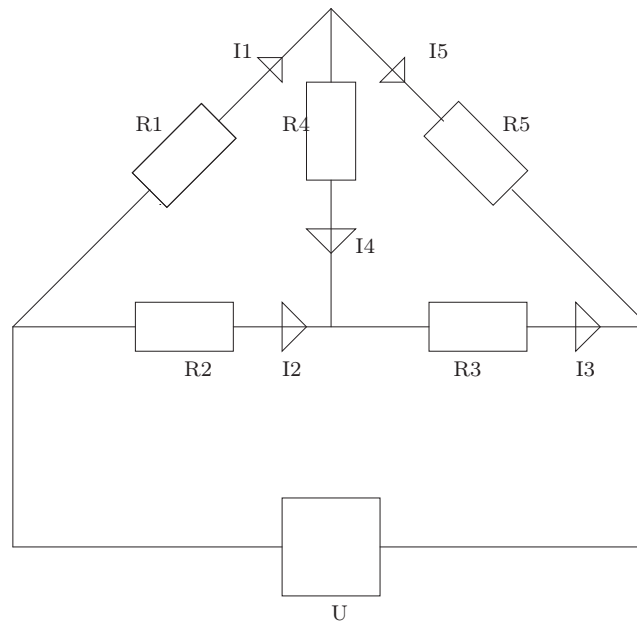
- Der elektrische Schwingkreis
- Das Pendel
- Das Biegen von Balken
- Die Bewegung des Kreisels
- Die Drehung eines Kreisels um seine stabilen und instabilen Achsen.
- Die Spektren von Atomen und Molekülen
- Die Darstellung, Bearbeitung und Kompression digitaler Bilder
- usw.

Die Kraft der Mathematik und ganz besonders der Linearen Algebra liegt darin, dass es sich um eine Struktur abstrakter Regeln handelt, die auf die verschiedensten Objekte angewandt werden können. So werden wir beispielsweise die Addition von Vektoren einführen und die zugehörigen Regeln, die Ihr bereits vom Rechnen mit Zahlen kennt, *postulieren*. Damit kann dann mit dem abstrakten Konzept Vektor gerechnet werden, ganz gleich ob wir mit Vektor eine Zahl, eine Gruppe von Zahlen, ein Polynom, eine Funktion, eine Matrix, eine Abbildung etc. meinen. Immer wird gleich gerechnet. Dies ist am Anfang schwierig, da zum Beispiel das Additionssymbol $+$ (obwohl immer gleich geschrieben) immer wieder anders meint. Es ist wichtig und sinnvoll die Addition immer gleich zu schreiben, denn es sind *die Regeln wie man damit umgeht*, die im Vordergrund stehen und die sind immer gleich. Dies macht zu Beginn das Verstehen der Linearen Algebra nicht gerade leicht. Kaum ein Gebiet der Mathematik ist am Anfang so fremd wie die Lineare Algebra - kaum eines wird so bald zum vertrauten Terrain.

1.2 BEISPIELE

Um Euch einen Eindruck von der Kraft der Linearen Algebra zu vermitteln, stellen wir drei Anwendungsbeispiele vor, die an dieser Stelle allerdings noch nicht wirklich zu verstehen sind. Dies wird sich jedoch ändern. Zum Schluss des Kurses werdet Ihr sie alle bestens verinnerlicht haben!

Beispiel 1.2.1. [Die Wheatstonesche Brücke] Dies ist ein Instrument zur Messung eines elektrischen Widerstandes. Es besteht aus 5 Widerständen und einer Batterie der Spannung U und sieht folgendermaßen aus:



Es gelten die sogenannten Kirchhoffschen Gesetze

$$\begin{array}{rcccccccl}
 R_1 I_1 & & & & + R_5 I_5 & = & U & \\
 & R_2 I_2 & + R_3 I_3 & & & = & U & \\
 R_1 I_1 & - R_2 I_2 & & + R_4 I_4 & & = & 0 & \\
 & & - R_3 I_3 & - R_4 I_4 & + R_5 I_5 & = & 0 & \\
 I_1 & & & - I_4 & - I_5 & = & 0 & \\
 & I_2 & - I_3 & + I_4 & & = & 0 &
 \end{array} \quad (1.1)$$

Dies ist ein lineares Gleichungssystem: Gegeben sind R_1, \dots, R_5 und U . Gesucht die Ströme I_1, \dots, I_5 .

Nun füttern wir dies in das Computer-Algebra-Programm „Maple“ und lassen dieses für uns rechnen. „Maple“ liefert sofort eine Lösung und verwendet dazu das sog. Gaußsche Verfahren, das wir bald kennenlernen werden.

Frage: Ist das Resultat richtig? Können wir ihm vertrauen? Um diese wichtige Frage beantworten zu können, muss der Lösungsprozess *verstanden* werden. Es ist ein wichtiges Ziel dieses Kurses, Euch diese Fähigkeit zu vermitteln. Die Frage nach der Richtigkeit der Antwort zu stellen, ist nicht müßig, denn *auch Maple macht Fehler!*

Beispiel 1.2.2. [Bilder, Bilddatenkompression] Digitale Bilder erzeugen riesige Files. Ein Fernsehbild entspricht etwa 1000 Seiten Text und braucht oft sehr lange, um übertragen und angezeigt zu werden. Ein einfacher Trick der linearen Algebra bringt uns einen guten Schritt weiter auf dem Weg „Dicke Bilder durch dünne Leitungen, z.B. Telefonleitungen, zu schicken“.

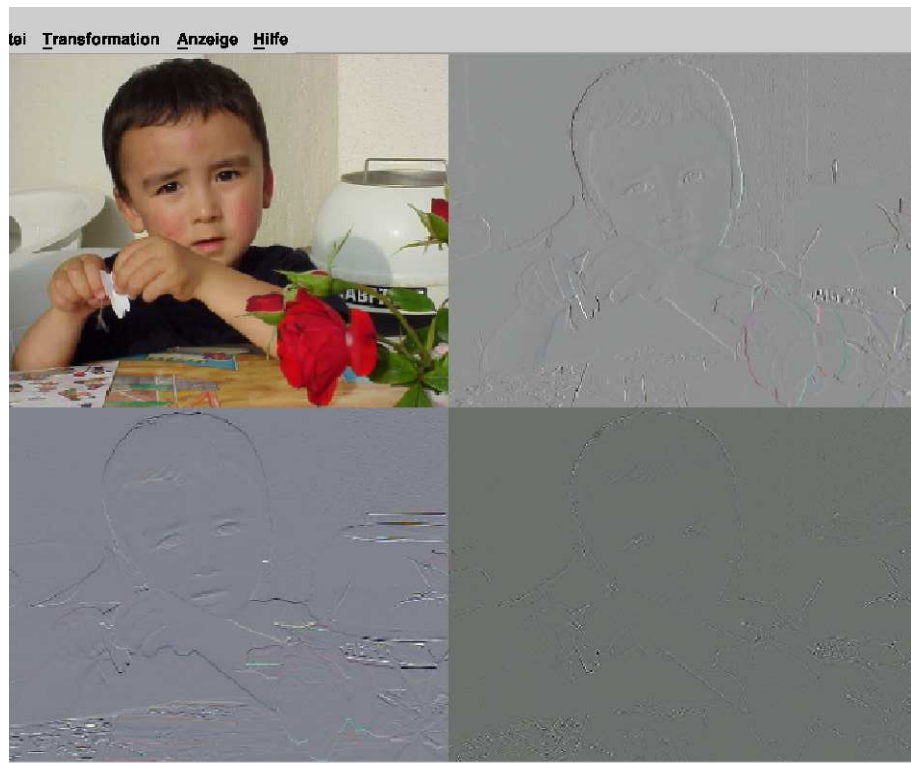
Bilder können wie Vektoren addiert werden indem die Helligkeitswerte jedes Pixels addiert werden. So wird ein Bild mit einer Katze und ein zweites mit einer Maus unter Addition zu

einem Bild mit einer Katze und einer Maus. Bilder können also durch Vektoren modelliert werden.

Hier ist ein Bild von Manuel.



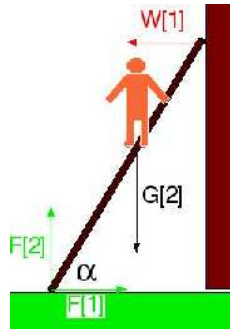
Um das Bild zu komprimieren wählen wir eine geschickte Transformation des Bildes. Daraus wird ein neues Bild, das nun vier sehr unterschiedliche Segmente enthält.



Unter den vier Feldern gibt es drei, die sozusagen keine „Information“ enthalten. Sie können – ohne die Bildqualität wesentlich zu beeinträchtigen – einfach weggelassen werden. Der Rest des Bildes ist nur ein Viertel so groß wie das ursprünglichen Bild. (Was „Information“ bedeutet, ist Thema des Kurses Wahrscheinlichkeitstheorie oder Informationstheorie.)

Die hier gewählte Transformation ist die einfachste Variante einer sogenannten Wavelettransformation und gehört zu der Klasse der linearen Transformationen, die uns eine ganze Weile in Atem halten werden.

Beispiel 1.2.3. [Die Leiter an der Wand] Eine Leiter steht an eine Wand angelehnt. Ein Häuslebauer möchte sich daraufstellen, ohne wegzurutschen. Wie groß muss der Reibungskoeffizient mindestens sein?



Auch dieses Problem führt – ähnlich wie im ersten Beispiel – auf ein lineares System von Gleichungen und eine Ungleichung. Es kann von Hand oder aber schneller mit Maple gelöst werden. Soviel zu den Beispielen.

1.3 ZIELE UND STRUKTUR DES KURSES

Es geht also in diesem Kurs darum, die Sprache und Denkweise der Linearen Algebra zu erlernen, um Probleme, wie wir sie in den Beispielen gesehen haben, zu lösen. Wie dies organisatorisch vonstatten geht, sei hier kurz beschrieben.

Der Kurs besteht aus zwei Teilen, der Vorlesung und den Übungen in kleinen Gruppen. In der Vorlesung werden die mathematischen Konzepte erklärt. Jede Vorlesung beginnt damit, dass das Thema und die Frage erklärt werden. Dies dauert ca. 10 Minuten. Wer zu spät kommt und diese Einführung nicht mitbekommt, wird es sehr schwer haben, den Rest der Vorlesung zu verstehen. Der Stoff ist umfangreich und die Zeit sehr kurz bemessen, so dass für Wiederholungen kein Platz bleibt. Den Text zur Vorlesung werden wir Euch elektronisch zur Verfügung stellen. Die wenigsten Leute kommen aber damit alleine aus, da das mathematische Denken und Argumentieren zum guten Teil über das „Nachmachen“ erlernt wird: In der Vorlesung könnt Ihr dies live erleben.

Allerdings gibt die Vorlesung einen verzerrten Eindruck, weil die Argumentation poliert und fast fehlerfrei ist. Im praktischen Umgang mit Mathematik ist dies nie so.

Die Übungen sind dazu da, Euch beim eigenständigen Lösen der Hausaufgaben zu helfen. Hier werdet Ihr angeleitet, einfache Probleme zum Stoff der Vorlesung zu lösen. Die Fähigkeit, Probleme der angewandten Mathematik zu lösen, ist das, was Ihr später eigentlich braucht. Deshalb sind die Übungen so eminent wichtig. Und noch einmal: Das Lösen von Aufgaben ist zum Erfolg des Kurses, zum Bestehen der Prüfung und für Eure spätere Tätigkeit in Studium und Beruf von ganz großer Bedeutung. Das Zuhören in der Vorlesung und das aktive Lösen von Aufgaben und Problemen sind zwei sehr unterschiedliche intellektuelle Tätigkeiten, und wir wollen beides mit Euch so intensiv wie möglich betreiben.

1.4 VOM LÖSEN MATHEMATISCHER UND ANDERER PROBLEME

Hier ist eine kleine Heuristik (Anleitung) zum Vorgehen beim Lösen von Problemen. Sie soll Euch helfen die Aufgaben, die wöchentlich zu bearbeiten sind und einen wesentlichen Teil dieses Kurses ausmachen, systematisch und erfolgreich zu lösen.

Die folgende Heuristik – das werdet Ihr sofort sehen – ist durchaus auch auf andere Gebiete der Wissenschaften anwendbar, bis hin zum Lösen alltäglicher Probleme. Dieser Kurs und auch die Analysis sind also eine hervorragende Spielwiese, um Problemlösungsstrategien zu erlernen. Die Bedeutung des mathematischen Denkens und die Fähigkeit Probleme schnell knacken zu können, geht weit über die Mathematik hinaus. Deshalb seid Ihr für die Wirtschaft ganz allgemein und nicht nur für diejenigen Betriebe, die sich mit Eurem eigentlichen Fachgebiet beschäftigen, so interessant – falls Ihr die Technik des Lösen von Übungen in der Linearen Algebra gelernt habt!!

Das Lösen von Problemen verläuft in vier Schritten (vgl. G. Polya: Die Schule des Denkens: Vom Lösen mathematischer Probleme):

- Die Formulierung des Problems in eigenen Worten: Wenn Ihr eine Aufgabe bearbeiten wollt oder eine Sache in der Vorlesung zu verstehen sucht, sprecht darüber! Formuliert das Problem für Euch, so lange es geht ohne Papier, vielleicht vor einer Tafel. Sprecht miteinander über das Problem beim Kaffee oder Spazieren.
- Formuliert eine Lösungsstrategie. Zerlegt das Problem in Teile. Versucht Euch an ähnliche Fragen zu erinnern. Versucht Euch vorzustellen, wie die Antwort aussehen könnte. Überprüft gedanklich die Lösungsstrategie z.B. so: Enthält die formulierte Lösungsstrategie alles was über das Problem bekannt ist? Auch dieser Schritt geht oftmals sehr gut im Gespräch zu zweit oder zu dritt.
- Jetzt muss der Plan durchgeführt werden. Dazu ist meistens Ruhe nötig, d. h. dieser Teil muss oftmals allein durchgeführt werden.
- Im letzten Schritt geht es darum, darüber nachzudenken, was sagt mir die Lösung, ist sie richtig? Stimmen die physikalischen Dimensionen? Stimmt die Größenordnung? Was geschieht, wenn die Parameter etwas geändert, die Aufgabe etwas variiert wird? Habe ich alle mir über das Problem bekannten Informationen verwendet?

Ziel der Übungen ist es

- die beiden ersten Schritte beim Lösen der Hausaufgaben zu gehen und
- einzelne Probleme exemplarisch ganz zu bearbeiten.

In den Übungen habt Ihr das Wort, die Übungsleiterin wird Euch helfen und den Ablauf der Übungen steuern, aber möglichst wenig zur Lösung der Probleme beitragen.

Wieviel Zeit ist für Übungen und Prüfung nötig? Natürlich ist dies von Person zu Person verschieden. Hier sind Durchschnittswerte: Für das Lösen der Übungsaufgaben und das Nacharbeiten der Vorlesung ist 7 Stunden pro Woche nötig. Für die Vorbereitung der Klausur sind es dann noch einmal 30 Stunden.

1.5 TIPPS

Zum Schluss noch einige Ratschläge und Warnungen aus langjähriger Erfahrung:

- Die Darstellung des Stoffes baut aufeinander auf. Auch nur eine einzige Vorlesung oder Übung nicht mitzumachen, wird Euch große Schwierigkeiten bereiten. Vermeidet es!
- Wenn Ihr glaubt, den Stoff bereits zu kennen, bleibt trotzdem dran. Das Tempo in der Vorlesung ist hoch und es könnte dann sein, dass plötzlich der Zug abgefahren ist.
- In den Vorlesungen werdet Ihr nicht immer alles verstehen. Dies geht allen so. Natürlich könnt Ihr fragen, und wir werden, soweit es die Zeit gestattet, darauf eingehen. Was nicht eintreffen darf, ist, dass Ihr „den roten Faden“ verliert. Ihr müsst immer wissen, wovon die Rede ist. Die Einzelheiten spielen dabei vorerst noch keine entscheidende Rolle. (Natürlich wäre es besser, auch die zu verstehen. Manche Leute müssen aber zu Hause oder im Cafe noch einmal darüber nachdenken.)
- Mathematik ist eine abstrakte Sprache mit einer effizienten Schrift: Um sie zu erlernen, müsst Ihr mit uns und unter Euch über Mathematik reden und Hausaufgaben schreiben. Wenn einige Leute versuchen, Euch glaubhaft zu machen, Hausaufgaben seien zum Verstehen der Mathematik nicht nötig, glaubt Ihnen nicht. Sie wissen nicht, wovon sie reden!

2. KAPITEL

DIE VEKTORRÄUME \mathbb{R}^n , \mathbb{C}^n , \mathbb{K}^n

Diese Vorlesung dient dazu, die grundlegenden Begriffe der Vektorrechnung einzuführen. Sie formen die Sprache, in der wir uns in den nächsten Wochen unterhalten werden.

Neben den Begriffen, die in sogenannten Definitionen eingeführt werden, sind es die mathematischen Resultate, die es zu verstehen und zu lernen gilt. Um sie vom Rest der Vorlesung abzugrenzen, werden sie in sogenannten Theoremen oder Lemmas zusammenfassend dargestellt. Theoreme enthalten anspruchsvolle Resultate, Lemmas dagegen Einfacheres.

2.1 DEFINITION UND EIGENSCHAFTEN

Wir stellen zunächst die Hauptdarsteller der ersten Vorlesung vor und diskutieren ihre grundlegenden Eigenschaften.

Fundamental sind die reellen und komplexen Zahlen die wir mit den Symbolen \mathbb{R} und \mathbb{C} bezeichnen. Sie bilden einen Körper, d.h. man kann sie addieren und multiplizieren, wobei diese Operationen den – mindestens für die reellen Zahlen – wohlbekannten Regeln genügen. In der Analysis wird gezeigt, wie diese Zahlen aus den natürlichen Zahlen \mathbb{N} und ganzen Zahlen \mathbb{Z} konstruiert werden.

Jetzt kommen wir zum wichtigsten Konzept dieser Vorlesung, dem Vektorraum der reellen n -Tupel:

Definition 2.1.1 (Vektorraum \mathbb{R}^n). Ein n -Tupel reeller Zahlen

$$\vec{v} := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad v_1, v_2, \dots, v_n \in \mathbb{R} \quad (2.1)$$

heißt *n -dimensionaler reeller Vektor*, im folgenden kurz *Vektor*. Die Einträge v_1, v_2, \dots, v_n heißen *Komponenten*, *Standardkoordinaten* oder kurz *Koordinaten* von \vec{v} . (In diesem Kurs und ebenso in der Analysis werden noch andere Koordinaten eine Rolle spielen.)

Der spezielle Vektor

$$\vec{0} := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.2)$$

heißt *Nullvektor*.

Zwei Vektoren gleicher Größe können addiert werden. Die Summe ist wieder ein Vektor und folgendermassen definiert:

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} := \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}. \quad (2.3)$$

Diese Vorschrift nennt man *Vektoraddition*.

Eine reelle Zahl $\alpha \in \mathbb{R}$ kann man mit einem Vektor multiplizieren und erhält wiederum einen Vektor. Diese Operation ist so definiert:

$$\alpha \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} := \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}. \quad (2.4)$$

Das nennt man *Multiplikation mit einem Skalar*, da α auch als *Skalar* bezeichnet wird. Wir verwenden auch die folgende Schreibweise: $(-1)\vec{v} = -\vec{v}$.

Die Menge aller n -dimensionalen reellen Vektoren, zusammen mit der oben definierten Addition von Vektoren und der Multiplikation von Vektoren mit reellen Zahlen, nennt man den *Vektorraum* \mathbb{R}^n . Die Vektoraddition und die Multiplikation mit Skalaren heissen *Vektorraumoperationen*.

Bemerkung 2.1.2. Der Doppelpunkt vor dem Gleichheitszeichen in der obigen Gleichung heisst, dass die linke Seite durch die rechte definiert ist. Diese Notation werden wir im Folgenden oft benutzen. Sie gestattet Größen neu einzuführen ohne die umständlichere Form der formalen Definition zu verwenden.

Bemerkung 2.1.3. Man kann sich die Vektoren des \mathbb{R}^2 (bzw. des \mathbb{R}^3) in einem Koordinatensystem in der Ebene (bzw. im Raum) als Pfeile, die vom Ursprung aus starten, vorstellen.

Beispiel 2.1.4. [$n = 2$: \mathbb{R}^2] Einige Vektoren:

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \pi \\ e \end{bmatrix}, \begin{bmatrix} \sin(\frac{\pi}{4}) \\ \cos(\frac{\pi}{4}) \end{bmatrix} \quad (2.5)$$

Einige Summen:

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad (2.6)$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (2.7)$$

$$\begin{bmatrix} \pi \\ e \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \pi \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \pi \\ e \end{bmatrix} \quad (2.8)$$

Einige Multiplikationen mit Skalaren:

$$2,5 \cdot \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \quad \sqrt{2} \cdot \begin{bmatrix} -\sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \quad (2.9)$$

Beispiel 2.1.5. [$n = 3$: \mathbb{R}^3] Einige Vektoren:

$$\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \pi \\ e \\ \sqrt{2} \end{bmatrix} \quad (2.10)$$

Einige Summen:

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix} \quad (2.11)$$

$$\begin{bmatrix} \pi \\ e \\ \sin(\frac{\pi}{4}) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \pi \\ e \\ \sin(\frac{\pi}{4}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \pi \\ e \\ \sin(\frac{\pi}{4}) \end{bmatrix} \quad (2.12)$$

Einige Multiplikationen mit Skalaren:

$$2, 5 \cdot \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix}, \quad \sqrt{2} \cdot \begin{bmatrix} -\sqrt{2} \\ \sqrt{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ \frac{\sqrt{2}}{2} \end{bmatrix} \quad (2.13)$$

Beispiel 2.1.6. [$n = 1: \mathbb{R}$] Auch einzelne Zahlen $v \in \mathbb{R}^1 = \mathbb{R}$ kann man als Vektoren $\vec{v} := [v]$ auffassen.

Bemerkung 2.1.7. Vektoren können nicht nur aus reellen Zahlen aufgebaut werden. So bilden die n -Tupel komplexer Zahlen mit der Addition und Multiplikation mit komplexen Zahlen den Vektorraum \mathbb{C}^n . Noch allgemeiner schreibt man \mathbb{K}^n für den Vektorraum der n -Tupel von Zahlen aus einem beliebigen Körper \mathbb{K} . In dieser Vorlesung verwenden wir die Bezeichnung \mathbb{K} oft in Definitionen oder Theoremen. Wir wollen damit ausdrücken, dass die Definition für beide Körper gleichermassen lautet und dass die Aussage des Theorems für beide Körper gilt.

Theorem 2.1.8 (Eigenschaften). *Für die Addition von Vektoren $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ gilt:*

$$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}) \quad (\text{VR1.1})$$

$$\vec{v} + \vec{0} = \vec{v} \quad \vec{v} + (-\vec{v}) = \vec{0} \quad (\text{VR1.2})$$

$$\vec{u} + \vec{v} = \vec{v} + \vec{u} \quad (\text{VR1.3})$$

Für die Multiplikation von Skalaren $\alpha, \beta \in \mathbb{R}$ mit Vektoren $\vec{u}, \vec{v} \in \mathbb{R}^n$ gilt:

$$\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v} \quad (\text{VR2.1})$$

$$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} \quad (\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v} \quad (\text{VR2.2})$$

$$1\vec{v} = \vec{v} \quad (\text{VR2.3})$$

Beweis. Die obigen Eigenschaften folgen unmittelbar aus den Definitionen für die Addition von Vektoren und die Multiplikation mit einem Skalar.

Bemerkung 2.1.9. Eine Menge mit Addition und Multiplikation mit Skalaren, für die eine entsprechende Aussage wie in Theorem 2.1.8 gilt, heißt *Vektorraum*. Also besagt das Theorem hier: der \mathbb{R}^n ist ein *Vektorraum*. Sobald wir uns an den \mathbb{R}^n gewöhnt haben, werden wir weitere Vektorräume kennenlernen und feststellen, dass mit ihnen genau so gerechnet werden kann wie mit \mathbb{R}^n .

2.2 BASIS UND DIMENSION

Ist der \mathbb{R}^3 in irgendeinem Sinne „größer“ als der \mathbb{R}^2 ? Beide haben unendlich viele Elemente. Trotzdem gibt es eine endliche Größe, die beide in ihrer „Kompliziertheit“ unterscheidet. Mathematiker nennen diese Größe „Dimension“, Physiker

oder Ingenieure oft auch „Freiheitsgrade“. In diesem Abschnitt werden wir die dafür nötigen Begriffe diskutieren.

Definition 2.2.1 (Linearkombination). Ein Vektor $\vec{u} \in \mathbb{R}^n$ heißt *Linearkombination* der Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$, wenn $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ existieren mit

$$\vec{u} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k. \quad (2.14)$$

Man sagt: \vec{u} lässt sich aus den $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ *linear kombinieren*. Die $\alpha_1, \alpha_2, \dots, \alpha_k$ heißen *Koeffizienten* der Linearkombination.

Beispiel 2.2.2. Der Vektor $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ ist Linearkombination der Vektoren $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ wegen

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.15)$$

Beispiel 2.2.3. Der Vektor $\begin{bmatrix} \pi \\ e \\ e \end{bmatrix}$ ist Linearkombination der Vektoren $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ und $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, denn

$$\begin{bmatrix} \pi \\ e \\ e \end{bmatrix} = e \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \pi \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad (2.16)$$

Beispiel 2.2.4. Der Vektor $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ist Linearkombination einer jeden Menge von Vektoren in \mathbb{R}^2 : man setzt einfach alle Koeffizienten auf Null.

Beispiel 2.2.5. Der Vektor $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ ist **nicht** Linearkombination der Vektoren $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Warum nicht?

Nehmen wir an, $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ wäre doch Linearkombination von $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Nach Definition der Linearkombination gäbe es dann reelle Zahlen α_1 und α_2 mit:

$$\alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \xLeftrightarrow{\text{Vektorrechnung}} \quad \begin{cases} \alpha_1 \cdot 1 + \alpha_2 \cdot 0 = 0 \\ \alpha_1 \cdot 0 + \alpha_2 \cdot 1 = 0 \\ \alpha_1 \cdot 0 + \alpha_2 \cdot 0 = 1 \end{cases}. \quad (2.17)$$

Die untere Gleichung auf der rechten Seite kann offensichtlich nicht erfüllt werden: Widerspruch.

Beispiel 2.2.6. Der Vektor $\begin{bmatrix} i \\ -i \end{bmatrix}$ (wobei $i = \sqrt{-1}$) ist eine (komplexe) Linearkombination der Vektoren $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ wegen

$$\begin{bmatrix} i \\ -i \end{bmatrix} = i \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-i) \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.18)$$

Definition 2.2.7 (Lineare Hülle). Die *lineare Hülle* der Menge $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$ ist die Menge aller Vektoren, die sich aus den $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ kombinieren lassen:

$$\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} := \{ \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k \mid \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R} \} \quad (2.19)$$

Statt lineare Hülle verwendet man oft auch den Begriff *Spann*.

Definition 2.2.8 (Erzeugendensystem des \mathbb{R}^n). Eine Menge $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ von Vektoren im \mathbb{R}^n heißt *Erzeugendensystem* des \mathbb{R}^n , wenn

$$\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} = \mathbb{R}^n. \quad (2.20)$$

Man sagt, die $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ *spannen* den \mathbb{R}^n *auf*.

Beispiel 2.2.9. Die Vektoren $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ spannen den \mathbb{R}^2 auf. Warum? Sei $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ ein beliebiger Vektor im \mathbb{R}^2 . Dann gilt:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \cdot 1 + v_2 \cdot 0 \\ v_1 \cdot 0 + v_2 \cdot 1 \end{bmatrix} = \underbrace{v_1}_{=: \alpha_1 \in \mathbb{R}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \underbrace{v_2}_{=: \alpha_2 \in \mathbb{R}} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.21)$$

Beispiel 2.2.10. Die Vektoren $\begin{bmatrix} 1-i \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1-i \end{bmatrix}$ spannen den \mathbb{C}^2 auf, denn für einen beliebigen Vektor $\begin{bmatrix} a+bi \\ c+di \end{bmatrix} \in \mathbb{C}^2$ gilt

$$\begin{bmatrix} a+bi \\ c+di \end{bmatrix} = \underbrace{\left(\frac{a-b}{2} + \frac{(a+b)}{2}i \right)}_{=: \alpha_1 \in \mathbb{C}} \cdot \begin{bmatrix} 1-i \\ 0 \end{bmatrix} + \underbrace{\left(\frac{c-d}{2} + \frac{(c+d)}{2}i \right)}_{=: \alpha_2 \in \mathbb{C}} \cdot \begin{bmatrix} 0 \\ 1-i \end{bmatrix}. \quad (2.22)$$

Definition 2.2.11 (*Lineare Unabhängigkeit*). Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ in \mathbb{R}^n heißt genau dann *linear unabhängig*, wenn die Vektorgleichung

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0}$$

für die Unbekannten $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ nur die Lösung: $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ hat, d.h. die Lösung $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ ist eindeutig.

Definition 2.2.12 (*Lineare Abhängigkeit*). Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ in \mathbb{R}^n heißt genau dann *linear abhängig*, wenn die Vektorgleichung

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0}$$

für die Unbekannten $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ nicht nur die Lösung: $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ hat; d.h. die Lösung ist *nicht eindeutig* und es gibt eine Lösung sodass nicht alle $\alpha_1, \alpha_2, \dots, \alpha_k$ gleich Null sind.

Beispiel 2.2.13. Die Vektoren $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ sind linear unabhängig, denn

$$\alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{0} \quad \text{Vektorrechnung} \quad \Leftrightarrow \quad \begin{cases} \alpha_1 \cdot 1 + \alpha_2 \cdot 0 = 0 \\ \alpha_1 \cdot 0 + \alpha_2 \cdot 1 = 0 \end{cases}. \quad (2.23)$$

Aus der oberen Gleichung liest man $\alpha_1 = 0$ ab, aus der unteren $\alpha_2 = 0$.

Bemerkung 2.2.14. Wenn die Vektoren komplizierter sind, braucht man anspruchsvollere Hilfsmittel (Lösen linearer Gleichungssysteme), um lineare Unabhängigkeit zu testen. Das ist Gegenstand der vierten Vorlesung.

Definition 2.2.15 (Basis des \mathbb{R}^n). Ein linear unabhängiges Erzeugendensystem des \mathbb{R}^n heißt *Basis* des \mathbb{R}^n .

Theorem 2.2.16 (Standardbasis des \mathbb{R}^n). *Die Vektoren*

$$\underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}_{=: \vec{e}_1}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}_{=: \vec{e}_2}, \dots, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{=: \vec{e}_n} \quad (2.24)$$

bilden eine Basis des \mathbb{R}^n , genannt Standardbasis des \mathbb{R}^n .

Die folgende Aussage folgt aus dem berühmten Steinitzschen Austauschsatz, den wir hier für \mathbb{R}^n formulieren aber nicht beweisen:

Theorem 2.2.17 (Steinitzschen Austauschsatz). *Alle Basen des \mathbb{R}^n haben die selbe Anzahl von Elementen, und diese ist - weil die Standardbasis n Elemente hat - gleich n .*

Damit sind wir soweit, den wichtigen Begriff der *Dimension* einzuführen:

Definition 2.2.18 (Dimension des Vektorraums \mathbb{R}^n). Die *Dimension* des \mathbb{R}^n ist die Anzahl n der Elemente in einer Basis des \mathbb{R}^n .

2.3 TEILRÄUME

Manche Teilmengen des \mathbb{R}^n haben analoge Eigenschaften wie der \mathbb{R}^n selbst:

Definition 2.3.1 (Teilraum). Eine Teilmenge T des \mathbb{R}^n heißt *Teilraum* oder *Untervektorraum* des \mathbb{R}^n , wenn folgendes gilt:

- (i) T ist nicht leer.
- (ii) Für alle $\vec{u}, \vec{v} \in T$ ist $\vec{u} + \vec{v}$ wieder in T .
- (iii) Für alle $\vec{v} \in T$ und alle $\alpha \in \mathbb{R}$ ist $\alpha \vec{v}$ wieder in T .

Man sagt, T ist *abgeschlossen* bzgl. der Vektorraumoperationen.

Bemerkung 2.3.2. Aus der Definition folgt sofort:

1. Jeder Teilraum muss den Nullvektor enthalten, denn: da T nicht leer ist, gibt es ein $\vec{v} \in T$. Dann muss nach (iii) auch $\vec{0} = 0 \cdot \vec{v}$ in T sein.
2. Die Menge $\{\vec{0}\}$ und der \mathbb{R}^n sind Teilräume des \mathbb{R}^n .

Die Begriffe *Erzeugendensystem*, *Basis*, der *Satz von Steinitz* und *Dimension* haben wir im letzten Abschnitt für die Situation des \mathbb{R}^n formuliert. Sie lassen sich aber ohne weiteres auf nichttriviale Teilräume T (d.h. $T \neq \{\vec{0}\}$) des \mathbb{R}^n übertragen. Für $T = \{\vec{0}\}$ definieren wir die leere Menge $\{\}$ als Basis. Es folgt $\dim(\{\vec{0}\}) = 0$.

Beispiel 2.3.3. Die Menge der Vektoren

$$T := \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3 \mid v_3 = 0 \right\} \quad (2.25)$$

ist ein Teilraum des \mathbb{R}^3 . Da sowohl die Vektoraddition als auch die Multiplikation mit Skalaren koordinatenweise erfolgen, bleibt bei allen diesen Operationen die dritte Koordinate Null, wie gefordert.

Eine Basis dieses Teilraums bilden z. B. die Vektoren

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \quad (2.26)$$

Daher hat T die Dimension Zwei.

Beispiel 2.3.4. Die Menge der Vektoren

$$T := \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2 \mid v_1 = v_2 \right\} = \left\{ \begin{bmatrix} v \\ v \end{bmatrix} \mid v \in \mathbb{R} \right\} \quad (2.27)$$

ist ein Teilraum des \mathbb{R}^2 . Warum? Seien $\begin{bmatrix} u \\ u \end{bmatrix}, \begin{bmatrix} v \\ v \end{bmatrix} \in T$ beliebig. Dann ist

$$\begin{bmatrix} u \\ u \end{bmatrix} + \begin{bmatrix} v \\ v \end{bmatrix} = \begin{bmatrix} u+v \\ u+v \end{bmatrix} \in T. \quad (2.28)$$

Ebenso prüft man für ein beliebiges $\alpha \in \mathbb{R}$:

$$\alpha \begin{bmatrix} v \\ v \end{bmatrix} = \begin{bmatrix} \alpha v \\ \alpha v \end{bmatrix} \in T. \quad (2.29)$$

Eine Basis von T bildet z. B. der Vektor $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Beispiel 2.3.5. Die Menge der Vektoren

$$T := \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2 \mid v_1 = v_2^2 \right\} = \left\{ \begin{bmatrix} v^2 \\ v \end{bmatrix} \mid v \in \mathbb{R} \right\} \quad (2.30)$$

ist **kein** Teilraum des \mathbb{R}^2 . Warum nicht? Betrachte den Vektor $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \in T$. Der Vektor $2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ ist aber nicht in T , da $2 \neq 2^2$.

Bemerkung 2.3.6. Alle Begriffe in diesem Kapitel sind analog auch für komplexe Zahlen \mathbb{C} , ja sogar für jeden beliebigen Körper \mathbb{K} definiert.

3. KAPITEL

MATRIZEN

In dieser Vorlesung werden wir lernen was Matrizen sind und wie man mit Matrizen rechnet. Diese Erkenntnisse werden wir im Verlauf der Vorlesung immer wieder verwenden.

Warum Matrizen? Matrizen ermöglichen eine kompakte Darstellungsform vieler mathematischer Strukturen. Zum Beispiel: Lineare Gleichungssysteme können in einer einzigen Matrixgleichung geschrieben werden; lineare Abbildungen (z. B. lineare Filter in der Bildbearbeitung) können durch Matrizen beschrieben werden und digitale Bilder werden durch Matrizen im Rechner codiert. Rechnungen mit diesen Objekten werden durch Rechnungen mit Matrizen beschrieben: weiß man, wie man mit Matrizen rechnet, so weiß man, wie man mit linearen Abbildungen rechnet.

3.1 ANWENDUNGEN DIE AUF MATRIZEN FÜHREN

Digitale Bilder werden durch m Zeilen und n Spalten beschrieben. Jede Zeile besteht aus n , jede Spalte aus m Pixeln. Jedes der Pixel eines Schwarz-Weißbildes (sw-Bild) hat einen bestimmten Helligkeitswert, der durch eine reelle Zahl charakterisiert ist. Mittelgrau entspricht etwa dem Helligkeitswert 0, ein großer positiver Helligkeitswert bedeutet (fast) weiß, ein negativer Helligkeitswert mit großem Betrag ist dann (fast) schwarz. (Bei Farbbildern gibt es drei Helligkeitswerte, je einen für rot, grün und blau). Ein sw-Bild ist also ein Zahlenschema der Gestalt

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}.$$

Ein solches Zahlenschema nennt man Matrix (vgl. unten). Bilder mit der selben Anzahl von Zeilen und Spalten, d. h. vom gleichen Format, können addiert werden indem die Helligkeitswerte der sich entsprechenden Pixel addiert werden. Aus einem Bild mit einer Katze und einem zweiten mit einem Hund wird dann ein Bild mit einer Katze und einem Hund (Doppelbelichtung). Wenn jeder Helligkeitswert mit ein und derselben Zahl größer Eins multipliziert wird, ergibt sich wieder ein Bild: Es ist kontrastreicher. Fazit: Bilder (Matrizen) können addiert und mit einer reellen Zahl multipliziert werden, so wie wir dies in der vorangehenden Vorlesung über das Rechnen mit Vektoren aus \mathbb{K}^n kennen gelernt haben.

Das nächste Beispiel zeigt, wie ein Zahlenschema von der obigen Gestalt zur Beschreibung einer Aufgabe benutzt werden kann, die von einer grundsätzlich

anderen Natur ist. Wir haben bereits gesehen, dass der Raum, in dem wir leben, durch den Vektorraum \mathbb{R}^3 modelliert werden kann. Jetzt betrachten wir eine Fläche durch den Ursprung, und darauf einen senkrecht stehenden Vektor, die Normale \vec{n} . Die Sonne erzeugt ein elektromagnetisches Feld, das auf die Fläche auftrifft und einen Druck (= Kraft pro Flächeneinheit) $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$ ausübt. Der Druck ist durch folgenden mathematischen Ausdruck berechenbar:

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^3 t_{1,j} n_j \\ \sum_{j=1}^3 t_{2,j} n_j \\ \sum_{j=1}^3 t_{3,j} n_j \end{bmatrix} \quad (3.1)$$

Die Zahlen $t_{i,k}$ bilden den sogenannten Maxwellschen Spannungstensor und sind durch das elektrische Feld \vec{E} und das magnetische Feld \vec{B} ausdrückbar. Zum Beispiel ist $t_{1,2} = 1/4\pi(E_1 E_2 + B_1 B_2)$. Die obige Formel bildet also den Normalenvektor \vec{n} auf den Druckvektor \vec{p} ab. Bei schönem Wetter ist die Kraft der Sonnenstrahlung auf ein Feld von einem Quadratkilometer so groß wie die Schwerkraft einer Masse von einem halben Kilo.

Nun gehen wir dazu über, das Konzept der Matrizen korrekt zu definieren und zu erklären, wie damit gerechnet wird. Im folgenden steht \mathbb{K} entweder überall für \mathbb{R} oder überall für \mathbb{C} .

3.2 DEFINITION UND BEISPIELE VON MATRIZEN

Definition 3.2.1 (Matrix). Für Zahlen $a_{i,j} \in \mathbb{K}$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, heißt das Zahlenschema

$$A := \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} := (a_{i,j}) := (a_{i,j})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \quad (3.2)$$

eine *Matrix vom Format $(m \times n)$ mit Einträgen in \mathbb{K} oder kurz $(m \times n)$ -Matrix*.

Die Menge aller Matrizen vom Format $(m \times n)$ mit Einträgen in \mathbb{K} wird mit $\mathbb{K}^{m,n}$ bezeichnet.

Bemerkung 3.2.2. Die Indizes setzt man nach der Regel: Zeile vor Spalte.

Beispiel 3.2.3. Einige Matrizen:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2,3}, \quad \begin{bmatrix} i & 2+3i \\ 1-i & 4 \end{bmatrix} \in \mathbb{C}^{2,2} \quad (3.3)$$

Beispiel 3.2.4. Vektoren aus \mathbb{K}^n können als $(n \times 1)$ -Matrizen aufgefasst werden:

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \in \mathbb{R}^{3,1} = \mathbb{R}^3 \quad (3.4)$$

Definition 3.2.5 (Quadratische Matrix).

Matrizen vom Format $(n \times n)$ heißen *quadratisch*.

Definition 3.2.6 (Einheitsmatrix).

Die Matrix $I_n \in \mathbb{K}^{n,n}$ mit

$$I_n := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (3.5)$$

heißt *n-dimensionale Einheitsmatrix*.

Definition 3.2.7 (Diagonalmatrix). Quadratische Matrizen der Gestalt

$$\begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{bmatrix} \quad (3.6)$$

mit $a_1, a_2, \dots, a_n \in \mathbb{K}$ heißen *Diagonalmatrizen*.

Definition 3.2.8 (Obere Dreiecksmatrix).

Quadratische Matrizen der Gestalt

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{bmatrix} \quad (3.7)$$

heißen *obere Dreiecksmatrizen*.

Definition 3.2.9 (Die Transponierte Matrix).

Gegeben sei die Matrix

$$A = (a_{i,j}) \in \mathbb{K}^{m,n}.$$

Jetzt definieren wir eine Matrix in $\mathbb{K}^{n,m}$ folgendermaßen:

$$A^T := (b_{i,j}) \text{ mit } b_{i,j} := a_{j,i}.$$

Sie heißt die zu A *transponierte Matrix*.

Definition 3.2.10 (Symmetrische Matrizen).

Eine Matrix $A = (a_{i,j}) \in \mathbb{K}^{n,n}$ heißt *symmetrisch* falls gilt

$$A = A^T \text{ oder explizit } (a_{i,j}) = (a_{j,i}) \text{ für alle vorkommenden Indizes.}$$

Definition 3.2.11 (Antisymmetrische Matrizen).

Eine Matrix $A = (a_{i,j}) \in \mathbb{K}^{n,n}$ heißt *antisymmetrisch* (oder *schief-symmetrisch*) falls gilt

$$A = -A^T \text{ oder explizit } (a_{i,j}) = -(a_{j,i}) \text{ für alle vorkommenden Indizes.}$$

Definition 3.2.12 (Adjungierte Matrizen).

Für $A = (a_{i,j}) \in \mathbb{C}^{n,n}$ heißt die Matrix

$$A^* := \overline{A^T} \text{ oder explizit } (a_{i,j}^*) := \overline{(a_{j,i})} \text{ für alle vorkommenden Indizes,}$$

die zu A *adjungierte Matrix*.

Definition 3.2.13 (Selbstadjungierte Matrizen).

Eine Matrix $A = (a_{i,j}) \in \mathbb{C}^{n,n}$ heißt *selbstadjungiert* falls gilt

$$\begin{aligned} A &= A^* \text{ oder explizit} \\ a_{i,j} &= \overline{a_{j,i}} \text{ für alle vorkommenden Indizes} \end{aligned}$$

3.3 MATRIZEN ALS VEKTORRAUM

Mit Matrizen in $\mathbb{K}^{m,n}$ kann man rechnen wie mit Vektoren:

Definition 3.3.1 (Vektorraumoperationen). Gegeben: $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{K}^{m,n}, \alpha \in \mathbb{K}$.

Wir definieren die folgenden Matrizen in $\mathbb{K}^{m,n}$:

$$\begin{aligned} A + B &:= (a_{i,j} + b_{i,j}) && \text{(Summe)} \\ \alpha A &:= (\alpha a_{i,j}) && \text{(Multiplikation mit Skalaren)} \end{aligned}$$

Beispiel 3.3.2.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 5 & 8 & 11 \end{bmatrix}, \quad (3.8)$$

$$2 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix}, \quad (3.9)$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}. \quad (3.10)$$

In der letzten Vorlesung haben wir die Eigenschaften des Vektorraumes \mathbb{K}^n in Theorem (2.1.8) zusammengestellt. Eine analoge Aussage gilt auch hier:

Theorem 3.3.3 (Eigenschaften). Für die Addition von Matrizen $A, B, C, 0 \in \mathbb{K}^{m,n}$ wobei die Nullmatrix 0 , definitionsgemäß aus lauter Nullen besteht, gelten die Rechenregeln:

$$(A + B) + C = A + (B + C) \quad (\text{VR1.1})$$

$$A + 0 = A \quad (\text{VR1.2})$$

$$A + B = B + A \quad (\text{VR1.3})$$

Für die Multiplikation von Skalaren $\alpha, \beta \in \mathbb{K}$ mit A, B in $\mathbb{K}^{m,n}$ gilt:

$$\alpha(\beta A) = (\alpha\beta)A \quad (\text{VR2.1})$$

$$\alpha(A + B) = \alpha A + \alpha B \quad (\alpha + \beta)A = \alpha A + \beta A \quad (\text{VR2.2})$$

$$1A = A \quad (\text{VR2.3})$$

Bemerkung 3.3.4 (Vektorraum der Matrizen). Im Vorgriff auf das abstrakte Konzept “Vektorraum”, das im Kapitel 5 eingeführt werden wird, besagt das Theorem, dass die Menge der Matrizen $\mathbb{K}^{m,n}$ zusammen mit den beiden Operationen *Addition* und *Multiplikation mit einem Skalar* ein Vektorraum darstellt.

3.4 MATRIZENMULTIPLIKATION

Etwas komplizierter ist die Multiplikation von Matrizen. Achtung! Zwei Matrizen können nur dann multipliziert werden, wenn ihre Formate zusammenpassen, wie dies in der folgenden Definition ausgeführt ist:

Definition 3.4.1 (Multiplikation).

Gegeben: $A = (a_{i,j}) \in \mathbb{K}^{m,n}$, $B = (b_{j,k}) \in \mathbb{K}^{n,p}$.

Wir definieren folgende Matrix in $\mathbb{K}^{m,p}$ und nennen sie das *Produkt der Matrizen A und B*:

$$AB := \left(\sum_{j=1}^n a_{i,j} b_{j,k} \right)_{\substack{i=1,\dots,m \\ k=1,\dots,p}}$$

Ausgeschrieben bedeutet das:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} & \vdots & b_{1,p} \\ b_{2,1} & b_{2,2} & \vdots & b_{2,p} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n,1} & b_{n,2} & \vdots & b_{n,p} \end{bmatrix} \\ := \begin{bmatrix} \sum_{j=1}^n a_{1,j} b_{j,1} & \sum_{j=1}^n a_{1,j} b_{j,2} & \cdots & \sum_{j=1}^n a_{1,j} b_{j,p} \\ \sum_{j=1}^n a_{2,j} b_{j,1} & \sum_{j=1}^n a_{2,j} b_{j,2} & \cdots & \sum_{j=1}^n a_{2,j} b_{j,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^n a_{m,j} b_{j,1} & \sum_{j=1}^n a_{m,j} b_{j,2} & \cdots & \sum_{j=1}^n a_{m,j} b_{j,p} \end{bmatrix}$$

Beispiel 3.4.2. Einige Produkte:

$$\begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 3 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 4 + 3 \cdot 5 & 1 \cdot 2 + 2 \cdot 3 + 3 \cdot 6 \\ 6 \cdot 1 + 5 \cdot 4 + 4 \cdot 5 & 6 \cdot 2 + 5 \cdot 3 + 4 \cdot 6 \end{bmatrix} \\ = \begin{bmatrix} 24 & 26 \\ 46 & 51 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{bmatrix} \begin{bmatrix} 1 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} = \text{nicht definiert,} \quad (\text{Formate müssen stimmen!})$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad (I_n \text{ neutral})$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 30 \end{bmatrix}, \quad (\text{Vereinfachung bei Diag.-Matr.})$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (\text{es gibt Nullteiler})$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (\textbf{Achtung:} \text{ Kommutativitat } \dots)$$

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (\dots \text{ gilt } \textbf{nicht!})$$

Definition 3.4.3 (Inverse). Eine quadratische Matrix $A \in \mathbb{K}^{n,n}$ heit *invertierbar*, wenn es eine Matrix $B \in \mathbb{K}^{n,n}$ gibt mit

$$AB = I_n. \quad (3.11)$$

Die Matrix B ist dann eindeutig und heit *die Inverse* zu A . Sie wird mit A^{-1} bezeichnet.

Bemerkung 3.4.4. Vorsicht! Nicht alle Matrizen sind invertierbar, z.B. gibt es keine Matrix B , sodass $A^{-1} = B$ fur

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

Lemma 3.4.5. Fur eine invertierbare Matrix $A \in \mathbb{K}^{n,n}$ gilt auch $A^{-1}A = I_n$.

Beispiel 3.4.6. Einige inverse Matrizen:

$$\begin{aligned} I_n^{-1} &= I_n, & (I_n \text{ neutral}) \\ \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}^{-1} &= \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{3} \end{bmatrix}, & (\text{Vereinfachung bei Diag.-Matr.}) \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} &= ? & (3.12) \end{aligned}$$

In der nachsten Vorlesung werden wir den *Gaualgorithmus* lernen, mit dem wir die Inverse einer Matrix systematisch berechnen konnen.

Es gibt einige wichtige Gruppen von Matrizen, die eine Inverse zu berechnen besonders leicht zulassen:

Definition 3.4.7 (Orthogonale Matrizen).

Eine Matrix $Q = (q_{i,j}) \in \mathbb{R}^{n,n}$ heit *orthogonal* falls gilt

$$Q^T Q = I_n.$$

Definition 3.4.8 (Unitare Matrizen).

Eine Matrix $U = (a_{i,j}) \in \mathbb{C}^{n,n}$ heit *unitar* falls gilt

$$U^* U = I_n.$$

Dabei verwenden wir die oben eingefuhrte Bezeichnung: $U^* := \overline{U^T}$.

Wir fassen einige wichtige Eigenschaften der Matrix-Multiplikation in dem folgenden Satz zusammen.

Theorem 3.4.9. *Gegeben: $m, n, p, q \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} .*

Dann:

$$\begin{array}{lll}
 (AB)C = A(BC) & \forall A \in \mathbb{K}^{m,n}, & \text{(Assoziativitt)} \\
 & \forall B \in \mathbb{K}^{n,p}, & \\
 & \forall C \in \mathbb{K}^{p,q}; & \\
 A(B + C) = AB + AC & \forall A \in \mathbb{K}^{m,n}, & \text{(Distributivitt)} \\
 & \forall B \in \mathbb{K}^{n,p}, & \\
 & \forall C \in \mathbb{K}^{n,p}; & \\
 A(\alpha B) = \alpha(AB) & \forall A \in \mathbb{K}^{m,n}, & \\
 & \forall B \in \mathbb{K}^{n,p}, & \\
 & \forall \alpha \in \mathbb{K}; & \\
 (AB)^T = B^T A^T & \forall A \in \mathbb{K}^{m,n}, & \text{(Multiplikation-Transposition)} \\
 & \forall B \in \mathbb{K}^{n,p}, & \\
 (AB)^{-1} = B^{-1} A^{-1} & \forall \text{invertierbaren Matrizen } A, B \in \mathbb{K}^{n,n} & \\
 & \text{(Multiplikation-Inversion)} & \\
 AI_n = A & \forall A \in \mathbb{K}^{m,n}, & \text{(Neutrales Element)} \\
 I_m A = A & & \\
 (A^{-1})^T = (A^T)^{-1} & \forall \text{invertierbare Matrizen } A \in \mathbb{K}^{n,n}. & \\
 & \text{(Inversion-Transposition)} &
 \end{array}$$

Bemerkung 3.4.10. Wir haben schon gesehen, dass wir den \mathbb{K}^n auch als $\mathbb{K}^{n,1}$ auffassen knnen. Damit haben auch sofort die Definition der Multiplikation einer Matrix $A \in \mathbb{K}^{m,n}$ mit einem Vektor $\vec{v} \in \mathbb{K}^n$, indem wir den Vektor einfach

als Matrix $\vec{v} := \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{K}^{n,1}$ auffassen und dann liefert uns die Matrizenmultiplikation $\begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = A\vec{v}$ das Ergebnis in $\mathbb{K}^{m,1}$, welches wir dann wieder als Vektor $\vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \in \mathbb{K}^m$ auffassen knnen.

3.5 LINEARE ABBILDUNGEN UND MATRIZEN

Wir werden uns nun mit Abbildungen von \mathbb{K}^n nach \mathbb{K}^m befassen.

Definition 3.5.1 (Lineare Abbildung). Eine Abbildung

$$L : \begin{cases} \mathbb{K}^n & \rightarrow \mathbb{K}^m \\ \vec{v} & \mapsto L(\vec{v}) \end{cases} \quad (3.13)$$

heißt *linear*, wenn folgendes gilt:

$$L(\vec{u} + \vec{v}) = L(\vec{u}) + L(\vec{v}) \quad \forall \vec{u}, \vec{v} \in \mathbb{K}^n, \quad (\text{LA1})$$

$$L(\alpha \vec{v}) = \alpha L(\vec{v}) \quad \forall \vec{v} \in \mathbb{K}^n, \alpha \in \mathbb{K}. \quad (\text{LA2})$$

Lineare Abbildungen (auch Homomorphismen genannt) sind dadurch charakterisiert, dass das Bild der Summe von Vektoren gleich der Summe der Bilder der Vektoren ist, d. h. die lineare Struktur von Urbild (= Ausgangsraum) und Bildraum (= Zielraum) wird erhalten. Die Menge aller linearen Abbildungen L von \mathbb{K}^n nach \mathbb{K}^m wird mit $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ bezeichnet.

Beispiel 3.5.2. Aus der Definition erhalten wir:

1. Für alle linearen Abbildungen L gilt: $L(\vec{0}) = \vec{0}$, denn: $L(\vec{0}) = L(0\vec{v}) = 0 \cdot L(\vec{v}) = \vec{0}$, egal welches \vec{v} wir einsetzen.
2. Die folgenden Abbildungen von \mathbb{R}^3 nach \mathbb{R}^2 sind linear:

$$L \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) := \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad (3.14)$$

$$L \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) := \begin{bmatrix} v_1 + v_2 \\ v_2 + v_3 \end{bmatrix}, \quad (3.15)$$

$$L \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) := \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (3.16)$$

Warum? Man muss die Bedingungen aus der Definition prüfen. Zum Beispiel in (3.14):

(LA1): Seien $\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ und $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ aus \mathbb{R}^3 beliebig. Dann gilt:

$$\begin{aligned} L \left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) &= L \left(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix} \right) \\ &= \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix} \\ &= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= L \left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \right) + L \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) \end{aligned} \quad (3.17)$$

(LA2): Seien $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ aus \mathbb{R}^3 und $\alpha \in \mathbb{R}$ beliebig. Dann gilt:

$$\begin{aligned} L \left(\alpha \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) &= L \left(\begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \alpha v_3 \end{bmatrix} \right) \\ &= \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \end{bmatrix} \\ &= \alpha \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \alpha L \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) \end{aligned} \quad (3.18)$$

Damit ist L linear. □

Beispiel 3.5.3. Die folgenden Abbildungen von \mathbb{R}^3 nach \mathbb{R}^2 sind **nicht** linear:

$$N \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} := \begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} \quad (3.19)$$

$$N \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} := \begin{bmatrix} v_1 \cdot v_2 \\ v_2 \cdot v_3 \end{bmatrix} \quad (3.20)$$

$$N \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (3.21)$$

Warum nicht? Zum Beispiel gilt für (3.19):

$$\begin{aligned} N \left(2 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) &= N \left(\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} 4 \\ 0 \end{bmatrix} \\ &\neq \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \cdot N \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \end{aligned} \quad (3.22)$$

Damit gilt (LA2) nicht für N , und N ist nicht linear. \square

Zwei wichtige Teilräume, die zu einer linearen Abbildung gehören, sind Kern und Bild. Wir definieren vorerst Kern und Bild als Mengen. Danach ist zu zeigen, dass es sich um Teilräume von Urbild respektive Bildmenge handelt.

Definition 3.5.4 (Kern und Bild). Gegeben sei eine lineare Abbildung L von \mathbb{K}^n nach \mathbb{K}^m . Die Menge aller Vektoren im \mathbb{K}^n , die durch L auf den Nullvektor in \mathbb{K}^m abgebildet wird, also die Menge

$$\text{Kern}(L) := \{ \vec{v} \in \mathbb{K}^n \mid L(\vec{v}) = \vec{0} \} \quad (3.23)$$

heißt *Kern* von L .

Ferner heißt die Menge aller Vektoren im \mathbb{K}^m , die ein Urbild unter L haben, also die Menge

$$\text{Bild}(L) := \{ L(\vec{v}) \in \mathbb{K}^m \mid \vec{v} \in \mathbb{K}^n \} \quad (3.24)$$

das *Bild* von L .

Lemma 3.5.5. Sei $L : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ eine lineare Abbildung.

1. Der Kern von L ist ein Teilraum des Urbildraums \mathbb{K}^n .
2. Das Bild von L ist ein Teilraum des Bildraums \mathbb{K}^m .

Beispiel 3.5.6. Kern und Bild von $L \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) := \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$ sind

$$\text{Kern}(L) = \left\{ \begin{bmatrix} 0 \\ 0 \\ v \end{bmatrix} \mid v \in \mathbb{R} \right\} \quad (3.25)$$

$$\text{Bild}(L) = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} \mid v_1, v_2 \in \mathbb{R} \right\}. \quad (3.26)$$

Mit Hilfe der Matrizen lassen sich nun lineare Abbildungen von \mathbb{K}^n nach \mathbb{K}^m charakterisieren. Wir fassen dies in einem Satz zusammen:

Theorem 3.5.7. *Lineare Abbildungen von $\mathbb{K}^n \rightarrow \mathbb{K}^m$ lassen sich durch Matrizen in $\mathbb{K}^{m,n}$ darstellen. Genauer:*

(i) Sei $A \in \mathbb{K}^{m,n}$. Dann ist

$$L_A : \begin{cases} \mathbb{K}^n & \rightarrow \mathbb{K}^m \\ \vec{v} & \mapsto A\vec{v} \end{cases} \quad (3.27)$$

(mit $A\vec{v}$ aufgefasst als Matrizenprodukt) eine lineare Abbildung.

(ii) Sei $L \in \text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$. Dann gibt es genau eine Matrix $A_L \in \mathbb{K}^{m,n}$ mit

$$L(\vec{v}) = A_L \vec{v} \quad \forall \vec{v} \in \mathbb{K}^n. \quad (3.28)$$

Die Matrix A_L heißt die zu L zugehörige Matrix von L . Die Spaltenvektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ sind die Bilder der Standardbasisvektoren, d.h. $A_L \vec{e}_i = \vec{a}_i$ für $i = 1, \dots, n$.

(iii) Seien $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$. Dann ist die Komposition (Hintereinanderausführung)

$$L_A \circ L_B : \begin{cases} \mathbb{K}^p & \rightarrow \mathbb{K}^m \\ \vec{v} & \mapsto L_A(L_B(\vec{v})) \end{cases} \quad (3.29)$$

eine lineare Abbildung mit zugehöriger Matrix AB , d. h.

$$L_A \circ L_B = L_{AB}. \quad (3.30)$$

Bemerkung 3.5.8 (Wichtige Konvention). Die durch die Matrix A zugehörige lineare Abbildung L_A von \mathbb{K}^n nach \mathbb{K}^m wird zur Abkürzung im Folgenden mit A bezeichnet, d.h. die Matrix A und die zugehörige lineare Abbildung L_A wird identifiziert. Ebenso wird die lineare Abbildung L und die zugehörige Matrix A_L identifiziert. Wir schreiben also für A, L_A und A_L immer nur A . Um diese Situation hervorzuheben sprechen wir auch von einer *Matrixabbildung*.

Bemerkung 3.5.9. Das *Bild* einer Matrixabbildung (vgl. Bemerkung 3.5.8) ist die lineare Hülle der Spaltenvektoren, i.e. sei $A \in \mathbb{K}^{m,n}$ eine Matrix mit Spaltenvektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$. Dann gilt für die zugehörigen Matrixabbildung:

$$\text{Bild}(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$$

4. KAPITEL

DER GAUSSALGORITHMUS

In dieser Vorlesung werden wir lernen, wie man ein lineares Gleichungssystem (LGS) systematisch löst. Ein Verfahren (unter vielen) ist der *Gaußalgorithmus*. Es ist der wichtigste Algorithmus der linearen Algebra.

4.1 BEISPIELE

Wir rechnen zunächst Beispiele aus, auch wenn noch nicht alle Begriffe formal eingeführt worden sind. Das wird im nächsten Abschnitt nachgeholt.

Beispiel 4.1.1. Betrachte

$$\left| \begin{array}{rcl} 3x_1 + 5x_2 & = & -2 \\ x_1 - 6x_2 & = & 7 \end{array} \right|.$$

Gesucht ist die Menge aller reellen Zahlen x_1 und x_2 , sodass beide Gleichungen erfüllt sind. So etwas ist ein lineares Gleichungssystem (LGS), und was wir suchen, ist seine *Lösungsmenge*.

Wir ermitteln x_1 und x_2 durch geschickte Umformungen:

$$\begin{aligned} \left| \begin{array}{rcl} 3x_1 + 5x_2 & = & -2 \\ x_1 - 6x_2 & = & 7 \end{array} \right| &\longrightarrow \left| \begin{array}{rcl} x_1 - 6x_2 & = & 7 \\ 3x_1 + 5x_2 & = & -2 \end{array} \right| & \text{(Zeilen tauschen)} \\ &\longrightarrow \left| \begin{array}{rcl} x_1 - 6x_2 & = & 7 \\ 23x_2 & = & -23 \end{array} \right| & \text{(2. Zeile} - 3 \cdot \text{1. Zeile in 2. Zeile)} \\ &\longrightarrow \left| \begin{array}{rcl} x_1 - 6x_2 & = & 7 \\ x_2 & = & -1 \end{array} \right| & \text{(2. Zeile} \cdot \frac{1}{23}) \\ &\longrightarrow \left| \begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & -1 \end{array} \right| & \text{(1. Zeile} + 6 \cdot \text{2. Zeile in 1. Zeile)} \end{aligned}$$

Offensichtlich ist $x_1 = 1$, $x_2 = -1$, d.h. es gibt also genau eine Lösung oder anders ausgedrückt, die Lösungsmenge besteht aus einem einzigen Element. Oft sagt man auch das LGS hat eine eindeutige Lösung.

Wie haben wir das gemacht? Zunächst stellen wir fest, dass wir uns einige Schreibarbeit sparen können, wenn wir das LGS in Matrixschreibweise behandeln. Die Information steckt nämlich allein in den Koeffizienten vor den Variablen und in den rechten Seiten.

Beispiel 4.1.2. Dasselbe Beispiel nochmal in Matrixschreibweise. In der i -ten Spalte stehen die Koeffizienten von x_i im LGS. Als weitere Spalte fügen wir noch die rechten Seiten hinzu.

$$\begin{aligned} \left[\begin{array}{cc|c} 3 & 5 & -2 \\ 1 & -6 & 7 \end{array} \right] &\longrightarrow \left[\begin{array}{cc|c} 1 & -6 & 7 \\ 3 & 5 & -2 \end{array} \right] & \text{(Zeilen tauschen)} \\ &\longrightarrow \left[\begin{array}{cc|c} 1 & -6 & 7 \\ 0 & 23 & -23 \end{array} \right] & \text{(2. Zeile} - 3 \cdot \text{1. Zeile in 2. Zeile)} \\ &\longrightarrow \left[\begin{array}{cc|c} 1 & -6 & 7 \\ 0 & 1 & -1 \end{array} \right] & \text{(2. Zeile} \cdot \frac{1}{23}) \\ &\longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -1 \end{array} \right] & \text{(1. Zeile} + 6 \cdot \text{2. Zeile in 1. Zeile)} \end{aligned}$$

Beobachtung 4.1.3. Die sogenannten elementaren Zeilenoperationen, die wir in diesem Beispiel geschickt benutzt haben, sind:

- Vertauschen von Zeilen
- Addieren eines Vielfachen einer Zeile zu einer anderen Zeile
- Multiplikation einer Zeile mit einer Zahl ungleich Null

Die gesuchten Zahlen x_1, x_2 schreiben wir in einen Spaltenvektor $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ und stellen fest: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ ist genau dann eine Lösung des LGSs, wenn $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ auch eine Lösung des durch elementare Zeilenoperationen umgeformten LGSs ist.

Beobachtung 4.1.4. Angenommen, wir haben für ein LGS mit n Variablen und n Gleichungen durch elementare Zeilenumformungen folgende Matrix erzeugt:

$$\left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & b_1 \\ 0 & 1 & \dots & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & b_n \end{array} \right] \quad (4.1)$$

Dann ist

$$\begin{aligned} x_1 &= b_1 \\ x_2 &= b_2 \\ &\vdots \\ x_n &= b_n \end{aligned}$$

die eindeutige Lösung des LGSs.

Der Gaußalgorithmus ist eine Methode, die mit elementaren Zeilenoperationen die angenehme obige Situation (oder eine ähnliche) sicher und ohne Umwege erreicht.

Bevor wir das Verfahren im allgemeinen Fall beschreiben, zeigen wir noch, dass die Lösung eines LGSs nicht eindeutig sein muss.

Beispiel 4.1.5. Wir wollen das folgende LGS lösen:

$$\left| \begin{array}{l} 3x_1 + 6x_2 = 9 \\ 4x_1 + 8x_2 = 12 \end{array} \right|$$

Wir arbeiten in Matrixschreibweise:

$$\begin{aligned} \left[\begin{array}{cc|c} 3 & 6 & 9 \\ 4 & 8 & 12 \end{array} \right] &\longrightarrow \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 8 & 12 \end{array} \right] & (1. \text{ Zeile} \cdot \tfrac{1}{3}) \\ &\longrightarrow \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 0 \end{array} \right] & (2. \text{ Zeile} - 4 \cdot 1. \text{ Zeile}) \end{aligned}$$

Wie sollen wir das interpretieren? Die in der letzten Matrix enthaltene Information besagt:

$$\left| \begin{array}{l} x_1 + 2x_2 = 3 \\ 0 = 0 \end{array} \right|$$

Man drückt das wie folgt aus: x_2 ist ein frei wählbarer *Parameter* und $x_1 = 3 - 2x_2$. Benennt man x_2 noch in s um, so kann die Lösungsmenge als Menge von Vektoren geschrieben werden als

$$\left\{ \begin{bmatrix} 3-2s \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\}. \quad (4.2)$$

Beispiel 4.1.6. Wir wollen das folgende LGS lösen:

$$\left| \begin{array}{l} 3x_1 + 6x_2 = 9 \\ 4x_1 + 8x_2 = 13 \end{array} \right|$$

$$\begin{aligned} \left[\begin{array}{cc|c} 3 & 6 & 9 \\ 4 & 8 & 13 \end{array} \right] &\longrightarrow \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 8 & 13 \end{array} \right] & (1. \text{ Zeile} \cdot \tfrac{1}{3}) \\ &\longrightarrow \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 1 \end{array} \right] & (2. \text{ Zeile} - 4 \cdot 1. \text{ Zeile}) \end{aligned}$$
$$\left| \begin{array}{rcl} x_1 + 2x_2 & = & 3 \\ 0 & = & 1 \end{array} \right|$$

4.2 LINEARE GLEICHUNGSSYSTEME

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n & = & b_m \end{array} \quad (4.3)$$

Die Matrix

$$A := (a_{i,j})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \quad (4.4)$$

heißt die *Koeffizientenmatrix* des LGSs. Der Vektor

$$\vec{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (4.5)$$

heißt *Inhomogenität* oder *rechte Seite* des LGSs.

Die Matrix

$$[A|\vec{b}] := \left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array} \right] \quad (4.6)$$

heißt *erweiterte Koeffizientenmatrix* des LGSs.

Ein Vektor

$$\vec{x} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (4.7)$$

heißt *Lösung* oder *Lösungsvektor* des LGSs, wenn seine Koordinaten x_j , $j = 1, 2, \dots, n$, alle Gleichungen des LGSs erfüllen.

Die Menge aller Lösungsvektoren \vec{x} heißt die *Lösungsmenge* des LGSs.

Bemerkung 4.2.2. A, \vec{b}, \vec{x} spielen verschiedene Rollen: A, \vec{b} sind gegeben. \vec{x} ist gesucht.

Bemerkung 4.2.3. Mit Hilfe der Matrizenmultiplikation können wir das LGS nun kurz als

$$A\vec{x} = \vec{b} \quad (4.8)$$

schreiben.

Ein LGS lösen heißt von nun an, seine gesamte Lösungsmenge zu bestimmen. Der Rest der Vorlesung beschreibt ein Verfahren, mit dem man jedes LGS lösen kann. Dafür benötigen wir die drei elementaren Zeilenoperationen.

Definition 4.2.4 (Elementare Zeilenoperationen). Die *elementaren Zeilenoperationen* in einer Matrix A sind:

- Das Tauschen von Zeilen.
- Das Addieren eines Vielfachen einer Zeile zu einer anderen Zeile.
- Die Multiplikation einer Zeile mit einer Zahl ungleich Null aus \mathbb{K} .

Theorem 4.2.5 (Invarianz der Lösungsmenge). Sei $[A|\vec{b}]$ die erweiterte Koeffizientenmatrix des LGSs (4.3). Dann gilt: Elementare Zeilenoperationen verändern seine Lösungsmenge nicht.

[illegible]

Mit Hilfe der ZSF kann man eine wichtige Größe, den *Rang einer Matrix* einführen.

Definition 4.3.2. Der *Rang* einer Matrix ist die maximale Anzahl linear unabhängiger *Zeilen* oder – was dasselbe ist – die maximale Anzahl linear unabhängiger *Spalten* der Matrix.

Theorem 4.3.3. Der Rang einer Matrix A kann so berechnet werden: Bringe A auf ZSF. Dann ist der Rang gleich der Anzahl der nichtverschwindenden Zeilen (also die Zahl der Köpfe der ZSF von A).

Definition 4.3.4 (Anwendung der Konzepte ZSF und NZSF auf lineare Gleichungssysteme). Ein LGS ist in *ZSF* bzw. in *NZSF*, wenn seine erweiterte Koeffizientenmatrix in ZSF bzw. in NZSF ist. Ein LGS ist also in NZSF falls es folgende Gestalt hat:

$$\left[\begin{array}{cccccccccccccccccccc|c} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & b_1 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & b_2 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & b_3 \\ \vdots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \dots & \dots & \dots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & * & \dots & * & \dots & \dots & \dots & \dots & b_r \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \dots & \dots & b_{r+1} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \dots & \dots & 0 \\ \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \dots & \dots & 0 \end{array} \right] \quad (4.11)$$

Der Rang r der Koeffizientenmatrix $A \in \mathbb{K}^{m,n}$ und der Rang der erweiterten Matrix $[A|\vec{b}]$ wird in der Beschreibung der Lösungsmenge eines LGSs eine wichtige Rolle spielen.

Falls ein LGS in normierter Zeilenstufenform vorliegt, kann die Struktur der Lösungsmenge wie folgt abgelesen werden:

Theorem 4.3.5 (Struktur der Lösungsmenge für den Fall: Das LGS ist in NZSF). Gegeben sei ein LGS $[A|\vec{b}]$ in NZSF (siehe Gleichung (4.11)). Die Koeffizientenmatrix des Gleichungssystems hat Rang r . Dann gilt:

1. Das LGS hat keine Lösung, wenn b_{r+1} ungleich Null ist (es gibt eine widersprüchliche Zeile), d.h. falls der Rang der Koeffizientenmatrix A kleiner ist als der Rang der erweiterten Koeffizientenmatrix $[A|\vec{b}]$: $\text{Rang } A = \text{Rang}([A|\vec{b}]) - 1$.
2. Wenn $b_{r+1} = 0$, so hat das LGS mindestens eine Lösung.
 - (a) Wenn $r = n$, so hat das LGS genau eine Lösung (alle Spalten sind Kopfspalten, d.h. sie enthalten einen Kopf).
 - (b) Wenn $r < n$, so hat das LGS unendlich viele Lösungen. In einer allgemeinen Lösung sind die $n - r$ Nichtkopfvariablen, frei wählbare Parameter. Die restlichen r Variablen (Kopfvariablen) sind durch die Zeilen, in denen sie Kopfvariable sind, festgelegt.

Bemerkung 4.3.6. Wenn Fall 1 in Theorem 4.3.5 auftritt, so ist der Rang der erweiterten Koeffizientenmatrix Eins größer als der Rang von A . Im Fall 2 haben die beiden Matrizen denselben Rang. Also gilt der Merksatz:

„Das LGS $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn $\text{Rang } A = \text{Rang}[A|\vec{b}]$ “.

4.4 DER GAUßALGORITHMUS

Das folgende Verfahren heißt Gaußalgorithmus und ist dazu da, ein LGS durch elementare Zeilenoperationen in ZSF bzw. NZSF zu bringen. Das ist deshalb nützlich weil wir wissen: Die Lösungsmengen bleiben unverändert. Es gilt:

Die Beispiele, die wir zu Beginn dieser Vorlesung betrachtet haben, legen das folgende allgemeine Verfahren nahe:

Algorithmus 4.4.1 (Gaußalgorithmus: ZSF).

Input: $A \in \mathbb{K}^{m,n}$, $A \neq 0$.

Output: ZSF von A .

- 1.: „Linksten“ Kopf nach oben tauschen:** Suche eine Zeile, deren Kopf den kleinsten Spaltenindex j hat. (So eine Zeile existiert, denn nach Voraussetzung ist A nicht die Nullmatrix.) Falls dies nicht etwa bereits die 1. Zeile ist, tausche diese Zeile mit der 1. Zeile von A .
- 2.: „Radieren“ unter dem Kopf:** Subtrahiere ein geeignetes Vielfaches der ersten Zeile von allen darunter stehenden Zeilen, sodass unter dem Kopf der 1. Zeile nur noch Nullen stehen.
- 3.: Rekursion:** Die erste Zeile und die erste Spalte bleiben wie sie sind. Ist die Matrix noch nicht in ZSF, so bearbeite die restliche, kleinere Matrix $A' \in \mathbb{K}^{m-1,n-1}$ analog.

Theorem 4.4.2. Um eine $m \times n$ Matrix auf ZSF zu bringen, benötigt man nicht mehr als $\frac{m(m-1)}{2}$ elementare Zeilenoperationen.

Algorithmus 4.4.3 (Gaußalgorithmus: NZSF).

Input: $A \in \mathbb{K}^{m,n}$, $A \neq 0$ in ZSF.

Output: NZSF von A .

- 1.: Normieren der Köpfe:** Dividiere jede Zeile, die keine Nullzeile ist, durch ihren Kopf.
- 2.: „Eliminieren“ über den Köpfen:** Führe für jeden Kopf folgende Schritte aus: Subtrahiere ein geeignetes Vielfaches seiner Zeile von allen darüber stehenden Zeilen, sodass schliesslich über ihm nur noch Nullen stehen.

Beispiel 4.4.4. Aufgabe: Löse das folgende LGS:

$$\begin{array}{rcl} 2x_1 + x_2 + 3x_3 & = & 3 \\ 2x_1 & + & 2x_3 = 2 \\ 2x_1 + 2x_2 + 4x_3 & = & 4 \end{array} \quad (*)$$

Wir stellen zunächst die erweiterte Koeffizientenmatrix auf.

$$[A|\vec{b}] = \left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 2 & 0 & 2 & 2 \\ 2 & 2 & 4 & 4 \end{array} \right] \quad (4.12)$$

Nun starten wir den Gaußalgorithmus:

1. Schritt: Linksten Kopf nach oben tausche: Entfällt, da $a_{1,1} \neq 0$.

2. Schritt: Elimination unter dem Kopf:

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 2 & 0 & 2 & 2 \\ 2 & 2 & 4 & 4 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \end{array} \right]$$

3. Schritt: Rekursion: Wir betrachten jetzt die kleinere Matrix

$$\left[\begin{array}{cc|c} -1 & -1 & -1 \\ 1 & 1 & 1 \end{array} \right]$$

in der Matrix

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \end{array} \right]$$

An dieser Matrix führen wir den Algorithmus erneut durch:

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Damit haben wir die Matrix in ZSF transformiert.

Nun gehen wir weiter und transformieren die Matrix in die NZSF:

1. Schritt: Normierte Köpfe:

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

2. Schritt: Eliminieren über den Köpfen (hier nur über dem Kopf der 2. Zeile):

$$\left[\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Wir lesen ab:

$$x_3 = \text{frei wählbar} \quad (4.13)$$

$$x_2 = 1 - x_3 \quad (4.14)$$

$$x_1 = 1 - x_3 \quad (4.15)$$

Damit ist die Lösungsmenge

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 = 1 - x_3, x_2 = 1 - x_3 \right\} = \left\{ \begin{bmatrix} 1 - s \\ 1 - s \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\} \quad (4.16)$$

Beispiel 4.4.5. Aufgabe: Löse das folgende LGS:

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ 2x_1 &+ 2x_3 &= 0 \\ 2x_1 + 2x_2 + 4x_3 &= 0 \end{aligned} \quad (4.17)$$

Man nennt dies *das zu (*) gehörige homogene LGS*.

Der Gaußalgorithmus läuft genau wie oben. Die NZSF sieht wie folgt aus:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad (4.18)$$

Wir lesen ab:

$$x_3 = \text{frei wählbar} \quad (4.19)$$

$$x_2 = -x_3 \quad (4.20)$$

$$x_1 = -x_3 \quad (4.21)$$

Damit ist die Lösungsmenge

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 = -x_3, x_2 = -x_3 \right\} = \left\{ \begin{bmatrix} -s \\ -s \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\} \quad (4.22)$$

Beobachtung 4.4.6. Die „allgemeine Lösung“ $\begin{bmatrix} 1-s \\ 1-s \\ s \end{bmatrix}$ des LGSs aus Beispiel 4.4.4 ist die Summe der „allgemeinen Lösung“ $\begin{bmatrix} -s \\ -s \\ s \end{bmatrix}$ des zugehörigen homogenen LGSs aus Beispiel 4.4.5 und einer speziellen Lösung $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ des inhomogenen LGSs.

Das folgende Theorem besagt, dass das immer so ist. Einen Bezug zum Bild linearer Abbildungen fügen wir hinzu.

Theorem 4.4.7 (Struktur der Lösungsmenge: inhomogenes LGS). Die Lösungsmenge eines inhomogenen LGSs ist die Menge der Vektoren die sich auf folgende Weise ergibt: Sei \vec{x}_P eine Lösung des inhomogenen LGSs (eine sogenannte spezielle Lösung - auch partikuläre Lösung genannt). Jetzt addieren wir zu \vec{x}_P je ein Element der Lösungsmenge des homogenen Systems:

$$\{\vec{x}_P + \vec{x}_H \mid \vec{x}_H \text{ Lösung des homogenen Systems}\}.$$

Dies ist dann die Lösungsmenge des inhomogenen LGSs.

Es gibt genau dann eine spezielle Lösung des inhomogenen LGSs, wenn die rechte Seite im Bild der durch A beschriebenen linearen Abbildung von \mathbb{R}^n nach \mathbb{R}^m liegt.

Daher ist die Struktur der Lösungsmenge der homogenen LGSs besonders wichtig. Man prüft leicht anhand der Definitionen den folgenden Satz nach:

Theorem 4.4.8 (Struktur der Lösungsmenge: homogenes LGS). Die Lösungsmenge eines homogenen LGSs mit Koeffizientenmatrix $A \in \mathbb{K}^{m,n}$ ist ein Teilraum des \mathbb{K}^n der Dimension $n - \text{Rang } A$. Dieser Raum ist gleich dem Kern der durch A beschriebenen linearen Abbildung von \mathbb{R}^n nach \mathbb{R}^m .

Zur Beschreibung der Lösungsmenge eines homogenen LGSs genügt es eine Basis des Kerns der zugehörigen Matrixabbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m; \vec{v} \mapsto A\vec{v}$ anzugeben.

In Beispiel 4.4.5 ist $\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$ eine Basis (der Kern hat Dimension 1).

4.5 WAS MAN AUS DER NZSF ALLES BERECHNEN KANN

Wir geben hier einige Verfahren an, wie man die NZSF für immer wiederkehrende Rechnungen (mechanisch) benutzt. Die Korrektheit der Verfahren folgt stets aus den Sätzen der vorigen Abschnitte. Es lohnt sich aber aus Gründen der Zeitersparnis, sich die Verfahren in untenstehender Form zu merken und zu üben.

Algorithmus 4.5.1 (Allgemeine Lösung).

Input: $A \in \mathbb{K}^{m,n}$, $\vec{b} \in \mathbb{K}^m$.

Output: Lösungsmenge $\{\vec{x} \in \mathbb{K}^n \mid A\vec{x} = \vec{b}\}$.

1. Bringe $[A|\vec{b}]$ in die NZSF $[\tilde{A}|\tilde{b}]$.
2. Schreibe die Terme mit den ‚Nichtkopfvariablen‘ (von \tilde{A}) auf die rechte Seite und füge sie zur Inhomogenität (\tilde{b}) dazu. Damit werden die Kopfvariablen durch die Nichtkopfvariablen ausgedrückt. Die Kopfvariablen sind jetzt also abhängig von den Nichtkopfvariablen, die zu Parametern der Lösungsmenge werden. In Formeln heißt dies für die Kopfvariable x_i :

$$x_i = \tilde{b}_k - \sum_{j=i+1}^n \tilde{a}_{kj} x_j.$$

Algorithmus 4.5.2 (Basis des Kerns).

Input: $A \in \mathbb{K}^{m,n}$, $r := \text{Rang}(A) > 0$.

Output: Basis von $\text{Kern}(A) = \{\vec{b}_1, \dots, \vec{b}_{n-r}\} \subset \mathbb{K}^n$.

1. Bringe A in NZSF $\tilde{A} = [\tilde{a}_{i,j}]$. Betrachte das homogene LGS $\tilde{A}\vec{x} = \vec{0}$.
2. Zerlege die Variablen x_1, \dots, x_n in r Kopfvariable und $n-r$ Nichtkopfvariable.
3. Setze die 1. Nichtkopfvariable gleich 1, alle anderen gleich 0 und berechne die eindeutige Lösung des homogenen LGSs. Dies ist der 1. Basisvektor des Kerns.
4. Setze die 2. Nichtkopfvariable gleich 1, alle anderen gleich 0 und berechne die eindeutige Lösung des homogenen LGSs. Dies ist der 2. Basisvektor des Kerns.
5. usw. Auf diese Weise ergeben sich $n-r$ Basisvektoren des Kerns.

Algorithmus 4.5.3 (Basis des Bildes).

Input: $A \in \mathbb{K}^{m,n}$.

Output: Basis $\{\vec{b}_1, \dots, \vec{b}_r\} \subset \mathbb{K}^m$ des Bildes von A .

1. Bringe die Matrix auf ZSF \tilde{A} und bestimme die Kopfvariablen.
2. Die Spaltenvektoren von A , die zu einer Kopfvariablen in \tilde{A} gehören, bilden eine Basis des Bildes.

Bemerkung 4.5.4. Seien $A \in \mathbb{K}^{m,n}$ und \tilde{A} eine ZSF von A .

- Die Spaltenvektoren von A bilden ein Erzeugendensystem des Bildes von A .
- $\text{Rang}(A) = \dim(\text{Bild}(A)) = \text{Anzahl der Kopfvariablen von } \tilde{A}$
- $\dim(\text{Kern}(A)) = \text{Anzahl der Nichtkopfvariablen von } \tilde{A}$
- Die Anzahl der Kopfvariablen von \tilde{A} + die Anzahl der Nichtkopfvariablen von \tilde{A} ist = der Anzahl aller Variablen (= Anzahl der Spalten von A), d.h.

es gilt:

$$\dim(\text{Kern}(A)) + \dim(\text{Bild}(A)) = n = \dim(\text{Urbild}(A)) = \# \text{Spalten in } A$$

Algorithmus 4.5.5 (Inverse).

Input: $A \in \mathbb{K}^{n,n}$.

Output: A^{-1} falls A invertierbar ist.

1. Bringe $[A|I_n]$ in NZSF $[\tilde{A}|B]$.
2. Ist $\tilde{A} \neq I_n$, so ist A nicht invertierbar. (Das merkt man schon während des Algorithmus, wenn im bereits fertig bearbeiteten Teil der Matrix eine Null auf der Diagonale erscheint.)
3. Andernfalls ist $A^{-1} = B$.

Algorithmus 4.5.6 (Darstellung eines Vektors als Linearkombination von Basisvektoren).

Sei \vec{v} ein Vektor und T ein Teilraum des \mathbb{K}^n . Ferner sei $\mathcal{B} := \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_r\}$ eine Basis von T . Sei $B \in \mathbb{K}^{n,r}$ die Matrix, die die $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_r$ als Spalten enthält, also $T = \text{Bild}(B)$.

Dann läßt sich \vec{v} genau dann aus Vektoren in \mathcal{B} linear kombinieren, liegt also in T , wenn das LGS $B\vec{x} = \vec{v}$ eine Lösung $\vec{\alpha} \in \mathbb{K}^r$ hat. Die Lösung ist in diesem Falle eindeutig (weil \mathcal{B} linear unabhängig ist) und

$$\vec{v} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \dots + \alpha_r \vec{b}_r \quad (4.23)$$

mit

$$\underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_r \end{bmatrix}}_{\text{Koordinatenvektor von } \vec{v} \text{ bzgl. } \mathcal{B}} =: \vec{v}_{\mathcal{B}} \quad (4.24)$$

Spezialfall $r = n$: Hier gilt zusätzlich $\vec{\alpha} = B^{-1}\vec{v}$.

Bemerkung 4.5.7. Falls der Gaußalgorithmus numerisch auf dem Computer implementiert wird, gibt es ein paar zusätzliche Schwierigkeiten, z.B. kann es passieren dass man eine der Todsünden der Numerischen Mathematik begeht, die Subtraktion von Zahlen ungefähr gleicher Größenordnungen. (Dies kann zu sehr grossen Fehlern führen). Aus diesem Grunde muss der Gaußalgorithmus verfeinert werden. Dies wird in der Vorlesung über Numerische Mathematik erklärt.

5. KAPITEL

VEKTORRÄUME ÜBER DEM KÖRPER \mathbb{K}

Diese Vorlesung dient dazu, die grundlegenden Begriffe der Vektorrechnung einzuführen. Sie ist in Analogie zur zweiten über die Vektorrechnung im \mathbb{R}^n (bzw. \mathbb{C}^n) zu sehen und hat die folgenden beiden Ziele: Erstens sollen die dort eingeführten Begriffe und Resultate so verallgemeinert werden, dass sie zur Modellierung vieler Phänomene der Ingenieur- und Naturwissenschaften verwendet werden können. Zweitens wird gezeigt, wie mit dem Konzept der Koordinaten der Fall endlichdimensionaler Vektorräume (der Begriff Dimension wird in Definition 5.2.19 eingeführt) auf den Fall der Vektorräume \mathbb{R}^n und \mathbb{C}^n zurückgeführt werden kann.

Ein typisches Beispiel einer Anwendung des Konzeptes Vektorraum ist die Beschreibung digitaler Bilder durch Matrizen. Ein anderes Beispiel ist die Beschreibung, Analyse und Bearbeitung von Signalen durch Polynome. Im ersten Beispiel werden Bilder als Vektoren im Vektorraum der Matrizen aufgefasst (vergleiche auch Beispiel 5.1.4). Im zweiten werden die Signale in typische Grundmuster z.B. Cosinus- und Sinusfunktionen, d.h. in Potenzen einer komplexen Variablen, zerlegt und auf diese Weise bearbeitet. Hier kommt also der Vektorraum der Polynome ins Spiel. Beide Beispiele aus der Informations- und Signaltheorie werden uns in dieser Vorlesung und den nächsten als Leitmotive dienen.

5.1 DEFINITION, BEISPIELE, EIGENSCHAFTEN VON VEKTORRÄUMEN

Hauptdarsteller dieser Vorlesung sind die Vektoren. Sie sind abstrakte Verallgemeinerungen der Koordinatenvektoren aus \mathbb{R}^n oder \mathbb{C}^n und allein durch Regeln definiert, wie mit ihnen umzugehen ist. Diese Regeln sind so gemacht, dass sie auf viele Situationen passen, z. B. auf den Umgang mit $(m \times n)$ -Matrizen, mit Polynomen und Funktionen.

Definition 5.1.1 (Vektorraum). Das Quadrupel $\{V, \mathbb{K}, Plus, Mal\}$ heißt *Vektorraum* falls folgendes gilt: V ist eine Menge und \mathbb{K} der Körper der reellen oder komplexen Zahlen. *Plus* ist eine Abbildung von $V \times V$ nach V und *Mal* eine Abbildung von $\mathbb{K} \times V$ nach V . Zur Vereinfachung der Notation führen wir folgende Bezeichnung ein: Seien \vec{u}, \vec{v} Elemente von V und α, β Elemente von \mathbb{K} . Statt *Plus*(\vec{u}, \vec{v}) schreiben wir $\vec{u} + \vec{v}$ und statt *Mal*(α, \vec{v}) die Verkürzung $\alpha\vec{v}$. Es gelte:

1. Es gibt in V ein spezielles Element $\vec{0}$, genannt der Nullvektor, das für jedes Element $\vec{v} \in V$ die Gleichung $\vec{v} + \vec{0} = \vec{v}$ erfüllt.
2. Zu jedem Element $\vec{v} \in V$ gibt es ein Element $-\vec{v}$ in V mit $\vec{v} + (-\vec{v}) = \vec{0}$.

3. Es gelten die folgenden Rechenregeln: (vgl. Lemma 2.1.8): Für alle Elemente $\vec{u}, \vec{v}, \vec{w}$ aus V ist

$$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}) \quad (\text{VR1.1})$$

$$\vec{u} + \vec{v} = \vec{v} + \vec{u}. \quad (\text{VR1.2})$$

Für die Multiplikation von Skalaren $\alpha, \beta \in \mathbb{K}$ mit Elementen $\vec{u}, \vec{v} \in V$ gilt:

$$\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v} \quad (\text{VR2.1})$$

$$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} \quad (\text{VR2.2})$$

$$(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v} \quad (\text{VR2.3})$$

$$1\vec{v} = \vec{v}. \quad (\text{VR2.4})$$

Bemerkung 5.1.2. Da die Schreibweise $\{V, \mathbb{K}, \text{Plus}, \text{Mal}\}$ für einen Vektorraum sehr umständlich ist, begnügt man sich da wo es keine Verwechslung geben kann mit der Kurzform V .

Bemerkung 5.1.3. Obwohl mit Vektoren aus V so gerechnet wird wie mit Vektoren aus \mathbb{R}^n oder \mathbb{C}^n , ist es oft ungünstig dieses Bild vor Augen zu haben. Besser ist es, sich an dem folgenden Beispiel zu orientieren, das wir bereits in der Einleitung zu dieser Vorlesung als Leitmotiv dieser und der nächsten Vorlesung angekündigt haben.

Beispiel 5.1.4. [Leitmotiv] Als Modell oder konkrete Realisierung des Konzeptes „Vektorraum“ betrachten wir die Menge der reellen $(m \times n)$ -Matrizen, $V = \mathbb{R}^{m,n}$. Sie modellieren digitale Schwarz-Weiß-Bilder mit m Zeilen und n Spalten. Null bedeutet eine mittlere Helligkeit. In der dritten Vorlesung haben wir gesehen, wie Matrizen addiert und mit einem Skalar multipliziert werden können. Lemma 5.3.1 besagt, dass die dort eingeführte Addition von Matrizen und Multiplikation mit Skalaren die Eigenschaft eines Vektorraums besitzt. $\mathbb{R}^{m,n}$ ist also ein Vektorraum über dem Körper der reellen Zahlen \mathbb{R} . Der Nullvektor $\vec{0}$ ist in diesem Fall eine Matrix vom Format $m \times n$, wobei alle Einträge 0 sind.

Beispiel 5.1.5. Als zweites Modell betrachten wir reelle Polynome vom Grad n . Seien p, q Polynome vom Grad n mit reellen Koeffizienten in der reellen Variablen x :

$$\begin{aligned} p(x) &= p_n x^n + p_{n-1} x^{n-1} + \cdots + p_1 x + p_0, \\ q(x) &= q_n x^n + q_{n-1} x^{n-1} + \cdots + q_1 x + q_0. \end{aligned} \quad (5.1)$$

$p + q$ ist wiederum ein Polynom und so definiert:

$$(p + q)(x) := (p_n + q_n)x^n + (p_{n-1} + q_{n-1})x^{n-1} + \cdots + (p_0 + q_0) \quad (5.2)$$

Sei $\alpha \in \mathbb{R}$; damit ist das Polynom αp so definiert

$$(\alpha p)(x) := \alpha p_n x^n + \cdots + \alpha p_1 x + \alpha p_0. \quad (5.3)$$

Die Menge der reellen Polynome vom Grad $\leq n$ wird mit $\mathbb{R}_{\leq n}[x]$ bezeichnet. Es ist einfach nachzurechnen, dass diese Menge mit der oben eingeführten Addition von Polynomen und Multiplikation mit Elementen aus \mathbb{R} eine zweite Realisierung des abstrakten Konzeptes Vektorraum darstellt. (Dazu müssen die Vektorraumeigenschaften VR1.1–VR1.2 und VR2.1–VR2.4 nachgeprüft werden.)

Der Nullvektor in $\mathbb{R}_{\leq n}[x]$ ist das Nullpolynom: $\vec{0} = 0x^n + 0x^{n-1} + \cdots + 0x + 0$.

5.2 BASIS UND DIMENSION

Im Unterschied zur Vorlesung über die Vektorräume \mathbb{R}^n , \mathbb{C}^n und \mathbb{K}^n sprechen wir jetzt nicht zuerst vom Vektorraum als Ganzes und dann von Teilräumen, sondern gleich von Teilräumen wobei dann der Gesamttraum als Spezialfall aufzufassen ist. Teilräume sind spezielle Teilmengen von Vektorräumen:

Definition 5.2.1 (Teilraum).

Eine Teilmenge T eines Vektorraums V über dem Körper \mathbb{K} heißt *Teilraum* oder *Untervektorraum* von V , wenn folgendes gilt:

- (i) T ist nicht leer.
- (ii) Für alle $\vec{u}, \vec{v} \in T$ ist $\vec{u} + \vec{v}$ wieder in T .
- (iii) Für alle $\vec{v} \in T$ und alle $\alpha \in \mathbb{K}$ ist $\alpha\vec{v}$ wieder in T .

Diese definitorischen Eigenschaften drücken wir in Kurzform folgendermaßen aus: Ein Teilraum ist eine nicht leere Teilmenge, die bezüglich Addition von Vektoren und Multiplikation mit Skalaren *abgeschlossen* ist.

Bemerkung 5.2.2. Aus der Definition folgt sofort:

1. Jeder Teilraum muss den Nullvektor enthalten, denn: da T nicht leer ist, gibt es ein $\vec{v} \in T$. Dann muss auch $0\vec{v} = \vec{0}$ in T sein.
2. Die Mengen $\{\vec{0}\}$ und V sind Teilräume eines Vektorraums V .

Beispiel 5.2.3. Die folgende Menge von Polynomen

$$\{p \in \mathbb{R}_{\leq n}[x] \mid p(x=0) = 0\}$$

ist ein Teilraum des Vektorraums $\mathbb{R}_{\leq n}[x]$ der Polynome vom Grad nicht grösser als n .

Beispiel 5.2.4. Die folgende Teilmenge von $\mathbb{R}^{2,2}$

$$D := \left\{ \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \mid a_{11}, a_{22} \in \mathbb{R} \right\}$$

ist ein Teilraum im Vektorraum der reellen (2×2) -Matrizen. Er heißt Teilraum der *Diagonalmatrizen*.

Beispiel 5.2.5. Die folgende Menge von Polynomen

$$\{p \in \mathbb{R}_{\leq n}[x] \mid \int_{-\infty}^{\infty} p(x) e^{-x^2} dx = 0\}$$

ist ein Teilraum des Vektorraums $\mathbb{R}_{\leq n}[x]$ der Polynome vom Grad nicht größer als n .

Nun kommen wir zum Begriff der Linearkombination. Er ist wiederum in Analogie zu dem bereits bekannten Begriff für Vektoren in \mathbb{K}^n gefasst:

Definition 5.2.6 (Linearkombination). Ein Vektor $\vec{v} \in V$ heißt *Linearkombination* der Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in V$, wenn Zahlen $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$ existieren so, dass

$$\vec{v} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k. \quad (5.4)$$

Man sagt: \vec{v} lässt sich aus den $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ *linear kombinieren*. Die $\alpha_1, \alpha_2, \dots, \alpha_k$ heißen *Koeffizienten* der Linearkombination.

Beispiel 5.2.7. Der Vektor

$$x^2 + 2x + 4 \in \mathbb{R}_{\leq 2}[x]$$

ist Linearkombination der beiden Vektoren

$$x^2 \quad \text{und} \quad x + 2,$$

weil gilt

$$x^2 + 2x + 4 = 1x^2 + 2(x + 2)$$

Beispiel 5.2.8. Der Vektor

$$x^2 + 2x + 4 \in \mathbb{R}_{\leq 2}[x]$$

ist *nicht* Linearkombination der beiden Vektoren

$$x^2 \quad \text{und} \quad x + 1 :$$

$$\alpha_1 x^2 + \alpha_2 (x + 1) = x^2 + 2x + 4 \Rightarrow \alpha_1 = 1 \text{ und } \alpha_2 (x + 1) = 2x + 4 \text{ (Widerspruch)}$$

Die Menge aller Linearkombinationen zweier linear unabhängiger Vektoren in \mathbb{R}^3 ist anschaulich gesprochen eine Ebene im \mathbb{R}^3 die den Nullpunkt enthält. Diese Vorstellung gibt ein Konstruktionsverfahren an, wie auch in diesem allgemeineren Fall aus Vektoren Teilräume konstruiert werden können (siehe unten das Theorem 5.2.11) und führt auf den Begriff der linearen Hülle:

Definition 5.2.9 (Lineare Hülle). Die *lineare Hülle* der Menge $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq V$ ist die Menge der Vektoren, die sich aus den $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ linear kombinieren lassen:

$$\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} := \{ \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k \mid \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K} \} \quad (5.5)$$

Beispiel 5.2.10. Die lineare Hülle der Menge, die aus den drei Vektoren (Polynome)

$$x^2, \quad x, \quad x(x+1)$$

besteht, ist der Teilraum der bei Null verschwindenden Polynomen in $\mathbb{R}_{\leq 2}[x]$, den wir im Beispiel 5.2.3 eingeführt hatten.

Die Erfahrung mit diesem Beispiel motiviert das folgende

Theorem 5.2.11. Seien k Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ aus V . Dann ist die Menge $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ ein Teilraum von V .

Das Theorem ist einfach zu beweisen falls die Begriffe klar sind. Es empfiehlt sich ein Selbsttest!

Nun gehen wir in die andere Richtung, d. h. von Teilräumen zu Vektoren:

Definition 5.2.12 (Erzeugendensystem). Sei T ein Teilraum des Vektorraums V . Die Menge $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq V$ heißt ein *Erzeugendensystem* von T , falls ihre lineare Hülle gleich T ist.

Wir erläutern dies an folgendem

Beispiel 5.2.13. Wir betrachten den Teilraum $\{p \mid p \in \mathbb{R}_{\leq 2}[x], p(x=0) = 0\}$ von Polynomen vom Grad nicht größer als 2 und die Menge

$$\{x^2, \quad x, \quad x(x+1)\}.$$

Es ist leicht nachzuprüfen, dass die Menge ein Erzeugendensystem dieses Teilraums ist. Ja, es ist sogar so, dass je zwei ein Erzeugendensystem bilden. Ein Vektor ist also überflüssig. Man sagt (siehe unten): „Die drei Vektoren sind linear abhängig“.

Dies führt auf den Begriff der linearen Abhängigkeit der nicht viel mehr als eine Wiederholung dessen ist, was wir in der zweiten Vorlesung kennen gelernt haben:

Definition 5.2.14 (Lineare Unabhängigkeit). Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ des Vektorraums V über \mathbb{K} heißt genau dann *linear unabhängig*, wenn die Vektorgleichung

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0}$$

für die Unbekannten $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ nur die Lösung: $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ hat, d. h. die Lösung $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ ist eindeutig.

Definition 5.2.15 (Lineare Abhängigkeit). Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ des Vektorraums V über \mathbb{K} heißt genau dann *linear abhängig*, wenn die Vektorgleichung

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0}$$

für die Unbekannten $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ nicht nur die Lösung: $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ hat; d.h. neben der Lösung $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ gibt es noch andere Lösungen mit mindestens einem von Null verschiedenen α_i .¹ Die Lösung ist *nicht eindeutig*.

Bemerkung 5.2.16. Um umständliche Formulierungen zu vermeiden, sagen wir auch, dass eine Menge M linear (un)abhängig sei, falls die Vektoren in M linear (un)abhängig sind.

Beispiel 5.2.17. Die drei Vektoren im Vektorraum der Polynome

$$x^2, \quad x, \quad x(x+1)$$

sind linear abhängig.

Damit gelangen wir nun zu den zentralen Begriffen, Dimension und Basis eines Vektorraums über einem Körper \mathbb{K} . Als erstes unterscheiden wir zwischen endlichdimensionalen und unendlichdimensionalen Vektorräumen.

Definition 5.2.18 (Vektorräume endlicher Dimension). Ein Vektorraum V über \mathbb{K} heißt *endlichdimensional*, falls es ein (endliches) Erzeugendensystem $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq V$ gibt, das ihn aufspannt. Andernfalls heißt er *unendlichdimensional*.

Definition 5.2.19 (Basis eines Vektorraums). Ein linear unabhängiges Erzeugendensystem des Vektorraums V heißt *Basis des Vektorraums V* .

Bemerkung 5.2.20. Es gibt kein linear unabhängiges Erzeugendensystem des Vektorraums $\{\vec{0}\}$. Wir definieren die leere Menge $\{\}$ als (einzige) Basis dieses Raums.

Beispiel 5.2.21. In Beispiel 5.2.13 haben wir als Erzeugendensystem eine Menge bestehend aus drei Polynome eingeführt. Jede Menge, die aus zwei dieser Polynomen, sind linear unabhängig und stellt bereits für sich allein ein Erzeugendensystem des Teilraums der Polynome vom Grad nicht größer als zwei dar. Also gibt es hier drei offensichtliche Basen. Alle bestehen aus zwei Elementen. All dies nachzuprüfen ist recht einfach und wird empfohlen.

Das folgende Resultat ist von grundlegender Bedeutung und folgt aus dem Steinitzschen Austauschsatz. Ein Beweis ist nicht einfach und kann hier nicht ausgeführt werden.

Theorem 5.2.22. *Alle Basen eines Vektorraums haben dieselbe Anzahl von Elementen.*

Definition 5.2.23 (Dimension). Sei V ein endlichdimensionaler Vektorraum (vgl. Definition 5.2.15). Dann ist die *Dimension* von V gleich der Anzahl der Elemente einer Basis.

Beispiel 5.2.24. [Endlichdimensionale Vektorräume] Die Dimension des Teilraums D (vgl. Beispiel 5.2.4 und Beispiel 5.2.3) der diagonalen (2×2) -Matrizen ist Zwei.

Die Dimension des Vektorraums $\mathbb{K}^{m,n}$ ist mn . Denn die Matrizen mit einer Eins an irgendeiner Stelle und überall sonst Nullen bilden eine Basis.

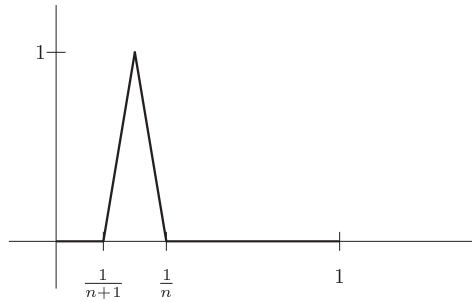
¹In der Tat gibt es dann beliebig viele Lösungen.

Beispiel 5.2.25. [Unendlichdimensionale Vektorräume] 1) Die Polynome in einer Variablen x mit reellen Koeffizienten haben die Struktur eines Vektorraums über \mathbb{R} . Die Dimension ist unendlich, denn zu jeder natürlichen Zahl m sind die m Monome $x^0, x^1, x^2, \dots, x^m$ linear unabhängig.

2) Sei $V = \mathcal{C}([0, 1], \mathbb{R})$ der Vektorraum der stetigen reellwertigen Funktionen auf $[0, 1]$, so ist $\dim V = \infty$.

Sei für $n = 1, 2, \dots$

$$f_n(x) = \begin{cases} 0, & x < \frac{1}{n+1}, \\ 0, & \frac{1}{n} < x, \\ 2n(n+1)x - 2n, & \frac{1}{n+1} \leq x \leq \frac{1}{2} \left(\frac{1}{n} + \frac{1}{n+1} \right), \\ -2n(n+1)x + 2n+2, & \frac{1}{2} \left(\frac{1}{n} + \frac{1}{n+1} \right) < x \leq \frac{1}{n}. \end{cases}$$



Jede Linearkombination $\sum_{j=1}^k \alpha_j f_j$ hat an der Stelle $\frac{1}{2} \left(\frac{1}{j} + \frac{1}{j+1} \right)$ den Wert α_j , also kann 0 nur vorkommen, wenn alle $\alpha_j = 0$ sind. Somit sind f_1, \dots, f_k für alle k linear unabhängig. Damit folgt, dass es keine endliche Basis geben kann.

Zum Schluss dieses Abschnittes geben wir – wieder ohne Beweis – ein Resultat an, das einen nützlichen und komplementären Blick auf das Konzept Basis wirft:

Theorem 5.2.26. *Sei V ein Vektorraum über \mathbb{K} . Dann ist die Dimension von V gleich der maximalen Anzahl linear unabhängiger Vektoren. Die Dimension ist eine natürliche Zahl oder Unendlich. Letzteres heißt definitionsgemäß: Zu jeder natürlichen Zahl m gibt es m linear unabhängige Vektoren.*

5.3 KOORDINATEN UND KOORDINATENABBILDUNG

In den vorangehenden Abschnitten haben wir uns damit befasst, das Konzept des Vektorraums \mathbb{R}^n zu verallgemeinern. Jetzt werden wir noch kurz erläutern, wie jeder endlich-dimensionale Vektorraum über \mathbb{K} mithilfe eines \mathbb{K}^n beschrieben werden kann. Das Schlüsselwort dazu heißt *Koordinaten* und *Koordinatenabbildung*:

Theorem 5.3.1. *Sei V ein endlichdimensionaler Vektorraum über \mathbb{K} und $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_{\dim(V)}\}$ eine Basis von V .*

Dann kann jeder Vektor $\vec{v} \in V$ in eindeutiger Weise als Linearkombination $\vec{v} = \alpha_1 \vec{b}_1 + \dots + \alpha_{\dim(V)} \vec{b}_{\dim(V)}$ geschrieben werden. (Eindeutig heißt: Die Zahlen $\alpha_1, \dots, \alpha_{\dim(V)} \in \mathbb{K}$ sind eindeutig.)

Damit kann eine Abbildung definiert werden, die sogenannte *Koordinatenabbildung* $K_{\mathcal{B}}$. Sie geht aus vom Vektorraum V und bildet in den Vektorraum $\mathbb{K}^{\dim(V)}$ ab. *Sie ist von der Wahl der Basis abhängig!*

Für eine Basis $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_{\dim(V)}\}$ von V heißt

$$K_{\mathcal{B}} : \left\{ \begin{array}{ccc} V & \rightarrow & \mathbb{K}^{\dim(V)} \\ \vec{v} = \sum_{i=1}^{\dim(V)} \alpha_i \vec{b}_i & \mapsto & \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{\dim(V)} \end{bmatrix} \end{array} \right. \quad (5.6)$$

die *Koordinatenabbildung* von V bzgl. \mathcal{B} .

Diese Abbildung ordnet jedem Element eines Vektorraums V genau ein Element von $\mathbb{K}^{\dim(V)}$ zu und gestattet demnach, den Vektorraum V mit $\mathbb{K}^{\dim(V)}$ zu identifizieren.

6. KAPITEL

LINEARE ABBILDUNGEN

Die Vorlesung von heute hat die linearen Abbildungen zum Thema. Wir hatten bereits in der dritten Vorlesung den Begriff im Zusammenhang mit dem speziellen Vektorraum \mathbb{K}^n und den Matrizen kennen gelernt. Jede $m \times n$ Matrix kann als lineare Abbildung von \mathbb{K}^n nach \mathbb{K}^m aufgefasst werden. Jetzt geht es darum, diesen Begriff auf die Situation allgemeiner Vektorräume auszudehnen, genau so wie wir in der letzten Vorlesung den Begriff der Vektoren im \mathbb{K}^n auf allgemeine Vektoren, z.B. auf Matrizen oder Polynome ausgedehnt haben.

Lineare Abbildungen treten in den verschiedensten Anwendungen als wesentliches mathematisches Instrument auf. Wir werden hier in erster Linie unser Leitmotiv, den Vektorraum der digitalen Bilder und Signale, sowie Varianten davon als Beispiel diskutieren. Die linearen Abbildungen sind dann lineare Filter und haben eine große Bedeutung in der Signaltheorie.

Besonders wichtig für die Analysis sind die linearen Abbildungen als Approximation allgemeiner differenzierbarer Abbildungen. Die beste lineare Approximation einer Abbildung in der Nähe eines Punktes heißt ihre Ableitung. Diese wichtige Aussage wird allerdings erst am Ende des zweiten Semesters verständlich werden.

6.1 DEFINITION UND BEISPIELE

Die linearen Abbildungen sind abstrakte Verallgemeinerungen der Matrizen. Eine Abbildung heißt dann linear („linea“ heißt lateinisch „Gerade“), wenn sie gewissen Regeln genügt. Diese können in folgender Weise aus der geometrischen Anschauung heraus motiviert werden: Sei f eine Abbildung von \mathbb{R} nach \mathbb{R} ; dann kann die Frage, ist die Abbildung linear oder nicht, am Graphen $\Gamma(f)$ abgelesen werden: Ist $\Gamma(f)$ in $\mathbb{R} \times \mathbb{R}$ eine Gerade durch den Ursprung, dann ist f linear und sonst nicht. Analytisch bedeutet dies: Das Bild der Summe von Vektoren ist gleich der Summe der Bilder und das Bild der Multiplikation eines Vektors mit einem Skalar ist gleich der Multiplikation des Bildes mit dem Skalar. In Formeln ausgedrückt heißt dies, für alle $x, y, \alpha \in \mathbb{R}$ gilt $f(x + y) = f(x) + f(y)$ und $f(\alpha x) = \alpha f(x)$. Man fasst dies folgendermaßen zusammen: Lineare Abbildungen erhalten die Strukturen der Vektorräume.

Definition 6.1.1. Seien V und W Vektorräume über demselben Körper \mathbb{K} . Dann heißt eine Abbildung L von V in W *linear*, wenn folgende Regeln gelten: Für alle Vektoren $\vec{u} \in V, \vec{v} \in V$ und alle $\alpha \in \mathbb{K}$ gilt

$$L(\vec{u} + \vec{v}) = L(\vec{u}) + L(\vec{v}), \quad (6.1)$$

$$L(\alpha \vec{v}) = \alpha L(\vec{v}). \quad (6.2)$$

Die Menge der linearen Abbildungen von V nach W bezeichnen wir mit dem Symbol $\text{Hom}(V, W)$, denn die linearen Abbildungen werden auch Homomorphismen genannt.

Falls die beiden Vektorräume zusammenfallen, d.h. falls $V = W$ ist, gibt es in $\text{Hom}(V, V)$ die Identität, die jedem $\vec{v} \in V$ das Bild \vec{v} zuordnet. Wir bezeichnen diese spezielle Abbildung mit I_V .

Bemerkung 6.1.2. $\vec{u} + \vec{v}$ ist die Addition in V und $L(\vec{u}) + L(\vec{v})$ die Addition in W . Beide werden mit demselben Symbol $+$ bezeichnet. Dies ist einerseits gefährlich, weil ja unterschiedliche Dinge nicht mit demselben Symbol bezeichnet werden sollten. Andererseits ist dies eine sehr effiziente Schreibweise, da sie auf den wesentlichen Punkt aufmerksam macht: In V und in W gelten dieselben Rechenregeln, die allerdings auf unterschiedliche Objekte angewandt werden. Man hat auf diese Weise einen wesentlichen Abstraktionsschritt getan und – vorausgesetzt man weiß worum es geht – beträchtlich an Übersicht gewonnen.

Jetzt gehen wir durch eine Reihe von Beispielen, an denen sich einige interessante Eigenschaften linearer Abbildungen erklären lassen.

Beispiel 6.1.3. Die Abbildung

$$H : \left\{ \begin{array}{ccc} \mathbb{R}^{2,2} & \rightarrow & \mathbb{R}^{2,2} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto & \begin{bmatrix} \frac{a_{11}+a_{12}}{2} & a_{11} - a_{12} \\ \frac{a_{21}+a_{22}}{2} & a_{21} - a_{22} \end{bmatrix} \end{array} \right. \quad (6.3)$$

ist linear; ebenso ist dies die Abbildung

$$V : \left\{ \begin{array}{ccc} \mathbb{R}^{2,2} & \rightarrow & \mathbb{R}^{2,2} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto & \begin{bmatrix} \frac{a_{11}+a_{21}}{2} & \frac{a_{12}+a_{22}}{2} \\ a_{11} - a_{21} & a_{12} - a_{22} \end{bmatrix} \end{array} \right. \quad (6.4)$$

H und V sind die diskreten horizontalen und vertikalen Haarfilter. Beide gehören zu einer wichtigen Klasse von Filtern, den diskreten Waveletfiltern.

Beispiel 6.1.4. Die Abbildung

$$K : \left\{ \begin{array}{ccc} \mathbb{R}^{2,2} & \rightarrow & \mathbb{R}^{2,2} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto & \begin{bmatrix} a_{11} + a_{12} + a_{21} + a_{22} & 0 \\ 0 & 0 \end{bmatrix} \end{array} \right. \quad (6.5)$$

ist linear; ebenso ist dies die Abbildung

$$T : \left\{ \begin{array}{ccc} \mathbb{R}^{2,2} & \rightarrow & \mathbb{R} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto & a_{11} + a_{12} + a_{21} + a_{22} \end{array} \right. \quad (6.6)$$

Beispiel 6.1.5. Wir betrachten die beiden Vektorräume \mathbb{R}^n und $\mathbb{R}_{\leq n-1}[z]$. Die Abbildung

$$\left\{ \begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R}_{\leq n-1}[z] \\ \vec{v} & \mapsto & p(z) = v_1 + v_2 z + v_3 z^2 + \dots + v_n z^{(n-1)} \end{array} \right. \quad (6.7)$$

ist linear. Hier sei noch einmal darauf hingewiesen, dass die Summe im Urbildraum \mathbb{R}^n und im Bildraum $\mathbb{R}_{\leq n-1}[z]$ zwei Operationen sind, die denselben Regeln genügen, aber auf unterschiedliche Objekte wirken. Im Urbildraum \mathbb{R}^n geht es um die Summe von Koordinatenvektoren, im Bildraum $\mathbb{R}_{\leq n-1}[z]$ um die Summe von Polynomen. Diese Abbildung heißt die z -Transformation und ist in der Theorie der Signalverarbeitung von Bedeutung. Sie erlaubt ein Signal \vec{v} durch ein Polynom $p(z)$ in eindeutiger Weise darzustellen. Man spricht von einer *Kodierung* des Signales.

Beispiel 6.1.6. [Ableitung] Das folgende Beispiel kennt ihr aus der Analysis. Es handelt sich um die Ableitung von Polynomen, die hier für den einfachen Fall von Polynomen zweiten Grades ausgeführt ist:

$$D : \left\{ \begin{array}{ccc} \mathbb{R}_{\leq 2}[x] & \rightarrow & \mathbb{R}_{\leq 2}[x] \\ a_0 + a_1 x + a_2 x^2 & \mapsto & a_1 + 2a_2 x \end{array} \right. \quad (6.8)$$

Die Ableitung gehört zu den wichtigsten linearen Abbildungen.

6.2 DER VEKTORRAUM DER LINEAREN ABBILDUNGEN

Mit linearen Abbildungen vom Vektorraum V in den Vektorraum W kann wie mit Vektoren gerechnet werden.

Definition 6.2.1 (Vektorraumoperationen). Seien V und W Vektorräume über \mathbb{K} und $L, M \in \text{Hom}(V, W)$, $\alpha \in \mathbb{K}$. Wir definieren die folgenden linearen Abbildungen als Elemente in $\text{Hom}(V, W)$:

$$L + M : \begin{cases} V & \rightarrow & W \\ \vec{v} & \mapsto & L(\vec{v}) + M(\vec{v}) \end{cases} \quad (\text{Summe})$$

$$\alpha L : \begin{cases} V & \rightarrow & W \\ \vec{v} & \mapsto & \alpha L(\vec{v}) \end{cases} \quad (\text{Multiplikation mit Skalaren})$$

Es ist leicht nachzuprüfen, dass mit den oben definierten Vektorraumoperationen die Menge der Abbildungen $\text{Hom}(V, W)$ zu einem Vektorraum über \mathbb{K} wird.

Beispiel 6.2.2. In Beispiel 6.1.3 hatten wir die beiden Abbildungen H und V eingeführt. Nun betrachten wir deren Summe:

$$H + V : \begin{cases} \mathbb{R}^{2,2} & \rightarrow & \mathbb{R}^{2,2} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto & \begin{bmatrix} \frac{2a_{11}+a_{12}+a_{21}}{2} & \frac{2a_{11}-a_{12}+a_{22}}{2} \\ \frac{2a_{11}-a_{21}+a_{22}}{2} & a_{21} + a_{12} - 2a_{22} \end{bmatrix} \end{cases} \quad (6.9)$$

Beispiel 6.2.3. In Beispiel 6.1.4 hatten wir die beiden Abbildungen K und T definiert und stellen jetzt fest, dass deren Summe nicht definiert ist, da die beiden Abbildungen nicht dieselben Urbild- und Bildräume haben. (Dies entspricht Matrizen, die nicht die für die Summenbildung passenden Formate haben.)

6.3 DIE KOMPOSITION LINEARER ABBILDUNGEN

Die *Komposition linearer Abbildungen* wird auch *Hintereinanderausführung* oder *Produkt* linearer Abbildungen genannt.

Definition 6.3.1. Seien U, V, W drei Vektorräume über \mathbb{K} und $L \in \text{Hom}(U, V)$, $M \in \text{Hom}(V, W)$. Dann ist die Komposition (auch genannt die Hintereinanderausführung oder das Produkt) ML eine lineare Abbildung von U nach W und folgendermaßen definiert: $(ML)\vec{v} := M(L\vec{v})$ für alle $\vec{v} \in U$. Statt ML wird oftmals auch $M \circ L$ geschrieben.

Definition 6.3.2 (Inverse Abbildung). Sei V und W Vektorräume über \mathbb{K} und $L \in \text{Hom}(V, W)$. Dann heißt L *invertierbar* und L^{-1} die *Inverse*- oder auch *Umkehrabbildung* von L , falls es eine Abbildung L^{-1} von W nach V gibt, so dass $L^{-1} \circ L = I_V$ und $L \circ L^{-1} = I_W$.

Bemerkung 6.3.3. Die inverse Abbildung einer linearen Abbildung ist linear. Falls V endlichdimensional ist und $L \in \text{Hom}(V, V)$ (d.h. $V = W$) gilt nicht nur $L^{-1} \circ L = I_V$, sondern auch $L \circ L^{-1} = I_V$. (Wir geben für diese Behauptung keinen Beweis. Sie folgt aus dem Dimensionssatz 6.4.5).

Definition 6.3.4. Seien $L \in \text{Hom}(V, V)$ und $n \in \mathbb{N}$. Dann ist L^n als n -fache Komposition von L mit sich selbst, wobei $L^0 := I_V$. Falls L invertierbar ist, ist L^{-n} die n -fache Komposition von L^{-1} mit sich selbst.

Beispiel 6.3.5. Die zu H inverse Abbildung ist:

$$H^{-1} : \begin{cases} \mathbb{R}^{2,2} & \rightarrow \mathbb{R}^{2,2} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \mapsto \begin{bmatrix} a_{11} + \frac{a_{12}}{2} & a_{11} - \frac{a_{12}}{2} \\ a_{21} + \frac{a_{22}}{2} & a_{21} - \frac{a_{22}}{2} \end{bmatrix} \end{cases} \quad (6.10)$$

Die oben behaupteten Eigenschaften können daran nachgerechnet werden.

Beispiel 6.3.6. [Projektion] Die Abbildung K , die wir in 6.1.4 eingeführt hatten, hat die Eigenschaft $K^2 = K$. Abbildungen mit dieser Eigenschaft heissen *Projektionen*.

Beispiel 6.3.7. [Nilpotent] Die Ableitung D auf den Polynomen $\mathbb{R}_{\leq 2}[x]$ (Bsp. 6.1.6) hat die Eigenschaft $D^3 = 0$, d.h. alle Bilder von D^3 sind gleich dem Nullpolynom. Falls es zu einer linearen Abbildung A ein $n \in \mathbb{N}$ gibt, so dass $A^n = 0$ ist, dann heißt A *nilpotent*. D ist also nilpotent.

Die Eigenschaften der Addition und Komposition linearer Abbildungen sind in dem folgenden Theorem zusammengefasst:

Theorem 6.3.8. Gegeben sind die Vektorräume U, V, W und X über dem Körper \mathbb{K} .

Dann gilt:

$$\begin{array}{lll} (LM)N = L(MN) & \forall L \in \text{Hom}(W, X), & \text{(Assoziativität)} \\ & \forall M \in \text{Hom}(V, W), \\ & \forall N \in \text{Hom}(U, V); \\ L(M + N) = LM + LN & \forall L \in \text{Hom}(V, W), & \text{(Distributivität)} \\ & \forall M \in \text{Hom}(U, V), \\ & \forall N \in \text{Hom}(U, V); \\ L(\alpha M) = \alpha(LM) & \forall L \in \text{Hom}(V, W), \\ & \forall M \in \text{Hom}(U, V), \\ & \forall \alpha \in \mathbb{K}; \\ LI_U = I_V L = L & \forall L \in \text{Hom}(U, V). & \text{(Neutrales Element)} \end{array}$$

6.4 KERN UND BILD EINER LINEAREN ABBILDUNG, LÖSUNGSRAUM LINEARER GLEICHUNGEN

Die Menge der Urbilder einer linearen Abbildung, die auf die Null abgebildet werden und die Menge der Bilder spielen eine wichtige Rolle in der linearen Algebra und der Anwendung auf lineare Gleichungen. Sie haben deshalb einen eigenen Namen:

Definition 6.4.1 (Kern). Seien V und W Vektorräume über \mathbb{K} und L eine lineare Abbildung von V nach W . Dann heißt die Menge $\{\vec{v} \in V \mid L(\vec{v}) = \vec{0}\}$ der *Kern* oder der *Nullraum* der linearen Abbildung L . Sie wird in der Regel mit $\text{Kern}(L)$ bezeichnet.

Definition 6.4.2 (Bild). Seien V und W Vektorräume über \mathbb{K} und L eine lineare Abbildung von V nach W . Dann heißt die Menge $\{L(\vec{v}) \mid \vec{v} \in V\}$ das *Bild* der linearen Abbildung L . Sie wird in der Regel mit $\text{Bild}(L)$ bezeichnet.

Beispiel 6.4.3. Das Bild der Abbildung K aus dem Beispiel 6.1.4 ist die Teilmenge von $\mathbb{R}^{2,2}$

$$\left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}.$$

Der Kern von K ist die Teilmenge von $\mathbb{R}^{2,2}$

$$\left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{11} + a_{12} + a_{21} + a_{22} = 0 \right\}.$$

Beides sind offenbar Teilräume der reellen 2×2 Matrizen.

Nun berechnen wir die Dimensionen von Kern und Bild der Abbildung K . Das Bild von K ist

$$\text{span} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

und deshalb von der Dimension eins. Um die Dimension des Kerns von K zu finden, argumentieren wir so: Die Dimension ist kleiner als 4, denn andernfalls könnte jede Matrix als Element des Kerns aufgefasst werden, $\mathbb{R}^{2,2}$ hat die Dimension 4. Dies ist aber nicht der Fall, z.B. liegt die angegebene Matrix, welche das Bild erzeugt, nicht im Kern von K . Andererseits ist es leicht drei linear unabhängige Vektoren im Vektorraum der reellen 2×2 -Matrizen anzugeben, z.B. ist folgendes eine mögliche Wahl:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$$

Damit ist gezeigt, dass die Dimension des Kerns von K gleich drei ist.

Das obige Beispiel macht das folgende Theorem plausibel:

Theorem 6.4.4. Seien V und W Vektorräume über \mathbb{K} und L eine lineare Abbildung von V nach W . Dann ist der Kern von L ein Teilraum der Urbildmenge V und das Bild von L ein Teilraum der Bildmenge W .

Beweis. Wir zeigen die Aussage für den Kern von L . Sei \vec{u} und \vec{v} im Kern von L dann ist $L(\vec{u} + \vec{v}) = L(\vec{u}) + L(\vec{v})$ (siehe Theorem 6.3.8). Da \vec{u} und \vec{v} im Kern von L sind, steht auf der rechten Seite der vorangehenden Gleichung der Nullvektor. Damit ist ein Teil der Behauptung gezeigt, nämlich die Gleichung (6.1) in der Definition 5.2.1. Der Rest des Beweises verläuft analog.

Theorem 6.4.5 (Dimensionssatz). Seien V und W Vektorräume über \mathbb{K} und L eine lineare Abbildung von V nach W (d.h. $L: V \rightarrow W$). Falls V endlichdimensional ist gilt die Gleichung

$$\dim(V) = \dim(\text{Kern}(L)) + \dim(\text{Bild}(L))$$

Der Beweis des Satzes ist nicht sehr schwierig, aber von der Zeit her zu aufwendig für diesen Kurs. In Beispiel 6.4.3 haben wir die Dimensionsformel bereits nachgerechnet.

Nun betrachten wir *abstrakte lineare Gleichungen*.

Definition 6.4.6 (Lineare Gleichung). Seien V und W Vektorräume über \mathbb{K} , L eine lineare Abbildung von V nach W und $\vec{b} \in W$. Die Menge $\{\vec{x} \mid \vec{x} \in V, L(\vec{x}) = \vec{b}\}$ heißt die Lösungsmenge der linearen Gleichung $L(\vec{x}) = \vec{b}$. Falls $\vec{b} \neq \vec{0}$ ist, heißt die Gleichung *inhomogen*, andernfalls *homogen*.

Die Lösungsmenge der homogenen Gleichung $L(\vec{x}) = \vec{0}$ ist gleich dem Kern von L . Damit können wir uns eine geometrische Vorstellung über die Lösungsmenge einer linearen Gleichung machen:

Theorem 6.4.7 (Struktur des Lösungsraumes). Sei $L(\vec{x}) = \vec{b}$ eine lineare Gleichung, wie in der obigen Definition 6.4.6 beschrieben. Sei \vec{x}_P eine Lösung der linearen Gleichungen. Dann ist die Lösungsmenge

$$\vec{x}_P + \text{Kern}(L) := \{\vec{x}_P + \vec{x}_H \mid \vec{x}_H \in \text{Kern}(L)\}.$$

Beispiel 6.4.8. Sei D die lineare Abbildung aus Beispiel 6.1.6, d.h. die Ableitung der Polynome zweiten Grades und sei $b(x) := 2x + 1$ die Inhomogenität in der linearen Gleichung $Dp = b$. Gesucht sind also alle reellen Polynome zweiten Grades, die dieser Gleichung genügen. Der Kern von D besteht aus allen Polynomen nullten Grades, also allen Konstanten. Eine partikuläre (oder spezielle) Lösung q_P ist beispielsweise das Polynom $q_P(x) := x^2 + x$. Somit ist die Lösungsmenge $\{x^2 + x + a \mid a \in \mathbb{R}\}$.

Bemerkung 6.4.9 (Berechnung von Kern und Bild). Kern und Bild einer linearen Abbildung können besonders einfach für Matrixabbildungen (vgl. Bemerkung 3.5.8) berechnet werden. Sei $A \in \mathbb{K}^{m,n}$. Dann ist der Kern von A die Lösungsmenge des *homogenen* Gleichungssystems $A\vec{x} = \vec{0}$ und das Bild von A gleich der linearen Hülle der Menge der Spaltenvektoren von A (vgl. Bemerkung 3.5.9).

Für *lineare* Abbildungen hängen die Begriffe *Kern* und *Bild* unmittelbar mit den aus der Analysis bekannten Konzepten *injektiv* und *surjektiv* zusammen. Zur Erinnerung:

Definition 6.4.10 (injektiv, surjektiv, bijektiv). Seien V und W Vektorräume über \mathbb{K} und L eine lineare Abbildung von V nach W . Dann heißt die Abbildung L *injektiv*, falls jeder Bildvektor nur gerade mal einen Urbildvektor besitzt. L heißt *surjektiv*, falls jeder Vektor von W Bildvektor ist und *bijektiv*, falls L surjektiv und injektiv ist. (Abstrakt gesehen, können im Falle dass L bijektiv ist, die beiden Räume mittels L identifiziert werden.)

Der Zusammenhang zwischen den genannten Begriffen und Konzepten ist in dem folgenden Satz formuliert:

Theorem 6.4.11 (injektiv-Kern, surjektiv-Bild). Sei L eine lineare Abbildung von V nach W . Dann gilt:

$$L \text{ ist injektiv} \iff \text{Kern}(L) = \{\vec{0}\} \iff \dim(\text{Kern}(L)) = 0 \quad (6.11)$$

Ist V endlichdimensional, dann gilt:

$$L \text{ ist surjektiv} \iff \text{Bild}(L) = W \iff \dim(\text{Bild}(L)) = \dim(W) \quad (6.12)$$

7. KAPITEL

KOORDINATEN UND DARSTELLENDEN MATRIZEN

In dieser Vorlesung soll gezeigt werden, wie ein Vektor als Koordinatenvektor und eine lineare Abbildung als Matrix aufgefasst, oder wie man sagt dargestellt, werden kann. Den Übergang von Vektor und linearer Abbildung zu Koordinatenvektor und darstellender Matrix vermittelt die *Koordinatenabbildung*.

Diese Brücke kann immer dann eingeführt werden, wenn in den betrachteten Vektorräumen Basen gegeben sind. Die Koordinatenabbildung ist explizit von der Wahl der Basis abhängig. Diese ist willkürlich. Davon wird oft mit großem Nutzen Gebrauch gemacht. Denn eine geschickte Wahl der Basen kann Rechnungen erheblich verkürzen.

In dieser Vorlesung werden nur endlichdimensionale Vektorräume betrachtet. Vieles kann ohne weiteres auf beliebige Vektorräume mit Basen übertragen werden, falls die Basiselemente mit natürlichen Zahlen durchnummeriert werden können. Manche der Anwendungen sind von dieser Art, z.B. die Darstellung stetiger Funktionen auf einem Intervall in der Basis von Cosinus und Sinus. Dies führt auf die sogenannten Fourierreihen, die in sehr vielen Anwendungen, z.B. in der Akustik, von besonderer Bedeutung sind.

7.1 DIE DARSTELLUNG VON VEKTOREN DURCH KOORDINATENVEKTOREN

Sei V ein endlichdimensionaler Vektorraum über dem Körper \mathbb{K} mit der Dimension n und $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis. In Theorem 5.3.1 und dem darauffolgenden Abschnitt haben wir die Koordinatenabbildung $K_{\mathcal{B}}$ eingeführt. Die Koordinatenabbildung $K_{\mathcal{B}}$ bildet V bijektiv auf \mathbb{K}^n ab und ist linear. Das folgende Diagramm veranschaulicht diese Situation:

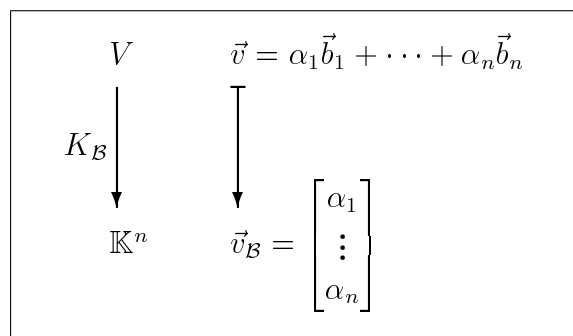


ABBILDUNG 7.1: Die Koordinatenabbildung $K_{\mathcal{B}}$ von V nach \mathbb{K}^n

Nach einer gewissen Zeit werden Sie sich an das Umgehen mit Koordinaten so gewöhnt haben, dass man rundweg den Vektorraum V mit dem Koordinatenvektorraum \mathbb{K}^n identifiziert.

Der Beweis von Theorem 5.3.1 ist leicht, da er unmittelbar aus der Definition 5.2.19 folgt. Er ist aber für die tatsächliche Berechnung des Koordinatenvektors nutzlos, da er eine reine Existenzaussage macht und keine Handhabe gibt, wie die Berechnung tatsächlich zu erfolgen hat. Deshalb geben wir jetzt einen Algorithmus an, wie Koordinatenvektoren zu berechnen sind:

Algorithmus 7.1.1 (Koordinatenvektor). :

Input:

Vektorraum V
 Basis $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ des n -dimensionalen Vektorraums V
 Vektor $\vec{v} \in V$.

Output: Koordinatenvektor

$$\vec{v}_{\mathcal{B}} := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Durchführung:

\vec{v} ist Linearkombination der Basisvektoren, d.h. es gibt Zahlen $\alpha_1, \alpha_2, \dots, \alpha_n$ so, dass

$$\vec{v} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \dots + \alpha_n \vec{b}_n.$$

Um diese zu berechnen, gehen wir so vor:

1. Schritt: Stelle ein LGS für die Koordinaten $\alpha_1, \alpha_2, \dots, \alpha_n$ auf.
2. Schritt: Löse das LGS z.B. mit dem Gaußalgorithmus.

Beispiel 7.1.2.

Input:

$$\begin{aligned} V &= \mathbb{R}^{2,2} \\ \mathcal{B}_1 &= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\ A &= \begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix} \end{aligned}$$

Output: Koordinatenvektor

$$A_{\mathcal{B}_1} := \begin{bmatrix} \alpha_1 = 1 \\ \alpha_2 = 2 \\ \alpha_3 = 3 \\ \alpha_4 = 2 \end{bmatrix}$$

Durchführung:

1. Schritt: Weil \mathcal{B}_1 eine Basis des $\mathbb{R}^{2,2}$ ist, gibt es nach Satz 5.3.1 eindeutige Zahlen $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (Koordinaten von A bezüglich \mathcal{B}_1), sodass

$$\alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix}.$$

Durch Komponentenvergleich ist diese Matrixgleichung zu dem folgenden LGS äquivalent:

$$\begin{aligned}\alpha_1 + \alpha_4 &= 3 \\ \alpha_2 &= 2 \\ \alpha_3 &= 3 \\ -\alpha_1 + \alpha_4 &= 1\end{aligned}$$

2. Schritt: Dieses Gleichungssystem kann mit dem Gaußalgorithmus gelöst werden und führt auf das im Output angegebene Resultat.

Jetzt betrachten wir ein zweites Beispiel. Es ist so gewählt, dass daraus die Bedeutung der Basiswahl erkennbar wird. Der Vektorraum und der Vektor ist so wie in Beispiel 7.1.2, nur die Basis ist anders gewählt.

Beispiel 7.1.3.

Input:

$$\begin{aligned}V &= \mathbb{R}^{2,2} \\ \mathcal{B}_2 &:= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\ A &:= \begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix}\end{aligned}$$

Output: Koordinatenvektor

$$A_{\mathcal{B}_2} := \begin{bmatrix} \beta_1 & = & 3 \\ \beta_2 & = & 5/2 \\ \beta_3 & = & -1/2 \\ \beta_4 & = & 1 \end{bmatrix}$$

Durchführung:

1. Schritt: Nach Satz 5.3.1 gibt es wiederum Zahlen $\beta_1, \beta_2, \beta_3, \beta_4$ (Koordinaten von A bezüglich \mathcal{B}_2), sodass

$$\beta_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \beta_3 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \beta_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix}$$

Diese Matrixgleichung ist zu dem folgenden LGS äquivalent:

$$\begin{aligned}\beta_1 &= 3 \\ \beta_2 + \beta_3 &= 2 \\ \beta_2 - \beta_3 &= 3 \\ \beta_4 &= 1\end{aligned}$$

2. Schritt: Dieses Gleichungssystem führt auf das im Output angegebene Resultat.

Fazit: *Derselbe Vektor hat ganz unterschiedliche Koordinatenvektoren, je nachdem auf welche Basis sich die Koordinatenvektoren beziehen.*

7.2 DIE TRANSFORMATION DER KOORDINATENVEKTOREN BEI BASISWECHSEL

Wir haben soeben gesehen, dass ein und derselbe Vektor durch verschiedene Koordinatenvektoren dargestellt wird, falls sich diese auf verschiedene Basen beziehen. Diese Situation kann sehr übersichtlich in dem folgenden Diagramm zusammengefasst werden:

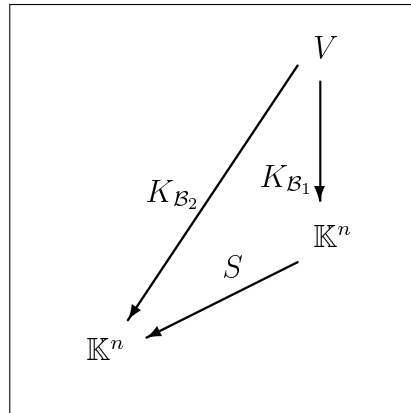


ABBILDUNG 7.2: Die Transformation der Koordinatenvektoren bei Basiswechsel

Aus diesem sogenannten *kommutativen Diagramm* kann sofort abgelesen werden, wie der erste Koordinatenvektor \vec{v}_{B_1} in den zweiten \vec{v}_{B_2} umgerechnet werden kann:

$$\vec{v}_{B_2} = S \vec{v}_{B_1}, \quad S := K_{B_2} \circ K_{B_1}^{-1}$$

S ist eine lineare Abbildung von \mathbb{K}^n auf \mathbb{K}^n und damit eine $n \times n$ Matrix.

Wir überprüfen dies an dem vorangehenden Beispiel nach der folgenden Strategie:

1. Berechnung der Abbildung $K_{B_1}^{-1}$.
2. Berechnung der Abbildung K_{B_2} .
3. Berechnung der Komposition $S := K_{B_2} \circ K_{B_1}^{-1}$
4. Verifikation der Gleichung

$$\begin{bmatrix} 3 \\ 5/2 \\ -1/2 \\ 1 \end{bmatrix} = S \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}$$

Sie drückt aus, dass der Koordinatenvektor aus Beispiel 7.1.3 das Bild des Koordinatenvektors in Beispiel 7.1.2 ist.

zu 1: Die Abbildung $K_{B_1}^{-1}$ geht von \mathbb{R}^4 nach $\mathbb{R}^{2,2}$. Sei also der Koordinatenvektor $A_{B_1} \in \mathbb{R}^4$ der allgemeinen Matrix A gegeben. Dann ist

$$A := \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = K_{B_1}^{-1}(A_{B_1}),$$

das heißt konkret

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Somit ist die Abbildung $K_{\mathcal{B}_1}^{-1}$ durch folgendes Gleichungssystem beschrieben:

$$\begin{aligned} a_{1,1} &= \alpha_1 + \alpha_4 \\ a_{1,2} &= \alpha_2 \\ a_{2,1} &= \alpha_3 \\ a_{2,2} &= -\alpha_1 + \alpha_4 \end{aligned} \quad (7.1)$$

zu 2: Die Abbildung $K_{\mathcal{B}_2}$ geht von $\mathbb{R}^{2,2}$ nach \mathbb{R}^4 . Sei also $A := \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ vorgegeben. Dann ist

$$A_{\mathcal{B}_2} = K_{\mathcal{B}_2} \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right).$$

Um $A_{\mathcal{B}_2}$ aus $\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ zu berechnen, muss das lineare Gleichungssystem

$$\beta_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \beta_3 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \beta_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

nach $\beta_1, \beta_2, \beta_3, \beta_4$ aufgelöst werden. Dies ist äquivalent zu dem LGS

$$\begin{aligned} \beta_1 &= a_{1,1} \\ \beta_2 + \beta_3 &= a_{1,2} \\ \beta_2 - \beta_3 &= a_{2,1} \\ \beta_4 &= a_{2,2}. \end{aligned}$$

Daraus folgt

$$\begin{aligned} \beta_1 &= a_{1,1} \\ \beta_2 &= (a_{1,2} + a_{2,1})/2 \\ \beta_3 &= (a_{1,2} - a_{2,1})/2 \\ \beta_4 &= a_{2,2}. \end{aligned} \quad (7.2)$$

zu 3: Zur Berechnung der Abbildung S müssen die beiden Abbildungen $K_{\mathcal{B}_2}$ und $K_{\mathcal{B}_1}^{-1}$ hintereinander ausgeführt werden, d.h. die Gleichungen 7.1 müssen in die Gleichungen 7.2 eingesetzt werden. Das Resultat ist:

$$\begin{aligned} A_{\mathcal{B}_2} &= S A_{\mathcal{B}_1} \\ S &= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

zu 4: Jetzt verifizieren wir die Gleichung

$$\begin{bmatrix} 3 \\ 5/2 \\ -1/2 \\ 1 \end{bmatrix} = S \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}$$

Allgemein wird die Matrix S nach dem folgenden Algorithmus berechnet:

Algorithmus 7.2.1 (Koordinatentransformation S). :

Input:

Vektorraum	V
zwei Basen \mathcal{B}_1	$= \{\vec{b}_1, \dots, \vec{b}_n\}$
\mathcal{B}_2	$= \{\vec{c}_1, \dots, \vec{c}_n\}$
des	n -dimensionalen Vektorraums V

Output: Koordinatentransformation S

Durchführung: \vec{b}_1 ist Linearkombination der Basisvektoren $\mathcal{B}_2 = \{\vec{c}_1, \dots, \vec{c}_n\}$. Die Koeffizienten bilden den Koordinatenvektor von \vec{b}_1 bezüglich der Basis \mathcal{B}_2 . Dieser Koordinatenvektor ist der erste Spaltenvektor der Koordinatentransformation S . Nun betrachten wir \vec{b}_2 . Auf die gleiche Weise ergibt sich daraus der zweite Spaltenvektor usw.

7.3 DIE DARSTELLUNG LINEARER ABBILDUNGEN DURCH MATRIZEN

Die Ausgangssituation ist – analog zum ersten Abschnitt in dieser Vorlesung – folgende: Sei V ein endlichdimensionaler Vektorraum über dem Körper \mathbb{K} mit der Dimension n , $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis und L eine *lineare Abbildung* von V nach V . Wir zeigen jetzt, dass L durch eine quadratische Matrix $L_{\mathcal{B}}$ dargestellt werden kann, ganz analog zur Darstellung eines Vektors \vec{v} durch einen Koordinatenvektor $\vec{v}_{\mathcal{B}}$.

Die darstellende Matrix lässt sich am einfachsten durch ein Diagramm erklären: Es zeigt, wie $L_{\mathcal{B}}$ aus drei linearen Abbildungen aufgebaut ist und motiviert die folgende

Definition 7.3.1. Sei L eine lineare Abbildung des Vektorraums V der Dimension n nach V , $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis und $K_{\mathcal{B}}$ die Koordinatenabbildung. Dann heißt

$$L_{\mathcal{B}} := K_{\mathcal{B}} \circ L \circ (K_{\mathcal{B}})^{-1} \in \mathbb{K}^{n,n}$$

die *darstellende Matrix* von L bezüglich der Basis \mathcal{B} .

Um die darstellende Matrix einer linearen Abbildung effizient auszurechnen, formulieren wir unseren nächsten

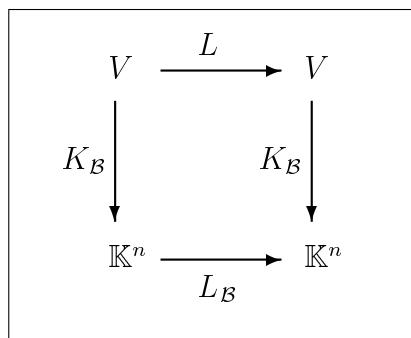
Algorithmus 7.3.2. :

Input:

Vektorraum	V
Basis	$\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ des n -dimensionalen Vektorraums V
Lineare Abbildung	$L \in \text{Hom}(V, V)$.

Output: Darstellende Matrix $L_{\mathcal{B}}$.

Durchführung:

ABBILDUNG 7.3: Die darstellende Matrix L_B

1. Schritt: Berechne die Bildvektoren $\{\vec{l}_1 := L(\vec{b}_1), \dots, \vec{l}_n := L(\vec{b}_n)\}$
2. Schritt: Verwende den Algorithmus 7.1.1 zur Berechnung der Koordinatenvektoren $(\vec{l}_1)_B, \dots, (\vec{l}_n)_B$
3. Schritt: Die Matrix L_B ist jetzt

$$L_B = [(\vec{l}_1)_B \cdots (\vec{l}_n)_B],$$

wobei $(\vec{l}_k)_B$, $k = 1, \dots, n$, den k -ten Spaltenvektor von L_B bezeichnet.

Bemerkung 7.3.3. Dass dieser Algorithmus tut was er soll, kann folgendermassen eingesehen werden: Nach Definition ist L_B die Hintereinanderausführung der drei Abbildungen $(K_B)^{-1}$, dann L , und schließlich K_B . Die Matrix L_B hat die Spaltenvektoren $\{(\vec{l}_1)_B, \dots, (\vec{l}_n)_B\}$. Diese sind die Bilder der Standardbasisvektoren $\vec{e}_1, \dots, \vec{e}_n$ des Vektorraums \mathbb{K}^n . Sei $k \in \{1, \dots, n\}$; dann ist der Basisvektor $\vec{b}_k = (K_B)^{-1}\vec{e}_k$. Der Vektor \vec{b}_k wird durch L auf \vec{l}_k abgebildet. K_B macht daraus wieder einen Koordinatenvektor $(\vec{l}_k)_B$. Dieser ist der k -te Spaltenvektor von L_B . Somit haben wir die Richtigkeit des Algorithmus verifiziert.

Nun demonstrieren wir den Algorithmus an Hand des folgenden Beispiels:

Beispiel 7.3.4.

Input:

Vektorraum	$V = \mathbb{R}_{\leq 3}[x]$
Basis	$\mathcal{B} = \{p_1(x) := x^3, p_2(x) := x^2, p_3(x) := x, p_4(x) := 1\}$
Lineare Abbildung	$D : \begin{cases} \mathbb{R}_{\leq 3}[x] & \rightarrow \mathbb{R}_{\leq 3}[x] \\ q(x) & \mapsto \frac{dq}{dx}(x) \end{cases}$

Output:

$$\text{Darstellende Matrix} \quad D_B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Durchführung:

1. Schritt:

$$\begin{aligned}
 \vec{l}_1 &= D(p_1(x)) = 3x^2 = 3p_2(x) \\
 \vec{l}_2 &= D(p_2(x)) = 2x = 2p_3(x) \\
 \vec{l}_3 &= D(p_3(x)) = 1 = 1p_4(x) \\
 \vec{l}_4 &= D(p_4(x)) = 0 = 0
 \end{aligned}$$

2. Schritt: Die Koordinatenvektoren in der Basis \mathcal{B} der Bildvektoren können ohne Rechnung aus den obigen Gleichungen abgelesen werden:

$$(\vec{l}_1)_{\mathcal{B}} = \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} \quad (\vec{l}_2)_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} \quad (\vec{l}_3)_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (\vec{l}_4)_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

3. Schritt: Damit kann jetzt die darstellende Matrix hingeschrieben werden:

$$D_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Nun kommen wir zu der Frage wie würde die darstellende Matrix aussehen, wenn wir statt der obigen Basis \mathcal{B} die folgende

$$\mathcal{B}_2 = \{r_1(x) := \frac{1}{3!}x^3, r_2(x) := \frac{1}{2!}x^2, r_3(x) := \frac{1}{1!}x, r_4(x) := \frac{1}{0!}1 = 1\}$$

verwendeten? (Ist n eine positive ganze Zahlen, so ist $n!$ das Produkt aller positiven Zahlen kleiner gleich n , also $n! := 1 \cdot 2 \cdot \dots \cdot (n-1) \cdot n$. Per Definition ist $0! = 1$.)

Eine Rechnung, die fast gleich aussieht wie die soeben durchgeführte, zeigt, dass dann die darstellende Matrix noch etwas einfacher ist:

$$\text{Darstellende Matrix} \quad D_{\mathcal{B}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

7.4 DIE TRANSFORMATION DER DARSTELLENDEN MATRIX BEI BASISWECHSEL

Wir haben soeben gesehen, dass ein und dieselbe lineare Abbildungen durch verschiedene Matrizen dargestellt wird, falls sich diese auf verschiedene Basen beziehen. Diese Situation kann übersichtlich in dem folgenden Diagramm zusammengefasst werden:

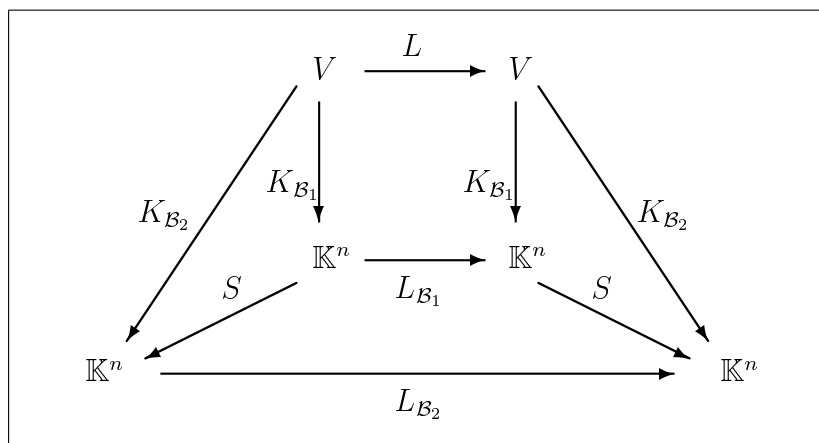


ABBILDUNG 7.4: Die Transformation der darstellenden Matrix bei Basiswechsel

Aus diesem Diagramm kann wiederum abgelesen werden, wie die erste darstellende Matrix in die zweite umgerechnet werden kann.

$$\begin{aligned} L_{\mathcal{B}_2} &= S \circ L_{\mathcal{B}_1} \circ S^{-1} \\ S &= K_{\mathcal{B}_2}(K_{\mathcal{B}_1})^{-1} \end{aligned}$$

Jetzt verifizieren wir diese Formel an Hand der Beispiele ?? und ??: Dazu muss der entsprechende Ausdruck für die Koordinatentransformation S berechnet werden. Da dies bereits im Abschnitt 2 vorgeführt wurde, sei hier nur gerade das Resultat angegeben:

$$\begin{aligned} S &= K_{\mathcal{B}_4}(K_{\mathcal{B}_3})^{-1} \\ S &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \\ S^{-1} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \end{aligned}$$

Nun kann die Gleichung

$$L_{\mathcal{B}_2} = S \circ L_{\mathcal{B}_1} \circ S^{-1}$$

leicht nachgerechnet werden.

Bemerkung 7.4.1. Um die Darstellung übersichtlich zu gestalten, haben wir hier nur lineare Abbildungen eines Vektorraumes V in sich und ihre Matrixdarstellungen diskutiert. In manchen Anwendungen kommt jedoch ein etwas allgemeinerer Fall vor, wo es um lineare Abbildungen aus einem Vektorraum V in einen zweiten W geht. Auch in dem Fall kann eine Matrixdarstellung gefunden werden, falls in beiden Räumen Basen zur Verfügung stehen. Die Matrizen sind dann nicht mehr unbedingt quadratisch sondern vom Format $\dim(W) \times \dim(V)$.

8. KAPITEL

EUKLIDISCHE UND UNITÄRE RÄUME

Die heutige Vorlesung hat die geometrischen Konzepte *Länge*, *Abstand*, also die Länge der Differenz zweier Vektoren, und *Winkel* zwischen zwei Vektoren zum Thema. Dies sind Strukturen, die zu denen des Vektorraumes dazukommen und oft zur Modellierung von Phänomenen sinnvoll sind – manchmal aber auch nicht.

Hier sind drei Anwendungsbeispiele, die sich dadurch unterscheiden, dass in den einen Fällen Länge, Abstand und Winkel sinnvoll sind, in den anderen nicht. Das erste Beispiel ist die Modellierung eines idealen Gases in einem abgeschlossenen Gefäß. Seine Zustände sind durch Vektoren $\vec{x} = (p, T) \in \mathbb{R}^2$ charakterisiert, wobei p für Druck und T für Temperatur steht. Es macht keinen Sinn vom Winkel zweier Gaszustände zu sprechen. Auch der Abstand ist als Begriff nicht unmittelbar einsichtig. Im zweiten Beispiel betrachten wir – wie bereits oft – die Modellierung von Bildern durch Matrizen. Hier ist es durchaus sinnvoll von Abständen zu sprechen, da man beispielsweise die Qualität einer Kopie zum Ausdruck bringen will. Diese ist dann als Abstand vom Original zur Kopie aufzufassen. Das letzte Beispiel ist die Modellierung der Punkte in einer Ebene mit einem ausgezeichneten Punkt. Wir stellen uns vor, dass eine Gerade gegeben ist und ein Punkt außerhalb der Geraden. Es geht nun um die Aufgabe, den Abstand zu der Gerade zu beschreiben. Dazu muss das Lot auf die Gerade gefällt und ihre Länge gemessen werden. Hier sind also sowohl Abstand wie Winkel von Bedeutung. Vektorräume dieser Art heißen euklidische oder unitäre Räume, je nach dem ob es sich um Vektorräume über den reellen oder den komplexen Zahlen handelt.

8.1 NORM, BEISPIELE

Der Begriff der Norm, oder wie man auch sagt der Länge, ist aus der geometrischen Anschauung heraus motiviert. Die Länge einer Strecke ist immer positiv, sie ist genau dann Null, wenn die Strecke nur aus einem Punkt besteht. Es gilt die Dreiecksungleichung, d.h. die Länge der Seite eines Dreiecks ist immer kleiner als die Summe der Länge der beiden anderen. Diese heuristischen Vorstellungen werden jetzt abstrahiert und zum Anlass der folgenden Definition genommen:

Definition 8.1.1 (Norm, Länge). Sei V ein linearer Raum über \mathbb{K} .

$$\begin{aligned} \text{Eine Abbildung } \|\cdot\| : V &\rightarrow \mathbb{R}^+ = \{x \in \mathbb{R} \mid x \geq 0\} \\ \vec{v} &\mapsto \|\vec{v}\| \end{aligned}$$

heißt Norm, wenn folgende Eigenschaften gelten:

1. $\|\vec{v}\| \geq 0$ für alle $\vec{v} \in V$, $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$ für alle $\vec{v} \in V$
2. $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ (Dreiecksungleichung) für alle $\vec{u}, \vec{v} \in V$
3. $\|\alpha \vec{v}\| = |\alpha| \cdot \|\vec{v}\|$ für alle $\alpha \in \mathbb{K}$, $\vec{v} \in V$

Nun betrachten wir eine Reihe von Beispielen. Dabei wird ein wichtiger Aspekt deutlich werden: Es ist in der linearen Algebra üblich, für eine Norm immer das Symbol $\|\cdot\|$ zu verwenden, obwohl es sich natürlich bei weitem nicht immer um dieselbe Norm-Funktion handelt. Da einzig und allein die Rechenregeln, die in der Definition festgelegt sind, in den Argumentationen eine Rolle spielen, ist die genaue Ausgestaltung der Norm irrelevant. Wir hatten diesen Standpunkt bereits bei der Addition kennen gelernt, wo für die unterschiedlichsten Arten der Addition immer dasselbe Symbol $+$ verwendet wurde. Diese zu Beginn etwas verwirliche Notation ist in Wahrheit eine der großen Leistungen der linearen Algebra. Man konzentriert sich hier ganz auf die abstrakten *Spielregeln* und lässt es egal sein, womit man spielt. Dies macht das Schema so leistungsfähig.

Beispiel 8.1.2. [Standardnorm] Als Vektorraum V verwenden wir \mathbb{R}^n mit der üblichen Summe und skalaren Multiplikation. Folgende Funktion auf V heißt die *Standardnorm* auf \mathbb{R}^n

$$\|\vec{x}\| := (x_1^2 + \dots + x_n^2)^{1/2}, \quad \vec{x} \in V.$$

Alle Eigenschaften der Norm sind einfach nachzuprüfen bis auf die Dreiecksungleichung. Diese folgt aus der sogenannten *Ungleichung von Cauchy-Schwarz* 8.3.1, 8.3.4, die wir gegen Ende der Vorlesung formulieren und beweisen wollen. Zum besseren Verständnis der Bedeutung des durch die Standardnorm gegebenen Längenbegriffs betrachten wir in \mathbb{R}^2 die Menge der sog. Einheitsvollkugel $\{\vec{x} | \vec{x} \in \mathbb{R}^2, \|\vec{x}\| \leq 1\}$. Sie hat die Gestalt einer Kreisscheibe, vgl. folgende Abbildung 8.1.

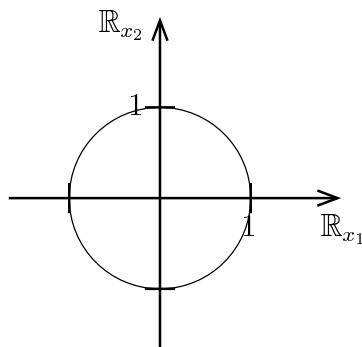


ABBILDUNG 8.1: Die Einheitskugel für die Standardnorm

Beispiel 8.1.3. [Maximumsnorm] Der Vektorraum ist so wie im letzten Beispiel 8.1.2 gewählt. Die Norm ist jetzt so definiert:

$$\|\vec{x}\|_{max} := \max\{|x_1|, |x_2|, \dots, |x_n|\}, \quad \vec{x} \in V.$$

Diesmal ist es einfach nachzuprüfen, dass alle Eigenschaften der Norm erfüllt sind. Die Einheitskugel in \mathbb{R}^2 hat nun die Gestalt eines Quadrates, vgl. folgende Abbildung 8.2. Sie wird zum Beispiel beim Platinenbohren verwendet.

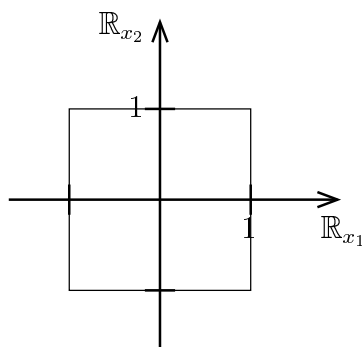


ABBILDUNG 8.2: Die Einheitskugel für die Maximumsnorm

Beispiel 8.1.4. [L^1 Norm] Der Vektorraum ist so wie im Beispiel 8.1.2 gewählt. Die Norm ist jetzt so definiert:

$$\|\vec{x}\| := |x_1| + \dots + |x_n|, \quad \vec{x} \in V.$$

Diesmal ist es einfach nachzuprüfen, dass alle Eigenschaften der Norm erfüllt sind. Die Einheitskugel in \mathbb{R}^2 hat wieder die Gestalt eines Quadrates, allerdings steht es auf der Spitze, vgl. folgende Abbildung 8.3.

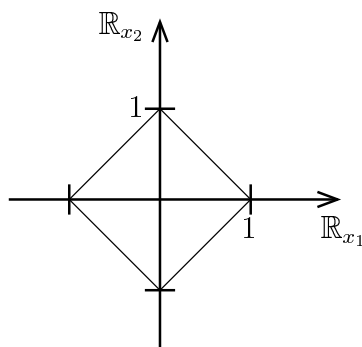


ABBILDUNG 8.3: Die Einheitskugel für die Norm im Beispiel 8.1.4

Beispiel 8.1.5. Sei A ein Element aus dem Vektorraum der komplexen 2×2 Matrizen $\mathbb{C}^{2,2}$. Wir definieren zwei Normen wie folgt:

$$\|A\|_1 = \sum_{i,k} |a_{ik}| \quad (8.1)$$

$$\|A\|_\infty = \max |a_{ik}| \quad (8.2)$$

Beispiel 8.1.6. Sei N eine natürliche Zahl und $p = \sum_{n=0}^N a_n x^n$ ein Element aus dem Vektorraum $\mathbb{R}_{\leq N}[x]$ der reellen Polynome in der Variablen x vom Grad nicht größer als N . Wir definieren eine Norm wie folgt:

$$\|p\| = \sum_{n=0}^N |a_n| \quad (8.3)$$

8.2 SKALARPRODUKT, BEISPIELE

Wir deduzieren wieder aus der Anschauung heraus und setzen voraus, dass wir die Länge von Vektoren und den Winkel zwischen Vektoren messen sowie den zugehörigen Cosinus ausrechnen können. Damit kann dann das Skalarprodukt der Vektoren \vec{u} und \vec{v} folgendermaßen definiert werden:

$$\langle \vec{u}, \vec{v} \rangle := \|\vec{u}\| \|\vec{v}\| \cos \angle(\vec{u}, \vec{v})$$

Speziell gilt dann für einen Vektor \vec{v} der Länge eins, d.h. für $\|\vec{v}\| = 1$ die Gleichung

$$\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \cos \angle(\vec{u}, \vec{v}).$$

Dies lässt sich zeichnerisch veranschaulichen: Abbildung 8.4.

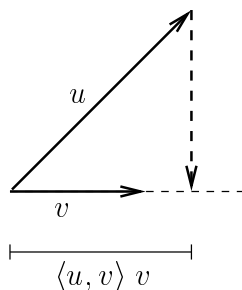


ABBILDUNG 8.4: Geometrische Bedeutung des Standard-Skalarproduktes

Die positive Zahl $|\langle \vec{u}, \vec{v} \rangle|$ ist die Länge der orthogonalen Projektion von \vec{u} auf die von \vec{v} aufgespannte Gerade. Der Vektor $\langle \vec{u}, \vec{v} \rangle \vec{v}$ heißt die orthogonale Projektion von \vec{u} auf \vec{v} . Für zwei orthogonale Vektoren \vec{u}, \vec{v} , d.h. für zwei Vektoren die aufeinander senkrecht stehen, ist $\cos \angle(\vec{u}, \vec{v}) = 0$ und damit *das Skalarprodukt gleich Null*.

Nun gehen wir von der Anschauung zur abstrakten Definition des Skalarproduktes über.

Definition 8.2.1 (euklidisches Skalarprodukt). Sei V ein Vektorraum über \mathbb{R} . Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ heißt *euklidisches Skalarprodukt*, oder kurz *Skalarprodukt*, wenn für alle $\vec{u}, \vec{v}, \vec{w} \in V, \alpha \in \mathbb{R}$ folgende Eigenschaften gelten:

1. $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$
2. $\langle \alpha \vec{u}, \vec{v} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle$
3. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$
4. $\langle \vec{v}, \vec{v} \rangle \geq 0, \langle \vec{v}, \vec{v} \rangle = 0 \Leftrightarrow \vec{v} = \vec{0}$

Bemerkung 8.2.2. In Übertragung der Sprechweise wie sie in der euklidischen Geometrie üblich ist heißen zwei Vektoren, deren Skalarprodukt gleich Null ist, *orthogonal*.

Nun fügen wir wiederum eine Reihe von Beispielen an.

Beispiel 8.2.3. [Standard-Skalarprodukt auf \mathbb{R}^n] Der Vektorraum sei wie in Beispiel 8.1.2 gewählt. Dann ist die folgende Funktion auf $V \times V$ ein Skalarprodukt

$$\langle \vec{u}, \vec{v} \rangle := u_1 v_1 + \dots + u_n v_n, \quad \vec{u}, \vec{v} \in V.$$

Dass diese Funktion alle Eigenschaften des Skalarproduktes erfüllt, ist einfach nachzuprüfen.

Beispiel 8.2.4. [Kinetische Energie] Der Vektorraum sei wie in Beispiel 8.1.2 gewählt. Seien m_1, m_2, \dots, m_n Zahlen größer Null. Dann ist die folgende Funktion auf $V \times V$ ein Skalarprodukt

$$\langle \vec{u}, \vec{v} \rangle := m_1 u_1 v_1 + \dots + m_n u_n v_n, \quad \vec{u}, \vec{v} \in V.$$

Dass diese Funktion alle Eigenschaften des Skalarproduktes erfüllt, ist wiederum einfach nachzuprüfen. Dieses Beispiel kommt unter anderem in der Mechanik im Zusammenhang mit der kinetischen Energie vor.

Beispiel 8.2.5. [Trigonometrische Polynome] Sei N eine natürliche Zahl. Dann heißt

$$V = \{p \mid p(x) = a_0 + \sum_{k=1}^N (a_k \cos(kx) + b_k \sin(kx)), \quad a_k \text{ und } b_k \text{ aus } \mathbb{R}\}$$

die Menge der trigonometrischen Polynome der Ordnung N . Die Summe im Sinne der Summe von Funktionen und die Multiplikation mit einer reellen Zahl macht daraus einen Vektorraum. Darauf definieren wir ein Skalarprodukt wie folgt:

$$\langle p, \tilde{p} \rangle := a_0 \tilde{a}_0 + \sum_{k=1}^N (a_k \tilde{a}_k + b_k \tilde{b}_k).$$

In der Analysis wird dann gezeigt, dass gilt

$$\langle p, \tilde{p} \rangle = 1/\pi \int_0^\pi p(x) \tilde{p}(x) dx. \quad (8.4)$$

Definition 8.2.6 (Euklidischer Vektorraum). Ein Vektorraum über den reellen Zahlen mit Skalarprodukt heißt *euklidischer Vektorraum*.

Skalarprodukte auf Vektorräumen über den komplexen Zahlen sind fast gleich definiert, aber nicht ganz:

Definition 8.2.7 (unitäres Skalarprodukt). Sei V ein Vektorraum über \mathbb{C} . Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ heißt *unitäres Skalarprodukt*, oder kurz *Skalarprodukt*, wenn für alle $\vec{u}, \vec{v}, \vec{w} \in V, \alpha \in \mathbb{C}$ folgende Eigenschaften gelten:

1. $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$
2. $\langle \alpha \vec{u}, \vec{v} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle$
3. $\langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$
4. $\langle \vec{v}, \vec{v} \rangle \geq 0, \langle \vec{v}, \vec{v} \rangle = 0 \Leftrightarrow \vec{v} = \vec{0}$

Definition 8.2.8 (unitärer Vektorraum). Ein Vektorraum über den komplexen Zahlen mit Skalarprodukt heißt *unitärer Vektorraum*.

Beispiel 8.2.9. [Standard-Skalarprodukt auf \mathbb{C}^n] Der Vektorraum sei wie in Bemerkung 2.1.7 gewählt. Dann ist die folgende Funktion auf $\mathbb{C}^n \times \mathbb{C}^n$ ein Skalarprodukt

$$\langle \vec{u}, \vec{v} \rangle := u_1 \bar{v}_1 + \dots + u_n \bar{v}_n, \quad \vec{u}, \vec{v} \in \mathbb{C}^n.$$

Dass diese Funktion alle Eigenschaften des Skalarproduktes erfüllt, ist einfach nachzuprüfen.

8.3 EIGENSCHAFTEN VON SKALARPRODUKTEN UND NORMEN

Zu jedem Skalarprodukt gehört eine assoziierte Norm:

Theorem 8.3.1. *Sei V ein euklidischer oder unitärer Vektorraum. Dann ist die Abbildung*

$$\begin{aligned}\|\cdot\| : V &\rightarrow \mathbb{K} \\ \vec{v} &\mapsto \|\vec{v}\| := \sqrt{\langle \vec{v}, \vec{v} \rangle}\end{aligned}$$

eine Norm. Sie heißt die zu dem Skalarprodukt assoziierte Norm.

Bemerkung 8.3.2. Alle Eigenschaften einer Norm sind leicht nachzuprüfen, bis auf die Dreiecksungleichung. Diese folgt aus der *Cauchy-Schwarzschen Ungleichung*.

Bemerkung 8.3.3. Im Zusammenhang mit euklidischen und unitären Vektorräumen wird meistens einfach von der Norm gesprochen. Es ist dann immer die assoziierte Norm gemeint.

Skalarprodukt und assoziierte Norm genügen einer wichtigen Ungleichung:

Theorem 8.3.4 (Cauchy-Schwarzsche Ungleichung). *Sei $\langle \cdot, \cdot \rangle$ das Skalarprodukt eines euklidischen oder unitären Vektorraums und $\|\cdot\|$ die assoziierte Norm. Dann gilt für alle $\vec{u}, \vec{v} \in V$ die Ungleichung*

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|. \quad (8.5)$$

Beweis. 1. Wir betrachten hier den Fall eines euklidischen Vektorraums. Der andere Fall eines unitären Vektorraums geht analog. Betrachte das quadratische Polynom in α ,

$$\begin{aligned}p(\alpha) &= \langle \vec{u} + \alpha \vec{v}, \vec{u} + \alpha \vec{v} \rangle \\ &= \alpha^2 \langle \vec{v}, \vec{v} \rangle + 2\alpha \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{u} \rangle\end{aligned}$$

2. Aus den definitiven Eigenschaften des Skalarproduktes 8.2.1, 8.2.7 folgt, dass das Polynom keine negativen Werte annehmen kann, d.h. für alle $\alpha \in \mathbb{R}$, ist $p(\alpha) \geq 0$.

3. Die obige Aussage gilt insbesondere für den Wert von α für den die Ableitung p' verschwindet, d.h. für die Nullstelle von

$$\begin{aligned}p'(\alpha) &= 2\alpha \langle \vec{v}, \vec{v} \rangle + 2\langle \vec{u}, \vec{v} \rangle \\ \alpha_{min} &= -\langle \vec{u}, \vec{v} \rangle / \langle \vec{v}, \vec{v} \rangle \\ \langle \vec{v}, \vec{v} \rangle \langle \vec{u}, \vec{u} \rangle &\geq \langle \vec{u}, \vec{v} \rangle^2\end{aligned}$$

Daraus folgt durch Wurzelziehen die Behauptung.

Bemerkung 8.3.5. Wenn die Ungleichung 8.5 mit 2 multipliziert und auf beiden Seiten $\|\vec{u}\|^2 + \|\vec{v}\|^2$ addiert wird, folgt unmittelbar die Dreiecksungleichung für die assoziierte Norm.

8.4 ORTHONORMALBASIS, GRAM-SCHMIDTSCHES VERFAHREN

Motivation 8.4.1 (Orthonormalbasis). Sei $\{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ eine Basis im \mathbb{R}^3 .

Um einen Vektor $\vec{v} \in \mathbb{R}^3$ bzgl. der Basis darzustellen, muss man das folgende LGS lösen:

$$\vec{v} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \alpha_3 \vec{b}_3 \quad (\alpha_i \in \mathbb{R} \text{ gesucht}) \quad (8.6)$$

Angenommen es gelte $\langle \vec{b}_i, \vec{b}_j \rangle = 0$, $\langle \vec{b}_i, \vec{b}_i \rangle = 1$ für $i \neq j$, dann könnte man die α_i einfach berechnen durch: $\langle \vec{v}, \vec{b}_i \rangle = \dots = \alpha_i$.

Definition 8.4.2 (ONB). Sei V ein euklidischer oder unitärer Vektorraum, $\{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis. Sie heißt ONB (**O**rthonormal**b**asis), wenn $\langle \vec{b}_i, \vec{b}_j \rangle = \delta_{i,j}$ ist ($\delta_{i,j}$ ist das sog. *Kroneckersymbol*, das definiert ist durch $\delta_{i,i} := 1, \delta_{i,j} := 0$ für $i \neq j$).

Da man, wie anfangs gesehen, mit ONB's leichter rechnen kann, wollen wir nun aus einer beliebigen Basis eine ONB des Vektorraums berechnen.

Es gibt dafür ein allgemeines Verfahren, das wir zunächst im \mathbb{R}^2 erläutern wollen:

Beispiel 8.4.3. Sei eine Basis $\{\vec{b}_1, \vec{b}_2\}$ im \mathbb{R}^2 gegeben. Wir konstruieren jetzt eine "dazu passende" Orthonormalbasis: $\{\vec{q}_1, \vec{q}_2\}$.

Erster Schritt: Der 1. Vektor entsteht durch Normierung von \vec{b}_1 : $\vec{q}_1 := \frac{\vec{b}_1}{\|\vec{b}_1\|}$

Zweiter Schritt: Der zweite Vektor steht senkrecht auf dem ersten (Lot auf die von \vec{q}_1 erzeugte Gerade) und ist auf die Länge 1 normiert.

$$\vec{q}_2 = \frac{\vec{b}_2 - \langle \vec{b}_2, \vec{q}_1 \rangle \vec{q}_1}{\|\vec{b}_2 - \langle \vec{b}_2, \vec{q}_1 \rangle \vec{q}_1\|} \quad (8.7)$$

Auf dieser Idee baut das folgende Orthonormalisierungsverfahren auf:

Algorithmus 8.4.4 (Orthonormalisierungsverfahren von Gram-Schmidt). :

Input: Euklidischer oder unitärer Vektorraum V , Basis von V : $\{\vec{b}_1, \dots, \vec{b}_n\}$.

Output: ONB von V : $\{\vec{q}_1, \dots, \vec{q}_n\}$.

1. Normierung des ersten Basisvektors \vec{b}_1 :

$$\vec{q}_1 := \frac{\vec{b}_1}{\|\vec{b}_1\|} \quad (8.8)$$

2. Lot von \vec{b}_2 auf die von \vec{q}_1 erzeugte Gerade:

$$\vec{l}_2 := \vec{b}_2 - \langle \vec{b}_2, \vec{q}_1 \rangle \vec{q}_1 \quad (8.9)$$

3. Normierung des Lots:

$$\vec{q}_2 := \frac{\vec{l}_2}{\|\vec{l}_2\|} = \frac{\vec{b}_2 - \langle \vec{b}_2, \vec{q}_1 \rangle \vec{q}_1}{\|\vec{b}_2 - \langle \vec{b}_2, \vec{q}_1 \rangle \vec{q}_1\|} \quad (8.10)$$

4. Lot von \vec{b}_k auf die von $\vec{q}_1, \dots, \vec{q}_{k-1}$ erzeugte Ebene, $k = 3, \dots, n$:

$$\vec{l}_k := \vec{b}_k - \langle \vec{b}_k, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle \vec{b}_k, \vec{q}_{k-1} \rangle \vec{q}_{k-1} \quad (8.11)$$

5. Normierung des Lots \vec{l}_k , $k = 3, \dots, n$:

$$\vec{q}_k := \frac{\vec{l}_k}{\|\vec{l}_k\|} \quad (8.12)$$

Beispiel 8.4.5. Betrachte den $(N+1)$ -dimensionalen Vektorraum der reellen Polynome $\mathbb{R}_{\leq N}[x]$ mit der Standardbasis $\{b_1 = 1, b_2 = x, \dots, b_{N+1} = x^N\}$. Nun führen wir zwei unterschiedliche Skalarprodukte ein. Das Gram-Schmidt-Verfahren führt dann auf die orthogonalen Polynome von Legendre und Chebyshev:

a) $\langle p_1, p_2 \rangle_L := \int_{-1}^{+1} p_1(x) p_2(x) dx.$

Dann ist $\{q_0, q_1, \dots, q_N\}$, $q_0(x) := 1$, $q_n(x) := \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$, ($n = 1, 2, \dots, N$), eine Menge paarweise orthogonaler Polynome und es gilt: $\langle q_n, q_n \rangle = \frac{2}{2n+1}$ und $q_n(1) = 1$. Daraus kann leicht eine Orthonormalbasis konstruiert werden indem jedes Polynom q_k , $k = 1, \dots, n$ mit einem geeigneten Faktor multipliziert wird (Normierung).

b) $\langle p_1, p_2 \rangle_{Ch} := \int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}} p_1(x) p_2(x) dx$

Dann ist $\{q_0, q_1, \dots, q_N\}$, $q_0(x) := 1$, $q_n(x) := \frac{(-2)^n n!}{(2n)!} \sqrt{1-x^2} \frac{d^n}{dx^n} (1-x^2)^{n-\frac{1}{2}}$, ($n = 1, 2, \dots, N$) ein (nicht normiertes) Orthogonalsystem. Es gilt ferner

$$\begin{aligned} \langle q_0, q_0 \rangle_{Ch} &= \pi \\ \langle q_n, q_n \rangle_{Ch} &= \frac{\pi}{2}, \quad (n = 1, 2, \dots, N) \end{aligned} \quad (8.13)$$

q_n kann auch anders beschrieben werden:

$$q_n(x) = \cos(n \arccos(x)) \quad , (n = 0, 1, \dots, N) \quad (8.14)$$

8.5 DAS VEKTORPRODUKT

Im \mathbb{R}^3 gibt es noch ein weiteres wichtiges Produkt, welches eine große Bedeutung in der Anwendung hat.

Definition 8.5.1. Das Vektorprodukt in \mathbb{R}^3 ist definiert durch

$$\vec{u} \times \vec{v} = \begin{bmatrix} u_2 v_3 & - & u_3 v_2 \\ u_3 v_1 & - & u_1 v_3 \\ u_1 v_2 & - & u_2 v_1 \end{bmatrix}, \quad \text{für alle } \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3,$$

wobei u_i, v_i die Koordinaten von \vec{u}, \vec{v} in der Standardbasis sind.

Lemma 8.5.2. Es seien $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$. Dann gilt

- (a) $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$,
- (b) $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$,
- (c) $(\alpha \vec{u}) \times \vec{v} = \alpha(\vec{u} \times \vec{v})$,
- (d) $\langle \vec{u}, \vec{u} \times \vec{v} \rangle = \langle \vec{v}, \vec{u} \times \vec{v} \rangle = 0$,
- (e) $\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \cdot \|\vec{v}\|^2 - \langle \vec{u}, \vec{v} \rangle^2$.

Aussage (d) bedeutet: Das Vektorprodukt von \vec{u} und \vec{v} steht senkrecht auf \vec{u} und \vec{v} .

9. KAPITEL

ORTHOGONALE UND UNITÄRE ABBILDUNGEN

In dieser Vorlesung werden wir Abbildungen von euklidischen und unitären Räumen betrachten, die alle Strukturen erhalten. Neben der Addition und Multiplikation mit Skalaren sind dies die in der letzten Vorlesung eingeführten euklidischen und unitären Skalarprodukte; d.h. wir werden lineare Abbildungen untersuchen, die Längen und Winkel unverändert lassen.

Typische Beispiele derartiger Abbildungen sind Drehungen und Spiegelungen der Ebene \mathbb{R}^2 oder des Raums \mathbb{R}^3 . Sie lassen das standard-euklidische Skalarprodukt invariant. Abbildungen mit dieser Eigenschaft heißen orthogonal. Drehungen und Spiegelungen sind leicht zu veranschaulichen und deshalb besonders nützlich.

Neben den Drehungen gibt es eine Reihe linearer Abbildungen, die abstrakterer Natur sind, wie z.B. die diskrete Fouriertransformation, die in der Signalanalyse eine wichtige Rolle spielt und dort die z-Transformation heißt. Sie gehört zu der Klasse der unitären Abbildungen, d.h. solchen, die ein unitäres Skalarprodukt unverändert lassen.

Drehungen des dreidimensionalen Raums treten an sehr vielen Orten auf, z.B. in der Mechanik, der Geologie und Astronomie. Es geht dort beispielsweise darum, ein Koordinatensystem, das fest mit den Fixsternen verbunden ist, in eines zu überführen, das mit der Erde fest verbunden ist. Eine andere Anwendung ist bei der Beschreibung eines Kreisels von großem Nutzen. Dort verwendet man eine Drehung, die das Koordinatensystem des Labors in ein Koordinatensystem, das fest mit dem Kreisel verbunden ist, überführt. Dieses zweite Thema ist eng mit dem Namen von Leonhard Euler verbunden, weshalb diese Abbildungen auch nach ihm benannt sind (Eulersche Winkel).

9.1 DEFINITIONEN DER GRUNDLEGENDEN BEGRIFFE UND BEISPIELE

Wir beginnen mit den Abbildungen euklidischer Vektorräume, d.h. mit den Vektorräumen über den reellen Zahlen mit Skalarprodukt und definieren dafür die strukturerhaltenden linearen Abbildungen wie folgt:

Definition 9.1.1. Sei V ein euklidischer Vektorraum (vgl. Definition 8.2.6) und Q eine invertierbare lineare Abbildung von V nach V . Die Abbildung Q heißt eine *orthogonale Transformation*, falls für alle $\vec{u}, \vec{v} \in V$ gilt:

$$\langle Q\vec{u}, Q\vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle \quad (9.1)$$

Bemerkung 9.1.2. Falls V endlichdimensional ist, folgt aus (9.1) die Existenz der Inversen Abbildung Q^{-1} , denn wäre Q nicht invertierbar so würde aus $Q\vec{v} = \vec{0}$ und $\vec{v} \neq \vec{0}$ sofort ein Widerspruch folgen, wenn man in (9.1) einsetzt:

$\vec{v} = \vec{u}$. Somit kann die Definition für den Fall endlichdimensionaler Vektorräume vereinfacht werden, i.e. die Voraussetzung “ Q sei invertierbar”, kann fallengelassen werden.

Bemerkung 9.1.3. Für den Fall des euklidischen Vektorraums $V = \mathbb{R}^n$ mit Standardskalarprodukt, sind orthogonale Transformationen durch *orthogonale Matrizen* im Sinne der Definition 3.4.7 gegeben. Diese Matrizen genügen definitionsgemäß der Gleichung:

$$Q^T Q = I_n \quad (9.2)$$

Die obige Gleichung (9.2) kann folgendermassen interpretiert werden: *Die Matrix Q ist genau dann orthogonal, falls ihre Spaltenvektoren die Elemente einer Orthonormalbasis sind. Diese Sichtweise ist für viele Anwendungen besonders nützlich!*

Bemerkung 9.1.4. Das Bild einer ONB unter einer orthogonalen Transformation ist eine ONB. Dies folgt unmittelbar aus der obenstehenden Definition 9.1.1.

Wir betrachten jetzt den Vektorraum \mathbb{R}^2 mit Standardskalarprodukt $\langle \vec{u}, \vec{v} \rangle = u_1 v_1 + u_2 v_2$, d.h. den *euklidischen Vektorraum der Ebene*, und definieren die Drehung $R(\phi)$ um den Winkel $\phi \in \mathbb{R}$ folgendermassen:

$$R(\phi) := \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \quad (9.3)$$

Offensichtlich haben die beiden Spaltenvektoren die Länge eins und stehen senkrecht aufeinander. Sie sind die Bilder der kanonischen Basisvektoren und bilden wie sie eine ONB von $V = \mathbb{R}^2$. All dies ist in der Abbildung 9.1 veranschaulicht. Statt “Drehung” sagt man oftmals auch “Rotation”.

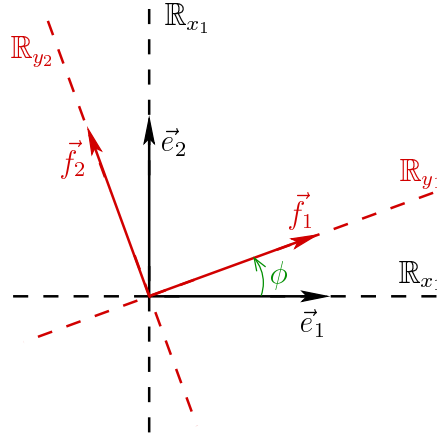


ABBILDUNG 9.1: Rotation in \mathbb{R}^2

Die Drehung $R(\phi)$ ist eine Funktion von $\phi \in \mathbb{R}$ mit Werten in den 2×2 -Matrizen. Die Eigenschaften der Drehungen um einen Winkel sind in folgendem Theorem zusammengefasst:

Theorem 9.1.5. Für alle $\phi_1, \phi_2 \in \mathbb{R}$ gilt

$$R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2). \quad (9.4)$$

Insbesondere ist

$$R(\phi_1)R(-\phi_1) = R(0) = I_2. \quad (9.5)$$

Bemerkung 9.1.6. Theorem 9.1.5 besagt, dass die Komposition zweier Drehungen um Winkel ϕ_1, ϕ_2 wieder eine Drehung vom selben Typ ist, dass es zu jeder Drehung $R(\phi)$ eine Umkehrabbildung gibt und dass die Identität zu der Familie der Drehungen um einen Winkel gehört. Eine Menge mit dieser Struktur heißt eine einparametrische kommutative (oder auch abelsche) Gruppe, kommutativ deshalb, weil es in der Gleichung 9.4 nicht auf die Reihenfolge der Faktoren ankommt.

Eine weitere Klasse orthogonaler Transformationen der Ebene sind die *Spiegelungen* an einer Geraden durch den Ursprung. Die Spiegelung S an der \vec{e}_1 -Achse ist folgendermassen definiert

$$S : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ \vec{x} & \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \vec{x} \end{cases} \quad (9.6)$$

(vgl. Abbildung 9.2).

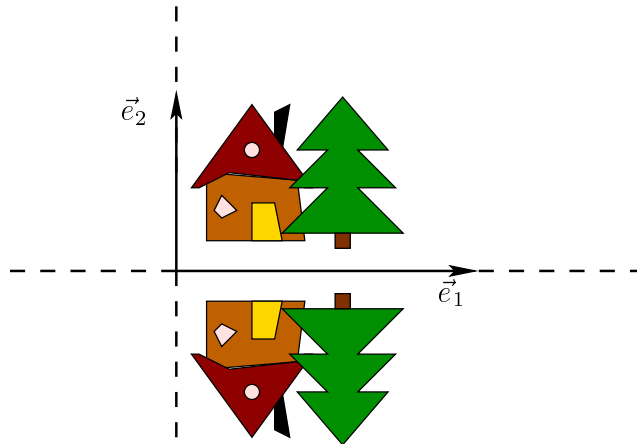


ABBILDUNG 9.2: Spiegelung in \mathbb{R}^2 an der \vec{e}_1 -Achse

Nun kommen wir zu den komplexen Vektorräumen:

Definition 9.1.7. Sei V ein unitärer Vektorraum (vgl. Definition 8.2.7) und U eine invertierbare lineare Abbildung von V nach V . Die Abbildung U heißt eine *unitäre Transformation*, falls für alle $\vec{v}, \vec{w} \in V$ gilt:

$$\langle U\vec{v}, U\vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle.$$

Bemerkung 9.1.8. Man beachte in diesem Fall die etwas andere Definition des Skalarprodukts als im reellen Fall, siehe Definition 8.2.7.

Auch für diesen Fall gilt die Bemerkung 9.1.2, d.h. im Falle endlichdimensionaler Vektorräume kann die obige Definition vereinfacht werden und die Voraussetzung, "U sei invertierbar" weggelassen werden.

Bemerkung 9.1.9. Die unitären Transformationen von \mathbb{C}^n mit dem Standardskalarprodukt sind *unitäre Matrizen*, die definitionsgemäß der folgenden Gleichung genügen:

$$U^*U = I_n. \quad (9.7)$$

Die obige Gleichung (9.7) kann - analog zur Gleichung (9.2) folgendermaßen interpretiert werden: *Die Matrix U ist genau dann unitär, falls ihre Spaltenvektoren die Elemente einer Orthonormalbasis von \mathbb{C}^n sind.*

Beispiel 9.1.10. [unitäre 2×2 -Matrizen] Sei $V = \mathbb{C}^2$ mit dem Standardskalarprodukt $\langle \vec{v}, \vec{w} \rangle = v_1 \overline{w_1} + v_2 \overline{w_2}$ der zweidimensionale unitäre Vektorraum. Darauf definieren wir die Abbildung U so, dass die beiden Spaltenvektoren senkrecht zueinander stehen und die Länge eins haben. Die Spaltenvektoren, die ja immer als die Bilder der orthonormierten Standardbasis aufgefasst werden können, bilden also eine ONB von $V = \mathbb{C}^2$: Sei

$$\begin{aligned} \vec{z} &:= \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad \text{ein Vektor der Länge 1, d.h. } |z_1|^2 + |z_2|^2 = 1 \\ U &:= \begin{bmatrix} z_1 & -\overline{z_2} \\ z_2 & \overline{z_1} \end{bmatrix} \end{aligned}$$

Sie kann - mit etwas mehr Anstrengung - als Drehung des dreidimensionalen euklidischen Raums aufgefasst werden.

Beispiel 9.1.11. Es gibt natürlich auch Drehungen und Spiegelungen im \mathbb{R}^n , $n > 2$.

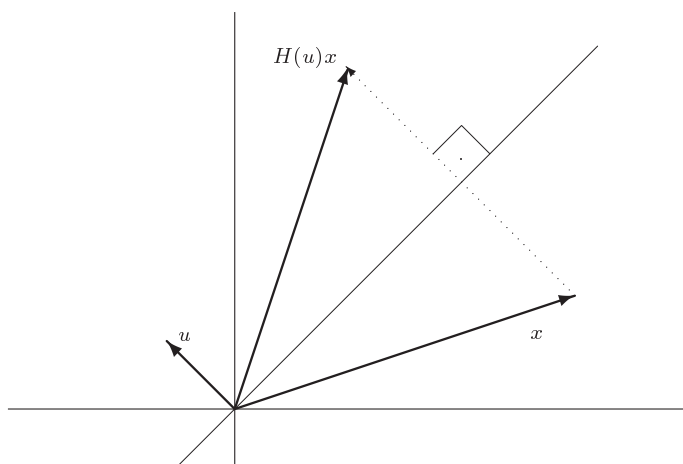
(a) Jede Matrix der Form

$$H(\vec{u}) := I - 2 \frac{\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}}, \quad \vec{u} \in \mathbb{R}^n \quad \vec{u} \neq \vec{0},$$

ist orthogonal, denn

$$\begin{aligned} H(\vec{u})^T H(\vec{u}) &= \left(I_n - \frac{2\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} \right)^T \left(I_n - \frac{2\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} \right) \\ &= \left(I_n - \frac{2\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} \right) \left(I_n - \frac{2\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} \right) = H(\vec{u})^2 \\ &= I_n - \frac{4\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} + 4 \frac{\vec{u} (\vec{u}^T \vec{u}) \vec{u}^T}{(\vec{u}^T \vec{u}) (\vec{u}^T \vec{u})} \\ &= I_n - \frac{4\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} + 4 \frac{\vec{u} \vec{u}^T}{\vec{u}^T \vec{u}} = I_n. \end{aligned}$$

Die Multiplikation von $H(\vec{u})$ mit einem Vektor $\vec{v} \in \mathbb{R}^n$ bewirkt eine Spiegelung von \vec{v} an der Hyperebene die senkrecht zu \vec{u} steht $\{\vec{w} \in \mathbb{R}^n \mid \vec{u}^T \vec{w} = 0\}$.



Wie man es bei einer Spiegelung erwarten sollte gilt $H(\vec{u}) \cdot (H(\vec{u})\vec{v}) = (H(\vec{u}))^2 \cdot \vec{v} = \vec{v}$.

(b) Jede Matrix der Form

$$R_{ij}(\alpha) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} & & & & \\ & \cos \alpha & & -\sin \alpha & \\ & & 1 & & \\ & & & \ddots & \\ & \sin \alpha & & & \cos \alpha \\ & & & & & 1 \end{bmatrix} \begin{matrix} \leftarrow i \\ \leftarrow j \\ \end{matrix}$$

$\uparrow \qquad \qquad \uparrow$
 $i \qquad \qquad j$

ist orthogonal, denn

$$R_{ij}(\alpha)^T R_{ij}(\alpha) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} & & & & \\ & \cos^2 \alpha + \sin^2 \alpha & & 0 & \\ & & 1 & & \\ & & & \ddots & \\ & 0 & & & \cos^2 \alpha + \sin^2 \alpha \\ & & & & & 1 \end{bmatrix}$$

$= I_n.$

Die Multiplikation von $R_{ij}(\alpha)$ mit $\vec{v} \in \mathbb{R}^n$ beschreibt eine Drehung in der (i, j) -Koordinatenebene.

9.2 DIE QR -ZERLEGUNG

Eine der wichtigsten Anwendungen der orthogonalen (unitären) Abbildungen ist die Lösung von *Ausgleichsproblemen*, die zum Anpassen von Funktionen an gemessene Daten in der Modellierung verwendet wird, wie dies im Beispiel (9.2.4) am Ende dieses Abschnittes an Hand des *linearen Ausgleichsproblems* vorgestellt wird.

Das lineare Ausgleichsproblem führt auf ein lineares Gleichungssystem, das wir hier nicht mit dem Gaußalgorithmus sondern mit einer neuen Methode, der Methode der QR -Zerlegung lösen wollen. Die Idee ist kurz und unpräzise folgende: Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ kann so umgeschrieben werden: $QR\vec{x} = \vec{b}$, wobei die Matrizen Q und R so gewählt werden, dass sie *einfach invertierbar sind*: Damit ist die Lösung: $\vec{x} = R^{-1}Q^{-1}\vec{b}$.

Definition 9.2.1. Eine Zerlegung einer Matrix $A \in \mathbb{R}^{m,n}$ in ein Produkt zweier Matrizen $A = Q \cdot R$, mit $Q \in \mathbb{R}^{m,m}$ orthogonal, und $R \in \mathbb{R}^{m,n}$ obere Dreiecksmatrix, heißt QR -Zerlegung.

Es gilt der Satz:

Theorem 9.2.2. Jede Matrix $A \in \mathbb{R}^{m,n}$ kann QR zerlegt werden.

Wie erhalten wir eine QR -Zerlegung? In der Tat haben wir den wesentlichen Schritt bereits kennen gelernt, denn die QR -Zerlegung erweist sich bei genauerer Betrachtung als eine Variante des Gram-Schmidt-Verfahrens. Am einfachsten lässt sich die Idee der QR -Zerlegung im Falle einer quadratischen Matrix A mit linear unabhängigen Spaltenvektoren (d.h. A ist invertierbar) erklären. Wir beschränken uns auf diesen Fall.

Die n Spaltenvektoren $\{\vec{a}_1, \dots, \vec{a}_n\}$ der quadratischen Matrix A sind nach Voraussetzung linear unabhängig und bilden deshalb eine Basis des \mathbb{R}^n , den wir als euklidischen Vektorraum bzgl. des Standardskalarprodukts betrachten. Sie können nach dem Verfahren von Gram-Schmidt in eine ONB $\{\vec{q}_1, \dots, \vec{q}_n\}$ transformiert werden und definieren - als Spaltenvektoren aufgefasst - eine orthogonale Matrix, die wir mit Q bezeichnen. Die a -Vektoren lassen sich als Linearkombination der q -Vektoren schreiben. In Matrixform ausgedrückt ist dies die QR -Zerlegung von A .

Wir erklären die Methode der QR -Zerlegung am folgenden Beispiel:

Beispiel 9.2.3 (QR -Zerlegung (oder -Faktorisierung)). Ausgangspunkt ist die quadratische Matrix

$$A : \quad = \begin{bmatrix} 3 & 1 & 4 \\ 0 & 0 & -1 \\ 4 & 7 & -3 \end{bmatrix} \quad \text{mit den Spaltenvektoren}$$

$$\vec{a}_1 = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 7 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 4 \\ -1 \\ -3 \end{bmatrix}$$

Jetzt wenden wir das Gram-Schmidt-Verfahren auf die drei Spaltenvektoren an und schreiben den k -ten Spaltenvektor als Linearkombination von k -vielen

orthonormierten Vektoren, die wir schrittweise konstruieren, vgl. Algorithmus 8.4.4:

1. *Schritt*: Den ersten Spaltenvektor \vec{a}_1 schreiben wir als Vielfaches r_{11} des auf eins normierten Vektors \vec{q}_1 :

$$\begin{aligned}\vec{a}_1 &= r_{11}\vec{q}_1 \\ r_{11} &= 5, \quad \vec{q}_1 = \begin{bmatrix} 3/5 \\ 0 \\ 4/5 \end{bmatrix}\end{aligned}$$

2. *Schritt*: Den zweiten Spaltenvektor \vec{a}_2 schreiben wir als Linearkombination von *zwei* orthonormierten Vektoren: \vec{q}_1 kennen wir schon, \vec{q}_2 ist das auf die Länge eins normierte Lot von \vec{a}_2 auf die von \vec{q}_1 erzeugte Gerade:

$$\begin{aligned}\vec{a}_2 &= r_{22}\vec{q}_2 + r_{12}\vec{q}_1 \\ r_{22} &= 17/5, \quad r_{12} = 31/5, \quad \vec{q}_2 = \begin{bmatrix} -4/5 \\ 0 \\ 3/5 \end{bmatrix}\end{aligned}$$

3. *Schritt*: Den dritten Spaltenvektor \vec{a}_3 schreiben wir als Linearkombination von *drei* orthonormierten Vektoren: \vec{q}_1 und \vec{q}_2 sind bereits bekannt; der dritte Vektor \vec{q}_3 ist das auf eins normierte Lot von \vec{a}_3 auf die von den ersten beiden aufgespannte Ebene:

$$\begin{aligned}\vec{a}_3 &= r_{33}\vec{q}_3 + r_{23}\vec{q}_2 + r_{13}\vec{q}_1 \\ r_{33} &= 1, \quad r_{23} = -5, \quad r_{13} = 0, \quad \vec{q}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}\end{aligned}$$

4. *Schritt*: Jetzt werden die obigen Resultate in Matrixform geschrieben:

$$\begin{aligned}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \\ \begin{bmatrix} 3 & 1 & 4 \\ 0 & 0 & -1 \\ 4 & 7 & -3 \end{bmatrix} &= \begin{bmatrix} 3/5 & -4/5 & 0 \\ 0 & 0 & -1 \\ 4/5 & 3/5 & 0 \end{bmatrix} \begin{bmatrix} 5 & 31/5 & 0 \\ 0 & 17/5 & -5 \\ 0 & 0 & 1 \end{bmatrix}\end{aligned}$$

Damit haben wir die Matrix A QR -faktoriert: $A = QR$ mit

$$Q = \begin{bmatrix} 3/5 & -4/5 & 0 \\ 0 & 0 & -1 \\ 4/5 & 3/5 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 5 & 31/5 & 0 \\ 0 & 17/5 & -5 \\ 0 & 0 & 1 \end{bmatrix}.$$

Falls A mehr Zeilen als Spalten hat ($m > n$) so kann man die QR -Zerlegung dennoch mit dem Gram-Schmidt-Verfahren machen, aber man muss dann zu den Spalten von A noch $m - n$ Vektoren $\vec{a}_{n+1}, \dots, \vec{a}_m$ dazu nehmen bis wir eine Basis des Raums \mathbb{R}^m erhalten.

Eine bessere Methode, weil wir nicht diese Vektoren dazu nehmen müssen, ist es die Matrix mit Hilfe von Multiplikationen mit Spiegelungs- oder Drehungsmatrizen auf Dreiecksform zu bringen. Das geht im Prinzip genau so wie die Transformation auf Zeilenstufenform.

Der Unterschied zum Gaußalgorithmus ist, dass durch die Transformation mit orthogonalen Matrizen Längen und Winkel erhalten bleiben und dies lässt sich z.B. nutzen um die folgende wichtige Aufgabe zu lösen.

Beispiel 9.2.4 (Lineares Ausgleichsproblem). Angenommen wir haben Messwerte (t_i, y_i) , $i = 1, \dots, m$ und suchen eine Gerade $y = a_1 t + a_2$, die diese Punkte am besten repräsentiert.¹ Mathematisch formulieren wir "am besten" mittels einer geeigneten Fehlerfunktion, die wir ad hoc als Summe der Fehlerquadrate einführen:

$$F(\vec{a}) := \sum_{i=1}^m ((a_1 t_i + a_2) - y_i)^2, \quad \vec{a} := \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad (9.8)$$

$F(\vec{a})$ ist genau dann Null, wenn die gesuchte Gerade durch alle Messpunkte geht. Andernfalls ist sie grösser als Null. Wir suchen nun diejenigen Parameterwerte a_1, a_2 , die die Fehlerfunktion minimal machen. Dies führt auf ein Problem, das sehr schön mit der Methode der QR -Zerlegung gelöst werden kann.

Als erstes bringen wir die Fehlerfunktion F in eine Gestalt, die einer geometrischen Interpretation zugänglich ist und führen dazu die Matrix A mit den Spaltenvektoren \vec{v}_1 und \vec{v}_2 und der QR -Zerlegung $A = QR$ ein:

$$A := [\vec{v}_1, \vec{v}_2], \quad \vec{v}_1 := \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{bmatrix}, \quad \vec{v}_2 := \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Zusätzlich setzen wir $\vec{y} := \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$. Damit kann $F(\vec{a})$ folgendermassen geschrieben

und umgeschrieben werden:

$$F(\vec{a}) = \|A\vec{a} - \vec{y}\|^2 \quad (9.9)$$

$$\begin{aligned} &= \|Q^T(A\vec{a} - \vec{y})\|^2 \\ &= \|R\vec{a} - Q^T\vec{y}\|^2 \\ &= \left\| \begin{bmatrix} r\vec{a} \\ \vec{0} \end{bmatrix} - \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \end{bmatrix} \right\|^2, \quad \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \end{bmatrix} := Q^T\vec{y} \\ &= \|r\vec{a} - \vec{b}_1\|^2 + \|\vec{b}_2\|^2 \end{aligned} \quad (9.10)$$

Dabei haben wir verwendet, dass die Abbildung Q^T die Länge eines Vektors unverändert lässt und die Matrix R die folgende Gestalt hat:

$$R = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} r \\ \vec{0} \end{bmatrix}$$

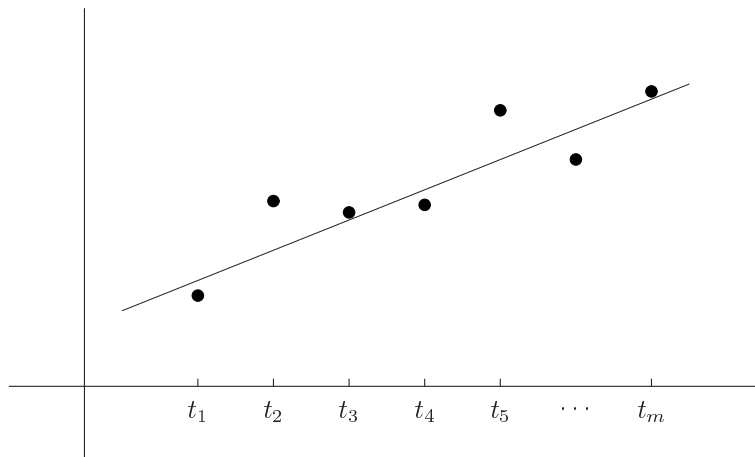
¹Wir setzen voraus nicht alle t -Werte sind gleich.

mit $r \in \mathbb{R}^{2,2}$, $\vec{0} \in \mathbb{R}^{m-2,2}$.

Die Gleichung (9.10) eignet sich hervorragend um das Minimum von $F(\vec{a})$ zu berechnen. Die rechte Seite besteht aus 2 positiven Termen; der erste ist von den Parametern \vec{a} abhängig der zweite nicht. Das Minimum wird offensichtlich angenommen, wenn der erste Term verschwindet, d.h. wenn \vec{a} das LGS $r\vec{a} = \vec{b}_1$ löst. Um das Minimum zu finden benötigen wir also die 2×2 Matrix r und den Vektor \vec{b}_1 - ausgedrückt durch die vorgegebenen Daten (t_i, y_i) , $i = 1, \dots, m$ - um damit das LGS

$$\begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \vec{q}_1^T \\ \vec{q}_2^T \end{bmatrix} \cdot \vec{y} = \vec{b}_1$$

zu lösen. Nachzutragen ist, dass \vec{q}_1, \vec{q}_2 die ersten beiden Spalten der orthogonalen Matrix Q der QR-Zerlegung von $A = [\vec{v}_1, \vec{v}_2]$ bezeichnen.



Nun kommen wir zur Berechnung der Matrix r und der beiden ersten Spaltenvektoren \vec{q}_1, \vec{q}_2 von Q , d.h. zur QR-Zerlegung von A und wenden dazu das

Gram-Schmidt-Verfahren auf die Spaltenvektoren \vec{v}_1 und \vec{v}_2 an:

$$\begin{aligned} q_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|}, \\ r_{11} &= \|\vec{v}_1\| = \sqrt{\sum_{j=1}^m t_j^2}, \\ r_{12} &= \langle \vec{v}_2, \vec{q}_1 \rangle = \frac{1}{\|\vec{v}_1\|} \langle \vec{v}_2, \vec{v}_1 \rangle = \frac{\sum_{j=1}^m t_j}{\sqrt{\sum_{j=1}^m t_j^2}}, \\ \vec{q}_2 &= \frac{\vec{v}_2 - \langle \vec{v}_2, \vec{q}_1 \rangle \vec{q}_1}{\|\vec{v}_2 - \langle \vec{v}_2, \vec{q}_1 \rangle \vec{q}_1\|} = \frac{\vec{v}_2 - r_{12} \vec{q}_1}{\|\vec{v}_2 - r_{12} \vec{q}_1\|}, \end{aligned}$$

$$\begin{aligned} \vec{v}_2 - r_{12} \vec{q}_1 &= \vec{v}_2 - \frac{\vec{v}_1}{\|\vec{v}_1\|} \cdot \frac{1}{\|\vec{v}_1\|} \langle \vec{v}_2, \vec{v}_1 \rangle = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} \cdot \frac{1}{\sum_{j=1}^m t_j^2} \cdot \sum_{j=1}^m t_j \\ &= \begin{bmatrix} 1 & - & t_1 & \frac{\sum_{j=1}^m t_j}{\sum_{j=1}^m t_j^2} \\ & & \vdots & \\ 1 & - & t_m & \frac{\sum_{j=1}^m t_j}{\sum_{j=1}^m t_j^2} \end{bmatrix}, \\ r_{22} &= \langle \vec{v}_2, \vec{q}_2 \rangle = \left\langle \vec{v}_2, \frac{1}{\|\vec{v}_2 - r_{12} \vec{q}_1\|} (\vec{v}_2 - r_{12} \vec{q}_1) \right\rangle. \end{aligned}$$

Zum Schluss ein kurzer Rückblick auf die in diesem Beispiel durchgeführte Rechnung: $F(\vec{a})$ ist das Quadrat des Abstandes von $\vec{y} \in \mathbb{R}^m$ und $A\vec{a}$. Die Menge $\{A\vec{a} | \vec{a} \in \mathbb{R}^2\}$ ist ein zweidimensionaler Teilraum von \mathbb{R}^m . Das Minimum von F ist also - geometrisch gesehen - das Quadrat der Länge des Lotes von \vec{y} auf die Fläche $\{A\vec{a} | \vec{a} \in \mathbb{R}^2\}$. Wir haben also mit Hilfe der QR -Zerlegung das Lot auf eine Fläche und dessen Länge berechnet.

Ähnlich geht man auch vor, wenn man komplizierte Parameterschätzungen aus Messwerten erzeugen will.

10. KAPITEL

DETERMINANTE

10.1 EINLEITUNG

Die Determinante ordnet jeder quadratischen Matrix $A \in \mathbb{K}^{n,n}$ eine Zahl $\det A \in \mathbb{K}$ zu. Etwas anders ausgedrückt – aber konzeptionell besser: Die Determinante ist eine Funktion von n Vektoren $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{K}^n$ mit Werten in \mathbb{K} . Die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ können als Spalten einer quadratischen Matrix A aufgefasst werden.

¹

Eine der wichtigsten Eigenschaften der Determinante ist, dass sie genau dann den Wert Null annimmt, wenn die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ linear abhängig sind. Geometrisch heißt dies, dass das von $\vec{a}_1, \dots, \vec{a}_n$ aufgespannte Parallelepiped das Volumen Null hat.

Falls die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ linear unabhängig sind, d.h. eine Basis bilden, ist der Zahlenwert der Determinante von Null verschieden. Der Betrag der Determinante hat die Bedeutung des Volumens des von $\vec{a}_1, \dots, \vec{a}_n$ aufgespannten Parallelepipeds, das Vorzeichen ist die Orientierung von $\vec{a}_1, \dots, \vec{a}_n$ (im Verhältnis zur Standardbasis).

Zur Definition der Determinante ist keine der geometrischen Strukturen wie Skalarprodukt oder Norm notwendig. Wir hätten sie also durchaus früher einführen können. Dies erst jetzt zu tun, macht jedoch interessantere Anwendungen möglich.

Die Geschichte der Determinanten beginnt bereits vor mehr als 200 Jahren. Seit dem sind Determinanten in verschiedenen Zusammenhängen immer wieder aufgetaucht. Eingeführt wurden sie ursprünglich, um lineare Gleichungssysteme und deren Lösungen zu untersuchen, insbesondere zur Entscheidung der Frage, ob ein LGS mit n Gleichungen und n Unbekannten eine eindeutige Lösung besitzt. Es gilt das folgende Kriterium: Der Rang der Matrix A ist genau dann maximal und damit gleich n , falls die Determinante von Null verschieden ist; es gibt also genau dann eine eindeutige Lösung, wenn die Determinante der Koeffizientenmatrix des LGS nicht verschwindet oder anders ausgedrückt, falls die Koeffizientenmatrix A invertierbar ist. Dies kann dadurch explizit gemacht werden, dass es eine Formel für die Inverse der Matrix A gibt in der die Determinante im Nenner eines Bruchs vorkommt (vgl. Theorem 10.2.5).²

In der nächsten Vorlesung werden wir von Eigenwerten linearer Abbildungen sprechen. Dies führt auf das Problem die Nullstellen des charakteristischen Polynoms einer Matrix zu berechnen. Auch hier geht es um eine Determinante.

¹Statt der Spaltenvektoren könnten auch Zeilenvektoren betrachtet werden.

²Computeralgebra-Programme verwenden allerdings andere sehr viel effizientere Algorithmen.

Integrale, wie sie in der Analysis auftreten, können oft durch eine Transformation der Integrationsvariablen vereinfacht werden. Dabei tritt die sogenannte Funktionaldeterminante auf. Dies ist die Determinante der Ableitungsmatrix (Jacobi-Matrix) der Transformation der Variablen.

10.2 DEFINITION UND EIGENSCHAFTEN DER DETERMINANTE

Um die Definition der Determinante zu motivieren, beginnen wir mit dem anschaulichen Fall der Dimension $n = 3$ und den oben angesprochenen geometrischen Eigenschaften.

Seien

$$\vec{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{bmatrix} \quad (j \in \{1, 2, 3\}) \quad (10.1)$$

drei Elemente des \mathbb{R}^3 , die nicht in einer Ebene liegen und deren “Drehrichtung” $\vec{a}_1 \rightarrow \vec{a}_2 \rightarrow \vec{a}_3$ so ist wie beim Koordinatensystem.

Das Volumen V des von den 3 Vektoren aufgespannten Spates (Parallelepipedes) ist dann $V = F \cdot h$:

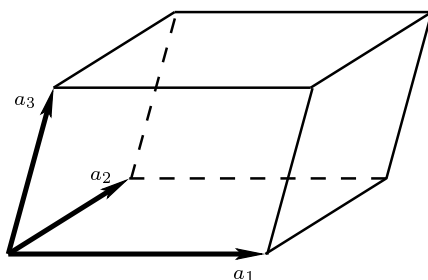


ABBILDUNG 10.1: Parallelepiped 3-dimensional

Es ist

$$\begin{aligned} F &= \|\vec{a}_1 \times \vec{a}_2\| \\ h &= \|\vec{a}_3\| \cos \gamma \end{aligned} \quad (10.2)$$

also

$$\begin{aligned} V &= \|\vec{a}_1 \times \vec{a}_2\| \cdot \|\vec{a}_3\| \cos \gamma \\ &= \langle \vec{a}_1 \times \vec{a}_2, \vec{a}_3 \rangle, \end{aligned} \quad (10.3)$$

wobei γ der Winkel zwischen den beiden Vektoren ist. Sei nun für beliebige

$$\vec{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{bmatrix} \in \mathbb{R}^3 \quad (j \in \{1, 2, 3\})$$

$$A := [\vec{a}_1, \vec{a}_2, \vec{a}_3] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (10.4)$$

die quadratische Matrix mit den \vec{a}_j als Spalten. Dann ist die Determinante von A folgendermassen definiert:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (10.5)$$

$$:= \langle \vec{a}_1, \vec{a}_2 \times \vec{a}_3 \rangle \quad (10.6)$$

Statt $\det A$ schreiben wir auch $\det[\vec{a}_1, \vec{a}_2, \vec{a}_3]$ oder

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

$\det A$ ist also bis auf das Vorzeichen gleich dem Volumen des von $\vec{a}_1, \vec{a}_2, \vec{a}_3$ aufgespannten Spates. Aus den Eigenschaften des Kreuz- und Skalarproduktes oder auch durch geometrische Argumentation erhält man dann unmittelbar folgende fünf Eigenschaften für die Abbildung:

$$\det : \mathbb{R}^{3,3} \rightarrow \mathbb{R} \quad (10.7)$$

(1) Für $\vec{a}_j \in \mathbb{R}^3$ ($j \in \{1, 2, 3\}$) $\vec{b} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$:

$$\begin{aligned} \det[\vec{a}_1 + \vec{b}, \vec{a}_2, \vec{a}_3] &= \det[\vec{a}_1, \vec{a}_2, \vec{a}_3] + \det[\vec{b}, \vec{a}_2, \vec{a}_3] \\ \det[\vec{a}_1, \vec{a}_2 + \vec{b}, \vec{a}_3] &= \det[\vec{a}_1, \vec{a}_2, \vec{a}_3] + \det[\vec{a}_1, \vec{b}, \vec{a}_3] \\ \det[\vec{a}_1, \vec{a}_2, \vec{a}_3 + \vec{b}] &= \det[\vec{a}_1, \vec{a}_2, \vec{a}_3] + \det[\vec{a}_1, \vec{a}_2, \vec{b}] \end{aligned} \quad (10.8)$$

und

$$\begin{aligned} \det[\lambda \vec{a}_1, \vec{a}_2, \vec{a}_3] &= \det[\vec{a}_1, \lambda \vec{a}_2, \vec{a}_3] \\ &= \det[\vec{a}_1, \vec{a}_2, \lambda \vec{a}_3] = \lambda \det[\vec{a}_1, \vec{a}_2, \vec{a}_3] \end{aligned} \quad (10.9)$$

also: $\det[\cdot, \cdot, \cdot]$ ist linear in jeder Spalte.

(2) Vertauschen von 2 Spaltenvektoren ergibt einen Faktor (-1) :

$$\det[\vec{a}_2, \vec{a}_1, \vec{a}_3] = -\det[\vec{a}_1, \vec{a}_2, \vec{a}_3] \quad (10.10)$$

(3) Das Volumen des von den Basisvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ aufgespannten Einheitswürfels ist 1, also

$$\det[\vec{e}_1, \vec{e}_2, \vec{e}_3] = \det I_3 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1. \quad (10.11)$$

(4) Addieren eines Vielfachen von \vec{a}_2 oder \vec{a}_3 zu \vec{a}_1 ändert die Determinante nicht:

$$\begin{aligned} \det[\vec{a}_1 + \lambda \vec{a}_2, \vec{a}_2, \vec{a}_3] \\ = \det[\vec{a}_1, \vec{a}_2, \vec{a}_3] + \lambda \underbrace{\det[\vec{a}_2, \vec{a}_2, \vec{a}_3]}_{=0 \text{ nach (5)}}. \end{aligned} \quad (10.12)$$

Analog im 2. und 3. Eingang (in der 2. und 3. Spalte).

- (5) $\det[\vec{a}_1, \vec{a}_2, \vec{a}_3] \neq 0 \Leftrightarrow \vec{a}_1, \vec{a}_2, \vec{a}_3$ linear unabhängig, also $\det[\vec{a}_1, \vec{a}_2, \vec{a}_3] = 0$ genau dann, wenn $\vec{a}_1, \vec{a}_2, \vec{a}_3$ linear abhängig sind, also in einer Ebene liegen.

Ganz analog zum gerade besprochenen Fall $n = 3$ ergibt sich für $n = 2$:

Sei

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \in \mathbb{R}^2 \quad (10.13)$$

dann ist

$$a_{11}a_{22} - a_{12}a_{21} \quad (10.14)$$

- bis auf das Vorzeichen - der Flächeninhalt des von \vec{a}_1 und \vec{a}_2 aufgespannten Parallelogramms (vgl. Fig.10.2). Somit ist es natürlich die Determinante der Matrix

$$A = [\vec{a}_1, \vec{a}_2] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (10.15)$$

so einzuführen:

$$\det A := a_{11}a_{22} - a_{12}a_{21} \quad (10.16)$$

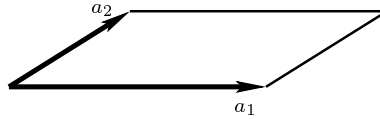


ABBILDUNG 10.2: Parallelepiped 2-dimensional

Man verifiziert dann leicht, daß die Abbildung $\det : \mathbb{R}^{2,2} \rightarrow \mathbb{R}$ Eigenschaften hat, die ganz analog sind zu den gerade besprochenen Eigenschaften (1) bis (5) sind. Schließlich sei noch

$$\begin{aligned} \det &: \mathbb{R}^{1,1} \rightarrow \mathbb{R} \\ \det[a_{11}] &= a_{11} \end{aligned} \quad (10.17)$$

also

$$\det[a_{11}] = \begin{cases} \text{Länge des Intervalls } [0, a_{11}], & a_{11} > 0 \\ - \text{Länge des Intervalls } [a_{11}, 0], & a_{11} < 0 \end{cases} \quad (10.18)$$

Damit haben wir explizite Formeln für die Determinante in den Fällen $n = 3, 2, 1$. Um dies auf beliebiges n zu verallgemeinern, benötigen wir die folgende

Definition 10.2.1 (Streichungsmatrix). Sei $n \in \mathbb{N}$, $n > 1$.

Für

$$A \in \mathbb{R}^{n,n}, \quad (10.19)$$

sei

$$S_{ij}(A) := \begin{bmatrix} a_{11} & \cdots & \overline{a_{1j}} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & \overline{a_{ij}} & \cdots & \overline{a_{in}} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \quad (10.20)$$

die $(n-1) \times (n-1)$ Matrix, die aus A durch Streichen der i -ten Zeile und j -ten Spalte hervorgeht.

Jetzt soll am Beispiel der Determinante von 3×3 Matrizen erklärt werden, weshalb Streichungsmatrizen in der Definition der Determinante eine Rolle spielen: Sei $A = [\vec{a}_1, \vec{a}_2, \vec{a}_3] \in \mathbb{R}^{3,3}$. Dann ist nach Definition (10.5) und (10.16)

$$\begin{aligned} \det A &:= \langle \vec{a}_1, \vec{a}_2 \times \vec{a}_3 \rangle \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} \underbrace{(a_{22}a_{33} - a_{23}a_{32})}_{= \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}} - a_{21} \underbrace{(a_{12}a_{33} - a_{13}a_{32})}_{= \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}} + a_{31} \underbrace{(a_{12}a_{23} - a_{13}a_{22})}_{= \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}} \\ &= a_{11} \cdot \det S_{11}(A) - a_{21} \cdot \det S_{21}(A) + a_{31} \cdot \det S_{31}(A). \end{aligned}$$

Für den Fall von 2×2 Matrizen $A = [\vec{a}_1, \vec{a}_2] \in \mathbb{R}^{2,2}$ gilt analog:

$$\begin{aligned} \det A &= a_{11}a_{22} - a_{12}a_{21} \\ &= a_{11} \det S_{11}(A) - a_{12} \det S_{12}(A) \end{aligned} \quad (10.21)$$

Damit ist also die Definition der Determinante von 3×3 -Matrizen auf die Definition der Determinante von 2×2 -Streichungs-Matrizen zurückgeführt, und diese auf die von 1×1 -Matrizen. Das Bildungsgesetz ist jetzt erkennbar, das zur rekursiven Definition von Determinanten von quadratischen Matrizen beliebiger Zeilenzahl benützt wird. Hierbei wird auch \mathbb{C} oder (irgendein anderer) Zahlkörper \mathbb{K} zugelassen.

Definition 10.2.2 (Determinante). Für $n \in \mathbb{N}$ ist

$$\det : \mathbb{K}^{n,n} \rightarrow \mathbb{K} \quad (10.22)$$

rekursiv definiert: Falls $n = 1$ ist

$$\begin{aligned} \det : \mathbb{K}^{1,1} &\rightarrow \mathbb{K} \\ a_{11} &\rightarrow \det[a_{11}] := a_{11}, \end{aligned} \quad (10.23)$$

andernfalls (d.h. für den Fall $n > 1$) ist:

$$\underbrace{\det A}_{\text{Det } n \times n \text{ Matrix}} = \sum_{i=1}^n a_{i1} (-1)^{1+i} \underbrace{\det S_{i1}(A)}_{\text{Det } (n-1) \times (n-1) \text{ Matrix}} \quad (10.24)$$

Man prüft leicht nach, dass die obige Definition der Determinante für den Fall $n = 1, 2, 3$ mit dem übereinstimmt, was wir zu Beginn des Kapitels bereits eingeführt hatten. Erst für den Fall $n = 4$ ergibt sich zum ersten mal was neues.

Bemerkung 10.2.3. Die Definition der Determinante ist so formuliert, dass sie als Algorithmus zur numerischen Berechnung von $\det A$ verwendet werden kann. Es zeigt sich allerdings, dass es oftmals einfachere Wege gibt, die auf dem sogenannten Laplace'schen Entwicklungssatz beruhen, welchen wir in Teil 9 des Satzes 10.2.5 kennen lernen werden.

Beispiel 10.2.4. Eine berühmte 4×4 Matrix ist die elektromagnetische Feldstärke. Sie ist antisymmetrisch und hat als Einträge die Komponenten der dreidimensionalen reellen Vektoren \vec{E} und \vec{B} . Sie bezeichnen die elektrische und die magnetische Feldstärke an einem Punkt im Ort zu einer festen Zeit:

$$F = \begin{bmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_1 \\ E_2 & -B_3 & 0 & B_2 \\ E_3 & B_1 & -B_2 & 0 \end{bmatrix}$$

Die obige Definition der Determinante ergibt für diesen Fall:

$$\begin{aligned} \det F &= -E_1 \begin{vmatrix} -E_1 & -E_2 & -E_3 \\ -B_3 & 0 & B_2 \\ B_1 & -B_2 & 0 \end{vmatrix} + E_2 \begin{vmatrix} -E_1 & -E_2 & -E_3 \\ 0 & B_3 & -B_1 \\ B_1 & -B_2 & 0 \end{vmatrix} - E_3 \begin{vmatrix} -E_1 & -E_2 & -E_3 \\ 0 & B_3 & -B_1 \\ -B_3 & 0 & B_2 \end{vmatrix} \\ &= \langle \vec{E}, \vec{B} \rangle^2 \end{aligned}$$

Dabei bezeichnet $\langle \vec{E}, \vec{B} \rangle$ das Standardskalarprodukt in \mathbb{R}^3 .

Es zeigt sich, daß die Eigenschaften (1) bis (5), die wir für die Determinante im Fall $n = 2$ und $n = 3$ oben aufgeschrieben haben, auch im Fall beliebiger $n \in \mathbb{N}$ gültig sind. Sie sind über vollständige Induktion beweisbar und Grundlage wichtiger Eigenschaften der Determinante, wie sie im folgenden Satz formuliert werden.

Theorem 10.2.5 (Eigenschaften der Determinante, 1. Teil).

Sei $A \in \mathbb{K}^{n,n}$ eine quadratische Matrix mit den Spaltenvektoren $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{K}^n$ und $S_{i,j}(A)$ bezeichne die i, j -te Streichungsmatrix von A . Sei ferner $\vec{c} \in \mathbb{K}^n, \lambda \in \mathbb{K}$ und $k, l \in \{1, 2, \dots, n\}$.

Dann gilt:

1. Linearität in jeder Spalte:

$$\det(\vec{a}_1, \dots, (\vec{a}_k + \lambda \vec{c}), \dots, \vec{a}_n) = \det(\vec{a}_1, \dots, \vec{a}_k, \dots, \vec{a}_n) + \lambda \det(\vec{a}_1, \dots, \vec{c}, \dots, \vec{a}_n)$$

2. Antisymmetrie: Die Determinante einer Matrix ändert das Vorzeichen beim Vertauschen zweier Spalten.

3. Normierung: $\det I_n = 1$

4. $k \neq l \Rightarrow \det(\vec{a}_1, \dots, (\vec{a}_k + \lambda \vec{a}_l), \dots, \vec{a}_n) = \det(\vec{a}_1, \dots, \vec{a}_k, \dots, \vec{a}_n)$

5.

$$\begin{aligned} \det A \neq 0 &\Leftrightarrow \text{Die Spaltenvektoren sind linear unabhängig} \\ &\Leftrightarrow \text{Die Zeilenvektoren sind linear unabhängig} \\ &\Leftrightarrow \text{Rang } A = n \\ &\Leftrightarrow A^{-1} \text{ existiert} \end{aligned}$$

$$\Leftrightarrow A^{-1} = \frac{1}{\det A} \begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \\ A_{n,1} & \cdots & A_{n,n} \end{bmatrix}^T$$

Dabei wurde die Abkürzung $A_{i,k} := (-1)^{i+k} \det S_{i,k}(A)$ verwendet.³

³Die Matrix mit diesen Einträgen wird auch als "Adjunkte" von A bezeichnet.

6.

$$\begin{aligned} \det A = 0 &\Leftrightarrow \text{es existiert ein } \vec{x} \in \mathbb{K}^n \text{ mit } \vec{x} \neq 0 \text{ und } A\vec{x} = 0 \\ &\Leftrightarrow \text{Rang } A < n \end{aligned}$$

7.

$$\det A = \det A^T$$

8. Sei $\mathbb{K} = \mathbb{C}$ dann ist $\det \bar{A} = \overline{\det A}$.

9. (Laplace'scher Entwicklungssatz)

$$\begin{aligned} \det A &= \sum_{k=1}^n a_{ik} (-1)^{i+k} \det S_{ik}(A) \quad \text{für alle } i \in \{1, \dots, n\} \quad (\text{Entw. nach Zeilen}) \\ &= \sum_{i=1}^n a_{ik} (-1)^{i+k} \det S_{ik}(A) \quad \text{für alle } k \in \{1, \dots, n\} \quad (\text{Entw. nach Spalten}) \end{aligned}$$

10. $\frac{d}{dt} \det(I_n + tA)|_{t=0} = \text{Spur}(A)$ mit $\text{Spur}(A) := \sum_{i=1}^n a_{ii}$

Die obige Formel für die Inverse einer Matrix ist nur für theoretische Zwecke geeignet. In der Praxis wäre eine Berechnung der Inversen viel zu teuer auf diese Weise, dazu verwendet man besser den Gaußalgorithmus.

Algorithmus 10.2.6 (Determinantenberechnung). :*Input:* $A \in \mathbb{K}^{n,n}$ *Output:* $\det A$ *Durchführung*

1. Schritt: Wähle diejenige Zeile oder Spalte aus, die am meisten Nullen enthält.
2. Schritt: Verwende den Laplaceschen Entwicklungssatz um die Berechnung von $\det A$ auf die Berechnung von Determinanten von Matrizen aus $\mathbb{K}^{n-1,n-1}$ zurückzuführen.
3. Gehe zu 1. falls es sich nicht bereits um die Berechnung der Determinante von 2×2 Matrizen handelt, die von Hand ausgerechnet werden können.

Es ist oft sehr nützlich zu wissen, welche Operationen mit einer Matrix durchgeführt werden können, ohne dass sich dabei der Wert der Determinante ändert oder sich der Wert nur um ein Vorzeichen ändert. Dazu sagen die Aussagen (2),(4) und (7) des obigen Theorems was aus, die wir in der folgenden Bemerkung in geeigneter Form noch einmal zusammenfassen:

Folgerung 10.2.7. (Invarianz der Determinante)Sei $n \in \mathbb{N}$

- (1) Die Determinante einer Matrix $A \in \mathbb{K}^{n,n}$ ändert sich nicht beim
 - (i) Addieren eines Vielfachen einer Spalte zu einer anderen.
 - (ii) Addieren eines Vielfachen einer Zeile zu einer anderen.
 - (iii) Transponieren.
- (2) Die Determinante einer Matrix $A \in \mathbb{K}^{n,n}$ verändert sich um den Faktor (-1) beim

(i) Vertauschen zweier Spalten.

(ii) Vertauschen zweier Zeilen.

Im Folgenden betrachten wir die Determinante eines Produktes zweier quadratischer Matrizen und machen plausibel, dass dies gleich dem Produkt der Determinanten der einzelnen Faktoren ist. Dazu schauen wir uns zuerst einmal die Determinanten von Dreiecksmatrizen an.

Sind $A, B \in \mathbb{K}^{n,n}$ zwei obere Dreiecksmatrizen, dann ist auch das Produkt eine obere Dreiecksmatrix:

$$\begin{aligned}
 A \cdot B &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ 0 & \ddots & \vdots \\ 0 & 0 & b_{n,n} \end{bmatrix} \\
 &= \begin{bmatrix} a_{1,1} \cdot b_{1,1} & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & a_{n,n} \cdot b_{n,n} \end{bmatrix} \quad (10.25)
 \end{aligned}$$

Man erhält so für die Determinante des Produktes der beiden Dreiecksmatrizen:

$$\begin{aligned}
 \det(AB) &= a_{11} \cdot b_{11} \cdot a_{22} \cdot b_{22} \cdot \dots \cdot a_{nn} \cdot b_{nn} \\
 &= a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn} \cdot b_{11} \cdot b_{22} \cdot \dots \cdot b_{nn} \\
 &= \det A \cdot \det B \quad (10.26)
 \end{aligned}$$

Dass Determinanten von Dreiecksmatrizen einfach zu berechnen sind, haben wir gerade eingesehen. Das kann dazu verwendet werden, die Determinante eines Produktes von beliebigen quadratischen Matrizen zu berechnen: Mit dem Gaußalgorithmus kann eine beliebige quadratische Matrix systematisch zu einer Dreiecksmatrix transformiert werden, wobei sich der Wert der Determinante höchstens um ein Vorzeichen oder einen bekannten Faktor ändert.⁴

Theorem 10.2.8. (Eigenschaften von Determinanten, Teil 2) Sei $n \in \mathbb{N}$ und $A, B \in \mathbb{K}^{n,n}$. Dann gilt

(1) $\det(AB) = \det A \cdot \det B$

(2) Ist A invertierbar, dann ist

$$\det(A^{-1}) = (\det A)^{-1} \quad (10.27)$$

(3) Ist A invertierbar, dann ist

$$\det B = \det(ABA^{-1}) \quad (10.28)$$

Es ist klar, daß (2) aus (1) folgt. Denn

$$\begin{aligned}
 1 &= \det I_n = \det(AA^{-1}) \\
 &= \det A \cdot \det(A^{-1}) \quad (10.29)
 \end{aligned}$$

(3) erhält man dann mit (1) und (2).

⁴Die Zeilenstufenform kann durch Multiplikation von links mit Elementarmatrizen (vgl. Anhang) erreicht haben. Ferner ist $\det(P_{i,j}) = (-1)$, $\det(G_{i,j}) = 1$ und $\det(M_i(\lambda)) = \lambda$.

Bemerkung 10.2.9. Es ist früher gezeigt worden, dass mit dem Gaußalgorithmus quadratische Matrizen in Dreiecksmatrizen umgeformt werden können, und dies geht durch Multiplikation mit den Elementarmatrizen (siehe Anhang), deren Determinante wir kennen.

Da sich bei diesen Umformungen der Wert der Determinanten höchstens um Faktoren (-1) ändert, der Wert der Determinante einer Dreiecksmatrix, wie im Beispiel oben gezeigt, gleich dem Produkt der Diagonalelemente ist, gibt der Gaußalgorithmus - zusammen mit dem vorangehenden Beispiel - also eine alternative Methode zur Berechnung von Determinanten.

Bemerkung 10.2.10 (Determinante einer linearen Abbildung). Eine spezielle Anwendung der obigen Bemerkung und Punkt 7. ist folgende bemerkenswerte Tatsache: Sei L eine lineare Abbildung eines Vektorraums über \mathbb{K} in sich. Dann gilt für zwei darstellende Matrizen in zwei Basen die Beziehung

$$L_{\mathcal{B}_2} = S L_{\mathcal{B}_1} S^{-1} \quad (10.30)$$

$$S = K_{\mathcal{B}_2} (K_{\mathcal{B}_1})^{-1} \quad (10.31)$$

Somit sind die Determinanten der beiden darstellenden Matrizen gleich.

$$\det L_{\mathcal{B}_2} = \det L_{\mathcal{B}_1} \quad (10.32)$$

Damit ist der Wert der Determinante eine Eigenschaft der linearen Abbildung und von der Darstellung durch eine Matrix unabhängig.

11. KAPITEL

EIGENWERTE, EIGENVEKTOREN UND CHARAKTERISTISCHES POLYNOM

Eigenvektoren und Eigenwerte linearer Abbildungen treten in fast allen Gebieten der Mathematik und ihrer Anwendungen auf. Von besonderer Bedeutung sind sie für die Theorie der *Linearen Differentialgleichungen*, die wir im letzten Teil dieser Vorlesung diskutieren.

Als Beispiel betrachten wir die Exponentialfunktion. Sie ist Eigenvektor der linearen Abbildung D , wie wir im Folgenden erklären: Sei V der Vektorraum der auf \mathbb{R} definierten und beliebig oft differenzierbaren Funktionen und D die Abbildung

$$D : \begin{cases} V & \rightarrow V \\ f & \mapsto Df = f' \end{cases} .$$

Für ein $\alpha \in \mathbb{C}$ definieren wir die Exponentialfunktion

$$f_\alpha(x) := e^{\alpha \cdot x}, \quad x \in \mathbb{R},$$

Dann gilt:

$$Df_\alpha = \alpha \cdot f_\alpha,$$

d.h. die Funktion f_α wird durch D auf ein Vielfaches von sich selbst abgebildet. Sie heißt *Eigenvektor* von D . Der Wert α ist der dazugehörige *Eigenwert*.

Dieses und ähnliche Beispiele treten bei elektrischen und mechanischen Schwingungen auf, wie der mathematischen Modellierung einer Brücke oder der Saiten eines Musikinstrument. Im ersten Fall werden wir darauf achten, dass keine Eigenschwingungen angeregt werden, damit die Brücke nicht einstürzt; im zweiten Fall wird die Eigenschwingung mit Absicht angeregt, z.B. mit dem Bogen der Violine. Mathematisch eng verwandt damit ist die Modellierung der Farben unserer Umwelt. Auch sie werden durch Eigenwerte einer linearen Abbildung beschrieben die in der Quantenmechanik auftritt (*Erwin Schrödinger* 1927). Wichtig sind Eigenwerte und Eigenvektoren auch für die Datenanalyse und Datenkompression. Ein Stichwort ist hier die *Markovkette*, ein Konzept der Wahrscheinlichkeitstheorie.

Für den Fall endlichdimensionaler Vektorräume werden wir ein Verfahren kennen lernen, wie Eigenwerte systematisch berechnet werden können. Für den Fall von Vektorräumen mit Dimension unendlich ist dies viel schwieriger und gehört nicht zu den Themen der Linearen Algebra.

Zur systematischen Berechnung von Eigenwerten verwenden wir die Methode des charakteristischen Polynoms. Es ist vorerst einmal nur für quadratische $n \times$

n -Matrizen – oder was dasselbe ist, für Abbildungen von \mathbb{K}^n nach \mathbb{K}^n – definiert (vgl. Def. 11.2.1). Das Konzept kann jedoch auf lineare Abbildungen von V nach V ausgedehnt werden, wobei V einen endlichdimensionalen Vektorraum bezeichnet (vgl. Def. 11.2.4).

Wir werden uns hier und in den folgenden Vorlesungen auf den Fall komplexer Vektorräume konzentrieren, denn die Berechnung und Diskussion von Eigenwerten und Eigenvektoren für Vektorräume über \mathbb{C} ist einfacher als der Fall $\mathbb{K} = \mathbb{R}$.

Dies bedeutet kein Verlust für die praktische Anwendung, da ja jeder reelle Vektorraum als Teilmenge eines komplexen Vektorraums und jede reelle lineare Abbildung als ein Spezialfall einer komplexen aufgefasst werden kann; typisches Beispiel ist der $\mathbb{R}^n \subset \mathbb{C}^n$ und eine Abbildung, die durch eine reelle Matrix beschrieben wird.

Weshalb Theorie und praktisches Rechnen für Vektorräume über den komplexen Zahlen einfacher ist als im reellen, wird später klar werden, wenn wir den Zusammenhang von Eigenwerten und Eigenvektoren mit Nullstellen des charakteristischen Polynoms besprechen.

Die Strategie der systematischen Berechnung der Eigenwerte für endlichdimensionale Vektorräume ist folgende: Wir wissen wie algebraisch entschieden werden kann, ob der Kern einer linearen Abbildung nur aus dem Nullvektor besteht oder nicht: Es gibt einen vom Nullvektor verschiedenen Vektor im Kern genau dann, wenn die Determinante verschwindet. Nun stellen wir die Eigenwertgleichung

$$A\vec{v} = \lambda\vec{v}$$

um und schreiben

$$A\vec{v} - \lambda\vec{v} = \vec{0},$$

oder was das damit gleichbedeutend ist:

$$(A - \lambda \cdot I_{\dim V})\vec{v} = \vec{0}$$

Damit geht es jetzt um den Kern der quadratischen Matrix $A - \lambda \cdot I_{\dim V}$. Er enthält einen von Null verschiedenen Vektor genau dann, wenn die Determinante von $A - \lambda \cdot I_{\dim V}$ verschwindet oder anders ausgedrückt, wenn das Polynom $p_A(z) := \det(A - z \cdot I_{\dim V})$ in der Variablen z an der Stelle $z = \lambda$ eine Nullstelle hat. $p_A(z)$ heißt *charakteristisches Polynom* von A . Das Verfahren, wie die Eigenwerte einer linearen Abbildung zu berechnen sind, ist in Theorem 11.3.1 und dem anschließenden Algorithmus 11.3.4 formuliert.

Die Methode kann auf allgemeine lineare Abbildungen eines endlichdimensionalen Vektorraums verallgemeinert werden, da jede solche Abbildung durch eine Matrix dargestellt werden kann. Das dabei auftretende Konsistenzproblem – die Matrixdarstellung ist nicht eindeutig – wird in Theorem 11.2.2 gelöst.

11.1 DEFINITION, BEISPIELE UND GRUNDLEGENDE EIGENSCHAFTEN

In dem einleitenden Beispiel haben wir gesehen, wie eine lineare Abbildung auf einen Vektor so wirkt, wie ein Vielfaches der Identität. Solche Vektoren sind besonders nützlich und haben deshalb einen Namen:

Definition 11.1.1. Sei V ein Vektorraum über dem Körper \mathbb{C} , A eine lineare Abbildung von V nach V und $\lambda \in \mathbb{C}$. λ heißt *Eigenwert* von A , falls es einen Vektor $\vec{v} \in V$ gibt, der vom Nullvektor verschieden ist und für den gilt

$$A\vec{v} = \lambda\vec{v}.$$

Der Vektor \vec{v} heißt *Eigenvektor* von A zum Eigenwert λ .

Analog dazu definiert man den Begriff Eigenwert und Eigenvektor für den Fall von Vektorräumen über den reellen Zahlen:

Definition 11.1.2. Sei V ein Vektorraum über dem Körper \mathbb{R} , A eine lineare Abbildung von V nach V und $\lambda \in \mathbb{R}$. λ heißt *Eigenwert* von A , falls es einen Vektor $\vec{v} \in V$ gibt, der vom Nullvektor verschieden ist und für den

$$A\vec{v} = \lambda\vec{v}$$

gilt. Der Vektor \vec{v} heißt *Eigenvektor* von A zum Eigenwert λ .

Verwandte Begriffe sind *Eigenraum* und *geometrische Vielfachheit* (vgl. Definition (11.1.15)).

Bemerkung 11.1.3. Wenn \vec{v} Eigenvektor von A zum Eigenwert λ ist, dann ist für alle $\alpha \in \mathbb{C} \setminus \{0\}$ auch $\alpha\vec{v}$ Eigenvektor von A zum *selben* Eigenwert λ . Der Eigenvektor \vec{v} spannt also einen 1-dimensionalen Teilraum $\text{span}\{\vec{v}\}$ (vgl. Def. 2.2.7) auf. Darin sind alle Vektoren bis auf den Nullvektor Eigenvektoren zum selben Eigenwert λ .

Bemerkung 11.1.4. Jeder von $\vec{0}$ verschiedene Vektor des Kerns von A ist Eigenvektor von A zum Eigenwert $\lambda = 0$.

Nun kommen wir zu einer Reihe von Beispielen:

Beispiel 11.1.5. Sei $V = \mathbb{C}^2$ und $\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. Dann ist $\begin{bmatrix} 1 \\ i \end{bmatrix}$ Eigenvektor von σ_2 zum Eigenwert $\lambda = 1$ und $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ Eigenvektor zum Eigenwert $\mu = -1$.

Beispiel 11.1.6. Sei $V = \mathbb{C}^2$ und $\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Dann ist $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ Eigenvektor von σ_3 zum Eigenwert $\lambda = 1$ und $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ zum Eigenwert $\mu = -1$.

Beispiel 11.1.7. Sei $V = \mathbb{R}^2$ und $A = \frac{1}{6} \begin{bmatrix} 15 & -9 \\ -1 & 15 \end{bmatrix}$.

Dann ist $\vec{v}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ (normierter) Eigenvektor zum Eigenwert $\lambda_1 = 2$ und $\vec{v}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ zum Eigenwert $\lambda_2 = 3$. Für die Eigenvektoren bewirkt A lediglich eine Skalierung mit dem zugehörigen Eigenwert als Skalierungsfaktor. Andere Vektoren werden jedoch skaliert und zusätzlich gedreht.

Beispiel 11.1.8. Sei $V = \mathbb{R}^2$ und $A = \frac{1}{2} \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$.

Man beachte, dass A *symmetrisch* ist. Dann ist $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (normierter) Eigenvektor zum Eigenwert $\lambda_1 = 2$ und $\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ zum Eigenwert $\lambda_2 = 3$. Für die Eigenvektoren bewirkt A lediglich eine Skalierung mit dem zugehörigen Eigenwert als Skalierungsfaktor. Andere Vektoren werden jedoch skaliert und zusätzlich gedreht. Man beachte, dass die (betragsmäßig) größte bzw. (betragsmäßig) kleinste Skalierung in Richtung der Eigenvektoren zum (betragsmäßig) größten bzw. (betragsmäßig) kleinsten Eigenwert erfolgt.

Beispiel 11.1.9. Jetzt betrachten wir den Vektorraum der Polynome vom Grad kleiner oder gleich n , $\mathbb{R}_{\leq n}[x]$, und die lineare Abbildung

$$\sigma : \begin{cases} \mathbb{R}_{\leq n}[x] & \rightarrow \mathbb{R}_{\leq n}[x] \\ p(x) = a_n x^n + \dots + a_1 x + a_0 & \mapsto q(x) = x \cdot \frac{dp}{dx}(x) \end{cases}$$

Dann ist das Monom x^k Eigenvektor zum Eigenwert $k \in \{0, 1, \dots, n\}$.

Beispiel 11.1.10. Jetzt betrachten wir die trigonometrischen Polynome vom Grad n , $\text{Trig}_{\leq n}[x]$, wie sie in Beispiel 8.2.3 eingeführt wurden und darauf als lineare Abbildung die zweite Ableitung:

$$\Delta : \begin{cases} \text{Trig}_{\leq n}[x] & \rightarrow \text{Trig}_{\leq n}[x] \\ p(x) := a_0 + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)) & \mapsto p''(x). \end{cases}$$

Dann sind $\cos(kx)$ und $\sin(kx)$, $k \in \{0, 1, \dots, n\}$ Eigenvektoren von Δ zum Eigenwert $-k^2$.

Beispiel 11.1.11. Wir betrachten die Situation von Beispiel 6.1.4 und die Abbildung K . Sei $a \in \mathbb{R} \setminus \{0\}$, dann ist $\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$ ein Eigenvektor von K zum Eigenwert 1.

Falls $b, c \in \mathbb{R}$ und nicht beide gleich Null sind, dann ist $\begin{bmatrix} 0 & b \\ c & -(b+c) \end{bmatrix}$ ein Eigenvektor von K zum Eigenwert Null.

Beispiel 11.1.12. Sei $V = \mathbb{C}^2$ und $A = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Dann ist $\vec{e}_+ := \begin{bmatrix} 1 \\ -i \end{bmatrix}$ ein Eigenvektor von A zum Eigenwert $\lambda_+ := e^{i\pi/4}$. Analog ist $\vec{e}_- := \begin{bmatrix} 1 \\ i \end{bmatrix}$ ein Eigenvektor von A zum Eigenwert $\lambda_- := e^{-i\pi/4}$.

Beispiel 11.1.13. Sei $V = \mathbb{R}^3$ und $P := \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Dann sind $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ und $\vec{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ Eigenvektoren von P zum Eigenwert 1. P hat ferner den

Eigenwert Null, denn $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ ist im Kern von P .

Beispiel 11.1.14. Sei V Vektorraum über \mathbb{K} . Die Abbildung A , die jeden Vektor auf den Nullvektor $\vec{0}$ abbildet, hat jedes $\vec{v} \in V \setminus \{\vec{0}\}$ als Eigenvektor mit Eigenwert Null.

Die Beispiele 11.1.10, 11.1.11 und 11.1.13 zeigen, dass es zu einem Eigenwert λ durchaus mehrere linear unabhängige Eigenvektoren geben kann. Dies gibt Anlass zu der

Definition 11.1.15 (Eigenraum, geometrische Vielfachheit). Sei V ein Vektorraum über dem Körper \mathbb{C} , A eine lineare Abbildung von V nach V und λ ein Eigenwert von A . Die Teilmenge $V_\lambda := \{\vec{v} | \vec{v} \in V, A\vec{v} = \lambda\vec{v}\}$ von V heißt *Eigenraum* von A zum Eigenwert λ . Die Dimension des Eigenraums ist die *geometrische Vielfachheit* des Eigenwertes.

Ein damit verwandter Begriff ist die *algebraische Vielfachheit*, die wir weiter unten kennen lernen werden (vgl. Definition 11.3.6).

Bemerkung 11.1.16. Die geometrische Vielfachheit eines Eigenwertes λ ist also die maximale Anzahl linear unabhängiger Eigenvektoren zu diesem Eigenwert.

Bemerkung 11.1.17. Es gilt offensichtlich $\text{Kern}(A) = V_{\lambda=0}$.

Bemerkung 11.1.18. In der folgenden Tabelle fassen wir die Ergebnisse verschiedener vorstehender Beispiele zusammen.

Beispiel	Eigenwert	geometrische Vielfachheit
11.1.10	$-k^2$	2
11.1.13	1	2
11.1.13	0	1

Bemerkung 11.1.19. Nach Konstruktion ist die Einschränkung von A auf einen Eigenraum V_λ gerade so gemacht, dass gilt: $A|_{V_\lambda} = \lambda \cdot \text{Identität}_{V_\lambda}$. Auf Eigenräumen ist A also besonders einfach zu berechnen. Dies eröffnet noch einmal eine nützliche Veranschaulichung des Konzeptes *Eigenvektor* und *Eigenwert*. Eine lineare Abbildung A kann als eine Abbildung verstanden werden, die auf den von Eigenvektoren aufgespannten Teilräumen eine *Skalierung*, d.h. Streckung um einen Faktor, nämlich um den Eigenwert, darstellt. Falls die lineare Abbildung eine Basis von Eigenvektoren besitzt, kann sie auf diese Weise vollständig beschrieben werden.

Jetzt betrachten wir einen wichtigen Spezialfall, der nicht direkt in den bis jetzt geschilderten Rahmen passt, den Fall einer orthogonalen Abbildung. Wir werden etwas später sehen, wie dieser Fall mit dem zuvor behandelten zusammenhängt. Wir erinnern uns: Eine orthogonale Abbildung wirkt auf einen euklidischen Vektorraum, d.h. auf einen Vektorraum über \mathbb{R} mit Skalarprodukt. Sei also Q eine orthogonale Transformation des euklidischen Vektorraums V und $\lambda \in \mathbb{R}$ Eigenwert von Q mit Eigenvektor \vec{v} . Achtung, λ muss reell sein, da sonst die Eigenwertgleichung

$$Q\vec{v} = \lambda\vec{v}$$

gar nicht erst geschrieben werden kann. Da Q orthogonal ist gilt die Gleichung

$$\begin{aligned} \langle Q\vec{v}, Q\vec{v} \rangle &= \lambda^2 \langle \vec{v}, \vec{v} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle. \end{aligned}$$

Da $\langle \vec{v}, \vec{v} \rangle \neq 0$ ist, muss $\lambda^2 = 1$ sein, d.h. λ ist gleich 1 oder gleich -1 .

Für den Fall unitärer Abbildungen ist die Situation ähnlich wie für orthogonale Transformationen. Sei also U eine unitäre Abbildung des unitären Raums V und $\lambda \in \mathbb{C}$ ein Eigenwert von U mit Eigenvektor $\vec{v} \in V$. Dann ist

$$\begin{aligned} \langle U\vec{v}, U\vec{v} \rangle &= |\lambda|^2 \langle \vec{v}, \vec{v} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle. \end{aligned}$$

Da $\langle \vec{v}, \vec{v} \rangle \neq 0$ ist, muss $|\lambda| = 1$ sein, d.h. λ liegt auf dem Einheitskreis in \mathbb{C} .

Nun kommen wir auf die orthogonale Transformation euklidischer Vektorräume und deren Einordnung in die hier besprochene Theorie der Eigenwerte und Eigenvektoren von Transformationen komplexer Vektorräume zurück. Es ist einfach einzusehen, dass ein reeller Vektorraum als Teilmenge (reeller Teilraum) eines komplexen Vektorraums aufgefasst werden kann. Es gilt sogar, dass jeder euklidische Vektorraum als Teilmenge eines unitären aufgefasst werden kann, wobei dann das euklidische Skalarprodukt die Einschränkung des unitären auf

die reelle Teilmenge darstellt. Wir kennen dies von dem Fall $\mathbb{R}^N \subset \mathbb{C}^N$ und den reellen Matrizen, die eine Teilmenge der komplexen darstellen $\mathbb{R}^{N,N} \subset \mathbb{C}^{N,N}$ und den dazugehörigen Standardskalarprodukten.

Falls die Abbildung A des unitären Raums V nach V selbstadjungiert ist (vgl. 3.2.13), sind die Eigenwerte von A reell, denn es gilt:

$$\begin{aligned}\langle \vec{v}, A\vec{v} \rangle &= \lambda \langle \vec{v}, \vec{v} \rangle \\ &= \langle A\vec{v}, \vec{v} \rangle \\ &= \overline{\lambda} \langle \vec{v}, \vec{v} \rangle.\end{aligned}$$

Eine nützliche Eigenschaft von Eigenvektoren ist folgendes

Theorem 11.1.20. *Seien \vec{v}_1, \vec{v}_2 Eigenvektoren von A mit Eigenwerten λ_1 und λ_2 . Dann gilt die Aussage:*

$$\lambda_1 \neq \lambda_2 \implies \vec{v}_1 \text{ und } \vec{v}_2 \text{ sind linear unabhängig.}$$

Falls zusätzlich A selbstadjungiert oder unitär ist, sind die Eigenvektoren nicht nur linear unabhängig sondern stehen senkrecht aufeinander.

Beweis. *Der Beweis der ersten Aussage ist offensichtlich, wenn man verwendet, dass die obige Aussage zur folgenden äquivalent ist:*

$$\vec{v}_1 \text{ und } \vec{v}_2 \text{ sind linear abhängig} \implies \lambda_1 = \lambda_2.$$

Beispiel 11.1.21. Wir betrachten noch einmal Beispiel ?? . σ_1 hat die Eigenvektoren $\vec{e}_+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{e}_- = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ zu den Eigenwerten $\lambda_+ = +1$ bzw. $\lambda_- = -1$. Somit sind \vec{e}_+ und \vec{e}_- linear unabhängig und bilden eine Basis. σ_1 besitzt also eine Basis aus Eigenvektoren.

Von obigem Satz 11.1.20 gilt folgende Verallgemeinerung:

Theorem 11.1.22. *Sei k eine natürliche Zahl und $\vec{v}_1, \dots, \vec{v}_k$ Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_k$. Dann gilt:*

$$\lambda_1, \dots, \lambda_k \text{ paarweise verschieden} \implies \vec{v}_1, \dots, \vec{v}_k \text{ linear unabhängig.}$$

Beweis. *Die Aussage wird mit Hilfe vollständiger Induktion gezeigt. Der letzte Satz 11.1.20 ist gerade der Induktionsanfang, sodass wir hier nur noch den Induktionsschritt $j \rightarrow j+1$ ausführen müssen:*

Die Aussage des Satzes sei wahr für ein $j \in \{1, \dots, k-1\}$. Jetzt zeigen wir, dass sie dann auch für $j+1$ richtig ist. Wiederum ist es einfacher, die äquivalente Aussage

$$\vec{v}_1, \dots, \vec{v}_{j+1} \text{ linear abhängig} \implies \text{mindestens zwei Eigenwerte sind gleich}$$

zu zeigen.

O.B.d.A. nehmen wir an, dass

$$\vec{v}_1 = \sum_{i=2}^{j+1} \alpha_i \vec{v}_i.$$

Dann ist

$$\begin{aligned} A\vec{v}_1 &= \lambda_1 \sum_{i=2}^{j+1} \alpha_i \vec{v}_i \\ &= A \sum_{i=2}^{j+1} \alpha_i \vec{v}_i \\ &= \sum_{i=2}^{j+1} \alpha_i \lambda_i \vec{v}_i. \end{aligned}$$

Daraus folgt

$$\sum_{i=2}^{j+1} (\lambda_i - \lambda_1) \alpha_i \vec{v}_i = \vec{0}.$$

Entweder sind alle Koeffizienten $(\lambda_2 - \lambda_1)\alpha_2, \dots, (\lambda_{j+1} - \lambda_1)\alpha_{j+1}$ Null oder sie sind es nicht. Im ersten Fall gibt es ein $\lambda_l \in \{\lambda_2, \dots, \lambda_{j+1}\}$ so, dass $\lambda_l = \lambda_1$, denn nicht alle α_i können verschwinden (sonst wäre $\vec{v}_1 = \vec{0}$). Andernfalls sind $\vec{v}_2, \dots, \vec{v}_{j+1}$ linear abhängig, und es gibt ein Paar in $\{\lambda_2, \dots, \lambda_{j+1}\}$, das gleiche Werte hat (Induktionsvoraussetzung).

Aus dem obigen Satz folgt unmittelbar die

Folgerung 11.1.23. Eine lineare Abbildung A auf einem (endlichdimensionalen) Vektorraum V hat höchstens so viele Eigenwerte wie $\dim V$.

Um die Eigenwerte und Eigenvektoren einer linearen Abbildung systematisch zu berechnen, benötigt man eine Technik, die wir jetzt erklären (vgl. Algorithmus 11.3.4).

11.2 DEFINITION UND ELEMENTARE EIGENSCHAFTEN CHARAKTERISTISCHER POLYNOME

Als erstes betrachten wir eine lineare Abbildung von \mathbb{C}^n nach \mathbb{C}^n , d.h. eine quadratische $n \times n$ - Matrix und definieren das zugehörige charakteristische Polynom wie folgt:

Definition 11.2.1. Sei A aus $\mathbb{C}^{n,n}$. Dann heißt das Polynom in der Variablen $z \in \mathbb{C}$

$$p_A(z) := \det(A - z \cdot I_n)$$

das *charakteristische Polynom* von A .

Bemerkung 11.2.2. Das charakteristische Polynom von $A \in \mathbb{C}^{n,n}$ ist vom Grad n . Beweisen lässt sich dies zum Beispiel durch Induktion nach n .

Beispiel 11.2.3. Sei $A := \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ eine lineare Abbildung von $V := \mathbb{C}^2$ nach V . Dann ist das zugehörige charakteristische Polynom

$$p_A(z) = \det(A - z \cdot I_2) = (1 - z)(3 - z).$$

Das charakteristische Polynom hat folgende Invarianzeigenschaften:

Theorem 11.2.1.

1. Sei $A \in \mathbb{C}^{n,n}$ und A^T die transponierte Matrix (3.2.9). Dann gilt $p_A(z) = p_{A^T}(z)$.
2. Seien $A, S \in \mathbb{C}^{n,n}$ und sei S invertierbar. Dann gilt $p_A(z) = p_{SAS^{-1}}(z)$.

Beweis. Beide Behauptungen folgen unmittelbar aus Theorem 10.2.8 über die Invarianz der Determinante.

Theorem 11.2.2 (Folgerung aus Theorem 11.2.1). Sei L eine lineare Abbildung eines endlichdimensionalen Vektorraums V nach V und $L_{\mathcal{B}_1}$ und $L_{\mathcal{B}_2}$ darstellende Matrizen von A zu zwei verschiedenen Basen. Dann ist $p_{L_{\mathcal{B}_1}}(z) = p_{L_{\mathcal{B}_2}}(z)$.

Beweis. Da es ein invertierbares $S \in \mathbb{C}^{n,n}$ gibt (siehe Abschnitt 7.2), so, dass $L_{\mathcal{B}_2} = SL_{\mathcal{B}_1}S^{-1}$, folgt die Behauptung unmittelbar aus dem obigen Satz 11.2.1.

Damit kann der Begriff *charakteristisches Polynom* auf lineare Abbildungen ausgedehnt werden.

Definition 11.2.4. Sei V ein endlichdimensionaler Vektorraum und L eine lineare Abbildung von V nach V . Sei ferner $L_{\mathcal{B}}$ eine Matrixdarstellung von L . Dann heißt folgendes Polynom das *charakteristische Polynom* von L :

$$p_L(z) := \det(L_{\mathcal{B}} - z \cdot I_{\dim V}).$$

Bemerkung 11.2.5. Zur obigen Definition 11.2.4 gehört die Wahl einer darstellenden Matrix. Trotz der Freiheit dieser Wahl ist das charakteristische Polynom eindeutig definiert und von der Wahl der darstellenden Matrix unabhängig (vgl. Theorem 11.2.2).

Beispiel 11.2.6. Sei $V = \mathbb{C}_{\leq 2}[x]$ und D die Ableitung (vgl. Beispiel 6.1.6). Jetzt verwenden wir in $\mathbb{C}_{\leq 2}[x]$ die Basis der Monome $\mathcal{B} = \{\vec{m}_0 = 1, \vec{m}_1 = x, \vec{m}_2 = x^2\}$. Nach Algorithmus 7.3.2 ist jetzt $D\vec{m}_i$ zu berechnen:

$$D\vec{m}_0 = 0, \quad D\vec{m}_1 = \vec{m}_0, \quad D\vec{m}_2 = 2\vec{m}_1.$$

Für die darstellende Matrix erhalten wir:

$$D_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Daraus berechnet sich das charakteristische Polynom zu

$$p_D(z) = \det \begin{bmatrix} -z & 1 & 0 \\ 0 & -z & 2 \\ 0 & 0 & -z \end{bmatrix} = -z^3.$$

11.3 BERECHNUNG DER EIGENWERTE ALS NULLSTELLEN DES CHARAKTERISTISCHEN POLYNOMS

Die hier dargestellte Methode zur Berechnung von Eigenwerten geht über die Bestimmung der Nullstellen des *charakteristischen Polynoms*. Sie ist die konzeptionell einfachste und klarste. Aus der Sicht der numerischen Bestimmung von Eigenwerten ist dies keine gute Wahl.

Theorem 11.3.1.

Sei $A \in \mathbb{C}^{n,n}$ eine lineare Abbildung von \mathbb{C}^n nach \mathbb{C}^n und $\lambda \in \mathbb{C}$. Dann gilt:

Es existiert ein Eigenvektor \vec{v} zum Eigenwert $\lambda \iff \det(A - \lambda I_n) = 0$

Beweis. Der Satz ist eine unmittelbare Folge aus der Aussage 6 in Theorem 10.2.5.

Bemerkung 11.3.1. Eine zum obigen Satz analoge Aussage gilt für den Fall des reellen statt des komplexen Zahlkörpers: Sei $A \in \mathbb{R}^{n,n}$ eine lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^n und $\lambda \in \mathbb{R}$. Dann gilt:

Es existiert ein Eigenvektor \vec{v} zum Eigenwert $\lambda \iff \det(A - \lambda I_n) = 0$

Folgerung 11.3.2. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} der Dimension $n \in \mathbb{N}$. Weiter sei L eine lineare Abbildung von V nach V , \mathcal{B} eine Basis von V und $\lambda \in \mathbb{K}$. Dann gilt:

Es existiert ein Eigenvektor \vec{v} zum Eigenwert $\lambda \iff \det(L_{\mathcal{B}} - \lambda I_n) = 0$

Beweis. In der 7. Vorlesung haben wir erklärt, wie von einem allgemeinen Vektorraum endlicher Dimension auf den Vektorraum $\mathbb{K}^{\dim V}$ der Koordinatenvektoren übergegangen werden kann, sofern eine Basis von V vorhanden ist. Damit kann gezeigt werden, dass $\vec{v} \in V$ genau dann Eigenvektor von L zum Eigenwert λ ist, wenn der Koordinatenvektor $\vec{v}_{\mathcal{B}} \in \mathbb{K}^{\dim V}$ Eigenvektor der Matrix $L_{\mathcal{B}}$ zum Eigenwert λ ist. Damit ist die Behauptung in der Folgerung auf den vorausgehenden Satz zurückgeführt.

Bemerkung 11.3.3. Nach dem Fundamentalsatz der Algebra hat das charakteristische Polynom *mindestens* eine *komplexe Nullstelle*. Somit hat eine lineare Abbildung eines endlichdimensionalen Vektorraumes über \mathbb{C} immer mindestens einen Eigenwert λ und einen zugehörigen Eigenvektor. Dies ist nicht so für Vektorräume über den *reellen* Zahlen, denn ein reelles Polynom muss nicht unbedingt reelle Nullstellen haben wie das Beispiel $p(z) = z^2 + 1$ zeigt. Dieser Umstand ist der wesentlichste Grund weshalb das Rechnen mit komplexen Vektorräumen deutlich einfacher ist als mit reellen.

Jetzt können wir die Resultate zu einem Algorithmus für die Berechnung von Eigenwerten und Eigenvektoren zusammenstellen:

Algorithmus 11.3.4.

Input: $A \in \mathbb{C}^{n,n}$

Output: Alle Eigenwerte und zugehörige Eigenvektoren.

Durchführung:

1. Schritt: Berechne das charakteristische Polynom $p_A(z)$.
2. Schritt: Berechne die Nullstellen von $p_A(z)$.¹ Dies sind die Eigenwerte.
3. Schritt: Berechne für jede Nullstelle λ eine Basis des linearen Teilraums $\{\vec{v} \mid \vec{v} \in \mathbb{C}^n, A\vec{v} = \lambda\vec{v}\}$. Dies ist eine Basis des Eigenraums V_{λ} von A zum Eigenwert λ .

¹Für $n > 4$ kann dies im Allgemeinen nur numerisch gemacht werden.

Als letztes wenden wir uns der Frage zu, ob es zu jeder linearen Abbildung eines endlichdimensionalen Vektorraums in sich eine Basis aus Eigenvektoren gibt. Die Antwort darauf lautet zwar nein, wie das folgende Beispiel 11.3.5 zeigt, es gibt jedoch große und wichtige Klassen linearer Abbildungen, auf die dies zutrifft. Dies wird das Thema der nächsten Vorlesung sein, wo es um sog. *diagonalisierbare Matrizen* geht.

Beispiel 11.3.5. Sei $V = \mathbb{C}^2$ und $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Das charakteristische Polynom von A ist

$$p(z) = \det(A - zI_2) = z^2.$$

Es gibt also nur den einen Eigenwert $\lambda = 0$. Der Eigenraum dazu, bzw. der Kern von A , ist $\text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$. Er ist eindimensional, d.h. es gibt keine Basis von Eigenvektoren für diese lineare Abbildung.

Dieses Beispiel zeigt, dass der Eigenwert 0 eine doppelte Nullstelle des charakteristischen Polynoms $p_A(z)$ darstellt. Dies führt auf den Begriff der algebraischen Vielfachheit, der – zusammen mit seinem Partner, der geometrischen Vielfachheit – in der nächsten Vorlesung eine nützliche Rolle spielt:

Definition 11.3.6. Sei $p_A(z)$ das charakteristische Polynom der quadratischen Matrix A und λ eine Nullstelle (=Eigenwert). Dann kann $p_A(z)$ in der Form $(z - \lambda)^r q(z)$ geschrieben werden, wobei $q(z)$ ein Polynom in z ist. Die größte natürliche Zahl r für die es eine solche Faktorisierung gibt, heißt *algebraische Vielfachheit* des Eigenwertes λ .

Eng verwandt mit diesem Begriff ist die *geometrische Vielfachheit* eines Eigenwertes λ (Definition 11.1.15).

Im nächsten Kapitel werden wir uns der Frage zuwenden, wann eine lineare Abbildung eine Basis von Eigenvektoren besitzt. Falls dies der Fall ist, sind wir in einem besonders einfachen Fall, da die darstellende Matrix bezüglich der Basis von Eigenvektoren dann offensichtlich diagonal ist. Leider trifft dieser Fall nicht immer zu. Die beste Strukturaussage für solche Fälle ist dann der Satz von Jordan (siehe ?? im Anhang). Ein glücklicher Umstand ist allerdings, dass dieser Fall selten eintritt, im selben Sinne wie sich zwei Geraden im \mathbb{R}^3 selten treffen.

Der Jordansche Satz kann dazu verwendet werden folgende Bemerkung zu beweisen:

Bemerkung 11.3.7. Die algebraische Vielfachheit eines Eigenwertes ist größer oder gleich der geometrischen.

DIAGONALISIERBARKEIT VON MATRIZEN

Mit diagonalen Matrizen lässt sich sehr einfach rechnen. Multiplizieren, Addieren, Berechnen von Eigenwerten, Eigenvektoren, all dies ist mit diagonalen Matrizen ein Leichtes. Dasselbe trifft für diagonalisierbare Matrizen zu, d.h. für solche Matrizen A , die durch Ähnlichkeitstransformation mit einer invertierbaren Matrix S auf Diagonalgestalt gebracht werden können:

$$\text{Diagonalmatrix} = S^{-1}AS$$

oder – was dazu äquivalent ist – die als Ähnlichkeitstransformation einer Diagonalmatrix geschrieben werden können

$$A = S \text{ Diagonalmatrix } S^{-1}.$$

Wie wir bereits im letzten Kapitel gesehen haben können nicht alle Matrizen diagonalisiert werden. Nach Satz ?? (siehe Anhang) muss dazu die Jordan Normalform diagonal sein, was heißt, dass keine Blöcke grösser als 1 auftreten dürfen. Für wichtige Klassen von Matrizen trifft dies immer zu; z.B. können alle symmetrischen, alle selbstadjungierten und alle unitären Matrizen diagonalisiert werden. Weiterhin sind alle $n \times n$ Matrizen diagonalisierbar, die – als lineare Abbildung aufgefasst – n paarweise verschiedene Eigenwerte haben.

Die Antwort auf die Frage der Diagonalisierbarkeit von Matrizen ist sehr verschieden, je nach dem, ob sie für reelle oder komplexe Matrizen gestellt wird. Dies hat unmittelbar damit zu tun, dass Polynome mit reellen Koeffizienten nicht unbedingt auch reelle Nullstellen haben wie das folgende Beispiel einer Drehung von \mathbb{R}^2 zeigt:

Beispiel 12.0.8. Sei $V = \mathbb{R}^2$ und $A = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ (vgl. Beispiel 11.1.12).

Das charakteristische Polynom von A ist

$$\begin{aligned} p_A(z) &= \det \begin{bmatrix} \frac{1}{\sqrt{2}} - z & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - z \end{bmatrix} \\ &= \left(z - \frac{1}{\sqrt{2}}\right)^2 + \frac{1}{2}. \end{aligned}$$

Es hat die beiden Nullstellen

$$\frac{1}{\sqrt{2}}(1 \pm i) = e^{\pm i\pi/4}.$$

Insbesondere hat das charakteristische Polynom keine reelle Nullstelle und somit gibt es keine Eigenwerte und Eigenvektoren von A als lineare Abbildung von \mathbb{R}^2 nach \mathbb{R}^2 .

Wie wir zuvor gesehen haben (vgl. Beispiel 11.1.12), hat die Matrix $A = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ als lineare Abbildung von \mathbb{C}^2 nach \mathbb{C}^2 sehr wohl Eigenvektoren zu den komplexen Eigenwerten $e^{\pm i\pi/4}$, den komplexen Nullstellen des charakteristischen Polynoms.

Beispiel 12.0.9. Sei $R(\varphi)$ eine Drehung um den Winkel $\varphi \in [0, 2\pi[$ des \mathbb{R}^2 . Dann ist das charakteristische Polynom

$$\begin{aligned} p_{R(\varphi)}(z) &= \det \begin{bmatrix} \cos \varphi - z & -\sin \varphi \\ \sin \varphi & \cos \varphi - z \end{bmatrix} \\ &= z^2 - 2(\cos \varphi) \cdot z + 1. \end{aligned}$$

$p(z)$ hat genau dann eine reelle Nullstelle, wenn $\varphi = 0$ oder π ist. Das bedeutet insbesondere, dass es für jede Drehung um den Winkel φ , die von der Identität oder minus der Identität verschieden ist, keine reellen Eigenvektoren und Eigenwerte gibt.

Dieser Unterschied zwischen dem Fall reeller und dem Fall komplexer Matrizen hat Konsequenzen für das praktische Rechnen. Denn er legt die Strategie nahe, auch reelle Matrizen grundsätzlich als spezielle komplexe Matrizen aufzufassen; so gelingt die Diagonalisierung fast immer und die Rechnungen können damit oftmals enorm verkürzt werden. *Wir werden deshalb in diesem Kapitel grundsätzlich alle Matrizen - auch reelle - als komplexe auffassen*, genau so wie reelle Zahlen als Spezialfälle komplexer verstanden werden können.

Ein Beispiel für den Erfolg dieser Strategie werden wir später bei der Behandlung des elektrischen Schwingkreises kennen lernen, der – mathematisch gesehen – das selbe ist wie das Federpendel mit Reibung.

Die Diagonalisierung von Matrizen ist aber nicht nur eine wichtige Rechenmethode sondern sie ist von großer Bedeutung in manchen Anwendungen, z.B. in der Modellierung der elektrischen Polarisierung von Materie, der Beschreibung der Dynamik eines Kreisels, der Beschreibung elastischer Materie mit Spannungs- und Verzerrungstensor, der Beschreibung von Atomen und Molekülen um nur einige wenige Beispiele zu nennen.

Zur Diagonalisierung von Matrizen werden wir einen Algorithmus angeben, von dem gezeigt werden kann, dass er bei diagonalisierbaren Matrizen immer zum Ziel führt (vgl. Theorem 12.4.4). In der Praxis d.h. auf dem Rechner verwendet man allerdings einen anderen wichtigen Ansatz, die Schur-Form, die wir auch kurz vorstellen werden. Diese Methode ist numerisch stabiler als die auf dem Algorithmus (12.3.1) basierende Methode d.h. sie ist unempfindlicher gegen Rundungsfehler, wie sie bei numerischen Rechnungen unvermeidlich sind.

12.1 DEFINITION UND WEITERE BEISPIELE

Zunächst einmal eine kleine Erinnerung an einen Begriff, der uns bereits in der dritten Vorlesung über Matrizen begegnet ist: Eine quadratische Matrix $D = [d_{ik}] \in \mathbb{R}^{n,n}$ heißt diagonal, wenn $d_{i,k} = 0$ ist für alle Indizes $i \neq k$. (vgl. Definition 3.2.7).

Nun kommen wir zur Definition des zentralen Begriffs dieser Vorlesung:

Definition 12.1.1 (Diagonalisierbare Matrix). Eine Matrix $A \in \mathbb{C}^{n,n}$ heißt diagonalisierbar, falls es eine invertierbare Matrix $S \in \mathbb{C}^{n,n}$ gibt, sodass die Matrix $D = S^{-1}AS$ diagonal ist.

Bemerkung 12.1.2. Im Folgenden wird wichtig sein: Für jede Diagonalmatrix $D = \text{diag}\{d_1, \dots, d_n\} \in \mathbb{C}^{n,n}$ ist die Standardbasis von \mathbb{C}^n eine Basis von Eigenvektoren. Die Elemente $\{d_1, \dots, d_n\}$ auf der Diagonale sind die zugehörigen Eigenwerte.

Beispiel 12.1.3. [Diagonalisierbare Matrizen] Die beiden Matrizen (Paulimatrizen σ_1 und σ_2) sind diagonalisierbar:

$$\begin{aligned}\sigma_1 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ \sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}\end{aligned}$$

12.2 DIAGONALISIERUNG, EIGENVEKTOREN UND EIGENWERTE

Eine Matrix $A \in \mathbb{C}^{n,n}$ identifizieren wir wie üblich mit der linearen Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$. Damit kann die Frage der Diagonalisierbarkeit unter einem anderen Gesichtspunkt betrachtet werden wie dies in dem folgenden Satz formuliert ist, der unmittelbar aus der Jordanschen Normalform (siehe Anhang) folgt. Darin wird ein Zusammenhang zwischen Diagonalisierbarkeit und der Existenz einer Basis von Eigenvektoren formuliert:

Theorem 12.2.1 (Diagonalisierbarkeit und Eigenvektoren). *Sei $A \in \mathbb{C}^{n,n}$. Dann gilt:*

1. *Die Matrix A ist diagonalisierbar. \Leftrightarrow Die lineare Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ hat eine Basis von Eigenvektoren.*
2. *Die Matrix $A \in \mathbb{C}^{n,n}$ ist diagonalisierbar. \Leftrightarrow Für alle Eigenwerte der Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ gilt: algebraische Vielfachheit = geometrische Vielfachheit.*
3. *Sei $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von Eigenvektoren der linearen Abbildung A . Dann ist die Matrixdarstellung von A in der Basis \mathcal{B} diagonal.*
4. *Sei A diagonalisierbar, d.h. es existiert eine invertierbare Matrix $S \in \mathbb{C}^{n,n}$, sodass $S^{-1}AS$ diagonal ist. Dann bilden die Spaltenvektoren von S eine Basis von Eigenvektoren der Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$.*
5. *Sei A diagonalisierbar, dann sind die Einträge auf der Diagonalen ihrer Diagonalisierung D die Eigenwerte der Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$:*

$$D = \text{diag}\{\lambda_1, \dots, \lambda_1; \lambda_2, \dots, \lambda_2; \dots; \lambda_k, \dots, \lambda_k\}$$

12.3 ALGORITHMUS ZUR BERECHNUNG DER DIAGONALISIERUNG

Nach dem obigen Satz ist die Diagonalisierung äquivalent zur Konstruktion einer Basis von Eigenvektoren. Der Gedanke liegt dem folgenden Algorithmus zu Grunde, den wir für den Fall komplexer Matrizen formulieren:

Algorithmus 12.3.1 (Diagonalisierung von Matrizen).

Input: Quadratische Matrix $A \in \mathbb{C}^{n,n}$

Output: Falls die Matrix A diagonalisierbar ist: Diagonalisierende Matrix S und Diagonalmatrix D

Andernfalls: Abbruch.

1. Schritt: Berechne die Nullstellen $\lambda_1, \dots, \lambda_k$ des charakteristischen Polynoms. Dies sind nach Theorem 11.2.1 die Eigenwerte der linearen Abbildung $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$.
2. Schritt: Berechne für $i = 1, \dots, k$ die Lösungsmengen der Gleichungssysteme

$$\text{Kern}(A - \lambda_i I_n) = \{\vec{v} \mid (A - \lambda_i I_n)\vec{v} = \vec{0}\}$$

z.B. mit dem *Gaußalgorithmus*.

3. Schritt: Prüfe ob für alle Eigenwerte die algebraische Vielfachheit mit der geometrischen übereinstimmt. Wenn ja gehe zum nächsten Schritt, andernfalls Abbruch.
4. Schritt: Wähle eine Basis \mathcal{B}_i in $\text{Kern}(A - \lambda_i I_n)$, $i = 1, \dots, k$.
5. Schritt: $\mathcal{B} := \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ ist eine Basis von Eigenvektoren. Die darstellende Matrix $A_{\mathcal{B}}$ ist nach Konstruktion diagonal.

12.4 RECHNEN MIT DIAGONALISIERBAREN MATRIZEN

In diesem Abschnitt betrachten wir das charakteristische Polynom und einige spezielle Funktionen diagonalisierbarer Matrizen, wie Polynome und die Exponentialfunktion.

Für diagonalisierbare Matrizen gilt offensichtlich (vgl. Theorem 11.2.1)

$$A = SDS^{-1} \Rightarrow p_A(z) = p_D(z) = \prod_{i=1}^k (\lambda_i - z)^{r_i},$$

r_i ($i = 1, \dots, k$) bezeichnet dabei die Vielfachheit (= geometrische Vielfachheit = algebraische Vielfachheit) des Eigenwertes λ_i .

Beispiel 12.4.1. [Charakteristisches Polynom diagonalisierbarer Matrizen]

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} &= 1/\sqrt{2} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} 1/\sqrt{2} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \\ p_A(z) &= (z-1)^2(z+1). \end{aligned}$$

Jetzt zeigen wir, wie in natürlicher Weise gewisse einfache Funktionen von Matrizen eingeführt werden können. Dies steht im Einklang mit der Analysis von Funktionen mehrerer Veränderlichen. Die Betonung liegt hier jedoch auf der Algebra. Die Begriffe haben vorerst nichts mit der Diagonalisierbarkeit der Matrizen zu tun. Was hier vorgeführt werden soll ist vielmehr, dass im Falle diagonalisierbarer Matrizen diese Funktionen einfach berechnet werden können.

Sei p ein Polynom N -ten Grades in der Unbestimmten x und Koeffizienten in \mathbb{C} ,

$$p(x) = \sum_{l=0}^N a_l x^l.$$

Sei $A \in \mathbb{C}^{n,n}$ eine quadratische Matrix. Dann definieren wir die neue Matrix

$$p(A) := \sum_{l=0}^N a_l A^l \in \mathbb{C}^{n,n}.$$

Beispiel 12.4.2. [Polynom von Matrizen] Sei A dieselbe Matrix wie im vorangehenden Beispiel 12.4.1 und $p(x) = x^l$ ein Monom vom Grad l . Dann ist

$$\begin{aligned} p(A) &= A^l = (SDS^{-1})^l \\ &= SDS^{-1} SDS^{-1} \dots SDS^{-1} \\ &= SD^l S^{-1} \\ &= S \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (-1)^l \end{bmatrix} S^{-1}. \end{aligned}$$

Das obige Beispiel kann leicht verallgemeinert werden und führt zu dem Resultat:

Theorem 12.4.3 (Polynom diagonalisierbarer Matrizen). *Seien $A \in \mathbb{C}^{n,n}$ und $p(x) = \sum_{l=0}^N a_l x^l \in \mathbb{C}_{\leq N}[x]$. Dann gilt:*

1. $A = SDS^{-1} \Rightarrow p(A) = Sp(D)S^{-1}$
2. $p(D) = \text{diag}\{p(\lambda_1), \dots, p(\lambda_1); p(\lambda_2), \dots, p(\lambda_2); \dots; p(\lambda_k), \dots, p(\lambda_k)\}$

Die Exponentialfunktion einer quadratischen Matrix $A \in \mathbb{C}^{n,n}$ ist so definiert:

$$e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n \quad (12.1)$$

Diese Reihe ist – wie in der Analysis – über die Folge der Partialsummen definiert. Es kann gezeigt werden, dass für jedes Paar von Indizes i, k aus $1, \dots, n$ die Folge

$$\left\{ \left(\sum_{n=0}^N \frac{1}{n!} A^n \right)_{i,k} \right\}_{N=0}^{\infty}$$

konvergiert.

Im Falle diagonalisierbarer Matrizen ist die Situation wiederum besonders einfach. Es gilt

Theorem 12.4.4 (Exponentialfunktion diagonalisierbarer Matrizen). *Sei $A \in \mathbb{C}^{n,n}$ diagonalisierbar, d.h. es gibt $S \in \mathbb{C}^{n,n}$ invertierbar, sodass $A = SDS^{-1}$ mit der Diagonalmatrix*

$$D = \text{diag} \{ \lambda_1, \dots, \lambda_1; \lambda_2, \dots, \lambda_2; \dots; \lambda_k, \dots, \lambda_k \}.$$

Dann ist $e^A = Se^D S^{-1}$, wobei

$$e^D = \text{diag}\{e^{\lambda_1}, \dots, e^{\lambda_1}; e^{\lambda_2}, \dots, e^{\lambda_2}; \dots; e^{\lambda_k}, \dots, e^{\lambda_k}\}.$$

Bemerkung 12.4.5. Da wir in diesem Kurs nicht über genügend Instrumente aus der Analysis verfügen, um diesen Satz zu beweisen, kann er nur plausibel gemacht werden. Dazu dient das folgende Beispiel.

Beispiel 12.4.6. [Exponentialfunktion von Matrizen] Die Aussage des obigen Satzes kann an folgendem Beispiel überprüft werden: Die Matrix

$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$$

ist diagonalisierbar, wie in Beispiel 12.4.1 gezeigt wurde.

Als erstes berechnen wir $e^{-i\phi\sigma_2}$ für $\phi \in \mathbb{R}$ und zwar über die Potenzreihe. Danach werden wir den Ausdruck $Se^D S^{-1}$ berechnen und zeigen, dass die beiden Ausdrücke identisch sind.

Die Exponentialreihe

$$e^{-i\phi\sigma_2} = \sum_{n=0}^{\infty} \frac{1}{n!} (-i\phi)^n \sigma_2^n$$

kann sofort summiert werden, weil $\sigma_2^2 = I_2$. Damit kann die Reihe in gerade und ungerade Exponenten von σ_2 zerlegt werden:

$$\begin{aligned} e^{-i\phi\sigma_2} &= \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \phi^{2n} \right) I_2 - i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \phi^{2n+1} \sigma_2 \\ &= (\cos \phi) I_2 - i(\sin \phi) \sigma_2 \end{aligned}$$

Nun berechnen wir den Ausdruck $Se^D S^{-1}$ und verwenden dazu die Gleichung

$$Se^D S^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} e^{-i\phi} & 0 \\ 0 & e^{i\phi} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -i \\ -i & 1 \end{bmatrix}.$$

Wenn diese drei Matrizen ausmultipliziert werden, ergibt sich das gewünschte Resultat.

12.5 KLASSEN DIAGONALISIERBARER UND NICHTDIAGONALISIERBARER MATRIZEN

In diesem Abschnitt gehen wir durch die wichtigsten Klassen diagonalisierbarer und nicht diagonalisierbarer Matrizen. Das erste Resultat, das in dem folgenden Satz ausgedrückt wird, ist eine unmittelbare Folge von Theorem 11.2.2, das besagt, dass es zu jeder Nullstelle in \mathbb{C} des charakteristischen Polynoms von $A \in \mathbb{C}^{n,n}$ mindestens einen Eigenvektor gibt und dass ferner Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind.

Theorem 12.5.1. Sei $A \in \mathbb{C}^{n,n}$. Dann gilt:

Die lineare Abbildung $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ hat n paarweise verschiedene Eigenwerte \Rightarrow die Matrix A ist diagonalisierbar.

Bemerkung 12.5.2. Die Umkehrung der obigen Aussage gilt nicht!

In der folgenden Tabelle sind die wichtigsten Klassen diagonalisierbarer Matrizen zusammengestellt. In manchen Fällen kann die diagonalisierende Matrix S speziell einfach gewählt werden:

Matrixklasse	diagonalisierbar bzw nicht diagonalisierbar
$A \in \mathbb{C}^{n,n}$ n paarweise verschiedene EW	diagonalisierbar mit $S \in \mathbb{C}^{n,n}$
$A \in \mathbb{R}^{n,n}$ symmetrisch d.h. $A = A^T$	diagonalisierbar mit S Drehung
$A \in \mathbb{C}^{n,n}$ selbstadjungiert d.h. $A = A^* := \overline{A^T}$	diagonalisierbar mit S unitär
$U \in \mathbb{C}^{n,n}$ unitär d.h. $U^* U = I_n$	diagonalisierbar mit S unitär
$A \in \mathbb{C}^{n,n}, \quad A \neq 0$ $A^m = 0$, für ein $m \in \mathbb{N}$	nicht diagonalisierbar

Bemerkung 12.5.3. Orthogonale Matrizen können i.A. nicht als reelle Matrizen diagonalisiert werden jedoch als Spezialfälle von unitären; d.h. die diagonalisierende Matrix S ist dann nicht reell wählbar sondern komplex.

Beispiel 12.5.4. [Symmetrische Matrix] Ein Satellit wird als ein starrer Körper modelliert, auf den keine Kräfte wirken. Sein Schwerpunkt folgt der Bahn eines freien Massenpunktes mit der Gesamtmasse des Satelliten. Der Satellit selbst dreht sich um seinen Schwerpunkt. Die Drehung wird durch den Vektor $\vec{\Omega} \in \mathbb{R}^3$ beschrieben, der in die Drehachse zeigt und die Länge der Winkelgeschwindigkeit hat (siehe folgende Abbildung). Die Energie ist durch das

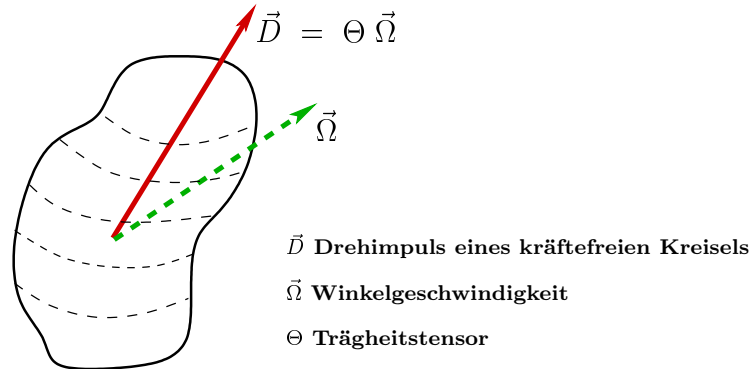


ABBILDUNG 12.1: Starrer Körper mit einwirkenden Kräften

Skalarprodukt $\langle \vec{\Omega}, \Theta \vec{\Omega} \rangle$ gegeben, wobei Θ den sogenannten Trägheitstensor bezeichnet. Er ist für einen einzigen Massenpunkt am Ort \vec{x} der Masse m durch folgende symmetrische 3×3 Matrix gegeben:

$$\Theta = m \cdot \begin{bmatrix} x_2^2 + x_3^2 & -x_1 x_2 & -x_1 x_3 \\ -x_1 x_2 & x_3^2 + x_1^2 & -x_2 x_3 \\ -x_1 x_3 & -x_2 x_3 & x_1^2 + x_2^2 \end{bmatrix}$$

Für mehrere Massenpunkte ist der Trägheitstensor als Summe der einzelnen Teile definiert, für ein Kontinuum wird dies ein Integral.

Die Eigenvektoren und Eigenwerte sind entscheidend für das dynamische Verhalten des Satelliten. Seien $\alpha_1, \alpha_2, \alpha_3$ die der Größe nach angeordneten Eigenwerte von Θ (nach der obigen Tabelle ist Θ diagonalisierbar) und $\vec{v}_1, \vec{v}_2, \vec{v}_3$ die zugehörigen Eigenvektoren. Dann ist die Bahn des Drehimpulses $\vec{D} := \Theta \vec{\Omega}$ geometrisch wie folgt zu ermitteln (siehe folgende Abbildung): \vec{D} muss sich nach den Gesetzen der Mechanik immer auf einer Fläche konstanter Energie bewegen. Die Energie ist folgendermaßen definiert: $E := \langle \vec{D}, \Theta^{-1} \vec{D} \rangle$. Für eine vorgegebene Energie

bilden die möglichen Drehimpulse \vec{D} einen Ellipsoid. Nun gibt es aber noch eine weitere Bedingung an den Vektor \vec{D} . Er muss immer dieselbe Länge haben. Das heißt er liegt während des gesamten Bewegungsablaufes auf einer Kugel mit Radius r . Somit ist die Bahn von \vec{D} im Schnitt der Kugel und des Ellipsoids enthalten. Die Figur 12.2 zeigt, dass es nicht klug

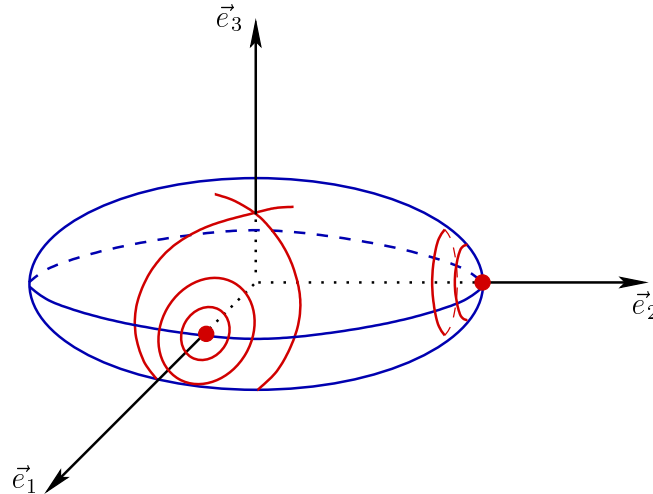


ABBILDUNG 12.2: Bahnen des Drehimpulses \vec{D} eines kräftefreien Kreisels (Satelliten) sind Schnittlinien eines Ellipsoids und den Sphären für verschiedene Radien (Energien) in \mathbb{R}^3

ist, einen Satelliten so zu bauen, dass sein Drehmoment in die Richtung des zweiten Eigenvektors zeigt, d.h. in die Richtung des Vektors \vec{v}_3 , es sei denn man ist am Torkeln des Satelliten interessiert.

Wir haben gesehen, dass eine Diagonalisierung von Matrizen viele Aufgaben wie z.B. die Berechnung von Matrixfunktionen, wie Polynomen oder der Matrix-Exponentialfunktionen erheblich erleichtert. Allerdings ist die Diagonalisierung nicht immer möglich, insbesondere auf dem Rechner ist sie sehr gefährlich, weil durch Rundungsfehler die Eigenwerte gestört werden. Nun ist es fast so einfach mit Dreiecksmatrizen zu rechnen, wie mit Diagonalmatrizen. Daher verwenden heutzutage fast alle numerischen Methoden zur Berechnung von Eigenwerten, Eigenvektoren und von Matrixfunktionen stattdessen das folgende Resultat.

Theorem 12.5.5. [Schur-Form]

Sei $A \in \mathbb{C}^{n,n}$, so existiert eine unitäre Matrix $U \in \mathbb{C}^{n,n}$, so dass

$$T = U^{-1}AU = \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{bmatrix},$$

wobei die Eigenwerte auf der Diagonale in jeder beliebigen Reihenfolge stehen können.

Diese Form kann man auch dem Rechner sehr gut ausrechnen weil die Transformation mit unitären Matrizen (wegen der Erhaltung von Längen und Winkeln) auch die Fehler nicht verstärkt. Den Algorithmus zur Berechnung der Schur-Form einer Matrix erklären wir an dem folgenden Beispiel:

Beispiel 12.5.6. Wir bringen die Matrix

$$A: = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$$

auf *Schur-Form* und stellen dies so dar, dass das Verfahren auf beliebige quadratische Matrizen ausgedehnt werden kann.

Als erstes wird eine Eigenvektor von A berechnet z.B. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Er wird normiert,

d.h. durch seine Länge geteilt: $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Damit haben wir einen normierten Eigenvektor \vec{u}_1 zum Eigenwert 2 konstruiert. Er wird als erster Spaltenvektor einer unitären Matrix U aufgefasst. Um ganz U zu konstruieren muss ein zweiter Vektor der Länge eins gefunden werden, der auf dem ersten senkrecht steht, z.B. $\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Er wird als zweiter Spaltenvektor von U aufgefasst:

$$U: = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Im allgemeinen Fall einer $n \times n$ Matrix müssten $n - 1$ orthonormierte Vektoren gefunden werden, die den ersten zu einer Orthonormalbasis ergänzen. Sie werden als Spaltenvektoren einer unitären Matrix aufgefasst.

Nun zurück zu unserem Beispiel einer 2×2 Matrix. Nach Konstruktion ist

$$\vec{u}_1 = U \vec{e}_1$$

wobei \vec{e}_1 den ersten Standardbasisvektor von \mathbb{C}^2 bezeichnet. Somit ist

$$AU \vec{e}_1 = 2U \vec{e}_1$$

Multiplikation von links mit U^{-1} ergibt:

$$U^{-1}AU \vec{e}_1 = 2\vec{e}_1$$

d.h. die Matrix $U^{-1}AU$ ist eine obere Dreiecksmatrix. Dies wird durch folgende Rechnung bestätigt:

$$\begin{aligned} U^{-1}AU: &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

Damit ist die Matrix A auf *Schur-Form* gebracht worden.

Bemerkung 12.5.7. In einigen in der Anwendung sehr wichtigen Spezialfällen, wie bei selbstadjungierten, orthogonalen oder unitären Matrizen liefert der Satz von Schur ganz offensichtlich sogar gleich die Diagonalisierung, denn ist beispielsweise $A \in \mathbb{C}^{n,n}$ selbstadjungiert, d.h. $A = A^*$, so gilt dies auch für die Dreiecksmatrix der Schur-Form $T = U^*AU = U^*A^*U = T^*$ und damit ist T diagonal.

LINEARE DIFFERENTIALGLEICHUNGEN (TEIL 1)

Differentialgleichungen unterscheiden sich wesentlich von den Gleichungen, wie wir sie bis jetzt in diesem Kurs angetroffen haben. Lösungen von Differentialgleichungen sind nicht mehr Zahlen sondern Funktionen mit Werten in \mathbb{R}^n oder \mathbb{C}^n . Eine Lösung $\vec{x}(t)$ gibt beispielsweise an, an welchem Ort sich ein System zur Zeit t befindet.

Alle fundamentalen Naturgesetze können als Differentialgleichungen geschrieben werden. Ein typisches Beispiel ist das Newtonsche Gesetz, das so lautet: Die Kraft \vec{k} , die auf einen Massenpunkt am Ort $\vec{x}(t)$ zur Zeit t wirkt, ist gleich seiner Masse m mal seiner Beschleunigung:

$$m \frac{d^2 \vec{x}}{dt^2}(t) = \vec{k}(\vec{x}(t))$$

Bei konstanter Schwerkraft mg in Fallrichtung (g bezeichnet die Erdbeschleunigung $g = 9,81 m/s^2$) führt dies auf das Galileische Gesetz $x(t) = g \frac{t^2}{2}$, worin $x(t)$ die Fallhöhe zur Zeit t bezeichnet. Das Galileische Gesetz ist eine *Lösung* der Newtonschen Differentialgleichung.

Die beiden wichtigsten Probleme in der Theorie der Differentialgleichungen sind das *Anfangswertproblem* und das *Randwertproblem*. Beide sollen vorerst an typischen Beispielen skizziert werden.

Als Beispiel für ein Anfangswertproblem betrachten wir ein Fahrzeug, das sich in der Ebene bewegen kann. An jedem Ort \vec{x} der Ebene ist die Geschwindigkeit $\vec{v}(\vec{x})$ vorgeschrieben. Ferner ist die Startzeit t_0 und der Startort \vec{x}_0 vorgegeben. Gesucht ist der Weg des Fahrzeuges $\{\vec{x}(t) | t \geq 0\}$, sodass gilt

$$\frac{d\vec{x}}{dt}(t) = \vec{v}(\vec{x}(t)), \quad (13.1)$$

d.h. das Fahrzeug ist zur Zeit t am Ort $\vec{x}(t)$ und hat da die vorgeschriebene Geschwindigkeit $\vec{v}(\vec{x}(t))$. Man sagt $\vec{x}(\cdot)$ ist Lösung der Differentialgleichung 13.1.

Ein typisches Randwertproblem ergibt sich aus der mathematische Modellierung der Schwingung einer Saite der Länge L , die an beiden Enden festgeklemmt ist. Die Saite schwinde um die Ruhelage mit der Frequenz ω , d.h. die Auslenkung um die Ruhelage zur Zeit t am Ort x sei von der Form $\sin(\omega t)y(x)$. Mit anderen Worten, die Auslenkung faktorisiert in den zeitabhängigen Term $\sin(\omega t)$ und den ortsabhängigen Teil $y(x)$. Unter diesen Voraussetzungen ergibt sich für die Amplitudenfunktion $y(x)$ das Randwertproblem:

$$\frac{d^2 y}{dx^2}(x) = -\omega^2 y(x) \quad \text{und} \quad y(0) = y(L) = 0.$$

Eine Lösung der obigen Differentialgleichung ist $y(x) = \sin(\omega x)$. Die Randbedingungen werden von $y(x)$ aber nur für ganz spezielle Werte von ω erfüllt,

nämlich dann wenn ω ein ganzzahliges Vielfaches von π/L ist, d.h. das Randwertproblem hat nicht für alle ω s eine Lösung. Wir werden sehen, dass die Werte für die es eine Lösung gibt, als Eigenwerte einer linearen Abbildung aufgefasst werden können.

Wir werden uns ausführlicher mit den sogenannten *linearen* Differentialgleichungen befassen. Sie zu lösen kann als ein Problem der linearen Algebra verstanden werden. Das Anfangswertproblem kann für lineare Differentialgleichungen immer *eindeutig* gelöst werden. Die eindeutige Lösbarkeit des Anfangswertproblems gilt auch für den nicht linearen Fall unter sehr allgemeinen Voraussetzungen. Den linearen Fall zeichnet aus, dass dafür die lineare Algebra einen einfachen *Lösungsalgorithmus* liefert.

Zum Schluss noch ein Wort zur Bedeutung der *Theorie* der Differentialgleichungen, als Aufmunterung, sie zu lernen: Die wenigsten Differentialgleichungen, die in der Praxis angetroffen werden, können gelöst werden. Deshalb sind Approximationsverfahren und qualitative Methoden von besonderer Bedeutung für die Praxis.

Ein besonders nützliches Approximationsverfahren ist die sogenannte Linearisierung einer Differentialgleichung, die wir an der Pendelgleichung im Abschnitt 13.2 kennen lernen werden. Wie der Name sagt, wird das vorerst explizit nicht lösbare Problem durch eine lineare Differentialgleichung ersetzt, die lösbar ist. Andere Approximationsverfahren, wie etwa die Diskretisierung der Differentialgleichung (Euler-Verfahren, Runge-Kutta Verfahren) werden in der Vorlesung Numerische Mathematik besprochen.

Zu den qualitativen Methoden gehören die Beantwortung von Fragen folgender Art:

- Existiert eine Lösung und wenn ja wie viele?
- Was ist die Struktur der Lösungsmenge?
- Wie verhalten sich Lösungen für große Zeiten?
- Wie steht es mit der Stabilität von Lösungen unter kleinen Störungen?

Diese letzte Frage ist für Anwendungen in den Ingenieurwissenschaften von besonders grossem Interesse, wenn es zum Beispiel um die Frage geht, ob eine schnell rotierende Turbine, den Ansprüchen kleiner Störungen genügt. Es ist interessant, dass Fragen dieser Art in sehr vielen Fällen beantwortet werden können, ohne dass explizite Lösungen berechnet werden müssen!

Diese zugegebenermassen recht theoretischen Fragen, sind wichtig und nützlich, wenn es um die Beurteilung der Vertrauenswürdigkeit von Lösungen geht, die numerisch mit Hilfe von Softwareprogrammen gefunden wurden. Hier kommt eine besondere Verantwortung auf Sie zu, die typisch ist für unsere Zeit: Bald kann jedes Kind Differentialgleichungen mit dem Computer numerisch lösen und damit komplizierte Vorgänge modellieren. Manche dieser Modellierungen betreffen Anwendungen, die durchaus nicht harmloser Natur sind. Es ist deshalb in solchen Fällen unumgänglich sich bei der numerischen Lösung von Differentialgleichungen Rechenschaft zu geben, ob die Lösungen korrekt sind. Hierbei spielt das theoretische Verständnis der qualitativen Eigenschaften von Lösungen insbesondere die Stabilität von Lösungen eine herausragende Rolle.

Differentialgleichungen können in lineare und nichtlineare, in Differentialgleichungen 1-ter Ordnung und n -ter Ordnung, in gewöhnliche und partielle gruppiert werden. Wir werden diese systematische Klassifizierung auf das Ende der

nächsten Vorlesung - der letzten - verschieben, wenn uns bereits eine gewisse Zahl von Beispielen zur Verfügung steht.

13.1 DEFINITION UND BEISPIELE

In diesem Abschnitt definieren wir die grundlegenden Begriffe der Theorie linearer Differentialgleichungen mit konstanten Koeffizienten.

Definition 13.1.1. Sei $A \in \mathbb{K}^{n,n}$ eine quadratische Matrix. Dann heißt

$$\frac{d\vec{y}}{dt}(\cdot) = A\vec{y}(\cdot) \quad (13.2)$$

eine *homogen lineare Differentialgleichung erster Ordnung mit konstanten Koeffizienten*.

Sei ferner $t_0 \in \mathbb{R}$ und $\vec{y}_0 \in \mathbb{K}^n$. Dann heißt

$$\begin{aligned} \frac{d\vec{y}}{dt}(\cdot) &= A\vec{y}(\cdot) \\ \vec{y}(t_0) &= \vec{y}_0 \end{aligned} \quad (13.3)$$

das *Anfangswertproblem* der linearen Differentialgleichung (13.2) mit den Anfangsdaten t_0 und \vec{y}_0 .

Eine auf einem Intervall \mathbb{I} definierte differenzierbare Funktion $\vec{y}(t)$ mit Werten in \mathbb{K}^n heißt eine *Lösung der Differentialgleichung (13.2)*, falls die Vektoren $\vec{y}(t)$ und $\frac{d\vec{y}}{dt}(t)$ die Gleichung 13.2 für alle $t \in \mathbb{I}$ erfüllt.

$\vec{y}(\cdot)$ heißt eine *Lösung des Anfangswertproblems (13.3)*, falls zusätzlich die vorgegebenen Anfangsdaten angenommen werden, d.h. falls die Gleichungen (13.3) erfüllt sind.

Bemerkung 13.1.2. Die Differentialgleichung (13.2) sieht ausgeschrieben so aus:

$$\begin{bmatrix} \frac{dy_1}{dt}(\cdot) \\ \vdots \\ \frac{dy_n}{dt}(\cdot) \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \begin{bmatrix} y_1(\cdot) \\ \vdots \\ y_n(\cdot) \end{bmatrix}$$

Beispiel 13.1.3. Die Newtonsche Gleichung für das Federpendel der Frequenz $\omega > 0$ ist

$$\frac{d^2x}{dt^2}(\cdot) + \omega^2 x(\cdot) = 0. \quad (13.4)$$

Dabei bezeichnet $x(t)$ die Auslenkung aus der Ruhelage zur Zeit t . Sie kann mit der Substitution

$$\vec{y}(\cdot) := \begin{bmatrix} \omega x(\cdot) \\ \frac{dx}{dt}(\cdot) \end{bmatrix} \quad (13.5)$$

in eine lineare Differentialgleichung mit konstanten Koeffizienten umgeschrieben werden, wie dies in den Übungen erklärt wird. Damit ergibt sich die äquivalente Differentialgleichung erster Ordnung

$$\frac{d\vec{y}}{dt}(\cdot) = \omega \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \vec{y}(\cdot). \quad (13.6)$$

Äquivalent bedeutet folgendes: Eine Lösung der Differentialgleichung zweiter Ordnung 13.4 erzeugt über die Substitution 13.5 eine Lösung der Differentialgleichung erster Ordnung 13.1.3. Umgekehrt ist die erste Koordinate einer Lösung der Differentialgleichung 13.1.3 eine Lösung der Differentialgleichung 13.4.

Es ist leicht nachzuprüfen, dass für jeden Vektor \vec{y}_0 aus \mathbb{R}^2 die für alle $t \in \mathbb{R}$ definierte Funktion

$$\vec{y}(t) := \begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix} \vec{y}_0 \quad (13.7)$$

eine Lösung der Differentialgleichung (13.1.3) ist. Die Matrix bezeichnet eine Drehung der Ebene um den Winkel ωt entgegen den Uhrzeigersinn (also $\vec{y}(t) = R(-\omega t) \vec{y}_0$ mit (9.3)).

13.2 DIFFERENTIALGLEICHUNGEN UND LÖSUNGEN ZEICHNEN, DAS PHASENRAUMPORTRAIT

Eine der besten Methoden, Differentialgleichungen zu studieren, ist sie zu “zeichnen”. Natürlich kann eine solche Methode keine numerisch exakten Resultate liefern. Sie gestattet aber dennoch, ein qualitatives Verständnis von Lösungen und stellt damit eine wichtige Methode zur Überprüfung von Resultaten, die durch mathematische Software gewonnen wurden, dar. Wir machen dies an der Pendelgleichung vor.

In diesem Abschnitt begeben wir uns also ganz in das Gebiet der Heuristik! Zeichnungen und Berechnungen der Lösungen sind mit dem Rechner erstellt.

Der Zustand eines Pendels zur Zeit t ist durch den Auslenkwinkel $\phi(t)$ aus der Ruhelage und seine Winkelgeschwindigkeit $\frac{d\phi}{dt}(t)$ beschrieben (vgl. folgende Abbildung 13.1). Wir fassen diese beiden Größen in einen Vektor

$$\vec{y}(t) := \begin{bmatrix} \phi(t) \\ \frac{d\phi}{dt}(t) \end{bmatrix}$$

zusammen. Die Newtonsche Gleichung für das Pendel, kurz die Pendelgleichung, mit der Reibungskonstanten α , der Länge L und der Erdbeschleunigung g ist

$$\frac{d^2\phi}{dt^2}(\cdot) + \alpha \frac{d\phi}{dt}(\cdot) + (g/L) \sin \phi(\cdot) = 0.$$

Sie ist nichtlinear da die Sinusfunktion nicht linear ist.

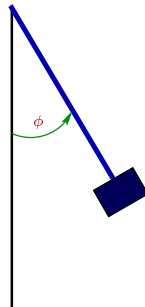


ABBILDUNG 13.1: Das Pendel

Zur Vereinfachung der Formeln wählen wir $\alpha = 0.3$ und $g/L = 1$. Umgeschrieben in \vec{y} wird aus der obigen Differentialgleichung zweiter Ordnung eine Differentialgleichung erster Ordnung

$$\begin{bmatrix} \frac{dy_1}{dt}(\cdot) \\ \frac{dy_2}{dt}(\cdot) \end{bmatrix} = \begin{bmatrix} y_2(\cdot) \\ -\sin y_1(\cdot) - 0.3 y_2(\cdot) \end{bmatrix} \quad (13.8)$$

Die obige Gleichung bedeutet: Zu jeder Zeit t ist die Geschwindigkeit $\begin{bmatrix} \frac{dy_1}{dt}(t) \\ \frac{dy_2}{dt}(t) \end{bmatrix}$ gleich dem Vektor $\begin{bmatrix} y_2(t) \\ -\sin y_1(t) - 0.3 y_2(t) \end{bmatrix}$. *Er hängt nur von $\vec{y}(t)$ ab.*

Dies kann in der folgenden Weise gezeichnet werden: An jedem Ort mit dem Ortsvektor \vec{y} in \mathbb{R}^2 fügen wir den Geschwindigkeitsvektor $\begin{bmatrix} y_2 \\ -\sin y_1 - 0.3 y_2 \end{bmatrix}$ an. Auf diese Weise erhält man ein sogenanntes *Vektorfeld*, das in der Abbildung 13.2 rot eingetragen ist. Wie man sich leicht überzeugt, ist das Vektorfeld periodisch in der y_1 -Richtung, da der Sinus periodisch ist, d.h. weil für alle reellen y_1 , $\sin(y_1) = \sin(y_1 + 2\pi)$ gilt. Eine Lösung der Differentialgleichung kann jetzt leicht gezeichnet werden.

Nun kommen wir zur Linearisierung der Differentialgleichung, die dadurch gewonnen wird, dass die nichtlineare Funktion $F(y_1) = \sin y_1$ durch ihre Linearisierte bei $y_1 = 0$, d.h. durch die lineare Funktion $f(y_1) = F'(0)y_1 = y_1$ ersetzt wird:

$$\begin{bmatrix} \frac{dy_1}{dt}(\cdot) \\ \frac{dy_2}{dt}(\cdot) \end{bmatrix} = \begin{bmatrix} y_2(\cdot) \\ -y_1(\cdot) - 0.3 y_2(\cdot) \end{bmatrix} \quad (13.9)$$

$$= \begin{bmatrix} 0 & 1 \\ -1 & -0.3 \end{bmatrix} \begin{bmatrix} y_1(\cdot) \\ y_2(\cdot) \end{bmatrix} \quad (13.10)$$

Wir zeichnen wiederum an jeden Ortsvektor den Geschwindigkeitsvektor

$$\begin{bmatrix} y_2 \\ -y_1 - 0.3 y_2 \end{bmatrix}$$

in blauer Farbe. Wie die Abbildung 13.2 zeigt, sind sich die Vektorfelder in der Nähe des Ursprungs sehr ähnlich und damit ist es anschaulich klar, dass dies auch für die jeweiligen Lösungen zutrifft. All dies gilt – wie die Skizze zeigt – nur falls sich alles in der Nähe des Ursprungs abspielt oder anders ausgedrückt, falls die Auslenkungen aus der Ruhelage des Pendels klein sind. Die Methode der Linearisierung versagt, wenn wir beispielsweise den Überschlag des Pendels beschreiben wollen. Dennoch ist die Qualität der Linearisierung in der Nähe des Ursprungs sehr gut, wie die Abbildung 13.2 zeigt, worin die beiden Vektorfelder und eine numerisch berechnete Lösung des Anfangswertproblems gezeichnet sind, die auf der Exp-Methode beruht, die im nächsten Abschnitt erklärt wird.

13.3 DIE EXP-METHODE ZUR BERECHNUNG VON LÖSUNGEN

Die Exponentialmethode zur Lösung des Anfangswertproblems für eine lineare Differentialgleichung mit konstanten Koeffizienten für n Funktionen $y_1(\cdot)$,

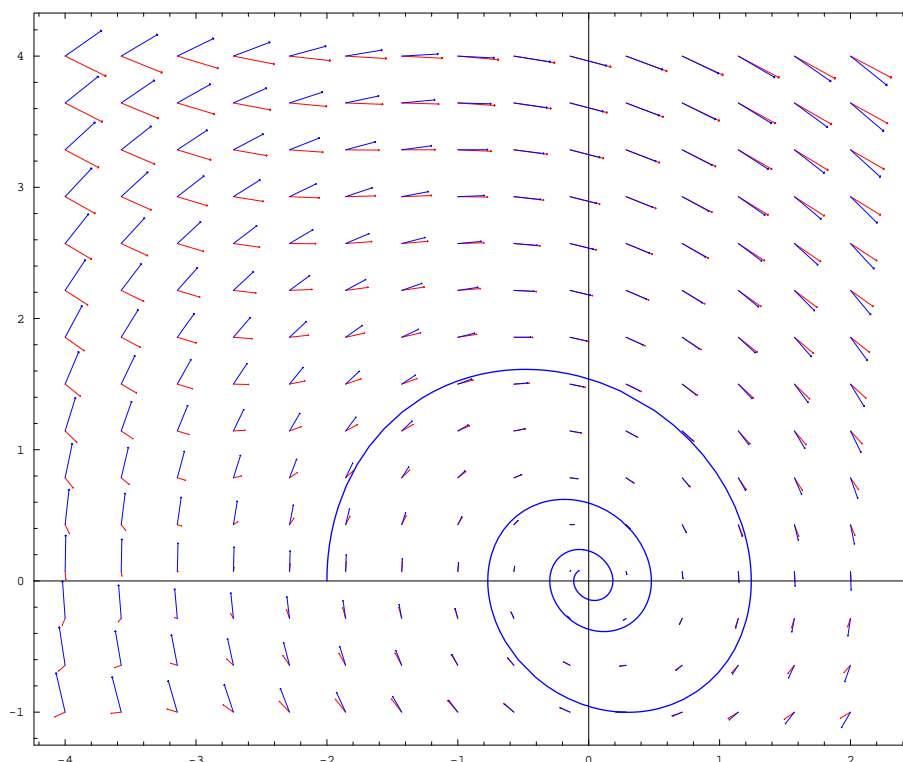


ABBILDUNG 13.2: Das Vektorfeld des Pendels (rot), des linearisierten Pendels (blau) und die Lösung der AWP's für das linearisierte Pendel nach der Exp-Methode

$y_2(\cdot), \dots, y_n(t)$ verallgemeinert die bekannte Situation einer einzigen Funktion wie folgt:

Betrachte das Anfangswertproblem

$$\begin{aligned} \frac{dy}{dt}(\cdot) &= ay(\cdot) \\ y(0) &= y_0. \end{aligned} \quad (13.11)$$

Aus der Analysis ist bekannt, dass die Lösung dieser Gleichung durch die Exponentialfunktion

$$y(t) = e^{ta} y_0$$

gegeben ist. Der folgende Satz ist deshalb keine Überraschung:

Theorem 13.3.1. *Das Anfangswertproblem (13.3) hat die eindeutige für alle $t \in \mathbb{R}$ definierte Lösung*

$$\vec{y}(t) := e^{(t-t_0)A} \vec{y}_0.$$

Dabei ist die Exponentialfunktion der quadratischen Matrix über die Potenzreihe erklärt (vgl. 12.1).

Beweis: Der Beweis des Satzes geht ganz analog zu dem entsprechenden Satz in der Analysis. Die Ableitung der rechten Seite kann ausgerechnet und so die Behauptung verifiziert werden.

Bemerkung 13.3.2. In der letzten Vorlesung wurde die Exponentialfunktion einer quadratischen Matrix definiert und gezeigt, wie diese für diagonalisierbare Matrizen besonders einfach berechnet werden kann.

Folgerung 13.3.3. *Theorem 13.3.1 hat eine bemerkenswerte Konsequenz für den Fall, dass \vec{y}_0 ein Eigenvektor von A zum Eigenwert λ ist. Dann gilt für die Lösung des Anfangswertproblems:*

$$\vec{y}(t) = e^{\lambda(t-t_0)} \vec{y}_0.$$

Umgekehrt ist es ein Leichtes zu verifizieren: Falls $\vec{y}(t) := e^{\lambda(t-t_0)} \vec{y}_0$ eine Lösung der obigen Differentialgleichung ist, dann ist \vec{y}_0 Eigenvektor zum Eigenwert λ .

Analog gilt für den Fall, dass \vec{y}_0 als Linearkombination von Eigenvektoren $\vec{v}_1, \dots, \vec{v}_k$ mit Eigenwerten $\lambda_1, \dots, \lambda_k$ geschrieben werden kann,

$$\vec{y}_0 = \sum_{i=1}^k \alpha_i \vec{v}_i.$$

Dann ist

$$\vec{y}(t) = \sum_{i=1}^k \alpha_i e^{\lambda_i(t-t_0)} \vec{v}_i.$$

Nun formulieren wir einen (wiederum theoretischen) Algorithmus für reelle lineare Differentialgleichungen, der in manchen Anwendungen nützlich ist:

Algorithmus 13.3.4 (Lösung des AWP's für reelle DGL). Sei $A \in \mathbb{R}^{n,n}$ eine Matrix, die – aufgefasst als komplexe Matrix – diagonalisierbar ist. Betrachte das zugehörige Anfangswertproblem

$$\begin{aligned} \frac{d\vec{y}}{dt}(\cdot) &= A\vec{y}(\cdot) \\ \vec{y}(t_0) &= \vec{y}_0 \in \mathbb{R}^n. \end{aligned} \tag{13.12}$$

Um dies zu lösen, verwenden wir folgende Strategie:

1. Schritt: Diagonalisiere die Matrix im Komplexen:

$$A = SDS^{-1}.$$

2. Schritt: Berechne für alle $t \in \mathbb{R}$

$$\vec{y}(t) = Se^{(t-t_0)D}S^{-1}\vec{y}_0.$$

Dies ist die Lösung des Anfangswertproblems.

Bemerkung 13.3.5. Obwohl weder S noch D reelle Matrizen sind, ist $\vec{y}(t)$ dennoch reell und damit eine Lösung des Anfangswertproblems. Dies folgt einfach aus folgendem:

1. Wenn A eine reelle quadratische Matrix ist, so trifft dies auch für e^{tA} zu.
2. Die Kombination der drei quadratischen Matrizen $Se^{tD}S^{-1}$ ist gleich e^{tA} und damit reell. e^{tA} bildet also den reellen Vektor \vec{y}_0 in einen reellen ab.

Beispiel 13.3.6. Betrachte das Anfangswertproblem

$$\begin{aligned} \frac{d\vec{y}}{dt}(\cdot) &= A\vec{y}(\cdot) \\ A &:= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ \vec{y}(0) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned} \quad (13.13)$$

Die Lösung kann nach dem Algorithmus 13.3.4 konstruiert werden:

1. Schritt: Diagonalisierung der Matrix im Komplexen:

Die Matrix A hat die Eigenwerte $\lambda_1 = i, \lambda_2 = -i$ und eine zugehörige Basis von Eigenvektoren

$$\vec{v}_1 := \begin{bmatrix} 1 \\ i \end{bmatrix}, \vec{v}_2 := \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Eine kurze Rechnung zeigt:

$$\begin{aligned} A &= SDS^{-1} \\ S &:= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \\ D &:= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \end{aligned}$$

2. Schritt: Für alle $t \in \mathbb{R}$ ist:

$$\begin{aligned} \vec{y}(t) &= Se^{tD}S^{-1}\vec{y}_0 \\ &= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \vec{y}_0 \\ &= \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \vec{y}_0 \\ &= \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} \end{aligned}$$

Bemerkung 13.3.7. Das AWP kann oft am Einfachsten nach dem in der Folgerung 13.3.3 beschriebenen Verfahren gelöst werden:

Schreibe \vec{y}_0 als Linearkombination der Eigenvektoren von A . Dann gilt für alle $t \in \mathbb{R}$

$$\begin{aligned} \vec{y}(t) &= e^{it} \frac{1}{2} \vec{v}_1 + e^{-it} \frac{1}{2} \vec{v}_2 \\ &= \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix}. \end{aligned}$$

LINEARE DIFFERENTIALGLEICHUNGEN (2. TEIL)

Wie kann aus einem Antennenschwingkreis das Signal mit der richtigen Trägerfrequenz herausgezogen werden? Welchen Ton muss man treffen, um ein Glas zum Zerspringen zu bringen? Auf was muss geachtet werden, wenn eine Gruppe im Gleichschritt über eine Hängebrücke geht, welche Tourenzahlen müssen vermieden werden, damit einem die Rotorblätter eines Triebwerkes nicht um die Ohren fliegen? All dies kann mit der Schwingungsgleichung mathematisch formuliert und untersucht werden. Ihre einfachste Form ist

$$x''(t) + \omega^2 x(t) = A \cos(\Omega t) \quad (14.1)$$

mit der speziellen Lösung

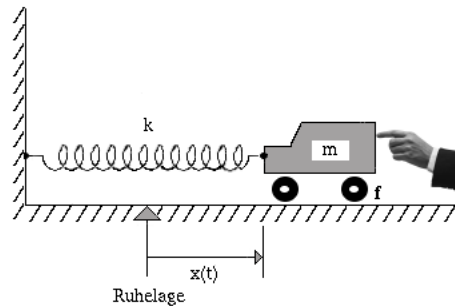
$$x(t) = \frac{A}{\omega^2 - \Omega^2} \cos(\Omega t). \quad (14.2)$$

Hier steht $x(t)$ für die Auslenkung aus der Ruhelage zur Zeit t und $A \cos(\Omega t)$ die treibende äußere Kraft, die mit der Frequenz Ω auf das System wirkt (die Frequenzen sind hier in Einheiten Winkel pro Sekunde gemessen). Bereits an dieser speziellen Lösung ist ein charakteristisches Phänomen erkennbar: Falls die sogenannte Eigenfrequenz ω gleich der Frequenz Ω wird, geschieht etwas ganz besonderes *Resonanz*: Die Amplitude $\frac{A}{\omega^2 - \Omega^2}$ der Schwingung divergiert, das Glas zerspringt, die Brücke stürzt ein, das Rotorblatt reisst sich los, das Signal wird empfangen....

Wie kommt es zur Schwingungsgleichung, die für Mechanik und Elektrodynamik von herausragender Bedeutung ist? Die Antwort haben wir eigentlich bereits in der letzten Vorlesung kennen gelernt. Wir hatten das gezeigt, wie die Schwingungsgleichung aus der Pendelgleichung durch Linearisierung hergeleitet werden kann.

In dieser Vorlesung werden wir am Beispiel der Schwingungsgleichung die Theorie der linearen Differentialgleichungen n -ter Ordnung entwickeln. Dabei beschränken wir uns im wesentlichen auf den Fall konstanter Koeffizienten.

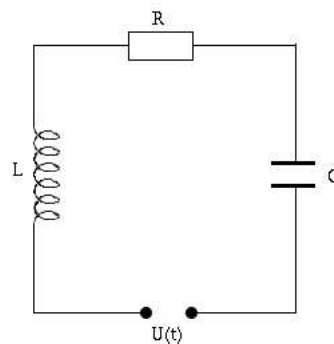
Nun wollen wir noch etwas genauer auf die physikalische Bedeutung der Schwingungsgleichung eingehen. Die Schwingungsgleichung modelliert die mechanische Schwingung einer Masse m an einer Feder mit der Federkonstanten k und der Reibung f mit oder ohne einer Kraft, die von außen auf den Massenkörper wirkt. t bezeichnet die Zeit und ist als reelle Variable aufgefasst; die anderen Grössen sind in der untenstehenden (Fig. 14.15) erklärt.



$$m\ddot{x}(t) + f\dot{x}(t) + kx(t) = A \cos(\Omega t) \quad (14.3)$$

Figur 1: Mechanische Schwingung einer Masse an einer Feder mit äußerer Kraft.

Die gleiche Schwingungsgleichung modelliert auch den elektrischen Schwingkreis, der aus einer Spule mit Selbstinduktion L , einem Widerstand R und Kapazität C besteht mit oder ohne einer von außen angelegten Spannung $U(t)$ (Fig 1).



Figur 2: Elektrischer Schwingkreis mit äußerer Spannung $U(t)$.

Für spezielle periodische äußere Spannungen oder Kräfte kann der für die Anwendung besonders wichtige Fall der Resonanz auftreten, wie wir ihn zu Beginn der Vorlesung kurz beschrieben haben.

Um die Schwingungsgleichung zu lösen und deren Lösungsmengen zu studieren, verwenden wir unter anderem Methoden der Linearen Algebra, obwohl dies nicht mehr ganz so offensichtlich ist wie in der letzten Vorlesung; denn Schwingungsgleichung sind nicht von dem Typ den wir im letzten Kapitel betrachtet haben: Schwingungsgleichungen sind Differentialgleichung *zweiter Ordnung*.

Es ist grundsätzlich möglich, jede Differentialgleichung höherer Ordnung als ein System von Differentialgleichungen erster Ordnung aufzufassen und sie dann mit Hilfe der linearen Algebra - so wie in der letzten Vorlesung zu lösen.

Wir werden heute einen anderen Weg darstellen, der direkt von den Differentialgleichungen zweiter Ordnung ausgeht.

Noch ein Wort zur Notation: Lösungen von Differentialgleichungen sind Funktionen. Sie werden in Anwendungen oft als Funktionen der Zeit oder des Ortes aufgefasst. Im ersten Fall ist es üblich die Notation $x(t)$ zu verwenden im zweiten $y(x)$; t bezeichnet dann die Variable "Zeit". x bezeichnet einmal die Funktion $x(t)$ aber dann auch wieder die Variable "Ort". Dies ist sicherlich alles etwas unglücklich - gehört aber zu den aus der Tradition gewachsenen Standardverwirrspielen der Mathematiker und ihre Anwender. Um diese Problematik zu vermeiden, werden wir standardmäßig die gesuchte Funktion mit $y(t)$ bezeichnen. Die Ableitungen sind dann $y'(t)$ oder $y''(t)$.

Die Theorie der linearen Differentialgleichungen ähnelt sehr stark den linearen Gleichungssystemen und ihrer Theorie. Dieser Zusammenhang ist sehr nützlich. Um ihn hervorzuheben, werden wir im folgenden oft von der "Gleichung" sprechen, wenn eigentlich der Begriff "Differentialgleichung" gemeint ist. Als Abkürzung für den Begriff Lineare Differentialgleichung verwenden wir auch die Notation DGL.

14.1 DIE LINEARE DIFFERENTIALGLEICHUNG ZWEITER ORDNUNG

Jetzt kommen wir zur Definition des Begriffs "Differentialgleichung zweiter Ordnung", zur Struktur ihrer Lösungsmenge und zum Anfangswertproblem. Das wichtigste Beispiel einer solchen Differentialgleichung ist die Schwingungsgleichung.

14.1(a) Definitionen und Begriffe

Die lineare Differentialgleichung zweiter Ordnung ist ähnlich definiert wie die Differentialgleichung erster Ordnung, die wir im vorangehenden Kapitel eingeführt hatten (Def.13.1.1). Im Unterschied zu dort enthält sie aber auch die zweite Ableitung der gesuchten Funktion.

Definition 14.1.1 (Lineare Differentialgleichung zweiter Ordnung). Seien $a_1(t)$, $a_2(t)$, $b(t)$ stetige Funktionen auf einem Intervall $I \subset \mathbb{R}$. Dann nennt man

$$y''(t) + a_1(t)y'(t) + a_2(t)y(t) = b(t) \quad (14.4)$$

inhomogene lineare Differentialgleichung zweiter Ordnung.

Die Funktion $b(t)$ ist die sogenannte *Inhomogenität*. Die Funktionen $a_1(t)$, $a_2(t)$ heißen *Koeffizientenfunktionen*. Alle auftretenden Funktionen haben Werte in den reellen oder komplexen Zahlen. Man spricht dann von *reellen* oder *komplexen Differentialgleichungen*.

Falls $b(t) = 0$ ist, spricht man von einer *homogenen linearen Differentialgleichung zweiter Ordnung*:

$$y''(t) + a_1(t)y'(t) + a_2(t)y(t) = 0 \quad (14.5)$$

Falls die Koeffizientenfunktionen t -unabhängige Konstanten a_1, a_2 sind, spre-

chen wir von einer *linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten* oder von der *Schwingungsgleichung*.

Lineare Differentialgleichungen zweiter Ordnung – so einfach sie aussehen – können im Allgemeinen nicht explizit gelöst werden. Sie gehören zu den sogenannten integrierbaren Systemen, d.h. solchen deren Lösungen durch Integrale dargestellt werden können. Diese führen allerdings nicht immer auf elementare Funktionen. Weil Differentialgleichungen im Allgemeinen keine expliziten Lösungen haben, sind qualitative Resultate über die Lösungen von großem praktischem Nutzen.

Die homogene Schwingungsgleichung beschreibt die *freie Schwingung*, die inhomogene die *erzwungene*. Beide sind von zentraler Bedeutung für Mechanik und Elektrotechnik. Wir werden sie deshalb besonders genau beschreiben.

Hier ist ein erstes Beispiel diesen Typs:

Beispiel 14.1.2. [Harmonischer Oszillator]: Sei ω eine reelle Zahl und

$$y''(t) + \omega^2 y(t) = t \quad (14.6)$$

eine inhomogene Differentialgleichung. Sie heißt die harmonische Oszillatorgleichung mit der Kreisfrequenz oder Winkelgeschwindigkeit ω . Sie ist ein Spezialfall der Gleichung 14.4 und hat viele Lösungen auf deren Struktur wir im nächsten Abschnitt eingehen. Eine spezielle Lösung, oder wie in diesem Zusammenhang oft gesagt wird eine "partikuläre" Lösung, ist:

$$y_P(t) = \frac{t}{\omega^2} \quad (14.7)$$

Die homogene DGL zu der inhomogenen Gleichung 14.6 ist

$$y''(t) + \omega^2 y(t) = 0. \quad (14.8)$$

Eine Lösung dieser Gleichung ist beispielsweise

$$y_H(t) = \cos(\omega t).$$

14.1(b) Struktur der Lösungsmenge

Die Lösungsmengen der Gleichungen 14.4 und 14.5 haben die gleiche Struktur wie wir sie bei linearen inhomogenen und homogenen Gleichungssystemen in der Vorlesung formuliert haben (Theorem 6.4.7 über die Struktur des Lösungsraums):

Die Lösungsmenge der homogenen Gleichung 14.5 ist ein Teilraum eines Vektorraums. Um dies einzusetzen, betrachten wir den Vektorraum $\mathcal{C}^2(\mathbb{I})$ der mindestens zweimal stetig differenzierbaren Funktionen und die lineare Abbildung L von $\mathcal{C}^2(\mathbb{I})$ in den Vektorraum $\mathcal{C}^0(\mathbb{I})$ der stetigen Funktionen, wobei die Notation der obigen Definition 14.1.1 verwendet wurde:

$$\begin{aligned} L : \mathcal{C}^2(\mathbb{I}) &\rightarrow \mathcal{C}^0(\mathbb{I}) \\ y &\mapsto y'' + a_1 y' + a_2 y \end{aligned}$$

Der Kern von L ist nach Konstruktion der Lösungsraum der homogenen DGL 14.5 und damit ein Teilraum von $\mathcal{C}^2(\mathbb{I})$; d.h. die Summe zweier Lösungen von 14.5 ist wieder eine Lösung. Genauso ist das Produkt einer Lösung mit einer Zahl aus \mathbb{R} oder \mathbb{C} wieder eine Lösung.

Die Lösungsmenge der inhomogenen Gleichung ist *kein* Teilraum von $\mathcal{C}^2(\mathbb{I})$, sondern ein sogenannter *affiner Raum*: Das heißt die Differenz zweier Lösungen der inhomogenen Gleichung ist eine Lösung der homogenen Gleichung, oder anders ausgedrückt: Sei $y_P(t)$ eine Lösung der inhomogenen Gleichung und $y_H(t)$ eine Lösung der homogenen Gleichung, dann ist $y(t) := y_P(t) + y_H(t)$ eine neue Lösung der inhomogenen Gleichung 14.4.

Beispiel 14.1.3. Wir betrachten wiederum die inhomogene und homogene Gleichung aus Beispiel 14.1.2 und die darin angegebenen Lösungen.

$$\begin{aligned} y''(t) &+ \omega^2 y(t) = t \\ y(t) &= y_P(t) + y_H(t) \\ &= \frac{t}{\omega^2} + \cos(\omega t) \end{aligned} \quad (14.9)$$

Wie groß ist der Kern von L , d.h. was ist die Dimension des Teilraums $\text{Kern}(L)$? Die Antwort auf die Frage kann nur ohne Beweis berichtet werden:

Theorem 14.1.4 (Struktur des Lösungsraums).

- 1.) Die Dimension des Kerns von L ist 2.
- 2.) Sei $y_P(t)$ eine partikuläre Lösung der inhomogenen Gleichung 14.4. Dann kann jede Lösung der Gleichung als Summe der partikulären Lösung und einer Lösung der homogenen Gleichung geschrieben werden, d.h. die allgemeine Lösung ist von der Form:

$$y(t) = y_P(t) + y_H(t),$$

wobei $y_H(t)$ die allgemeine Lösung der homogenen Gleichung ist, d.h. die Lösungsmenge ist:

$$y_P(t) + \text{Kern}(L)$$

Folgerung: Falls wir zwei linear unabhängige Lösungen der homogenen Differentialgleichung 14.5 haben, d.h. falls eine **Lösungsbasis** vorliegt, ist jede Lösung der homogenen Gleichung als Linearkombination der beiden darstellbar.

14.1(c) Das Anfangswertproblem

Oftmals ist man an einer Lösung der Gleichung 14.4 oder 14.5 interessiert, die zu einem vorgegebenen Zeitpunkt t_0 (Anfangszeit) einen bestimmten Anfangswert y_0 und eine bestimmte Anfangsgeschwindigkeit y_1 hat, d.h. wir suchen eine Lösung $y(t)$ der homogenen oder der inhomogenen Gleichung, sodass gilt

$$\begin{aligned} y(t_0) &= y_0 \\ y'(t_0) &= y_1. \end{aligned} \quad (14.10)$$

Die Aufgabe eine Lösung einer Differentialgleichung zu finden, die diesen Anfangswerten genügt, heißt Anfangswertproblem der Differentialgleichung.

Theorem 14.1.5 (Anfangswertproblem). Gegeben sei die Differentialgleichung

$$y''(t) + a_1(t)y'(t) + a_2(t)y(t) = b(t) \quad (14.11)$$

mit stetigen reell- oder komplexwertigen Funktionen a_1, a_2, b auf einem Intervall $\mathbb{I} \subset \mathbb{R}$. Seien weiter $t_0 \in \mathbb{I}$ und y_0, y_1 reelle oder komplexe Zahlen. Dann gibt es genau eine auf ganz \mathbb{I} definierte Lösung $y(t)$, die den Anfangsbedingungen

$$\begin{aligned} y(t_0) &= y_0 \\ y'(t_0) &= y_1 \end{aligned}$$

genügen.

Bemerkung 14.1.6. Wir hatten bereits in der letzten Vorlesung vom Anfangswertproblem einer Differentialgleichung erster Ordnung bzw. eines Systems von Differentialgleichungen erster Ordnung gesprochen. Die Anfangsdaten waren da durch den einen Koordinatenvektor \vec{y}_0 gegeben. Da wir es hier mit einer Differentialgleichung zweiter Ordnung zu tun haben, bestehen die Anfangswerte nicht nur aus dem Anfangsort y_0 sondern auch aus der Anfangsgeschwindigkeit y_1 .

Beispiel 14.1.7. Die Differentialgleichung

$$y''(t) + \omega^2 y(t) = 0 \quad (14.12)$$

mit den Anfangsdaten

$$t_0 = 0, \quad y(0) = 1, \quad y'(0) = 0 \quad (14.13)$$

hat die Lösung: $y(t) = \cos(\omega t)$.

14.2 DIE LINEARE DGL. 2-TER ORDNUNG MIT KONSTANTEN KOEFFIZIENTEN: DIE SCHWINGUNGSGLEICHUNG

Jetzt betrachten wir die für sehr viele Anwendungen besonders wichtige Differentialgleichung

$$y''(t) + a_1 y'(t) + a_2 y(t) = b(t) \quad (14.14)$$

mit konstanten Koeffizienten a_1, a_2 . Sie heißt auch *Schwingungsgleichung*.

Um die Lösungsmenge dieser Gleichung zu beschreiben, formulieren wir in einem ersten Abschnitt 14.2(a) einen Algorithmus, mit dem eine Lösungsbasis der homogenen Gleichung berechnet werden kann. Um danach eine partikuläre Lösung der inhomogenen Gleichung zu finden, geben wir in Abschnitt 14.2(b) zwei Verfahren an: Den Algorithmus der Variation der Konstanten und den Algorithmus des Exponentialansatzes. Letzterer kann nur dann angewandt werden, wenn die Anregung (=Inhomogenität) vom Typ $\cos(\Omega t)$ oder $e^{i\Omega t}$ ist. Damit kann dann das Anfangswertproblem gelöst werden. Es führt auf ein System von zwei Gleichungen für zwei Unbekannte und kann elementar gelöst werden.

Homogene Schwingungsgleichungen und Schwingungsgleichungen mit einer Inhomogenität von der Form $e^{i\Omega t}$ oder $\cos \Omega t$ können explizit gelöst werden im Unterschied zu den Differentialgleichungen zweiter Ordnung mit variablen Koeffizientenfunktionen.

Typische Anwendungen sind:

Beispiel 14.2.1. [Gedämpfte Schwingung eines massiven Körpers] Die gedämpfte Schwingung eines massiven Körpers an einer Feder mit einer äußeren Kraft (Figur 1) wird durch die DGL

$$m\ddot{x}(t) + f\dot{x}(t) + kx(t) = A \cos(\Omega t) \quad (14.15)$$

beschrieben. Dabei bezeichnet m die Masse, f den Reibungskoeffizienten, k die Federkonstante und $x(t)$ die Auslenkung aus der Ruhelage zur Zeit t . $A \cos(\Omega t)$ steht für eine von außen auf den Körper wirkende periodische Kraft mit der Kreisfrequenz Ω . Die Gleichung ist nichts anderes als die Ausformulierung des Newtonsche Kraftgesetz für den Fall einer Masse an einer Feder und einer äußeren Kraft.

Da die Masse eine positive Zahl ist, kann die Differentialgleichung durch m dividiert werden und ist dann von der Form (14.14).

Beispiel 14.2.2. [Der gedämpfte elektrische Schwingkreis] Der gedämpfte elektrische Schwingkreis (Figur 2) wird durch die gleiche Differentialgleichung beschrieben wie im vorangehenden Beispiel:

$$L\ddot{q}(t) + R\dot{q}(t) + \frac{1}{C}q(t) = U_0 \cos(\Omega t) \quad (14.16)$$

L ist die Selbstinduktion, R der ohmsche Widerstand und C die Kapazität, $q(t)$ beschreibt die Ladung im Kondensator, $U_0 \cos(\Omega t)$ ist die von außen angelegte Spannung. Die Gleichung ist die Ausformulierung der Energieerhaltung für den Fall des gedämpften elektrischen Schwingkreises.

14.2(a) Die homogene Gleichung: Allgemeine Lösung für die freie Schwingung

Die homogene Differentialgleichung

$$y''(t) + a_1 y'(t) + a_2 y(t) = 0 \quad (14.17)$$

hat eine Lösungsbasis, die mit einem Exponentialansatz gefunden werden kann:

Algorithmus 14.2.3 (Lösungsbasis der homogenen Schwingungsgleichung).

Input: Freie Schwingungsgleichung 14.17 d.h. homogene reelle oder komplexe Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten.

Output: Reelle bzw. komplexe Lösungsbasis $\{y_1(\cdot), y_2(\cdot)\}$.

1. Schritt: Sei $\lambda \in \mathbb{C}$. Betrachte die Ansatzfunktion $y(t)$ und ihre beiden ersten Ableitungen:

$$\begin{aligned} y(t) &= e^{\lambda t}, \\ y'(t) &= \lambda e^{\lambda t}, \\ y''(t) &= \lambda^2 e^{\lambda t}. \end{aligned}$$

2. Schritt: Den Ansatz $y(t)$ in die Differentialgleichung 14.17 einsetzen ergibt die sogenannte *charakteristische Gleichung*:

$$\lambda^2 + a_1 \lambda + a_2 = 0$$

Die Nullstellen sind

$$\begin{aligned} \lambda_+ &= -\frac{a_1}{2} + \sqrt{\left(\frac{a_1}{2}\right)^2 - a_2}, \\ \lambda_- &= -\frac{a_1}{2} - \sqrt{\left(\frac{a_1}{2}\right)^2 - a_2}. \end{aligned}$$

3a. Komplexer Fall :

Falls $a_1^2 - 4a_2 \neq 0$, so ist $\{y_1(t) = e^{t\lambda_+}, \quad y_2(t) = e^{t\lambda_-}\}$ Lösungsbasis.

Falls $a_1^2 - 4a_2 = 0$, so ist $\{y_1(t) = e^{t\lambda_+}, \quad y_2(t) = te^{t\lambda_+}\}$ Lösungsbasis.

3b. Reeller Fall

Falls $a_1^2 - 4a_2 > 0$, so ist $\{y_1(t) = e^{t\lambda_+}, y_2(t) = e^{t\lambda_-}\}$ Lösungsbasis.

Falls $a_1^2 - 4a_2 < 0$, so ist $\{y_1(t) = \cos(t \operatorname{Im} \lambda_+) e^{t \operatorname{Re} \lambda_+} = \operatorname{Re} e^{t\lambda_+},$

$y_2(t) = \sin(t \operatorname{Im} \lambda_+) e^{t \operatorname{Re} \lambda_+} = \operatorname{Im} e^{t\lambda_+}$ Lösungsbasis.

Falls $a_1^2 - 4a_2 = 0$ ist $\{y_1(t) = e^{t\lambda_+}, y_2(t) = t e^{t\lambda_+}\}$ Lösungsbasis.

Bemerkung 14.2.4.

1. Die Aussagen des Algorithmus sind leicht nachzurechnen. Zu beachten ist, dass der Trick mit der charakteristischen Gleichung nur für DGL mit konstanten Koeffizienten sinnvoll und erfolgreich ist!
2. Auch im Falle reeller Differentialgleichungen kann die charakteristische Gleichung komplexe Nullstellen haben.

Beispiel 14.2.5. [Freie elektrische Schwingung] Der freie gedämpfte Schwingungskreis ist durch drei Parameter, die Selbstinduktion $L > 0$, den Ohmschen Widerstand $R \geq 0$ und die Kapazität $C > 0$ charakterisiert. Die Ladung im Kondensator genügt der folgenden Differentialgleichung:

$$Lq''(t) + Rq'(t) + \frac{1}{C}q(t) = 0 \quad (14.18)$$

Zur Berechnung einer Lösungsbasis benützen wir den obenstehenden Algorithmus für den *reellen Fall*. Als erstes muss die charakteristische Gleichung aufgestellt werden, indem die Ansatzfunktion $e^{\lambda t}$ in die DGL eingesetzt wird:

$$\lambda^2 + \frac{R}{L}\lambda + \frac{1}{LC} = 0$$

Die Nullstellen sind

$$\lambda_+ = -\frac{R}{2L} + \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}},$$

$$\lambda_- = -\frac{R}{2L} - \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}}.$$

Sie charakterisieren das qualitative Verhalten der Lösungen. In der Notation des Algorithmus gilt:

Fall 3b: $\left(\frac{R}{L}\right)^2 - \frac{4}{LC} > 0$: Exponentieller Abfall.

Fall 3c: $\left(\frac{R}{L}\right)^2 - \frac{4}{LC} < 0$:

$R > 0$: Exponentieller Abfall und Oszillationen.

$R = 0$: Oszillationen.

Fall 4b: $\left(\frac{R}{L}\right)^2 - \frac{4}{LC} = 0$: Exponentieller Abfall.

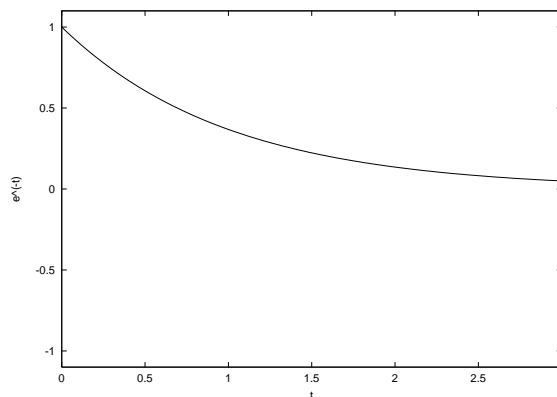
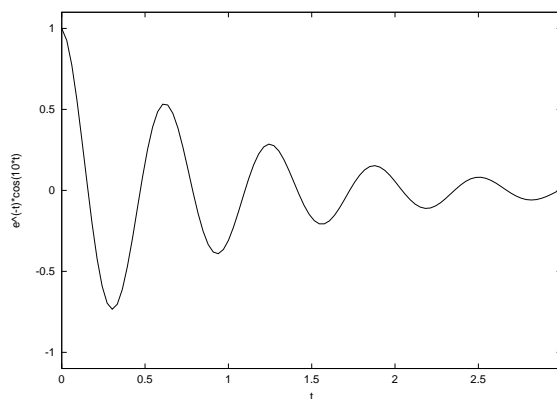


Fig.3 Reiner exponentieller Abfall: Fall 3b und 4b

Fig.4 Exponentieller Abfall mit Oszillation: Fall 3c, $R > 0$.

14.2(b) Die inhomogene Gleichung: Die erzwungene Schwingung

Wir haben bereits die Strukturen der Lösungsmenge der inhomogenen DGL (vgl. Satz 14.1.4) verstanden und wissen daher, dass es genügt eine partikuläre Lösung zu finden, um alle Lösungen der inhomogenen Schwingungsgleichung durch Addition einer Lösung der homogenen Gleichung zu finden.

Wir werden zwei verschiedene Methoden angeben, um eine partikuläre Lösung zu konstruieren. Die erste funktioniert in jedem Fall und heißt die Methode der Variation der Konstanten. Die Idee ist von einer Linearkombination der Elemente einer Lösungsbasis auszugehen und die Koeffizienten nicht als Konstanten, sondern als Funktionen aufzufassen,

$$y(t) = y_1(t)c_1(t) + y_2(t)c_2(t) \quad (14.19)$$

Die Koeffizientenfunktionen $c_1(t)$ und $c_2(t)$ können so gewählt werden, dass $y(t)$ zu einer partikulären Lösung wird (vgl. Algorithmus 14.2.6). Die zweite Methode ist auf den speziellen Fall einer Inhomogenität der Form $\cos(\Omega t)$, $\sin(\Omega t)$, $e^{i\Omega t}$ zugeschnitten, der aber für die Anwendung von besonderer Bedeutung ist. Diese zweite Methode fußt auf einem Ansatz für die partikuläre Lösung, der physikalisch motiviert ist.

1. Methode:

Algorithmus 14.2.6 (Variation der Konstanten).

Input: Inhomogene Schwingungsgleichung,

Lösungsbasis der homogenen Gleichung $\{y_1, y_2\}$

Output: Partikuläre Lösung der inhomogenen Gleichung.

1. Schritt, Gleichungssystem für $c'_1(t), c'_2(t)$:

Löse das Gleichungssystem

$$\begin{bmatrix} y_1(t) & y_2(t) \\ y'_1(t) & y'_2(t) \end{bmatrix} \begin{bmatrix} c'_1(t) \\ c'_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ b(t) \end{bmatrix} \quad (14.20)$$

nach $c'_1(t), c'_2(t)$ auf (das ist immer eindeutig möglich). Die 2×2 Matrix heißt die *Wronskimatrix*.

2. Schritt : Integriere das Resultat. Daraus ergibt sich $c_1(t), c_2(t)$.

3. Schritt : Die Funktion

$$y(t) = c_1(t) y_1(t) + c_2(t) y_2(t)$$

ist eine partikuläre Lösung.

Wie es zu diesem Algorithmus kommt, erklären wir am Beispiel der Gleichung

$$\frac{d^2x}{dt^2}(t) + x(t) = b(t).$$

Die homogene Gleichung haben wir bereits in der letzten Vorlesung diskutiert und sie in ein System einer DGL erster Ordnung umgeschrieben. Dies tun wir nun auch für den inhomogenen Fall: Mit der Schreibweise

$$\vec{y}(t) := \begin{bmatrix} x(t) \\ x'(t) \end{bmatrix}$$

ergibt sich

$$\frac{d\vec{y}}{dt}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \vec{y}(t) + \begin{bmatrix} 0 \\ b(t) \end{bmatrix}.$$

Die homogene Gleichung

$$\frac{d\vec{y}}{dt}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \vec{y}(t)$$

hat die Lösungen

$$\vec{y}(t) = S(t) \vec{y}_0,$$

wobei

$$S(t) := \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} = \begin{bmatrix} \cos t & -\sin(-t) \\ \sin(-t) & \cos t \end{bmatrix} = R(-t)$$

eine Drehung um den Winkel $-t$ und $\vec{y}_0 \in \mathbb{R}^2$ den Vektor der Startwerte bezeichnet. Nun machen wir den Ansatz

$$\vec{y}(t) = S(t) \vec{c}(t)$$

und setzen ihn in die inhomogene Gleichung ein:

$$\begin{aligned} \frac{d\vec{y}}{dt}(t) &\stackrel{\text{DGL}}{=} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} S(t) \vec{c}(t) + \begin{bmatrix} 0 \\ b(t) \end{bmatrix} \\ &= S(t)' \vec{c}(t) + \begin{bmatrix} 0 \\ b(t) \end{bmatrix}. \end{aligned}$$

$\frac{d\vec{y}}{dt}(t)$ kann ein zweites Mal mit Hilfe der Produktregel berechnet werden:

$$\frac{d\vec{y}}{dt}(t) \stackrel{\text{Prod. Rgl}}{=} S(t)' \vec{c}(t) + S(t) \vec{c}'(t)$$

Daraus folgt durch Vergleich

$$S(t)\vec{c}'(t) = \begin{bmatrix} 0 \\ b(t) \end{bmatrix}. \quad (14.21)$$

Dies ist die Gleichung (14.20) im obigen Algorithmus für den hier betrachteten Fall.

Jetzt wenden wir diese Methode auf die Schwingungsgleichung ohne Dämpfung an und treffen dann auf den Begriff der *Eigenfrequenz*:

Beispiel 14.2.7. Als erstes betrachten wir den elektrischen Schwingkreis ohne Dämpfung, d.h. ohne Ohmschen Widerstand, angetrieben durch eine äußere Spannung (z.B. Antennenspannung) unter der Annahme $\Omega \neq \omega$ (ω ist weiter unten definiert). Der Fall $\Omega = \omega$ heißt *Resonanzfall* und sollte für sich betrachtet werden (!):

$$Lq''(t) + \frac{1}{C}q(t) = U_0 \cos(\Omega t) \quad (14.22)$$

Es ist üblich die DGL umzuschreiben und zwar so:

$$\begin{aligned} q''(t) + \omega^2 q(t) &= A \cos(\Omega t) \\ \omega^2 &:= \frac{1}{LC}, \quad A := \frac{U_0}{L} \end{aligned} \quad (14.23)$$

mit der Abkürzung $\omega := \sqrt{\frac{1}{LC}}$ und $A := \frac{U_0}{L}$.

Als Basis der homogenen Gleichung verwenden wir die komplexwertigen Funktionen

$$\begin{aligned} y_1(t) &:= e^{i\omega t} \\ y_2(t) &:= e^{-i\omega t}. \end{aligned} \quad (14.24)$$

Die Wronskimatrix sieht dann so aus

$$W(t) := \begin{bmatrix} e^{i\omega t} & e^{-i\omega t} \\ i\omega e^{i\omega t} & -i\omega e^{-i\omega t} \end{bmatrix}. \quad (14.25)$$

Nun ist das Gleichungssystem

$$W(t) \begin{bmatrix} c_1'(t) \\ c_2'(t) \end{bmatrix} = \begin{bmatrix} 0 \\ A \cos(\Omega t) \end{bmatrix} \quad (14.26)$$

zu lösen und es ergibt sich:

$$\begin{aligned} \begin{bmatrix} c_1'(t) \\ c_2'(t) \end{bmatrix} &= W^{-1}(t) \begin{bmatrix} 0 \\ A \cos(\Omega t) \end{bmatrix} \\ &= \frac{A}{4i\omega} \begin{bmatrix} e^{i(\Omega-\omega)t} + e^{-i(\Omega+\omega)t} \\ -e^{i(\Omega+\omega)t} - e^{-i(\Omega-\omega)t} \end{bmatrix} \end{aligned} \quad (14.27)$$

Integration ergibt:

$$\begin{bmatrix} c_1(t) \\ c_2(t) \end{bmatrix} = -\frac{A}{4\omega} \begin{bmatrix} \frac{e^{i(\Omega-\omega)t}}{\Omega-\omega} - \frac{e^{-i(\Omega+\omega)t}}{\Omega+\omega} \\ -\frac{e^{i(\Omega+\omega)t}}{\Omega+\omega} + \frac{e^{-i(\Omega-\omega)t}}{\Omega-\omega} \end{bmatrix} \quad (14.28)$$

Somit ergibt sich für die partikuläre Lösung:

$$\begin{aligned}
 y(t) &= c_1(t)y_1(t) + c_2(t)y_2(t) \\
 &= -\frac{A}{4\omega} \left(\frac{e^{i(\Omega-\omega)t}}{\Omega-\omega} - \frac{e^{-i(\Omega+\omega)t}}{\Omega+\omega} \right) e^{i\omega t} \\
 &\quad - \frac{A}{4\omega} \left(-\frac{e^{i(\Omega+\omega)t}}{\Omega+\omega} + \frac{e^{-i(\Omega-\omega)t}}{\Omega-\omega} \right) e^{-i\omega t} \\
 &= \frac{A}{\omega^2 - \Omega^2} \cos(\Omega t)
 \end{aligned} \tag{14.29}$$

Die Größe $\frac{\omega}{2\pi}$ ist die sogenannte *Eigenfrequenz* des Schwingkreises. Sie ist für die folgende Beobachtung wichtig: Die Lösung hat folgende Eigenschaft: Für $\Omega = \omega$ divergiert die Amplitude. Dies ist der Fall der **Resonanz**. Falls Ω betragsmäßig größer ist als ω , dann ist das Vorzeichen der Auslenkung dasselbe wie das der Anregung. Sonst läuft das System in "Gegenphase".

Wie sich im folgenden zeigen wird, ist die Divergenz der Amplitude für den Fall $\Omega = \omega$ eine Folge davon, dass wir die Dämpfung durch Ohmsche Widerstände im Schwingkreis nicht berücksichtigt haben. Dies werden wir jetzt nachholen.

2. Methode:

Den gedämpften elektrischen Schwingkreise, werden wir mit einer neuen Methode behandeln, die - im Unterschied zu der Methode von vorhin - nur dann funktioniert, wenn der Schwingkreis von einer Schwingung der Form $\cos \Omega t$, $\sin \Omega t$ oder $e^{i\Omega t}$ angetrieben wird.

Die Methode beruht auf einem Ansatz, der durch folgende Intuition motiviert ist. Der Schwingkreis wird von außen mit einer Frequenz $\frac{\Omega}{2\pi}$ angetrieben. Wir starten den Schwingkreis zu einer gewissen Zeit. Nach einem Einschwingvorgang wird der Schwingkreis mit der Frequenz $\frac{\Omega}{2\pi}$ der von außen wirkenden Kraft bzw. Spannung schwingen. Dies gibt zu dem untenstehenden Ansatz Anlass.

Die zu lösende inhomogene Differentialgleichung ist folgende:

$$Lq''(t) + Rq'(t) + \frac{1}{C}q(t) = U_0 \cos(\Omega t) \tag{14.30}$$

Sie unterscheidet sich von der Gleichung 14.22 durch den Term Rq' . Statt der obigen Gleichung 14.30 betrachten wir zuerst die Gleichung

$$L\hat{q}''(t) + R\hat{q}'(t) + \frac{1}{C}\hat{q}(t) = U_0 e^{i\Omega t}. \tag{14.31}$$

Offensichtlich gilt: Wenn $\hat{q}(t)$ eine Lösung von 14.30 ist; dann ist $q(t) := \text{Re}\hat{q}(t)$ eine Lösung von 14.31!

Jetzt suchen wir eine Lösung von 14.31 und machen dazu den Ansatz: $\hat{q}(t) = \hat{q}e^{i\Omega t}$. Eingesetzt in 14.31 ergibt sich:

$$(L(i\Omega)^2 + R(i\Omega) + \frac{1}{C})\hat{q}e^{(i\Omega)t} = U_0 e^{(i\Omega)t} \tag{14.32}$$

woraus durch Division mit $e^{(i\Omega)t}$ sofort folgt:

$$\hat{q} = \frac{U_0}{-L\Omega^2 + iR\Omega + \frac{1}{C}}. \tag{14.33}$$

Damit ist sofort schnell und einfach:

$$\hat{q}(t) = \frac{U_0}{-L\Omega^2 + iR\Omega + \frac{1}{C}} e^{(i\Omega)t} \quad (14.34)$$

eine partikuläre Lösung der komplexen Gleichung 14.31. Der Realteil davon ist eine Lösung der reellen Gleichung 14.30, von der wir ausgegangen sind. Mit den Abkürzungen $\omega^2 := \frac{1}{LC}$ und $\gamma := \frac{R}{L}$ erhalten wir:

$$\begin{aligned} q(t) &= \operatorname{Re} \hat{q}(t) \\ &= \frac{U_0}{L} \operatorname{Re} \frac{1}{\omega^2 - \Omega^2 + i\gamma\Omega} e^{(i\Omega)t} \\ &= \frac{U_0}{L} \frac{1}{\sqrt{(\omega^2 - \Omega^2)^2 + (\gamma\Omega)^2}} \cos(\Omega t - \delta) \\ &\quad \text{hier ist } \delta \text{ so definiert:} \\ \sin \delta : &= \frac{\gamma\Omega}{\sqrt{(\omega^2 - \Omega^2)^2 + (\gamma\Omega)^2}} \end{aligned} \quad (14.35)$$

Die Formel sagt, dass die Amplitude $q(t)$ sehr groß wird, falls $\gamma\Omega$ klein und ω^2 sehr nahe an Ω^2 ist.

Offensichtlich kann alles was wir für den "elektrischen Schwingkreis" gesagt haben auf den Fall des "mechanischen Schwingkreises" übertragen werden (Figur 1 und Formel 14.15) falls folgender Übersetzungsschlüssel verwendet wird:

$$\begin{aligned} L &\Rightarrow m \\ R &\Rightarrow f \\ \frac{1}{C} &\Rightarrow k \end{aligned}$$